

THE REGULATION AND RESTRUCTURING OF ELECTRICITY

Robert J. Michaels^{*}

I. Introduction

The twentieth century's many technological revolutions shared a common predecessor: the electrification of the planet. Electricity transformed the machine age and helped create the information age, as a once-undependable novelty became a necessity of life and commerce. Not surprisingly, almost from its inception governments have actively intervened in its production and distribution. Their activities produced a remarkable diversity of electrical systems. State-owned Electricité de France is a near-total monopolist, but known for its effective nuclear technology and its pioneering efforts in economically efficient pricing. Consumers in the United Kingdom can choose from competing sellers who buy or generate power in lightly regulated markets. Japan and the United States are dominated by large corporate utilities with exclusive territories and regulated rates, but the U.S. also has over 2,000 small systems owned by local governments. Almost everywhere, but at varying paces, electricity is moving to the market. Government-owned transmission and generators are being privatized and regulators in nations with private ownership are introducing competition in production and distribution.

Governed by a complex mix of federal and state regulation, American electricity is in a transition being shaped by technology, economics, and politics. All three of them provide rationales for regulation. The next section of this chapter examines the attributes of electricity and its delivery system that have motivated regulation or government ownership. It continues with a brief outline of the cost-based procedures that continue as the basis of regulatory practice even as markets arise in some areas of the industry. A further understanding of its evolution and prospects next requires a look at federal and state jurisdictions, whose conflict has become acute with the arrival of markets. This discussion is followed by a look at some successes and failures of competition, and how technology and regulation have influenced them. Finally, completeness requires examination of some remaining and developing regulatory issues, including standards for competition and a possible crisis in reliability.

Professor of Economics, California State University, Fullerton. rmichaels@fullerton.edu The author has served as consultant for utilities, independent power producers, power marketers, and others in the industry. The views expressed in this paper are solely those of the author.

II. Technology and Industrial Organization

A. Production

A generator transforms the energy of combustion or falling water into electrical current. Burning fuel or nuclear radiation in a boiler creates high-pressure steam that turns a turbine, whose electromagnetic oscillations induce an alternating current at 60 hertz (cycles per second) into any connected lines. That power (measured in megawatts, MW) can then be applied to resistive loads (light bulbs) or inductive ones (electric motors, which “engineered” it).¹ Coal-fired generators produced 50.1 percent of America’s power in 2002. Most of the remainder came from natural gas (17.9 percent), oil (2.5), nuclear (20.2), and hydroelectric (6.6) sources. Windmills, solar collectors, biomass conversion, geothermal and other sources usually called “renewables” produced the remaining 2 percent. (U.S. EIA 2003a, 6) Electricity is the economy’s most capital-intensive major industry. Between the 1920s and the 1970s the minimum efficient scale of new fossil-fueled generators rose steadily, a trend which has since reversed itself with important consequences for markets. A new 600 MW coal-fired generator, inclusive of pollution controls, currently costs approximately \$1.7 million per MW of capacity. A new 250 MW combined-cycle gas-fired plant costs \$0.7 million per MW, but has higher operating costs at current gas and coal prices. (U.S. DOE 2001) A new coal-fired unit requires an average of seven years from proposal to operation, while a gas-fired one takes approximately three. Coal and nuclear generators are usually “base-loaded” to operate at all times, while gas and hydroelectric plants operate during hours when demand is higher.

Electricity is a singular commodity, or possibly a service. Users value both its reliability and their options to change consumption without notifying a supplier. Both reliability and optionality require that the supplier have reserve generators operating at all times, ready to instantly increase or decrease output to meet changes in demand or unexpected outages of other generators or transmission lines. With minor exceptions, economically relevant amounts of electricity cannot be stored. Over 24 hours a region’s late afternoon peak demand can be twice its predawn minimum demand, requiring the supplier to optimize its mix of operating generators and to have others in readiness to meet unexpected conditions. A human operator (or possibly a market) determines the roster of available generators, and computers adjust their output to match changes in production and loads. Failure to track load exactly for as little as one second will black out a large region, whether the mismatch between supply and demand is positive or negative.

Electricity’s scale economies and network properties favor supply by large vertically integrated firms, or by unintegrated firms under contracts that mimic the effects of integration. Reliability requires the simultaneous control of generation, high-voltage transmission, and low-voltage distribution, but centralized operation need not require that one entity own all of the

¹ Roughly, a megawatt of power is sufficient to serve residences inhabited by 1,000 persons, but no businesses. Statistics on customers appear below.

system's assets. The operator, however, must have ultimate and immediate control over them at all times. A generator operating independently of the system would endanger area-wide reliability by greatly complicating coordination between production and demand. Network characteristics add to the complexity of electricity's vertical relationships. A region's operator cannot direct power to flow down an individual transmission line like water or gas in a pipe. Once power is injected, it flows through the entire interconnection in accordance with Ohm's law. The inability to direct its flows brings operational complexity. Avoidance of a transmission overload may require that the operator run a high marginal cost generator while a low cost one remains idle. The affected lines need not be directly connected to the generator. In the western U.S. less than half of the power sold by an Oregon generator to a California buyer will take lines that cross the boundary between them. Some if it will ultimately reach California by flowing over lines as far away as Utah (at the speed of light). This externality affects the ability of producers and consumers in Utah to construct their own economical power supplies. For efficiency, both power and transmission must carry prices that reflect their scarcities. The optimization of production with reliability constraints and transmission externalities requires complex modifications to a simple marginal cost rule for generator operation. As discussed later, economists are also attacking the harder problem of pricing transmission congestion in the presence of these externalities.

The vertical chain of production and delivery is usually divided into high- and low-voltage activities. High-voltage lines move power from distant generators to areas of load. The largest of them have capacities as high as 1500 MW. (Effective capacity differs with system conditions, and simultaneous bidirectional flows on a line are possible.) Imports and exports must be coordinated with a region's generation. Scale economies in transmission are extensive. Construction costs per mile increase less than proportionately with a line's voltage. Doubling the voltage of a line more than doubles its carrying capacity, and a smaller proportion of the larger line's power is lost in transit due to resistance. (National Regulatory Research Institute 1987) The fact that a transmission line is a technological "natural monopoly" does not, however, preclude the possible allocation of its capacity by markets.

Low-voltage distribution of power to final consumers is known as retail service. Residences used 36.6 percent of delivered power in 2002, commercial establishments 32.2 percent, and industrial plants 28.1 percent.² Duplicated low-voltage lines are redundant and unattractive, so most retail distributors are utilities with monopoly franchises granted by local governments, obligated to serve all customers in their territory. The distribution arms of approximately 230 vertically integrated corporate utilities accounted for 72.5 percent of total U.S. sales revenues in 2002, 2007 governmental utilities (municipalities and special districts) for 18.4 percent, and 830 cooperatives for 9.1 percent. (American Public Power Association 2003) Governmental utilities include cities as large as Los Angeles and San Antonio, but most are too small to own generation of efficient scale and instead purchase their requirements at regulated rates from corporate utilities. Many of the others have access to regional transmission owned by others that they use to convey power from generators of their choice. Small distribution cooperatives often get supplies from larger generation-transmission cooperatives that they collectively fund. Most state regulators have little control over municipals and cooperatives,

² EIA 2003a, 7. The remainder of deliveries are for miscellaneous uses like street lighting and agricultural pumping.

whose rates are set by their own governing boards. Tax exemption of some debt and other legal protections allow municipal utilities to contribute to city government while charging rates similar to corporate utilities.

The distribution of production differs significantly from that of retail sales. Governmental utilities and cooperatives, many of them small, on balance generate less than they distribute. They purchase some of the shortfall from corporate systems and small amounts from non-utility power producers. Much of the remainder comes from federal facilities (9 percent of U.S. production, mostly hydroelectric) and is preferentially allocated to them by law at below-market prices. Their legal entitlement, however, has not guaranteed delivery, since until recently federal law also prohibited regulators from ordering transmission-owning utilities to transport it to them. These rent-seeking issues became competitive matters in 1972, when the Supreme Court decided *Otter Tail Power Co., v. U.S.*³ The Court held that failure to transmit power for a potential competitor (a new municipal utility) violated the antitrust laws, and a court could order the utility to do so. *Otter Tail's* findings shaped public policy by requiring the introduction of competitive considerations into the regulation of entities that had been generally viewed as monopolies. Competition between corporate and municipal utilities, however, remains largely in stasis, with fewer than twenty franchises changing hands in a typical year. (Michaels 1997) The principles of competitive access would matter only after the emergence of non-utility generating industry in the 1980s. Understanding that industry and related developments requires a closer look at regulation.

III. Electricity's Regulatory Environment

A. State Regulation

Numerous regulatory statutes affect electric utilities as they do most other businesses, including those on labor, environment, and safety. As the nation's largest burners of coal they are particularly affected by environmental regulation and localized environmental pressure, most importantly over the siting of powerplants and transmission. Regarding rates and service policies, electricity's regulatory regime is singularly complex. It is also increasingly at odds with economic reality.

Commercial electricity began in 1882 with Thomas Edison serving 85 customers from the Pearl Street generating station in lower Manhattan. (Kufahl and Hazan 2000) By the early twentieth century his direct current technology had been eclipsed by alternating current, which could cover wider areas with more complex and reliable grids. Generators grew in size and efficiency, and longer high-voltage transmission lines were developed to move power from distant generators to consuming areas. Further growth of service required that suppliers obtain easements and rights of way from local governments, who quickly understood their prospective value and the options for extortionate ("opportunistic") behavior after the utilities had put their

³ 410 U.S. 366 (1972).

assets in place. (Priest 1993) Some governments allowed duplicative suppliers, but by the 1920s they usually awarded franchises to a single operator, which might be a corporation or a municipally-owned utility. City governments learned to regulate rates, sometimes in accordance with economic principles and sometimes in accordance with political expediency.

The first state regulatory commissions to deal with electricity began their work in 1907 in Wisconsin and New York. Institutional economist John R. Commons helped design Wisconsin's commission, which became a template for other states. By 1940 almost all states regulated electricity rates and terms of service. Theories to explain the origin and spread of regulation have proliferated. Public interest arguments have not fared well: at least one economist argues that it came earlier to states whose utilities wanted protection from competition rather than states whose consumers faced high prices. (Jarrell 1978) Others argue that utilities found state regulation a superior alternative to municipal franchising, citing the lobbying efforts of midwestern utility consolidator Samuel Insull. (Priest 1993). Definitive explanation appears impossible, but Insull's expectation was largely fulfilled: state commissions tried to apply uniform and predictable ratemaking principles, while permanently allocating large territories to single suppliers in order to minimize risks of opportunism and to facilitate the exploitation of scale economies.

Most state regulation continues to use "cost of service" principles that it began with. In a "general rate case" the utility presents regulators with data (or occasionally projections) to justify a "revenue requirement" that meets its expected expenses and capital costs, including returns it believes are adequate to attract investors. The company then asks for rates that reflect the costs of serving different customer classes. Costs to serve a class will vary with its consumption characteristics. Industrial plants with steady loads around the clock can be served almost entirely from base-loaded generation, while residential users whose demand is highest in late afternoon force the utility to invest in powerplants that it will operate only at those hours. Regulators take testimony from the utility and from "intervenors" who support or protest the proposed rates, the most common protests being that costs are allocated inequitably or inefficiently. The process has an unavoidable arbitrariness and a clear political tinge because cost allocation is a zero-sum activity with no single correct solution. The same facilities often produce joint products (off-peak and on-peak power) as well as collective goods (reliability). Regulators have the power to disallow costs prospectively or retrospectively, but the only in such rare cases as nuclear cost overruns of the 1980s have many of them done the latter.

Investors in a utility earn returns on its "rate base," the book or replacement cost of its authorized facilities net of depreciation. Most state regulators only allow a facility to enter rate base after they have determined that a need exists or that it is otherwise "used and useful." Some econometric studies have found that the input choices of a utility under a rate-of-return constraint will be biased toward excessive capital, while others have not. (Bailey 1973) Utilities theoretically have incentives to cut costs if "regulatory lag" leaves rates temporarily unaffected, but no useful estimates of its benefits exist. As fuel costs became volatile in the 1970s, regulators began to allow rate adjustments without formal proceedings. They also introduced performance-based ratemaking, promising utilities higher returns for exceeding predefined standards on such measures as outage frequencies or service call response times. Regulation makes principal/agent relationships between shareholders and management particularly intractable. The more a utility spends the more it makes (for a given authorized return), and managers of larger firms generally earn more. Utility shareholders may not monitor their managements as closely as

they would if regulation did not limit profits, and the Public Utility Holding Company Act makes takeovers of utilities by outsiders very unlikely. Share ownership in utilities is more dispersed than in most other industries, and institutional investors take relatively smaller positions in them.⁴

Service obligations are the other side of exclusive territories and captive ratepayers. Most states nominally require that utilities sell power at regulated rates to all retail customers in their territories who want service. In practice, during the first half of the twentieth century regulators often excused utilities from serving remote customers if doing so would significantly raise the bills paid by less remote ones. The refusal to serve some inconvenient users encouraged some of those users to form municipal utilities and rural cooperatives. Service obligations also require that the utility arrange for resources ahead of expected demand growth, with differing procedures for regulatory pre-approval among the states. Some state regulators also enforce environmental standards, such as California's legislated requirements on renewable resources and demand management activities. State regulators also control the siting of transmission, a legal holdover from the days when utilities were small and self-contained.

B. Federal Regulation

The Federal Power Act of 1935 established the Federal Power Commission (FPC), which in 1977 became the Federal Energy Regulatory Commission (FERC). Its original duties of hydroelectric licensing and the allocation of power from federal dams are minor fractions of today's workload. Most of the commission's electrical activity (it also regulates gas and oil pipelines) has become the setting of rates for "wholesale" transactions. Those are legally defined as sales of any size intended for ultimate resale to retail customers, whose rates remain under state regulation. Wholesale transactions include those for power and transmission services between utilities (both corporate and governmental), and those with non-utility generators and marketers, discussed below. Roughly 30 percent of all delivered electricity costs are federally jurisdictional. Under the "Filed Rate Doctrine," state regulators must allow utilities to pass federally approved charges on to retail customers.

The law nominally requires that wholesale power prices be cost-justified, but FERC believes that wholesale competition is now sufficiently strong that it has granted almost all utilities, non-utility generators, and marketers a blanket authority to trade at negotiated prices, known as market-based rates. In transmission, FERC continues to base rates on cost. Until quite recently, the letter of the law severely limited its authority to order the transmission of power by utilities for other eligible entities. Competition has emerged less from legislation than from FERC's initiatives to define transmission access, enforce standards for it, and price it rationally. These efforts reflect FERC's history and culture, which appear to be unique among regulatory agencies. Commissioners' terms overlap administrations, and at most three of the five may be members of the President's party. Under both parties the agency has consistently favored the expansion of markets where competition is feasible, and attempted to make regulation more streamlined and

⁴ Insull encouraged dispersed shareholding because it facilitated resistance to expropriation of utility assets and interference with regulatory rate-setting.

predictable where it is not. No other regulator's web site contains a seven-word "mission statement" that does not mention regulation: "Dependable, affordable energy through sustained competitive markets." The development of electricity markets has been difficult, but not because anticompetitive interests have "captured" its federal regulators.

IV. From Regulation to Markets

A. Into the 1970s: the Golden Age

Regulation looked easy in the postwar decades. Technological progress made it possible for state regulators to satisfy normally conflicting interests without politically painful tradeoffs.⁵ Generators steadily became more fuel-efficient as electric utilities enjoyed higher rates of technical progress than any other industry. Demand grew with rising incomes and falling prices to accommodate more efficient plants, while generator fuel prices remained stable. The happy confluence allowed regulators to cut consumer bills while also satisfying shareholders. Utility managements wanted to build ever-larger plants embodying the improved technologies, and consumers had no reason to question them. The trend would terminate with nuclear generators whose power was touted as "too cheap to meter" but turned out to be very costly. Most utilities remained self-sufficient, but some areas, including New England, initiated central dispatch of all generators in their regions and centralized financial settlements among their utilities. The FPC regulated the relatively small volume of power then under its jurisdiction and oversaw the continuing but muted conflict between public and private power. The 1972 *Otter Tail* decision allowed it to order transmission service on antitrust grounds, but few orders were issued as unallocated federal power supplies diminished. Soon, however, the FPC and FERC would begin applying competitive principles more widely.

B. The 1970s: almost Everything Goes Wrong

By 1970 technical progress in generation had halted, as large new coal-fired plants failed to deliver promised economies and became unreliable. Nuclear facilities proved to be too costly an alternative. Forsaking the economical and standardized "turnkey" plants built by contractors in the 1960s, inexperienced utilities often embarked on customized projects that encountered cost overruns and new environmental opposition. Many utilities found themselves with billions invested in unfinished plants that earned them no returns under traditional ratemaking. As state regulators began to allow earnings on construction in progress, customer bills began rising.

Fuel prices compounded the utilities' problems. The 1973 embargo and rise of OPEC brought higher and less stable prices, while price controls and new environmental restrictions brought declines in domestic production. Oil fueled only a few generators (in the northeast and

⁵ See Stalon (1991) for a fuller description of this period. A former FERC Commissioner, he is responsible for the term "golden age."

Florida), but its domestic decline affected natural gas, which was often found in association with it. Price controls on gas had existed since 1955, but it only fell into shortage as oil declined. New federal allocation rules for gas gave powerplants low priority.⁶ Coal remained abundant but became a major target of a growing environmental movement, as new laws imposed high compliance costs on generators that burned it. Nuclear plants were victims of both their own costs and environmentalism, their fate sealed by the 1979 accident at Three Mile Island. As conventional generation became less feasible some state regulators began approving experiments with unorthodox technologies, as well as demand management programs that might delay construction of new generation. Programs ranged from technological research to the planting of shade trees. They required ratemaking innovations to stabilize utility revenues in the presence of declining consumption.

The chaos of the 1970s also brought markets. Plant cancellations and delays left some utilities with generation deficits while their neighbors had capacity that exceeded local demand. Fluctuating relative fuel prices also provided exchange opportunities. Utilities developed extra-high voltage (EHV) lines and control technologies that could maintain reliability over wide areas, opening up the possibility of transactions based on differences in demand. The opening of the Pacific Intertie in 1970 allowed Californians to import inexpensive hydropower from as far as Canada to meet their summer peak, while California and the southwest profitably sent power northward during the northwest's winter peak. Utilities everywhere would soon reconstitute themselves as both producers and traders.

C. The 1980s and beyond: from Trades to Markets

The Carter administration's 1977 energy policy focused on conservation to cope with what it saw as imminent resource exhaustion. New laws would phase out the "inessential" use of gas for power production and replace it with coal, conservation, and renewable resources. True to its culture, FERC creatively interpreted the Natural Gas Policy Act of 1978. In 1985 it successfully ordered the decontrol of all gas, including large quantities the law would have perpetually controlled. (Michaels 1993) The shortage ended immediately, real gas prices fell steadily through 2002, and FERC would go on to reform pipeline regulation to further encourage gas use. Decontrolled gas unexpectedly allowed the Public Utility Regulatory Policies Act of 1978 (PURPA) to revolutionize generation. PURPA's intent was to encourage the use of industrial heat to "cogenerate" power by requiring utilities to purchase it at the "avoided cost" of self-production, as determined by state regulators. Inaccurate and politicized forecasts led California to offer payments based on the assumption that 1990 oil prices would exceed \$100 per barrel, and led New York to set a very high minimum payment of 6 cents per kilowatt-hour for any cogenerated power.

Expecting long-term gas shortages, PURPA's authors believed its effects would be minimal, but as gas unexpectedly became abundant the rush to independent power was on.

⁶ Gas that did not cross state lines was exempt from controls and remained the dominant generator fuel in producing states.

Plentiful gas brought in its wake the development of advanced generators to burn it. Gas-fired plants of under 200 MW were now more efficient than coal-fired plants four times their size. By the late 1980s, entrepreneurial organizations were building larger plants (some not eligible for PURPA treatment) and selling their power at negotiated prices to utilities that wished to avoid regulatory and environmental uncertainty. Some of the producers were unregulated affiliates of utilities such as Southern California Edison's Mission Energy and others were newcomers like Enron and Dynegy. By 2002 there were 2,800 independent power companies and industrial cogenerators (often single-project LLCs), and non-utility generation was 24.8 percent of power production. (EIA 2003, 1). A flurry of new construction and divestitures brought unregulated generation to 45 percent of total capacity by mid-2004. (FERC, June 10, 2004, 5) A contract-based market for wholesale power unexpectedly came into being. At one end were agreements for interruptible power that would flow over the next hour at market prices, and at the other were commitments to deliver reliable power for years at predictable prices. Everything in between was also possible. A new industry of marketers, often utilizing their experience in gas, arose as middlemen and risk managers.

Markets grew despite FERC's lack of legal power to order transmission prior to the Energy Policy Act of 1992. Soon after that law's passage, the commission again tested its limits. In 1996 it issued Order 888, requiring all transmission owners to offer service to eligible wholesale users at predetermined rates under nondiscriminatory "open access" rules. A utility had to deal with outsiders under the same rules it followed to deliver power to its own retail customers. Beyond some residual discrimination in transmission, two major issues remained for the new century. First, transmission rates continued to be based on fictitious "contract paths" that did not acknowledge the grid-wide characteristics of actual power flows. The charge for a contract path through several systems was the sum of their rates, often a figure that foreclosed transactions which were both technically feasible and economically efficient. Increased market activity required reform of transmission pricing. Second, retail consumers of electricity were almost universally excluded from the markets, remaining under state-regulated rates as "captive" utility customers. Some industrial users were larger than municipal systems that had arranged economical supplies for themselves, and advances in metering technology might allow the aggregation of smaller users who wished to go to market. The interests of incumbent utilities ensured that retail competition would not arrive easily.

V. Restructuring Retail Markets

A. Stranded Costs and Market Design

As of 2003, 24 states and the District of Columbia had either enacted legislation or regulations allowing retail customers "direct access" to non-utility suppliers.⁷ (EIA 2003) These

⁷ Direct access is the current term of choice, used interchangeably with "retail wheeling."

states usually had high rates, while those that remain without choice often have rights to inexpensive hydroelectric power (northwest) or are dominated by relatively efficient holding company utilities (southeast). High rates embodied the costs of unexpectedly uneconomic assets such as nuclear facilities and utility power purchase contracts whose prices were now above-market. Reluctant to face retail competition, utilities claimed a right to recover the costs of these “stranded” assets from both departing and remaining customers. They testified that regulation was a metaphorical “compact” that promised modest but near-certain profits in return for service to all customers at cost-based rates. With the approval of regulators they had made good faith investments in the stranded assets, and non-recovery might be an unconstitutional taking of their property. Opponents of stranded cost recovery saw regulation not as a contract but as a political bargain with no earnings guarantees.⁸ If retail competition was coming, numerous interested parties intended to shape a post-transition environment that advantaged them. Utilities would do so by exploiting their long familiarity with politics and regulation. They also formed coalitions of convenience with their labor unions, consumer advocates whose cross-subsidies would vanish with competition, and environmentalists who would lose access to regulators.⁹ On the other side were industrial users who threatened plant closings and relocations, allied with independent power producers and marketers who wished to transact with them.

The first rulemaking on retail competition was the California Public Utilities Commission’s (CPUC) in April 1994.¹⁰ The state’s three major corporate utilities had two differing market visions. One company supported a contract regime to grant retail customers transmission access similar to wholesale users, but with surcharges for stranded cost recovery. Utilities would remain as “suppliers of last resort” for customers who chose not to go to market or who wished to return to regulated service. To avoid inefficiency and inequity, rates would require surcharge provisions for returning customers. Two other companies wanted to follow the United Kingdom’s model of a compulsory regional energy market that took bids from all generators (and offers to reduce demand from buyers) to set a time-varying price. Utilities would remain retail monopolists whose rates passed market prices through to customers, along with stranded cost and other surcharges. Users or marketers could not contract directly with individual producers, but would be able to hedge energy prices and thereby (the utilities claimed) receive all of the benefits of wholesale competition. California’s two-year debate on market design would be repeated elsewhere.

B. The States’ Experiences

⁸ Compare Sidak and Spulber (1996) and Michaels (1994).

⁹ Not all environmentalists sided with utilities. In California, the Environmental Defense Fund favored retail access on grounds that it would foster efficient pricing and that it would allow those favoring environmentally friendly technologies to buy power produced by them.

¹⁰ CPUC 1994. This docket followed a 1992 report by the Commission’s staff that laid the blame for prices 50 percent above the national average on the Commission itself. The Commission has removed the 1994 document from its web site.

California's 1996 legislation, passed unanimously, was a compromise between the two market visions with add-ons to satisfy other interests. It more closely resembled a settlement agreement than a blueprint for efficient markets. The law embodied a massive gamble for the utilities, whom it gave six years to recover their stranded costs in the difference between retail rates that the law had frozen and fluctuating market energy prices. The west appeared to have substantial excess generating capacity that would press those prices downward toward marginal cost and allow the recovery. The law authorized a Power Exchange (PX) to determine day-ahead hourly energy prices, and an Independent System Operator that would schedule the use of utility-owned transmission in a nondiscriminatory manner. The ISO would also operate a real-time balancing market and markets for "ancillary services," which included reserves of varying degrees of readiness. Users who chose direct access to individual suppliers could contract outside of the PX and ISO, but remained liable for their shares of stranded costs. Utilities, however, had to buy all their energy at the prices prevailing in those markets and resell it at frozen rates. The law in effect prohibited them from hedging these purchases and further required that they divest much of their fossil-fueled generation.

The PX and ISO opened in April 1998, and for two years prices covered little more than variable generation costs. By mid-2000, however, the gamble turned against the utilities. A poor hydroelectric season, high demand conditions, extreme weather, high natural gas prices, and possible market power by owners of divested plants all combined to drive wholesale prices above retail rates for nearly a year. The legislature refused to alter the price freeze law, and instead chose to suspend direct access, which was 20 percent of load at its peak. The state's largest utility was bankrupted, the second-largest became seriously insolvent. State government took over power purchasing in January 2001, financing it with bond issues. The state then created a second generation of stranded costs by signing fixed-price contracts of up to fifteen years, immediately before market fundamentals drove prices back down in mid-2001.¹¹ Legislation passed in 2003 has reimposed planning, reserve, and renewable resource requirements on utilities quite similar to those that existed before 1994, while conditions for the possible resumption of direct access are currently under debate.

Other states chose different market structures and avoided California's disaster. Instead of leaving stranded cost recovery to an unhedgeable gamble, Pennsylvania regulators negotiated settlements with individual utilities and set rate surcharges to recover them over ten years. Also unlike California, Pennsylvania already had extensive experience with the organized markets and independent operation of the Pennsylvania-New Jersey-Maryland Interconnection (PJM). It did not require divestiture of generation, but gave utilities the option.¹² Regulators calculated cost-based utility bill reductions ("shopping credits") that gave competitive suppliers a well-defined

¹¹ Numerous other events of interest took place over this period, including FERC's imposition of weak price caps on the short-term markets (but not contracts), a proceeding there to set refunds for alleged overcharges (not resolved as of this writing), and proceedings to analyze certain types of energy transactions that allegedly manipulated market prices.

¹² The one Pennsylvania utility that chose to divest (GPU) is also the only one currently in poor financial health. Intermediate between California and Pennsylvania, New York required divestitures but also required that utilities enter contracts with generators for a three-year transition period. This period is ending without any important market disruptions.

target price to beat, again unlike California, where they had to compete against an unpredictable PX price.¹³ In January 2004, 24.4 percent of Pittsburgh area customers (including 131,000 residences) were taking service from competitive suppliers.¹⁴ Texas, which instituted a similar “price-to-beat” system, has seen 25 percent of its major utilities’ loads go to alternative providers, including 6 percent of all residences. In some territories an average residential customer can save as much as \$166 per year by doing so.¹⁵ By comparison with the pervasive switching of telephone services, electricity’s near-term response to retail competition has been modest. By comparison with California, the relative success of Pennsylvania and Texas graphically illustrates the importance of rationally designed institutions.

VI. Other Regulatory Issues

A. FERC

1. Wholesale Trading and Market Power

Most electricity is either produced by vertically integrated utilities for their own customers or purchased under transaction-specific contracts. FERC generally grants *pro forma* approval to such contracts if the transaction occurs in a market that meets its criteria for competition. Relatively small amounts of power also trade in spot markets for delivery a day ahead or less. FERC monitors those markets with care, out of concern that producers with market power might be able to affect price by exploiting the requirement that supply equal demand every instant. The Commission sees short-term prices as indicators of efficiency and as signals for investment in facilities. In areas without ISOs (also called Regional Transmission Organizations or RTOs) markets are often thin and trades are sometimes infeasible due to “pancaking” of rates along contract paths. FERC thus views unified regional operation and area-wide transmission tariffs as important for the future of competition. Legally, however, FERC may not have the power to order utilities into RTOs. Rather than gamble on a court test, the Commission has chosen an indirect form of coercion. It now intends to deny market-based rates to utilities that are not members of RTOs. Prior to 2002 FERC granted them to any utility whose area could pass a relatively supply

¹³ Pennsylvania customers also received a rate freeze which is currently ending. The state’s official consumer guide appears at <http://www.utilitychoice.org/Choice.cfm?cid=e&lid=eced>

¹⁴ Pennsylvania PUC, *Keystone Choice*, Winter 2004, 2. Some of the state’s lower-cost utilities saw very few customer switches. http://www.puc.state.pa.us/general/publications_reports/pdf/KC_Winter_04.pdf

¹⁵ Public Utility Commission of Texas, *Scope of Competition in Electric Markets in Texas*, Jan. 2003, 86-90. http://www.puc.state.tx.us/electric/reports/scope/2003/2003scope_elec.pdf

concentration screen. In 2002 it proposed strengthening the test to one that few utilities could pass, but a failing utility would be given market-based rates if it belonged to an RTOs.

Like the antitrust agencies, FERC is often a forum for debates on market power. A utility merger or market-based rate application must meet several legal standards for commission approval, one of which is its effect on competition. Here FERC faces the same problems as the Department of Justice and the Federal Trade Commission when the law requires that they pre-screening mergers for any potential adverse effects on competition. The DO and FTC (1992) have created “Guidelines” in part based on seller concentration that help distinguish cases that are almost surely harmless from those that might be harmful. The Guidelines provide standards to define relevant product and geographic markets and summarize seller concentration by a Herfindahl-Hirschman Index (HHI), whose value is compared with thresholds of concern.¹⁶ Since the effects of a merger depend on more than just market structure, the HHI is only a tool for identifying cases that warrant further investigation. FERC uses the same basic scheme, but faces an additional interpretation problem: pervasive economies of scale may render high market shares desirable, and utilities with service obligations may have little ability to monopolistically withhold output.

2. Transmission Pricing and Investment

Spot markets price energy and allow the imputation of shadow values to transmission links. For efficiency, the marginal costs of generators at both ends of an unconstrained line must be equal. With limited capacity, generators at each end should be scheduled by marginal cost (as occurs in a market), and the difference between those costs will measure the value of incremental transmission capacity. Approximation algorithms allow RTOs to calculate these “nodal” or “locational marginal” prices at numerous production, consumption, and junction points several times an hour.¹⁷ Purchasers pay these prices to the RTOS, Unpredictable variations in transmission prices can be hedged by “transmission congestion contracts.” The holder of a TCC between two points buys the right from the RTOS to collect congestion charges between them from other users, or to pay such charges to itself.¹⁸ (Stoft 2002)

A persistent difference between prices can provide incentives to build additional transmission, with the builder getting rights to the additional congestion charges that the new capacity makes feasible. Transmission, however, is subject to great economies of scale. Constructing a line of efficient size at the efficient time may leave it uncongested for years, during

¹⁶ The HHI is the sum of squares of the market shares of the individual firms, times 100. There are extenuating circumstances, but a merger that raises a market’s HHI above 1800 begins to trigger concerns by the antitrust authorities.

¹⁷ The PJM Interconnection’s prices are calculated every five minutes for over 3,000 nodes, a number that will double as it expands to the midwest.

¹⁸ In most existing RTOs, the rights of transmission-owning utilities to capacity and congestion contracts have been “grandfathered” so that they in effect continue to charge retail customers for transmission as they have in the past.

which any rights to congestion revenue will be valueless. Similarly, an owner of a line that earns congestion revenue from other parties may have little interest in expanding its capacity. Competition to upgrade lines is in principle possible, but experience with it is thus far limited. Joint planning and collective funding under the auspices of RTOs may be necessary when a new line provides substantial benefits to all users of a grid. For cases where beneficiaries are few and easy to identify (e.g. a radial link between a new generator and the grid) FERC is currently investigating proposals for “participant funding” by the beneficiaries alone.

3. Reliability

Governments regulate electricity’s production and pricing, but standards for reliability are set by associations with little enforcement power. The nonprofit North American Electric Reliability Council (www.nerc.com) oversees ten regional councils that set standards and collect operating data for the U.S., Canada and northern Mexico. (Hirst 2004, 4) Its membership and governance includes both utilities and other industry stakeholders. The organization also monitors adequacy of generation and transmission but has no powers to order construction. Some of its operating criteria are based solely on engineering factors and at times foster economic inefficiency. Transmission Loading Relief (TLR) procedures that are invoked in overload situations often redispatch generators without regard to cost or value of load. In one 1998 episode, TLR procedures to relieve a 30 MW overload on a single line in central Wisconsin rearranged generator operations and curtailed at least 1,900 MW of wholesale transactions throughout the midwest. (Michaels and Ellig 1999) The blackout of August 14, 2003 triggered investigations into NERC’s problems in disciplining utilities whose actions degrade reliability. Some proposed legislation may give NERC and FERC some enforcement powers, but today’s basic reliability situation is the same as before the blackout.

Reliability depends on investments in new transmission, but those investments are matters of state and federal regulatory conflict, further complicated by an environmental lobby that resists almost any new lines. Power flows do not respect state lines, but the law does. The Federal Power Act became law when utilities were largely insular, and left state regulators to oversee transmission siting and construction, including eminent domain. A line located entirely within one state benefits (and sometimes harms) consumers and producers in others, producing free rider and environmental issues that negotiations often cannot resolve. Wholesale transactions and retail loads have grown steadily since the 1970s, but transmission has not. Transmission capacity as a percentage of peak load reached its highest value in 1982. By 2002 it was 70 percent of that level. (Hirst 2004, 6) New construction steadily declined between 1975 and 1999, and its upward trend since then has not yet noticeably affected capacity to load. (Hirst 2004, 7) Most new lines have been short urban connections rather than cross-country links. Major TLR emergencies increased by 600 percent between 1998 and 2002. (Hirst 2004, 8) Even after the 2003 blackout, Congress failed to pass legislation that would give FERC backstop eminent domain authority for lines meeting certain “national interest” standards.

B. The States

1. Jurisdictional Competition

Power markets and grid operations cover wider areas, but state regulators have consistently resisted federal encroachment on their jurisdictions. More than a simple rivalry between bureaucracies, it is also a confrontation between a federal agency that favors competition and state interests that are more reflective of local politics and rent seeking. The appellate courts have granted FERC some additional jurisdiction, but the likelihood that future laws will increase the federal scope does not appear high. There is in fact some likelihood that any new law will take powers away from FERC. Legislation that failed to pass in 2003 would have forced FERC to delay the introduction of new market designs, and to limit the scope of open access by giving transmission-owning utilities rights to discriminate in favor of their own retail customers ("native load") when allocating capacity on their lines.

2. Demand Sensitivity

Inelastic short-run demand is a major cause of energy price volatility. If customers do not pay the marginal cost of power when it is consumed they will inefficiently overuse it during system peaks and underuse it when demand is low. Lack of price signals also necessitates excessive generation in peaking generation that operate only a few hours per year. Until recently only large industrial customers saw their power costs at the time of consumption, and only in states whose regulators had approved real-time rate schedules would they actually pay those costs. Meters that can provide price information to smaller customers and calculate their bills as consumption varies have recently become economic and their dispersal may induce some time-shifts in consumption. Regulators, however, may also lose political support under such a pricing scheme. Interruptible rates for industrial users are a less precise and less efficient way to induce demand response. As implemented in California, the utility will give the customer one or two hours notice when system conditions are strained in exchange for a discount.¹⁹ The customer has discretion over the portions of its load that are served at interruptible rates.

Some regulators, most importantly in California and New York, have attempted to encourage non-price demand management programs by allowing utilities to earn returns on investments in them, and redesigning rates to break the link between power production and revenue. Determining the effectiveness of demand management programs is econometrically difficult, and the time pattern of findings in different jurisdictions suggests that some research may be politically influenced. More generally, however, demand management may be a policy without an economic rationale. Environmental externalities are more efficiently dealt with directly (e.g. by taxation of a powerplant's pollutants) than by attempting to change the consumption of power through by non-price means. Demand management is also rationalized by claims that government's information about costs and benefits is superior to that of consumers and investigations by individual consumers would be costly and duplicative. Hence its advocates believe that energy efficiency can be attained at lower cost by limiting consumer choices to those

¹⁹ Prior to the 2000-2001 crisis California interruptible customers usually saw at most one interruption a year. Regulators also gave them the option to maintain service through the interruption by paying ten to twenty times regular rates.

that others have determined as efficient. A usual example refers to the literature showing that consumers who reject energy-efficient appliances sometimes appear to have very high implicit discount rates since the added investment in them has a quick payback. Such seemingly irrational behavior, however, does not generally extend to their other financial choices.

3. Distributed Generation

In some states final customers have the right to choose suppliers, and large industrial users can cogenerate power for themselves and sell the remainder to utilities. Economical self-generation technologies are becoming cheaper and their efficient scale is shrinking. It is now feasible in some areas for small generators to arrange deliveries to nearby customers and meter them while not interfering with local reliability. Such “distributed generation” technologies may become serious competitive threats for utilities, and some state regulators have initiated RULEMAKING and rate cases for them. Distributed generation is not a complete minus for utilities because small powerplants close to users can eliminate the need for some investments in low-voltage facilities. It does, however, pose an important threat to the local networks that are the last important monopolies still held by vertically integrated utilities.

VII. Conclusions

Electricity is so singular an industry that it is remarkable that markets have made any headway at all into it. It is also not surprising that it has been the last and most difficult of the major deregulations. The product is a necessity of life whose production can take a substantial toll on the environment. Its technology requires exacting and centralized operation coordination of a capital-intensive process over an entire region. Once produced, power flows at the speed of light in directions that cannot be controlled. Transmission and distribution systems are neither economically nor politically feasible to duplicate. Almost from the industry’s inception regulation has been comprehensive, but with a split between state and federal responsibilities that reflects the technology and the politics of a bygone era. Further, few perceived that exposure to markets might improve performance in light of the seeming efficiency of everyday utility operations. Regulated systems dispatched generators and arranged their reserves on marginal cost principles, and their operators were watchful for opportunities to buy and sell energy. There were no analogues of the natural gas shortage that price controls had caused, or the inefficient route systems and operations of airlines before deregulation. Even mistakes such as nuclear cost overruns were viewed as lapses rather than as intrinsic to the industry.

The most common term for electricity’s current transformation is not deregulation, but restructuring. The distinction is more than jargon. Once-monolithic utilities are in a slow process of vertical reintegration, but both technology and politics foreclose any shift to fully unregulated markets. Contracts with terms that are effectively unregulated will continue to dominate wholesale markets, but short-term pricing of both energy and transmission is necessary for efficiency. The short-term markets cannot be fully unregulated because they will always require some centralized operating authority. It will exist alongside a continuing residual competition for some of the rents

that regulation once provided in greater quantity. There is no market to decide the characteristics that efficient short-term markets should have, like there are competing exchanges on which the same stock can be traded. There is not even an obvious way to determine the goods that should be traded in power markets. (How many different types of reserves should be tradeable, and what happens if there are too many or too few?) Economists have made important contributions to the design of these markets. (Wilson 2002) In some cases, however, they have neglected to acknowledge the undercurrents of rent-seeking that constrain their work, and in others they have contributed to the success of rent-seekers. Markets will only prove their value by producing more efficient investments and contracts than were available under the old regime. We will probably not know their full value for a long time.

REFERENCES

- American Public Power Association, *Number of Electric Providers and Customers*, 2003
<http://appanet.files.cms-plus.com/PDFs/owngencap2002.pdf> and
<http://appanet.files.cms-plus.com/PDFs/ppdataystate2002.pdf>
- Bailey, Elizabeth, *Economic Theory of Regulatory Constraint*. Heath Lexington, 1973.
- California, Public Utilities Commission, Order Instituting Rulemaking and Order Instituting Investigation, Docket Nos. R.94-04-031 and I.94-04-032.
- Hirst, Eric. *U.S. Transmission Capacity: Present Status and Future Prospects*, Report prepared for Edison Electric Institute, Washington, 2004.
<http://www.ehirst.com/PDF/TransmissionCapacityFinal.pdf>
- Jarrell, Gregg A., "The Demand for State Regulation of the Electric Utility Industry," *Journal of Law and Economics* 21 (Oct. 1978), 269-296.
- Kufahl, Pam, and Earl Hazan, "An Industry Takes Shape," *Transmission and Distribution World*, Dec. 1, 2000. <http://www.keepmedia.com/Register.do?olilID=225>
- Michaels, Robert J., "The New Age of Natural Gas: How the Regulators Brought Competition," *Regulation* 16 (Win. 1993), 68-79.
- Michaels, Robert J. "Unused and Useless: the Strange Economics of Stranded Investment," *Electricity Journal* 7 (Oct. 1994), 12-22.
- Michaels, Robert J. "Would Anyone Invent Public Power Today? Can Anyone Reinvent It?" *Electricity Journal* 10 (Nov. 1997), 52-58.
- Michaels, Robert J., and Jerry Ellig, "Price Spike Redux," *Public Utilities Fortnightly*, Feb. 1, 1999, 40-47.
- National Regulatory Research Institute, *Some Economic Principles for Pricing Wheeled Power*, Ohio State University, Columbus, 1987.
- Priest, George L. "The Origins of Utility Regulation and the 'Theories of Regulation' Debate," *Journal of Law and Economics* 36 (April 1993, Pt. 2), 289-324.
- Sidak, J. Gregory, and Daniel Spulber, "Deregulatory Takings and Breach of the Regulatory Contract," *New York University Law Review* 71 (Oct. 1996), 851-999.
- Stalon, Charles G., "The Significance of the FERC's Transmission Task Force Report in the Evolution of the Electric Industry," *Research in Law and Economics* 13 (1991), 105-149.

Stoft, Steven, *Power System Economics*. Wiley, 2002.

U.S., Department of Energy, *Tracking Coal-Fired Power Plants*, (2001)
<http://www.illinoisbiz.biz/coal/pdf/DOE%20New%20Coal%20Plants%2010-25-01.pdf>

U.S., Department of Justice and Federal Trade Commission, 1992 Horizontal Merger Guidelines, amended 1997. <http://www.ftc.gov/bc/docs/horizmer.htm>

U.S., Energy Information Administration, *Electric Power Annual 2002*, (2003a)
<http://www.eia.doe.gov/cneaf/electricity/epa/epa.pdf>

U.S., Energy Information Administration, "Status of State Electric Industry Resstructuring Activity as of February 2003." http://www.eia.doe.gov/cneaf/electricity/chg_str/regmap.html

U.S., FERC, Docket No. PL04-9-000, Transcript of Technical Conference, June 10, 1994, Issuance 20040610-4069. <http://www.ferc.gov/docs-filing/elibrary.asp>

Wilson, Robert, "Architecture of Power Markets," *Econometrica* 70 (July 2002), 1299-1340.