TOO MUCH IS NEVER ENOUGH: CONSTRUCTING ELECTRICITY CAPACITY MARKET DEMAND

Todd Aagaard[*] & *Andrew N. Kleit*[**]

Synopsis: Some regions of the United States have created institutions known as capacity markets in an effort to use competitive market forces to ensure adequate electricity supply at lowest cost. But capacity markets are driven more by political and bureaucratic judgments than by competition. The manner in which the capacity market is designed to determine demand exemplifies this observation. As there is no natural demand for capacity, Regional Transmission organizations (RTOs) administratively create demand in capacity markets. RTOs derive capacity demand from three components: the capacity requirement based on forecasted peak demand plus an additional margin, the net cost of new entry based on the cost of new facilities entering the market, and the shape of the demand curve. The processes that RTOs use to generate each of these components lack theoretical or analytical justification and tend to produce biased results. As a result, electricity customers are paying billions of dollars for excess capacity that is unnecessary to maintain adequate grid reliability. Capacity markets should address these shortcomings so that demand reflects the actual value of capacity.

[*] Professor of Law, Villanova University Charles Widger School of Law.

[**] Professor of Energy and Environmental Economics and MICASU Faculty Fellow, The Pennsylvania State University. For a more detailed and comprehensive examination of capacity markets, *see* TODD AAGAARD & ANDREW N, KLEIT, ELECTRICITY CAPACITY MARKETS (Cambridge Univ. Press 2022). <u>This article draws from but builds on Chapter 8 of that book.</u>

I. INTRODUCTION

The United States economy runs on enormous quantities of energy, much of which is used in the form of electricity that is generated, transmitted, and distributed through the electricity grid. More than eight thousand power plants produce almost four trillion kilowatt hours of electricity annually, which five million miles of transmission and distribution lines deliver to 150 million electricity customers.[1] The modern electricity grid is so complex that it seems almost impossible that it is able to function as well and as reliably as it does. American households and businesses take for granted that the grid will provide power on demand at all times and in virtually any circumstances short of a catastrophic natural disaster.

In fact, however, the electricity sector and its state and federal regulators invest considerable effort and substantial resources via a variety of policies to ensure that the economy has reliable access to electricity. In some areas of the country that transact electricity through competitive markets, the system operators, known as regional transmission organizations (RTOs), have created capacity markets to support the reliability of the electricity grid. The reliability of the grid encompasses two distinct aspects, resource adequacy and operating reliability.[2] Adequacy, associated with long-term reliability, refers to the electricity system's ability to provide sufficient supply to electricity consumers, even during conditions of peak demand. Operating reliability, associated with short-term reliability, is the system's ability to withstand sudden disturbances such as unexpected outages at large generation plants.[3] When referring to reliability, this article focuses on resource adequacy, as concerns about adequacy create the impetus for capacity markets.

RTOs created capacity markets to ensure that the grid will have sufficient generation capacity to satisfy peak demand in the future, so that the grid continuously provides a reliable supply of electric power. Capacity markets do this by creating an additional revenue stream for resources that, in return for receiving

1. *See* ENERGY INFORMATION ADMINISTRATION, ELECTRIC POWER ANNUAL (2018); ENERGY INFORMATION ADMINISTRATION, FREQUENTLY ASKED QUESTIONS (2019); Jennifer Weeks, *U.S. Electrical Grid Undergoes Massive Transition to Connect to Renewables*, SCI. AM. (Apr. 28, 2010).

2. *See* NORTH AMERICAN ELECTRIC RELIABILITY COUNCIL, 2018 LONG-TERM RELIABILITY ASSESSMENT 5 (2018).

3. *See id*.

payments, incur an obligation to be available to provide power on demand. This additional revenue stream of capacity payments is supposed to encourage construction of new generation and to allow some existing generation to remain in operation, to the extent necessary to achieve adequate reliability.

A capacity market operates as an auction that matches demand and supply to provide the desired level of capacity at the lowest price. Sellers in the market are generators and other resources that supply electric power.[4] Buyers in the market are load-serving entities that sell electricity to end users in retail electricity markets. Each buyer's purchase of capacity in the market increases the overall reliability of the grid. Reliability benefits every user of electricity from the grid. Grid reliability is thus a public good, as the benefit of a capacity purchase accrues to all those who rely on the grid.

What is known as the 'missing money' theory posits that the market cannot rely on private preferences of load-serving entities to purchase the amount of capacity sufficient to attain the optimal level of grid reliability.[5] Accordingly, an RTO that operates a capacity market creates demand in the market administratively so that the market will provide the desired level of capacity, sufficient to meet overall system-wide peak demand for electricity.[6] The RTO then obligates each load-serving entity to buy its quota of capacity in the market, an amount that reflects its share of the system's overall peak electricity demand.

Originally, in what we term *first generation capacity markets*, capacity markets set demand as a fixed quantity deemed sufficient to meet long-term resource adequacy. This created a vertical demand curve, with demand for the capacity product at the same quantity regardless of price. But a vertical demand curve (along with other pathologies) severely impaired the early capacity markets, leading to reforms in the early 2000s that created *second generation capacity markets* with downward-sloping demand curves in which the quantity demanded increases as price decreases. The downward slope comports with the general economic principle of diminishing returns, which posits that the marginal value of a product decreases as the quantity increases.[7]

In organically arising markets, a demand curve should represent the marginal value of the product demanded—that is, a buyer in the market should be

4. In addition to traditional generation such as gas-fired power plants, other contributors to resource adequacy such as demand response, energy efficiency, and transmission can sell into capacity markets. *See* PJM INTERCONNECTION, L.L.C., PJM MANUAL 18: PJM CAPACITY MARKET 15 (Revision 49, Aug. 1, 2021). The premise of this inclusion is that reducing power demand through demand response and enhanced efficiency and adding transmission capability can contribute to resource adequacy just as traditional power generation can.

5. Separately from the resource adequacy policies of RTO capacity markets, many states apply their own resource adequacy policies to regulated utilities, often through what is known as Integrated Resource Planning. *See* Charles B. Howland, *Brightfields: Sustainable Opportunities for Renewable Energy Projects on Environmentally Impaired Lands*, NAT. RESOURCES & ENV'T, Fall 2014, at 41, 43.

6. In addition, the RTOs also operate zonal submarkets for capacity so that areas impacted by transmission congestion will have adequate supply to meet peak demand. *See*, e.g., NextEra Energy Res., LLC v. FERC, 898 F.3d 14, 20 (D.C. Cir. 2018) (noting ISO New England's use of zonal demand curves in its capacity market).

7. *See* N. GREGORY MANKIW, PRINCIPLES OF MICROECONOMICS 443 (7th ed. 2012).

willing to pay as much for the product as the product is worth to the buyer.[8] Application of this framework to capacity markets poses a challenge, because system operators lack reliable information about the marginal value of capacity to electricity consumers.[9]

This leaves capacity market design in a quandary. Markets are formed by the interaction of supply and demand. Here, natural demand for capacity is inadequate, consistent with the missing money theory, because system capacity is a public good.[10] Capacity market design accordingly must create demand administratively. Theoretically demand should reflect marginal value—that is, the incremental benefit to consumers of an additional unit of capacity. But RTOs do not know the marginal value of capacity. So RTOs creating capacity markets have tended to design demand curves based on three factors: a capacity requirement that reflects forecasts of future demand, the cost of new entry, and a slope (which is a function of the desired price elasticity).[11]

The decisions to use these factors to determine capacity market demand were pragmatic judgments unmoored to economic theory.[12] Moreover, each of these factors entails a series of administrative judgments that are subject to dis-

---

8. *See* Todd S. Aagaard & Andrew N. Kleit, *The Complexity Dilemma in Policy Market Design*, 30 DUKE ENVTL. L. & POL'Y F. 1, 87 (2019).

9. The value of capacity is sometimes measured as the value of lost load, representing the cost of an electricity outage. The value of lost load, however, is notoriously difficult to determine. *See* Andreas Bublitz, *A Survey on Electricity Market Design: Insights from Theory and Real-World Implementations of Capacity Remuneration Mechanisms*, 80 ENERGY ECON. 1059, 1060 (2019). The value of lost load depends on various factors such as the length of the relevant outage, how much advance notice consumers receive of the outage, what sector (residential, commercial, industrial) the outage affects, the characteristics and demographics of the affected consumers, as well as the specific location studied. *See* Abhishek Shivakumar et al., *Valuing Blackouts and Lost Leisure: Estimating Electricity Interruption Costs for Households Across the European Union*, 34 ENERGY RESEARCH & SOCIAL SCI. 39, 40 (2017).

10. *See* MANKIW, *supra* note 7, at 216 (defining a public good as something that is non-excludable, people cannot be prevented from using it, and non-rivalrous ,one person's use does not diminish another person's ability to use). National defense and basic research are examples of public goods. *See id*. at 219-20. Markets tend to undersupply public goods because people know they can free ride—that is, obtain the benefit of a public good without paying for it. *See id*. at 218.

11. *See, e.g.*, MANASA KOTHA, CAPACITY ZONES FORMATION AND DEMAND CURVES 24 (ISO New England 2019); JOHANNES PFEIFENBERGER ET AL., REVIEW OF PJM'S RELIABILITY PRICING MODEL (RPM) 43 (Brattle Group 2008).

12. When the RTOs submitted their demand curves to FERC for approval, they provided supporting materials from economists. *See, e.g.*, Supplemental Affidavit of Benjamin F. Hobbs on Behalf of PJM Interconnection, L.L.C. on the September 29, 2006 Settlement Capacity Demand Curve, PJM Interconnection, L.L.C., FERC Docket Nos. ER05-1410-000 & EL05-148-000 (Sept. 29, 2006); *see also* Ming-Che Hu & Benjamin Hobbs, *Dynamic Analysis of Demand Curve Adjustment and Learning in Response to Generation Capacity Cost Dynamics in the PJM Capacity Market*, IEEE Power and Energy Society 2008 General Meeting: Conversion and Delivery of Electrical Energy in the 21st Century (2008) (publishing analysis from affidavit). For the most part, however, the analyses in these supporting materials merely modeled the results of different simulated capacity auctions to determine which alternative demand curves produced the desired balance of capacity procured and cost. In other words, the analyses did not attempt to support the derivation of the demand curves by reference to economic principles. ISO New England did link its demand curve to an analysis of the marginal impact of additional capacity on reliability. *See* Kotha, *supra* note 11, at 23. Because increases in reliability are a measure of the benefit of capacity, this idea comports with economic theory. But ISO New England then alters the reliability-quantity curve "to convert it into a price-quantity curve," without reference to economic principles. *See id*. at 24.

cretion, arbitrariness, and error. The consequences of these judgments are significant. Even small changes in demand curve parameters can cause large changes in capacity market prices and revenues[13].

Given the importance of capacity market demand for outcomes in electricity markets, how demand is determined deserves more attention. Capacity market operators and regulators should assess the methods used to determine capacity market demand and look for ways to bring more accuracy and accountability to those methods. Fortunately, some obvious opportunities exist for improvement. Unfortunately, so far there is no indication that either the RTOs or their regulator, the Federal Energy Regulatory Commission (FERC), will avail themselves of these opportunities.[14] FERC has, however, recently demonstrated a willingness to correct other unwise existing policies, such as its much-maligned expansion of Minimum Offer Price Rules.[15] These developments suggest that the time may be ripe for FERC to initiate a rulemaking to develop a coherent approach to determining capacity market demand based on economic principles.

This article proceeds in four parts. Part I provides the necessary background, explaining the development of competitive electricity markets, the rationale for capacity markets, and the basic elements of capacity markets. The remainder of the article examines the three key components of capacity market demand. Part II addresses the capacity requirement. Part III reviews the cost of new entry. Part IV examines the shape and slope of the demand curve. Each component of capacity market demand exhibits similar flaws: a lack of foundation in economic theory, a prevalence of questionable administrative judgments, and a history of statistically biased outcomes. Regulators should hold the RTOs accountable for these shortcomings and require better market design.

II. BACKGROUND

Capacity markets are embedded in a complex larger system of electricity markets. To understand capacity demand therefore requires some knowledge of the basics of capacity market design and how capacity markets fit into the overall electricity sector. This Part provides that necessary background.

---

13. For example, the model used in Section III.B.3 implies that reducing the demand forecast by one percent would have decreased revenues by approximately 4.7 percent, or about $470 million.

14. FERC primarily regulates RTO capacity markets under sections 205 and 206 of the Federal Power Act, 16 U.S.C. §§ 824d, 824e. Section 205 requires public utilities to file changes to market rules with FERC, which the agency approves if they are "just and reasonable." *Id.* § 824d(a). Section 206 allows FERC to reject an existing rule and impose a new one if it determines the existing rule is unjust and unreasonable. *Id.* § 824d(a). *See generally* ADAM VANN, THE LEGAL FRAMEWORK OF THE FEDERAL POWER ACT 2 (Congressional Research Serv. 2020).

15. *See* Ethan Howland, *PJM's 'Focused' MOPR Takes Effect, Boosting Renewables and Nuclear as FERC Commissioners Deadlock*, UTILITY DIVE (Sept. 30, 2021) (reporting that a deadlock among FERC commissioners resulted in the default approval of PJM's proposal to narrow its Minimum Offer Price Rule). For examples of critiques of FERC's orders expanding the Minimum Offer Price Rules, see, *e.g.*, Todd S. Aagaard & Andrew N. Kleit, *A Road Paved with Good Intentions?: FERC's Illegal War on State Electricity Subsidies*, 33 ELEC. J. 1 (2020); Joshua Macey & Robert Ward, *MOPR Madness*, 42 ENERGY L.J. 67 (2021). In addition to changing its policy on Minimum Offer Price Rules, FERC recently has reversed prior decisions regarding PJM's reserve markets. *See, e.g.*, *PJM Interconnection, L.L.C.*, 177 FERC ¶ 61,209 (2021).

A. *Electricity Market Restructuring*

Capacity markets are a relatively recent addition to the electricity sector. They arose in the late 1990s as part of a larger transition in the industry away from traditional regulation of public utility monopolies and toward competition in wholesale power markets.

Electricity grids first developed as the property of vertically integrated utility companies that owned the power plants, transmission lines, and distribution systems that comprise the electricity system.[16] Although neighboring utilities were interconnected physically, each utility company owned and operated the grid within its service territory.[17] State public utility commissions granted monopoly rights in these territories to utility companies, in exchange for which the companies incurred obligations, including duties to provide uninterrupted service to the public at 'just and reasonable' rates.[18] Federal and state regulators applying the 'just and reasonable' standard employed a cost-of-service approach that set rates for monopolist public utilities based on predicted fixed and variable costs and a reasonable rate of return on capital investments.[19] Recoverable costs and capital investments included the expense of having enough available generation capacity to meet peak demand. The public utility model operating under a cost-of-service regulatory approach is still intact, with modification, in some parts of the United States, primarily in states in the Mountain West, Plains, and Southeast.[20]

Over the last three decades, a series of complementary legal and economic developments at the federal and state level in many states led to the breakup of vertically integrated monopolies and the creation of competitive wholesale generation markets for electricity, in which power plants sold their output to still-regulated electricity distribution companies.[21] In these 'restructured' competitive markets, regulators ensured rates were 'just and reasonable' by creating competitive market conditions rather than by directly regulating rates.[22] Restructuring was intended to harness competitive market forces to reduce electricity prices

16. *See* FERC v. Elec. Power Supply Ass'n, 577 U.S. 260, 267 (2016).

17. *See* Joseph P. Tomain, *The Past and Future of Electricity Regulation*, 32 ENVTL. L. 435, 438 (2002).

18. *See* Alison Gocke, *Nodal Governance of the U.S. Electricity Grid*, 29 DUKE ENVTL. L. & POL'Y F. 205, 216 (2019).

19. *See id*.

20. *See* ENERGY INFORMATION ADMINISTRATION, STATUS OF STATE ELECTRIC INDUSTRY RESTRUCTURING ACTIVITY (Feb. 2003) (showing map with status of each state's electricity sector). As an example of one such modification, states in the Southeast recently established the Southeast Energy Exchange Market, which facilitates bilateral trading among utilities. *See Duke Energy Progress, LLC*, 177 FERC ¶ 61,080, P 2 (2021).

21. *See* Jim Lazar, ELECTRICITY REGULATION IN THE US: A GUIDE 9-10 (Regulatory Assistance Project, 2d ed. 2016).

22. *See Elec. Power Supply Ass'n*, 577 U.S. at 267.

and improve service,[23] but it also added complexity to the grid, which became a network of transactions among numerous firms.[24]

As part of the restructuring process, institutions known as regional transmission organizations (RTOs) were formed to operate the grid and coordinate transactions in competitive markets.[25] The seven RTOs in the United States now encompass all or parts of thirty-eight different states.[26] RTOs are nonprofit membership organizations that decide how to operate the grid within the RTO service territory, subject to regulatory approval from FERC.[27] Many RTO members are buyers and sellers in RTO-governed electricity markets and therefore have a financial interest in the RTO's decisions. Each RTO is responsible for grid stability in its region.[28] With literally billions of dollars at stake, the rules of capacity and other markets are often fiercely disputed.[29] This article focuses on capacity markets in the three RTOs of the Northeast United States—ISO New England, New York ISO (NYISO), and the PJM Interconnection.

RTOs operate several different wholesale electricity markets, including energy markets, ancillary services markets, and capacity markets.[30] In energy markets, generators sell electric power to load-serving entities.[31] Ancillary services markets transact power services that maintain grid stability and security, such as reserve power, reactive power, frequency regulation, and voltage support.[32] The purpose of capacity markets—which are the focus of this article—is to ensure adequate generation resources are available to meet demand for electricity at almost all times.[33]

---

23. *See* REISHUS CONSULTING, LLC, ELECTRIC RESTRUCTURING IN NEW ENGLAND—A LOOK BACK 6 (Dec. 2015).

24. *See* JAMES BUSHNELL, ERIN T. MANSUR & KEVIN NOVAN, REVIEW OF THE ECONOMICS LITERATURE ON US ELECTRICITY RESTRUCTURING 6 (Feb. 23, 2017).

25. *See* Regional Transmission Organizations (Order 2000), 65 Fed. Reg. 810, 811 (Jan. 6, 2000).

26. *See* FEDERAL ENERGY REGULATORY COMMISSION, REGIONAL TRANSMISSION ORGANIZATIONS (Nov. 2015). In addition, two other wholesale market structures, the Western Energy Imbalance Market and the Southeast Energy Exchange Market, operate with some resemblance to RTO markets. *See Alabama Power Co*., 178 FERC ¶ 61,048 (2022); *Cal. Indep. Sys. Operator Corp*., 153 FERC ¶ 61,087 (2015).

27. *See* Regional Transmission Organizations (Order 2000), 65 Fed. Reg. at 811.

28. *See* Del. Div. of Pub. Advoc. v. FERC, 3 F.4th 461, 467 (D.C. Cir. 2021).

29. *See, e.g*., *PJM Interconnection, L.L.C*., 161 FERC ¶ 61,197 P 185 (2017) (noting the Illinois Attorney General's argument that PJM's capacity market rules were "fundamentally unfair to Illinois"); *Astoria Generating Co. L.P*., 140 FERC ¶ 61,189 (2012) (noting the Independent Power Producers' of New York argument that NYISO's proposed capacity market rules "would product an absurd result" and had "no logical basis").

30. *See* Joel B. Eisen, *FERC's Expansive Authority to Transform the Electric Grid*, 49 U.C. DAVIS L. REV. 1783, 1793 n.44 (2016).

31. *See* TransCanada Power Mktg. Ltd. v. FERC, 811 F.3d 1, 10 (D.C. Cir. 2015).

32. *See* Regional Transmission Organizations (Order 2000), 65 Fed. Reg. at 874 (noting that ancillary services "maintain grid reliability"); *New York Indep. Sys. Operator, Inc*., 129 FERC ¶ 61,164 (2009) (listing categories of ancillary services).

33. *PJM Interconnection, L.L.C*., 155 FERC ¶ 61,157 P 28 (2016).

B. Grid Reliability and the 'Missing Money' Problem

In the traditional public utility model, the regulated utilities kept their systems operating by balancing power supply and demand and were responsible for maintaining grid reliability. Monopolist utilities maintained generation capacity sufficient to meet peak demand, including a reserve margin, and recovered the costs of maintaining this capacity and a 'reasonable' rate of return on their capital investments through the regulator-approved rates they charged their customers.[34] Shortages in capacity were not generally a problem.[35] Indeed, one of the critiques of the regulated utility model is that, by virtually assuring regulated firms a return on their capital investments in generation, the model induces over-investment in capacity.[36]

Over time, regulated utilities created physical interconnections, allowing them to transfer power from one utility company to another in a process known as *wheeling*.[37] Sometimes utility companies formed *power pools* to coordinate their generation and transmission operations.[38] These steps enabled utilities to share their generating reserves, which increased efficiency and reliability.[39] But coordination also created interdependencies in reliability across utilities, foreshadowing issues that the advent of electricity competition would pose directly.

In those parts of the country with competitive wholesale electricity markets managed by an RTO, the move to competition forced a change in managing reliability. Because operations and transactions were now occurring across firms rather than within a single firm, no single firm could be held responsible for maintaining adequate capacity. Moreover, with the advent of competition, utilities were no longer assured of earning a return on investments in capacity and instead faced incentives to cut costs. A combination of new regulatory requirements and market forces would have to be established to ensure adequate capacity.

Distinctive features of electricity—the difficulty of storing electricity, a need to balance supply and demand continuously and instantaneously, demand that is not responsive to the costs of production, the use of price caps, and the fact that the reliability of the grid must be managed system-wide, along with other factors—are perceived by scholars, policymakers, and those in the industry as contributing to an underinvestment in generation resources that threatens grid re-

34. *See* JAMES BUSHNELL, MICHAELA FLAGG & ERIN MANSUR, CAPACITY MARKETS AT A CROSSROADS 8 (Energy Institute at Haas, Apr. 2017).

35. *See id*.

36. See Harvey Averch & Leland L. Johnson, *Behavior of the Firm Under Regulatory Constraint*, 52 AM. ECON. REV. 1052 (1962); Richard J. Pierce, Jr., *The Regulatory Treatment of Mistakes in Retrospect: Canceled Plants and Excess Capacity*, 132 U. PA. L. REV. 497, 506 (1984).

37. *See* Joseph P. Tomain, *Electricity Restructuring: A Case Study in Government Regulation*, 33 TULSA L.J. 827, 840 (1998).

38. *See* Mark E. Haedicke, *Competitive-Based Contracts for the New Power Business*, 17 ENERGY L.J. 103, 111-12 (1996).

39. *See* Richard P. Bonnifield & Ronald L. Drewnowski, *Transmission at a Crossroads*, 21 ENERGY L.J. 447, 449 n.6 (2000).

liability.[40] This underinvestment is widely known as the 'missing money' problem.[41] In many, but not all, competitive electricity markets, RTOs have created capacity markets to address the 'missing money' problem by ensuring that the grid will have sufficient generation capacity to satisfy peak demand and thereby to avoid widespread grid failure.[42] The revenues from capacity markets are essentially an incentive payment for capital investment aimed at enhancing grid reliability. Whether or not the 'missing money' problem actually exists, and whether capacity markets are the appropriate means of addressing the "missing money" problem if it does exist, are controversial questions.[43]

C. Capacity Market Demand

Demand in capacity markets poses a fundamental design problem for system operators.[44] The RTOs and their regulator, FERC, created capacity markets to facilitate the procurement of capacity at quantities adequate to attain system reliability. Because system reliability is managed on a system-wide basis, private demand may undervalue it.[45] Without adequate private demand, the RTOs

---

40. *See, e.g.*, Peter Cramton P & Axel Ockenfels, *Economics and Design of Capacity Markets for the Power Sector*, 36 ZEITSCHRIFT FÜR ENERGIEWIRTSCHAFT 115–23 (2012). As technological innovations and falling costs of battery storage add storage capacity to the grid, this may eventually alleviate at least some of the missing money problem by adding flexibility to the supply of electricity.

41. *See, e.g.*, MICHAEL HOGAN, HITTING THE MARK ON MISSING MONEY: HOW TO ENSURE RELIABILITY AT LEAST COST TO CONSUMERS 3 (2016); Peter Cramton, Axel Ockenfels & Steven Stoft, *Capacity Market Fundamentals*, 2 ECON. ENERGY & ENVTL. POL'Y, Sept. 2013, at 27, 30; Emily Hammond & David B. Spence, *The Regulatory Contract in the Marketplace*, 69 VAND. L. REV. 141, 169-70 (2016); David B. Spence, *Naïve Energy Markets*, 92 NOTRE DAME L. REV. 973, 1015 (2017).

42. *See* Jay Morrison, *Capacity Markets: A Path Back to Resource Adequacy*, 37 ENERGY L.J. 1, 44 (2016).

43. *See* TODD AAGAARD & ANDREW N, KLEIT, ELECTRICITY CAPACITY MARKETS (Cambridge Univ. Press 2022).

44. Unlike administratively constructed demand, supply in a capacity market is set by private companies in the market rather than by regulators. Electricity suppliers receive revenues from selling products into several different electricity markets, and therefore capacity market revenues are one of several potential revenue streams for a supplier. If supply in the capacity market is competitive, suppliers can be expected to bid the money they expect they will need to reach a zero economic profit, which can be thought of as the market rate of return.

45. In an ideal market, the reliability of electricity supply would be valued in the market. Reliability would have a price, and each customer would be able to purchase the amount of reliability it desired. In the context of the current electricity grid, however, reliability depends inherently on the overall electricity network. When a generator adds capacity to the grid, this enhances the reliability of the entire network, to every user's benefit. When an electricity user draws power from the grid, this reduces the reliability of the entire network, to every user's detriment. Thus, everyone using the network shares its reliability, and consumers cannot be excluded from sharing reliability. Reliability also is a non-rivalrous good, as the benefits one user receives from the network's reliability do not reduce the benefits another user gains from reliability. These characteristics of non-excludability and non-rivalry mean that reliability is what economists call a public good. *See* Malcolm Abbott, *Is the Security of Electricity Supply a Public Good?*, ELECTRICITY J., Aug./Sept. 2001, at 31; Cramton & Ockenfels, *supra* note 40, at 116-17; Dominique Finon & Virginie Pignon, *Electricity and Long-Term Capacity Adequacy: The Quest for Regulatory Mechanism Compatible with Electricity Market*, 16 UTILITIES POL'Y 143, 143-44 (2008); Paul L. Joskow, *Capacity Payments in Imperfect Electricity Markets: Need and Design*, 16 UTILITIES POL'Y 159, 165 (2008); *see generally* MANKIW, *supra* note 7, at 216. Because the benefits of reliability inure to the entire network, economic theory predicts that the market will underpro-

must create demand administratively. To create demand in capacity markets, the RTOs form demand curves that are intended to produce outcomes similar to a well-functioning competitive market. Like all bureaucratic processes, however, the reality in practice diverges from the theoretical ideal.

Capacity market demand curves in the three Northeast RTOs—PJM, NYISO, and ISO New England—are a function of three elements: the capacity requirement, the Net Cost of New Entry (Net CONE), and the shape of the curve. The capacity requirement drives the horizontal dimension of the curve (quantity), the Net CONE drives the vertical dimension of the curve (price), and the shape determines the relationship between quantity and price. The remainder of this article explains how these elements come together to create capacity market demand and summarizes and evaluates the methodology used to determine each element.

III. CAPACITY REQUIREMENT

Capacity requirements are based on peak electricity demand, with an additional margin for safety. Because the primary objective of a capacity market is to ensure that sufficient capacity exists to satisfy peak levels of demand for electricity, capacity requirements are a key component of the constructed demand curve for capacity. Capacity requirements aim to achieve a level of reliability as measured by the *loss of load expectation*.[46] The loss of load expectation represents the expected frequency of outages caused by supply that does not meet demand.[47] In the United States, a common loss of load expectation is one outage in ten years.[48] This reliability goal is a widely accepted engineering-based standard that has been used for decades, with little inquiry into whether it appropriately balances the benefits and costs of achieving reliability.[49]

Three basic elements comprise a capacity requirement: annual peak load forecast, reserve margin, and resource outage rate.[50] A peak load forecast estimates peak electricity demand over the period in which capacity will be delivered.[51] The reserve margin reflects a judgment as to the amount of capacity beyond the peak load forecast that is necessary to provide the desired level of reliability that meets the loss of load expectation.[52] A typical reserve margin is fifteen percent, meaning that the RTO will seek fifteen percent more capacity

---

vide electricity reliability. *See* Abbott*, supra*, at 33; Cramton & Ockenfels, *supra* note 40, at 116-17; Finon & Pignon, *supra*, at 143-44; Joskow, *supra*, at 65.

46. *See* PJM INTERCONNECTION, L.L.C., PJM MANUAL 18: PJM CAPACITY MARKET 21 (Revision 49, Aug. 1, 2021).

47. *See* PJM Interconnection, L.L.C., PJM Glossary, https://www.pjm.com/Glossary.

48. *See* NORTH AMERICAN ELECTRIC RELIABILITY COUNCIL, 2018 LONG-TERM RELIABILITY ASSESSMENT 17 (2018).

49. *See* JOHANNES PFEIFENBERGER ET AL., RESOURCE ADEQUACY REQUIREMENTS: RELIABILITY AND ECONOMIC IMPLICATIONS 83-84 (Brattle Group, Sept. 2013).

50. *See* PJM INTERCONNECTION, L.L.C., PJM MANUAL 20: PJM RESOURCE ADEQUACY ANALYSIS 31-32 (Revision 11, Aug. 1, 2021).

51. *See* PJM INTERCONNECTION, L.L.C., PJM MANUAL 19: LOAD FORECASTING AND ANALYSIS 13 (Revision 31, June 1, 2016).

52. *See* PJM INTERCONNECTION, L.L.C., *supra* note 46, at 14.

than it thinks it actually will need.[53] The resource outage rate reflects the probability that some resources will not be available to contribute their output.[54] The higher the outage rate, the more capacity the RTO will require to meet its reliability objective—that is, the higher the capacity requirement. Traditionally, the U.S. electricity grid has maintained a high level of reliability. Whether that is appropriate, and in particular whether the benefits of such high levels of reliability justify the costs of maintaining a system with so much reserve capacity, is a matter of some dispute.[55]

Peak demand forecasting has been a common element of electricity regulation for decades.[56] In traditional regulated electricity markets, monopoly utilities maintain enough capacity to meet peak demand, plus a reserve margin, and to recover the costs of maintaining this capacity and a 'reasonable' rate of return on their capital investments through the regulator-approved rates they charge their customers.[57] The regulated utility has incentives to overestimate peak demand, thereby increasing the amount of capacity, in order to justify higher revenues from the ratepayers.[58] Generators have similar incentives in restructured markets with capacity markets. The higher that forecasted peak demand is, the more capacity an RTO will purchase in its capacity market, and the more revenue generators earn.

A. Methodology

Accurate demand forecasting is difficult and requires a series of judgments. A forecasting model entails creating a statistical model that predicts demand peaks as a function of historical variables such as weather (including temperature, humidity, windspeed), population, number of residential households, employment, economic output, day of the week, whether or not the day in question is a holiday, and the stock and efficiency of various electrical equipment.[59] The actual underlying determinants of demand are inherently unknown, so the model uses these data variables as proxies for the actual determinants. Modelers at-

53. *See, e.g.*, *Independent Market Monitor for PJM v. PJM Interconnection, L.L.C.*, 174 FERC ¶ 61,212, P 44 (2021) (noting PJM's target reserve margin of 15.8%). In May 2021, the California Public Utility Commission granted the California ISO's request to increase its planning reserve margin to 17.5% to increase reliability. *See* Hudson Sangree, *CPUC, CAISO Take Major Steps for Summer Reliability*, RTO INSIDER (May 25, 2021).

54. *See* PJM INTERCONNECTION, L.L.C., *supra* note 46, at 28-29.

55. *See, e.g.*, Iulia Gheorghiu, *PJM, NYISO and ISO-NE Pay $1.4B Annually for Excess Capacity: Report*, UTILITY DIVE (Nov. 22, 2019) (noting an estimate that "approximately $1.4 billion per year in total is wasted by the Northeast regional transmission operators and independent system operators by securing a combined 34.7 GW of excess capacity"); NORTH AMERICAN ELECTRIC RELIABILITY CORP., 2021 LONG-TERM RELIABILITY ASSESSMENT (Dec. 2021) (contending that "[c]apacity-based estimates [of reliability] . . . can give a false indication of resource adequacy").

56. *See, e.g.*, Derel W. Bunn, *Forecasting Loads and Prices in Competitive Power Markets*, 88 PROCEEDINGS OF THE IEEE 163 (2000).

57. *See* JOHANNES PFEIFENBERGER, KATHLEEN SPEES & ADAM SCHUMACHER, A COMPARISON OF PJM'S RPM WITH ALTERNATIVE ENERGY AND CAPACITY MARKET DESIGNS 13 (Brattle Group 2009); *see also supra* Part I.A (describing traditionally regulated electricity markets).

58. *See* BUSHNELL, FLAGG & MANSUR, *supra* note 34, at 9-10.

59. *See* PJM INTERCONNECTION, L.L.C., *supra* note 46, at 22-25.

tempt to choose the variables that will best predict the outcome—here, peak demand—but there is no guarantee that they choose the right variables. There is no definitive rule for determining which variables to use, or the form of those variables. In addition to using variables to create a model, forecasters may consider opinions about how demand in an area might change in ways not accounted for by the historical variables in the model.[60]

Forecasters first estimate demand peaks for each zone within the RTO territory.[61] The overall system peak demand is not, however, merely the sum of these zonal peaks.[62] This is because not all zones will reach their peaks at the same time—the peaks are said to be 'non-coincident.'[63] For example, the hour and day at which demand peaks in Chicago may not be the hour and day at which demand peaks in Baltimore. Thus, calculating total peak demand as the sum of individual zonal peaks would bias the estimated system peak upward. To adjust for this problem, forecasters model weather patterns as they affect the entire RTO area, allowing them to model the relationships between demand in different areas within the RTO and how those relationships affect total system demand.[64] This allows for the estimation of the RTO-wide system peak.

Once forecasters have created the model establishing the historical relationship between the chosen independent variables and the dependent variable to be modeled (here, overall system demand), they use the model to predict future system demand based on estimates of the independent variables used in the model—such as weather, population, and number of residential households—for the forecast year.[65] As noted above, to do so requires estimates of economic growth, growth in distributed generation and simulations of weather conditions for the future period in question, including variability in weather conditions.[66] Choosing the values for future variables is inherently difficult, and errors reduce the accuracy of a model's forecasts.

Depending on the RTO, as many as three additional steps may be required to finalize the capacity requirement. First, every RTO includes a safety margin, generally termed the *installed reserve margin*, in case generators are not available, transmission is unusually congested, or demand is above the forecasted peak. For PJM for delivery year 2018/19, for example, the forecasted peak was

---

60. *See* PJM INTERCONNECTION, L.L.C., RESOURCE ADEQUACY PLANNING DEPARTMENT, 2020 LOAD FORECAST SUPPLEMENT 21-28 (Jan. 2020), https://www.pjm.com/-/media/planning/res-adeq/demand-forecast/2020-demand-forecast-supplement.ashx?la=en.

61. *See* PJM INTERCONNECTION, L.L.C., *supra* note 46, at 15.

62. *See id*.

63. *See id*.

64. *See id*.

65. *See* PJM INTERCONNECTION, L.L.C., *supra* note 46, at 17-25.

66. *See* PJM INTERCONNECTION, L.L.C., DEMAND FORECAST DEVELOPMENT PROCESS (2019), https://www.pjm.com/planning/resource-adequacy-planning/load-forecast-dev-process.aspx.

161,418 MW.[67] Since the reserve margin was 15.7 percent, forecasted demand was increased by 15.7 percent to 186,761 MW.[68]

Second, if the forecast model has predicted demand based on installed capacity but the capacity market is conducted in terms of unforced capacity, then the forecast must be translated into unforced capacity.[69] For 2018/19, the estimated forced outage rate in PJM was 6.35 percent.[70] Together the reserve margin and the forced outage rate imply an adjustment to the forecasted peak capacity of $(1 + 0.157) * (1 - 0.0635) = (1.157 * 0.9735)$, which PJM rounded to 1.0835. This results in a capacity requirement of $161{,}418\ \mathrm{MW} * 1.0835 = 174{,}897\ \mathrm{MW}$.

Finally, if the RTO has an opt-out for some resources, such as the fixed resource requirement option in PJM, the capacity that has opted out of the market must be taken out of the capacity requirement.[71] In 2018/19 in PJM, 8.17 percent of resources opted out via the fixed resource requirement, and so the final capacity requirement was reduced by 8.17 percent (or, multiplied by 0.9183) to yield a final capacity requirement of 160,607 MW.[72]

67. *See* PJM INTERCONNECTION, L.L.C., 2018/2019 RPM BASE RESIDUAL AUCTION PLANNING PERIOD PARAMETERS, https://www.pjm.com/-/media/markets-ops/rpm/rpm-auction-info/2018-2019-planning-parameters-report.ashx.

68. *See id*.

69. Installed capacity reflects a generator's theoretical availability based on nameplate output. *See* Keyspan-Ravenswood, LLC v. FERC, 474 F.3d 804, 807 (D.C. Cir. 2007). Unforced capacity reflects a generator's actual availability—that is, it discounts for the probability of an outage that renders a generator unavailable. *See id*.

70. *See* PJM INTERCONNECTION, L.L.C., *supra* note 46. A forced outage is an outage that cannot be controlled, such as a mechanical failure. *See, e.g.*, PJM INTERCONNECTION L.L.C., PJM MANUAL 22: GENERATOR RESOURCE PERFORMANCE INDICES 12 (Revision 18, Mar. 26, 2020) (defining "forced outage" as "[a] complete reduction in the capability of a generating unit due to a failure that cannot be postponed beyond the end of the next weekend"). Forced outages can be distinguished from scheduled outages for inspection or maintenance. PJM INTERCONNECTION L.L.C., PJM MANUAL 10: PRE-SCHEDULING OPERATIONS x (Revision 39, Nov. 19, 2020) ("A Generator Planned Outage is the scheduled removal from service, in whole or in part, of a generating unit for inspection, maintenance or repair with the approval of PJM"). The forced outage rate refers to the percentage of time that a power plant should have been running but was not. *See* G. MICHAEL CURLEY, RELIABILITY ANALYSIS OF POWER PLANT UNIT OUTAGE PROBLEMS 31 (General Consulting Services, LLC, 2013) (defining "forced outage rate" as "[t]he percent of scheduled operating time that a unit is out of service due to unexpected problems or failures" and noting that it "[m]easures the reliability of a unit during scheduled operation"). Scheduled outages and other periods during which a plant is not expected to run are excluded from the forced outage rate calculation.

71. When PJM created its mandatory capacity market in 2006, it included an alternative for load-serving entities that want to opt out of participation in the centralized capacity market. This opt-out option was known as the fixed resource requirement. Instead of purchasing capacity in the market, a load-serving entity can demonstrate to PJM that it owns or procures enough supply to meet its capacity obligation. *See PJM Interconnection, L.L.C.*, 117 FERC ¶ 61,331 (2006). ISO New England and NYISO have somewhat similar programs that allow load-serving entities to meet their capacity obligations outside of the auctions with self-supplied capacity resources. *See* ISO NEW ENGLAND INC., MARKET RULE 1: STANDARD MARKET DESIGN § III.13.1.6 (2020); NEW YORK INDEPENDENT SYSTEM OPERATOR, MANUAL 4: INSTALLED CAPACITY MANUAL 5, 36 (Version 6.46, Apr. 23, 2020).

72. *See* PJM INTERCONNECTION, L.L.C., *supra* note 46. The example calculations in the text apply to the RTOs' system-wide capacity requirement calculations. In addition, the RTOs also analyze demand within specific zones of the RTO territories and then assess, based on the availability of supply through local generation and transmission, whether to operate submarkets within the broader system-wide capacity market. *See*

B. Evaluation

1. Forecasted Demand Consistently Exceeds Actual Demand.

Two different comparisons can be used to evaluate peak demand forecasts. The first is to compare forecasted peak demand to actual peak demand. Because the purpose of a forecast model is to predict outcomes, the most obvious evaluation of the model is to compare its predicted outcome to the actual outcome. For example, for the 2018/2019 delivery year, PJM's model predicted peak demand would be 161,128 MW, and actual peak demand turned out to be 150,565 MW.[73] Thus, actual peak demand was approximately seven percent below forecasted peak demand.

A second evaluation compares forecasted peak demand to weather-normalized peak demand. Weather-normalized demand reflects forecaster's determinations of what actual demand would have been under normal weather conditions.[74] Since demand depends on weather, and weather fluctuates unpredictably, it is in some sense inappropriate to evaluate the model based on how well it predicts demand that has been influenced by weather. Weather-normalized demand attempts to remove the effect of fluctuations in weather, allowing what may be a fairer comparison of predicted peak demand to weather-normalized peak demand. For example, during the 2018/2019 PJM delivery period, weather-normalized peak demand was 149,593 MW,[75] about 7.3 percent below the forecast level of 161,418 MW.[76]

Figure 1 compares PJM's forecasted peak demand to actual and weather-normalized demand since the advent of PJM's current capacity market, known as the Reliability Pricing Model in 2009.[77] For each of the nine years in question, forecasted peak demand is above both actual peak demand and weather-normalized peak demand. On average, forecasted demand is 9 percent above actual peak demand (standard deviation of 5.36 percent) and 7.5 percent above

AAGAARD & KLEIT, *supra* note 43, at 98; *See* Kathleen Spees, Samuel A. Newell & Johannes P. Pfeifenberger, *Capacity Markets—Lessons Learned from the First Decade*, 2 ECON. OF ENERGY & ENVTL. POL'Y 1, 9 (2013).

73. *See* PJM INTERCONNECTION, L.L.C., PJM LOAD FORECAST (1999-2019) (reporting data on PJM forecast, actual, and weather-normalized peak demand); *see also* PJM INTERCONNECTION, L.L.C., LOAD FORECAST DEVELOPMENT PROCESS, https://www.pjm.com/planning/resource-adequacy-planning/load-forecast-dev-process.aspx (reporting forecasts since 2011). Thanks to James Wilson for supplying PJM forecasts from 1999 to 2010.

74. *See* J. STEWART MCMENAMIN, DEFINING NORMAL WEATHER FOR ENERGY AND PEAK NORMALIZATION 3 (Itron Forecasting 2008).

75. *See* PJM INTERCONNECTION, L.L.C., WEATHER-NORMALIZED PEAKS, https://www.pjm.com/-/media/planning/res-adeq/load-forecast/weather-normalized-peaks.ashx .

76. *See* PJM INTERCONNECTION, L.L.C., *supra* note 46.

77. *See* PJM Interconnection, LLC, Proposal for Reliability Pricing Model, FERC Docket Nos. ER05-1410-000 and EL05-148-000 (Aug. 31, 2005); PJM INTERCONNECTION, L.L.C., PJM LOAD FORECAST (1999-2019) (reporting data). No actual or weather-normalized peak demand data are available for the forecasted regions for 2011 and 2012. Duke Energy Ohio and Kentucky, as well as American Transmission Systems, Inc. joined PJM during those years. *See* PJM INTERCONNECTION, L.L.C., PJM HISTORY, https://www.pjm.com/about-pjm/who-we-are/pjm-history.aspx?p=1. PJM did not publish peak demand figures for the areas that matched the previous forecasts, and so no direct comparison of forecasted demand to actual demand is possible. We also exclude 2020 from our analysis, as demand was unexpectedly low during that year due to the COVID-19 pandemic.

weather-normalized peak demand (standard deviation of 2.59 percent). Both differences are statistically significant (t-statistics of 5.04 and 8.65, respectively).

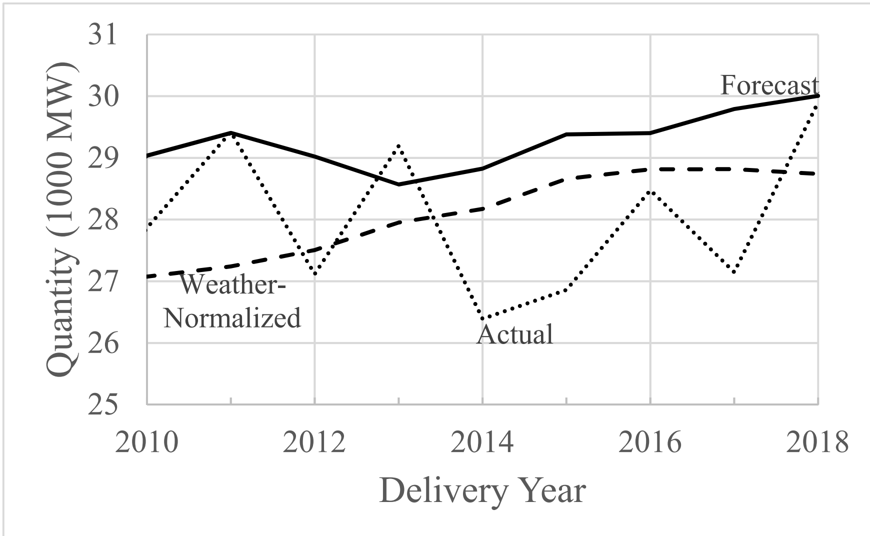

Figure 1: PJM Forecasted vs. Actual Peak Loads, 2008-2019

Figure 2 compares ISO New England's forecasted peak demand to actual and weather-normalized peak demand since the advent of ISO New England's second-generation market in 2010.[78] Forecasted peak demand exceeds actual peak demand in seven of the nine years in question. Forecasted peak demand is above weather-normalized peak demand in each of the nine years examined. On average, forecasted peak demand is 4.55 percent above actual peak demand (standard deviation of 4.52 percent) and 4.2 percent above weather-normalized demand (standard deviation of 2.26 percent). Both differences are statistically significant (t-statistics of 3.05 and 5.55, respectively).

78. *See* ISO NEW ENGLAND, FORECAST REPORT OF CAPACITY, ENERGY, LOADS, AND TRANSMISSION (various years), https://www.iso-ne.com/system-planning/system-plans-studies/celt.

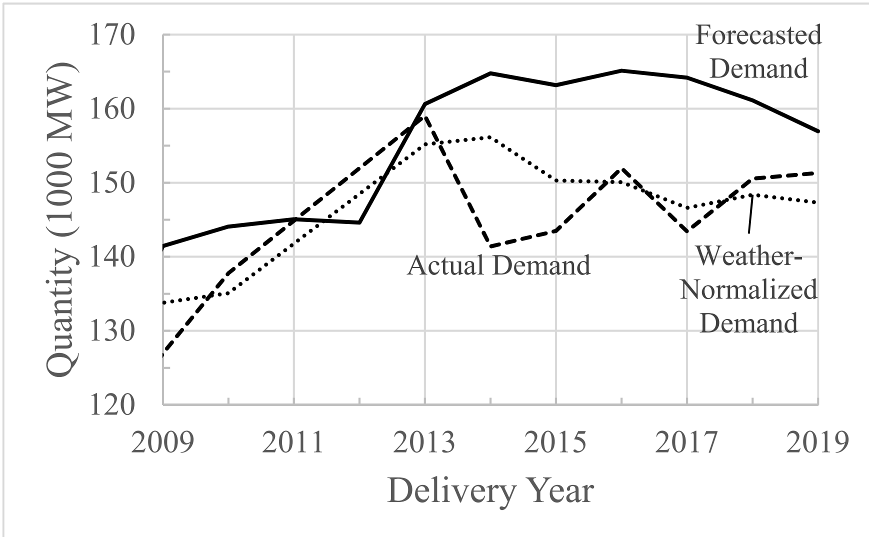

Figure 2: ISO New England Forecasted vs. Actual Peak Demand, 2010-2018

Figure 3 compares NYISO's forecasted peak demand to actual and weather-normalized peak demand since the beginning of NYISO's second-generation market in 2006.[79] The mean difference between forecasted peak demand and actual peak demand is 4.22 percent (standard deviation of 5.23 percent), a statistically significant difference (t-statistic of 3.01). Forecasted peak demand exceeds actual peak demand in ten of the fourteen relevant years.

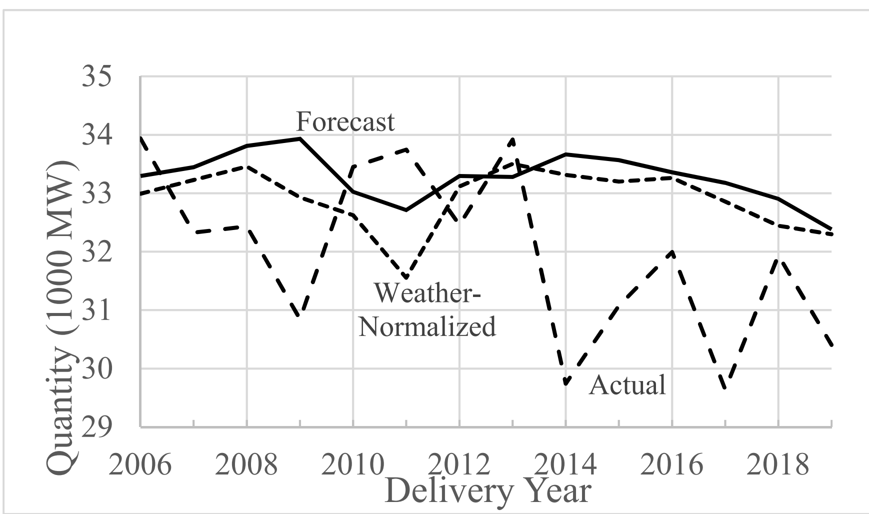

Figure 3: NYISO Forecasted vs. Actual Peak Demand, 2006-2019

79. *See* NEW YORK STATE RELIABILITY COUNCIL, L.L.C., NYSRC NEW YORK CONTROL AREA INSTALLED CAPACITY REQUIREMENT REPORTS, http://ww.nysrc.org/NYSRC_NYCA_ICR_Reports.html.

Although weather-normalized peak demand is greater than forecasted peak demand in all but one of fourteen years, the mean difference is only 1.11 percent. While the difference is statistically significant (standard deviation of 1.09 percent, with a t-statistic of 3.80), this difference is far less than the 7.5 and 4.41 percent found for PJM and ISO New England. The mean forecast errors are 5.38 percent versus actual peak demand and 1.21 percent for weather-normalized peak demand.

2. Demand Forecasts Are Inaccurate and Biased.

Forecast models can be evaluated based on whether they exhibit accuracy and a lack of statistical bias. Accuracy measures how well a model's predictions match actual outcomes.[80] Because no forecast is perfect, accuracy is measured in comparative terms—that is, as among two or more forecast models, which exhibits less error. One common method of measuring forecast errors is to examine a forecast's average percentage absolute error—that is, the mean of the absolute value of the percentage errors of each forecasted value.[81]

A model also can be evaluated based on whether it exhibits statistical bias in its predictions. A good forecast is unbiased—that is, as likely to overestimate as it is to underestimate. Bias is commonly measured as the average percentage by which a model's forecasted values deviate from the actual values.[82]

Based on the data above,[83] PJM's forecasts can be compared to NYISO's forecasts to determine which forecast is more accurate and whether the forecasts are statistically biased. As noted above, PJM overestimated peak demand in every year since the advent of its RPM. Under these circumstances, when all forecasting errors are in the same direction, the average percentage absolute error of the forecasts (a measure of accuracy) equals the average percentage bias of its forecasts (a measure of bias). Thus, PJM's average percentage absolute error and average percentage bias are 9 percent against actual demand and 7.5 percent against weather-normalized demand.

By comparison, NYISO's average overestimation bias is only 4.2 percent against actual peak demand and 1.11 percent against weather-normalized peak demand. These differences are far less than those found for PJM. Because NYISO had one year when its forecast underestimated demand, its average percentage forecast error is slightly higher than its bias. The forecast error is 5.5 percent against actual demand and 1.2 percent against weather-normalized demand. Again, these differences are far less than those for PJM. Thus, NYISO's

80. *See, e.g.*, Jin Li, *Assessing Accuracy of Predictive Models for Numerical Data: Not R nor R^2, Why Not? Then What?*. 12 PLOS ONE 8 e0183250 (Aug. 24, 2017) ("Predictive accuracy should be measured based on the difference between the observed values and predicted values.").

81. *See, e.g.*, Rob J. Hyndman & Anne B. Koehler, *Another Look at Measures of Forecast Accuracy*, 22 INT'L J. FORECASTING 679, 682 (2006).

82. *See, e.g.*, E.L. Lehmann, *A General Concept of Unbiasedness*, 22 ANNALS MATHEMATICAL STATISTICS 587 (1951).

83. *See supra* Figure 1 and Figure 3.

demand forecasts are both much more accurate and far less biased than PJM's forecasts.[84]

With capacity markets, the consequences of bias can be especially harmful. A forecast of peak demand that is biased downward results in a capacity market that potentially procures less capacity than necessary, posing an increased risk of insufficient capacity. A forecast of peak demand that is biased upward results in a capacity market that procures more capacity than necessary, at an increased cost to consumers and potentially increased profits for generators.[85]

The timing of PJM's, ISO New England's, and NYISO's forecasts may explain the differences in their accuracy and bias. PJM and ISO New England operate forward capacity market auctions that run three years before the relevant commitment periods, so demand forecasts for those markets must occur three years ahead of time.[86] In contrast, NYISO's monthly capacity auctions occur just before the commitment periods in question, so its forecasts can occur much closer to the relevant times.

Forecasting demand three years in advance is inherently more difficult than forecasting less than one year in advance. Forecasting for a single year ahead versus forecasting two years earlier means an additional two years' worth of data is available for use in projections. Perhaps more important, the elements in the forecasting model that themselves need to be forecasted, such as employment and economic growth, are likely to be more accurate when determined less than a year in advance.

Based on this data for PJM, ISO New England, and NYISO, it appears running capacity market auctions three years before the delivery years in question poses a serious disadvantage. Thus, if the PJM and ISO New England markets were to eliminate the three-year period between the market auctions and the capacity delivery period, their peak demand forecasts might be more accurate.

3. The Effects of Forecast Errors Are Costly.

As explained above,[87] capacity market demand curves are anchored to the forecasted peak demand. If that forecast is in error, there are real consequences for electricity markets. The capacity market may procure too much or not enough capacity. Forecast errors also may affect the price of capacity and total

84. ISO New England's forecasts are also biased upward. See *supra* note 78 and accompanying text (reporting that ISO New England's forecasted peak demand was 4.55 percent above actual peak demand (standard deviation of 4.52 percent) and 4.2 percent above weather-normalized demand (standard deviation of 2.26 percent), with t-statistics of 3.05 and 5.55, respectively).

85. RTOs sometimes readily admit to reaching judgments that upwardly bias demand, arguing that the costs of procuring too much capacity are much lower than the costs of procuring too little. *See, e.g.*, *PJM Interconnection, L.L.C.*, 149 FERC ¶ 61,183 (2014). The reserve margin embedded in the capacity requirement, however, already protects against uncertainties in demand. Moreover, reliability depends on more than just long-term resource adequacy, as problems with natural gas supply during the February 2021 electricity crisis in Texas illustrated. *See* AAGAARD & KLEIT, *supra* note 43, at 242-45.

86. Theoretically PJM's Incremental Auctions, which it operates between the Base Residual Auctions and the delivery year, could correct some of the overestimation. PJM has used its Incremental Auctions to adjust demand downward, but these are generally not enough to fully correct its overestimations in the Base Residual Auctions. *See id*. at 168-74.

87. *See supra* Part II.A.

cost of capacity, which in turn affects revenues to generators and costs to electricity consumers. Indeed, because of the inelasticity of capacity market demand curves, small changes in demand can lead to relatively large changes in capacity market prices and therefore revenues. In an effort to illustrate some of these effects, this section estimates the impacts of forecast errors on the PJM capacity market.

If the quantities and prices bid by firms were available, a supply curve could be calculated that would allow precise estimation of how forecast errors in capacity requirements affect capacity market prices and quantities. Unfortunately, most likely to protect confidential business data, PJM does not make such data available. The Brattle Group, however, has published pictorial demand curves for various delivery years.[88] (Brattle apparently had access to PJM's internal data.) We derived an approximation of PJM's supply curve from Brattle's pictorial demand curve.

Figure 4 illustrates the approximated PJM supply curve for delivery year 2018-2019. It also represents the demand curve PJM used for the capacity market, based on a capacity requirement of 160,607 MW. In the actual auction, the resulting equilibrium quantity was 166,830 MW and the market clearing price was $164.88/MW-day. This implies capacity market revenues of slightly over $10 billion.[89]

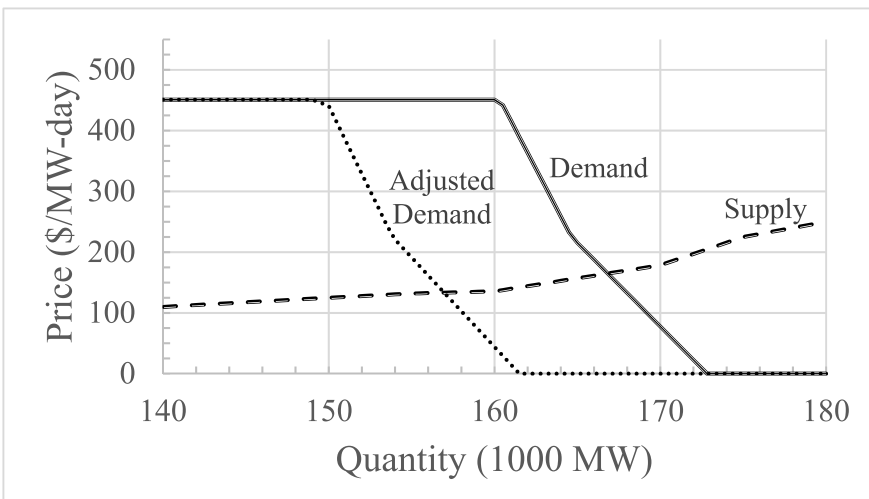

Figure 4: PJM 2018/19 Capacity Market with Adjusted Demand

Figure 4 also illustrates the adjusted demand curve that would have applied had PJM accurately forecasted peak demand at 150,565 MW, or 93.28 percent of its actual forecast. Using the method described above to translate forecasted

88. *See* SAMUEL A. NEWELL ET AL., FOURTH REVIEW OF PJM'S VARIABLE RESOURCE REQUIREMENT CURVE 42 (Brattle Group Apr. 18, 2018).

89. For simplicity purposes, here we model PJM as having one price across its various zones. In fact, some zones in PJM had prices higher than $164.88/MW-day, resulting in higher actual total capacity market revenues for 2018-19.

peak demand to an overall capacity requirement,[90] this would have resulted in a final capacity requirement of 149,808 MW. Combining the adjusted demand curve (based on this capacity requirement) with the approximated supply curve to model the market, the market-clearing outcome would have been an equilibrium quantity of 156,968 MW, with a price of $133.18/MW-day and resulting revenues of $7.62 billion. Thus, using an adjusted demand curve based on accurately forecasted demand, rather than forecasted demand that overestimated by 7 percent, would have decreased quantity by 6 percent, price by 19 percent, and annual revenues by 24 percent, or about $2.4 billion. The market-clearing quantity of 156,968 MW still would have substantially exceeded the capacity requirement of 149,809 MW, maintaining adequate reliability.

Collectively, these data and results suggest that the RTOs—and especially PJM and New England ISO—are systematically overestimating peak demand for electricity. This overestimation leads to an excess quantity of capacity, at costs to consumers of billions of dollars per year.

IV. COST OF NEW ENTRY

The cost of new entry (CONE) is meant to represent the long-run marginal cost of supply in the capacity market.[91] More specifically, CONE attempts to reflect the annualized cost—fixed costs and capital investment—of constructing and operating new generation resources that will add capacity to the grid.[92] The Northeast RTOs use CONE to shape their capacity market demand curves.[93] They define the price ceiling in each capacity market—represented in the demand curve as the horizontal top section—as a multiple of CONE.[94] In addition, they calculate the price level at other points on the downward-sloping portion of the demand curves as multiples or fractions of CONE.[95]

Gross CONE represents the total annual net revenue that a new generation resource would need to recover its capital and fixed costs.[96] This revenue can be

---

90. *See supra* Part II.A.

91. *See* Spees, Newell & Pfeifenberger, *supra* note 72, at 9.

92. *See New York Indep. Sys. Operator, Inc*., 136 FERC ¶ 61,192 P 3 (2011).

93. *See, e.g*., *ISO New England Inc*., 155 FERC ¶ 61,319 P 38 (2016); *New York Indep. Sys. Operator, Inc.*, 156 FERC ¶ 61,039 P19 (2016); *PJM Interconnection, L.L.C*., 167 FERC ¶ 61,029 P 2 (2019).

94. *See New York Indep. Sys. Operator, Inc*., 103 FERC ¶ 61,201 (2003); *ISO New England Inc*., 161 FERC ¶ 61,035 P 16 (2017); *PJM Interconnection, L.L.C*., 167 FERC ¶ 61,029 P 4 (2019).

95. *See ISO New England Inc*., 155 FERC ¶ 61,319 P 35 (2016); NEW YORK INDEP. SYS. OPERATOR, INC., *supra* note 88, at 30; PJM INTERCONNECTION, L.L.C., *supra* note 33, at 39-41 (Revision 49, Aug. 1, 2021). In addition to helping position the demand curves, Net CONE values are also used to set offer price screens for Minimum Offer Price Rules, which attempt to prevent exercises of buyer-side market power. *See* NEWELL ET AL., *supra* note 88, at 1; *see generally* Macey & Ward, *supra* note 15.

Although the Net CONE shapes the current downward-sloping demand curves, the CONE predates the second-generation capacity markets. In the first generation of capacity markets, which employed vertical demand curves, load-serving entities that failed to meet their capacity requirements were assessed deficiency charges based on CONE. *See New York Indep. Sys. Operator, Inc*., 103 FERC ¶ 61,201 (2003); *PJM Interconnection, L.L.C*., 115 FERC 61,079 (Apr. 20, 2006).

96. *See* NEWELL ET AL., *supra* note 88, at iii. Gross CONE excludes variable costs. *See Panda Stonewall L.L.C.*, 174 FERC ¶ 61,266 P 181 n.371 (2021). Variable costs are expected to be recovered in the energy market, because resources should not bid into the energy market at less than their variable costs.

earned in any of the available electricity markets, including the energy market, capacity market, and ancillary services market.[97] To determine the *capacity market revenue* necessary to sustain a new generation resource, Net CONE subtracts from Gross CONE the annual revenues above its variable costs that the new resource would be expected to earn in the energy and ancillary services markets:[98]

$$\text{Net CONE} = \text{Gross CONE} - \text{Energy and Ancillary Services Revenues.}$$[99]

Net CONE thus is intended to equal the amount of 'missing money' that necessitates capacity market remuneration.[100]

A 2013 FERC staff report on capacity market design aptly observed that "calculating a value for CONE requires a number of estimations and assumptions that can be contentious."[101] At least two factors stoke controversies over CONE estimation. First, the stakes are high. The CONE value significantly affects capacity market outcomes, especially capacity prices, and therefore capacity revenues. Capacity market sellers tend to favor calculations that will lead to higher CONE values and therefore higher capacity market prices. Capacity market buyers tend to favor lower CONE values that result in lower capacity market prices. Second, as explained below, calculating CONE is also controversial because it involves numerous judgments on questions for which there is no clear answer.

A. Methodology

The stated objective of the CONE is to estimate the costs of a new plant in the capacity market.[102] The RTOs estimate Net CONE administratively by evaluating the costs of constructing and operating a hypothetical new generation resource.[103] The determination of CONE thus depends on all the factors that influence the costs of a new plant, such as plant location, technology, and configuration; engineering, procurement and construction costs; other development costs; and the cost of capital. Each of these factors involves multiple judgments that may affect the overall Net CONE estimate.

97. *See supra* Part I.A (discussing the organized electricity markets operated by RTOs).

98. *See* PJM INTERCONNECTION, L.L.C., *supra* note 33, at 39.

99. *See PJM Interconnection, L.L.C.*, 149 FERC ¶ 61,183 (2014).

100. *See* JAMES F. WILSON, OVER-PROCUREMENT OF GENERATING CAPACITY IN PJM: CAUSES AND CONSEQUENCES 6 (2020); Feng Zhao, Tongxin Zheng & Eugene Litvinov, *Constructing Demand Curves in Forward Capacity Market*, 33 IEEE TRANSACTIONS ON POWER SYSTEMS 525, 531 (2018); *see also supra* Part I.B (describing the 'missing money' problem).

101. *See* FEDERAL ENERGY REGULATORY COMMISSION STAFF, REPORT NO. AD13-7-000, CENTRALIZED CAPACITY MARKET DESIGN ELEMENTS 10 (Aug. 23, 2013).

102. *See* SAMUEL A. NEWELL ET AL., PJM COST OF NEW ENTRY COMBUSTION TURBINES AND COMBINED-CYCLE PLANTS WITH JUNE 1, 2022 ONLINE DATE 2 (Brattle Group and Sargent & Lundy Apr. 19, 2018).

103. *See* NEWELL ET AL., *supra* note 88, at 5; NEWELL ET AL., *supra* note 72, at 4; NEW YORK INDEPENDENT SYSTEM OPERATOR, PROPOSED NYISO INSTALLED CAPACITY DEMAND CURVES FOR CAPABILITY YEAR 2017/2018 AND ANNUAL UPDATE METHODOLOGY AND INPUTS FOR CAPABILITY YEARS 2018/2019, 2019/2020, AND 2020/2021: NYISO STAFF FINAL RECOMMENDATIONS 1 (Sept. 15, 2016).

CONE is intended to reflect the costs of new generation in a generic sense, not the situation of any specific project.[104] Thus, estimating CONE involves creating a hypothetical competitive generation resource—known as a *reference resource*—that is meant to reflect past industry experience as well as projected future market conditions.[105] The analysis calculates the costs of the resource using a 'bottom-up' approach, so called because it estimates the total fixed costs of a resource as the sum of all expenditures required to construct the resource and bring it into operation.[106] The calculation requires numerous highly specific judgments that affect the cost of the reference resource. Some of these questions include the following:

- where the plant will be located;
- whether the plant will be constructed on a greenfield or a brownfield site;
- how much site preparation will be necessary;
- what technology the plant will use;
- how the plant will be interconnected to gas and electric infrastructure;
- whether the plant will utilize a backup fuel;
- what pollution control equipment and practices the plant will employ;
- what kind of evaporative cooling technology the plant will include; and
- how much of each type of material (e.g., concrete, masonry, steel, piping, electrical, instrumentation, insulation, painting, furnishings) will be used.[107]

These judgments collectively yield detailed specifications for the reference resource, which then must be converted into costs, requiring an additional suite of judgments. Once all these calculations are complete, the total fixed costs of a project are translated to an annualized value, which is the Gross CONE.

The 'bottom-up' tabulation of costs proceeds step by step. First, the analysis estimates the capital costs of the reference resource. The analyst selects locations for the reference resource based on areas in which new power plants have been built recently and are likely to be built in the future.[108] The analyst then selects the technical specifications for the reference resource, including characteristics such as turbine model, size, net heat rate, and environmental controls.[109] The analyst estimates the plant capital costs associated with constructing and developing a resource with the chosen characteristics. Plant capital costs are those costs incurred when constructing the power plant before the plant begins operat-

104. *See* NEWELL ET AL., *supra* note 88, at 2.

105. *See ISO New England Inc.*, 161 FERC ¶ 61,035 P 16-17 (2017).

106. *See PJM Interconnection, L.L.C.*, 167 FERC ¶ 61,029 (2019)

107. *See* CONCENTRIC ENERGY ADVISORS, ISO-NE CONE and ORTP Analysis 20-27 (Jan. 13, 2017).

108. *See* CONCENTRIC ENERGY ADVISORS, *supra* note 107, at 17; NEWELL ET AL., *supra* note 88, at 10-12.

109. *See* CONCENTRIC ENERGY ADVISORS, *supra* note 107, at 17-19; NEWELL ET AL., *supra* note 88, at 13; NEW YORK INDEP. SYS. OPERATOR, INC., *supra* note 92, at 3-10.

ing. Plant capital costs include owner-furnished equipment, such as the gas turbines; engineering, procurement, and construction (EPC) costs, which include other equipment, labor, and materials; and non-EPC costs, which include development costs, startup costs, interconnection costs, and inventories.[110]

Second, the CONE analysis estimates annual fixed operating and maintenance costs of the relevant power plant.[111] These costs include plant operation and maintenance, property taxes, insurance, and asset management. Variable operating and maintenance costs are not included in CONE, but they are relevant to the revenues a resource needs to earn in the energy and ancillary services markets to be financially viable.[112]

Third, the analysis uses discounting to translate the total upfront capital costs and other fixed costs of the plant into an annualized value, which is the Gross CONE. A discount rate converts the uncertain flows of future costs into a net present value. To select an overall discount rate, the RTOs use the after-tax weighted-average cost of capital method, which considers factors such as the corporate income tax rate, debt-equity ratio of project financing, cost of debt, and return on equity.[113]

Finally, the CONE analysis estimates the expected annual revenues the reference resource would earn in the energy and ancillary services markets beyond recouping its variable costs.[114] This value, known as the energy and ancillary services offset, is subtracted from the Gross CONE to yield the Net CONE. Estimating revenues in these markets depends on factors such as energy prices, ancillary services prices, fuel prices, the heat rate of the reference resource, and assumptions about how the reference resource would bid and be dispatched in these markets. None of these questions is simple to answer, and the process and standards by which RTOs address these issues is complex and often opaque.

Once the RTO's analysts have completed their CONE analysis, the analysis is reviewed through the RTO's stakeholder decision making processes.[115] The RTO may make significant changes to its analyst's recommendations.[116] In reviewing proposed changes to a CONE, FERC accords broad discretion to an RTO's judgments and seldom rejects a proposed CONE.[117]

---

110. *See* CONCENTRIC ENERGY ADVISORS, *supra* note 107, at 20-27; NEWELL ET AL., *supra* note 88, at 21-22; NEW YORK INDEPENDENT SYSTEM OPERATOR, INC., *supra* note 92, at 13-17.

111. *See* CONCENTRIC ENERGY ADVISORS, *supra* note 107, at 29-30; NEWELL ET AL., *supra* note 88, at 30-31.

112. See SAMUEL A. NEWELL ET AL., GROSS AVOIDABLE COST RATES FOR EXISTING GENERATION AND NET COST OF NEW ENTRY FOR NEW ENERGY EFFICIENCY at ii (Mar. 17, 2020).

113. *See* CONCENTRIC ENERGY ADVISORS, *supra* note 107, at 38; NEWELL ET AL., *supra* note 88, at 35; NEW YORK INDEPENDENT SYSTEM OPERATOR, INC., *supra* note 92, at 20.

114. *See* CONCENTRIC ENERGY ADVISORS, *supra* note 107, at 49-65; NEWELL ET AL., *supra* note 88, at 19-30; NEW YORK INDEPENDENT SYSTEM OPERATOR, INC., *supra* note 92, at 22-29.

115. *See* ISO New England Inc., ISO New England Inc., Demand Curve Design Improvements, FERC Docket No. ER16-1434, at 14 (Apr. 15, 2016); PJM INTERCONNECTION, L.L.C., QUADRENNIAL REVIEW OF VRR CURVE PARAMETERS: PJM PRELIMINARY RECOMMENDATIONS 2 (May 2, 2018).

116. *See* PJM Interconnection, Periodic Review of Variable Resource Requirement Curve Shape and Key Parameters, PJM Interconnection, LLC, FERC Docket No. ER19-105, at 17 (Oct. 12, 2018).

117. *See, e.g.*, *PJM Interconnection, L.L.C.*, 149 FERC ¶ 61,183 (2014) (accepting PJM's proposed CONE values over the objections of intervenors); *ISO New England Inc.*, 175 FERC ¶ 61,172 P 16 (2021) (ac-

Given the burdens of estimating CONE, RTOs do not undertake new estimations every year. PJM, for example, estimates new CONE values every four years, with escalation rates applied to adjust the CONE in the intervening years.[118] The escalation rates are based on historical real escalation rates for various costs such as land, equipment and materials, and labor, which are then added to a forecasted inflation rate to yield nominal escalation rates.[119]

B. Evaluation

The highly detailed and specific methodology that RTOs follow in developing CONE estimates can give the process an air of scientific rigor. In reality, however, CONE estimations are rife with potentially arbitrary judgments and prone to serious error. The entire CONE process would benefit from significant changes so that capacity market demand curves can better reflect market forces and contribute to the integrity of the markets.

1. Estimating CONE Involves Indeterminate Judgments.

The process of estimating CONE requires making a series of discretionary judgments on which there is no clear answer. Take, for example, the selection of a technology for the reference resource on which the CONE is based. Much of the controversy over a CONE estimate often centers on the choice of technology for the reference resource.[120] But guidance regarding how to select a reference technology is sparse and scattered.

Methodological uncertainty regarding selection of a reference technology appears rooted in ambiguity about precisely what the CONE is supposed to represent. Conceptually, because Net CONE is meant to embody the long-run marginal cost of supply in the capacity market, the reference resource should correspond to the long-run marginal market-clearing resource in the capacity market. Commenters, including a 2013 FERC staff report, have accordingly linked CONE to the marginal capacity resource that clears the market.[121] Moreover, the RTOs use their Net CONE estimations as if this were the case. Each of the capacity market demand curves includes a point—sometimes called a reference point—near the quantity of the capacity requirement and the price of Net CONE.[122] The expectation is that this reference point will be the long-term equilibrium market-clearing price and quantity. This implies that Net CONE is the

cepting ISO New England's proposed CONE values over the objections of intervenors); *New York Indep. Sys. Operator, Inc*., 136 FERC ¶ 61,192 P 71 (2011) (accepting NYISO's proposed CONE values over the objections of intervenors).

118. *See* NEWELL ET AL., *supra* note 88, at 1.

119. *See* NEWELL ET AL., *supra* note 88, at 29. The RTOs estimate Net CONE locally for each zone as well. *See id*. at 16.

120. *See, e.g*., *PJM Interconnection, L.L.C*., 167 FERC ¶ 61,029 P 31-62 (2019); *New York Indep. Sys. Operator, Inc*., 125 FERC ¶ 61,299 (2008).

121. *See* Jonathan Falk, *Capacity Markets: Prices vs. Quantities*, 7 NERA ECONOMIC CONSULTING ENERGY MARKET INSIGHTS 1, 2 (2010); FEDERAL ENERGY REGULATORY COMMISSION STAFF, *supra* note 101, at 32.

122. *See New York Indep. Sys. Operator, Inc*., 144 FERC ¶ 61,126 P 7, 31 (2013).

annual capacity market revenue needed for the marginal market-clearing resource to be economically viable.

The criteria that RTOs articulate to guide their selection of a reference technology do not, however, point to choosing a marginal market-clearing resource. Indeed, despite the strong similarities in the processes they use to estimate CONE, each RTO has described the selection of a reference technology differently, and none of the RTOs appears to have explained or justified its selection criteria by reference to any discernible economic theory. Consider the following examples:

- PJM has indicated that the technology should be "representative of a peaking unit in the energy market that derives a significant portion of its revenues from the capacity market."[123]
- PJM's independent consultant, the Brattle Group, has recommended selecting a reference technology based on five factors: (1) the technology is technically feasible; (2) the technology is economically viable; (3) the technology has a standard set of characteristics and costs; (4) the costs of the technology can be estimated with confidence; and (5) the technology will stay viable as a reference technology.[124]
- ISO New England has articulated a three-factor inquiry for choosing a reference technology: (1) the technology is "likely to be developed in New England"; (2) the RTO "could develop cost and revenue estimates . . . with confidence"; and (3) the technology "should produce prices high enough to meet the reliability standard but not so high as to add unnecessary costs."[125]
- ISO New England also has stated that it chose a reference resource that represents "the technology that is expected to be the most economically efficient and that is commercially available to new capacity suppliers."[126]
- NYISO's tariff specifies that the reference technology must be an economically viable peaking unit with the lowest fixed costs and highest variable costs.[127]

FERC has approved CONE estimations based on these various formulations, without adding its own clarification or unified standard.[128]

---

123. *See* PJM INTERCONNECTION, L.L.C., QUADRENNIAL REVIEW OF VRR CURVE PARAMETERS: PJM PRELIMINARY RECOMMENDATIONS 4 (May 2, 2018).

124. *See* JOHANNES P. PFEIFENBERGER ET AL., THIRD TRIENNIAL REVIEW OF PJM'S VARIABLE RESOURCE REQUIREMENT CURVE 28 (Brattle Group, May 15, 2014).

125. *See ISO New England Inc.*, 161 FERC ¶ 61,035 P 38 (2017).

126. *See ISO New England Inc.*, 170 FERC ¶ 61,052 P 7 (2020).

127. *See New York Indep. Sys. Operator, Inc.*, 125 FERC ¶ 61,299 (2008).

128. The lack of scrutiny FERC gives to an RTO's choice of reference technology is aptly illustrated by a 2017 order in which FERC observed that "ISO-NE's Tariff is not prescriptive as to how the reference technology should be chosen" but nevertheless concluded that ISO New England's selection of a combustion turbine as the reference technology was "consistent with the requirements of the Tariff." *ISO New England Inc.*, 161 FERC ¶ 61,035 P 37 (2017). In other words, ISO New England's choice complied with its nonexistent requirements.

None of these criteria is based on market principles, and none points clearly to marginal market-clearing plants. Several of the criteria refer to choosing a reference technology that is "feasible," "viable," "likely to be developed," "economically efficient," and "commercially available"—characteristics that describe any resource that clears the market, not marginal market-clearing resources in particular. Other criteria focus on whether costs can be estimated with confidence. That may help to estimate costs associated with a given technology, but it does nothing to ensure that the RTO has chosen the right technology to analyze in the first place. Further, the references to peaking units, or similarly to high variable costs, do not necessarily focus on marginal market-clearing plants. The rationale for capacity markets suggests a focus on peaking units,[129] but a peaking unit is not necessarily marginal in the capacity market just because it clears the market.[130]

Some economists have posited that, in theory, the long-term cost of each technology should be equal for the marginal unit for each technology that will be competitive in the market.[131] If that holds true, then the selection of a resource technology might not affect the CONE estimation. But the results of actual CONE estimations do not reflect this theory, as costs can vary significantly across technologies. For example, in the PJM market in 2018, the Net CONE for a simple combustion turbine was 8-12 percent higher than a Net CONE for a combined-cycle plant.[132] These intermediate judgments can cumulatively make a large difference in the CONE, and there is no clear framework guiding RTOs in exercising their discretion in making the judgments. The result is a cacophony of decisions without clear rationales.

CONE estimations at PJM and ISO New England illustrate the discordance. A core question in each process was the selection of a reference technology. When ISO New England selected a reference technology in its 2017 Net CONE calculation, it chose a simple combustion turbine.[133] In recommending a combustion turbine as the reference technology, ISO New England's consultant, Concentric Energy Advisors, noted that a combustion turbine was "substantially less expensive" than other technologies, was well established in the New England region, and had participated in and cleared recent capacity auctions in New England.[134] Concentric concluded that "the simple cycle technology is a cost-effective technology that has gained commercial acceptance and is economically viable in New England." ISO New England adopted Concentric's recommendation, noting the combustion turbine's low cost relative to other technologies and

129. The 'missing money' theory that serves as a justification for capacity markets tends to focus on inadequate revenues for peaking plants in the energy market. *See, e.g.*, PETER CRAMTON & STEVEN STOFT, THE CONVERGENCE OF MARKET DESIGNS FOR ADEQUATE GENERATING CAPACITY 3 (2006).

130. Here, the marginal unit in the capacity market will be the resource with the highest accepted bid, which therefore sets the market price.

131. *See* PFEIFENBERGER ET AL., *supra* note 124, at 27.

132. *See* NEWELL ET AL., *supra* note 88, at iv.

133. *See ISO New England Inc.*, 161 FERC ¶ 61,035 P 19-21 (2017).

134. *See* CONCENTRIC ENERGY ADVISORS, supra note 107, at 5-6.

market conditions in the New England region that favored the development of combustion turbine plants.[135] FERC approved the choice.[136]

By contrast, PJM also chose a combustion turbine as its reference technology in 2018 but for entirely different reasons and under much different circumstances. In its 2018 review of PJM's CONE, the Brattle Group recommended adopting a natural gas-fired combined-cycle turbine as the reference technology for CONE, because of its lower costs and prevalence in new generation.[137] Combined-cycle turbines have dominated new generation in PJM since 2005.[138] As of 2018, over the previous ten capacity auctions—that is, since the auction held in 2010 for Delivery Year 2012/13—28,181 MW of new combined-cycle plants had entered the market, versus just 3725 MW of combustion turbine plants.[139] The Brattle Group's recommended Net CONE for combined-cycle turbines was, depending on the zone, between 25 and 63 percent below Brattle's recommended updated Net CONE for a combustion turbine.[140] Given the cost disparity, Brattle noted that going forward combustion turbines might not even remain competitive in the PJM market.[141]

PJM rejected the Brattle Group's recommendation and decided to continue basing its Net CONE on the combustion turbine technology.[142] Despite the lower per-megawatt cost and prevalence of combined-cycles, PJM reasoned that combustion-turbine plants, as peaking plants, depend on capacity market revenue more than combined-cycle plants do.[143] In addition, PJM argued that combustion-turbine plants could be brought to market less expensively (on a per-plant basis) and more quickly than combined-cycles; that cost estimates of combined-cycles were more uncertain than for combustion turbines; and that both NYISO and ISO New England continued to use combustion turbines as their reference technology. FERC deferred to PJM's position, which it deemed reasonable.[144] It noted that two combustion turbine plants had entered the PJM market since 2014—but did not note the over seven and a half times as many megawatts of combined-cycle plants that were added during the same period.

That ISO New England and PJM could reach such different decisions after applying such different standards in choosing their reference technologies demonstrates the indeterminacy of the CONE process. In choosing its reference technology, ISO New England relied heavily on precisely the factors—lower CONE and commercial viability—that PJM rejected in making its choice. FERC

---

135. *See ISO New England Inc.*, 161 FERC ¶ 61,035 P 19-21 (2017).

136. *See ISO New England Inc.*, 161 FERC ¶ 61,035 P 36-39 (2017).

137. *See* NEWELL ET AL., *supra* note 88, at 17.

138. *See id*. at 5.

139. PJM INTERCONNECTION, L.L.C., 2021/2022 RPM BASE RESIDUAL AUCTION RESULTS 22 (May 23, 2018).

140. *See* NEWELL ET AL., *supra* note 88, at 17.

141. *See id*. at 33.

142. *See* PJM INTERCONNECTION, L.L.C., QUADRENNIAL REVIEW OF VRR CURVE PARAMETERS: PJM PRELIMINARY RECOMMENDATIONS 4 (May 2, 2018); *PJM Interconnection, L.L.C.*, 167 FERC ¶ 61,029 P 33 (2019).

143. *See* PJM INTERCONNECTION, L.L.C., *supra* note 116, at 128.

144. *See PJM Interconnection, L.L.C.*, 167 FERC ¶ 61,029 P 59, 61 (2019).

readily approved both decisions, without trying to clarify the standard or explain how both decisions could be permissible interpretations of CONE.

State public advocates and an environmental organization petitioned the D.C. Circuit for review of FERC's approval of PJM's proposed CONE, raising two arguments.[145] First, the petitioners argued that FERC "erred by not applying its 'established framework' for evaluating an RTO's choice of Reference Resource."[146] According to petitioners, in previous cases FERC had evaluated a proposed reference resource based on three factors: "(1) whether the unit is likely to be developed in the region, (2) whether cost and revenue estimates for that unit can be developed with confidence and (3) whether the [demand curve] produces prices high enough to meet the reliability standard while not adding unnecessary costs."[147] The court rejected this argument, noting that FERC had applied these factors in just one prior case and therefore they did not represent an established framework the Commission was obligated to apply in every case involving a choice of reference resource.[148] FERC's lack of a consistent or coherent approach to choosing a reference resource thus allowed the agency to continue acting inconsistently without exceeding its statutory discretion.

Second, the petitioners argued that FERC's approval of a combustion turbine as the reference resource was unjust and unreasonable.[149] The court rejected this argument as well. Just as FERC had deferred to PJM's selection despite the predominance of combined cycle natural gas plants among new capacity in the PJM region, the court deferred to FERC's approval of a combustion turbine as the reference resource.[150] The court reached this conclusion while conceding that "PJM's proposed combustion turbine plant resulted in a VRR Curve over four times more protective than the Reliability Requirement envisions" and "costs consumers $140 million more each year."[151] If FERC wants to approve an outcome that adds excess reliability at significant cost, the Federal Power Act gives it discretion to make that policy decision.[152]

145. *See* Del. Div. Public Advocate v. FERC, 3 F.4th 461 (D.C. Cir. 2021); *see also, e.g.*, PJM Interconnection, L.L.C., 128 FERC ¶ 61,157, 61,778 (2009) (rejecting a similar challenge arguing that PJM had overestimated Net CONE by choosing the wrong reference technology). In addition, the petitioners also argued that FERC erred in approving a ten percent adder for the Net CONE value. *See* 3 F.4th at 468. PJM rules allowed generation resources to bid into the energy market at 10% above their variable costs to account for cost uncertainties, and in developing its Net CONE PJM included a ten percent adder. The court agreed with the petitioners that the adder was arbitrary and capricious, because the record did not support the conclusion that a resource of the type represented by the reference resource would bid above its variable costs. *See id*. at 469.

146. *See id*. at 465.

147. *See id*. (citing ISO New England Inc.,147 FERC ¶ 61,173 PP 32–33 (2014)).

148. *See id*. The DC Circuit further noted that FERC had concluded that PJM's proposal was just and reasonable even applying the factors. *See id*. at 465 (citing *PJM Interconnection, L.L.C*., 171 FERC ¶ 61,040 P 14 (2020).

149. *See id*.

150. *See id*. at 467; *see generally* Emily Hammond, *Double Deference in Administrative Law*, 116 COLUM. L. REV. 1705 (2016) (evaluating deferential judicial review of agency decisions that in turn defer to nongovernmental standard-setting organizations).

151. *See id*. at 467-68.

152. *See id*. at 468.

Given FERC's broad discretion under section 205 of the Federal Power Act, the D.C. Circuit may have been correct in holding that the Commission acted within its authority in approving PJM's Net CONE, despite the inconsistency between PJM's approach and ISO New England's approach to its Net CONE just two years earlier.[153] But even if FERC's decisions were arguably legally defensible, they were bad policy. In the face of uncertainties in estimating Net CONE, FERC should not simply always err on the side of more reliability; this approach costs electricity customers hundreds of millions of dollars each year, with little in the way of benefit to show for it.

FERC and the RTOs need a coherent policy for how to determine an appropriate value for Net CONE. Given the frequency with which this issue arises—each RTO reassesses its Net CONE every few years, and each reassessment requires FERC approval—and the lack of a consistent methodology for choosing a reference resource, FERC should undertake a rulemaking process to develop a predictable and thoughtful approach to choosing a reference resource.

2. CONE Is Consistently Overestimated.

CONE estimations also can be evaluated on how well they match market results. The long-term capacity market-clearing price should equal the Net CONE.[154] This is because, if the capacity market is meeting its objective of inducing new resources to enter the market with the quantity of capacity necessary to meet capacity requirements, then the capacity price should equal the additional revenue—beyond that earned in other electricity markets—necessary to induce new resources to enter the market.

Again, reality does not match the theory. Capacity prices in all three Northeast RTOs are consistently well below Net CONE. The Brattle Group calculates that market-clearing prices in the PJM capacity auction have been on average sixty percent below PJM's Net CONE.[155] Wilson similarly calculates that, over the period from Delivery Year 2015 to 2021, PJM's Net CONE was more than double the three-year running average market-clearing price.[156]

Figure 5 compares Net CONE values to market prices for two areas within PJM. The first is the generally unconstrained area in PJM (generally referred to as "RTO" or "rest of RTO"), usually centered on Ohio. The second is the Mid-Atlantic Area Council ("MAAC") area, centered on Philadelphia.

---

153. Although under section 205 FERC does not have authority to revise an RTO's proposal, *see* Advanced Energy Mgmt. All. v. FERC, 860 F.3d 656, 662 (D.C. Cir. 2017), the Commission has the power to adopt substantive standards under section 205 and section 206 and then evaluate proposals under section 205 based on whether they comport with the agency's standard. *See, e.g.*, Promoting Wholesale Competition Through Open Access Nondiscriminatory Transmission Services by Public Utilities, Order No. 888, 61 Fed. Reg. 21,540, 21, 541 (1996).

154. *See* NEWELL ET AL., *supra* note 88 at 15 (Apr. 18, 2018); WILSON, *supra* note 100, at 6.

155. *See* NEWELL ET AL., *supra* note 88, at 4.

156. *See* WILSON, *supra* note 100, at 6.

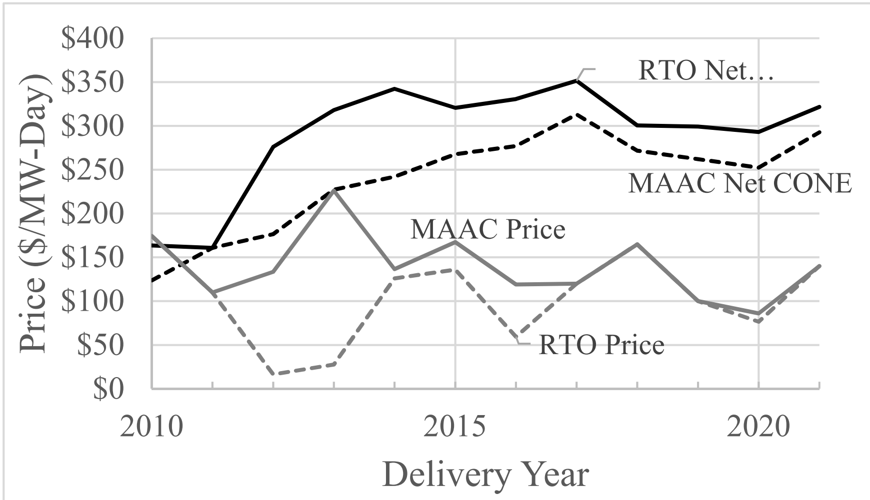

Figure 5: PJM Net CONE vs. Market Price, 2010-2021

As Figure 5 indicates, at the beginning of PJM's second-generation market, Net CONE and market prices were fairly close to each other. This quickly changed, however. Over the time period examined, RTO market prices averaged only 32.5 percent of RTO Net CONE values (t-statistic 7.59). Over the same period, MAAC market prices averaged less than 55 percent of Net CONE values (t-statistic 4.98). Net CONE values were greater than market prices for every market year in both zones since 2010.

NYISO has reported that its market-clearing prices "have been consistently below forty percent" of its Net CONE.[157] Our analysis is consistent with that observation. Since the outset of its capacity market, NYISO has divided its territory into three regions for capacity purposes: Upstate New York, New York City, and Long Island.[158] In 2013, NYISO split off an area in the Lower Hudson Valley from the Upstate region.[159] Figure 6 presents the Net CONE values and the twelve-month running average capacity market prices for Upstate and New York City since the outset in 2005 of the NYISO second-generation market.[160]

157. *See* NEW YORK INDEPENDENT SYSTEM OPERATOR, INC., 2018 ANNUAL INSTALLED CAPACITY REPORT 6 (2018).

158. *See Central Hudson Gas & Electric Corp.*, 88 FERC ¶ 61,138 (1999).

159. *See* New York Indep. Sys. Operator, Inc., 144 FERC ¶ 61,126 (2013).

160. NYISO operates monthly capacity auctions. *See* NEW YORK INDEPENDENT SYSTEM OPERATOR, MANUAL 4: INSTALLED CAPACITY MANUAL 157 (Version 6.49, May 2021). Because these prices are highly seasonal, we present twelve-month running averages in Figure 6.

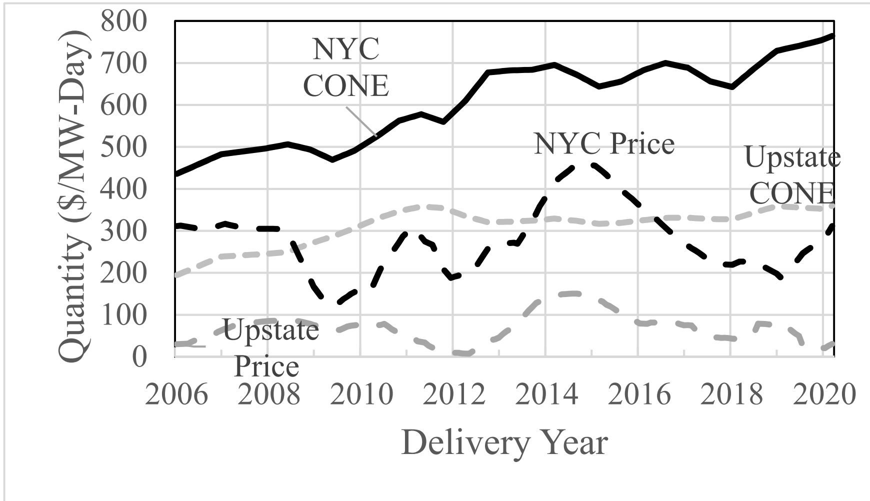

Figure 6: NYISO Net CONE vs. Market Price, 2005-2020

As Figure 6 indicates, the twelve-month running average prices (as well as monthly prices) for New York City and for Upstate New York have been below their corresponding Net CONE for each of the 177 months since 2005. During this period, the average capacity market price for New York City was $278.19/MW-day, 45.8 percent of the average Net CONE value of $607.73/MW-day. (The t-statistic for the differences is 11.73.) For Upstate New York, the average capacity price was $67.79/MW-day, only 21.9 percent of the average CONE value. (The t-statistic for the difference is 18.77.)[161]

Interpreting ISO New England market data is somewhat more complex. Prior to delivery year 2018/2019, ISO New England used a vertical market demand curve with minimum and maximum prices. The result was that in each year the capacity market price was determined administratively. Since that time, there have been six annual auctions. In these auctions the market price averaged slightly less than 55 percent of Net CONE. The market prices ranged from 24 to 86 percent of Net CONE, as displayed in Figure 7.[162]

161. For simplicity of presentation, we do not present prices or Net CONE values for either Long Island or the Lower Hudson Valley area. For Long Island, the average price of $126.10/MW-day was 33.2 percent of the average CONE of $379.87 (t-statistic 10.25). For the Lower Hudson Valley, the average market price of $192.42 was 38.3 percent of the average CONE of $502.28.

162. *See* ISO NEW ENGLAND INC., FORWARD CAPACITY MARKET PARAMETERS (Mar. 6, 2020), https://www.iso-ne.com/markets-operations/markets/forward-capacity-market/; ISO NEW ENGLAND INC., RESULTS OF ANNUAL FORWARD CAPACITY AUCTIONS (2020), https://www.iso-ne.com/about/key-stats/markets#fcaresults.

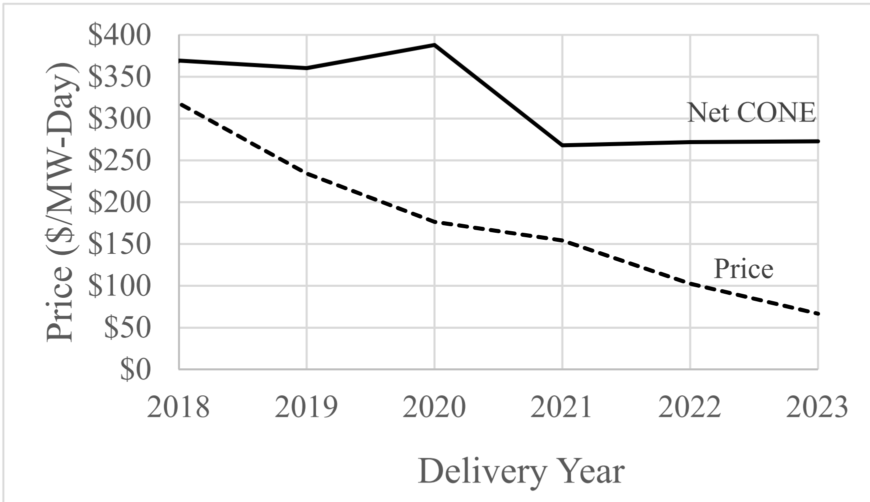

Figure 7: ISO New England Net CONE vs. Market Price, 2018-2023

The RTO capacity markets are exceeding their reliability objectives. Substantial quantities of new generation have cleared the auctions, resulting in overall capacity levels more than sufficient to meet reliability requirements.[163] Given that the capacity markets are clearing plenty of capacity at prices much lower than Net CONE, the RTOs' Net CONE values are obviously too high.

3. The Effects of CONE Overestimation Are Costly.

As with peak demand forecasts, models can estimate the effects of overestimating Net CONE on market outcomes. Our model employs the same approximated PJM supply curve described above that was used in modeling the impact of different peak demand forecasts.[164] Using this supply curve, we compared the market results of the demand curve PJM used for the 2018/2019 auction with the modeled results of an adjusted demand curve based on a Net CONE that was 50 percent lower. We reduced the value of Net CONE by 50 percent as a rough approximation of the proper CONE. As our purpose is to show the sensitivity of market results to the value of CONE, the exact level of CONE is not critical for this analysis.

Figure 8 illustrates the modeled demand and supply curves. Recall that, in the actual 2018/2019 PJM auction, the equilibrium quantity was 166,830 MW and the price was $164.88/MW-day, implying total capacity market revenues of slightly over $10 billion. Using the adjusted demand curve (based on a 50 percent Net CONE value), the market-clearing outcome would have been an equilibrium quantity of 163,233 MW with a price of $149.40/MW-day, resulting in

163. *See* NEW YORK INDEPENDENT SYSTEM OPERATOR, INC., 2018 ANNUAL INSTALLED CAPACITY REPORT 6 (2018); WILSON, *supra* note 100, at 7.

164. *See supra* Part II.B.3.

revenues of slightly less than $9 billion. Thus, using an adjusted demand curve based on a 50 percent Net CONE level, rather than the overestimated actual Net CONE, would have decreased quantity by 2.2 percent, price by 9.4 percent, and annual revenues by 11.3 percent, or about $1.1 billion. The equilibrium quantity of 163,233 MW still would have significantly exceeded PJM's capacity requirement of 160,607 MW.

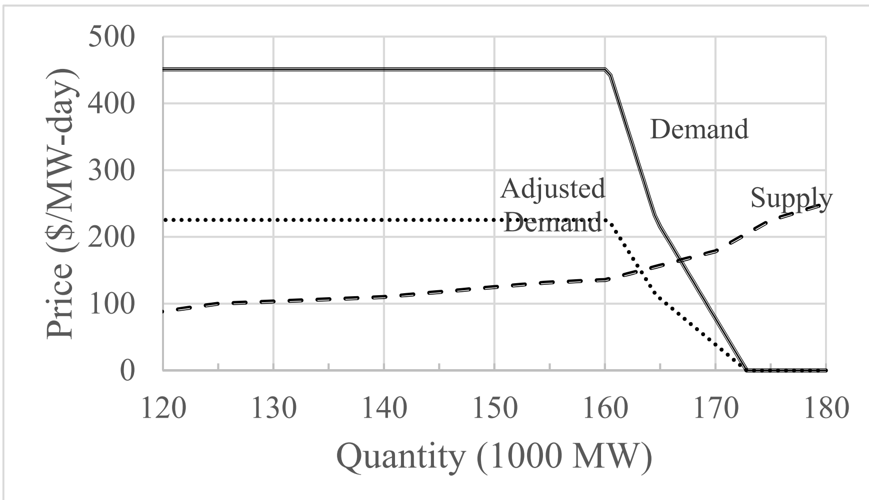

Figure 8: PJM 2018/2019 Capacity Market with Adjusted Demand

Although substantial, the effects of a lower Net CONE value on market outcomes are smaller than the effects of an accurate peak demand forecast as modeled above. As Figure 4 illustrates, an accurate peak demand forecast shifts the demand curve to the left, which affects price and especially quantity significantly. By contrast, as Figure 8 illustrates, a lower Net CONE value shifts the demand curve down, which has a smaller effect on price and quantity.[165]

4. The Entire CONE Methodology Is Flawed.

Consistent overestimation of Net CONE values across all three Northeast RTOs strongly suggests that the entire methodology for calculating Net CONE is fundamentally flawed. The RTOs calculate Net CONE through a bottom-up, engineering-based administrative process based on judgments about a hypothetical new power plant. This process is filled with challenging decisions that are difficult to review and appears to yield Net CONE values that are consistently biased upward.

The administrative method that RTOs use to estimate Net CONE in many ways resembles the process that FERC and state public utility commissions em-

165. If our model uses both actual peak demand rather than forecasted peak demand and a Net CONE equal to 50 percent of the Net CONE PJM used, the resulting equilibrium would have a price of $129.53/MW-day, a quantity of 153,233 MW, and revenues of $7.24 billion. This would have been nearly 28 percent below actual revenues.

ployed to set cost-of-service rates prior to restructuring and that state commissions continue to use in those states that have not restructured their electricity markets.[166] Net CONE estimations and cost-of-service rate cases have different objects of analysis—Net CONE estimations analyze the fixed costs of a single power plant, whereas rate cases analyze the total costs of an entire public utility—but their similarities are otherwise remarkable. Like the CONE analysis, a cost-of-service ratemaking uses a bottom-up administrative process to estimate the costs of production in order to determine the revenues necessary for financial viability.[167] Like the CONE analysis, a cost-of-service ratemaking relies on hypothetical expenses and projected market conditions.[168] Like the CONE analysis, complicated questions regarding the cost of capital financing play a major role in rate cases. The answers to these questions are difficult to unravel and play a major role in rate cases.[169] While it is understandable that utility regulators have adopted new processes that resemble their traditional methods, this continuity runs contrary to the goal of restructuring, which was to replace complicated administrative processes with markets.

The Net CONE is intended to represent an annualized amount of money that would induce a competitive new generation resource to enter the market. CONE estimation uses a cumbersome and opaque administrative process to estimate the costs of a hypothetical plant. This entire complex process is unnecessary. Instead, the value of the Net CONE could be determined more accurately and easily by reference to an empirical measure of the actual cost of new entry.

An empirically derived Net CONE would bring several advantages over the current method of administrative CONE estimation. Estimating CONE empirically—for example, as a multi-year running average market-clearing price[170]—would add integrity to capacity markets. Net CONE is a crucial parameter driving the shape and position of the capacity market demand curve, and therefore an important determinant of market outcomes. Because an administrative Net CONE estimates the future costs of a stylized hypothetical plant, it is inherently unclear whether the process has yielded the 'right' answer. An empirical Net CONE based on market-clearing prices has by definition cleared the capacity market at quantities sufficient to meet reliability requirements. In other words,

166. *See supra* Part I.A (discussing the transition from traditional ratemaking to restructured electricity markets). In addition, FERC still uses cost-of-service ratemaking for transmission, and states use cost-of-service ratemaking for distribution, neither of which is conducive to competition. *See* James W. Moeller, *Public Utilities and Environmental Justice: Electric Restructuring and Deregulation and Low-Income Communities*, 21 U.D.C. L. REV. 1, 4 (2019)

167. *See* A. Lawrence Kolbe & William B. Tye, *The Duquesne Opinion: How Much "Hope" Is There for Investors in Regulated Firms?*, 8 YALE J. REG. 113, 117 (1990); Michael E. Small, *A FERC Electric Rate Primer*, 5 ENERGY L.J. 107 (1984).

168. *See* Inara Scott, *Teaching an Old Dog New Tricks: Adapting Public Utility Commissions to Meet Twenty-First Century Climate Challenges*, 38 HARVARD ENVTL. L. REV. 371, 383 (2014).

169. *See* David A. Lander, *Public Utility Rate Design: The Cost of Service Method of Pricing*, 19 ST. LOUIS UNIV. L.J. 36, 37 (1974).

170. *See* WILSON, *supra* note 100, at 6. Wilson critiques PJM's Net CONE values by comparing them to a more accurate empirical Net CONE calculated from a three-year running average capacity price, but he does not directly propose that PJM should use an empirical CONE methodology.

an empirical Net CONE has proved adequate to induce investment in new generation at least cost—which is exactly what the Net CONE is supposed to do.

An empirical CONE would greatly simplify the process for selecting CONE values. Instead of a complicated calculation built on numerous judgments about a hypothetical future power plant, CONE would be easily calculated from already existing market data.

Taking the judgments out of the CONE process also would reduce the politics in the process. Under the current method, administrative CONE estimations are often highly controversial, because (a) they involve numerous judgments as to which there is no clear answer; and (b) CONE significantly affects capacity market prices, which in turn affects the revenues of generators and the costs of load-serving entities. The existing administrative process for estimating CONE requires RTOs to make a series of controvertible judgments, to which FERC has given the RTOs wide latitude. Although putatively designed as an expert-driven, bottom-up calculation, the stakeholder politics of RTO decision making create opportunities for rent-seeking and political jockeying in the CONE estimation. Switching to an empirical CONE would remove these arbitrary judgments and replace them with an empirical calculation, and in doing so would take away the opportunity for stakeholders to influence the CONE to their advantage.

A crucial shortcoming of the current CONE process is that RTOs are not being held accountable for consistently overestimating CONE. This lack of accountability allows the overestimation to continue unabated. An empirical CONE, by contrast, would be self-correcting, thus automatically adding built-in accountability to the process. An empirical CONE would admittedly not be entirely accurate—costs change, and previous costs do not perfectly predict future costs. But given the dismal record of current estimation methods, an empirical CONE likely would do better. At the very least, an empirical methodology would be less subject to manipulation and rent seeking through the political stakeholder process.

The RTOs do have some experience with using empirical data to estimate CONE inputs. Until 2020, for example, PJM based its energy and ancillary services offset on the three previous years of historical data.[171] In 2020, FERC required PJM to change this approach to an estimate based on forecasted revenues in the energy and ancillary services market, noting that using an empirical estimate "based on three years of historical data is easily distorted by anomalous market conditions in one year that are not representative of what market participants can expect in future delivery years."[172] Then, in 2021, FERC, now under Chairman Glick's leadership, reversed that conclusion, concluding that the record did not support a finding that PJM's backward-looking historical approach was unjust and unreasonable, and restored PJM's empirical approach.[173]

Even if the RTOs are not going to replace their administrative CONE estimations with empirical CONE measures, FERC should use empirical market results to hold the RTOs accountable for their CONE estimations. FERC should

171. *See* PJM Interconnection, L.L.C., 171 FERC ¶ 61,153 P 282-83 (2020)

172. *See id*. at P 313.

173. *See PJM Interconnection, L.L.C.*, 177 FERC ¶ 61,209 P 25 (2021).

require RTOs to explain divergences between their administrative estimates and empirical results and to take concrete steps toward achieving more accurate forecasts. Regrettably, FERC has not shown either the willingness or ability to carry out this modest oversight task. Instead, when presented with objections noting the inaccuracies of CONE forecasting, FERC has claimed the need to protect reliability—contrary to results showing that the RTOs have more than enough capacity to meet their reliability goals, and without reference to the harm to consumers.[174]

There is some basis for optimism here, however. Now-FERC Chairman Glick dissented from the agency's 2020 order approving PJM's proposed Net CONE, noting that "[t]he last few years have provided mountains of evidence that PJM's Net CONE figure is much too high."[175] Under Chairman Glick's leadership, FERC has reversed some of its previous decisions.[176] Perhaps Chairman Glick can invigorate FERC's review of the RTOs' Net CONE proposals with greater scrutiny.

V. SHAPE AND SLOPE

Part II and Part III examined the capacity requirement and the Net CONE, both of which are important inputs in creating a capacity market demand curve. The third crucial component of capacity demand is the shape and slope of the demand curve, which determine the quantity of capacity demanded at each price. Despite their importance, however, the shape and slope of capacity market demand curves have received little explanation and justification from RTOs and FERC.

A. Methodology

Price elasticity—that is, the responsiveness of demand to price changes—determines the slope of the demand curve. The more sensitive demand is to price, the flatter the curve. The less sensitive demand is to price, the steeper the demand curve. In organically arising demand curves, each point on the curve represents the marginal value of the good in question to consumers at a certain quantity.[177] Demand curves slope downward—that is, the marginal value decreases as quantity increases—because the benefits to consumption decrease as quantity increases.[178]

In contrast, a fixed capacity requirement in which the amount of capacity does not vary with the price of capacity creates the equivalent of a vertical demand curve.[179] In such a capacity market, the quantity is known, and the market-

174. *See supra* notes 132-134 and accompanying text.

175. *See* Dissenting Statement of Commissioner Richard Glick, PJM Interconnection, L.L.C., 171 FERC ¶ 61,040, P 2 (2020).

176. *See supra* note 97 and accompanying text.

177. See Ian Ayres & John Braithwaite, *Partial-Industry Regulation: A Monopsony Standard for Consumer Protection*, 80 CAL. L. REV. 13, 31 (1992).

178. *See* W.E. Johnson, *The Pure Theory of Utility Curves*, 23 ECON. J. 483, 492 (1913).

179. See Benjamin Hobbs et al., *A Dynamic Analysis of a Demand Curve-Based Capacity Market Proposal: The PJM Reliability Pricing Model*, 22 IEEE TRANSACTIONS ON POWER SYSTEMS 3, 4 (2007).

clearing price will be the bid of the marginal supplier necessary to meet the requisite demand. Early capacity markets tended to use a fixed capacity requirement.[180]

Capacity markets have moved to a downward-sloping demand curve, which offers advantages over a vertical demand curve. First, a sloped curve more accurately reflects the reality that the marginal contribution of a unit to reliability declines as the amount of capacity in the market increases.[181] Second, creating a demand curve in which demand decreases as price increases reduces price volatility and makes it more difficult for suppliers to earn monopoly profits by withholding capacity in the hopes of inflating the market-clearing price.[182] That said, a downward-sloping demand curve is more difficult to create and to administer than a vertical demand curve, as it requires judgements about CONE and the specific shape of the curve. The difference between a downward-sloping demand curve and a vertical demand curve is not as great as it may sound at first; most downward-sloping demand curves for capacity markets have a steep slope centered on the capacity requirement quantity.

Putting each of these factors together, Figure 9 illustrates an example of a demand curve for a capacity market, showing the role of each factor—capacity requirement, Net CONE, and slope—in determining the curve. At capacity levels well below the capacity requirement, demand is constant at a price equal to a multiple of Net CONE—here, 150% of Net CONE. As the quantity approaches the capacity requirement, the curve slopes downward very steeply—but not entirely vertically. At these quantities, demand is highly unresponsive to price—that is, quantity changes only slightly in response to changes in price. The demand price at the capacity requirement is equal to Net CONE. At levels of capacity substantially above the capacity requirement, price is equal to zero, reflecting the very low marginal value of entirely excess capacity.

---

180. *See* FEDERAL ENERGY REGULATORY COMMISSION STAFF, *supra* note 101, at 5.

181. *See New York Indep. Sys. Operator, Inc*., 103 FERC ¶ 61,201 (2003).

182. *See id.*; *PJM Interconnection, L.L.C.*, 117 FERC ¶ 61,331 (2006).

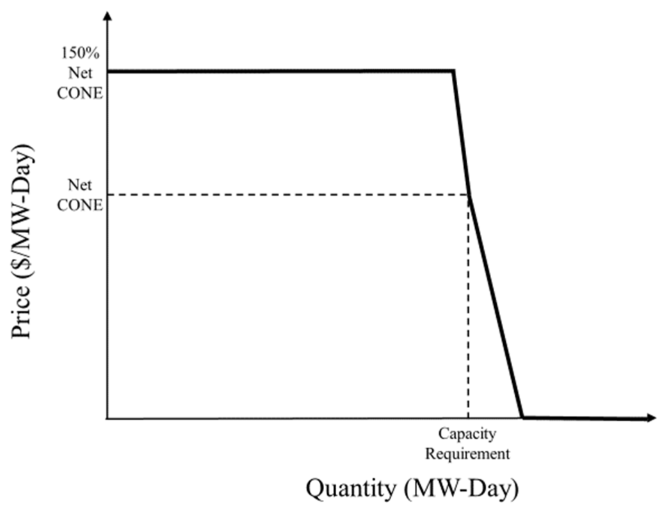

Figure 9: Example of Capacity Market Demand Curve

Downward-sloping demand curves were a key innovation of the Northeast RTOs' second-generation capacity markets, intended to reduce the problems that completely price-insensitive fixed capacity requirements caused in first-generation capacity markets. The specific shapes of the PJM, NYISO, and ISO New England demand curves shared important commonalities as well as differences. Overall, they exhibited a similar shape, with three distinct regions. At quantities from zero to near the capacity requirement, the curves were horizontal at a price fixed to a multiple of Net CONE. At quantities from slightly below the capacity requirement to slightly above it, the demand price decreased steeply. At quantities significantly above the capacity requirement, the demand price was zero.

The horizontal tops of the demand curves essentially set a price cap on the capacity market. The price cap advances a pragmatic objective—to limit price spikes in the event of low supply. For example, in an uncompetitive auction in which all supply resources clear the market without meeting demand, prices could skyrocket without a price cap.[183] The price cap is not, however, rooted in economic theory or evidence of the actual value of capacity. Whereas the price-capped demand curves represent the marginal value of capacity as constant at lower quantities, the actual marginal value of capacity presumably continues to increase significantly as quantities decrease, to values well above the cap. There is no reason to believe that the actual marginal value of capacity is the same at, for example, ninety percent of the capacity requirement, which would be sufficient to meet demand almost every day of the year, as it is at fifty percent, which would likely cause almost daily outages.

183. *See* ISO New England Inc., ISO New England Inc., Demand Curve Changes, FERC Docket No. ER14-1639, at 545-46 (Apr. 1, 2014).

B. Evaluation

1. Capacity Demand Curves Differ Arbitrarily.

The RTOs adopted downward-sloping demand curves as an improvement over the vertical demand curves of the first-generation capacity markets.[184] But although the demand curves adopted a downward-sloping shape, the slopes are quite steep, over a relatively narrow range of quantities. Thus, while the downward-sloping curves introduced some price responsiveness, the effect was limited. RTOs were not willing to let the market-clearing quantity fall much below the capacity requirement deemed necessary to meet reliability standards. They also were wary of creating demand for quantities much in excess of the capacity requirement. Overall, these concerns narrowed the range of capacity quantities deemed acceptable, which dictated a steeply sloped demand curve.

Economists measure the sensitivity of demand to price by calculating the relevant arc elasticity of demand.[185] Arc elasticities of less than one are considered inelastic and represent demand for which quantity is relatively insensitive to price.[186] Over the range of quantities on the demand curve where price is responsive to quantity, the arc elasticity of demand for PJM is 0.0375, and for ISO New England it is 0.057. For upstate New York, the arc elasticity of demand is 0.087, while for the flatter New York City curve (see below) the arc elasticity of demand is 0.129. Thus, each of the demand curves used in the second-generation capacity markets are highly inelastic.

While the demand curves of the three Northeast RTOs have their similarities, they also differ substantially. NYISO's demand curve is flatter than PJM's and ISO New England's. PJM's original second-generation demand curve was slightly concave,[187] whereas both NYISO's and ISO New England's were linear. As discussed below, ISO New England has since moved toward creating a convex demand curve based on the marginal impact of capacity on reliability. PJM has adopted a slightly convex demand curve.

In addition, the relationship between the demand curves in capacity-constrained zones and the total demand in the RTOs differs across RTOs. In PJM, each zonal demand curve has the same shape as the system demand

---

184. *See supra* note 168 and accompanying text; *see also New York Indep. Sys. Operator, Inc.*, 103 FERC ¶ 61,201 (2003) (explaining the rationale for NYISO's creation of a downward-sloping demand curve); *PJM Interconnection, L.L.C.*, 117 FERC ¶ 61,331 (2006) (explaining the rationale for PJM's creation of a downward-sloping demand curve).

185. *See* R.G.D. Allen & A.P. Lerner, *The Concept of Arc Elasticity of Demand*, 3 REV. ECON. STUD. 226 (1934).

186. Arc elasticity measures the percentage change in quantity as a function of the percentage change in price. This measure is unitless and defined as follows: If P_1 and Q_1 are the price and quantity, respectively, of one point on a demand curve, and P_2 and Q_2 are the price and quantity of another point on the same curve, then the arc elasticity of demand between the two points is $[(Q_2-Q_1)/(Q_1+Q_2)]/[(P_1-P_2)/((P_1+P_2)]$. In the context here, let (P_2,Q_2) be the zero-price point. That implies that P_2 equals zero, the denominator $(P_1-P_2)/((P_1+P_2)$ equals one, and the relevant arc elasticity of demand is $(Q_2-Q_1)/(Q_1+Q_2)$. Arc elasticity is defined similarly for a supply curve. *See* PETER M. SCHWARZ, ENERGY ECONOMICS 64 (2018).

187. A concave demand curve becomes steeper (less price sensitive) as quantity increases; a convex curve becomes flatter (more price sensitive) as quantity increases.

curve.[188] In ISO New England, zonal demand curves and the system demand curve have different shapes, but both are based on the same methodology of analyzing the marginal reliability impacts of capacity.[189] In NYISO, the overall shape of the zonal and system-wide demand curves are similar, but the slopes of the different demand curves are determined through stakeholder negotiation.[190]

2. Capacity Demand Curves Are Not Supported by Economic Theory.

Beyond the general concept of a downward slope consistent with declining marginal benefits of additional capacity, there seems to be no theory supporting the shape of capacity market demand curves. Theoretically a demand curve should represent the marginal benefit to the buyer of the product, which is how much the buyer should be willing to pay at the margin for the product.[191] Because system reliability is a public good, however, electricity consumers are not willing to pay the full value of reliability to them—instead they can free ride off others. This free rider problem could cause reliability to be undersupplied.[192]

System operators nevertheless could approximate the marginal benefit of capacity by estimating the marginal value of additional reliability at different capacity quantities. When ISO New England considered such an approach early in the second-generation capacity markets, however, it rejected it on the ground that the value of reliability (measured in terms of the Value of Lost Load) was too difficult to determine and that the curve might not meet traditional reliability standards.[193]

As an alternative to deriving a demand curve from estimates of the marginal benefit of capacity, the RTOs chose to adopt curves built on the capacity requirement and Net CONE.[194] The premise of the curves was that the long-term equilibrium price and quantity should be close to the additional revenue necessary to attract into the market a plant whose capacity was necessary to meet reliability objectives. Beyond the difficulties of overestimating both the capacity requirement and CONE,[195] this methodology creates two problems. First, it yields only a single point—the intersection of the capacity requirement and Net CONE—and a single point does not create a curve. Second, it is based on cost, rather than benefit, and cost is a factor underlying supply, not demand. Instead of tying their capacity market demand curves to economic theory, the RTOs supported their demand curves by showing that they yielded acceptable results in

---

188. *See* PJM INTERCONNECTION, L.L.C., *supra* note 33, at 41.

189. *See ISO New England Inc*., 155 FERC ¶ 61,319 P 6 (2016).

190. *See New York Indep. Sys. Operator, Inc*., 110 FERC ¶ 61,201 (2005).

191. *See* SCHWARTZ, *supra* note 186, at 40.

192. *See supra* note 34 and accompanying text (discussing reliability as a public good).

193. *See* Spees, Newell & Pfeifenberger, *supra* note 72, at 11; Steven Stoft, ISO New England, Inc., Prepared Direct Testimony on Behalf of ISO New England, FERC Docket No. ER03-563, at 11 (2004). ISO New England subsequently in 2016 adopted a demand curve based on the marginal value of reliability, as described below.

194. *See* Stoft, *supra* note 193, at 10.

195. *See supra* Part II.B.1 and Part III.B.2.

terms of ease of administration, cost, and reliability under a range of likely conditions.[196]

The NYISO zonal and system-wide demand curves illustrate this pragmatic approach. The NYISO linear curves are anchored at two points—the reference point, which is at the capacity requirement for quantity and Net CONE for price, and the zero crossing point, which is at a quantity at which the value of additional capacity has been asserted to be zero.[197] Further, the shape of the demand curves differs across zones within NYISO. Thus, since the beginning of its second-generation market, the zero-crossing point for the upstate zone has been 112 percent of the capacity requirement, while the zero-crossing points for New York City and Long Island have been at 118 percent of the capacity requirement.[198]

There is no indication that the different zero crossing points were based on differences in the marginal value of capacity. Instead of actually attempting to determine the quantity at which the marginal value of capacity reaches zero, the zero crossing points "were established through stakeholder negotiations to balance concerns over price volatility, market power and the relative sizes of marginal generators and owner portfolios as compared to locality size."[199] This process of stakeholder negotiation is vulnerable to decisions to achieve political compromise rather than any economic or analytical justification based on the marginal value of capacity that demand is supposed to represent.

When NYISO consultants subsequently recommended changing the zero crossing point (and therefore the slope) of the capacity market demand curve to better reflect the incremental reliability value of capacity, NYISO rejected the recommendation because a change could "introduce undue volatility and uncertainty into the market."[200] Similarly, when NYISO needed to create a zonal demand curve for the new Lower Hudson Valley zone, rather than taking an analytical approach, it chose a crossing point of 115 percent of the zonal capacity requirement, because that number was midway between two existing zero crossing points for Upstate and New York City zones.[201]

---

196. *See* ISO New England Inc., ISO New England Inc., Demand Curve Changes, FERC Docket No. ER14-1639, at 546 (Apr. 1, 2014); *New York Indep. Sys. Operator, Inc*., New York Independent System Operator, Inc., Filing of Revisions to the ISO Market Administration and Control Area Services Tariff: ICAP Demand Curve, FERC Docket No. ER03-647, at 155-56 (Mar. 21, 2003); PJM Interconnection, L.L.C., PJM Interconnection, L.L.C., Proposal for Reliability Pricing Model, FERC Docket Nos. ER05-1410-000 and EL05-148-000, at 64-69 (Aug. 31, 2005).

197. *See New York Indep. Sys. Operator, Inc*., 110 FERC ¶ 61,201 (2005).

198. *See, e.g.*, *New York Indep. Sys. Operator, Inc*., 134 FERC ¶ 61,058 (2011).

199. *See New York Indep. Sys. Operator, Inc*., 110 FERC ¶ 61,201 (2005).

200. *See New York Indep. Sys. Operator, Inc*., Proposed Tariff Revisions to Implement Revised ICAP Demand Curves and a New ICAP Demand Curve for Capability Years 2014/2015, 2015/2016 and 2016/2017 and Request for Partial Phase-In And for any Necessary Tariff Waivers, FERC Docket No. ER14-500, at 41 (Nov. 27, 2013)

201. *See* New York Independent System Operator, Inc., New York Independent System Operator, Inc., Proposed Tariff Revisions to Implement Revised ICAP Demand Curves and a New ICAP Demand Curve for Capability Years 2014/2015, 2015/2016 and 2016/2017 and Request for Partial Phase-In and for any Necessary Tariff Waivers, FERC Docket No. ER14-500, at 40 (Nov. 27, 2013).

Since the Northeast RTOs first instituted their sloped demand curves, PJM and ISO New England have changed the shape of their curves. PJM made slight alterations, whereas ISO New England made more significant changes.[202]

PJM changed its demand curve in 2014, extending the horizontal portion of the curve out closer to the capacity requirement and moving from a slightly concave shape to a slightly convex shape.[203] PJM argued that extending the horizontal portion of the curve was necessary to increase reliability. In proposing the change to a convex shape, PJM correctly noted that a convex curve was more consistent with the incremental value of capacity, which should decrease at the margin as quantity increases. Some capacity market buyers opposed the changes, based on their practical consequences—namely that shifting the curve would increase capacity costs unnecessarily and that a convex shape would increase price volatility and the ability to exercise market power in the region of the curve below the capacity requirement.[204] FERC approved a new demand curve, which included a change to a convex shape, on the ground that it would increase reliability at reasonable cost.[205]

ISO New England introduced greater changes to its demand curve for 2019. Acknowledging that its then-existing linear demand curve was "not a function of any specific design principle," ISO New England derived the shape of its new convex demand curve by modeling the marginal increase in reliability from each unit of additional capacity.[206] The resulting curve exhibits clear convexity that reflects the diminishing marginal reliability impact of adding capacity. Once ISO New England determined the shape of the curve through this modeling, it positioned the curve on the price-quantity axes so that the curve intersected the reference point at the capacity requirement quantity and Net CONE price.[207] Due to the convex shape, demand prices were lower at most quantities on the curve, as compared with the prior linear curve.[208] ISO New England retained its existing price cap for capacity quantities well below the capacity requirement. FERC approved ISO New England's new demand curve on the ground that it would meet reliability objectives more cost effectively than its previous curve.[209]

Thus, FERC approved both PJM's and ISO New England's proposed demand curve changes, even though PJM had no apparent underlying theoretical rationale for its proposal whereas ISO New England supported its new curve with a fairly elaborate theory. FERC has not made any attempt in subsequent

---

202. NYISO has maintained the same basic shape for its curve, although early in its second-generation market it lowered the horizontal top of its curve to 1.5 times Net CONE (from twice Net CONE). *See New York Independent System Operator, Inc.*, 110 FERC ¶ 61,201 (2005).

203. *See PJM Interconnection, L.L.C.*, 149 FERC ¶ 61,183 (2014). The then-existing curve's horizontal segment ended at 3 percent below the capacity requirement. The new curve's horizontal segment extended to 1 percent below the capacity requirement. *PJM Interconnection, L.L.C.*, 149 FERC ¶ 61,183 (2014).

204. *See* James F. Wilson, PJM Interconnection, L.L.C., Affidavit in Support of the Protest of PJM Load Group, FERC Docket No. ER14-2940, at 35-39 (2014).

205. *See PJM Interconnection, L.L.C.*, 149 FERC ¶ 61,183 (2014).

206. *See* ISO New England Inc., *supra* note 105, at 2, 6.

207. *See id*. at 7.

208. *See id*. at 710.

209. *See ISO New England Inc*., 155 FERC ¶ 61,319 P 21 (2016).

years to resolve this discrepancy. Despite approving ISO New England's curve, and therefore implicitly its underlying theory as well, FERC has not indicated to PJM or NYISO that it expects these RTOs to develop an analytical basis for their capacity market demand curves.[210]

Although more rooted in economic principles than the linear second-generation capacity market demand curves, the new convex curves in PJM and ISO New England constitute only a modest improvement over prior curves. PJM's curve is still almost linear and more important, like previous curves, it has a shape that is determined through judgments that are almost entirely opaque and justified only by modeling results comparing the curve to a few other equally arbitrarily chosen options. ISO New England's new curve represents a significant improvement in that it is derived from modeling based on economic theory—rather than just tested with modeling to ascertain the acceptability of its outcomes—but both the modeling itself and the process of translating the model results into a demand curve involve administrative judgments.

In particular, the methodology requires ISO New England to translate the model results, which plot reliability as a function of capacity, to a demand curve that plots the value (price) of capacity against the quantity of capacity. ISO New England readily admits that its translation methodology is not based on an assessment of the marginal value of capacity.[211] The values embedded in the curve cast doubt on how well it reflects the actual value of capacity. ISO New England economists report that the curve implies a value of lost load of $216,048/MWh.[212] This estimate exceeds other estimations of value of lost load, often by more than an order of magnitude.[213] Although the shape of the ISO New England demand curve has an analytical rationale, the extremely high value of lost load implied by ISO New England's curve calls into question the economic validity of the methodology used to position the curve on the price-quantity axes.

210. This is not to say that the different RTOs must adopt the same analytical basis for their demand curves. Federal Power Act section 205 gives FERC considerable leeway, in judging whether a market rule is just and reasonable, to allow different approaches in different circumstances. *See* Am. Pub. Power Ass'n v. Fed. Power Comm'n, 522 F.2d 142, 146 (D.C. Cir. 1975). But here FERC is not merely allowing different analytical approaches, it is allowing PJM and NYISO to proceed with demand curves that have no apparent underlying economic rationale at all other than their overall downward slope.

211. *See id*. at 702.

212. *See* Feng Zhao, Tongxin Zheng & Eugene Litvinov, *Constructing Demand Curves in Forward Capacity Market*, 33 IEEE TRANSACTIONS ON POWER SYSTEMS 525, 533 (2018).

213. *See id.* at 530 (2018). A value of over $200,000/MWh for Value of Lost Load is much higher than most estimates. For example, in a report prepared for ERCOT, the Brattle Group assumed an average VOLL of $9000/MWh, while noting that values can differ considerably across categories of consumers. *See* SAMUEL NEWELL ET AL., ERCOT INVESTMENT INCENTIVES AND RESOURCE ADEQUACY 6, 77 (Brattle Group, June 1, 2012). Another report prepared for ERCOT around the same time stated that the average Value of Lost Load for industrialized countries ranges from $9000 to $45,000. *See* JULIA FRAYER, SHEILA KEANE & JIMMY NG, ESTIMATING THE VALUE OF LOST LOAD 9 (London Economics International June 17, 2013). Similarly, the ISO New England External Market Monitor has, in a different context, referred to $30,000/MWh as "a relatively high value of lost load." DAVID B. PATTON ET AL., 2018 ASSESSMENT OF THE ISO NEW ENGLAND ELECTRICITY MARKETS at ix (Potomac Economics, Ltd. June 2019).

Thus, more than fifteen years after the creation of second-generation capacity markets with downward-sloping demand curves, the origins and basis of the shapes of the curves remain a mystery. The process for creating the shapes is apparently driven more by stakeholder politics and concerns about price volatility than by any sense of whether the curves accurately reflect the value of capacity.

3. Differences in Demand Curve Shapes Affect Market Outcomes.

Differences in demand curve shapes have significant consequences for market outcomes. Figure 10 reports the results of our modeling to illustrate how the different curves used in the Northeast RTOs affect the market-clearing price and quantity. To focus on the shape of the curve, we normalized each curve to the PJM forecast requirement and Net CONE for delivery year 2018-2019. At all quantities greater than the capacity requirement, the ISO New England curve has lower prices than either upstate New York or New York City. The ISO New England curve also yields lower prices than the PJM curve, until quantity reaches about 107 percent of the capacity requirement. The PJM curve results in lower prices than either of the New York state demand curves when the quantity is greater than about 102 percent of the capacity requirement. As long as the quantity is greater than the capacity requirement, New York City prices are higher than upstate New York prices.

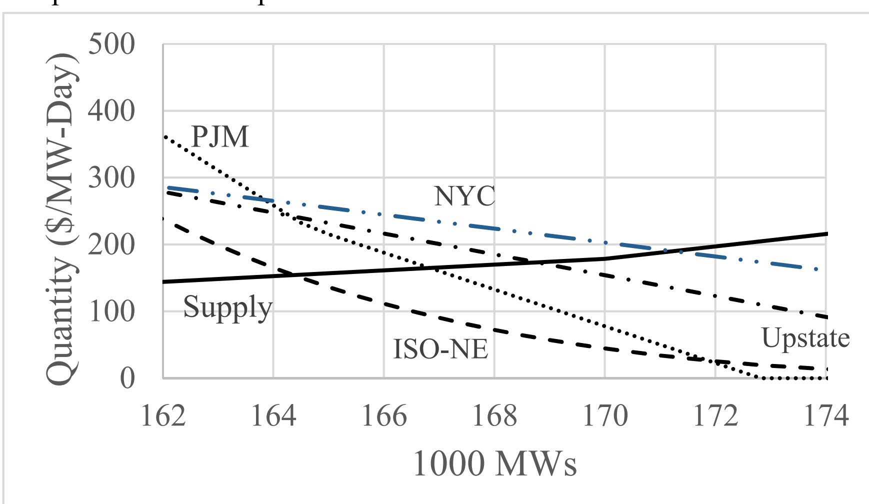

Figure 10: Market Results with Different Demand Curves

Table 1 reports the results of the modeling, employing the supply curve for PJM used above. Using the ISO New England curve results in the lowest price, slightly over $154/MW-day. The New York City curve results in the highest price, $190. Capacity revenues for one year under the New York City curve are $2.6 billion more than under the ISO New England demand curve.

Demand Shape	ISO-NE	PJM	Upstate New York	New York City
Quantity (1000 MW)	164.37	166.83	168.77	171.239
Price ($/MW-day)	$154.29	$164.88	$173.22	$190.02
Revenue ($ billion)	$9.26	$10.04	$10.67	$11.88

Table 1: Market Outcomes with Different Demand Shapes

Each RTO has developed its own demand curve, through its own administrative process, with at most limited investigation of what makes an optimal shape for a capacity market demand curve. Despite the difference in the curves, which can result in large differences in capacity revenues collected, and the scarcity of underlying theoretical justifications, FERC has deemed each of these demand functions 'just and reasonable.' There appears to be little consistency in FERC's rulings on these matters, other than deferential support for the RTOs' proposals.

VI. CONCLUSION

Capacity market demand curves derive from three factors: the capacity requirement, the Net CONE, and the shape of the demand curve. Each factor presents serious challenges for the RTOs and FERC. PJM and ISO New England have systematically overestimated peak demand for their capacity markets, which in turn inflates their capacity requirements. PJM, NYISO, and ISO New England all have overestimated Net CONE for their markets. Finally, the shapes of the demand curves are based on potentially arbitrary political compromises among stakeholders rather than economic valuations of capacity.

Justifications of administratively determined capacity market demand are quite thin. There appears to be little or no reason to believe that capacity market demand reflects the actual value of capacity. Existing methods for forecasting peak demand, estimating Net CONE, and setting the shape of demand curves require numerous administrative judgments on which there is little guidance or analytical clarity. Yet each of these decisions can have substantial consequences. The bias toward higher quantities and prices increases revenues to generators and costs to consumers. FERC, which must review and approve RTO decisions that determine demand, has largely abandoned its role, conducting its reviews with great deference to the RTO stakeholder-based process instead of identifying a coherent standard and then evaluating proposals based on whether they comport with the standard. As a result, demand in capacity markets depends more on the vagaries of RTO stakeholder politics than on market forces or theoretically grounded design principles.

Capacity market demand thus replicates many of the pathologies of traditional utility regulation through cost-of-service ratemaking. Demand in capacity markets is created through complex administrative processes assembled through an array of opaque decisions that involve discretionary judgments. These administrative processes resemble the decision-making processes used in traditional public utility regulation and, like those processes, yield statistically biased results that are not consistent with competitive market outcomes. This is ironic, because the electricity restructuring movement that birthed capacity markets arose as a rejection of traditional utility regulation in favor of competitive markets. Scrutiny reveals that capacity market demand shares far more with cost-of-service ratemaking than one would expect from a competitive market.

These problems could be alleviated, at least in part, with fixes to the administrative processes. Reducing the time between a market auction and the relevant delivery year could improve the accuracy of forecasting. An empirical Net CONE could reduce bias in administrative estimations. A modeled demand curve shape, like ISO New England's, together with a more reasonable value of lost load could more accurately represent the marginal value of capacity in the shape of demand curves.

Although FERC has to date addressed issues of capacity market demand in individualized adjudicatory decisions, the critiques raised in this article suggest that a more systematic approach is warranted. FERC has broad discretion under Federal Power Act section 205 to proceed by either adjudication or rulemaking.[214] A rulemaking process in which various stakeholders are able to propose methodologies for setting capacity market demand, with an accompanying justification rooted in economic analysis and sound empirical footing, could significantly reduce the costly errors of the current process that lead to excessive purchases in capacity markets. Rulemakings are time and resource intensive and should not be undertaken lightly. But the stakes here—billions of dollars of excessive costs imposed on electricity customers—are sufficient to justify a considerable investment by the agency.

Yet even a comprehensive rulemaking might not suffice. The administrative apparatuses employed by the RTOs and approved by FERC exhibit a strong bias toward conservatively protecting reliability in ways that inflate capacity market prices. The process allows for stakeholders to protect their vested interests. Modifications that operate within the system are unlikely to change the system fundamentally, and it is unclear that anything less than fundamental change can bring real improvement that would make capacity market demand look more like a real market.

214. *See* Towns of Concord, Norwood & Wellesley, Mass. v. FERC, 729 F.2d 824, 830 (D.C. Cir. 1984).