How reliable is our electric system under partial deregulation?

Keeping the Power On

Benjamin Zycher
RAND

Deregulation of electricity generation — partial deregulation in that transmission and distribution remain regulated — has changed the nature of our time-honored electric "reliability" measures. The traditional approach to system "reliability" in the electric utility sector has emphasized the difference between the maximum ("peak") capacity to produce electricity and peak demand ("load"). This approach provides a crude measure of the ability of the system to withstand unexpected outages of system components without the loss of service. Accordingly, the reliability "measure" used traditionally as an index of blackout/brownout potential essentially is an accounting exercise: Will generation units be available to provide backup power in the event of an adverse occurrence? This approach is important as a short-term question, given the possibility of severe heat waves or unexpected outages of generating units. Such questions forever will be with us as long as it is costly (that is, technologically difficult) to store electricity.

However, these questions do not examine the likely evolution of the system over time. In the current infancy of deregulated, competitive electricity markets, this dynamic question is most relevant for predicting the long run evolution of the system. And that long run evolution will be heavily influenced by the incentives faced by producers, consumers, regulators, and others — that is, by the economic environment within which market participants make their decisions. In this sense, the "reliability" question is largely an economic one. It is not, for the most part, a technological one of system engineering although, obviously, technical factors influence the underlying economics heavily.

Accordingly, this discussion examines the economic meaning of "reliability." As discussed in detail below, the competitive market price of electricity will contain two components: (1) the costs of providing optimal reliability, and (2) a price premium that would be analogous to a surety bond that the utility will lose if it fails to provide the optimal reliability reflected in the market price. Analyses of future electricity prices that focus only upon capital, fuel, labor, and other directly observable costs are likely to understate competitive market prices.

Competitive market forces will lead producers to invest the price premium in firm-specific investments that, ironically, will increase scale economies in the industry by some factor and will lead utilities to compete over broader geographic regions and in unfamiliar markets. The result is that the use of interstate transmission is likely to increase over time. But decision-makers pondering state-regulated transmission investments will have net incentives to obtain free rides on the willingness of other states to pay for transmission system reliability. And the Independent System Operators (ISOs) — the nonprofit entities that operate but have no ownership interest in transmission systems — are not likely to achieve efficient pricing of, or optimal investment in, transmission services. States pursuing a rationalization of their electric power markets through deregulation should turn instead to deregulated transmission companies because their historical market power will be constrained by rivalry from decentralized natural gas generation and transmission.

We turn first to the economic meaning of "reliability," one very important component of service quality. Of particular importance are price incentives for optimal system reliability under market competition. Optimal reliability is not "perfect" reliability, and in principle may not be even "high" reliability (although it almost certainly is in practice). Of particular interest (and concern) are the differing incentives facing deregulated generation companies competing with one another on the one hand, and regulated owners of transmission and distribution systems on the other.

THE ECONOMICS OF RELIABILITY

Like any dimension of product quality, reliability — more precisely higher reliability, however measured, rather

Benjamin Zycher is a RAND senior economist and an adjunct scholar of the Cato Institute. He has written on such topics as insurance, the environment and energy economics, and the economics of defense. He can be contacted via e-mail at bzycher@rand.org.

than lower — is costly. An additional increment of reliability may or may not be worth what it costs in terms of other goods and services. Accordingly, "perfect" reliability is never observed, not so much because it cannot be achieved but instead because it is not worth what it would cost. "Near-perfect" reliability certainly is achievable, however expensive, and, obviously, choices have to be made over time between lesser and greater degrees of near-perfection. Lesser near-perfection still is pretty good, and decision-makers since the electric age began rationally have opted for lesser near-perfection rather than greater.

So the question is not whether reliability under deregulated competition will be perfect or even high; the question is whether it will be optimal or efficient in light of the marginal benefits and costs of reliability to consumers and producers. There is nothing new about this question. But the changed incentives facing market participants under partial deregulation lead us to ask whether optimal reliability is the outcome that we can predict. Market incentives under truly deregulated generation will provide optimal reliability. However, continued regulation of transmission and distribution — particularly if conducted on a multi-state or federal basis — is likely to yield less-than-optimal reliability for the system as a whole as long as costs and political responsibility for outages can be shifted to others.

Utility Incentives To Prevent Blackouts Economists for many years have viewed reputations and brand names as private institutions providing incentives to honor contracts. This enforcement mechanism relies upon repeat sales to satisfied customers who would desert a given producer for competitors if the producer failed to honor promises. But repeat sales are not enough; prices must be sufficiently high that the gain to the producer from failing to provide the level of reliability contracted by consumers would be less than the present value of the returns to provision of higher quality. In other words, electricity consumers are willing to pay for a certain (high) level of quality but producers will likely "cheat" — provide an amount of quality lower than that contracted for — if doing so is worth more than the stream of returns from providing the contractual level of quality over time. This means that in a deregulated market in which regulators no longer are responsible for the level of reliability, market prices will be bid up so as to ensure quality. As long as producers can capture those returns — as long as property rights to the profits earned from selling electricity are enforced — the generation market will produce optimal reliability.

Assume, realistically, that consumers of electricity (whether industrial, commercial, or residential) can measure reliability *ex ante* only with some (substantial) difficulty. But consumers can perceive the delivery of a low-quality product easily *ex post*, in that a blackout is obvious. On the margin, consumers are willing to pay some amount for higher reliability, which can be provided only at some cost. If the marginal value is higher than the marginal cost, then the profit-maximizing producer who invests in capital assets will receive returns in the stream of higher prices paid over time for more reliable service. Let us ignore the possibility that producers prefer to deal honestly as a matter of course; the decision on whether actually to make the investment in higher reliability — that is, whether to deliver the higher-quality product promised — hinges on whether the producer perceives a higher one-time return from "cheating" than from delivering the more-reliable electric services over time. The gain is one-time in this simple world because once the blackout occurs, consumers realize that the producer has failed to provide optimal quality, and so they will shift their business to other producers.

In a more complex world, consumers cannot distinguish easily between such "cheating" and simple bad luck because the optimal level of reliability for which they contract is less than perfect. In other words, if the one-time gain from cheating is greater than the present value of the higher returns to more-reliable generation, the producer will not make the promised investments (or will not maintain them), and reliability will be inefficiently low. If, on the other hand, the latter present value is greater than the one-time gain from cheating, the profit-maximizing producer will be led to provide optimal quality, and so reliability in the context of the electricity market will be provided at an optimal level.

But optimal reliability is not perfect reliability. Once a blackout occurs, consumers must decide whether it was the result of cheating on the part of the producer, or whether it was a "fluke" that occurred even though the producer made optimal investments. In large part, this issue will be decided in light of the actual circumstances of the blackout. If it occurred on a mild spring day with no other independent sources of problems, consumers might suspect that something is amiss, although simple misfortune emerging in such forms as equipment failures obviously can occur anytime. In large part as well, consumers will develop knowledge over time: If a certain producer experiences outages at a much higher rate than other producers facing similar conditions, consumers will conclude that something is wrong. Perhaps most important, competitors will advertise, highlighting their own relative strengths and others' relative weaknesses.

The Market Price For Reliability These observations carry important implications for the evolution of electric utilities and their market behavior. Acceptance by a producer of the market price for electricity, including the price premium for optimal reliability, carries an important, but subtle, opportunity cost. Because the utility could earn a one-time return by failing to provide optimal reliability (a fact that would not be known by customers until a failure of the system), honest dealing — provision of optimal reliability — prevents the utility from earning that one-time return. That forgone opportunity is a real economic cost to the producer, however hidden it may be. Accordingly, the competitive market price of electricity must compensate utilities not only for the costs of providing optimal reliability, but also for the opportunity cost of not exploiting the difficulty faced by consumers in measuring the reliability of the electric services that they purchase. This price premium is analogous to a surety bond that the utility loses if it fails to provide the optimal quality reflected in the market price. Under competitive pressures, the present value of that stream equals the one-time gain from "cheating" — that is, a failure to deliver optimal quality.

Because electricity cannot be stored economically, the price premium is not likely to be trivial. If consumers could store large amounts of electricity cheaply, reliability on the part of utilities would not be worth a great deal, and so the competitive price premium for reliability would be small. But because electricity cannot be stored cheaply, and because reliability as a crude generalization is worth a great deal, the price premium is likely to be important in terms of overall revenues. Accordingly, one conclusion can be stated as follows: Analyses of future electricity prices that focus only upon capital, fuel, labor, and other directly observable costs are likely to understate competitive market prices. And analyses of capital investment predicated upon those "competitive" prices are likely to underestimate market incentives to provide reliability.

Reliability and Firm-Specific Capital The price premium earned for provision of optimal reliability would appear to yield economic returns for utilities that are higher than normal (or competitive). This might suggest that more competitors would enter the market over time. But that conclusion cannot be correct because entry would drive prices below the equilibrium level offered by consumers in order to ensure optimal reliability.

Instead, competition to dissipate the above-normal returns must take place in nonprice dimensions of the business. Only one such dimension is consistent with competitive market forces: Utilities will be driven to make investments in firm-specific capital. To say that the capital is "firm-specific" means that its full value cannot be salvaged through sale to another firm. Such investments must be firm-specific because the incentive to provide optimal reliability hinges upon the greater loss anticipated from a failure to deliver the promised reliability than the gain anticipated from "cheating." If the capital were not firm-specific, the utility simply could sell it to another firm, thus avoiding any loss from "cheating." In other words, in order for the higher market price to provide efficient incentives to the utilities and the correct signals to consumers, the utility must actually lose the economic value of the firm-specific assets if it fails to provide optimal reliability.

Accordingly, the utility will invest an amount in firm-specific capital equal to the present value of the stream of payments earned for supplying optimal reliability. Investments in generating capacity and in advertising are obvious forms that such firm-specific investment is likely to take. Investment in generation capital assets clearly is a fixed cost and advertising costs also are fixed in the sense that they do not vary in any necessary way with output. Our second central conclusion, therefore, can be stated as follows: Competition, perhaps ironically, will increase scale economies in the industry by some factor and will lead utilities to compete over broader geographic regions and in unfamiliar markets.

Interestingly enough, the firm-specific capital that the utility would lose were it to "cheat" does not have to provide information or other services to consumers directly. It can consist of non-salvageable production assets; large units of equipment are an obvious example because they often are designed to a specific firm's specifications and because resale involves substantial transaction costs. The non-salvageable portion of the value of such assets is likely to be greater when the assets are owned rather than leased. Accordingly, we can predict that the degree to which utility assets are owned rather than leased will be higher than would be predicted in a simple analysis of utility operations under competition. Whether direct ownership will be higher or lower than is the case under the historical incentives (whether positive or negative) yielded by traditional rate regulation is an empirical issue that cannot be predicted on *ex ante* grounds.

INCENTIVES UNDER PARTIAL DEREGULATION ARE PROBLEMATIC

COMPETITIVE *generation* MARKETS INTRINSICALLY WILL have efficient incentives to provide optimal reliability. The presence of information costs for consumers does not alter that conclusion, although it has important implications for the evolution of the market. The real reliability problem is likely to emerge in *transmission and distribution*. Because attendant price regulation will be multi-state or federal in nature, states will have net incentives to obtain free rides on the willingness of others to pay for reliability — the benefits of which accrue not only to themselves, but to others as well.

Consider the international market for pharmaceuticals. The long run supply of medical drugs depends upon the long run price, and thus upon patent protection. In the short run, most costs for a given drug are sunk in development. Production of the drug will continue as long as the price does not fall below marginal production cost, which for many drugs is trivial. Many small nations, such as Spain and New Zealand, have imposed stringent price controls on pharmaceuticals. They have done this in the knowledge that they will be able to continue to obtain the drugs even at very low prices, and that they are a sufficiently small part of the market that their price policies are unlikely to appreciably affect research and development. Other nations, unfortunately, also are adopting "lowest price" policies, in which the prices that they are willing to pay are some weighted average of the lowest prices observed in the world. This means that the free-riders, in effect, are driving the market. This will create adverse long-term consequences for research, development, and the ability of the medical sector to supply cures through the pharmaceutical market.

Similarly, from the viewpoint of individual states, investment in optimal reliability of the transmission and distribution system is just such a collective good. In the short run, each state has incentives to allow others to pay for that investment, particularly in interstate transmission. The benefits of such investment in part accrue to others, while the adverse effects of breakdowns caused by suboptimal investment within the state in part will be borne by others. Accordingly, each state participating in an interstate transmission grid will have incentives to obtain free rides. This is true even if there is multistate cooperation of whatever sort; conditions of weather, geography, and the like inevitably will vary across states, and so individual states will attempt to understate their demands and overstate their costs relative to others in order to bear as small a proportion of the total costs themselves.

Again, scale economies in generation are likely to increase under competition, and so interstate transmission is likely over time to increase as well. This means that this incentive to obtain free rides can be predicted to grow, with consequences for system reliability that can be predicted to be adverse qualitatively, even as the quantitative magnitude of this problem is yet to be revealed.

In short, it is the “joint” components of the newly emerging competitive system that will represent the most important source of future reliability problems, defined as reliability that is inefficiently low. This raises an important question: How might institutions of joint transmission and distribution be designed to overcome or minimize this problem?

Efficient Transmission Prices Essential In the traditional and familiar world of regulated electricity markets, electric power has been supplied by firms integrated vertically so that they provide generation, transmission, and distribution services. Because consumers have purchased a bundled good — delivered electricity — there has been no need for markets to price the components of the bundled good separately (although endless accounting contortions have been pursued for purposes of regulatory analysis). Since the central legacy of the traditional regulatory system is an inefficient and costly stock of generation capital, most discussions of the deregulation of electricity markets have concentrated on competition and pricing in the generation sector.

But — again — the “reliability” issue, as a problem of market economics, is likely to arise in transmission because of continued regulation and an attendant free-rider problem across jurisdictions. As discussed below, this problem is likely to be exacerbated by the status, much-discussed but little explained, of the ISOs as “non-profit.” Most current proposals in the transmission context emphasize nondiscriminatory open access rather than efficient pricing. But it is efficient pricing that is at the heart of the “reliability” problem, defined properly. If “reliability” as efficiency is the primary goal then, from the standpoint of economic analysis, rationalization of the electric power industry’s capital stock is a central objective. This must mean that pricing of transmission services also is central because of the close relationship between generation and transmission. Generation and transmission obviously are complements in the production of delivered electricity, but they also increasingly are substitutes — a trend that will expand as geographic competition rises in importance.

Interregional transmission has the economic effect of reducing regional differences in generation costs as cheaper imported power can substitute for more expensive electricity produced locally. More subtly, transmission also serves the economic function of reducing the cost of achieving a given level of reliability; imported power can be substituted for local generation assets. Efficient pricing of transmission services means that prices must equal marginal cost. (In the short run, marginal cost is electricity lost during transmission and — in the event of transmission capacity constraints — transmission congestion “rents” or competitive market prices sufficiently high to clear the market for the transmission capacity. Long run marginal cost is not relevant to efficient pricing except in the sense that additional investment is efficient when price rises above long run marginal cost.) In the short run, efficient transmission pricing is necessary for optimal geographic allocation of current electricity production. Thus, it also is necessary for optimal reliability, regardless of whether the existing stock of generation capital is efficient. If transmission services are priced below marginal cost, incentives are provided for inefficient importation of electricity from distant generating sources even though acquisition of power from more local sources would consume fewer resources. If transmission services are priced above marginal cost, analogously, the geographic market for power would be localized artificially with adverse implications for resource costs and for the competitiveness of the market.

In the long run — which does not necessarily mean a long period of time — efficient pricing of transmission services is required for correct signals to the market in terms of expansion and location of both transmission and generation capital. It thus is necessary for optimal reliability. In particular, if transmission prices do not reflect transmission costs correctly — including the economic cost of capacity constraints, and thus the marginal economic value of additional capacity — the inevitable result will be distortion of choices between expansion of transmission capacity and expansion of generation capacity geographically closer to electricity demands.

ISOS EXACERBATE INCENTIVES PROBLEM

THE ISOs WILL NOT ACTUALLY FACE A MARKET TEST IMPOSED by consumers or producers; indeed, the ISOs will have no ownership interest — no property rights — in generation or transmission assets. There is little reason to believe that efficient pricing of transmission services will be an important goal of the ISOs. (See “Can Nonprofit Transmission Be Independent?” in *Regulation*, Vol. 23, No. 3.) The California experience suggests that decisions within the ISO framework will be affected heavily by political considerations having little to do with efficiency and optimal reliability. For example, the Western Power Exchange proposal to emphasize importation of power from distant coal-fired generators is a political competitor to the proposal of the Coalition for

Comparable Transmission, which is more favorable to local gas-fired generation. Whatever the merits of the respective proposals, they are motivated largely by distributional (wealth transfer) considerations rather than by efficiency concerns. Moreover, various restrictions imposed upon the California ISO are likely to create efficiency and reliability problems. A recent analysis finds that operational rules will prevent efficient (least cost) dispatch of generation, will reduce reliability by causing operation of lines at their security limits even if operation at less than full load would be cheaper, and will price congestion perversely.

Constraints On T & D Market Power Accordingly, a central issue emerges: How can efficient transmission pricing be achieved? The conventional wisdom is that transmission pricing must be regulated and coordinated centrally because of supposed monopoly considerations and because of the engineering interrelationships inherent in electricity networks. But technological advance is eroding the ability of transmission owners to exercise monopoly power. Combined-cycle gas turbine technology has reduced the capital cost/fuel cost ratio sharply, thus increasing the geographic mobility of competitive electricity generation fired by natural gas. This means that the gas pipeline network is becoming a competitor to the electricity transmission network; gas pipelines in effect will transport "electricity" in another form.

In a larger sense, the likelihood that any regulatory system will yield "efficient" transmission prices must be viewed as exceptionally low. Quite apart from the sharp conceptual issues involved, we do have considerable experience in this context. "Efficiency" is not what drives the choices made by regulators, whatever their rhetoric of the day.

That transmission is naturally monopolistic and, thus, that market prices for transmission services would be inefficiently high is virtually an article of faith in most of the relevant literature. But, as noted above, technological advance is eroding this view even on its own terms. In effect, new technology is reducing sharply the cost of installing small-scale generation capacity anywhere that natural gas can be made available, thus bypassing or reducing reliance upon the transmission grid. By the time that legislation is implemented to restructure the electricity market (that is, to ratify the market forces at work), achievement of scale economies in the production of small generators will make them highly competitive in some, perhaps many, markets. The ongoing movement toward market pricing of electricity will speed this process, thus increasing competitive pressures on transmission pricing. Restructuring proposals that fail to take this dynamic process into account are likely to be out of date when the legislation is implemented.

Even in the absence of new generation technologies, competitive forces still constrain transmission prices. The availability of alternative transmission routes in many parts of the United States — particularly the eastern half of the nation — will constrain prices. Also important is the ability of customers to respond to price increases by reducing power purchases and by investing in conservation measures. More important in the long run — which, again, is not necessarily a lengthy period — is the ability of customers to move facilities or to avoid locating in geographic regions lacking significant competitive protection. Such market forces will act as powerful constraints on inefficiently high transmission prices.

CONCLUSION

THE CURRENT DEBATE IN THE CONTEXT OF DEREGULATED electricity markets has emphasized such important but ancillary issues as the problem of historical ("stranded") generation costs. It is important to change this orientation toward the long run organization of transmission services. It is clear that market forces will increase scale economies in generation, yielding greater geographic competition and increasing the importance of efficient transmission pricing. More efficient pricing will influence important choices between local generation and the importation of power, among other parameters. What is also clear is that old habits die hard; the casual assumption of important monopoly power on the part of transmission and distribution systems, deeply imbedded in historical thinking, is creating real problems in terms of the organization of the non-generation sectors of electric power supply. But even that framework is proving incorrect, as market forces drive the development of new technologies designed in part to circumvent the obstacles and rigidities inherent in systems that matured under such political institutions as public utility regulation. The current partial deregulation of electric power services provides a clear example of the economic benefits available to societies willing to trust in market institutions. R

r e a d i n g s

- Severin Borenstein and James Bushnell. "Electricity Restructuring: Deregulation or Reregulation?" *Regulation*, Vol. 23, No. 2 (2000).
- Energy Information Administration. *The Changing Structure of the Electric Power Industry: An Update*. Washington, D.C.: EIA, 1996.
- Paul Joskow. "Restructuring, Competition and Regulatory Reform in the U.S. Electricity Sector." *Journal of Economic Perspectives*, Vol. 11 (1997).
- Benjamin Klein. "Why Hold-Ups Occur: The Self-Enforcing Range of Contractual Relationships." In John R. Lott, ed., *Uncertainty and Economic Evolution: Essays In Honor of Armen A. Alchian*. New York: Routledge, 1997.
- Benjamin Klein and Keith Leffler. "The Role of Market Forces in Assuring Contractual Performance." *Journal of Political Economy*, August 1981.
- Vernon L. Smith. "Regulatory Reform In the Electric Power Industry." *Regulation*, Vol. 19, No. 1 (1996).
- Catherine Wolfram. "Electricity Markets: Should the Rest of the World Adopt the United Kingdom's Reforms?" *Regulation*, Vol. 22, No. 4 (1999).