Electricity Restructuring: Deregulation or Reregulation?

By Severin Borenstein and James Bushnell

Restructuring in the electricity industry is spreading across the United States and around the world. Some of these initiatives are well under way—for instance, in California, Pennsylvania, Australia, Norway, and the United Kingdom—but far more are in the design and early implementation stages. There also are bills in Congress to encourage the formation of open markets for electricity. The restructuring process has achieved some significant successes—most notably keeping the lights on. But serious problems—some predictable, others not—have also arisen.

As other jurisdictions refashion their electricity markets, they seem to be incorporating little of the experience from the markets that are furthest along in the process. That is unfortunate, because there are important and quite general lessons that can be gleaned. Probably the two most salient lessons are that the short-run benefits are likely to be small or nonexistent, and the long-run benefits, while compellingly supported in theory, may be very difficult to document in practice.

More concretely, market power among generators is likely to be a more serious and ongoing concern than has been anticipated by most observers. As a result, the roles of transmission capacity and demand-side elasticity are likely to be even more important than previously suggested. In general, the nonstorability of electricity, combined with very little demand elasticity and the need for real-time supply/demand balancing to keep the grid stable, has made restructuring of electricity markets a much greater challenge than was inferred from experience with natural gas, airlines, trucking, telecommunications, and a host of other industries.

Severin Borenstein is director of the University of California Energy Institute and professor of business at the University of California Berkeley's Haas School of Business. **James Bushnell** is director of research at the University of California Energy Institute.

THE ORIGIN OF ELECTRICITY INDUSTRY RESTRUCTURING

Analysis of the electricity industry begins with the recognition that there are three rather distinct components of it: generation, transmission, and distribution. Once electricity is generated, whether by burning fossil fuels, harnessing wind, solar, or hydro energy, or through nuclear fission, it is sent through high-voltage, high-capacity transmission lines to the local regions in which the electricity will be consumed. When the electricity arrives in the region in which it is to be consumed, it is transformed to a lower voltage and sent through local distribution wires to end-use consumers.

In the United States, all three of these vertically related sectors have typically been tied together within a utility, which has been either investor-owned and state-regulated or owned by the local municipality. For many years, each

sector was thought of as a natural monopoly. In the transmission and distribution sectors, effective competition would require that rival firms duplicate one another's wire networks, which would be inefficient. If wires owned by different companies were allowed to interconnect to form a single network, the laws of physics demonstrate that there would be significant externalities: the flow on one line affects the capacity of other lines in the system to carry power. Generation was argued to be a natural monopoly because of the large scale of efficient generation plants and the losses that occurred with long-distance transmission, which made it more efficient to have local areas served by one or a small number of generating plants.

Few people argue that the basic economics of transmission and distribution have changed. But, over time, the optimal scale of generating plants has declined, not increased, as many thought it would in the 1960s and 1970s with the growth of nuclear power. In addition, technology improvements reduced the losses that occurred during transmission, making it more feasible for plants hundreds of miles apart to compete with one another.

Thus, in the 1980s, a movement began to increase the efficiency of the generation sector by letting independent entrepreneurs compete to supply power to the utility. This was encouraged by the federal government in 1978 with the Public Utility Regulatory Policy Act (PURPA). Under PURPA, utilities were required to buy power from "qualifying" independent power producers (mostly small generators or ones using renewable energy sources) at a price equal to the "avoided cost" of the utility. Many states, however, set very high levels for the avoided costs, levels that were certainly much higher than the actual marginal savings to the utility of not producing the power itself. The result was that many utilities signed long-term purchase contracts at very high prices. Those prices looked especially bad as the cost of natural gas fell in real terms through the 1980s and 1990s, making most other generation sources much less economic.

Over about the same period of time, accidents, unforeseen construction costs, increased safety regulation, and higher-than-anticipated upkeep and waste disposal costs changed nuclear power from the cheap, clean power source advocates had promised to expensive white elephants. Under the regulatory agreement between states and the utilities, consumers still had to pay for the plants despite the fact that they turned out to be unwise choices.

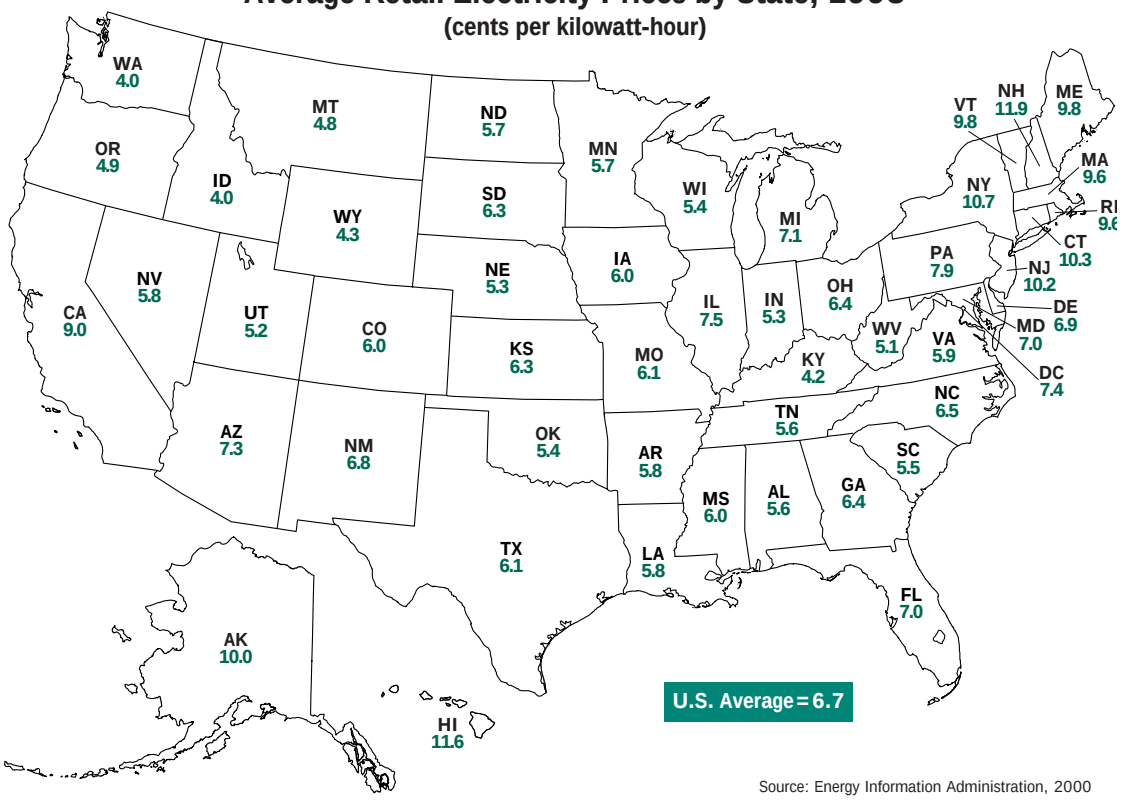
Uneconomic PURPA contracts and nuclear power investments were the primary reasons that some states found themselves in the 1990s with electricity prices that were well above the going-forward cost of building and operating new gas- or coal-fired power plants. Other states—those that had not pursued nuclear power and had been more cautious in signing long-term contracts under PURPA—retained relatively low prices. That contrast was probably the driving force behind the restructuring movement in the United States.

THE POLITICAL ECONOMY OF ELECTRICITY RESTRUCTURING

THE DISPARITY IN ELECTRICITY PRICES WITHIN THE UNITED States is evident from Figure 1. In regions such as California and the Northeast, residential rates have averaged as much as 10 cents per kilowatt-hour, while residents in some neighboring regions have been paying half that much. Some of these differences can be attributed to different natural resource endowments across regions—most importantly, hydroelectric opportunities—but a far larger share results from the need to pay for bad (at least in retrospect) investments and contract decisions made during the 1960s,

Figure 1

Average Retail Electricity Prices by State, 1998
(cents per kilowatt-hour)



Source: Energy Information Administration, 2000

1970s, and 1980s. Underlying those decisions had been the belief that natural gas prices would be far higher by the year 2000 than they have turned out to be, thanks to the deregulation of natural gas in the late 1970s and early 1980s.

Because these rate gaps stem from expenditures and commitments that are for the most part irreversible, they reflect differences in historical average costs, not marginal cost. A fairly robust, although imperfect, market for wholesale power trades between utilities, independent power producers, and a handful of large customers has operated through most of the 1990s. That wholesale market helped to equalize the marginal cost of production of utilities, subject to physical and institutional constraints on transmission access.

The success of the wholesale market also left many asking for more. While wholesale trades within a regulated environment allowed the customers of high-cost utilities to reap some marginal benefits, it did not allow them to escape the more significant burden of paying for the sunk costs of past investments, known in the industry as "stranded investment." Yet, the latter cost was the source of most of the rate disparities. The policy process has therefore largely been driven by the desire to cut costs that, being sunk, cannot be cut, only redistributed.

Economists and policy analysts have long argued that the most significant potential gains from electricity restructuring will stem from changing the way investment and consumption decisions are made.

This is not to say that there are not legitimate economic rationales behind restructuring the industry, only that such rationales do not get much play in the public discussions of these reforms. Deregulation or restructuring has the potential to produce gains in three broad sectors of the electric utility industry: operations, investment, and consumption.

The evidence indicates that regulated investor-owned utilities in the United States have done a reasonable job of efficiently generating power given their portfolio of generating assets. Labor accounts for only 12 percent of the total cost of supply electricity and is more productive in the United States than in other countries. Therefore the prospects for short-run efficiency gains on the supply side of this industry are quite modest. Indeed, there is evidence that operational efficiency has decreased under restructuring. This is in part due to significant alterations to operational practices that have been made to accommodate a generation market, and in part due to market power, which we discuss further below.

Economists and policy analysts have long argued that the most significant potential gains from electricity restructuring will stem from changing the way investment and consumption decisions are made. Since the bulk of the rate disparities in this country are due to investment decisions that turned out badly, it stands to reason that what is "broke" in this industry is the process that produced those poor investment decisions. Firms that do not have the security of a guaranteed rate of return on their investments will be more prudent in their capital expenditures and the way they manage risk.

The last, largely untapped, source of potential efficiency gains in the electricity industry lies on the consumer side. The cost of producing electricity fluctuates widely by hour or even by minute, and the market prices of deregulated power even more so. Consumers have had little opportunity to respond to these price fluctuations. They have instead paid a flat rate that reflects the average cost over months or years. Because consumers do not see prices that fluctuate with changes in the interaction of supply and demand, they cannot respond when there is a scarcity (such as on a very hot summer afternoon) by reducing their consumption (such as by turning their air conditioner from 75 to 78 degrees) and receiving savings that reflect the high value of the power they have foregone.

Note that it does not take deregulation to implement better consumer price response, although proponents of restructuring have argued that stodgy old regulated utilities have little incentive to pursue such initiatives. Thus far in the United States, real-time consumer price response has not developed either in regions that have restructured or in those that have remained under traditional regulation. However, as we discuss below, the need for price-responsive demand is much greater in a less-regulated environment.

THE COMPETITIVENESS OF ELECTRICITY MARKETS

AS WE HAVE MENTIONED, THE ELECTRICITY RESTRUCTURING movement has been driven in large part by changes in generation technology and fuel prices that have made it economic to generate with smaller units, and by the integration of separate utility systems into larger regional networks, which has increased the size of the market that could be served by individual generators. For some, those changes have been enough to conclude that electricity generation no longer requires any more government regulation than the markets for other commodities: natural gas, crude oil, gold, orange juice, or coffee.

Impediments to Competition But a number of factors make electricity a different, and much more difficult, case. First, electricity is extremely costly to store. The technologies for storage—for instance, hydroelectric pump storage (pumping water uphill to store as potential energy) or bat-

teries—are quite inefficient. Combined with the high cost of storage is the need to balance supply and demand second-by-second. A shortfall or surplus of electricity means not only that a few consumers do not get all the power they would like; it can endanger the stability of the entire electricity grid. It is as if a late delivery from UPS to one location caused flat tires to develop in all other UPS trucks in the region.

Exacerbating the challenges created by costly storage and the need for real-time supply/demand balancing is the fact that almost no end-use consumers of electricity even have the technology to observe, let alone respond to, real-time prices. Demand is virtually completely inelastic in the short run. Thus, little or none of the supply/demand balancing can be done on the demand side, unless the grid operator forcibly curtails consumption. In the rare instances that this blunt instrument is used, it imposes significant costs on the consumers whose electricity supply is curtailed and often has political repercussions. Short-run supply elasticities are not a great help either in these markets. Generating units have hard capacity constraints that imply marginal cost turns steeply upward at a certain output. The combination of very inelastic short-run demand and supply (at peak times) with the real-time nature of the market (costly storage and grid reliability requirements) makes electricity markets especially vulnerable to the exercise of market power.

To see why this is the case, think about the dreadful summer afternoon when the temperature and humidity are at peak levels and the grid needs virtually all resources in production in order to meet the tremendous demand for electricity to run air-conditioning units. If the grid has only a few percent margin of reserve capacity at that time and there is a producer that is supplying more than a few percent of the total output, then that producer is *pivotal* in meeting the demand. Put differently, that producer can ask for an extremely high price in order to deliver the power and consumers—more specifically, the local utility that represents them in the wholesale power market—will pay it.

In most markets, there are other constraints that keep a single firm with a fairly small percentage of production from driving up the price by a large amount. If the good is storable, the buyers, or marketers in the middle, can store product to defend against such vulnerability. If end-use consumers receive the price information before buying, their own hesitancy to pay extreme prices discourages the seller from asking such a price. If there is supply elasticity, one firm demanding a high price for its output will just shift market share to another supplier. Each of these attributes is much less prevalent in electricity markets than in most other industries. The result is that the ability of firms with even modest market shares to exercise market power is greater than in most markets. That is why concentration measures that are widely used to diagnose the potential for market power are not very informative when applied in electricity markets.

The Specter of Volatile Prices What complicates this analysis, and makes these market power discussions so controversial, is that the same factors that exacerbate market power in electricity would combine to produce volatile electricity prices even if there were no attempt by sellers to exercise market power. Even absent market power, inelastic demand and supply will naturally lead to high prices at peak times as demand rises above the production capacity of generators in a market and further price increases result in little additional supply or reduction of demand. The prices would then efficiently reflect the scarcity of supply relative to demand.

The ability of firms with modest market shares to exercise market power is greater than in most other industries. That is why concentration measures are not very informative when applied to electricity markets.

Unfortunately, it is easy to show that in such a situation a firm of more than microscopic size can almost always do better than passively accepting these scarcity rents, attractive as they might be. By withholding a bit of its supply (or offering it to the market at an extremely high price), such a firm can drive the price still higher while losing little demand, and boost its profits. Thus, while it is easy to argue that volatile prices would be seen in even a perfectly competitive market with these attributes, it is equally easy to demonstrate that if firms of notable size are not exercising market power, they are doing so out of the goodness of their heart and against the interest of their shareholders.

In industries where market power is present, government regulation certainly can still lead to outcomes that are even less attractive to consumers. One reason is that market power is usually self-correcting. Short-run exercise of market power will usually attract entry of new competitors. Even the threat of such entry can act to discourage incumbents from pushing prices too high, just as the threat of demand response in the longer run can discipline producer behavior. Unfortunately, these effects might not be particularly strong in the electricity industry, and the evidence from existing electricity markets is not encouraging.

The reason is the simple economics of time discounting. Neither new entry nor demand-side responsiveness is likely to happen quickly. With environmental and other licensing restrictions, new entry can easily take three to

five years. Likewise, the technology for real-time price notification and response by consumers is not widely available and not likely to be for a number of years. And a seller's payoff to exercising market power now can be extremely high. Furthermore, the uncertain regulatory environment may actually encourage sellers to take profits while they can, since they do not know what the competitive or regulatory landscape will look like a year or two from now.

Some industry participants and observers have responded to evidence of market power by arguing that because market power is present in every market, it should not be a concern here, or even that market power is *necessary* in order for firms in electricity markets to recover their full costs and earn a reasonable return on their investment. Both claims are incorrect.

There is simply no support in theory or practice for the claim that firms—even firms in capital-intensive industries—must exercise market power in order to cover their costs.

First, market power is widespread, but there are also many examples of industries in which there is virtually no producer market power—for example, gold, natural gas, orange juice, and soybeans. What these goods have in common with one another and with electricity is the nearly perfect homogeneity of the product. In such markets, market power is not a necessary result, but the special aspects of electricity markets discussed above make it much more difficult to prevent the exercise of market power in the electricity industry than in other industries.

Second, there is simply no support in theory or practice for the claim that firms—even firms in capital intensive industries—must exercise market power in order to cover their costs. Production of most of the goods listed on the commodity page of *The Wall Street Journal* is capital intensive, yet most exhibit virtually no market power and still producers continue in business. Producers in those markets are able to earn scarcity rents, which means that the price they receive is greater than their marginal cost for most units they sell. Scarcity rents are also available to producers of electricity. Whether they sell in a centralized exchange or through bilateral arrangements, producers receive prices that reflect the market conditions, not their own marginal cost. Furthermore, in all markets established to date, sellers have the opportunity to earn some payments beyond the market price for power in return for being available on a standby basis as an emergency resource for the grid operator. Finally, economic theory does not support an argument that price must exceed the competitive level for firms to break even. In fact, under reasonable conditions, the absence of market power leads to normal returns on investment with exactly the socially optimal quantity of electricity generation capacity.

REGULATION: OUT WITH THE OLD, IN WITH THE NEW

IN IMPLICIT OR EXPLICIT RECOGNITION OF THE VULNERAbility of electricity markets to market power, almost every organized electricity market currently operating around the world has in place some form of price or revenue cap. In some markets this is described as a "software" limit: the bidding software simply cannot accept too many digits in the bid. In others, such as Australia, prices are capped at a "value of lost load," a proxy for consumption value. In many regions, price caps have been considered a necessary expedient to bridge the gap between current conditions and a world in which electricity markets feature price-responsive demand.

One of the more notorious pricing incidents occurred in the California Independent System Operator's (ISO) market for replacement reserve, a form of standby power. Following the confusion surrounding an order from the Federal Energy Regulatory Commission (FERC) deregulating the prices of this form of reserve, prices surged from the previously regulated range of $10 per megawatt to $9,999 per megawatt. Some market participants apparently thought that the ISO could not accept bids exceeding four digits, when in fact the price during the period in question had no limit at all, and a bid in the millions of dollars per megawatt would have been acceptable under the existing market design. Shortly after that incident, the ISO requested from FERC the ability to cap prices, and subsequently they were limited to $250 per megawatt for all ancillary services.

Economists in these markets have periodically been asked to derive the "right" level of price cap for such markets. The trade-off this engenders is a familiar one to economists. If the price cap is set too low, in the short run it will discourage production from high-marginal-cost plants, whereas in the long run it will lead to disinvestment in the industry as producers are unable to cover their cost of capital. If the price cap is set too high, the exercise of market power will cause significant wealth transfers from consumers to producers. The negative consequences of market power are not limited to transfers, however. Prices that are raised to artificially high levels as a result of the exercise of market power, rather than actual scarcities, can stimulate inefficient entry and can depress the expansion of electricity intensive enterprises.

So, in most electricity markets we are left with regulation by price caps. The price caps restrict the exercise of market power, but they also implicitly determine the scarcity

value to be collected by suppliers. Some market participants—sellers—have bemoaned the use of any price caps, arguing that they do not exist in "real" markets so should not be used here. It is clear, however, that without some sort of "backstop" price cap, there would be times in most markets when the price could rise without bounds. Until real-time price-responsive demand is feasible on a significant scale, price caps will remain a necessary evil. The dirty secret of restructuring is that it is replacing old forms of regulation with new ones.

CAN TRANSMISSION SAVE RESTRUCTURING?

WHEREAS ENTRY OF NEW PRODUCERS TAKES YEARS IN THE electricity industry, importation of electricity from neighboring areas can happen instantaneously. The problem is that these imports are limited by the capacity of the transmission lines between regions. Once a transmission line reaches its capacity, there is effectively no further ability to import. The same problem arises within grid systems, where a small area can be isolated, on the margin, from the rest of the system when transmission lines into the area become congested. Such "load pockets" exist in New York City and San Francisco, for instance, making those regions dependent not only on imported energy but also on a handful of local generators whose output is necessary to sustain reliable network operations. This is not an isolated phenomenon: *more than half* of the 288 generation units in the California ISO system have been designated as "must-run" for reliability purposes under some conditions. What electrical engineers call reliability concerns economists call local market power.

Various forms of alternative regulation have been applied to deal with local market power. Some of these regulations make sense, others do not. In the Pennsylvania, New Jersey, Maryland pool, the operators are given wide-ranging powers to reset the bids of generators that have been deemed to be exercising local market power. In the United Kingdom, generators were threatened with a referral to the Monopolies and Mergers Committee. In California, the reliability must-run contracts have remained at the center of controversy for nearly two years.

One of the most straightforward, and probably economic, ways of promoting competition would be to increase the contestability of separate geographic markets by beefing up the transmission infrastructure that serves them. With sufficient transmission capacity, attempts to raise prices in smaller regions become unprofitable because such attempts simply draw supply from neighboring regions. In fact, adding transmission capacity can actually decrease its use: the *threat* of competitive imports can be sufficient to forestall the exercise of market power.

It seems then that construction of additional transmission capacity would be an important component of electricity restructuring. But transmission lines can be costly to upgrade. More importantly, by its very nature, transmission capacity creates huge winners and losers among suppliers and consumers of electricity. The distributional effects can easily be orders of magnitude greater than the efficiency effects. If your firm owned all the generation in New York City, how would you feel about proposals to beef up import capabilities into the city? The economic and political dynamics that have arisen in the electricity industry resemble those surrounding trade policy, as firms seek to preserve their advantages in serving local markets. Generators inside load pockets have much to learn from the U.S. steel industry.

Generically, the production of electricity is no longer a natural monopoly, but at specific locations and specific times, companies with even small market shares still can exercise substantial market power. To ensure viable competition, some combination of generation and transmission capacity increases must take place. These capacity increases may in themselves be inefficient. The question then becomes: are the costs of sustaining competition less than those of sustaining regulation?

Until real-time price-responsive demand is feasible on a significant scale, price caps will remain a necessary evil. The dirty secret of restructuring is that it is replacing old forms of regulation with new ones.

THE RECORD SO FAR

IN CALIFORNIA AND A NUMBER OF OTHER MARKETS, ADVOcates of restructuring have seen prices fall over the past few years and have declared victory. But independent of restructuring, electricity prices were expected to decline during this period in many parts of the United States. In California, for instance, as the sunk investment in nuclear plants was paid off and the high-cost PURPA contracts expired, even under old-style regulation consumers would have seen their prices fall. Restructuring could not make those stranded investments disappear, and old-style regulation could not make their impact continue indefinitely.

Instead of being distracted by historical accounting, one might ask just how competitive electricity markets have turned out to be. This is a first step in comparing the costs and benefits of restructuring. Catherine Wolfram, in "Measuring Duopoly Power in the British Electricity Spot Market" (*American Economic Review* 89), finds that prices exceeded competitive levels by around 20 to 25 percent in the England and Wales pool, and Severin Borenstein, James

Bushnell, and Frank Wolak (in their working paper, "Diagnosing Market Power in California's Deregulated Wholesale Electricity Market") estimate that energy purchase costs in California averaged about 16 percent above competitive levels during 1998 and 1999. These average figures mask large fluctuations in the severity of market power, which has been greatest during peak demand periods. In addition, evidence from California and New England, as well as from England, indicates that market power was felt more severely in markets for capacity and reserves than in markets for the energy itself.

These outcomes have transpired in a context of significant continued regulation. When California operated under a $250 price cap for energy before October 1999 (at which time it was raised to $750), the limit was binding on many occasions. Many generators were also earning significant revenues from reliability must-run contracts. In England, generators have been subjected to average-price caps, threats of antitrust enforcement, and periodic jawboning from regulators. Regulators in England also imposed sales of power under vesting contracts, set at administrated prices, on the incumbent suppliers, as well as minimum purchase quotas for British coal. In short, these results should not be construed as the extent of market power that would be experienced in a completely deregulated electricity market.

England provides perhaps the most serious cautionary tale about electricity restructuring. A high level of dissatisfaction with the outcomes in the British power market has led to a near total demolition of those market institutions. Much of the blame that has been placed upon the market's design should more properly be attributed to the market structure, which has remained fairly concentrated. The British experience with electricity markets over the past decade has so seriously eroded faith in markets that ominous "good behavior" clauses have been proposed as a requirement for generation firm licensing there. Such clauses have the potential to be far more arbitrary and intrusive than the traditional forms of regulation that have been employed in the United States during the twentieth century.

CONCLUSIONS

THE MOVEMENT TOWARD LESS REGULATION AND MORE reliance on market processes in the electricity industry has enormous potential benefits but also potential risks. A move toward deregulation that does not take the issue of market power seriously can undermine the goals of industry restructuring and even, as in the case of England, produce a regulatory backlash. Any restructuring initiative must recognize that the lack of economic storage and of price-responsive demand can produce serious market disruptions. Furthermore, levels of transmission capacity that may have been adequate under regulation may not be able to support effective competition.

In restructured electricity markets, some level of market power seems likely to persist. Given the enormous size of this industry, even small amounts of market power imply large transfers from consumers. This means that for consumers restructured electricity markets may in fact be *more* costly in the short run than were their regulated predecessors. For restructuring to benefit consumers, the long-term gains stemming from improved investment decisions on both the demand and supply sides of the industry must be sufficient to outweigh the potential short-run costs. Unfortunately, measuring those long-term benefits will be very difficult because it will rest on comparing the efficiency of investment under restructuring with the investment that would have occurred if traditional regulation had continued.

readings

- Severin Borenstein. "Understanding Competitive Pricing and Market Power in Wholesale Electricity Markets." *The Electricity Journal* 13 (forthcoming).
- Severin Borenstein, James B. Bushnell, and Frank A. Wolak. "Diagnosing Market Power in California's Deregulated Wholesale Electricity Market." University of California Energy Institute POWER Working Paper PWP-064r, 2000.
- Peter Cramton and Jeffrey Lein. "Eliminating the Flaws in New England's Reserve Markets." University of Maryland Working Paper, 2000.
- Paul Joskow. "Restructuring, Competition and Regulatory Reform in the U.S. Electricity Sector." *Journal of Economic Perspectives* 11 (1997): 119.
- Frank Wolak. "Market Design and the Behavior of Prices in Restructured Electricity Markets: An International Comparison." Stanford University Working Paper, 1997.
- Frank Wolak, Robert Nordhaus, and Carl Shapiro. "Preliminary Report on the Operation of the Ancillary Services Markets of the California Independent System Operator (ISO)." California Independent System Operator, Folsom, Calif., 1998.
- Frank Wolak and Robert H. Patrick. "The Impact of Market Rules and Market Structure on the Price Determination Process in the England and Wales Electricity Market." University of California Energy Institute POWER Working Paper PWP-047, 1997.
- Catherine Wolfram. "Measuring Duopoly Power in the British Electricity Spot Market." *American Economic Review* 89 (1999): 805.
- Catherine Wolfram. "Electricity Markets: Should the Rest of the World Adopt the United Kingdom's Reforms?" *Regulation* 22 (2000): 48.