THE PIG IN THE PYTHON: IS LUMPY CAPACITY INVESTMENT USED AND USEFUL?

William J. Baumol[*]
J. Gregory Sidak[**]

I. INTRODUCTION

A recurring question in the regulation of public utilities is whether the entity should be permitted to recover the cost of particular assets through its allowed rates. The traditional standards have been the backward-looking "prudency" test and the forward-looking "used-and-useful" test. Numerous state statutes and innumerable regulatory decisions since the early twentieth century have relied on the used-and-useful test to determine whether a particular asset belonging to a utility should be included in or excluded from the utility's rate base.[1]

Under the used-and-useful test, if regulators disallowed a particular asset as not used and useful, then the utility could not recover, through its rates, the capital invested in that asset. Nor could the utility earn, through its rates, any return on that invested capital. Thus, regulatory disallowance of a particular asset may be said to deny the utility a return *of*, and *on*, its investment in that asset. The used-and-useful test is the legal standard by which many regulators make this determination to permit or disallow a utility's recovery of, and on, its invested capital.

By the end of the twentieth century, the practice of disallowing certain investments as not used or not useful became more prevalent, particularly as interest groups organized politically to resist the construction of new nuclear

---

* Professor of Economics, New York University and Senior Research Economist and Professor Emeritus, Princeton University.

** F. K. Weyerhaeuser Fellow in Law and Economics Emeritus, American Enterprise Institute for Public Policy Research. We are grateful to the C.V. Starr Center for Applied Economics at New York University for support of this work. The views expressed here are solely our own, and not those of New York University, Princeton University, or the American Enterprise Institute, none of which takes institutional positions on specific legislative, regulatory, adjudicatory, or executive matters.

1. For a representative statute, see 66 PA. CONS. STAT. § 1315 (Supp. 2001) ("the cost of construction or expansion of a facility undertaken by a public utility producing . . . electricity shall not be made a part of the rate base nor otherwise included in the rates charged by the electric utility until such time as the facility is used and useful in service to the public"). For classic statements of the principle in case law, see *Duquesne Light Co. v. Barasch*, 488 U.S. 299, 308-09 (1989) ("To the extent utilities' investments turn out to be bad ones (such as plants that are canceled and so never used and useful to the public), the utilities suffer because the investments have no fair value and so justify no return."); Missouri *ex rel.* Southwestern Bell Tel. v. Public Serv. Comm'n, 262 U.S. 276, 292 n.3 (1923) ("In estimating replacement cost the first step is to determine what part of the property owned is used and useful in the public service. That involves, among other things. . . the question whether the size and capacity of the plant are, in part, excessive."). For a legal (rather than economic) analysis of the used-and-useful test, see James J. Hoecker, *"Used and Useful": Autopsy of a Ratemaking Policy*, 8 ENERGY L.J. 303 (1987).

power facilities for electricity generation.[2] The used-and-useful test is distinguished from the prudent-investment test, which permits the utility to recover, through its allowed rates, the historical cost of its investments, provided that they were prudent when made.[3]

In Part II, we ask, when is an investment beneficial? A utility's investment in seemingly "excess" capacity immediately provides an option to consumers, an option that has substantial economic value. In that sense, excess capacity is a capital investment that not only is currently used by the utility, but also is currently useful to consumers. Excess capacity is a form of insurance that protects consumers when demand unexpectedly surges, supply unexpectedly collapses, or both occur simultaneously. In addition, where construction entails substantial economies of scale and demand can be expected to grow, excess capacity during the period following construction can constitute another substantial benefit to consumers in the form of long-run cost savings.

In Part III, we examine the used-and-useful test from an economic perspective. An understanding of the option value of excess capacity helps to explain the difference between foresight and hindsight standards for the bearing of market risk in the regulated network industries. Such an understanding, moreover, suggests that regulators may well err on the side of declaring that an asset is not used and useful when it actually is. It is possible that greater appreciation by regulators and courts of the option value of seemingly "excess" capacity can lead cases in this area of law to be differently decided.[4]

In Part IV, we discuss the California electricity crisis of 2000-01 as an

---

2. The definitive scholarly examination of this development is Richard J. Pierce, Jr., *The Regulatory Treatment of Mistakes in Retrospect: Canceled Plants and Excess Capacity*, 132 U. PA. L. REV. 497 (1984). [hereinafter Pierce]. *See also* Paul Rodgers & Charles D. Gray, *State Commission Treatment of Nuclear Plant Cancellation Costs*, 13 HOFSTRA L. REV. 443 (1985); Melvin G. Dakin, *The Changing Nature of Public Utility Regulation: The Used and Useful Property Rate Base Versus the Capitalization Rate Base in the Nuclear Age*, 45 LA. L. REV. 1033 (1985); Claire A. Watkins, Comment, *Nuclear Power Rate Regulation After Eastern Enterprises: Are Ratepayers Being Taken for a Ride?*, 28 B.C. ENVTL. AFF. L. REV. 191 (2000). For further discussion of the prudent-investment rule, see *infra* notes 12-15 and accompanying text.

3. In practice, some states employ a rule that requires not only that the utility's investment be currently used and useful, but also that it was prudent when made. In *Duquesne*, for example, the Supreme Court noted:

> Pennsylvania has modified the system in several instances . . . when prudent investments will never be used and useful. For such occurrences, it has allowed amortization of the capital lost, but does not allow the utility to earn a return on that investment. The loss to utilities from prudent but ultimately unsuccessful investments under such a system is greater than under a pure prudent investment rule, but less than under a fair value approach.
>
> Pennsylvania's modification slightly increases the overall risk of investments in utilities over the pure prudent investment rule. Presumably the [public utilities commission] adjusts the risk premium element of the rate of return on equity accordingly.

488 U.S. at 312 n.7 (citations omitted). One commentator argues that the used-and-useful test and the prudent investment test are not substitutes. John Burritt McArthur, *Cost Responsibility or Regulatory Indulgence for Electricity's Stranded Costs*?, 47 AM. U. L. REV. 775, 882 (1998), "That each can exclude costs from the ratebase independently can be seen from the fact that a plant may be useful, but its costs have not been incurred imprudently." *Id.* (citing Pierce, *supra* note 2, at 513).

4. The leading Supreme Court case in this area is Duquesne Light Co. v. Barasch, 488 U.S. 299 (1989), from which the Court gave no indication of departing in Verizon Comm. Inc. v. FCC, 122 S. Ct. 1646, 1666 (2002).

illustration of the ways in which, by embracing a narrow interpretation of "used and useful," regulators made consumers unnecessarily vulnerable to outlier events.

II. WHEN IS AN INVESTMENT ECONOMICALLY BENEFICIAL?

Whether an investment is economically beneficial depends upon many factors. Obviously, if current capacity is insufficient to meet demand at prevailing prices, and if an investment in generating plant (in the case of an electric utility) yields added capacity, then the output generated by that added capacity unquestionably constitutes an economic benefit. Where capacity is not currently in short supply, the question of whether there are any economic benefits to utility customers and to the general public from capacity expansion requires further analysis.

A. The Economic Meaning of "Lumpy" Investment

Economists sometimes speak of an investment being "lumpy." Because that concept is critical to the discussion that follows, we take a moment now to define it. In the most simplistic depiction of investment in economic theory, a business entity can add productive capacity in infinitesimally small increments, so that it can increase its output one unit at a time. Whether it is rising, falling, or constant, the marginal cost curve is smooth over a range of output, rather than having a jerky, stair-step appearance.

In the real world, of course, few productive activities exhibit such continuity in capacity additions. If the python could digest a few pieces of the pig at a time, or if it could catch a little pig at will, then that python would not need to slither about with a lump in its gut. But the technology of pigs and pythons imposes certain physical constraints: if there is to be any python meal at all, it must consist of at least a minimum-sized pig. The pig provides the python current sustenance, but the pig is also the python's lumpy investment in future nourishment.

Generation capacity is our pig, and the electric utility our python. When capacity constrains the utility's output, the utility must add capacity in discrete amounts having some minimum efficient size. A utility, for example, cannot add one kilowatt of generation capacity at a time, but rather must add all of the capacity inherent in a single generator or a single power plant. This inability to add capacity in tiny, tailor-made increments means that new capacity will often give the utility more capacity than it needs for immediate purposes.

In terms of technological attributes that bear on capacity, an investment is said to be lumpy if technological attributes of the item mean that there is a capacity level, X, such that capital of the type in question is unavailable with capacity lower than X, or that lower capacity capital is so costly as to make its acquisition unattractive or impractical. A lumpy investment is one that is only available on a substantial scale; when acquired, the investment significantly expands the firm's total capacity. Lumpy capacity is not a continuous function of output, but rather one that takes discrete steps of considerable size relative to total current demand. A number of influences generate lumpiness. In the electric utility area, lumpiness clearly is affected by demand and its role in the load-forecasting and capacity-

planning process, as carried out to ensure adequate capacity; by the inherent uncertainty of the future course of demand; and by the need to deal with technology whose scale economy attributes require substantial additions to capacity as a matter of practicality.

Suppose that demand is growing, or that it is necessary to replace older plants and equipment periodically, or that small additions to capacity are prohibitively expensive, while large expansions are not. If any of these conditions holds, then, even in a world of perfect certainty about the future and zero lag in plant construction, the so-called excess capacity becomes an inescapable feature of the obligation to serve customers.[5] To illustrate why that is so, suppose that current capacity is just sufficient to generate *X* kilowatt hours of electricity per year and that an economically viable new plant comes with a minimum capacity of *Y* kilowatt hours. Then, with steadily growing demand that will exceed the *X*-kilowatt-hour capacity of the system in another year, it becomes necessary to acquire a new plant that will become operational on the day that quantity demanded first exceeds *X*. At first, there will unavoidably be nearly *Y* units of excess capacity, because only a small portion of the new plant's *Y*-unit capacity will be needed to meet the small increment in demand that exceeds the system's earlier *X*-unit capacity. There is no economical way, however, to avoid this excess capacity. It can only be avoided at a cost to consumers that is likely to be substantial—that of acquiring additions to plant in small and prohibitively expensive increments, or of requiring customers to suffer shortages and interrupted service.

B. Capacity Expansion in the Absence of Current Shortages

Several possible benefits can arise from the addition of capacity when current capacity is not yet fully in operation. These benefits include: (1) any resulting fuel savings; (2) any enhancement in reliability attributable to the fact that the plants are newer than those previously available; (3) adaptation to the requirements of rational investment policy in an arena in which technology and other factors such as marked scale economies render investment inherently lumpy; (4) reductions in emissions attributable to reduced reliance on fossil fuels; and (5) insurance against blackouts, brownouts, or longer-period capacity shortages that can be caused by unforeseeable increases in demand, decreases in supply from other sources, or both.[6] These are not necessarily all the benefits that the availability of new plants may yield, but they illustrate the considerable number of attributes that regulators must consider.

Electricity consumers can benefit from all of these possible consequences.

---

5. For a legal and economic discussion of the capacity obligations inherent in the obligation to serve, see J. GREGORY SIDAK & DANIEL F. SPULBER, DEREGULATORY TAKINGS AND THE REGULATORY CONTRACT: THE COMPETITIVE TRANSFORMATION OF NETWORK INDUSTRIES IN THE UNITED STATES 119-29 (1997).

6. The demand curve facing any one firm can be derived by subtracting from the market demand curve the aggregate supply curve of all *n*-1 other firms in the market. *See, e.g.*, DENNIS W. CARLTON & JEFFREY M. PERLOFF, MODERN INDUSTRIAL ORGANIZATION 100 (2d ed. 1994). In this sense, an unforeseeable surge in demand for the output of an individual firm can result from either an increase in market demand (accompanying a summer heat wave, for example) or a decrease in the available capacity of the firm's competitors (as when a competitor shuts down a generation plant for unexpected maintenance or premature decommissioning, for example).

Although at first glance it may appear otherwise, the avoidance of capacity shortages is a benefit not different in principle from a direct financial benefit, such as fuel-cost savings. Consumers clearly benefit if enough additional capacity is provided to reduce the risk of shortages, because shortages harm consumers. Provision against risk is a very tangible product, and in some measure it is bought and sold in a market at prices that are clearly observable. That is precisely the task that the insurance industry performs.[7] Electrical generation capacity that reduces risk frees the electric utility, and ultimately its customers, from the necessity of incurring the costs that would be entailed in those risks. It also frees business firms that are electric customers from incurring the cost of business-interruption insurance against any financial damages to them derived from a power shortage. Each of these burdens has an obvious financial cost whose magnitude can, at least in principle, be estimated.

C. The Benefits of Incremental Investment Creating "Excess" Capacity

The capacity of a plant is usually somewhat flexible and normally can be stretched in a variety of ways, though always at some cost. A plant can be worked overtime if it is not already being used twenty-four hours per day. Delays in maintenance can reduce periods of disuse, at least for a while. Obsolete and uneconomic capacity, previously retired, can be resuscitated. Demand growth may be retarded or interrupted fortuitously. In some or all of these ways, the utility may be able to accommodate a rise in demand that exceeds the optimal capacity utilization of current equipment. Therefore, it may be possible for the utility to postpone the acquisition of new equipment requiring lumpy investment, until demand grows still further and the excess capacity of the new equipment, on the date of its introduction, is somewhat reduced. But the utility can afford to stretch the initial capacity only for a limited time and to a limited degree before the process becomes too expensive to be practical.

Consequently, it often will benefit consumers, financially and otherwise, for the utility to undertake the lumpy investment in additional capacity even before that added capacity appears to be needed. By so doing, the utility can avoid the expense of accelerated construction if the capacity expansion is undertaken at the last minute and the new plant turns out to be needed sooner than anticipated. Early construction can also help insure consumers against the high real costs of power shortages. In rate proceedings, the lumpiness of investment can be at issue when considering the utility's right to recover its cost of financing the construction of new plants, a cost item known in regulatory parlance as "construction work in progress" (CWIP). Despite early resistance to the recovery of such costs, regulators have correctly recognized that "because consumers 'derive a present benefit—assurance of adequate future service—from construction work in progress,' including some of the plant's costs in the rate

7. In reality, of course, there are many risks against which insurance policies are not purchased or available. But that does not mean that such risks are costless. It only means that the persons who have decided to eschew insurance have undertaken to bear that risk themselves. Nevertheless, the insurance market provides an evaluation of the cost of those risks—an evaluation, incidentally, that insurance regulation almost certainly forces lower than the free market price of transfer of the risks from the individual to the insurance firm.

base [does] not conflict with the 'used and useful' principle."[8]

D. Can Excess Capacity Be Used and Useful?

Where an investment is lumpy, one cannot legitimately infer from the existence of excess capacity alone that the investment is not "used and useful." The excess capacity that is characteristic at the time of inauguration of lumpy plant or equipment *is* useful—indeed, it is *in use*. The purpose of that investment may not be to produce output immediately, although that may happen. Rather, the investment is intended to smooth the course of adaptation of plant capacity to the expected intertemporal trajectory of demand and, in that process, to keep down cost to the customer. In that role, the excess capacity *is* currently used. When investment is lumpy, such capacity is not only used and useful; it is an inescapable part of the requirements for efficiency in the investment and production process. A regulator's failure to recognize this role of new capacity can ultimately harm the consumer.[9]

The construction of capacity that appears currently to be excessive, and on that ground seems on the surface not to be used and useful, in actuality provides consumers a valuable option in addition to a possible saving of direct construction or operating cost.[10] In effect, consumers (that is, ratepayers) pay something extra today to avoid having to pay substantially higher prices in the future if a shortage of generation or transmission capacity should eventuate.

In the absence of such excess capacity, consumers would, during a surge in demand or a collapse in supply, face the steeply increasing range of the short-run supply curve. This same supply condition, of course, explains why peak-load pricing of electricity may result in substantial jumps in price when, during the summer season of highest air conditioning consumption, it may be far more expensive for residential customers to run a washing machine or similar appliance during peak hours rather than before or after. In effect, current construction of excess capacity reduces the likelihood that consumers will encounter year-round these costly supply constraints, when utilities must call into service even the least

---

8. Town of Norwood v. FERC, 80 F.3d 526, 532 (D.C. Cir. 1996) (quoting Mid-Tex Elec. Coop. v. FERC, 773 F.2d 327, 346 (D.C. Cir. 1985)).

9. By analogy, patent law recognizes such a meaning of "useful." For an invention to be patentable, it must be novel, nonobvious, and useful. 35 U.S.C. § 101 (2000). To be useful, "a product need not be better than other alternatives or essential to competition," but rather need only be shown to be capable "of serving some identified, beneficial purpose." Vornado Air Circulation Sys., Inc. v. Duracraft Corp., 58 F.3d 1498, 1506-07 (10th Cir. 1995) *vacated on other grounds*. That beneficial purpose may be one that does not lead to immediate consumption of the product. For example, in a patent case involving pharmaceuticals, the Federal Circuit said: "The stage at which an invention in this field becomes useful is well before it is ready to be administered to humans. Were we to require [Food and Drug Administration] Phase II testing in order to prove utility, the associated costs would prevent many companies from obtaining patent protection on promising new inventions, thereby eliminating an incentive to pursue, through research and development, potential cures in many crucial areas such as the treatment of cancer." *In re* Brana, 51 F.3d 1560, 1568 (Fed. Cir. 1995).

10. The option value of capacity in a network industry is well recognized and plays an important role in rate design—notably in two-part tariffs in which the fixed component compensates for the option value of capacity and the variable component compensates for actual capacity usage, as measured in throughput, kilowatts, minute of use, and the like. *See also* JOHN T. WENDERS, THE ECONOMICS OF TELECOMMUNICATIONS: THEORY AND POLICY 46-48 (1987); Paul W. MacAvoy & J. Gregory Sidak, *The Efficient Allocation of Proceeds from a Utility's Sale of Assets*, 23 ENERGY L.J. 233, 237-38 (2001). [hereinafter MACAVOY & SIDAK].

efficient of their generation plants. Although regulators may regard "excess capacity" as a pejorative term, they nonetheless allow recovery of the cost of holding short-term "margin reserve" (to use the more attractive expression).[11] A regulator's failure to consider such benefits from excess capacity will discourage future investments that offer these benefits. If, for example, capacity above that required to meet minimum service standards is not recognized as a benefit in determining whether an investment is used and useful, then underinvestment in reserve capacity will likely follow. Investors will not fund such undertakings, or they will do so only at rates well above competitive norms. Such underinvestment can lead to brownouts and blackouts or even persistent shortfall in supply. Those consequences, if they occur, will not be the fault of the regulated firm, but rather the easily foreseeable result of unreasonable regulatory action.

E. Cost Recovery for Lumpy Investments

In a sense, the usefulness of a lumpy new plant will be different at different dates over the course of its lifetime. In the early stage of its life, a greater share of the lumpy plant's total benefit will consist of its role as insurance against shortages caused by unanticipated surges in demand, unanticipated drops in supply from other sources, or both. Later, if the available capacity of the new plant is absorbed by growth in customer purchases, more of the benefit will be its direct contribution to output.

That description, however, overstates the difference between the two kinds of benefits. Early acquisition of the plant may provide savings to customers, present and future alike, by permitting the entity to keep down the cost of the investment. It is reasonable to expect that new plants will employ newer and more efficient designs than plants already in operation. Consequently, one would expect the new plant or equipment to be fully used early, with the excess capacity taking the form of withdrawal of some of the utility's older plants from service. The result is an immediate benefit to current customers in the form of reduced current production costs. Thus, lumpy investments are likely to yield different quantities and mixes of benefits at different stages in their lifetimes. It is incorrect, however, to infer that the bulk of their benefits will accrue to future customers alone.

In recovering the cost of a lumpy plant over its lifetime, the payments should be timed as they are in any competitive market. Thus, the sum of the revenues over the lifetime of the investment should be sufficient to cover all costs, including replacement of the investment when the time arrives, and the cost of the capital tied

11. *See generally* Application for Rate Increase in Pinellas County by Mid-County Services, Inc., Dkt. No. 971065-SU, Order No. PSC-99-1912-FOF-SU, 1999 Fla. PUC LEXIS 1733, at 27 (Fla. Pub. Serv. Comm'n 1999) ("Margin reserve allows a utility to expand prudently beyond current demands to enable it to meet reasonable projected short term growth. This practice allows the utility to include a reasonable cost of expansion in its rate base without placing an unreasonable burden on current customers to pay for long term growth."); Matter of the Petition of Hoosier Energy Rural Electric Cooperative, Inc. to Increase Its Rates and Charges, Cause No. 37294-J, 1985 Ind. PUC LEXIS 204, at 32 (Pub. Serv. Comm'n Ind. 1985) ("an excess capacity case . . . would involve the determination of an appropriate reserve margin and the potential exclusion of investment related to any excess reserve capacity from the 'used and useful' rate base").

up in the investment during its lifetime. This fundamental relationship means that the discounted present value of these revenues must constitute a sum equal to the discounted present value of the costs. The timing of the realization of these revenues, however, cannot be determined definitively by the regulatory agency—or by the courts or the firm's management, for that matter. The timing ultimately is affected, if not entirely determined, by the state of the market at different periods during the lifetime of the investment.

Regulation emulates competitive markets when it works to ensure that the regulated firm's prices replicate those that would emerge in an effectively competitive market. This axiom has two implications for repaying the costs of a lumpy asset. First, regulators should take into account the entire lifetime of the lumpy investment in determining whether it is appropriate to cover its cost or in determining the proportion of its cost that is appropriate to cover. Second, regulators should handle the timing of the revenues from which that cost is to be recovered with as much flexibility as possible. By doing so, they will permit the utility to adapt to the changing and unforeseeable course of market conditions over the lifetime of the investment. Thus, subject to the usual limitations on rate of return, price caps, and similar constraints, the regulated firm with lumpy investments should be allowed the opportunity to earn on them currently.

A critical issue here is the choice of time horizon to be used in the process of determining whether a lumpy investment is used and useful. Rational evaluation of that issue must take into account the entire lifetime of the investment. Understandably, it is tempting to proceed in a very different manner, dividing that lifetime into different subperiods, distinguishing those in which there is excess capacity from those in which the lumpy investment is fully consumed and then taking the position that during times of excess capacity the investment is not used and useful. While intertemporal variations in demand pressures can cause competitive prices to parallel such a path roughly, the approach is indefensible for several reasons. First, it ignores the considerable uses that available "excess capacity" can serve. Those uses include its potential ability to cut current production costs and its role as insurance against shortages in output when there is an unforeseeable surge in demand, collapse in supply from alternative sources, or both. Second, that approach ignores the fact that, when investment is lumpy, excess capacity is an inescapable part of the production process over time.

The third and most fundamental objection to the subperiod approach illustrates the view that an investment with excess capacity is one that is neither used nor useful. The prototypical history of a lumpy investment entails a large share of excess capacity at the date of its introduction. As demand grows over time, that excess capacity gradually shrinks. At the moment it disappears altogether, however, yet another such lumpy facility may be brought on line—and the excess capacity appears all over again. In other words, the typical history of lumpy investment is one in which so-called *excess capacity is almost never absent.* If that fact is distorted to mean that the total investment is never fully used and useful, the implication is that full recovery of the cost of a lumpy investment should never be permitted by regulation. Such a view would ensure that lumpy investments would never be undertaken and that consumers would be

deprived of the availability of such clearly useful investments.

III. FORESIGHT AND HINDSIGHT MODELS OF COST RECOVERY

The competitive market rewards the investor on the basis of the outcome of the investment process. For example, if there proves to be substantial demand for the product of the investment, the market will reward the investor handsomely. It will do so without regard to whether that felicitous outcome is fortuitous or the product of the investors efforts and competence. It will do so without regard to whether the initial investment decision was wise and the product of superior foresight or whether that decision was highly imprudent but luckily was redeemed by unanticipated developments. Thus, the market always rewards on the basis of hindsight. It gives no credit for decisions that were superior at the time they were made, but which turned out badly through no fault of the decision maker.

A. Biases Inherent in Regulatory Hindsight Criteria

The used-and-useful test also is a hindsight requirement. The regulator asks after the fact whether the investment worked out as might have been hoped. This characteristic distinguishes the used-and-useful test from one that might be called the foresight criterion, an example of which is the prudent-investment test. The classic articulation of the prudent- investment test is often credited to Justice Brandeis in his 1923 concurrence in *Missouri ex rel. Southwestern Bell Tel. Co. v. Pub. Serv. Comm'n.*[12] He wrote that a utility should receive from regulators the opportunity to earn "a fair return on the amount prudently invested" to provide service.[13] Similarly, in *Duquesne,* Justice Scalia, in his concurrence, described prudent investment as the "capital reasonably expended to meet the utility's legal obligation to assure adequate service."[14] The prudent-investment test has come to stand for the proposition that a utility may recover its investment, through allowed rates, if its investment decision was prudent in light of the information reasonably available to the utility when that decision was made.[15]

From the standpoint of public policy, neither the foresight nor hindsight test is clearly superior so long as only one is used and used consistently. The one clear difference between the two is the choice of the party that is to bear the risk of the investment in question. A hindsight test implies that the regulated company and its shareholders bear the risk. If subsequent events go well, whether by accident or design, the firm's investors benefit. If matters work out badly, the investors suffer

12. 262 U.S. 276, 289 (1923) (Brandeis, J., concurring).

13. *Id.*

14. Duquesne Light Co. v. Barasch, 488 U.S. at 317 (1989) (Scalia, J., concurring).

15. *See generally* Pierce, *supra* note 2, at 511. Professor Pierce observes that such investment decisions are "rarely blatantly imprudent when viewed in light of the knowledge and alternatives reasonably available to the utility's management at the time of the decision." *Id.* at 512. For a representative example of a regulator's application of the prudent-investment test, see In re Cent. Ill. Light Co., 57 P.U.R.4th 351, 358-62 (Ill. Commerce Comm'n 1983).

the consequences (and ratepayers avoid having the loss passed through in rates).[16]

A foresight test, in contrast, transfers the risk to the customers of the regulated company. The firm will be rewarded if it made the choice that was prudent at the time, no matter how later events develop. This distinction would make it appear that a foresight test is better for the firm and worse for its customers, but that is not so. If regulation is to select the minimum rates necessary to elicit the level of investment that best serves the long-run interests of consumers, a regulated company must be compensated *more* if regulation requires it to assume the risks. The entity must be permitted to earn a rate of return higher than it is allowed in a regime in which those risks are borne by customers. Risk is a real cost, and someone must pay for it. Thus, a hindsight test can lead to a higher cost of capital that eventually must result in higher rates for consumers.

A mixture of the hindsight and foresight tests is the worst of all worlds. It prevents the entity from earning adequate revenues by disallowing recovery whenever past errors of judgment occurred. And it prevents earnings above the cost of capital when performance is outstanding. The Supreme Court has recognized this problem of asymmetric treatment of risk when it said in the *Duquesne* case that "a State's decision to arbitrarily switch back and forth between methodologies in a way which required investors to bear the risk of bad investments at some times while denying them the benefit of good investments at others would raise serious constitutional questions."[17]

Two factors, however, bear on whether a hindsight or foresight test should be chosen. First, in both regulated and unregulated markets, some entities are more efficient risk bearers than others. Typically, a life insurance company is a more efficient bearer of the risk of premature death than the head of a middle or lower income family. That is why people purchase life insurance. The buyer of a life insurance policy reduces the real cost that he bears by transferring the risk to the more efficient risk bearer. Thus, the rational basis for choosing between a foresight test and a hindsight test is the evidence on whether the utility or its customers are the more efficient risk bearers. If a hindsight test is selected, so that the firm is required to bear the risk, it must be compensated for carrying out this task through a suitable addition to the rate of return that it is allowed. Nevertheless, the payment of this risk premium may be beneficial to all parties if the regulated firm and its investors are the more efficient bearers of the risks.

A second, more pressing distinction between the hindsight and foresight approaches arises from the nature of the regulatory process. Under regulation, a hindsight test is biased toward providing the regulated firm with inadequate earnings on average. In a competitive market, in contrast, the process is neutral

16. A similar question of efficient risk bearing arises when a utility proposes the sale of certain assets that have risen or fallen substantially in value since their acquisition. The question then is, how should regulators allocate those gains or losses among ratepayers and shareholders? *See also* MACAVOY & SIDAK, *supra* note 10. Even under the hindsight model, ratepayers rather than shareholders bear the risk of changes (either positive or negative) in the value of the utility's investments made to provide service to the public. *Id.* Stated differently, ratepayers bear the risk that regulatory changes rather than market changes (such as changes in demand, technology, or management) cause the utility's investment no longer to be used and useful.

17. *Duquesne*, 488 U.S. at 315.

because it creates no presumption that the firm can expect to earn either more or less than the risk-adjusted cost of capital. Whether the market is regulated or unregulated, events following a prudent investment decision can turn out to be favorable or unfavorable to the investing enterprise. That, of course, is the nature of risk. In an unregulated competitive market, if matters turn out unfavorably, the firm will suffer a loss. But if they go well, the firm will earn a relatively large profit—one that exceeds the cost of capital. If the percentage returns from good and bad outcomes are randomly distributed around the mean return, then, as time passes, the firm's gains and losses will, as a matter of basic probability theory, average out to the competitive return on capital.

Under regulation that uses a hindsight test, however, the firm cannot expect to be protected from losses if post-investment developments prove unfavorable. The firm can, under a strict regime of rate-of-return regulation, however, expect regulation to rule out supracompetitive profits on investments for which matters go favorably. Rather, profits will be limited to something like the cost-of-capital return on even those felicitous investments. Now, however, the average of the earnings associated with a stream of favorable and unfavorable events is no longer neutral. The *average* of a loss and a return equal to the cost of capital *must* be a return lower than the cost of capital. The incentive for appropriate investment will be undermined and the cost of capital increased to compensate investors for this bias, at the expense of the customers of the regulated firm in the long run.

Regulation can correct the effects of bias implicit in hindsight approaches to regulation, but that correction has its price. To restore the requisite incentive for the investments that long-run ratepayer interests require, the regulated firm must be permitted to earn a rate of return sufficiently above the competitive level that would be sustainable in the absence of bias.[18] This higher rate will, of course, come from the pockets of ratepayers. It is not a price that they must pay because of a risk that is entailed naturally in the activities of the regulated firm. Rather, it is an expense they must incur to offset the unfortunate effects of hindsight-based regulation.

B. Reasonable Returns Under A Hindsight Model

The competitive market provides guidance for determining a fair rate of return on capital generally and a reasonable return on equity under a hindsight regulatory regime. The test that determines whether a particular return on equity is "reasonable" is whether or not it is adequate to elicit investment funding for the

18. This point is not fundamentally different when the utility is subject to price-cap regulation rather than cost-of-service regulation. The regulator must set an initial price under the price-cap formula. That price, set not higher than the stand-alone cost of the service, must be calculated using cost data, including cost-of-capital data, of the sort used in a traditional rate case under cost-of-service regulation. *See generally* WILLIAM J. BAUMOL & J. GREGORY SIDAK, TOWARD COMPETITION IN LOCAL TELEPHONY 89 (1994). The advantages of price-cap regulation over cost-of-service regulation surely do not lie in the calculation of the initial price in the price-cap formula, but rather in the superior incentives for productivity improvements as the price cap automatically adjusts (usually) downward over time. *See generally* DAVID E.M. SAPPINGTON & DENNIS L. WEISMAN, DESIGNING INCENTIVE REGULATION FOR THE TELECOMMUNICATIONS INDUSTRY (1996).

particular entity.[19] Even if the regulator were, through miscalculation or for some other reason, to arrive at a permitted rate of return well below the true value of the reasonable rate of return on equity, *new* equity investors will still obtain such a reasonable return, because they will refuse to provide any financing to the firm or to purchase the firm's stock in the secondary market until the price of its stocks fall to the point where the rate of return that investors can expect to receive on the stocks at their bargain basement prices attains the reasonable level. The rate of return *must* reach that level. If it is any less, investors will send their money elsewhere—to unregulated and healthy competitive markets that do offer a reasonable return on equity. Thus, the actual rate of return on the actual market price of the regulated company will indicate the reasonable return on equity figure, because the forces of supply and demand will automatically adjust stock prices to yield that return.

The ratepayer is thus a victim of regulatory decisions not to permit the return on equity that is reasonable under a hindsight regulatory regime, although the interests of earlier purchasers of equity also are damaged. The reason that earlier investors are hurt is obvious. They suffer a loss in the market value of their stock. The damage to consumer interests is more subtle and perhaps more important.

When the firm borrows, it will have to pay a higher rate on its new debt. A forced reduction in the price of company securities means that the company must pay a higher real price for its new capital. Every share of equity in effect gives its owner proprietorship of a portion of the company's assets. If the company's share price falls 50 percent, a purchase of $1000 in stock will now give the new stockholder twice as large a share of the company's assets as it did at the earlier and higher stock price. Thus a fall in stock price is, in effect, a means to disguise a rise in the amount that the firm has to pay for its capital. Ultimately, the burden will fall on consumers who will be forced, because of the regulatory action, to pay higher prices for poorer service as the company is forced to reduce the amount of investment funding it acquires. Because it acquires funding by offering, in effect, a share of its assets in return, the company will be able to raise correspondingly less capital when the market value of those assets collapses through a fall in the price of the company's stock.

The net effect is intentional or unintentional victimization by the regulatory process not only of the preexisting shareowners, but also of consumers. Ironically, this harm is inflicted in a way that gives it the appearance of contributing to consumer welfare by keeping investors as a body from obtaining a reasonable level of earnings. Thus, the belief that consumers benefit thereby at the expense of investors is an illusion. What consumers may gain today they will pay a high price for tomorrow.

This analysis is not affected, moreover, by the fact that investors may have received information through analyst reports, company disclosure statements, and other materials indicating that used-and-useful disallowances were a possible

19. *See generally Duquesne*, 488 U.S. at 312. The rate of return for regulated utilities should not be so low as to "jeopardize the financial integrity of the companies, either by leaving them insufficient operating capital or by impeding their ability to raise future capital." *Id.*

risk. The market tends to ensure that new investors receive a reasonable rate of return on their equity no matter what is done. Old investors, however, will be benefited or harmed as stock prices are driven upward or downward by changes in the direction of regulation or by the belief that such changes will occur. If investors come to believe that regulation by direct or indirect means will cut the allowed rate of return on equity below the reasonable level, stock prices will fall and old equity holders will be damaged. If that forecast is not borne out by subsequent events, the people who bought stock at its low price will benefit when the later developments restore the stock price.

The fortuitous gains of that one set of stockholders, however, like the fortuitous losses of the other, are not the central issue for the welfare of the regulated company's customers. Their long-term interest is promoted by getting the right figure for the reasonable return on equity. If the regulated return on equity is set too high, consumers will reward equity holders too generously. If that figure is set too low, the firm's future investment will be condemned to inadequacy—to levels below those called for by the interests of consumers.

IV. AN ILLUSTRATION: THE CALIFORNIA ELECTRICITY CRISIS

We come now to the California electricity crisis of 2000-01, a daunting subject that already is spawning volumes of theoretical and empirical analysis in regulatory law and economics. The precise cause of the crisis is a question that we leave to others better suited to the task of supplying an answer.[20] Our purpose here is, instead, to highlight a key aspect of the episode upon two features of which there is evidently no disagreement—that a severe capacity shortage occurred and that it was unexpected by regulators. In such circumstances, virtually all generation capacity in California was truly and evidently used and useful.

California restructured but did not truly (or fully) deregulate its electricity market. As Professor Paul Joskow has observed, the motivation for California's restructuring of its electricity market was partly a concern over costly or excessive generation capacity:

> In early 1993, when the California Public Utilities Commission began a comprehensive review of the state's electricity industry—then running under the old pattern of regulated utility monopolies—the state was under pressure from industrial consumers to reduce prices that were among the highest in the nation. The high prices were attributed to the utilities' *costly nuclear power plants, expensive long-term contracts with independent power suppliers and excess generating capacity*, and to inefficient regulation. The commission's vision for reform,

20. *See generally* PAUL L. JOSKOW & EDWARD KAHN, A QUANTITATIVE ANALYSIS OF PRICING BEHAVIOR IN CALIFORNIA'S WHOLESALE ELECTRICITY MARKET DURING SUMMER 2000 (AEI-Brookings Joint Ctr. for Regulatory Studies, Working Paper 01-01, 2001); Severin Borenstein, *The Trouble with Electric Markets and How They Derailed California's Electricity Restructuring*, 16 J. ECON. PERSP. 191 (2002); SEVERIN BORENSTEIN, ELECTRICITY PRICING SHOULD CLUE CONSUMERS TO JUDICIOUS USE (AEI-Brookings Joint Ctr. for Regulatory Studies, Policy Matters 01-04, 2001); FRANK A. WOLAK, WHAT WENT WRONG WITH CALIFORNIA'S RESTRUCTURED ELECTRICITY MARKET? (Nov. 2000) *available at* http://siepr.stanford.edu/papers/briefs/policybrief_nov00.pdf; Frank A. Wolak, Proposed Market Monitoring and Mitigation Plan for California Electricity Market (Feb. 2001), unpublished manuscript, *at* http://stanford.edu/~wolak/ [hereinafter Wolak].

articulated in 1994, was built around a new industry structure.

> The generation of electricity by existing plants and the entry of new plants would no longer be regulated by the state, and their power would be sold in a new, competitive wholesale market. Homeowners, factories and businesses would have the choice of using the transmission and distribution wires of their old local utility companies for "direct" access to the power sold in these new wholesale markets or continuing to buy power from the utilities themselves.[21]

The legislative process, however, produced a market that not only had complex institutions and rules, but also seriously distorted economic incentives, as Professor Joskow notes:

> Then came four years of legislative, administrative and public debate. California's new electricity market ended up being designed in a highly politicized process, heavily influenced by people with little knowledge of the business and by middlemen who stood to benefit from an inefficient market. What eventually emerged was *the most complicated electricity market ever created, with many features that had never been tried.*[22]

As part of the restructuring, the state's utilities sold their generation facilities and accepted a price cap on retail rates, in the belief that a competitive generation market would lead to lower input prices. Meanwhile, the state deregulated the wholesale power market and thus, allowed the wholesale price of electricity to fluctuate with changes in supply and demand conditions. When all electricity in the retail market was priced at the level of the cap, this arrangement clearly eliminated any incentive for retail consumers to search for cheaper power in the competitive generation market. California also prohibited utilities from entering into long-term, forward contracts for power. In addition, over many years, there had been resistance on environmental grounds to the construction of new generation capacity in California. Yet the restructuring of the California market did nothing to add generation capacity or expedite the siting process for new generators.

The restructured market began operating in April 1998. By the summer of 2000, natural gas prices began to rise sharply following a period of low snowfall and rainfall, which reduced the supply of hydroelectric generation. Simultaneously, a number of California's oldest generation plants were shut down for maintenance, and the price of emission permits in California, which are necessary to undertake generation of electricity, rose. Finally, it is the view of some economists that certain wholesale suppliers of electricity exercised market power by withholding generation output.

This combination of forces produced the "perfect storm." The spot price of wholesale electricity in California rose sharply, far above the utilities' fixed retail prices.[23] California's regulators and politicians did not allow the retail price to

---

21. Paul L. Joskow, Editorial, *California Can Tame Its Crisis,* N.Y. TIMES, Jan. 13, 2001, at A13 (emphasis added). [hereinafter Joskow].

22. *Id.* (emphasis added).

23. Professor Frank Wolak of Stanford University wrote in early 2001: "Average wholesale rates reached, by far, their highest level in 2000, during December. The average wholesale cost of electricity and ancillary services in California was 32 cents per kilowatt-hour (kWh) of load. The implied wholesale price of energy and ancillary services in the frozen retail rates of the three [investor-owned utilities] IOUs

rise correspondingly. In the words of Professor Frank Wolak of Stanford University, who chairs the Market Surveillance Committee of the California Independent System Operator, a pass-through of the cost increases to retail rates "would almost certainly cripple the California economy."[24] Instead, California's regulators and politicians shielded the state's consumers from the price fluctuations at the wholesale level, which in turn deprived consumers of any incentive to conserve or to defer their electricity consumption to off-peak periods. Rolling blackouts followed. The obligation of California's utilities to sell retail power at a fixed retail price, which by then happened to be *below the wholesale price in the spot market*, caused enormous losses for the state's utilities and put them in danger of bankruptcy. This risk of bankruptcy made wholesale suppliers of power in other states (over which California had no regulatory jurisdiction) reluctant to sell power to the California utilities, exacerbating the state's shortage. Governor Gray Davis petitioned the Federal Energy Regulatory Commission (FERC) to issue a federal regulation compelling out-of-state generators to sell power to California, but the FERC declined. California then enacted legislation to subsidize retail consumption of electricity through the issue of long-term state bonds. Thus, future taxpayers would pay some of the costs of California's current energy consumption and its failed experiment in electricity restructuring. Current shareholders (and perhaps bondholders as well) would also bear some of that cost, as Pacific Gas and Electric filed for Chapter 11 bankruptcy in April 2001, just three years into the restructuring.[25]

Although economists differ in their assessments of the proximate cause of the California crisis and in their prescriptions for avoiding its recurrence, one point of agreement among all of them is that the supply shortages and accompanying wholesale price spikes were completely unexpected by regulators. Professor Joskow has offered the following prescription:

> The answer for California now is not to return to the old, costly system of regulated monopolies, but to apply the harsh lessons it has learned from designing a flawed system. Competitive electricity markets will not work if consumers are completely insulated from wholesale market prices. Long-term contracts can protect consumers from volatile prices and price manipulation by suppliers. *New generating plants must be built.*[26]

California believed that the cost per kilowatt hour of the generation capacity of the state's utilities would exceed the wholesale price of power in a competitive market. That belief may have been reasonable—in a word, prudent—at the time that it motivated California's legislature and regulators to decide to restructure in

---

is between 6 cents/kWh and 6.5 cents/kWh, depending on the IOU. If these December 2000 wholesale costs had been passed through in retail rates, this would have caused rate increases of more than 300%." Wolak, *supra* note 20, at 3.

24. *Id.* at 6.

25. Pacific Gas and Electric Company, Voluntary Petition No. 30923 SMF11 (Bankr. N.D. Cal., Apr. 6, 2001), *available at* http://www.pge.com/006_news/current_issues /reorganization/ court_docs/pdf/ 00000001 .pdf.

26. Joskow, *supra* note 21 (emphasis added).

the manner than they did. In hindsight, however, the belief was clearly incorrect in light of the market structure that California created.

This error in foresight on the part of California's architects of electricity restructuring is relevant to the proper standards for determining whether an investment, particularly one seemingly contributing "excess" generation capacity, is really used and useful. A regulator's view of whether the investment in a particular asset is used and useful is likely to be limited by personal experience and the institutional memory of the regulatory body. To borrow from the language of statistics, that personal experience and institutional memory aids the regulator in making in-sample predictions of whether consumers will, over some relevant period of time, exercise the option inherent in the utility's excess capacity. That experience, however, provides little if any guidance with respect to out-of-sample market conditions, such as the California electricity crisis. Yet it is especially for such outlier events that insurance confers its greatest advantage to the insured.

V. CONCLUSION

The California electricity crisis of 2000-01 is a lesson to other jurisdictions that notions of the lack of usefulness of an investment can be distorted by the limits of one's personal experience and the limits of the institutional memory of a regulatory body. Only a few years earlier, few if any consumers or regulators of electricity in the United States would have considered it likely that California energy users would face shortages associated with spiking prices in the wholesale power market. Investments in "excess" generation and transmission capacity might have appeared before the summer of 2000 not to be used and useful any longer in a restructured electricity market. Yet events suddenly proved excess capacity to be currently and unquestionably used and useful. A utility's investment in seemingly "excess" capacity provides an immediate option to consumers, an option having substantial economic value if demand unexpectedly surges, supply unexpectedly collapses, or both occur simultaneously. That option is analogous to insurance. It is especially true for an outlier event like the California electricity crisis that insurance confers its greatest advantage upon the insured that are the very consumers whom public utility regulation exists to protect.