

NOTICE

This report was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or any agency thereof.

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Available electronically at SciTech Connect <http://www.osti.gov/scitech>

Available for a processing fee to U.S. Department of Energy and its contractors, in paper, from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
OSTI <http://www.osti.gov>
Phone: 865.576.8401
Fax: 865.576.5728
Email: reports@osti.gov

Available for sale to the public, in paper, from:

U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Road
Alexandria, VA 22312
NTIS <http://www.ntis.gov>
Phone: 800.553.6847 or 703.605.6000
Fax: 703.605.6900
Email: orders@ntis.gov

Cover Photos by Dennis Schroeder: (left to right) NREL 26173, NREL 18302, NREL 19758, NREL 29642, NREL 19795.

NREL prints on paper that contains recycled content.

Acknowledgments

The authors thank Richard O’Neill (Federal Energy Regulatory Commission), Beth Garza (Potomac Economics), and Venkat Krishnan and Dan Steinberg (National Renewable Energy Laboratory) for their thoughtful reviews, comments, and suggestions. We especially thank Charlton Clark (U.S. Department of Energy) for sponsoring this work. We are grateful for the editorial support from Katie Wensuc (National Renewable Energy Laboratory) and Maureen McIntyre (McIntyre Communications). This research was funded by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Wind and Water Power Technologies Office under contract number DE-AC36-08GO28308. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes. Any and all errors are the sole responsibility of the authors.

List of Acronyms

ATC	average total cost
CAISO	California Independent System Operator
CRM	capacity remuneration mechanisms
ERCOT	Electric Reliability Council of Texas
FERC	Federal Energy Regulatory Commission
HHI	Herfindahl-Hirschman Index
MC	marginal cost
MR	marginal revenue
MISO	Midcontinent Independent System Operator
ORDC	operating reserve demand curve
PPA	power purchase agreement
RTO	regional transmission organization

Executive Summary

A common approach to regulating electricity is through auction-based competitive wholesale markets. The goal of this approach is to provide a reliable supply of power at the lowest reasonable cost to the consumer and to enable transparency and mitigate market power, defined as the ability of a market actor to influence the market price or exclude competition. This necessitates market structures and operating rules that ensure revenue sufficiency for all generators needed for resource adequacy purposes. Wholesale electricity markets employ marginal-cost pricing to provide cost-effective dispatch such that generators are compensated for their operational costs. However, marginal-cost pricing alone cannot guarantee cost recovery outside of perfect competition, and current electricity market structures have at least six attributes that preclude them from functioning as perfectly competitive markets. These attributes include externalities, public good attributes of the network, ineffective demand curve, market power, wholesale price caps, and lack of large-scale storage.

Some of these attributes—namely externalities and public good attributes (in this case, of the network)—are classic sources of market failure and exist in today’s electricity market construct. The ineffective demand curve feature is arguably the most impactful contributor to market failure and refers to the inability of consumers to express their actual demand for electricity. It results from the “demand-side flaws” of demand inelasticity and a lack of real-time physical enforcement of bilateral contracts.¹ Other attributes—primarily market power, lack of large-scale storage, and wholesale price caps—have compounding causes and effects that amplify underlying market failures. For example, market power exists in part because of significant barriers to entry that yield monopolistic tendencies. Electricity markets are regulated to minimize this market power, but the resulting regulatory-imposed average-price-based retail rate structures contribute to the observed inelastic demand mentioned above. Similarly, the wholesale price caps that are implemented to restrict market power may also prevent prices from reaching levels needed to ensure adequate revenues for generators. In addition, imperfect information related to the planning and operation of the power system, such as uncertainty in load growth and future economic and policy factors, further amplifies these current market structure failures.

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

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

¹ The price elasticity of demand is the sensitivity of demand to a change in price and can be separated into long-run and short-run elasticities. If a 10% increase in price causes a 20% reduction in demand, then the elasticity is 2 (technically, this is a negative value, but the common convention is to redefine it as positive) (Stoft 2002).

the capacity factors of conventional generators through the merit-order effect in the simplest case of a convex market; non-convexities can also suppress prices, even outweighing the impact of low marginal cost generation.²

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

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

²We use the term *convex market* to describe a market where the unit commitment decisions are not considered in the market solution and each unit has a monotonically increasing supply cost function. Sources of non-convexities in real markets are from no-load costs, start-up costs, and certain constraints on generator output such as minimum generation levels and operating time limits. Non-convex representations could include a discontinuous or nonmonotonic supply curve. Non-convexities result in deviations from perfect competition, including inefficient—or the absence of—market equilibrium prices.

Table of Contents

1	Introduction	1
2	Basic Market Types	3
2.1	Perfect Competition.....	4
2.2	Monopoly	6
2.3	Monopolistic Competition.....	8
2.4	Oligopoly.....	9
2.5	Market Outcomes	11
3	Market Structure Complications	12
3.1	Near-Zero Marginal-Cost Examples	12
3.2	Market Failures	14
3.2.1	Market Power	14
3.2.2	Public Goods	15
3.2.3	Externalities.....	16
4	Electricity Markets	18
4.1	Electricity Market Inefficiencies	18
4.2	Revenue Sufficiency	22
4.3	Resource Adequacy.....	26
4.4	Impact of Variable Generation	27
4.5	U.S. Wholesale Electricity Market Examples to Minimize Revenue Sufficiency and Resource Adequacy Challenges.....	29
4.6	Complicating Factors and Unintended Consequences for Market Design.....	30
4.7	Future Research.....	31
5	Summary and Conclusions	32
	References	33
	Appendix: Cost Definitions and Curves	40

List of Figures

Figure 1. Basic market types along a continuum	3
Figure 2. Short-run equilibrium of a perfectly competitive market. The left panel shows an individual representative firm, and the right panel shows the market.....	5
Figure 3. Long-run equilibrium in perfect competition	5
Figure 4. Monopolistic market equilibrium	7
Figure 5. Monopolist with declining average total cost during the relevant market range of output	8
Figure 6. Monopolistic competition in the short run (left) and long run (right) for a representative firm....	9
Figure 7. One potential diagram of oligopoly with economic profit	10
Figure 8. The derivation of the demand curve for a private good (left) and a public good (right).....	16
Figure 9. Negative externalities are not taken into account in the market supply curve, S, and equilibrium quantity, q, and price, p. Supply curve S' and associated p' and q' show the impact of the externality in the market outcome.....	17
Figure 10. Stylistic representation of multiple contributors to market failure in electricity markets	19
Figure 11. Impact of variable renewable generation on market prices. <i>Image from Frew et al. (2016a)</i> ...	28

List of Tables

Table 1. Comparison of Key Outcomes for Different Market Types	11
Table 2. Example of Market Power: Communications.....	13
Table 3. Pricing Example from Communications Industry	13
Table A-1. Cost Definitions and Curves Relevant to Electricity Markets	40

1 Introduction

Electricity, by its nature, requires some form of regulation. There are a variety of well-documented market failures³ that result in a classical “public good” problem (Joskow 1997, 119–138). Some of the most frequently mentioned market failures relate to the inelasticity of demand, nonexclusivity, market power, externalities, and pure monopoly of transmission.¹ In light of such market failures, state and federal regulators have developed two basic approaches to regulating the electricity industry: (1) cost-of-service regulation and (2) auction-based competitive wholesale markets. The common objectives are to ensure that sufficient generation and transmission resources are procured and operated to meet expected conditions at the least cost and to ensure that the owners and operators of the necessary system resources can earn a fair rate of return.

First, the classic approach to addressing the public good problem of the electricity industry is cost-of-service regulation. With this approach, electric utilities operate a regulated monopoly whereby all prudently incurred costs for electricity supply and transmission are entitled to a regulatory approved rate of return.

However, in the early 1990s, several countries developed an alternative regulatory paradigm to address market failures related to the supply of electricity. The fundamental premise is that electricity transmission is a pure monopoly and must be operated in an independent manner to enable competition among electricity generators. Instead of guaranteeing a rate of return for generation, this second approach enables generation owners to earn revenues based on the results of an independently operated, auction-based market. This approach is designed to address supply-side market failures related to planning reserve margins, efficient operation, and market power (Hogan 2008). A key argument for these markets—generally referred to as regional transmission organizations (RTOs) in the United States—is that they create a more competitive market by increasing transparency, reducing the impact of market power, and lowering costs to consumers by co-optimizing operations. In this paper, we focus on the structure and performance of auction-based markets implemented in RTOs.

Although there are differences in many of the details in wholesale electricity market design, a common premise is that a large number of competitive generators can deliver electricity supply while regulatory measures can be enacted to limit the exercise of market power and the associated distortions of competitive prices.⁴ Under this premise, marginal-cost pricing results in a cost-effective dispatch such that generators are compensated for their operational costs.⁵ Generators rely on times when the market clearing price is more than their marginal cost (MC),

³ The term *market failure* is standard in economics. The term does not necessarily refer to a full market collapse; rather, it means that one or more imperfections prevent the market from achieving an economically efficient solution—i.e., a solution that is both allocatively and productively efficient. See Section 3.2 for details on classic sources of market failure.

⁴ Throughout this paper, the term *price* refers to the price of market products at the bulk power system level.

⁵ For a given dispatch or market period, the resource is compensated for variable costs; however, additional out-of-market payments might also apply. These include uplift (or make-whole) payments to cover no-load or start-up costs and day-ahead profit guarantees to prevent generators from losing profits earned in the day-ahead market when those generators would lose money by performing actions that benefit the system in the real-time market. Possible causes for these out-of-market payments include misalignment of average and marginal-cost curves and discrepancies between day-ahead and real-time operational conditions (e.g., Wang 2016).

such as during scarcity pricing events, to provide additional revenue to cover fixed costs; however, such long-term, fixed-cost recovery is not guaranteed under the marginal-cost pricing premise, and this is a key focus of this paper. More broadly, this paper’s position is that there is no reason to expect that marginal-cost pricing alone (i.e., without improvements to underlying market failures or administrative measures to compensate for flaws in energy price formation) will necessarily lead to a stable long-term electricity market or ensure resource adequacy because:

- Existing markets with similar structures (monopolistic competition and oligopoly) have difficulties in the long run with marginal-cost pricing. Specifically, under marginal-cost pricing, when the marginal cost is less than average total cost (ATC), prices are insufficient to cover both fixed and variable costs to support current generators or signal new capacity that might be needed for long-term system reliability.⁶
- No link exists between reliability and pricing unless reliability (or capacity) is priced in addition to energy pricing.
- Even though some reserve pricing might be effective in the short to medium term, it is not clear that this current approach is robust in a future with high penetration levels of variable generation and large amounts of quick-start capability, when the reserve market is likely saturated due to excess capacity for large parts of the year.

Many markets have been influenced recently by depressed energy prices, driven by lower-than-expected demand, historically low natural gas prices, and the adoption of near-zero marginal-cost variable generation, such as wind and solar (Hibbard et al. 2017). The purpose of this paper is to disentangle these surface layer challenges from the underlying market failures to better understand the function and maximize the potential of electricity markets. We aim to bridge the gap—and motivate further dialogue—between the related but often separate topics of economics and power systems. We begin in Section 2 by describing the four basic market types: perfect competition, monopoly, monopolistic competition, and oligopoly. We then use economic principles to examine the basic outcomes of these four markets and show that marginal-cost pricing is economically sustainable only within a perfectly competitive market. In Section 3, we cite some evidence showing that most markets in the United States do not exhibit perfect competition, and we explore pricing structures in other industries that might provide insight. We then discuss the unique and complicating characteristics of electricity markets in Section 4, with a discussion of existing and plausible electric power market structures for an evolving power system. We do not advocate for any specific market structure but rather aim to present a sampling of market structure types and design considerations.

⁶ This reference to long-term reliability is separate and distinct from the common N-1 operational reliability standard. The reliability focus of this paper is on resource (or capacity) adequacy, specifically for long-term capacity needs, but also for shorter-term system balance and flexibility requirements. In addition, the term *sustainability* or *sustainable* can be applied to markets where revenues are sufficient to support a resource-adequate system; we jointly use the terms *revenue sufficiency* and *resource adequacy* throughout this paper to refer to a sustainable market. Price-responsive demand (i.e., downward sloping demand curve) is critical to achieving these objectives, and the lack of this price responsiveness is a key contributor to market failure preventing a sustainable or efficient market, as discussed in Section 4.1.

2 Basic Market Types

This section provides an overview and background of the basic market types according to microeconomic theory. The intent is to provide a basic background for readers who may not have a background in this subject and can be skipped by those who do. Recognizing the differences among these market types is critical to understanding, from a foundational level, the limitations of electricity markets; those readers interested only in the discussion on electricity markets are referred to Section 4.

The four basic market types are (1) perfect competition, (2) monopoly, (3) monopolistic competition, and (4) oligopoly. In reality, these four market types represent points along a continuum because real markets can sometimes consist of a blend of market types. This is shown in Figure 1. We ignore economic externalities and public goods in this section to focus directly on the market structure, but both topics are introduced in Section 3. Our primary focus is on the relationship between price and marginal cost, and we show that in most of these market types, price diverges from marginal cost.

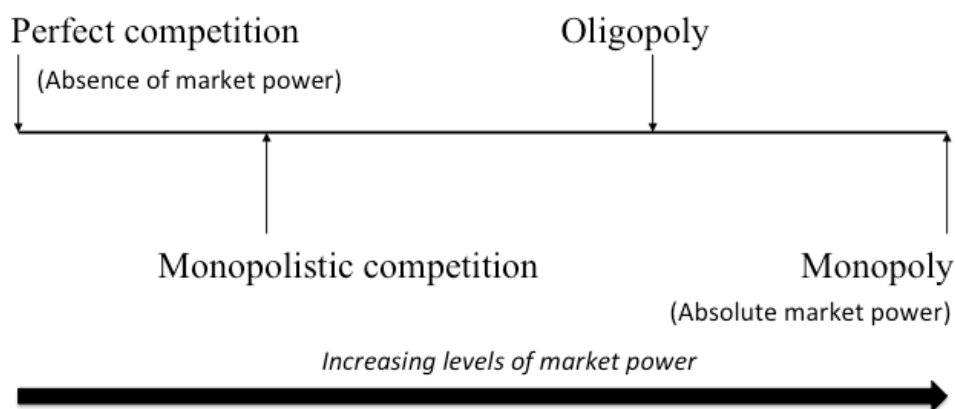


Figure 1. Basic market types along a continuum

As shown in Figure 1, a key characteristic of any market type is the degree of market power. Market power refers to the ability of a market actor to influence the market price or exclude competition. Two common ways to quantify market power are to use (1) four-firm concentration ratios (sometimes five firms are used instead) and (2) the Herfindahl-Hirschman Index (HHI).⁷ In both cases, the market share is defined as the percentage of all market transactions undertaken by one or more firms relative to the entire market. The four-firm concentration ratio is the sum of the top four firms' percentages of the total market. Conversely, the HHI is calculated as the sum of squares of all firms' market shares.

⁷ These market shares and concentration measures are applicable within the broader economics landscape. Concentration metrics such as HHI only account for one of five economic factors that determine the extent of market power; the other four are demand elasticity, style of competition, forward contracting, and geographic extent of the market (Stoft 2002). As a result, a better measure of market power is HHI/e where e is the demand elasticity. See Section 4.1 for a description of market power metrics used more commonly within electricity markets. These methods generally focus on generator offer prices and binding transmission constraints.

On the far left of the spectrum shown in Figure 1, a perfectly competitive market has many firms, none of which has significant market power. There could be, for example, 100 firms, each of comparable size. The HHI for this example would be 100. Moving to the right, a monopolistic competition market could be characterized by an HHI less than 1,000; a monopolistic HHI would be 10,000, and an oligopoly would lie somewhere in between. These HHI limits are not strictly defined, but they generally describe the standards that are used by the U.S. Department of Justice to screen and mitigate market power, specifically when evaluating mergers.⁸

2.1 Perfect Competition

The purest form of competition is known as perfect competition.⁹ Perfectly competitive markets consist of a large number of buyers and sellers to whom perfect information is always available. Neither producers nor consumers can individually impact price; thus, all market participants are price takers. All costs of production are known, and no external costs or benefits accrue to parties outside the transaction. Consumers are not able to discern any differences among goods supplied by different sellers (identical goods), and firms can freely enter or exit the industry. All goods or services bought and sold in this type of market are fully excludable; consumers who do not pay cannot consume the good/service. Consumption is rivalrous, and therefore consumption of the good/service by one consumer implies that another consumer will be required to purchase his/her own product to consume. All firms are assumed to produce the quantity of output that will maximize profits, subject to technical and market constraints and parameters.

To a noneconomist, these conditions might appear to be severely restrictive. In fact, they are, and this is why “perfect” is part of the name of this type of market. In reality, perfectly competitive markets do not exist in this pure form, although some markets might come close. But most markets are generally considered to be monopolistically competitive (buyers and/or sellers have some degree of market power, and products are not identical), oligopolistic (a small number of sellers), or monopolistic (one seller), all of which we discuss.

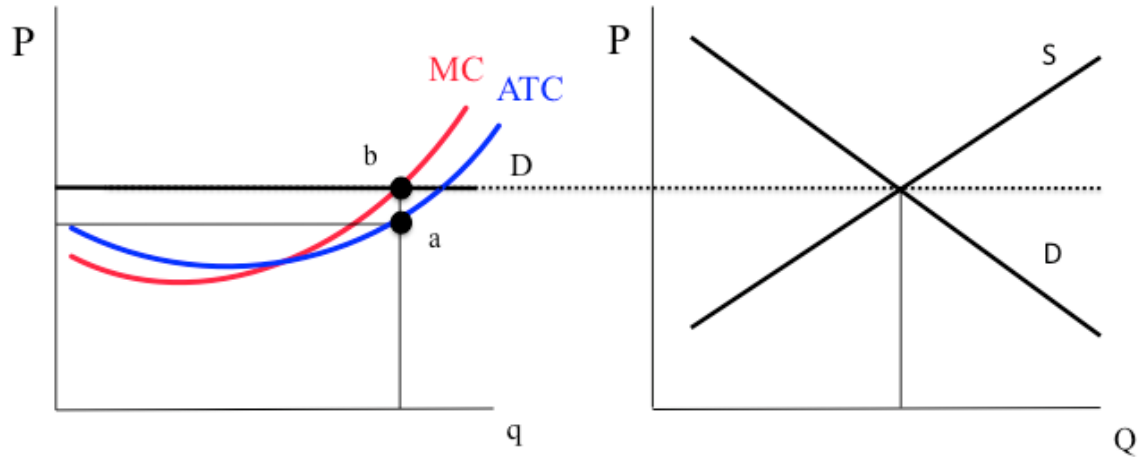
Perfectly competitive markets can be shown to achieve short-run equilibrium at the point at which price equals marginal cost ($P = MC$) (Bade and Parkin 2002). Although individuals have no discernable influence on price, the aggregation of buyers creates a downward-sloping demand curve; and the aggregation of sellers’ marginal-cost curve, which can be shown to represent the firms’ supply curve, creates an upward-sloping market supply curve. In the short run—defined as the period of time during which capital, or other productive factor, cannot be changed—some number of firms could earn positive economic profit.¹⁰ This is shown in Figure 2. The left panel shows the marginal cost and average total cost curve of a representative firm. P designates the price per unit, and q represents the quantity of output per period. Because the firm is a price taker, the demand curve appears perfectly inelastic (insensitive to price). The firm produces its output along the marginal-cost curve: if $P < MC$, then the cost of producing one more unit of output exceeds the price; and if $P > MC$, then the firm will expand output because the cost of producing one more unit of output will increase revenue and profit. The vertical distance

⁸ See <https://www.justice.gov/atr/15-concentration-and-market-shares>.

⁹ For now, we ignore public good attributes and externalities; these are discussed in Section 3.2.2 and Section 3.2.3, respectively.

¹⁰ Normal profit is defined as the minimum profit required to elicit the quantity of entrepreneurship required to run the firm. Profit in excess of this minimum is called economic profit.

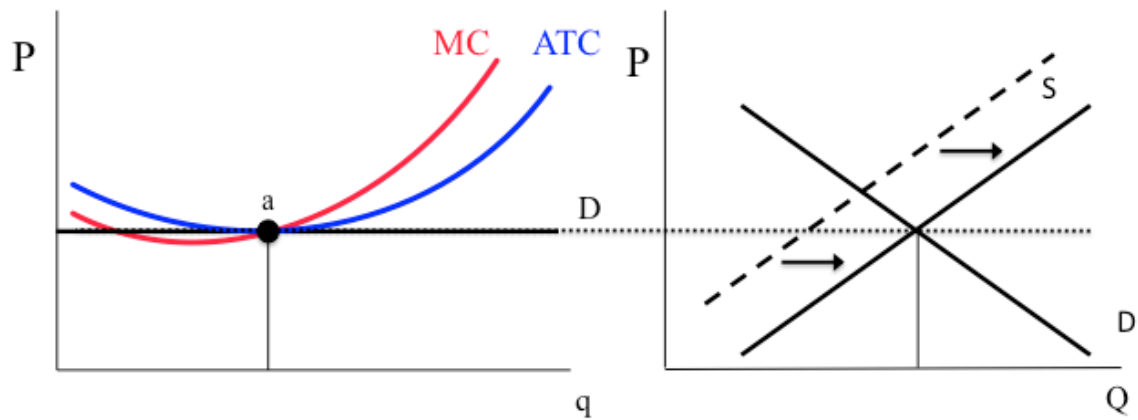
between points *a* and *b* represents the per-unit profit. The right panel shows the market, in which the supply curve is the aggregation of all firms' marginal-cost/supply curves. *Q* designates the market quantity, which generally can be an order of magnitude larger than *q*, the output level of a single firm.



(*D* = demand curve, *Q* = market quantity, *S* = supply curve, *q* = output of a single firm)

Figure 2. Short-run equilibrium of a perfectly competitive market. The left panel shows an individual representative firm, and the right panel shows the market.

In the long run, new firms are attracted to the industry by positive profits.¹¹ Supply expands (the supply curve shifts right), market price falls, and economic profits are no longer earned. Assuming that there are no other relevant changes in the industry, $P = MC$ and $P = \min(ATC)$, where *min* represents the minimum point of the *ATC* curve. This is shown in Figure 3; the left panel shows the representative firm, and the right shows the market.



(*D* = demand curve, *Q* = market quantity, *S* = supply curve, *q* = output of a single firm)

Figure 3. Long-run equilibrium in perfect competition

In the long run, perfect competition delivers two types of efficiency: productive and allocative. Productive efficiency means there are no alternative means of producing the given level of output that would achieve lower cost. Allocative efficiency means that resources are allocated to

¹¹ See Appendix for description of accounting versus economic profit.

activities that society most values because $P = MC$. If $P < MC$, then the cost of the marginal unit exceeds its value; conversely, if $P > MC$, then the cost of marginal production is less than its marginal value, and thus production should increase. Economic efficiency is the combination of allocative and productive efficiency, and perfect competition achieves both.

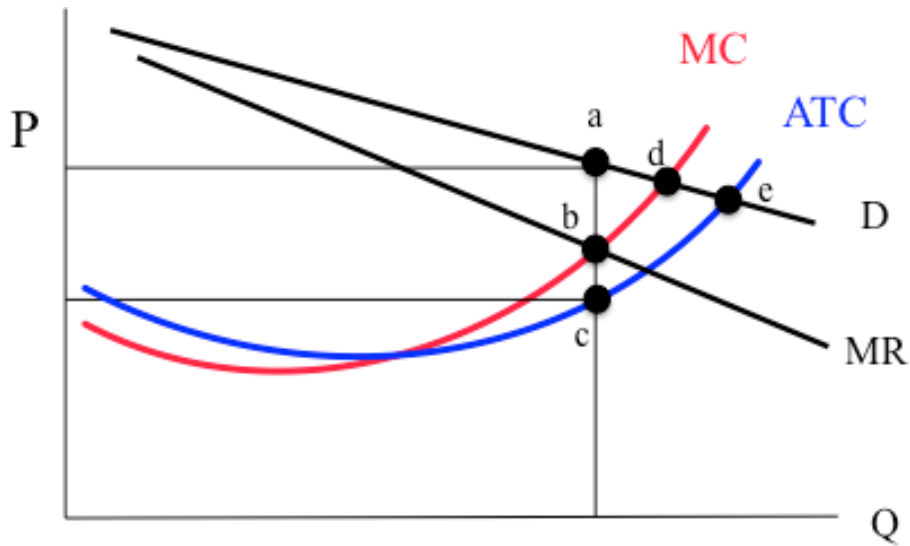
2.2 Monopoly

A monopoly is the polar opposite of perfect competition. Instead of having many buyers and sellers, a monopoly has a single seller and therefore unlimited market power. In a pure monopoly (which is unlikely to exist in pure form), there are significant barriers to the entry of new firms that would introduce competition to the single seller. These barriers might be economic—such as very large economies of scale—regulatory, or a combination. There are no products that are a close substitute for the monopolist's product. In practice, most developed and developing countries have some type of regulatory structure that prevents pure monopoly pricing. The pure monopolist maximizes profits by producing a quantity of output for which marginal revenue equals marginal cost. Economic profits exceed zero, and because there are significant barriers to entry, these excess profits persist in the long run.¹⁰ One important example is the regulated monopoly structure of the electric power system, which has undergone change in many parts of the world in recent years.

Figure 4 depicts the optimal output decision of a monopolist. Production occurs at the point at which $MC = \text{marginal revenue (MR)}$. If $MC < MR$, then the monopolist could produce one more unit of output at a marginal cost that is lower than the additional revenue it could earn, and it would therefore increase output. If $MC > MR$, then the additional revenue obtained from the last unit of output is less than the cost of producing this unit, and therefore the monopolist would decrease output. In the diagram, $MC = MR$ at Point b, and the monopolist can charge the price at Point a. The vertical distance between points a and c represents the per-unit profit earned by the monopolist. Because new firms cannot enter a perfect monopolistic market, the equilibrium depicted in the diagram persists into the long run.

The monopolist outcome features $P > MC$, and absent some regulatory framework, this price will persist. Using this simple graphical framework, we can explore the implications of alternative pricing rules that are enforced by a potential regulator. If the regulator were to impose the condition that $P = MC$, this would be represented by Point d on the graph. Because the price still exceeds ATC, the monopolist would continue to earn an economic profit although at a level less than in the unregulated case. However, note that at Point d $MC > MR$, and thus the monopolist could be persuaded only to produce at Point d under regulatory duress because it is losing money under the regulator.

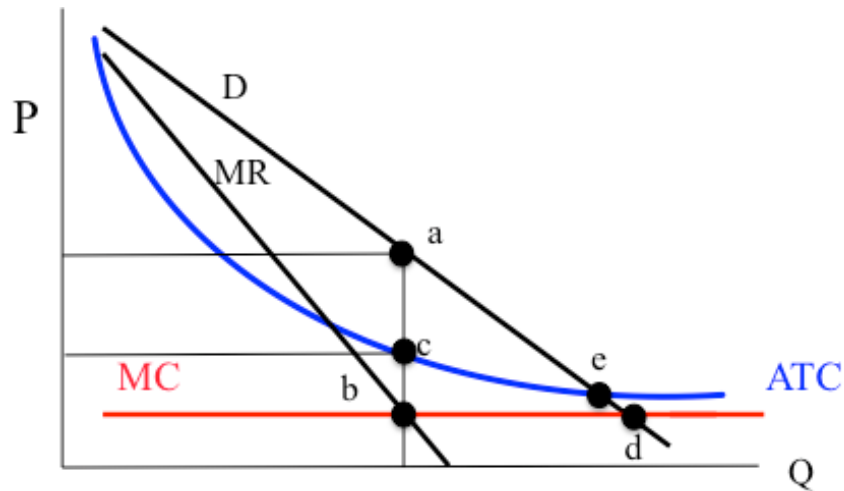
An alternative regulatory approach would be to set $P = ATC$. This would eliminate economic profit, but it would allow for the level of accounting profit that would elicit the necessary investment to support the firm. This is represented by Point e in the diagram. It is clear that $MC > MR$ and that this is not an optimal point as viewed by the monopolist, which would prefer to produce at the point where $MR = MC$.



(D = demand curve, Q = market quantity)

Figure 4. Monopolistic market equilibrium

The monopolist depicted above can be considered the “classic” monopolist, which is covered in standard economic textbooks. Through a large range of output levels, the ATC declines, but ultimately it increases at an output level that is less than optimal. An alternative cost structure for a monopolist is persistent economies of scale at, and above, equilibrium output levels. Figure 5 provides a graphical example in which the marginal cost is constant at all levels of output. This diagram can easily be modified to represent a firm with zero (or near-zero) marginal cost, such as an electronic book publisher, software publisher, or satellite communication service. Optimal output for this monopolist is at Point b, where $MC = MR$. Economic profit is positive and has a per-unit value of the distance between points a and c. If a regulator were to impose the condition that $P = MC$, the firm would produce at Point d. At this point, $P < ATC$, and thus this firm would lose money and eventually go out of business. Bankruptcy could be avoided in this case only if there were some form of subsidy. Alternatively, the regulator could require $P = ATC$, which results in market output at Point e. Although the firm would lose money on the output that exceeds its preferred level, overall the firm would break even, and thus it could remain in business without a subsidy. In this case, however, $P > MC$. Thus, without regulatory oversight, the price would exceed the marginal cost. With regulation that requires marginal-cost pricing, this firm would go bankrupt.



(D = demand curve, Q = market quantity)

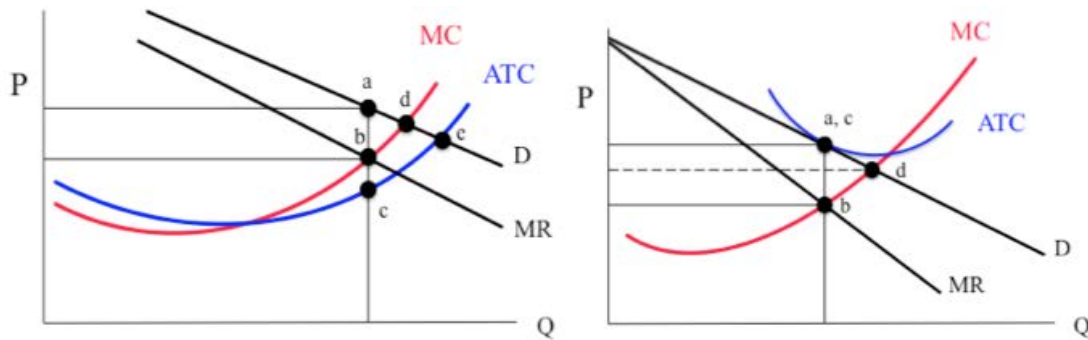
Figure 5. Monopolist with declining average total cost during the relevant market range of output

2.3 Monopolistic Competition

Between these extremes are monopolistic competition and oligopoly. Monopolistic competition is characterized by a relatively large number of firms that produce products that are similar to each other but consumers differentiate among products of different sellers. Examples of monopolistic competition include canned foods, sporting goods, soft drinks, and clothing industries. Other assumptions are similar to those of perfect competition; there are few barriers to entry/exit, and firms might have a relatively small but discernable impact on price. Thus, if economic profits are being earned in the short run, new firms will enter the industry, expanding supply.

In the long run, $P = ATC$, but this will not occur at the minimum ATC; thus, the long-run equilibrium is not efficient because there is some excess capacity. In the long run, $P > MC$. Figure 6 shows short-run and long-run outcomes. In the short run, the firm maximized profits by producing at the point where $MC = MR$, denoted as Point b on the left. Per-unit profit is the length of the segment from points a to c, the amount by which $P > ATC$. If a regulator required marginal-cost pricing, the firm would move to Point d, and it would continue to earn economic profit, but less than before. It would thus lose money and therefore have no incentive on its own to produce at Point d. Similarly, if a regulator required $P = ATC$, the firm would be at Point e, earning no economic profit, but it could remain in business without incurring economic losses.

If the firm is unregulated and therefore earning positive economic profit in the short run, the relatively low barriers to entry will allow new firms to enter the market. Similar to the supply curve shift in the right panel of Figure 3, the equilibrium price falls in the market. The volume of output that the representative firm can sell will therefore fall, resulting in a new, lower demand and MR curve, as shown in the right panel of Figure 6.



(D = demand curve, Q = market quantity)

Figure 6. Monopolistic competition in the short run (left) and long run (right) for a representative firm

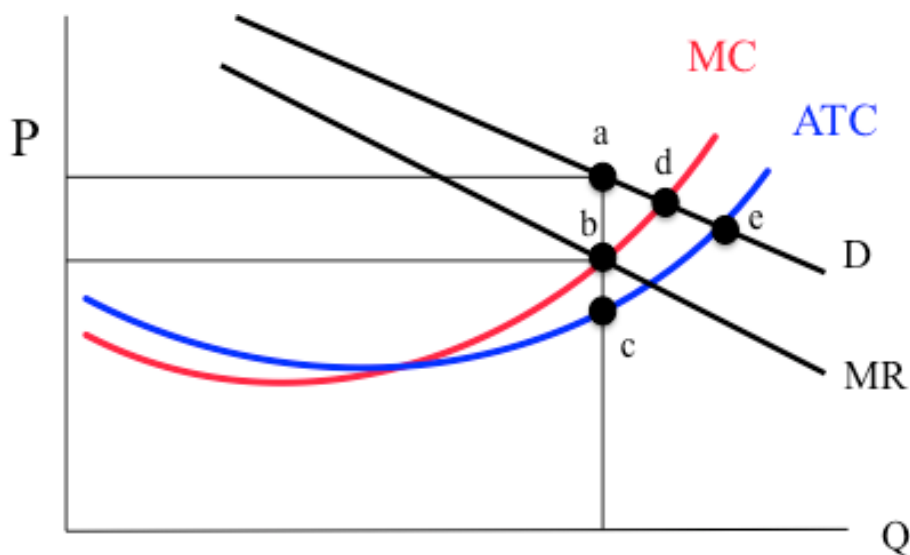
The firm continues to maximize profits, and therefore it produces the level of output at which $MC = MR$. But as the right panel shows, the point of profit maximization occurs when the ATC curve is tangent to the demand curve, and thus economic profit is zero in the long run. This means that this industry will have average-cost pricing in the long run, and, as shown by the diagram, $P > MC$. If a regulator requires marginal-cost pricing, the firm moves to Point d, at which it loses money and will not be able to remain in business. The revenue shortfall could be remedied by some form of subsidy; otherwise this is not a viable position in the long run. Note that in the absence of a regulator, this industry will have excess capacity—more than needed to fulfill market demand. Recall that this diagram is for a representative firm in the market, and thus there are many similar firms with similar diagrams and outcomes. At Point a on the right side of the diagram, the firm has not minimized ATC. This means that there is excess capacity not being used. To determine the excess capacity in the industry, one could aggregate the excess capacity of each firm. This problem cannot be solved by marginal-cost pricing, unless a subsidy makes up for the revenue shortfall.

2.4 Oligopoly

An oligopoly has a few large sellers. Products are differentiated among producers, and each seller has considerable but not unlimited market power. The number of sellers might vary, but it can be as small as 2 and is often as many as 10, although there is no limit. The actions of any individual seller impact other sellers, and thus there is a degree of interdependency among sellers. (This dependency is generally unwilling, unless there is collusion to coordinate price or quantity decisions, which is illegal in most parts of the world.) Predicting the market outcome for an oligopoly is difficult, and to the casual observer “virtually anything can happen” (Scherer 1970). In fact, a number of economic models have been constructed to explain the behavior of an oligopoly, ranging from the kinked demand curve to game theory and even agent-based modeling. The result is that the price-quantity outcome of this type of market cannot be analytically determined, as it is in others. In some cases, “ruinous competition,” caused by aggressive price cuts in an effort to force out competition, might result in an economic outcome where $P = MC$, as shown below; however, in this outcome, either at least one producer will eventually go out of business or all producers will gradually increase prices to avoid such a fate. At the other extreme, the outcome might approximate the monopolist outcome if firms collude, either explicitly (illegal) or implicitly. See Scherer (1970).

The oligopoly market outcome is a function of the degree of market power and the behavior of the firms, and it is the motivation for the many types of oligopolistic models. Because the focus of this paper is on marginal-cost pricing, we ignore the large literature on these models and instead look at potential pricing relative to marginal cost.

Some oligopolistic markets might have a relatively large number of firms, reducing the direct interdependence among individual firms. With more firms, one would expect less market power and potentially limited economies of scale. In such cases, a typical firm might have cost curves that resemble those of monopolistic competition, shown below in Figure 7. The curve structure is the same as the short-run monopolistically competitive market; however, with oligopoly there are significantly higher barriers to entry. Thus, this diagram represents the best possible outcome for an oligopolist, and there could potentially be long-run equilibrium if its competitors choose not to engage in competitive price reductions.



(D = demand curve, Q = market quantity)

Figure 7. One potential diagram of oligopoly with economic profit

Conversely, in the case where there are limited barriers to entry, new firms could enter the market. Eventually, this would result in an outcome similar to that of long-run equilibrium in a monopolistically competitive market, as depicted in Figure 6.

When there is a strong profit motive, oligopolists might attempt to gain market share from competitors by cutting price. Various game-theoretic approaches can be applied to determine potential outcomes. Regardless of the behavioral model used, we can predict a range of outcomes. Given aggressive price-cutting and retaliation, price can be driven so low that one or more firms go out of business. For that to happen, $P < ATC$, and this unsustainable position might result in an increase in market share and market power from the survivors. Thus, marginal-cost pricing should not be expected in the case of oligopoly.¹²

¹² Examples of alternative oligopolistic models include the kinked demand curve model and various game-theoretic approaches that include Bertrand models, Nash models, and others.

Oligopolists with zero/low marginal costs that do not increase can be represented by Figure 5. Because $MC < ATC$ and because marginal cost is constant (during the relevant range of output), marginal-cost pricing cannot be sustained because the firm will lose money absent some type of subsidy.

2.5 Market Outcomes

Table 1 summarizes some of the key outcomes of the four basic market types. Note that only one of these—perfect competition—features marginal-cost pricing. Because perfectly competitive markets are rare, industries where price equals marginal cost are also rare. This is an important concept for our discussion in subsequent sections on marginal-cost pricing in electricity markets.

Table 1. Comparison of Key Outcomes for Different Market Types

	Market Power	Number of Sellers	Product Differentiation	Short-Run Price	Long-Run Price	Long-Run Efficiency
Perfect Competition	None	Very large	None	$P = MC$	$P = MC$	Yes
Monopolistic Competition	Some	Large	Some	$P > MC$	$P > MC$	Excess capacity
Oligopoly	Considerable	Few interdependent	Some, possibly significant	Likely that $P > MC$ but indeterminate	Likely that $P > MC$ but indeterminate	Unlikely
Monopoly	Maximum	One	NA	$P > MC$	$P > MC$	No

3 Market Structure Complications

The previous section described the theory of the four basic market types; however, in reality these are much more complex than they are in theory. In practice, the boundaries between these market types are often blurred. In some cases, a dominant near-monopolist seller has some limited competition from a large number of small firms that are wholly dependent on the monopoly's behavior. In other cases, there could be an oligopolistic industry and many small companies. In both cases, this so-called “competitive fringe” acts like a perfect competitor, but there is a lead firm or group of firms that dominate market power. For details, see Scherer (1970).

In this section, we describe key complicating factors in real markets. These topics will be particularly important for the discussion in Section 4 on electricity markets. First, we describe the presence of near-zero marginal costs in several real-world industries and how those companies price products to recover costs. This is directly relevant to the conversation on marginal pricing of energy in electricity markets. We then describe key causes of market failure, which can lead to price distortions and other inefficiencies in any market. These causes will be revisited in our discussion on electricity market distortions in Section 4.

3.1 Near-Zero Marginal-Cost Examples

Several notable industries have very high fixed costs and near-zero marginal costs. These include cell phone companies and software. According to economic theory, these large up-front capital costs create barriers to entry, which limit the number of participating firms, resulting in market power.¹³ For example, although data is not publicly available, it reportedly cost Microsoft approximately \$5 billion to develop Windows Vista, but the marginal cost to produce one additional copy of the software is essentially zero, especially if downloaded from the Internet (Hill and Jones 2009). Vista sold for approximately \$200–300 per copy, depending on the package type. This price is significantly larger than the marginal cost, which by definition indicates that Microsoft displayed market power. It would be difficult or impossible to find a perfectly competitive market with zero marginal cost that could survive. Markets for software or books that are distributed electronically generally have high fixed costs and low variable costs, and therefore they are unlikely to feature marginal-cost pricing.

Another example of pricing comes from the U.S. cell phone industry. This industry could be characterized as an oligopoly with competitive fringe, with an HHI of nearly 2,800 and the top four firms dominating nearly 99% of the market share. Table 2 shows the HHI and rankings by customers/subscribers for the top seven U.S. firms based on data from Dano (2016). Company cost information is difficult or impossible to obtain; however, Table 3 shows the calculated

¹³ Some argue that, unless barriers to entry are *created* by one of the suppliers in order to raise the market price, barriers to entry are not in and of themselves a cause of market power (Stoft 2002). Barriers to entry, such as large up-front costs to build a generator, have an effect similar to market power by resulting in higher prices; however, barriers to entry also yield a capacity shortage—the opposite effect as withholding in the case of market power. The economically efficient higher prices resulting from barriers to entry will signal new deployment to meet the capacity shortfall, provided the barriers are not too significant. In practice, however, long-term planning efforts by load-serving entities aim to avoid such a capacity shortage caused by barriers of high cost and long project lead times.

consumer cell phone plan average price per GB/month and the marginal price paid for each additional GB (Verizon 2016).

Table 2. Example of Market Power: Communications

	Subscribers (Millions)	Percentage of Total	HHI
Verizon	140.0	35.4	1,251.1
AT&T	129.0	32.6	1,062.3
T-Mobile	63.0	15.9	253.4
Sprint	58.0	14.7	214.7
U.S. Cellular	5.0	1.3	1.6
Shentel	0.5	0.1	0.0
nTelos	0.3	0.1	0.0
Total	395.8	100.0	2,783.1
Four-Firm Concentration Ratio		98.5	
Five-Firm Concentration Ratio		99.8	

Table 3. Pricing Example from Communications Industry

Price (\$/month)	GB/month	Average Cost	Marginal Price
35	2	17.5	17.5
50	4	12.5	7.5
70	8	8.8	5
80	12	6.7	2.5
90	16	5.6	2.5
110	24	4.6	2.5

To Verizon, the marginal price is not the same as the marginal cost of an additional GB of data usage; however, the pricing pattern shows that it is likely that the first block of data download might be covering at least some (or all) of the fixed network cost, with declining effective average pricing that accounts for low (if any) marginal cost per GB. This type of pricing is sometimes called declining block pricing, which was prevalent in the United States for many years for electricity but is now discouraged because it provides an incentive to pay for higher usage levels.¹⁴

¹⁴ Key drivers for this trend away from declining block rates were the Public Utility Regulatory Policies Act of 1978 and subsequent amendments, which required states to examine the appropriateness of rate standards supporting any one of the act's purposes (Dillon 1979; National Action Plan for Energy Efficiency 2009).

It is common to find industries that do not price at marginal cost. A multitude of pricing approaches include subscription rates, entry/exit costs, discounts for large quantities (a form of declining-block pricing), and many others.¹⁵ The concept of marginal-cost pricing comes primarily from perfectly competitive markets, but it is not a common pricing method. This is because nearly all markets are not perfectly competitive, and thus they have some level of market power, product differentiation, and/or some type of market failure. Coupled with the fact that reliability, both in the planning and operational time frames, is not part of an electricity market per se, there is no reason to expect that marginal-cost pricing on its own will result in a sustainable market with revenue sufficiency that also meets the reliability objectives.

3.2 Market Failures

Another important concept in economic theory and practice is market failure. This refers to a situation where the market fails to most efficiently allocate resources. The presence of market failures can justify government intervention. Market power, public goods, and externalities are key textbook examples of market failures.

3.2.1 Market Power

As previously mentioned, market power refers to the ability of a market actor to lower the market efficiency. When sellers exercise market power, they alter prices to levels that are both inefficient and profitable. Market power consists of economic withholding (raising the price of output) and quantity withholding (reducing output or altering the terms of trade) to adversely affect customers (Stoft 2002). Market power typically results from a lack of adequate competition due to barriers to entry, economies of scale, or other factors.¹³

When there is significant market power, economists often refer to this as market failure. Market power is generally an issue that the U.S. Justice Department mitigates. For example, mergers among firms with significant market power are examined by the U.S. Justice Department and are generally not challenged unless the resulting HHI exceeds 1,800. Industries that might be susceptible to challenges include banking, airlines, cable television, communications, and others. In these cases, the definition of the market is more important than the HHI.

Some recent evidence suggests that, despite well-established antitrust institutions that are intended to mitigate anticompetitive conduct, many U.S. industries appear to be experiencing a decline in competition, which corresponds to an increase in market power, as evidenced by increasing profit rates, relatively low entries of new firms, and efficiencies of scale (Council of Economic Advisers 2016). The issues surrounding competition and market power are further discussed in a recent Economic Report of the President (Executive Office of the President and Council of Economic Advisers 2016), which suggests that an increase in market power is driving greater income inequality. Some suggest that the broader implications of such growing market power could extend beyond individual markets to result in slower overall economic growth and economy-wide inequality (Baker 2017).

¹⁵ For other examples, see <https://energyathaas.wordpress.com/2016/06/13/is-electricity-pricing-different-from-real-markets-should-it-be/>.

3.2.2 Public Goods

Economic theory defines common attributes of products. Private goods have two critical properties: excludability and rivalry. The former means that a seller can exclude potential buyers from consuming the product without paying. The latter means that the product is not sharable—i.e., if a given product is purchased for consumption by Buyer A, then Buyer B must purchase/consume his/her own product.¹⁶ Public goods, by contrast, do not possess one, or both, of these properties. For example, a streetlight might be paid for by a single resident, but anybody who is nearby might benefit. The light from the lamp is thus shared, and it is impractical to extract payments from all who walk or otherwise travel near enough to benefit from the light.¹⁷

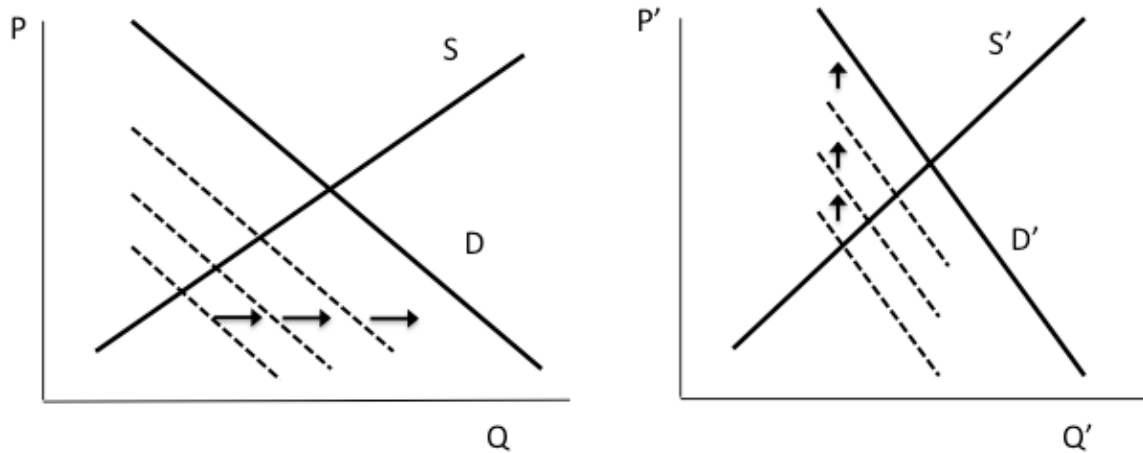
Calculating the market demand for a public good is fundamentally different than that for a private good. Because private goods are rival goods, the summation of each individual's quantity demanded at each price yields the market demand curve. Conversely, the demand curve for a public good is calculated by adding the price that each consumer is willing to pay for each level of output. This is illustrated in Figure 8.¹⁸

The shared nature of a public good means that, for a wide range of supply, the marginal cost is zero. The marginal cost of providing a park is zero for an additional user up to the point of sufficient congestion that renders the park unusable or unenjoyable. In practice, goods might not be entirely private or public; they might lie on a continuum between these attributes.

¹⁶ A good might be sharable in the sense of being divisible, but the standard use of the term *sharable* in this context means that the consumption of one unit of a good by one consumer implies there is nothing left for another consumer.

¹⁷ One could argue that this example of a street light is closer to a *club good*, which is nonrivalrous but excludable, because the light is a local object that typically only benefits residents in that area. In this paper, we assume the term *public good* applies to items for which at least the localized users of that good are not excluded from that good, though outsiders may not be strictly excluded or unimpacted, i.e., public lighting in this case, or as we discuss in Section 4.1, the system-wide reliability of an electricity network.

¹⁸ The number of gallons of gasoline demanded by each consumer at each price is another example of a private good. The “market” demand for a street light would be calculated by summing the amount that each consumer would be willing and able to spend for one streetlight, for two streetlights, etc.



(D = demand curve, Q = market quantity, S = supply curve)

Figure 8. The derivation of the demand curve for a private good (left) and a public good (right)

Because the usual market mechanism does a poor job of determining the efficient level of output of public goods, this is considered a market failure. Because of the market's inability to function properly, public goods are typically provided by government at state, local, or national levels. Examples include parks, street lights, national defense, and justice. Interested readers can consult standard economic texts such as Gruber (2016).

3.2.3 Externalities

Another market failure occurs when costs or benefits from a market transaction impact someone who is not a party to the transaction. This impact can be either positive or negative, and in the latter case a cost is imposed on the bystander. The most relevant example for this paper is emissions, which can cause damage via human health impacts or contamination of soil, water, and/or air.¹⁹ Solutions to the market failure posed by external costs are well known and easy to implement in principle. Once estimates are developed of the damage cost, these costs can be included in the cost of production by developing a market for emissions permits, or economic penalties per unit of harmful emissions can be assessed. In principle, either of these outcomes will produce the same result; and, if done properly, they can eliminate the market failure caused by the externality. In practice, there can be some combination of practical and political obstacles. If some or all of the externalities were to remain, this would cause a deadweight loss to society.²⁰

¹⁹ We do not explore positive externalities in this paper, but they are covered in standard economics text books.

²⁰ Deadweight loss is the decrease in consumer surplus and producer surplus that results from an inefficient level of production; it is a social loss, i.e., a cost borne by the entire society and not just the consumer or producer.

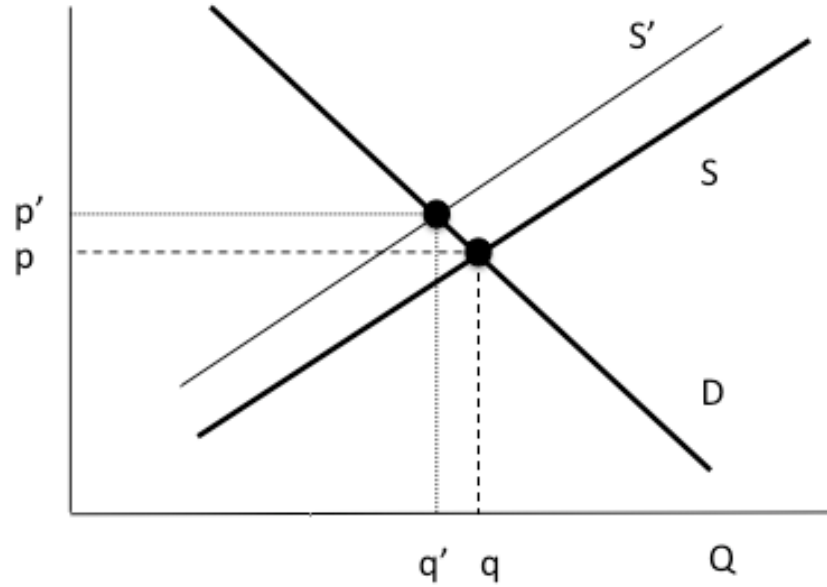


Figure 9. Negative externalities are not taken into account in the market supply curve, S , and equilibrium quantity, q , and price, p . Supply curve S' and associated p' and q' show the impact of the externality in the market outcome.

4 Electricity Markets

Electricity is an unusual product. It is not consumed for its own sake and therefore has limited, if any, intrinsic value. It has value because of the things it can do when coupled with various services, such as lighting, powering factory equipment, heating and cooling, and entertainment. The bulk power system in most developed and many developing countries consists of a large interconnected grid. Interconnected power systems have advantages and disadvantages. One benefit of these systems is that capacity and energy can be shared among regions, lowering total costs (King et al. 2012; Jordan and Piwko 2013; Milligan et al. 2013). Also, because they are large, they can withstand a variety of events that might otherwise take down a smaller system; however, because of the interconnected nature of the system, a significant event has the potential to collapse the entire system. Electricity cannot be stored economically at large scale or for long durations, although it is possible to store various fuels, including water behind a dam, for varying lengths of time.

Competitive wholesale electricity markets have evolved in many parts of the world, generally transforming from regulated monopolies that consisted of vertically integrated companies—utilities that provide electricity at the retail level after generating and transmitting the energy via facilities owned by the company. As discussed in Section 1, in these regulated systems, electricity is priced based on the cost of service rather than on a competitive market. In most parts of the world, some form of regulation or public ownership was thought to approximately solve concerns about market power and excessive reserve margins. Regulators promote the adoption of least-cost resources, subject to reliability constraints, in planning and operation.

The concepts of marginal cost were, and are, pervasive in utility economics. As markets began to replace parts of the regulated monopoly, the confluence of the desire for a competitive market outcome and the long tradition of least-cost dispatch based on marginal cost together resulted in market regulators adopting marginal cost as the basis of determining bids and offers in wholesale power markets. Price formation is based on marginal cost, but many factors can contribute to de facto deviations (Ela et al. 2014; Sauer 2014; Ela 2016).

4.1 Electricity Market Inefficiencies

Today's restructured electricity markets have several contributors to market failure. These attributes include externalities, public good attributes of the network, ineffective demand curve, market power, wholesale price caps, and lack of large-scale storage. Even regulated markets are likely to have at least some level of failure because it is difficult or impossible to fully eliminate these attributes.

Each of these six market attributes represents a different dimension of the market failure problem, which is illustrated in Figure 10. A perfectly competitive market for a pure private good with no externalities would be represented by a single point at the intersection of the various axes. A monopolistically competitive market for a pure private good and no externalities would have a single dimension of market failure, and it would be located somewhere along the upward vertical axis. Electricity markets would presumably have a nonzero location on each axis, depending on the empirically measured level of each attribute's contribution to market failure. Although electricity has some unique aspects, the contributors to market failures depicted

in the figure are present in other markets. In the remainder of this section, we further discuss each of these attributes within the context of electricity markets.

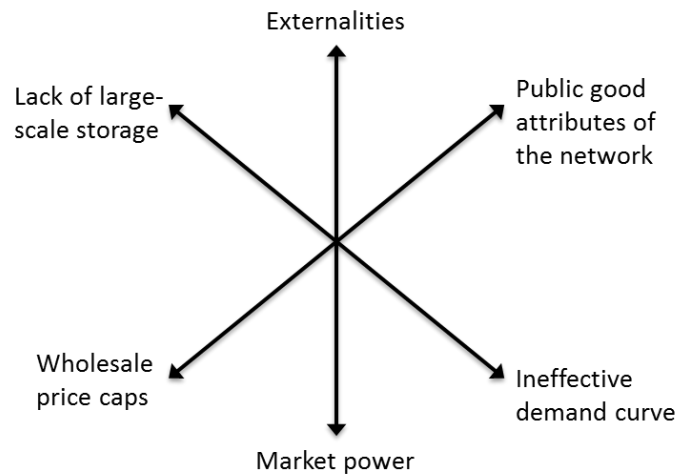


Figure 10. Stylistic representation of multiple contributors to market failure in electricity markets

First, externalities are a well-understood source of market failure, which exists in today’s wholesale electricity markets. Externalities in the production and transmission of electricity are well known. The primary negative externality is pollution, which is often regulated by government, including the U.S. Environmental Protection Agency. In principle, regulations can include a combination of limits, markets for tradeable permits, or imposing costs on pollution. Mitigation might not be perfect, but it sends signals to emitters and alters the level of externalities produced. In the United States, there are markets for sulfur-oxides emission permits, which aim to incorporate social costs into private costs. Its goal is to result in the socially optimal level of sulfur oxides; the effect is similar to what is depicted in Figure 9. At the time of this writing, there are no U.S.-wide markets for carbon dioxide or nitrogen oxides, though several regional markets do exist. To the extent that these externalities have not been corrected, some degree of market failure arises.

The second attribute—public good attributes of the network—is another classic cause of market failure. Houldin (2004) argues that the electricity network has attributes of a public good; either “all loads on a circuit consume (electricity) or none do.” It is also difficult to deliver different levels of reliability to two or more consumers on the same grid. If one customer decides to pay more for a higher level of reliability, all other customers on the same network will receive more reliable electricity whether or not they pay the increased cost. Therefore, reliability is nonexcludable. It is also nonrivalrous because the increase in network reliability can be consumed by everybody on the network. This implies some degree of market failure because of the public good attributes of the electricity network.

Arguably the most impactful contributor to wholesale electricity market failure—ineffective demand curve—is rarely mentioned in standard economic textbooks because it is fundamentally critical to all markets (Cramton et al. 2013, 27–46). This feature of current electricity markets results from the “demand-side flaws” of demand inelasticity and a lack of real-time physical enforcement of bilateral contracts (Stoft 2002). It is further reflected by the misalignment of

wholesale electricity prices with the average-cost-based prices paid by consumers at the retail level. When prices rise in a perfectly competitive market, perhaps as a result of shortages, some consumers withdraw from the market because they are not willing to pay the higher price. But when the price seen by the consumer is established by regulatory-approved tariffs, as is the case in most (if not all) retail electricity markets in the United States, the effective demand curve seen by the wholesale market is essentially vertical because consumers never see the true price.²¹ The lack of an effective demand curve rules out the possibility that the market can function normally, implying that out-of-market objectives are needed to ensure resource adequacy.

The fourth contributor to market failure is market power, which is sometimes considered a direct market failure (see Section 3.2.1 and footnote 13) and exists in part because of significant barriers to entry that yield monopolistic tendencies. Since the formation of large wholesale electricity markets around the world, a large body of work has been directed toward assessing and mitigating market power. In the United States, the Federal Energy Regulatory Commission (FERC) oversees interstate bulk electricity markets, with a key objective to minimize market power in wholesale markets by ensuring just and reasonable market rules and tariffs. Within this oversight, the potential for electricity suppliers to exercise market power is monitored and mitigated through (1) a combination of independent internal market monitoring units within each RTO/independent system operator and an independent external market monitor, (2) an independent external market monitor alone, or (3) mitigation protocols codified by regulatory bodies such as public utilities commissions (FERC 2014).

Because market power is a function of real-time demand, transmission congestion, planned maintenance, forced outages, and other factors that vary through time and location, it is difficult to quantify market power generally across the U.S. power grid. Nevertheless, two main types of approaches are used to identify and mitigate potential market power (FERC 2014).²² The first is called “structural,” which subjects resource offers to mitigation when prespecified conditions are met. An example is the three pivotal supplier test, which determines if out-of-merit resources are needed to relieve binding transmission constraints. The second type of approach is called “conduct and impact,” which looks at the resource offer amount and then assesses the impact of that offer on the market clearing price for energy. In both approaches, the mitigation action replaces the resource offers with precalculated reference levels (or some variant of them), usually based on the short-run marginal cost. Throughout the literature, more sophisticated methods for assessing market power have been proposed, such as explicitly evaluating the impact of market mergers or rate changes by simulating prices with quantitative market models (Helman and Hobbs 2010, 1,434–1,448).

The fifth contributor to electricity market failure is wholesale price caps, which are common measures to minimize the effects of market power in wholesale electricity markets and also substitute as an administrative demand curve in place of part of the wholesale demand curve. The

²¹ We differentiate between notional demand curves and effective demand curves. Notional demand curves are derived from consumers’ declining marginal utility, and effective demand curves actually function in the market. In the case of electricity, the effective retail demand curve is vertical or near vertical because the market price is prescribed by a tariff. With fully responsive demand, the notional demand curve would become the effective demand curve or very close to it.

²² At the time of this writing, FERC is seeking to potentially modify the methods used to assess market power (FERC 2016a).

result is that prices are constrained from rising above this limit during times of congestion or shortage. This is a well-known feature of wholesale markets, and much has been written about its impact—for example, Hogan (2005) provides a good discussion on this along with energy-only market design.

Finally, the sixth attribute of electricity that contributes to the market failure problem is the current inability of electricity to be stored in large quantities or for long periods of time.²³ This means that the traditional methods used by sellers (and buyers) in other markets—such as warehouse/storage as an arbitrage mechanism—are currently not available to electricity providers. Such warehousing serves as a buffer between oscillating supply and/or demand and can help stabilize prices. This feature of electricity is not a market failure in the economic sense of the term, but it does play a role in price formation in bulk electricity markets, thereby exacerbating many of the underlying market inefficiencies discussed above.

A common thread among these six wholesale electricity market features is imperfect information, which includes information asymmetries and uncertainty. For example, the lack of ubiquitous smart meters to receive price signals and in turn for demand to respond leads to an asymmetric flow of data that, in addition to underlying uncertainties in how load will respond in both the planning and operational timescales, contributes to the ineffective demand curve feature above. Storage helps to mitigate this dynamic, but the lack of large-scale storage in most operating areas severely limits the extent of its supply-demand buffering benefit.

To summarize:

- Externalities are currently unresolved.
- The electricity network has at least some elements of a public good.
- Retail electricity markets do not have a functioning effective demand curve.
- There might be some degree of market power in many markets, and the markets are generally unlikely to be perfectly competitive.
- Wholesale electricity markets typically have price caps that are binding during scarcity events. Electricity currently cannot be stored in large quantities for long periods of time in most operating areas.
- Taken as a whole, these observations imply that:
 - Marginal-cost pricing might not result in an economically efficient outcome.
 - Prices have a limited ability to signal shortages and do not have any effective impact on consumption.
 - Any one of these contributors to market failures will disrupt the market relative to a well-behaved competitive market; multiple contributors will result in additional disruption.

²³ Except for a few limited cases of pumped storage that can provide large quantities of seasonal storage, such as Norway and Quebec.

Price distortions and market inefficiencies resulting from the above market failures can potentially lead to situations in which current or new generators needed for reliability do not have sufficient opportunity to recover both their marginal and fixed costs from the electricity market. This is the revenue sufficiency problem, which is discussed in Section 4.2. Note that thus far we have ignored other aspects of the bulk system constraints, including stability, start-up costs, and the resulting out-of-market payments, yet we still find that marginal-cost pricing will likely (1) lead to inefficient outcomes and (2) cause revenue insufficiency.

4.2 Revenue Sufficiency

Revenues are sufficient when payments for the individual services required to maintain a reliable grid—including energy, capacity, and flexibility—cover the fixed and variable costs of providing those services. Revenue insufficiency—often called the “missing money” problem—results when payments are not adequate to (1) cover both fixed and variable costs incurred by existing generators (a necessary condition to remain in the market) and/or (2) justify investments in new capacity, in either case for resources that are needed for reliability.

Reliability includes both the investment and operational time horizons. The investment portion, referred to as resource (or capacity) adequacy, focuses on deploying sufficient installed generating capacity to ensure a very small probability, size, and duration of blackouts caused by insufficient capacity. Here, we focus on this investment-period objective, but, as we discuss later, this is not entirely separate from the operational timescale. Certain operational attributes within capacity—for example, flexibility—can be as important to ensuring long-term system reliability as the quantity of capacity itself. Resource adequacy is discussed further in Section 4.3.

Existing solutions to the revenue sufficiency and associated resource adequacy problem defined above generally fall into three categories: (1) capacity market or capacity payments, (2) supplementing the energy-only market with a reserve product and scarcity pricing, and (3) power purchase agreements or other contracting approaches. Each of these points is discussed below, and additional examples of hybrid approaches are provided in Section 4.5. These options are a means to compensate generators that provide services needed by the grid that are not explicitly incentivized through the day-ahead and real-time markets (Newbery 2016, 401–410). There is no wide consensus on which, if any, of these options are most effective to correct the underlying market failures and inefficiencies that distort prices so that prices can correctly reflect the desired and needed level of system reliability. Each of these options necessarily involves some form of administrative or out-of-market actions because, as discussed in Section 4.1, electricity markets will never be able to function as perfectly competitive markets because of their inherent attributes of market failure: externalities, public good attributes of the network, ineffective demand curve, market power, wholesale price caps, and lack of large-scale storage. The ideal solution for each operating area will likely depend heavily on the existing system portfolio, infrastructure, and market rules.

The first category of solutions to revenue sufficiency and resource adequacy is mandatory capacity remuneration mechanisms (CRMs), often referred to as capacity markets. A significant body of literature advocates for some form of capacity-based payments that creates a separate revenue stream for generators outside of energy and ancillary service markets to ensure system reliability (Jenkin et al. 2016; Milligan et al. 2016; Bushnell et al. 2017; Hogan 2017, 55–61). CRMs provide a direct signal for investors to build new capacity when and where it will be

needed at a future date to meet load in peak periods, with penalties if the developers fail to deliver. Different payment methods and structures are used in practice, including specifications regarding price caps and floors, the time horizon used to assess capacity needs, the qualification and compensation criteria for various resources and ages, the shape of the demand curve for capacity, the cost basis for the capacity, the alignment of state-level policies within multistate regional markets, market design aspects to ensure resiliency (e.g., polar vortex), and approaches to minimize market power. FERC has held a series of technical workshops to address and discuss these and related issues on capacity markets (FERC 2013b; FERC 2017).

A key debate with capacity markets is whether and how they can fully capture resource adequacy aspects, including the need for specific attributes of that capacity, such as flexibility. A wide range of approaches have been discussed for CRMs to achieve the ideal level of flexibility, including (1) a more traditional forward capacity market in which investors self-optimize for the level of flexibility capability within the new capacity, (2) a capacity market with a uniform fixed flexibility requirement for all new capacity, and (3) a capacity market with tiered levels of flexibility requirements. Examples of the first option include the existing capacity markets in the United States, such as PJM, Independent System Operator New England, and New York Independent System Operator. Flexible capacity requirements of either type (items 2 and 3 above) are uncommon. One example is the California Independent System Operator (CAISO) wholesale market in the United States, which requires load-serving entities not only to procure sufficient capacity to meet forecasted peak load but also to meet additional flexibility requirements within their capacity mix based on the expected future monthly 3-hour net load ramp (CAISO 2014; CAISO 2015).

Another important aspect of capacity markets—which is used to ensure long-term resource adequacy as well as revenue sufficiency—is the capacity value, or capacity credit, of variable generation. Capacity value is a measurement of the contribution of installed capacity to system resource adequacy. Capacity value is typically used by power system planners in long-term reliability assessments. It is also used within the context of capacity markets to determine the eligible capacity from variable generation resources that are qualified to participate in these markets. The geographic resolution, time period, and historical window of capacity value calculations vary by operating region. For example, at the time of this paper, PJM assigns a capacity value of 13% for wind and 38% for solar based on site-specific, performance-based capacity factors from the past 3 years during summer afternoons (PJM 2017c). Independent System Operator New England, on the other hand, has separate summer and winter capacity values, calculated as the average of the annual median output during predetermined hours per season during the past 5 years (Independent System Operator New England 2017).

A second approach to ensuring resource adequacy and revenue sufficiency involves improving energy and/or ancillary service markets. Instead of—or in addition to—capacity markets, other operating areas and researchers have focused on ways to supplement prices in energy-only markets to more fully reflect the cost of providing both energy and reliability (PJM 2017a; PJM 2017b; ERCOT 2017). Scarcity pricing might allow prices to exceed the variable costs of the most expensive operating resources for short periods of time when capacity or reserves are scarce so that those prices can help recover the fixed costs of the peaking units in addition to addressing the short-term reliability need. Scarcity pricing, also called shortage pricing, primarily includes administratively-set scarcity prices, which are set separately for both energy and

ancillary services; however, in typical power system operations, shortages will first result in insufficient reserve capacity while load balance is maintained. Therefore, in practice, scarcity prices typically depend on the ancillary service scarcity prices, which can be based on preset price curves or more dynamic, system-based conditions.

One example of an ancillary service scarcity pricing mechanism is the operating reserve demand curve (ORDC) within the Electric Reliability Council of Texas (ERCOT) region in the United States. In many markets, energy and ancillary services are co-optimized to ensure the optimal use of available units to meet energy, balancing, and other flexibility needs. As a result of this linkage, the prices for energy and ancillary services are directly coupled. For example, when reserves are short, the price of energy should increase. ERCOT, however, does not co-optimize in real time and has instead implemented the ORDC to pay resources that can provide reserves during the times of highest need. The ORDC is a continuous function of energy price adders derived from the operating reserve level in the current hour, the resulting expectation of loss of load probability, and the value of lost load. When operating reserves are low, this energy price adder increases, driving the energy price higher, and providing similar results to what would be expected in a co-optimized market. Despite concerns about the effectiveness of this mechanism, general consensus has found the ORDC to function as intended and designed, though some improvements could be made (ERCOT 2016). Other studies have shown the ORDC to be a preferred market design option compared to fixed reserve prices or fixed capacity payments for valuing incremental reserve capacity and helping to ensure revenue sufficiency (Levin and Botterud 2015, 392–406). Nevertheless, when any linkage between energy and ancillary services is not properly tuned or functioning, or if the reserve market in the case of ERCOT becomes saturated, thereby avoiding the triggering of the ORDC price adders, then the resources that are needed to ensure resource adequacy might not earn sufficient revenue to remain in the market long term.

Scarcity pricing for energy has historically included administratively-set caps to prevent price gouging or increases in prices that are too severe. In the United States, energy price caps are nominally set at \$2,000/MWh in most operating regions (FERC 2016b). ERCOT has applied a significantly higher price cap, currently \$9,000/MWh, in an attempt to allow energy prices to reach levels that appropriately signal needed investment in capacity²⁴; however, the current excess capacity in many systems, including ERCOT, has resulted in fewer hours with scarcity pricing and—along with low natural gas prices and lower growth of demand than anticipated—has contributed to the overall lower energy prices in recent years (e.g., Potomac Economics 2016a; Potomac Economics 2016b). In a more extreme example, Germany²⁵ is moving from its

²⁴ As an energy-only market, ERCOT relies on high-price periods that exceed the marginal unit cost to achieve revenue sufficiency. A higher price cap does not necessarily reflect the perception that market power challenges are not significant. Like other operating regions with wholesale electricity markets, the Public Utility Commission of Texas has rules for monitoring and mitigating market power (see <https://www.puc.texas.gov/agency/rulesnlaws/subrules/electric/Electric.aspx>). One interesting item within these rules is the so-called “Small Fish Rule,” which deems any electricity generating entity controlling less than 5% of the total installed generation capacity in ERCOT as not having ERCOT-wide market power. The rule protects small new entrants in the generation market from claims of market power abuse, providing the opportunity for a sufficient return on investment and removing potential uncertainty that might otherwise discourage the entry of new generation (Anderson 2015).

²⁵ The German wholesale electricity market consists of a forward market, day-ahead market, and intra-day market. While many wholesale transactions occur through bilateral contracts, a growing portion of these transactions are

feed-in tariffs to an auction system and has decided against both capacity markets²⁶ and any form of administratively-set energy price caps. Instead, Germany has implemented market-driven price formation, allowing energy prices to spike to unlimited levels, unfettered by any administrative action or price caps such as those throughout much of the United States (Bundestag 2016).

Finally, in an effort to hedge risks against volatile energy prices, many operating areas make use of bilateral or other types of contracts. Although the structure of these contracts can vary widely in practice, they typically establish predetermined price, quantity, time frame, and type of provision from a generator. In some operating regions in the United States, these contracts can provide a significant portion of guaranteed energy, ancillary services, and/or capacity services, and thus they serve an important role in ensuring long-term reliability. For example, the market monitoring group for CAISO has noted that long-term bilateral contracts are the primary means for ensuring adequate investment in new generation or retrofits of existing generation (CAISO 2016); however, if and when certain out-of-market contracts—along with various forms of self-scheduling²⁷—restrict operational flexibility through must-run and other price-taker positions, these transactions can contribute to the underlying price distortions that they aim to ameliorate. Thus, there appears to be an ideal balance that these contracts must achieve to provide a net positive benefit to market efficiency.

The above conversation focuses on challenges and potential solutions for restructured, competitive electricity markets; however, large portions of the United States still operate under traditional, vertically integrated markets in which retail prices explicitly capture all costs (i.e., cost-of-service pricing) and capacity is acquired as needed to meet a given resource adequacy target, thereby ensuring revenue sufficiency and resource adequacy. The downside—and a key justification for establishing competitive markets in other areas—is that such a centrally-controlled system inherently lacks competition among generators and thus cannot guarantee the least-cost system and operation for consumers; however, as discussed in this paper, competitive markets brought with them a suite of problems, most of which must still be resolved. Even with market reform, electricity markets will likely never be able to function as perfectly competitive markets because of inherent market failures, as addressed above. This has led some to suggest a return to traditional centrally-operated models; some claim that these have been more successful at ensuring long-term reliability (Morey et al. 2014). Others argue that some market design modifications, specifically CRMs, are effectively a step toward reregulation (Hogan 2016). Further, a viewpoint of electricity as a “public good” could justify a modified version of a central pricing system, such as a subscription-pricing structure like those of cell phones (Borenstein 2015; Borenstein and Bushnell 2015).

occurring through power exchanges, such as the European Power Exchange spot trading group, EPEX SPOT, and Energy Exchange Austria, which operate auction-based markets for energy at various time horizons and intervals (RAP 2015). Additional cross-border electricity exchanges can enable further trading across regions, for example, through the Nord Pool exchange group.

²⁶ However, a new “capacity reserve” product is being created, which ensures that a certain amount of out-of-market backup capacity is available during unforeseen events but does not distort competition in the actual market.

²⁷ Self-scheduled units submit a predefined generation schedule to the market operator and act as a price taker, and therefore they are very likely to earn lower profits. This also results in less efficient system dispatch. Bilateral contract and self-scheduled units can comprise a significant portion of an operating region, sometimes supplying nearly 75% of the energy (Monitoring Analytics 2015).

4.3 Resource Adequacy

Concerns about the impact of revenue sufficiency on resource adequacy extend beyond the challenge of co-optimizing energy and ancillary services. Signaling and providing adequate capacity is a necessary condition of planning and operating reliable power systems.

Resource adequacy is met when the portfolio of resources (generation and, if applicable, demand response, storage, etc.) for a given system meets the designated reliability target. The preferred method for calculating the contribution of these resources toward that target is based on loss-of-load probability and related probabilistic approaches. The reliability target used in practice is set by policy and is typically the result of an administrative action that establishes an acceptable level of reliability for long-term supply. In the United States, there is no universal resource adequacy target. Instead, each planning area²⁸ sets its own target, often imposed somewhat arbitrarily by policy. Each year, the North American Electric Reliability Corporation monitors, but does not enforce, long-term resource adequacy by reporting reserve margins for each assessment area. A common target for a loss-of-load expectation is 1 day/10 years,²⁹ but other targets can be used.

Whatever the resource adequacy target might be, it is fundamentally detached from the market process and outcomes, unless there is a long-term reliability component in electricity pricing (Cramton et al. 2013, 27–46; Ela et al. 2014). Because energy-only markets cannot explicitly take into account whether the target is 1 day/10 years, 1 day/2 years, or 1 day/20 years, there is no reason to expect that energy-only markets by themselves will simultaneously achieve an administered long-term reliability target, balance supply and demand, and provide sufficient revenue to all resources so that all costs (fixed and variable) are recovered³⁰; however, market and technology improvements to enable a more functional demand curve and allow customers to choose their level of reliability (Figure 10) might be able to provide price signals that bridge this gap between reliability and revenue sufficiency challenges. This disconnect between an energy-only market and long-term reliability is also a key argument for an additional capacity procurement mechanism to ensure sufficient planning reserves. At low loss-of-load expectation targets, this capacity might function more as an emergency reserve, and it can consist of combustion turbines that are relatively inexpensive (Cramton et al. 2013, 27–46).

One possible consequence of higher reliability is overcapacity.³¹ Long-term reliability targets might require additional capacity to be built, but this capacity might never need to run, or it might run for fewer hours of the year than it would have with a smaller reserve margin. In other words, excess capacity can be pushed up and/or off the merit-order curve, effectively keeping prices to lower portions of the supply curve. If periods of shortage do occur, prices can spike

²⁸ Typically the load-serving entity.

²⁹ Loss-of-load probability is calculated by a suitable convolution algorithm or Monte Carlo analysis using generator capacity and forced outage data along with load data; the typical application reports the probability of insufficient supply on an hourly or daily basis. Loss-of-load expectation is the expected value of insufficient supply, which is based on the loss-of-load probability and typically reported in days per year.

³⁰ A market with excess supply will likely have less total revenue because excess capacity will likely be distributed across the merit-order curve or to the left of the intersection of the demand curve. This will enable the system to operate at a lower portion of the merit-order curve, effectively suppressing prices.

³¹ However, over-capacity might not be a concern in cases where the cost of the extra capacity is less than the value to the customers of the improved reliability.

significantly, but administratively-set price caps might keep prices below levels that properly support peaking units during these shortage events. The result is that resources that might be needed for only a short time of the year (such as this extra capacity) must recover both capital and operational costs in a few hours or days, and therefore they might not have the opportunity to earn sufficient revenue to remain in the market (Cramton and Stoft 2006).

4.4 Impact of Variable Generation

Existing market challenges, such as revenue sufficiency and resource adequacy, can be exacerbated by the addition of low or near-zero marginal-cost resources. Low-cost natural gas and variable generation—namely wind and solar—are key contributors to what is known as the merit-order effect, which we describe below. Variable generation resources have high capital costs with near-zero marginal costs because of the lack of fuel costs. The introduction of these resources into the generation mix has two primary impacts: (1) price reduction and (2) energy sales reduction for other generators.

First, prices will be reduced in the market period because these low or zero-cost resources will displace one or more resources at the top of the dispatch stack, assuming a convex supply curve. When variable generation is the marginal resource, the price can be zero, or negative if production-based subsidies are present.³² This means that, on average, electricity prices will be reduced as low-cost natural gas generators and variable generation are introduced into the market. Power system economic studies and actual practice have shown such price suppression impacts during periods of, or in areas with, high variable generation output (Keane et al. 2011, 564–572; Milligan et al. 2012; Gallo 2016, 39–46; Levin and Botterud 2015, 392–406; Potomac Economics 2014, Würzburg et al. 2013, S159–S171; Zarnikau 2011, 3,906–3,913; Woo et al. 2011, 3,939–3,944; International Energy Agency Renewable Energy Technology Deployment Technology Collaboration Programme 2016; Brancucci Martinez-Anido et al. 2016, 474–487) as well as in regions with large amounts of natural gas generation when natural gas fuel prices and demand are low (Potomac Economics 2016a; Potomac Economics 2016b). Second, the remaining resources that are dispatched after these lower-cost resources are called on will run at lower capacity factors, further reducing revenues received by the generators with higher merit orders, again assuming a convex market.

These impacts on price and capacity factor suppression are illustrated in Figure 11 by the merit-order effect, which pushes more expensive resources up (or off) of the dispatch stack. The left panel of the figure shows a simplified supply curve for a small power system. Three demand curves (D1, D2, and D3) represent three different levels of electricity demand throughout the year; their intersections with the supply curve reveal hypothetical prices at 350 MW, 650 MW, and 900 MW.³³ Prices are determined by the generation cost of the marginal unit; for the three demand curves, the price is \$30/MWh, \$50/MWh, and \$80/MWh, respectively.

Now suppose that some combination of variable generation resources is added to the system. In this case, the original supply curve shifts right by the amount of renewable energy in any given

³² Negative prices can also occur if inflexible conventional units, such as coal or nuclear, are the marginal unit and bid in negative offers in order to stay online and avoid high shutdown and start-up costs. This has been seen recently in CAISO.

³³ For simplicity, the time period represented by each demand curve is 1 hour.

hour. For this example, assume 500 MW of variable generation output is added; the new supply curve is shown in the right panel of Figure 11. After this addition, the price at any demand level less than 500 MW is \$0/MWh. Applying the same demand curves, the left panel yields equilibrium prices of \$0/MWh, \$30/MWh, and \$35/MWh, respectively. For higher levels of variable generation, the region of the supply curve with the \$0/MW(h) price is larger, showing that (1) there will be more hours of zero (or negative) prices at higher levels of variable generation when these resources set the price, and (2) prices in other hours will be lower than they are in the absence of additional variable generation when all else is equal. Reductions in average price will make it more difficult for generators, especially those higher up on the merit-order curve, to recover all their costs because they will run less often and receive a lower average price when they do run. This is particularly relevant for some generators (e.g., peaking units) that are needed to ensure resource adequacy but do not earn sufficient revenue to remain in the market (Cramton and Stoft 2006; ECCO International 2015; Levin and Botterud 2015, 392–406). We note that similar price suppression effects, as well as merit-order reordering impacts, can result from the addition of low-cost natural gas generators.

However, many non-convexities exist in practice, such as those from no-load costs and start-up costs. In order to most economically meet the needs of the system, these non-convexities can necessitate a different merit order than that illustrated by a simple convex supply curve, such as in Figure 11 and described above. The effect of these non-convexities in most markets is price suppression, potentially outweighing, or at least masking, the price suppression impact of variable generation that has been noted in numerous studies cited above. More work is needed to develop robust pricing mechanisms to capture these non-convexities and enable market sustainability, i.e., revenue sufficiency to support resource adequacy.

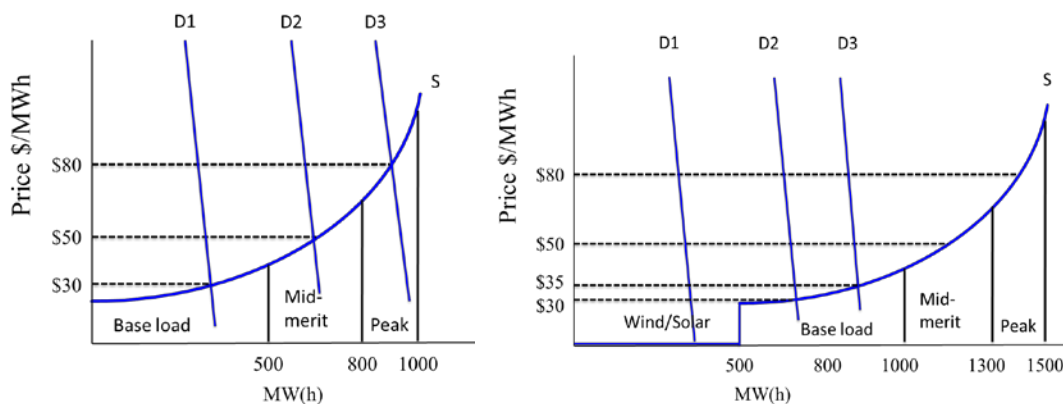


Figure 11. Impact of variable renewable generation on market prices. Image from Frew et al. (2016a)

It is well known that variable generation resources also increase the variability and uncertainty in the system, which can in turn require increased flexibility from the grid (Ela et al. 2016, 51–60). Although certain changes to short-term energy and ancillary service markets might be needed to ensure that the available flexibility is offered to the market, this might not guarantee that sufficient flexibility is built or available in the first place. This could necessitate new methods for evaluating and compensating all needed aspects of capacity adequacy, including how to quantify and value the flexible capability of capacity.

4.5 U.S. Wholesale Electricity Market Examples to Minimize Revenue Sufficiency and Resource Adequacy Challenges

Many variations in the ownership aspects of power systems have moved toward wholesale markets. Examples across this spectrum provide a range of possible ownership and market structures to minimize challenges with revenue sufficiency and resource adequacy. In the United States, CAISO and Midcontinent Independent System Operator (MISO) provide an interesting contrast. Note that although many regions in the United States have wholesale markets, the retail sector continues to function largely as a regulated monopoly (with a few exceptions).

The first example of a wholesale market structure is CAISO. When California moved to a wholesale market, the investor-owned utilities divested ownership in all generation. With increasing wind and solar generation, CAISO developed a flexible ramping product and incorporated that into the economic dispatch to ensure sufficient operational flexibility; this product was approved for both upward and downward ramping capability in 2016 (Ela 2016). To ensure resource adequacy, the investor-owned utilities under the purview of the California Public Utilities Commission are required to demonstrate sufficient contractual power purchases to meet adequacy requirements. From the perspective of the generating company, a contract should cover all, or at least part, of fixed costs, thus providing sufficient revenue. Revenues to pay for the contractual obligation are largely collected from retail rates, which are negotiated between each investor-owned utility, the California Public Utilities Commission, and other stakeholders. Markets with this type of hybrid—operational markets with regulatory oversight and tariffs designed to recover costs—can largely solve the revenue sufficiency question. The challenge might be to integrate a bilateral contract structure with operational flexibility requirements so that the market operator has sufficient access to and incentives for operational flexibility.

Our second example of wholesale ownership is Northern States Power, which is a subsidiary of Xcel Energy and is located within the MISO market and the State of Minnesota. MISO developed as a RTO and market—in contrast to CAISO, which was a state-wide initiative. Northern States Power continues to own generation that is bid into the MISO market. Power purchase agreements (PPAs), which are a type of contract with the generation owners, can be structured to ensure revenue sufficiency, with revenue streams from the wholesale market alongside those that are supported by retail rates set with the Minnesota Department of Commerce, the regulatory body for electricity in Minnesota. Although this represents a different ownership model than in CAISO, the revenue streams are similar.

Hybrid market/regulated paradigms such as those above will likely always exist, and these examples from CAISO and MISO might reduce or even eliminate the revenue sufficiency problem; however, they might do so at the expense of operational flexibility. This flexibility can in principle be maintained or increased by changing the structure of the PPAs so that operational flexibility is incentivized and/or explicitly required. For example, a modified long-term PPA could be structured to guarantee fixed-cost recovery while requiring the generator to participate in the energy market to recover variable costs and provide flexibility to the system. Ideally, this sort of modified PPA approach could allow for coordinated resource acquisition by the ISO to ensure both capacity and flexibility requirements at the system level; however, any contract faces risk from lack of foresight of future economic and power system changes. Fully unbundled markets with retail choice, independent market operators, and independent generators do not

currently exist in the United States outside of market-regulatory hybrids such as those discussed above. Instead, investor-owned utilities serve as the retail source for most electricity (ignoring public power, for now), and the regulatory compact with utility regulators can continue to enforce resource adequacy requirements with revenue sufficiency maintained by PPAs, providing that the PPAs do not strand the physical flexibility that could be available to operators.

4.6 Complicating Factors and Unintended Consequences for Market Design

We have laid out the attributes of wholesale electricity markets that suggest market failure: externalities, public good attributes of the network, ineffective demand curve, market power, wholesale price caps, and lack of large-scale storage; see Section 4.1. These failures, along with the array of possible corrective actions to mitigate the underlying market distortions and ensure revenue sufficiency and resource adequacy, make electricity markets fundamentally different and more complex than nearly any other market. This complexity grows with the number of related market products and policies, such as production or investment tax credits. A host of additional factors further complicate the already multifaceted challenges of ensuring revenue sufficiency and resource adequacy. Among these are a lack of consensus on ideal market structure and design(s), uncertainty surrounding the transition “path” to that elusive ideal market design, the sensitivity and unintended consequences of well-meaning regulations/designs, and the “creeping complexity” of certain market designs. Below we discuss each of these complicating factors as well as unintended consequences.

Market design, if not done carefully, can result in unintended consequences. One example is the declining frequency response in the eastern United States after energy markets were introduced. Many argued that this was a consequence of the economic disincentives to follow frequency and incur penalties for not meeting energy delivery obligations. Once this problem was recognized, the market design was revised to recognize the need for frequency response and develop a market mechanism to incentivize the need for both primary frequency response and performance-based secondary frequency response (i.e., regulating reserve) (Ela et al. 2016, 51–60).

Some proposed and implemented market designs—which were intended to offer greater simplicity and transparency—have instead resulted in additional reforms, more exceptions, and increasing complexity as a result of attempts to achieve an ideal balance of risk between investors and consumers. One notorious example of such creeping complexity is capacity markets in which CRMs have been criticized for insufficient penalties when capacity providers fail to perform, incentivizing overprocurement of capacity and the inability to differentiate capacity among various dimensions such as flexibility (Hogan 2016). An ideal CRM incentivizes capacity with the ideal quantity and attributes needed to supply energy and reserves during shortages; is technology neutral; and provides strong performance incentives to ensure reliability, improve efficiency of both the capacity and spot energy markets, and provide a level playing field—e.g., with a “no exceptions” policy (Cramton 2013; FERC 2013a). Many of these attributes might require revisiting the definition and method(s) for calculating resource adequacy, particularly as the power system continues to evolve into one with greater shares of variable renewable energy and demand response (Hogan 2013).

Ideally, a market design should be robust enough to provide the proper incentives during the transition from low to high variable generation penetration rates, although this might be difficult to achieve in practice. At existing variable generation penetration rates—nominally less than 20% in most regions—baseload generation might continue to operate economically. But as the penetration of variable generation becomes high enough to push out baseload generation on an economic basis, more flexible resources would be expected as a replacement. These new flexible resources might be capable of quickly starting and synchronizing to the grid. This could significantly change the unit commitment process and day-ahead market; and, given sufficiently flexible resources, it could even reduce the need for a nonspin product, at least when tied to the physical capability of the units. Thus, it is not clear whether energy markets supplemented by various types of reserve products would provide the revenue sufficiency critical to ensuring long-term reliability.

Finally, market designs should be technologically neutral. A market that is performance-based, with various requirements for start-up notification and deployment, length of deployment, etc., can acquire the appropriate response from any technology that can provide the needed capability. This opens the door for new technologies and promotes competition.

4.7 Future Research

Much work is needed to determine the ideal market design to ensure revenue sufficiency and reliability. The ideal electricity market would be robust to the penetration level of variable generation and reserve needs, would minimize boom/bust cycles, and would be effective in both the “target” state as well as in the transition from the current system to that ideal market design end point. This ideal market should incentivize operational efficiency and flexibility as well as investment in sufficient flexibility. In addition, the market should deliver sufficient revenue to support and incentivize the level of capacity needed for long-term reliability, but it should not deliver revenue sufficiency to an overbuilt system.

A rigorous market design analysis should involve capacity expansion models for long-term planning, production cost models for operations, and models with faster time steps for system stability; however, validation and calibration are often needed for model results to more accurately reflect actual market outputs and avoid providing misinformation to market designers or policymakers. For example, previous work has shown that it is important to accurately reflect reserve treatment and account for strategic generator bidding behavior in production cost models to produce more accurate price outputs (Frew et al. 2016b). Empirical evidence has further shown the importance of capturing operational factors and outages that can drive short-term price spikes or modify the optimal dispatch of the generator fleet. The former include generator and transmission outages, congestion, and contingencies that can result in localized shortage of generation and/or transmission flow limitations. The latter include generator self-scheduling operations that can shift other units off the offer curve or result in a higher need for more expensive flexible generation. Other operational factors, such as shutdown and start-up costs, have been shown to impact prices. For example, the surge in solar photovoltaic power in California is pushing many units to shut off midday when solar is peaking, but the high shutdown and start-up costs of many natural gas plants have led these units to prefer operating below their marginal cost—while still setting the price—during these hours (Nelson 2016).

5 Summary and Conclusions

The goal of any power system is to provide a reliable supply of power at the lowest reasonable cost to the consumer. Many of today's power systems are regulated through competitive wholesale electricity markets in order to provide transparency and mitigate market power. The underlying reliability objective necessitates market structures and operating rules that ensure revenue sufficiency for all generators needed for resource adequacy purposes.

Throughout this paper, we have discussed inherent attributes of current competitive wholesale electricity markets that contribute to market failure and thereby preclude these markets from functioning as perfectly competitive markets. These attributes include externalities, public good attributes of the network, ineffective demand curve, market power, wholesale price caps, and lack of large-scale storage. Some of these features are classic sources of market failure, while others have compounding effects. Imperfect information further contributes to and amplifies their collective impact. Outside of perfect competition, marginal-cost pricing alone cannot guarantee cost recovery; thus, this paper argues that marginal-cost pricing alone is not sufficient for the efficient investment in and operation of wholesale electricity markets.

Until (and unless) these contributors to market failure are ameliorated, some form of corrective action(s) will continue to be necessary to improve market efficiency so that prices can correctly reflect the needed level of system reliability. Many of these options necessarily involve some form of administrative or out-of-market actions, such as scarcity pricing, capacity payments, bilateral or other out-of-market contracts, or some hybrid combination. A hybrid approach could include a modified PPA contract that ensures fixed-cost recovery while requiring the generator to participate in the energy market for variable cost recovery and provision of flexibility to the system. A key focus with these options is to create a connection between the electricity market and long-term reliability/loss-of-load expectation targets, which are inherently disconnected in the native markets because of the aforementioned contributors to market failure.

The addition of low marginal cost resources, such as generators fueled by low cost natural gas and near-zero marginal cost wind and solar generators, can exacerbate revenue sufficiency and resource adequacy concerns caused by these underlying contributors to market failure. These low marginal cost resources effectively suppress energy prices and reduce the capacity factors of conventional generators through the merit-order effect in the simplest case of a convex market; non-convexities can also suppress prices, even outweighing the impact of low marginal cost generation.²

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

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

References

- Anderson, K. 2015. “RE: Notice of Proposed Order and Request for Comment on an Application for an Exemptive Order from Southwest Power Pool, Inc. From Certain Provisions of the Commodity Exchange Act Pursuant to the Authority Provided in Section 4(c)(6) of the Act.” http://www.puc.texas.gov/agency/about/commissioners/anderson/letters/PUCT_Comments_Regarding_SPP_Exemptive_Order.pdf.
- Bade, R., and M. Parkin. 2002. *Foundations of Microeconomics*. Boston, MA: Addison Wesley.
- Baker, J.B. 2017. “Market Power in the U.S. Economy Today.” *Washington Center for Equitable Growth*, March 20, 2017. Accessed May 2, 2017. <http://equitablegrowth.org/research-analysis/market-power-in-the-u-s-economy-today/>.
- Borenstein, S. 2016. “Is Electricity Different from ‘Real Markets’? Should it Be?” *Energy Institute Blog*, June 13, 2016. <https://energythaas.wordpress.com/2016/06/13/is-electricity-pricing-different-from-real-markets-should-it-be/>.
- Borenstein, S., and J. Bushnell. 2015. “The U.S. Electricity Industry after 20 Years of Restructuring (Working paper 252R).” Berkeley, CA: Energy Institute at Haas. <https://ei.haas.berkeley.edu/research/papers/WP252.pdf>.
- Brancucci Martinez-Anido, C., G. Brinkman, and B.-M. Hodge. 2016. “The Impact of Wind Power on Electricity Prices.” *Renewable Energy* 94.
- Bundestag. 2016. Gesetz zur Weiterentwicklung des Strommarktes (Strommarktgesetz). Bundesgesetzblatt, Part 1, No. 37, pp. 1786-1817, Bonn, Germany, July 29, 2016.
- Bushnell, J., M. Flagg, and E. Mansur. 2017. “Capacity Markets at a Crossroads” (White Paper). Berkeley, CA: Energy Institute at Haas.
- CAISO. 2016. *2015 Annual Report on Market Issues and Performance*. Folsom, CA: Department of Market Monitoring. <http://www.caiso.com/Documents/2015AnnualReportonMarketIssuesandPerformance.pdf>.
- CAISO, 2014. *Flexible Resource Adequacy Criteria and Must-Offer Obligation: Market and Infrastructure Policy Draft Final Proposal*. California ISO, Folsom, CA.
- CAISO, 2015. *Reliability Services Initiative—Phase 2 and Flexible Resource Adequacy Criteria and Must Offer Obligation—Phase 2: Issue Paper*. California ISO, Folsom, CA.
- Council of Economic Advisers. 2016. “Benefits of Competition and Indicators of Market Power. Issue Brief.” Washington, D.C. Accessed May 2, 2017. https://obamawhitehouse.archives.gov/sites/default/files/page/files/20160502_competition_issue_brief_updated_cea.pdf.
- Cramton, P. 2013. “Capacity Markets Going Forward.” Paper presented at the Federal Energy Regulatory Commission Technical Conference on Centralized Capacity Markets in Regional

Transmission Organizations and Independent System Operators, Washington, D.C., September 25, 2013. <http://www.ferc.gov/CalendarFiles/20130911144842-Cramton%20Comments.pdf>.

Cramton, P., A. Ockenfels, and S. Stoft. 2013. "Capacity Market Fundamentals." *Economics of Energy and Environmental Policy* 2 (2). <http://dx.doi.org/10.5547/2160-5890.2.2.2>.

Cramton, P., and S. Stoft. 2006. "The Convergence of Market Designs and Adequate Generating Capacity" (White Paper). Sacramento, CA: Electricity Oversight Board.

Dano, M. 2016. "How Verizon, AT&T, T-Mobile, Sprint, and More Stacked Up in Q4 2015: The Top 8 Carriers." *Fierce Wireless*, March 1, 2016. <http://www.fiercewireless.com/special-report/how-verizon-at-t-t-mobile-sprint-and-more-stacked-up-q4-2015-top-8-carriers>.

Dillon, T.P. 1979. "Title I of PURPA: The Effect of Federal Intrusion into Regulation of Public Utilities." *William & Mary Law Review* vol. 21, no. 2, article 6, rev. 491. Accessed August 16, 2017. <http://scholarship.law.wm.edu/cgi/viewcontent.cgi?article=2339&context=wmlr>.

ECCO International. 2015. *Modeling and Analysis of Wholesale Electricity Market Design: Understanding the Missing Money Problem* (Technical Report NREL/SR-5D00-64255). Golden, CO: National Renewable Energy Laboratory.

Ela, E. 2016. *Wholesale Electricity Market Design Initiative in the United States: Survey and Research Needs* (Technical Update 3002009273). Palo Alto, CA: Electric Power Research Institute.

Ela, E., M. Milligan, A. Bloom, A. Botterud, A. Townsend, and T. Levin. 2014. *Evolution of Wholesale Electricity Market Design with Increasing Levels of Renewable Generation* (Technical Report NREL/TP-5D00-61765). Golden, CO: National Renewable Energy Laboratory. <http://www.nrel.gov/docs/fy14osti/61765.pdf>.

Ela, E., M. Milligan, A. Bloom, A. Botterud, A. Townsend, T. Levin, and B. Frew. 2016. "Wholesale Electricity Market Design with Increasing Levels of Renewable Generation: Incentivizing Flexibility in System Operations." *The Electricity Journal* 20 (4).

ERCOT. 2016. "Discussion on ORDC Parameters with ERCOT Staff Analysis and Stakeholder Position Papers" (White Paper). Austin, TX: Supply Analysis Working Group. http://www.ercot.com/content/wcm/key_documents_lists/76817/10_SAWG_ORDC_Discussion_Analysis_and_Positions_1_11_16.pdf.

ERCOT. 2017. *Methodology for Implementing Operating Reserve Demand Curve (ORDC) to Calculate Real-Time Reserve Price Adder* (Technical Report). Accessed August 15, 2017. http://www.ercot.com/content/wcm/key_documents_lists/89286/Methodology_for_Implementing_ORDC_to_Calculate_Real-Time_Reserve_Price_Adder.zip.

Executive Office of the President and Council of Economic Advisers. 2016. "Economic Report of the President (2016)." Washington, D.C. Accessed May 2, 2017. <https://www.gpo.gov/fdsys/pkg/ERP-2016/content-detail.html>.

FERC. 2013a. *Centralized Capacity Market Design Elements* (Commission Staff Report AD13-7-000). Washington, D.C.

FERC. 2013b. *Centralized Capacity Markets in RTOs/ISOs: Docket No. AD13-7-000* (Technical Conference). Washington, D.C.

FERC. 2016a. *Modifications to Commission Requirements for Review of Transactions under Section 203 of the Federal Power Act and Market-Based Rate Applications under Section 205 of the Federal Power Act: Docket No. RM16-21-000; Item No. E-2*. Washington, D.C.
<https://www.ferc.gov/whats-new/comm-meet/2016/092216/E-2.pdf>

FERC. 2016b. *Offer Caps in Markets Operated by Regional Transmission Organizations and Independent System Operators: Docket No. RM16-5-000; Order No. 831*. Washington, D.C.

FERC. 2014. *Price Formation in Organized Wholesale Electricity Markets: Staff Analysis of Energy Offer Mitigation in RTO and ISO Markets: Docket No. AD14-14-000*. Washington, D.C.
<https://www.ferc.gov/legal/staff-reports/2014/AD14-14-mitigation-rto-iso-markets.pdf>

FERC. 2017. *State Policies and Wholesale Markets Operated by ISO New England Inc., New York Independent System Operator, Inc., and PJM Interconnection, L.L.C.: Docket No. AD17-11-000* (Technical Conference). Washington, D.C.

Frew, B.A., M. Milligan, G. Brinkman, A. Bloom, K. Clark, and P. Denholm. 2016a. “Revenue Sufficiency and Reliability in a Zero Marginal Cost Future” (NREL/CP-6A20-66935). Paper presented at the 15th International Workshop on Large-Scale Integration of Power into Power Systems as well as on Transmission Networks for Offshore Power Plants, Vienna, Austria, November 15–17, 2016.

Frew, B.A., G. Gallo, G. Brinkman, M. Milligan, K. Clark, and A. Bloom. 2016b. *Impact of Market Behavior, Fleet Composition, and Ancillary Services on Revenue Sufficiency* (Technical Report NREL/TP-5D00-66076). Golden, CO: National Renewable Energy Laboratory.
<http://www.nrel.gov/docs/fy16osti/66076.pdf>

Gallo, G. 2016. “Electricity Market Games: How Agent-Based Modeling Can Help under High Penetrations of Variable Generation.” *The Electricity Journal* 29 (2).
<http://dx.doi.org/10.1016/j.tej.2016.02.001>.

Gruber, J. 2016. *Public Finance and Public Policy*. New York: MacMillan.

Helman, U., and B. Hobbs. 2010. “Large-Scale Market Power Modeling: Analysis of the U.S. Eastern Interconnection and Regulatory Applications.” *IEEE Transactions on Power Systems* 25 (3).

Hibbard, P., S. Tierney, and K. Franklin. 2017. *Electricity Markets, Reliability and the Evolving U.S. Power System*. Analysis Group.
http://www.analysisgroup.com/uploadedfiles/content/insights/publishing/ag_markets_reliability_final_june_2017.pdf

Hill, C., and G. Jones. 2009. *Strategic Management Theory: An Integrated Approach—Ninth Edition*. Mason, OH; South-Western Cengage Learning.

Hogan, W.W. 2005. “On an “Energy Only” Electricity Market Design for Resource Adequacy.” Harvard University, Cambridge, MA. Accessed September 12, 2017. https://sites.hks.harvard.edu/fs/whogan/Hogan_Energy_Only_092305.pdf

Hogan, M. 2013. “Opening Statement: FERC Technical Conference on the Future of Centralized Capacity Markets.” Paper presented at the Federal Energy Regulatory Commission Technical Conference on Centralized Capacity Markets in Regional Transmission Organizations and Independent System Operators, Washington, D.C., September 25, 2013. www.ferc.gov/CalendarFiles/20130911144859-Hogan%20Comments.pdf.

Hogan, M. 2016. “Hitting the Mark on Missing Money: How to Ensure Reliability at Least Cost to Consumers.” *Regulatory Assistance Project*, September 12, 2016.

Hogan, M. 2017. “Follow the Missing Money: Ensuring Reliability at Least Cost to Consumers in the Transition to a Low-Carbon Power System.” *The Electricity Journal* 30.

Hogan, W.W. 2008. “Electricity Market Structure and Infrastructure.” Paper presented at the Conference on Acting in Time on Energy Policy, Harvard University, Cambridge, MA, September 25, 2008. https://www.hks.harvard.edu/fs/whogan/Hogan_Elec_r_092508.pdf.

Houldin, R.W. 2004. “Find the Public Good: Shedding Light on a Bulk Grid Electricity Card Trick.” *The Electricity Journal*. doi:/10.1016/j.tej.2004.09.008

Independent System Operator New England. 2017. *Forward Capacity Market: Docket # ER17-925-000*. Holyoke, MA. Accessed August 16, 2017. https://www.iso-ne.com/static-assets/documents/regulatory/tariff/sect_3/mr1_sec_13_14.pdf.

International Energy Agency Renewable Energy Technology Deployment Technology Collaboration Programme. 2016. *Electricity Market Design and RE Deployment (RES-E-MARKETS)* (Technical Report). Utrecht, The Netherlands. <http://iea-retd.org/wp-content/uploads/2016/09/IAE-RETD-RES-E-MARKETS-report-1.pdf>.

Jenkin, T., P. Beiter, and R. Margolis. 2016. *Capacity Payments in Restructured Markets under Low and High Penetration Levels of Renewable Energy* (Technical Report NREL/TP-6A20-65491). Golden, CO: National Renewable Energy Laboratory.

Jordan, G., and D. Piwko. 2013. *An Analysis of the Impact of Balancing Area Cooperation on the Operation of WECC and the Compounding Effect of Wind and Solar Generation* (Technical Report NREL/TP-5500-57115). Golden, CO: National Renewable Energy Laboratory.

Joskow, P.L. 1997. “Restructuring, Competition and Regulatory Reform in the U.S. Electricity Sector.” *Journal of Economic Perspectives* 11 (3). <http://kylewoodward.com/blog-data/pdfs/references/joskow-the-journal-of-economic-perspectives-1997A.pdf>.

- Keane, A., M. Milligan, C.J. Dent, B. Hasche, C. D’Annunzio, K. Dragoon, H. Holttinen, N. Samaan, L. Soder, and M. O’Malley. 2011. “Capacity Value of Wind Power.” *IEEE Transactions on Power Systems* 26 (2). doi: 10.1109/TPWRS.2010.2062543.
- King, J., B. Kirby, M. Milligan, and S. Beuning. 2012. *Operating Reserve Reductions from a Proposed Energy Imbalance Market with Wind and Solar Generation in the Western Interconnection* (Technical Report NREL/SR-5500-54660). Golden, CO: National Renewable Energy Laboratory.
- Levin, T., and A. Botterud. 2015. “Electricity Market Design for Generator Revenue Sufficiency with Increased Variable Generation.” *Energy Policy* 87.
- Milligan, M., H. Holttinen, L. Soder, C. Clark, and I. Pineda. 2012. “Markets to Facilitate Wind and Solar Energy Integration in the Bulk Power Supply: An IEA Task 25 Collaboration” (NREL/CP-5500-56212). Paper presented at the 11th Annual International Workshop on Large-Scale Integration of Wind Power into Power Systems as Well as on Transmission Networks for Offshore Wind Power Plants Conference, Lisbon, Portugal, November 13–15. <http://www.nrel.gov/docs/fy12osti/56212.pdf>.
- Milligan, M., K. Clark, J. King, B. Kirby, T. Guo, and G. Liu. 2013. *Examination of Potential Benefits of an Energy Imbalance Market in the Western Interconnection* (Technical Report NREL/SR-5500-53095). Golden, CO: National Renewable Energy Laboratory.
- Milligan, M., B. Frew, A. Bloom, E. Ela, A. Botterud, A. Townsend, and T. Levin. 2016. “Wholesale Electricity Market Design with Increasing Levels of Renewable Generation: Revenue Sufficiency and Long-Term Reliability.” *The Electricity Journal* 29 (2).
- Monitoring Analytics. 2015. *PJM State of the Market: 2014* (Technical Report). Eagleville, PA. http://www.monitoringanalytics.com/reports/PJM_State_of_the_Market/2014.shtml.
- Morey, M.J., L.D. Kirsch, B.K. Eakin, and R.J. Camfield. 2014. *Ensuring Adequate Power Supplies for Tomorrow’s Electricity Needs*. Madison, WI: Electric Markets Research Foundation.
- National Action Plan for Energy Efficiency. 2009. *Customer Incentives for Energy Efficiency Through Electric and Natural Gas Rate Design* (Technical Report). Prepared by William Prindle, ICF International, Inc. Washington, DC.: U.S. Environmental Protection Agency and U.S. Department of Energy. Accessed August 16, 2017. https://www.epa.gov/sites/production/files/2015-08/documents/rate_design.pdf.
- Nelson, W. 2016. “CAISO Gas Turbine Economics: Spark Spread Devolution.” *Bloomberg New Energy Finance*, October 25, 2016.
- Newbery, D. 2016. “Missing Money and Missing Markets: Reliability, Capacity Auctions and Interconnectors.” *Energy Policy* 94.
- PJM. 2017a. “PJM Shortage Pricing Fact Sheet.” Norristown, PA. Accessed August 10, 2017. <http://www.pjm.com/~media/about-pjm/newsroom/fact-sheets/shortage-pricing-fact-sheet.ashx>.

- PJM. 2017b. *PJM Manual 18: PJM Capacity Market*. Norristown, PA: Capacity Market Operations. Accessed August 10, 2017. <https://www.pjm.com/~media/documents/manuals/m18.ashx>.
- PJM. 2017c. *PJM Manual 21: Rules and Procedures for Determination of Generating Capability*. Norristown, PA: Systems Planning Department. Accessed August 16, 2017. <https://www.pjm.com/~media/documents/manuals/m21.ashx>.
- Potomac Economics, Ltd. 2014. *2013 State of the Market Report for the ERCOT Wholesale Electricity Markets*. Fairfax, VA. https://www.potomaceconomics.com/uploads/ercot_documents/2013_ERCOT_SOM_REPORT.pdf.
- Potomac Economics, Ltd. 2016a. *2015 State of the Market Report for the ERCOT Wholesale Electricity Markets*. Fairfax, VA. https://www.potomaceconomics.com/uploads/ercot_documents/2015_ERCOT_State_of_the_Market_Report_FINAL_update_6.21.16.pdf.
- Potomac Economics, Ltd. 2016b. *2015 State of the Market Report for the MISO Electricity Markets*. Fairfax, VA. <https://www.misoenergy.org/Library/Repository/Report/IMM/2015%20State%20of%20the%20Market%20Report.pdf>.
- Regulatory Assistance Project (RAP). 2015. *Report on the German Power System*. Version 1.0. Study commissioned by Agora Energiewende. Brussels, Belgium. https://www.agora-energiewende.de/fileadmin/downloads/publikationen/CountryProfiles/Agora_CP_Germany_web.pdf
- Sauer, W. 2014. *Staff Analysis of Uplift in RTO and ISO Markets* (Technical Report). Washington, D.C. Federal Energy Regulatory Commission.
- Scherer, F.M. 1970. *Industrial Pricing: Theory and Evidence*. Berlin, Germany: Rand McNally and Company.
- Stoft, S. 2002. *Power System Economics*. New York, NY: John Wiley and Sons, Inc.
- Verizon. 2016. <https://www.verizonwireless.com/plans/verizon-plan/>.
- Wang, Q., H. Wu, J. Tan, B.-M. Hodge, W. Li, and C. Luo. 2016. "Analyzing the Impacts of Increased Wind Power on Generation Revenue Sufficiency." Paper presented at the IEEE Power and Energy Society General Meeting (PESGM), Boston, MA, July 17-21, 2016.
- Woo, C.K., I. Horowitz, J. Moore, and A. Pacheco. 2011. "The Impact of Wind Generation on the Electricity Spot-Market Price Level and Variance: The Texas Experience." *Energy Policy* 39.
- Würzburg, K, X. Labandeira, and P. Linares. 2013. "Renewable Generation and Electricity Prices: Taking Stock and New Evidence for Germany and Austria." *Energy Economics* 40.

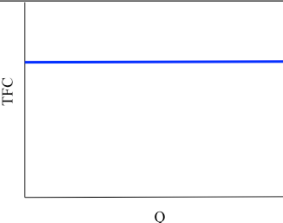
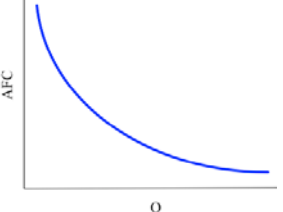

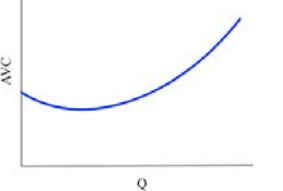
Zarnikau, J. 2011. “Successful Renewable Energy Deployment in a Competitive Electricity Market: A Texas Case Study.” *Energy Policy* 39.

Appendix: Cost Definitions and Curves

Standard microeconomic analysis is predicated on definitions of cost that are intended to be general enough to capture the key aspects of economic behavior. Accompanying these cost definitions is the notion of comparing short-run to long-run time periods. In the short run, it is not possible to alter the quantity of a productive resource. In the context of electricity markets, this means that in the short run it is not possible to add new resources or transmission. In the long run, changes can be made to the network and/or to operational practice; thus, the long run corresponds to the investment time frame. Table A-1 shows the basic cost definitions along with a simple graphic characteristic for each curve. In all cases, the Q axis denotes quantity of output—usually MWh—and all costs are either \$ or \$/MWh. Costs are shown as a function of output (MWh, or Q).

Economic profits differ from accounting profits. Accounting profit includes the return to businesses after deducting the cost of capital, labor, and rent from gross income. Economic profit also deducts the return to entrepreneurship, the rate of return required to elicit investment in the firm, and it represents the opportunity cost of profit; thus, in the long run, economic profit is zero in perfect competition.

Table A-1. Cost Definitions and Curves Relevant to Electricity Markets

Cost Definition	Description	Formula	Cost Curve Example
Total Fixed Cost (TFC)	Total investment cost and other fixed cost (might be annualized)	Sum of all fixed costs	
Average Fixed Cost (AFC)	Per unit of output (kWh)	$AFC = TFC/Q$	
Total Variable Cost (TVC)	Includes fuel, variable operation and maintenance, etc.	Sum of all variable costs	
Average Variable Cost (AVC)	Variable cost per unit of output	$AVC = TVC/Q$	

Total Cost	Sum of fixed and variable cost	$TC = TFC + TVC$	
Average Total Cost (ATC)	Average cost per unit of output	$ATC = AFC + AVC$	
Marginal Cost	Cost of producing one additional (the next) unit of output	$MC = d(TC)/dQ = d(TVC)/dQ$	
Cost Curves with Large Economies of Scale, Nonincreasing Marginal Cost		ATC, AVC, MC as defined above	