Journal of Economic Perspectives—Volume 11, Number 3—Summer 1997—Pages 119–138

Restructuring, Competition and Regulatory Reform in the U.S. Electricity Sector

Paul L. Joskow

Economical and reliable supplies of electricity make possible many of the services that we associate with modern life. From electric lights and microwave ovens to television, telephones and computers, electricity is a critical input supporting a wide range of consumption, transportation and production activities. The electricity sector is also a major manufacturing sector, accounting for about $210 billion of annual sales, about $40 billion in annual investment and 35 percent of U.S. primary energy use.

For nearly a century, the electricity sector in all countries has been thought of as a "natural" monopoly industry, where efficient production of electricity required reliance on public or private monopoly suppliers subject to government regulation of prices, entry, investment, service quality and other aspects of firm behavior. But dramatic changes are now taking place in the structure of electric power sectors around the world. The changes are designed to foster competition in the generating segment of the industry and to reform the regulation of the transmission and distribution functions, which continue to be viewed as natural monopolies. In the United States, reforms are being introduced most quickly in California and the Northeast, but many other states are moving quickly to introduce competition and reform regulation. Pilot programs that unbundle retail prices into separate generation, transmission, distribution and transition cost charges and that allow retail consumers to choose among a large number of competing generation service suppliers are already underway in New Hampshire, Massachusetts and Illinois. California, Rhode Island, Massachusetts, New Hampshire and perhaps other states are expected to give large numbers of retail consumers the opportunity to choose among competing electricity suppliers as early as 1998.

■ *Paul L. Joskow is Elizabeth and James Killian Professor of Economics, Massachusetts Institute of Technology, Cambridge, Massachusetts.*

Structural and regulatory reform of the electricity sectors in the U.S. and other countries is following the basic model previously applied to network industries such as telephones and natural gas. Potentially competitive segments (the generation of electricity) are being separated structurally or functionally from natural monopoly segments (the physical transmission and distribution of electricity). Prices for, entry to and exit from the competitive segments are being deregulated, and consumers are given the opportunity to choose among competing suppliers. Services provided by the natural monopoly segments are being unbundled from the supply of competitive services, nondiscriminatory access to "essential" network facilities mandated and prices for use of these facilities determined by new regulation mechanisms that are designed to control costs better than traditional rate-of-return regulation procedures.

While the basic model for structural and regulatory reform in electricity is fairly straightforward, the details of the institutional reforms that are necessary to improve on the performance of the present U.S. system are complex. Moreover, much of the pressure for reform in the United States reflects rent-seeking behavior by various interest groups pursuing private agendas that may not always be consistent with efficiency goals. At the same time, there are good public interest reasons to believe that structural and regulatory reforms that foster competition can lead to real cost savings in the long run *if* appropriate supporting institutional arrangements are put in place. Because of the critical role that economical and reliable supplies of electricity play in our economy, there is a profound public interest in ensuring that these reforms improve rather than degrade the performance of the electricity sector over the long run.

This paper discusses the electricity sector reforms that are taking place in the United States. The first half of the paper discusses the physical attributes of electric power networks: the industrial and regulatory structure that emerged during the last century to govern resource allocation in the sector, the performance attributes of the sector and the sources of the pressures for reform. The second half of the paper discusses a number of issues that must be confronted to create efficient competitive markets for generation services and to reform the regulation of the residual monopoly segments to support the efficient evolution of the competitive segments that must rely on them.

The Organization of the Electric Power Sector in the United States

The basic structure of the U.S. electricity sector and its regulation is discussed in detail elsewhere (Joskow and Schmalensee, 1983, 1986; Joskow, 1989, 1996). I provide only a very brief description here as background for the discussion of the reforms presently underway. Electricity is supplied to consumers in the United States by investor-owned or publicly owned (municipal, state and federal) utilities that have de facto exclusive franchises to sell electricity to retail customers in specific geographic areas. The discussion here focuses on the investor-owned segment of the industry, which accounts for over 75 percent of U.S. retail electricity sales and is the major focus of the reforms taking place in the United States.

Today, retail consumers must buy their electricity from the regulated monopoly supplier that has the legal right to distribute electricity at their locations. These franchised monopolies have a legal obligation to supply and to plan for the needs of all retail customers within their franchise areas and to make electricity available at prices approved by state regulatory commissions. Most utilities have historically met their obligations to supply by owning and operating all of the facilities required to supply a complete "bundled" electricity product to retail customers. That is, the typical utility is vertically integrated into four primary electricity supply functions: generation, transmission, distribution and retailing.

The *generation* of electricity involves the creation of electric energy using falling water, internal combustion engines, steam turbines powered with steam produced with fossil fuels, nuclear fuel and various renewable fuels, wind driven turbines and photovoltaic technologies. The *distribution* of electricity to residences and businesses at relatively low voltages relies on wires and transformers along and under streets and other rights of way. The distribution function typically involves both the provision of the services of the distribution "wires" to consumers as well as a set of *retailing* functions, including making arrangements for supplies of power from generators, metering, billing and various demand management services. These retailing functions have typically been viewed as an integral component of the distribution function. The *transmission* of electricity involves the use of wires, transformers and substation facilities to effect the high voltage "transportation" of electricity between generating sites and distribution centers, which includes the interconnection and integration of dispersed generating facilities into a stable synchronized AC (alternating current) network, the scheduling and dispatching of generating facilities that are connected to the transmission network to balance the demand and supplies of electricity in real time, and the management of equipment failures, network constraints and relations with other interconnected electricity networks.

Several key attributes of the supply and demand for electricity have important implications for whether and how competition can be introduced. The demand for electricity varies widely from hour to hour during an individual day and from day to day over the year, and electricity cannot be stored or inventoried economically by consumers or distributors. As a result, the generation and consumption of electrical energy must be balanced continuously to maintain the frequency, voltage and stability of an electric power network and to avoid sudden losses of power.

Although generation and transmission are typically discussed as separate segments of the vertical supply chain, there are important operating and investment complementarities between them that explain the evolution of an industry structure based on vertical integration of generation and transmission (Joskow and Schmalensee, 1983; Joskow, 1996a). The transmission system is not simply a transportation network that moves power from individual generating stations to demand centers, but a complex "coordination" system that integrates a large number of generating facilities dispersed over wide geographic areas to provide a reliable flow of electricity to dispersed demand nodes while adhering to tight physical requirements to maintain network frequency, voltage and stability.

Electric power networks are not switched networks like railroad or telephone networks, where a supplier makes a physical delivery of a product at point *A*, and it is then physically transported to a specific customer at point *B*. A free flowing AC network is an integrated physical machine that follows the laws of physics (Kirchoff's laws). When a generator turns on and off, it affects system conditions throughout the interconnected network. A failure of a major piece of equipment in one part of the network can affect the stability of the entire system. Efficient and effective remedial responses to equipment failures can involve coordinated reactions of multiple generators located far from the site of the failure. Finally, there is generally no meaningful direct physical relation between the electric power produced by a specific generator connected to the network and a specific customer taking energy from the network. This creates significant challenges for accurately measuring and settling consumer and generator financial obligations in a competitive electricity market.

The primary economic rationale for vertical integration between generation and transmission is that it internalizes within an organization the operating and investment complementarities between these supply functions, with their associated potential public goods and externality problems. Vertical integration also responds to challenges that decentralized market mechanisms face; for example, coordinating the efficient operation of generation and transmission capacity in real time in response to continuously changing demand and supply conditions, the need to balance the supply of generation and electricity consumption continuously at every point on the network, and the accurate measurement and billing of consumers and suppliers for injections to and withdrawals from the network. However, vertical integration between the network functions that have natural monopoly characteristics and the generation function effectively turns the supply of generating service into a monopoly as well, even if, as is the case in the United States, there are numerous generating plants connected to the network and limited economies of scale associated with generation per se in isolation from the coordination functions performed by the network (Joskow and Schmalensee, 1983). In turn, this leads to the extension of government regulation and any inefficiencies it entails, to the prices, costs and investment decisions related to the generation segment, which is potentially competitive.

It is sometimes argued that one reason that creating a separate competitive generation sector now makes sense is that the generation of electricity is no longer a natural monopoly as a consequence of technological change. This view is incorrect. Generation per se has not really been a strong natural monopoly requiring very large generating companies spanning a large fraction of regional wholesale power markets for many years (Joskow and Schmalensee, 1983). Just look at the United States, where hundreds of utilities own and operate generating plants, with little evidence that huge generating companies are necessary to exploit available economies of scale. Cheap natural gas and the new aero-derivative combined-cycle generating technology (CCGT) have certainly significantly reduced the minimum efficient scale of new generating facilities, reduced planning and construction lead times and facilitated siting as well. These developments have increased the feasibility of creating competitive generation markets quickly, but have not fundamentally transformed a sector with natural monopoly characteristics to one where these characteristics are completely absent. Rather, it is the attributes of the transmission network and its ability to aggregate and facilitate the efficient operation of generating facilities dispersed over wide geographic areas, over time frames from seconds to decades, that has played the most important role in defining the vertical and horizontal structure of this industry.

While the investor-owned utilities in the United States are typically vertically integrated into generation, transmission and distribution, there are over 100 of them serving specific geographic areas. They vary widely in size. In addition, thousands of relatively small unintegrated public and cooperative distribution entities exist that buy power from unaffiliated generating and transmission entities. The decentralized industry structure that has emerged in the United States is not ideally matched to the physical attributes of the electric power networks that have evolved over time. From a physical perspective, the U.S. sector (combined with portions of Canada and northern Mexico) is composed of three large synchronized AC networks, the Eastern Interconnection, the Western Interconnection and the Texas Interconnection. These three networks are not each under the physical control of a single network operator. Instead, there are roughly 150 separate "control areas" superimposed on the three networks where individual vertically integrated utilities or groups of utilities operating through power pooling arrangements are responsible for generator dispatch, network operations and maintaining reliability on specific portions of the networks.

To harmonize and rationalize the dispersed ownership and control of facilities that are physically interconnected and whose operations have impacts on facilities in remote control areas, the U.S. industry has developed a complex set of operating protocols—bilateral and multilateral agreements designed to maintain reliability, to facilitate coordinated operations, to facilitate trades of power between control areas and to minimize free-riding problems. These operating protocols, developed by a hierarchy of cooperative "technical" organizations, are essential for the reliable and efficient operation of synchronized networks when there are many hands on the wheel.

The decentralized structure of the U.S. electricity sector has also led to the development of competitive wholesale markets through which utilities buy and sell electricity among one another to reduce the costs of supplying their franchise customers. Wholesale power transactions and supporting transmission or "wheeling" arrangements are regulated by the Federal Energy Regulatory Commission (FERC). Wholesale trade expanded rapidly in the 1970s, initially in response to large differences in the short-run marginal cost of hydroelectric coal, oil and natural gas generating units, as well as variations in demand and capacity availability among utilities in the same region.

The Public Utility Regulatory Policy Act of 1978 (PURPA), which required utilities to buy power from cogenerators and small power producers using renewable fuels (Joskow, 1989, 1996), significantly spurred long-term contracts between vertically integrated utilities and certain types of independent generating companies. The Energy Policy Act of 1992 and state programs requiring utilities to meet additional generation needs through competitive bidding further expanded opportunities for independent power producers to sell electricity to utilities for resale. The independent producers account for only about 8 percent of U.S. generating capacity, but essentially all of that has been created since 1980, and in the last few years, independent producers have accounted for more than 50 percent of annual generating capacity additions.

Industry Performance and Pressures For Reform

To understand why electricity sector reform is taking place now, it is natural to look first at the performance of the industry as a stimulus for reform. However, the electric power sector in the United States has performed fairly well. In particular, it supplies electricity with high levels of reliability; investment in new capacity has been readily financed to keep up with (or often exceed) demand growth; system losses (both physical and those due to theft of service) are low; and electricity is available virtually universally. This contrasts sharply with the performance of the electricity sectors in many other countries. The average price of electricity in the United States today is about 6.9 cents/kWh. The average price charged to residential customers is about 8.4 cents/Kwh, and the price to industrial customers is about 4.7 cents/kWh. The difference between the residential and industrial prices largely reflects differences in load factor and the voltage level at which electricity is supplied. These prices are at the low end of the range of prices for OECD countries and have been falling in real terms for the last decade.

Despite these generally favorable performance attributes, there are a variety of apparent inefficiencies that are targets of opportunity for structural and regulatory reforms.

In the short run, the current system does a good job efficiently dispatching generating plants, making cost-reducing energy trades between generating utilities, maintaining network reliability, and dealing with congestion and emergencies. Restructuring for competition and regulatory reform is unlikely to lead to significant short-run cost savings.

However, medium-run efficiency gains may be associated with improving the operating performance of the existing stock of generating facilities and increasing the productivity of labor operating these facilities. The operating performance of both fossil and nuclear units varies widely even after controlling for age, size and fuel attributes, and some utilities have performance that lags behind industry norms

(Joskow and Schmalensee, 1987). In addition, regulatory cost recovery rules may encourage utilities to continue to operate generating plants even though it would be economical to close them. Other countries that have restructured their electricity sectors have experienced significant improvements in labor productivity. For example, the number of workers that have been shed by the electricity sector in England and Wales since the 1990 privatization and restructuring is quite impressive (Newbery and Pollitt, 1996). The potential gains from improvements in labor productivity and wage concessions must be kept in perspective, however. In the United States, wages and benefits account for only about 12 percent of the total cost of supplying electricity, and labor productivity is higher in the U.S. electricity sector than it was prior to reform in the countries that subsequently went through a restructuring process. Overall, my sense is that the opportunities for costs savings in the United States in the medium run are significant, but not enormous.

The most important opportunities for cost savings are associated with long-run investments in generating capacity. The cost of building reasonably comparable generating facilities varies significantly. These variations have been revealed most starkly in the context of nuclear generating facilities (Lester and McCabe, 1993), but appear as well in large fossil-fuel generating plants (Joskow and Rose, 1985). Significant variations also exist in the speed with which utilities have adopted new generating technologies (Rose and Joskow, 1990). Indeed, it is evident that PURPA's requirement that utilities contract with certain independent power suppliers, combined with competitive generation procurement programs in the late 1980s, helped to stimulate the technological innovation in combined-cycle generating technology (CCGT) using natural gas as a fuel. Finally, traditional regulatory pricing principles, based on the prudent investment standard and recovery of investment costs, implicitly allocates most of the market risks associated with investments in generating capacity to consumers rather than producers. Once regulators approve the construction costs of a generating plant or the terms of an energy supply contract, these costs (amortized in the case of capital investments) continue to be included in regulated prices over the life of the investment or contract, independent of whether the market values of these commitments rise or fall over time as energy prices, technology, and supply and demand conditions change. Accordingly, regulated prices reflect current market values of electricity only by accident.

While potential performance improvements in these and other dimensions represent plausible "public interest" motivations for structural and regulatory reform in the United States, they are not the primary stimulus to reform today. As White (1997) has explored in detail, the primary stimulus for reform of the U.S. electricity sector is the gap that exists in some parts of the United States between the implicit price of generation services embedded in regulated bundled electricity prices and the "unbundled" price of generation services that would be available in the wholesale market if consumers could buy it directly, paying the local utility only for transmission and distribution costs.

As noted earlier, the average price of electricity at retail is about 6.9 cents/kWh. Of that price, about 3.8 cents/kWh is associated with the costs of

Table 1
Average Electricity Prices for Selected States, 1995
(cents/kWh)

State	*All Sectors*	*Residential*	*Industrial*
Massachusetts	10.3	11.4	8.6
Connecticut	10.5	12.0	8.1
New York	10.8	14.0	5.6
Virginia	6.3	7.9	4.0
Florida	7.1	7.8	5.2
Indiana	5.3	6.8	3.9
Wisconsin	5.4	7.2	3.8
Illinois	7.7	10.4	5.3
Texas	6.1	7.7	4.0
Arizona	6.2	9.1	5.3
Oregon	4.7	5.5	3.5
California	9.9	11.6	7.5
U.S. Average	6.9	8.4	4.7

Source: U.S. Energy Information Administration, *Electric Power Annual 1995,* Volume 1, p. 39.

generating electricity and the rest with transmission, distribution, and unallocated general and administrative expenses. However, the averages reflect wide regional differences in regulated electricity prices, as shown in Table 1. In the Northeast and California, the average price of electricity is around 10 cents/kWh, while in Indiana it is about 5.5 cents/kWh, and in Oregon less than 5.0 cents/Kwh. Some of this variation in prices can be explained by regional differences in fuel costs, the mix of customers, average utilization rates and load factors, and differences in population density and construction costs. However, a large fraction of the variation in prices reflects differences in the sunk costs of generation investments and long-term purchase power contracts made during the 1970s and 1980s.

As already noted, regulated retail prices reflect the amortization of the sunk costs associated with past regulator-approved investments in generating plants (for example, nuclear power plants) and prices paid for energy under long-term purchase contracts mandated by PURPA signed many years ago, when expectations about fossil fuel prices and demand growth were very different from what eventually transpired. Thus, in much of the northeast and California, the average cost of generation services reflected in regulated retail prices is in the 6–7 cent/kWh range, reflecting historical investments in nuclear power plants and high-priced PURPA contracts that regulators required utilities to sign. In Indiana and Oregon, the average cost of generation services reflected in retail prices is 2–3 cents/kWh, reflecting low-cost coal-fired and hydroelectric generation resources, limited commitments to nuclear power and state regulatory policies that did not require utilities to sign expensive long-term PURPA power supply contracts.

More importantly for understanding the source of the interest group pressures for reform, the short-run *unregulated* price of electricity in the wholesale market today is about 2.5 cents/kWh, and the long-run marginal cost is in the 3–4 cents/kWh range, reflecting a combination of excess generating capacity, the abundant supply of cheap natural gas and the combined-cycle generating technology that can transform natural gas into electricity very efficiently. Thus, in areas like the northeast and California, there is a "price gap" of 3–4 cents/kWh between the price of generation service included in regulated retail rates and current and projected wholesale market prices in these areas, while in other parts of the country the gap is negligible or even negative. If generation services were instantly priced at current and projected market values in those areas where the price gap is positive, the net present value of the losses to utilities would be on the order of $100 billion.

Electricity sector reform efforts at the state level have been concentrated in the states where the gap is largest (White, 1997; Joskow, 1996b). They have been led by large industrial customers interested in lower electricity prices and by the independent power providers and new electricity marketers who can profit if reforms allow them to sell directly to end-use customers at prevailing wholesale market prices *and* if these customers are relieved of their responsibility to pay for generating plant investments and long-term contractual commitments their local utility made in conjunction with its historical public supply obligations. Not surprisingly, with $100 billion at stake, this in turn has led to a heated debate about the allocation of obligations for the existing sunk cost commitments between utilities, customers and independent power producers who signed high-price, long-term contracts—the so-called "stranded cost" problem (Sidak and Spulber, 1996). FERC and most state commissions that have dealt with stranded cost have allowed utilities to recover these stranded costs in the form of nonbypassable access charges in return for utility support and assistance in implementing competitive reforms quickly and, in a few cases, in return for "voluntary" generation divestiture.

Major Issues in Restructuring the U.S. Electricity Sector

Let me now turn to some of the major institutional issues that arise as the United States endeavors to implement regulatory and structural reforms aimed at creating a more competitive market for the generation of electricity, shrinking the domain of price and entry regulation and reforming the regulation of residual monopoly services. *The key technical challenge is to expand decentralized competition in the supply of generation services in a way that preserves the operating and investment efficiencies that are associated with vertical and horizontal integration, while mitigating the significant costs that the institution of regulated monopoly has created.*

What is the Right Model?

Two basic models for promoting competition in the electricity sector have been under discussion in the United States for the past several years. The first is the "portfolio manager model," in which the local distribution utility retains its traditional obligation to supply customers within its de facto exclusive franchise areas with bundled retail electricity service. However, in this model, the distributor relies on competitive procurement mechanisms to buy electricity from the lowest cost suppliers in competitive wholesale markets rather than building new generating facilities to serve growing electricity demand in its franchise area. The price for the electricity received by retail consumers continues to be regulated since the consumers must buy their electricity from the local monopoly distributor. But the regulation of the generation cost component of the retail price would presumably be based on market price indicia rather, as in the traditional method, than trying to track the underlying accounting costs and performance of generating plants owned by the distributor. This portfolio manager, or "wholesale competition model," was the framework envisioned by both the Public Utility Regulatory Policy Act of 1978 (PURPA) and the Energy Policy Act of 1992. It promotes both the continued growth of the independent power sector and associated competitive *wholesale* markets and retains the traditional *retail* monopoly over retail sales of electricity.

The second model is the "customer choice," or "retail wheeling" model. In this model, retail customers can access the wholesale market directly by purchasing unbundled distribution and transmission services from their local utility. Individual consumers take on the obligation to arrange for their own generation service supplies with independent competing electricity suppliers. The electricity suppliers can either be companies with physical generating assets or marketers that provide a bundled product of generation service procurement and risk management services (and no doubt will be calling us while we are having dinner). In this model, generators can sell energy in a competitive spot market, as well as arrange for longer term financial contracts with electricity supply intermediaries or directly with retail consumers. The role of local distributors is to provide "wires services" to retail customers for "access" to the power market, as well as metering and billing services. A network operator of some kind is responsible for operating (or owning and operating) the transmission network so that reliability is maintained and competition to supply energy from competing generators can proceed efficiently. The prices for these distribution and transmission services would still be subject to (better) regulation since they continue to be monopoly services.

Variations on the customer choice model have been adopted in England and Wales, Chile, Argentina, New Zealand, Norway and elsewhere, although the retail customers' freedom to choose has typically either been phased in over time or limited to large customers. The restructuring initiative that began in California in 1994 has stimulated much more interest in the customer choice model in the United States. This model is now guiding the restructuring for competition and regulatory reform initiatives in California, New England, New York, Pennsylvania, Illinois and other states and is the focus of legislation that has recently been introduced in the U.S. Congress.

The portfolio manager model involves the smallest changes in organizational arrangements and retains the largest continuing role for regulation. In this model, regulators will almost inevitably retain responsibility for supervising how utilities purchase generating capacity—especially if utilities continue to own and operate their existing generating facilities. While competitive procurement mechanisms for generation supplied by third parties could minimize regulatory supervision, there remains room for considerable regulatory intervention into decisions about the kinds of generating sources utilities will contract with and the prices that should be paid. This creates considerable opportunities for regulatory mischief driven by the kinds of interest group pressures that are partially responsible for the inefficiencies in the present system.

The customer choice model represents a much more dramatic change in utility and regulatory responsibilities and in organizational and financial arrangements. The major potential benefit of the customer choice model over the portfolio manager model is that by allowing end-use customers to manage their own electricity supply, this approach substantially reduces the ability of regulators to control the generation market, including service prices, entry to and exit from the generation segment and the forms of the contractual arrangements that support new generation investments. In theory, the customer choice model reduces the domain of regulation to the distribution and transmission segments and relies on market forces to govern the performance of the generation segment, the segment where performance has historically been the poorest and regulation-induced inefficiencies the largest.

Transmission Network Governance and Pricing Structures

All of the models for creating new competitive market structures in electricity being discussed in the United States recognize that there must be a single network operator responsible for controlling the physical operation of a control area, coordinating generator schedules, balancing demand for and supply of generation services flowing over the network in real time and coordinating with neighboring control areas. Also, the general agreement seems to be that it would be desirable to consolidate the roughly 150 control areas that now exist into a smaller number of regional control areas. However, there is much less agreement about precisely what the network operator's function should be, what information it needs to perform its tasks well, the ownership structure of the network operator and how it should be regulated.

Transmission pricing is a particularly challenging problem because of the existence of transmission constraints from time to time, complementarities between generation and transmission, and potential network externalities arising from the interrelationships between generators and demand at different locations on the network (Hunt and Shuttleworth, 1996). We must get transmission pricing right to decentralize competitive generation supply decisions efficiently over time and space on an AC network. Two "pure" approaches are being pursued for organizing the trading of energy on the network, the associated prices for transmission service and the management of network congestion and reliability standards.

The "tradable physical rights" framework involves defining physical transmission rights to inject energy at one or more points on the network for "receipt" at one or more other points on the network. In practice, it could work in this way. Engineering power flow models are used to determine the ''available transmission capacity'' (ATC) of a particular transmission system based on a variety of assumptions about system conditions and reliability. ATC is essentially the capacity a specific transmission interface has to accommodate generator schedules for 24 hours a day, 365 days a year, with high probability. The rights to use the ATC over a ''contract path'' from a set of injection points to one or more receipt points on the network are then sold to generators, distributors, retail marketers or directly to consumers, who can either use the rights themselves to buy and sell electricity on the network or trade them to third parties for their use. If the demand to use an interface rises beyond the ATC to handle all of the preferred schedules, the price for the fixed quantity of rights to use that transmission interface will rise to balance supply and demand.

The primary problem with this approach is that there is no unambiguous way to define a full set of contingent delivery and receipt property rights from one point on the network to another. The capacity to transfer power across one interface depends on demand, generation and power schedules elsewhere in the system, and thus it can vary widely with supply and demand conditions on the network. To avoid significant conflicts between rights for simultaneous use of different interfaces on the network, ATC must be defined conservatively to reflect a set of ''stressed'' system conditions. As a result, during many hours, more transmission capacity will be available for use than has been allocated to users. There will also be conflicts of rights under certain system conditions, when the capability of the network to accommodate schedules is less than the quantity of usage rights that have been allocated to use it. In addition, for this approach to work well, a set of transmission rights markets must evolve to operate in tandem with forward and real-time energy markets. The tradable physical rights approaches that are being proposed recognize these problems, but argue that the approach can be employed without significant efficiency losses since there is a relatively small number of transmission interfaces where congestion is a significant issue and that market mechanisms, monitoring and enforcement institutions can be created to assure that transmission rights markets and energy markets clear efficiently (Walton and Tabors, 1996; Tabors, 1996; Chao and Peck, 1996).

The second broad organizational framework, the ''nodal pricing'' approach, follows directly from the work of my late colleague Fred Schweppe (1988) and has been extended significantly by William Hogan (1992, 1993; Harvey, Hogan and Pope, 1996). Basically, the network operator runs a set of day-ahead and hour-ahead auction markets for energy (as well as the ancillary network support services required in any approach) and uses the bids submitted to it to derive a ''least cost'' merit order schedule of generators selected to supply energy and an associated set of market clearing prices, which take network constraints into account. On the supply side, the network operator accepts supply bids from generators offering supplies in the day-ahead and hour-ahead markets, or congestion reservation prices for generators or intermediaries seeking to schedule generators on the network but who do not want to participate directly in the auction markets run by the network operator.

On the demand side, customers articulate their willingness to pay for electricity, which includes their willingness to contract or expand their use at different times as price varies. The network operator then feeds the generator supply and customer demand bids into an optimization program that takes into account network operating constraints to determine prices and quantities for electricity to balance supply and demand continuously at each node. (A "node" is a point on the network where electricity is either supplied to the network or withdrawn from the network at a point of connection with a distribution system or large industrial consumer.) The transmission price for power physically flowing from one node to another is then the difference between the prices at each of the nodes. Ideally, these prices take all network interdependencies into account.

There has been a lot of controversy about which of these models is the best one to pursue, as well as how they might be combined effectively. Some of the controversy reflects reasonable differences of opinion about how best to create an efficient competitive electricity market that properly reflects all of the physical complexities of electric power networks. Some of the controversy also reflects rent-seeking behavior by market participants who envision financial opportunities that emerge by creating network operating and resource allocation institutions that have high transactions costs and are inefficient (Stoft, 1997).

Supporters of the tradable physical rights approach emphasize that it maximizes the freedom individual suppliers have to structure transactions and minimizes the role of the network operator, which they argue is a potential monopoly "central planner" that could abuse its authority. The network operator does not participate directly in the bulk of the electricity market transactions, does not determine market clearing prices in day-ahead and hour-ahead markets for energy and plays only a secondary role in managing network congestion *economically*. Its job is "limited" to maintaining the physical integrity of the network, enforcing transmission rights, managing conflicts between the exercise of rights to schedule generation and the actual capacity of the network to accommodate schedules, buying and selling a variety of network support and reliability services that are not self-supplied by energy traders using the network, and measuring and settling imbalances between those that have contracted to supply energy to serve consumers and their actual measured consumption.

The nodal pricing approach envisions a more active and central role for the network operator in the energy markets than does the tradable rights approach. In particular, the network operator runs day-ahead and hour-ahead markets for energy and ancillary network support services and uses the information obtained from these auctions to establish a least cost merit order generator dispatch schedule that matches demand and supply, to manage network constraints economically, and to define market clearing prices at each supply and demand node on the network consistent with network operating constraints. Basically, the network operator does what control area operators do today, except it relies on bids from competing generators as inputs into its least cost generator dispatch optimization programs. Proponents of the nodal pricing approach argue that the tradable rights approach does not solve the fundamental network externality problems that arise when congestion becomes important, will lead to inefficient allocations of scarce transmission capacity and further inefficiencies in the commitment and dispatch of generators, and will increase transactions costs for smaller consumers and generators seeking to participate in the market.

A related set of issues has arisen with regard to investments that increase capacity of the network to transfer energy from supply nodes to consumption nodes. Who should be responsible for identifying economical opportunities to expand transmission capacity, and who should pay for it? One approach would rely primarily on private parties to propose and pay for upgrades to the transmission network. The alternative vests responsibility for identifying needed transmission upgrades in the network operator and would share the costs of these facilities among those who use the system.

Transmission investment decisions do not immediately strike me as being ideally suited to relying entirely on the invisible hand. Transmission investments are lumpy, characterized by economies of scale and can have physical impacts throughout the network. The combination of imperfectly defined network property rights, economies of scale and long-lived sunk costs for transmission investments, and imperfect competition in the supply of generating services can lead to either underinvestment or overinvestment in transmission at particular points on the network if we rely entirely on market forces (Nasser, 1997). However, there is no reason why the primary initiative for transmission upgrades should not be left to private parties, especially if a reasonably good allocation of capacity rights, whether physical or financial, is created. The network operator could then determine whether proposed upgrades have adverse uncompensated effects on some users of the network, or whether there are inadequate private market incentives for investment because of scale economies or free-riding problems. In those cases the network operator could identify investment projects that the transmission owners would be obligated to build and the associate costs could be recovered from all network users. This appears to be the direction in which public policy is now moving.

Vertical Control Issues

Most of the transmission-owning utilities in the United States have ownership interests in generators that utilize these transmission facilities. In both models of restructuring the electricity industry, these firms would own and operate both competitive assets (generation) and regulated monopoly assets (transmission). As a result, a transmission network operator may have the incentive and ability to favor its own generators and disfavor competitors' generators when it makes decisions about the operation of and investments in the transmission network. Three types of "fixes" are being proposed to deal with these potential "self-dealing" problems.

One approach involves complete structural separation of generation, transmission and distribution by creating separate transmission or grid companies through vertical divestiture of generating plants. This is the structure in England and Wales, Norway and Argentina. In the United States, this step might be accompanied by horizontal integration of the pieces of the transmission network presently under separate ownership, creating a smaller number of regional "grids," each with a single control area operator. This vertical separation of ownership of generation from transmission assets was relatively easy to accomplish in other countries where restructuring took place as part of the privatization of state-owned assets. It is much harder to accomplish in the United States, since we are dealing primarily with private firms. Nevertheless, a few utilities are in the process of selling some or all of the generating capacity to resolve stranded cost recovery issues or to deal with market power problems (discussed below). More are likely to follow in the next few years.

A second approach would require functional separation of generation, transmission and distribution within existing vertically integrated firms—essentially, separating the regulated and competitive portions of the firms into separate divisions with separate cost accounting and limitations on communications between employees across divisions—combined with open access obligations and access pricing rules for use of the transmission and distribution networks approved and enforced by regulators. This approach requires the incumbent vertically integrated firms to unbundle the services they supply, separate costs attributable to different segments, post visible prices for these services and apply them to their own competitive transactions as well as to transactions involving third parties using their network. FERC's Open Access Rule (Order 888 issued in 1996) embodies this approach. The problem with this approach is that it may be difficult for regulators to enforce network access obligations, to unbundle properly the prices for competitive and monopoly services, and to specify appropriate access prices.

A third approach is a halfway house between the first two, which embodies the open access and unbundling requirements of the second approach while responding to residual self-dealing concerns without waiting for generation divestiture to be accomplished. Vertically integrated utilities in a region would turn the *operation* of their transmission systems over to an independent system operator (ISO); effectively, they would lease their system to the ISO. The ISO would be a nonprofit organization with an independent board of directors. Then, the ISO would be responsible for all network functions over a geographic expanse that more closely matches the physical characteristics of a synchronized alternating-current system. This approach is emerging as the primary transition mechanism for dealing with vertical control problems in the United States.

The reliance on ISOs raises all types of governance issues that would benefit from further analysis. Can the ownership of the transmission assets be completely separated from the use of these assets without distorting operating and investment decisions? Should the ISO be public or private? Should it be a separate company or a "cooperative" controlled by suppliers and customers? How is the ISO's board of directors selected, who does it represent, and what are the voting rules? How is the ISO's management selected, what objectives is it given, and how are good performance incentives provided to the management? What role does the ISO play in new investments in transmission facilities? How should the ISO be regulated? Policies governing the creation of ISOs are moving forward rapidly before these questions have been answered satisfactorily, and this quick movement will probably lead to performance problems in the future.

Horizontal Market Power in the Supply of Generation Services

All of the restructuring proposals hope to encourage a competitive generation sector that is largely free from price and entry regulation. Accordingly, issues associated with diagnosing and mitigating horizontal market power at the generation level are attracting a lot of attention.

Concerns about horizontal market power have been heightened among U.S. policymakers in part because of the experience gained from the restructuring in England and Wales since 1990, where various studies have shown that there is a significant horizontal market power problem at the generation level (Green and Newbery, 1992; Newbery, 1995; Wolfram, 1996a,b; Wolak and Patrick, 1996; von der Fehr and Harbord, 1993). The market power problems in England and Wales are generally attributed to the decision of the Thatcher government to divide the old state-owned generating assets into only three companies, two of which controlled all of the fossil plants. Moreover, some generators have strategic locations on the grid and, from time to time, "must run for reliability." Naturally, when the generators know that they will be called to run by the network operator to maintain network reliability (almost) regardless of what they bid, they submit high bids. Certain generating stations at strategic locations on the grid in England and Wales charged prices six times higher than those of other otherwise comparable generators before the regulator imposed a price ceiling on them (Office of Electricity Regulation, 1992).

Since the United States enters the restructuring process with a large number of companies with generating assets and with active wholesale markets for electricity, the challenge of creating a competitive generation sector should be less daunting. Moreover, the new CCGT technology is allowing generating plants to be built economically at relatively small scale and with much shorter planning and construction lead times, so entry from independent producers should play an important role in disciplining pricing behavior by incumbents.

Nevertheless, diagnosing horizontal market power associated with unregulated supplies of generation services must confront a number of significant analytical challenges (Borenstein et. al., 1995; Werden, 1996). It has long been recognized that an important factor in assessing horizontal market power at the generation level is the cost and availability of transmission capacity (Joskow and Schmalensee, 1983, ch. 12). The extent of congestion at points on a transmission network varies widely as supply and demand conditions change, both during a day and during a year, so that the relevant geographic markets change as well over time. Since electricity cannot be stored, considerable care must be taken in identifying what capacity is competitive under different supply and demand conditions. If demand is very inelastic, market power could be a potential problem even with a relatively large number of suppliers, under certain demand and supply conditions (Borenstein and Bushnell, 1997).

Creating a reasonably competitive generation market is certainly an important policy goal. However, creating a *perfectly* competitive generation market is not a realistic goal. The spatial attributes of generation markets and changing network conditions virtually assure that generation markets will never be perfectly competitive under all system conditions. But the test for deregulation of prices and entry into generation should not be whether competition is perfect—in the sense that prices must precisely equal marginal cost. If we applied such a test we would not have deregulated airlines, railroads, long distance telephone companies and many other industries. Clearly, policymakers will have to make some judgment about when there is enough competition so that any remaining costs of imperfect markets are less than the costs of continuing regulation.

If or when significant horizontal market power problems are identified, two primary mechanisms are available for mitigation. One is to continue to subject incumbent generators to some type of price regulation. This may be a necessary solution to certain types of "local" market power problems where specific generators or groups of generators "must run for reliability." The second alternative is to require horizontal divestiture of generating facilities as a way of creating additional independent competitive suppliers. This solution is now being pursued in England and Wales and in California.

Regulating Residual Monopoly Services

Most of the reform proposals being discussed in the United States anticipate that basic transmission and distribution services will continue to be provided by a monopoly that will be subject to government regulation. There is general agreement that whatever residual monopoly services are left will be subject to "incentive regulation" (Joskow and Schmalensee, 1986; Laffont and Tirole, 1993), or what has now come to be called in the regulatory arena "performance-based regulation."

The general approach being pursued is some type of price cap mechanism, which typically sets a base price, assumes that the real price will decline at some rate over time because of productivity gains and allows for adjustments in base prices over time for prespecified external factors, including input price inflation. If the regulated entity can raise productivity and cut costs more quickly than expected, it can earn additional profits. This approach has been applied to AT&T and to a growing number of local telephone companies in the United States in the last decade, and it is applied widely to all privatized network industries in England and Wales. But while this type of regulation can provide powerful incentives for cost reduction, it can also provide incentives to lower costs by allowing the quality of the services provided by the regulated firm to deteriorate. Accordingly, performance-based regulation mechanisms are now including a growing list of customer service, reliability criteria and other performance criteria.

Almost no thought has been given to the regulatory mechanisms that will govern the behavior of the network operator and grid owners, especially in the context of the independent system operator (ISO) structures that are emerging as the favored governance structure for operating the transmission network and guiding investments to expand it. This is especially surprising in light of problems that have emerged in both England and Wales and Argentina in stimulating appropriate investments in transmission capacity that properly take into account the costs of network congestion. The difficult task for these regulations is to encourage low-cost operation while also providing incentives to make investments in transmission capacity that can cost-effectively reduce congestion on the network. As the details of the ISO proposals are defined more clearly, the appropriate regulatory mechanisms to apply to the management of the ISO and transmission owners will likely emerge as an important issue.

Concluding Thoughts

Electricity restructuring and regulatory reform is likely to involve both costs and benefits. On the benefit side, a competitive generation market can significantly reduce many of the medium- and long-term inefficiencies discussed above. However, I think that it will be very difficult to replicate the efficiencies of central economic dispatch and network operations that characterize the operation of well-managed vertically integrated transmission and generation companies. There are also likely to be additional inefficiencies associated with decentralized investments in generation and transmission capacity due to complementarities between generation and transmission that will be difficult to capture fully in market mechanisms.

Because the motivations for electricity sector reform in the United States are being driven largely by distributional considerations—in particular efforts to reallocate responsibilities for paying for sunk investment costs and contractual commitments—a danger exists that, in the rush to implement reforms to satisfy the competing interest groups, longer run efficiency considerations will not be given adequate attention. If the restructuring of the electricity sector is done right, in a way that effectively addresses the challenges identified in this paper, the benefits of restructuring competition and regulatory reform can significantly outweigh the costs. But the jury is still out on whether policymakers have the will to implement the necessary reforms effectively.

References

Borenstein, Severin, and James Bushnell, "An Empirical Analysis of the Potential for Market Power in California's Electricity Industry," University of California Energy Institute, March

1997.

Borenstein, Severin, James Bushnell, Edward Kahn, and Steven Stoft, "Market Power in California Electricity Markets," *Utilities Policy,* 1995, *5*:3–4, 219–36.

Chao, Hung Po, and Stephen Peck, "Market Mechanisms for Electric Power Transmission," *Journal of Regulatory Economics,* 1996, *10*:1, 61–80.

Green, Richard, "The English Electricity Industry in the 1990s." In Olsen, Ole Jess, ed., *Competition in the Electricity Supply Industry.* Copenhagen: DJOF Publishing Company, 1995, pp. 107–36.

Green, Richard, and David Newbery, "Competition in the British Electricity Spot Market, *Journal of Political Economy,* 1992, *100*:5, 929–53.

Harvey, Scott, William Hogan, and Susan Pope, "Transmission Capacity Reservations Implemented Through a Spot Market With Transmission Congestion Contracts," *Electricity Journal,* 1996, *9*:9, 42–55.

Hogan, William, "Contract Networks for Electric Power Transmission," *Journal of Regulatory Economics,* 1992, *4*:3, 211–42.

Hogan, William, "Markets in Real Networks Require Reactive Prices," *Energy Journal,* 1993, *14*:3, 171–200.

Hunt, Sally, and Graham Shuttleworth, *Competition and Choice in Electricity.* West Sussex, Engl.: Wiley, 1996.

Joskow, Paul L., "Regulatory Failure, Regulatory Reform and Structural Change in the Electric Power Industry," *Brookings Papers on Economic Activity: Microeconomics,* 1989, 125–99.

Joskow, Paul L., "Introducing Competition in Network Industries: From Hierarchies to Markets in Electricity," *Industrial and Corporate Change,* 1996a, *5*:2, 341–82.

Joskow, Paul L., "Comments on Power Struggles: Explaining Deregulatory Reforms in Electricity Markets," *Brookings Papers on Economic Activity: Microeconomics,* 1996b, 251–64.

Joskow, Paul L., and Nancy Rose, "The Effects of Technological Change, Experience and Environmental Regulation on the Construction Costs of Coal-Burning Generating Units," *Rand Journal of Economics,* 1985, *16*:1, 1–27.

Joskow, Paul L., and Richard Schmalensee, *Markets for Power: An Analysis of Electric Utility Deregulation.* Cambridge: MIT Press, 1983.

Joskow, Paul L., and Richard Schmalensee, "Incentive Regulation for Electric Utilities," *Yale Journal on Regulation,* 1986, *4*:1, 1–49.

Joskow, Paul L., and Richard Schmalensee, "The Performance of Coal-Burning Electric Generating Units in the United States: 1960–1980," *Journal of Applied Econometrics,* 1987, *2*:2, 85–109.

Laffont, J. J., and Jean Tirole, *A Theory of Incentives in Procurement and Regulation.* Cambridge: MIT Press, 1993.

Lester, Richard K., and Mark J. McCabe, "The Effect of Industrial Structure on Learning by Doing in Nuclear Power Plant Operations," *Rand Journal of Economics,* 1993, *24*:3, 418–38.

Nasser, Thomas-Olivier, "Imperfect Markets for Power: Competition and Residual Regulation in the Electricity Industry," unpublished Ph.D. dissertation, MIT Department of Economics, April 1997.

Newbery, David, "Power Markets and Market Power," *Energy Journal,* 1995, *16*:3, 39–66.

Newbery, David, and Michael G. Pollitt, "The Restructuring and Privatization of the CEGB: Was It Worth It?," mimeo, Department of Applied Economics, Cambridge University, 1996.

Office of Electricity Regulation, "Report on Constrained-On Plant," Birmingham, England, 1992.

Rose, Nancy L., and Paul L. Joskow, "The Diffusion of New Technology: Evidence From the Electric Utility Industry," *Rand Journal of Economics,* 1990, *21*:3, 354–73.

Schweppe, Fred, et al., *Spot Pricing of Electricity.* Boston: Kluwer Academic Publishers, 1988.

Sidak, J. Gregory, and Daniel F. Spulber, "Deregulatory Takings and Breach of the Regulatory Contract," *New York University Law Review,* October 1996, *71*:4, 851–999.

Stoft, Steven, "What Should a Power Marketer Want?," *Electricity Journal,* June 1997, *10*:5, 34–45.

Tabors, Richard, "A Market-Based Proposal for Transmission Pricing," *Electricity Journal,* 1996, *9*:9, 61–67.

von der Fehr, Nils-Henrik Morch, and David Harbord, "Spot Market Competition in the UK Electricity Industry," *Economic Journal,* May 1993, *103,* 531–46.

Walton, Steve, and Richard Tabors, "Zonal Transmission Pricing: Methodology and Preliminary Results for the WSCC," *Electricity Journal,* 1996, *9*:9, 34–41.

Werden, Gregory, "Identifying Market Power in Electric Generation," *Public Utilities Fortnightly,* 1996, *134*:4, 16–21.

White, Mathew, "Power Struggles: Explaining Deregulatory Reforms in Electricity Markets," *Brookings Papers on Economic Activity: Microeconomics,* 1996, 201–50.

Wolak, Frank, and Robert H. Patrick, "The Impact of Rules and Market Structure on the Price Determination Process in the England and Wales Electricity Market," mimeo, Stanford University Department of Economics, 1996.

Wolfram, Catherine, "Measuring Duopoly Power in the British Electricity Spot Market," mimeo, MIT Department of Economics, September 1996a.

Wolfram, Catherine, "Strategic Bidding in a Multi-Unit Auction: An Empirical Analysis of Bids to Supply Electricity in England and Wales," mimeo, Harvard Department of Economics, July 1996b.