Electricity Capacity Requirements: Who Pays?

Timothy J. Brennan

August 2003 • Discussion Paper 03–39

Resources for the Future
1616 P Street, NW
Washington, D.C. 20036
Telephone: 202–328–5000
Fax: 202–939–3460
Internet: http://www.rff.org

Electricity Capacity Requirements: Who Pays?

Timothy J. Brennan

Abstract

Reserve requirements in electricity markets may get each producer to internalize the cost of grid-wide blackouts it might cause if unable to meet consumer demand. Markets for how such capacity might be procured have been studied. Less examined is how the costs of reserve capacity are covered. "Who pays" depends on how requirements are designed. If each producer has to provide peak capacity available to a grid operator at a below-spot price, requirements will increase volatility—that is, the gap between baseload and marginal peak prices. Requirements based on energy sales act as a tax on baseload to subsidize peak, reducing volatility. Finally, if requirements are designed to ensure that extreme-peak energy is available without scarcity rents, baseload prices remain unaffected, but (nonextreme) peak prices increase. Although this pattern seems unrelated to any economic or social goal, it replicates what one might see under crude seasonal or time-of-use pricing.

Key Words: capacity requirements, reserve requirements, electricity generation, utility regulation

JEL Classification Numbers: L94, L51, H22

Contents

Electricity Capacity Requirements: Who Pays?

Timothy J. Brennan*

Introduction

A fundamental design feature of electricity markets, in both the franchise monopoly and the open market eras, is the requirement to construct and maintain extra capacity for generating electricity.[1] In most businesses, firms may retain excess capacity, in either production itself or in inventory of already-produced goods, for a variety of reasons: to dissuade current customers from going elsewhere if demand is unexpectedly high, serve new customers if other suppliers fail to meet delivery timetables, ensure the ability to meet contractual commitments to supply a customer, or smooth out prices in the face of predictable swings in demand. But however important these justifications may be in any particular industry, they are rarely the focus of government regulation. Although courts can impose penalties for failures to meet contractual obligations, we do not typically have laws mandating extra seating in restaurants or keeping a full range of shoe sizes in the stockroom.

Consequently, at first blush, capacity requirements lead to inefficient pricing. Were potential unreliability the only problem, coping with outage risk could be internalized to the supplier. If customers wanted to smooth out payments, they could do so through borrowing during off-peak periods to cover costs during peak seasons. With uncertainty, consumers might

---

* Professor, Policy Sciences and Economics, University of Maryland, Baltimore County, and Senior Fellow, Resources for the Future. Email: brennan@umbc.edu. Thanks for helpful comments go to Jeff Church, Joseph Doucet, Douglas Hale, Andrew Kleit, Karen Palmer, and Jamie Wimberly. I also received numerous useful comments on an earlier version of this paper from those attending the Rutgers University Center for Research in Regulated Industries 16th Annual Western Conference, Advanced Workshop in Regulation and Competition, especially Hung-Po Chao, Robert Levin, Steven Ostrover, Cliff Rochlin, and Gary Stern. Special thanks go to Steven Stoft for maintaining links to useful materials on his website (www.stoft.com). Errors and omissions remain the author's sole responsibility.

[1] Federal Energy Regulatory Commission (FERC), *Remedying Undue Discrimination Through Open Access Transmission Service and Standard Electricity Market Design*, Docket No. RM01-12-000 (2002), at 252–90; Joskow, Paul, Comments before FERC, *Remedying Undue Discrimination Through Open Access Transmission Service and Standard Electricity Market Design*, Docket No. RM01-12-000, Jan. 10, 2003, at 14–48. FERC recently modified its proposal to eliminate specific nationwide reserve requirements, leaving it to states within regional transmission organizations. FERC, "White Paper: Wholesale Power Market Platform," at 11.

want to insure against high peak prices, but that preference could be satisfied directly through contract.[2]

The rationale for such requirements, seemingly inefficient or superfluous in most markets, is easy to address in theory, although using it to quantitatively set reserve margins remains difficult. As has been pointed out elsewhere,[3] electricity differs from other commodities in fundamental ways that could justify, among other things, capacity requirements. In addition to being an important commodity, electricity is both vulnerable to supply-demand imbalances (because storage costs are prohibitive) and interconnected (because switching costs are prohibitive). For that reason, if one supplier fails to meet the demands of its customers, not only will those customers lose service, but also all customers sharing the distribution grid will lose power.

The need to maintain balanced loads over an interconnected grid creates what we might call the "blackout externality."[4] As one observer puts it, "everyone is denied service in a shortage situation."[5] Accordingly, investments to maintain reliability and prevent demand from exceeding capacity have benefits to the public at large. The blackout externality can justify policies to encourage broader use of real-time metering, for example.[6] Regulatory policies that prevent retail

---

[2] Note that either with smoothing certain payment or insuring against peak prices, it would be more efficient, in principle, to divorce the borrowing or insurance premium from the price of electricity. Failures in capital or insurance markets thus could warrant policies that lead power companies to install extra capacity as a second-best response, but we ignore such possibilities here.

[3] Brennan, Timothy, Karen Palmer, and Salvador Martinez, *Alternating Currents* (Washington, DC: Resources for the Future, 2002); Steven Stoft, *Power System Economics* (New York: Wiley, 2002).

[4] Others call this a "free rider network externality" or "capacity externality." Ruff, Larry, "Competitive Electricity Markets: Why They Are Working and How to Improve Them," n/e/r/a, May 12, 1999, at 28; Jaffe, Adam, and Frank Felder, "Should Electricity Markets Have a Capacity Requirement? If So, How Should It Be Priced?" *Electricity Journal* (December 1996): 52–60.

[5] Duckworth, Jack, "The Fatal Flaw in Electric Power Deregulation," *Energy Pulse* (www.energypulse.net), accessed 2/18/03; Stern, Gary, "Resource Adequacy in Competitive Electricity Markets," presented at the Rutgers University Center for Research in Regulated Industries 16th Annual Western Conference, Advanced Workshop in Regulation and Competition, June 27, 2003.

[6] Brennan, Timothy, "Market Failures in Real-Time Pricing: A Theoretical Look," Resources for the Future Discussion Paper 02-53, October 2002. Severin Borenstein ("Frequently Asked Questions about Implementing Real-Time Pricing in California for Summer 2001," University of California Research Institute, March 2001) and Steven Stoft ("The Market Flaw California Overlooked," *New York Times*, Jan. 2, 2001, A19) cite inadequate real-time pricing in explaining the poor performance of some electricity markets. Absent the blackout externality, cost savings and intertemporal efficiency improvements in and of themselves need not justify policy to subsidize real-time meters. Their arguments might be better thought of as objections to regulatory policies that prevented customers from any exposure to real-time electricity costs.

or wholesale prices from rising high enough to conform to peak-period power costs can also justify policies to increase supply or limit demand, such as demand-side management subsidies of more efficient heating and cooling systems.[7] Capacity requirements can substitute for regulations or inadequate demand response that prevent prices from clearing electricity markets.

Not all industry observers believe that a blackout externality exists. They argue that if demand for electricity were responsive to prices, there would be no need for capacity requirements.[8] If demand response were sufficiently great to prevent any blackouts from occurring, the externality would in fact disappear. But in most cases involving risk, reducing the likelihood of an adverse outcome to zero is prohibitively expensive; tolerating some risk is efficient. If this general rule holds for electricity outages, a blackout externality would justify some policy intervention, be it capacity requirements or subsidies for real-time meters or other technologies that reduce consumption during periods of high demand.

Although the rationale for capacity requirements is widely appreciated, little analytical attention has been devoted to how the cost of capacity requirements is borne. Understanding how they work presents at least two difficulties. First is ascertaining what a reserve margin means. Must one have excess capacity whenever one operates? That would invite an obvious inefficiency. Why would one require capacity and then not use it? Effectively taxing capacity through reserve requirements just to pay for more capacity would seem to offer no effect on marginal capacity costs, and thus no effect on quantity of capacity or prices of electricity.

Second, the term "reserve margin" suggests that it is a call option provided to the grid manager to call upon as needed. If so, we need to know the strike price for the call option—that is, the price at which the energy from this reserve capacity would be sold to the grid operator or at its request. If the strike price is just the spot price, the requirement is meaningless; capacity could be procured in real time without any requirement. If the strike price is less than the spot price, who covers the cost of the call option? A related third point is determining what counts as eligible reserve capacity. If any called-upon capacity has to be paid the cost of actually bringing it online, those under the requirement can choose technologies with very low capital cost, down to hamsters in wheels.

---

[7] Brennan, Timothy, "Demand-Side Management Programs under Retail Electricity Competition," Resources for the Future Discussion Paper 99-02, October 1998.

[8] Fraser, Hamesh, "A Critique of the Resource Adequacy Requirement of FERC's SMD NOPR," *Electricity Journal* 16 (April 2003): 23–28.

To improve our sense of how capacity requirements play out, we assume perfect foresight in the quantity of power demanded during two periods, on and off-peak, and the fraction of time or probability at which demand is at peak levels. This assumption is less problematic than it might appear. Variations in demand induce considerable price *volatility*, but that is not the same thing as price *uncertainty*. Electricity prices can be much higher on hot August afternoons than on mild April mornings, and consumers may have some interest in smoothing out payments to avoid the summer spike. But these variations may be quite predictable. The more significant stochastic factor in electricity markets is probably not on the demand side but on the supply side—that a generation unit may go down unexpectedly. An outage of that kind is the sort of thing against which consumers (or liable load-serving entities) might want to insure through paying for capacity requirements. The magnitude of that risk, and the size of the blackout externality, would determine how much extra capacity should be put in place to mitigate against such risk.

As noted earlier, it is beyond the scope of this paper to estimate what the reserve margins should be. We also omit discussion of how a reserve requirement might be implemented and reallocated through capacity markets, and related technical engineering issues.[9] However, the question remains of who pays for the insurance provided by these reserve requirements. To answer the "who pays" question rather than the "how much" question, we need not get into the specifics of the uncertainty that underlies the requirements in the first place. The question before us is the incidence of the costs of a reserve requirement, not whether they exceed the benefits. Incorporating a specific model of uncertainty would not provide any additional insight into understanding who pays; we can infer from the underlying uncertainty whether the incidence of payments for capacity requirements is likely to reduce volatility or reproduce a stream of payments that consumers may prefer.

To understand who pays, we specify two generation technologies, peak and off-peak; later we add a third "extreme peak" technology. We also adopt three modeling approaches. After

---

[9] For more on those, see Stoft, n. 3 *supra* at Part 2; Hirst, Eric, and Brendan Kirby, "Technical and Market Issues for Operating Reserves," Oct. 19, 1998, available at http://www.ornl.gov/ORNL/BTC/Restructuring/Operating_Reserves.pdf; Cramton, Peter, "Review of the Reserves and Operable Capacity Markets: New England's Experience in the First Four Months," Nov. 17, 1999, available at www.cramton.umd.edu; Hobbs, Benjamin, Javier Iñon, and Steven Stoft, "Installed Capacity Requirements and Price Caps: Oil on the Water or Fuel on the Fire?" *Electricity Journal* 14 (July 2001): 23–34; Shankar, Roy, "Is an Adequacy Market Necessary?" Electric Power Generation Association presentation, Oct. 16, 2002.

setting a benchmark, we first portray requirements as a mandate that all suppliers allow the grid manager to purchase energy from some fraction of the installed capacity at a preset strike price during peak periods. Baseload prices would remain unchanged, but peak prices would rise for those who cannot purchase energy at the strike price. *If capacity requirements are based on installed capacity, they are thus likely to increase volatility—that is, the difference between peak and off-peak prices.*

Under a second approach, reserve requirements are based on production rather than on installed capacity. In that scenario, requirements create a tax on electricity use to subsidize peaking capacity, raising baseload rates and reducing peak-period prices. To help untangle the effects of the requirements, it can be helpful to treat reserve requirements not in quantity terms but in price terms, such as a subsidy to encourage more peak capacity, where the tax is paid per unit of output. This can be translated back into a quantity-based capacity target, where one specifies a tax sufficient to lead to the desired amount of additional capacity.[10] With that, the first of these may be partly translated back into a "fractional increase of capacity/reserve margin" requirement. In this setting, volatility decreases; *peak prices fall from the capacity subsidy, and baseload prices rise to pay for it—a pattern consumers might prefer because it smoothes out payments*. This suggests that the obligation to provide capacity should be based not on capacity itself but on energy production, imposing a higher requirement on baseload plants.

Finally, we examine a setting in which electricity demand can take on normal baseload, higher peak, and infrequent but very high extreme values with different probabilities. In this setting, we follow analyses of electricity markets that presume that electricity should be priced at the average variable cost of the marginal supplier.[11] All baseload and peak producers install sufficient extra extreme capacity (as a fraction of their normal capacity) to meet such demand at the variable costs of that capacity.[12] *We show that this requirement leaves baseload prices unaffected but raises peak prices (if it does not drive peak producers out of business).* It is

---

[10] Jaffe and Felder, n. 4 *supra*, discuss how uncertainty can affect the equivalence of price-based and quantity-based capacity requirements in electricity.

[11] Borenstein, Severin, James Bushnell, and Frank Wolak, "Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market," *American Economic Review* 92 (2002): 1376–1405; Joskow, Paul, and Edward Kahn, "A Quantitative Analysis of Pricing Behavior in California's Wholesale Electricity Market during Summer, 2002" *Energy Journal* 23 (2002): 1–35; Federal Energy Regulatory Commission, Order on Rehearing of Monitoring and Mitigation Plan for the California Wholesale Electricity Markets, Establishing West-Wide Mitigation, and Establishing Settlement Conference. EL00-95-031 et al., June 19, 2001, at 27–28, 34.

[12] Under this requirement, no one would enter just to provide service in the event of extreme demand.

difficult to ascertain a demand for insurance against high prices that would be paid during peak periods but not at baseload. However, policies that impose seasonal or time-of-use variations in electricity demand may embody a kind of averaging of peak and extreme-peak prices, leaving baseload rates unaffected.

The benchmark model

Following the models of peak-load pricing in Dansby and in Crew and Kleindorfer,[13] assume that there are two generating technologies, baseload and peak, indicated by subscripts B and P. Peak plants, however, operate only a fraction R of the time. The long-run average (hence marginal) cost of each technology is constant, including a per/MWh capital component K and variable component V. Baseload plants operate full-time. The per capacity unit cost of producing baseload power is $K_B + V_B$. The per capacity unit cost of producing peak power for the fraction R a peak-period power plant operates is $K_P + RV_P$.

We suppose that baseload technology is uneconomical if used only during peak periods, and peak-load technology is uneconomical if run all the time. The first of these implies that

$$K_B + RV_B > K_P + RV_P;$$

the second implies that

$$K_P + V_P > K_B + V_B.$$

These expressions can be combined as

$$V_P – V_B > K_B – K_P > R[V_P – V_B]. \qquad (1)$$

A baseload plant has higher per unit capital cost than a peaking plant, but that difference is less than the difference between the variable cost of the peaking plant and the variable cost of the baseload plant, were both operated full-time. That makes it uneconomical to operate a peaking plant as a baseload unit. However, that difference in capital costs exceeds the difference in operating costs for plants operated only R of the time, making it uneconomical to construct a peaking plant using baseload technology.

---

[13] Dansby, Robert, "Capacity Constrained Peak Load Pricing," *Quarterly Journal of Economics* 92 (1978): 387–98; Crew, Michael, and Paul Kleindorfer, *The Economics of Public Utility Regulation* (Cambridge, MA: MIT Press, 1986).

Let P_P be the peak energy price. Peaking plants will be constructed up to the point at which the revenue obtained from an additional unit of capacity, RP_P, just equals the cost of constructing and operating that capacity, $K_P + RV_P$. In equilibrium,

$$RP_P = K_P + RV_P$$

or

$$P_P = K_P/R + V_P.^{14} \tag{2}$$

Baseload capacity will be installed up to the point at which revenues earned from such capacity equal the costs of construction and full-time operation. A baseload plant charges P_P for the fraction R of the time, plus the baseload price P_B the rest of the time, $1 - R$. Thus, revenues equal costs when

$$[1 - R]P_B + RP_P = K_B + V_B.$$

Substituting from (2) and rearranging terms gives

$$P_B = V_B + \frac{K_B + RV_B - [K_P - RV_P]}{1 - R}. \tag{3}$$

If peak and baseload technologies are identical—that is, $K_B = K_P$ and $V_B = V_P$—then the baseload price P_B just equals V_B, the variable cost. All capital costs are recovered during the peak period. If the technologies differ, equation (1) implies

$$V_P > P_B > V_B.$$

The baseload price exceeds the baseload plant's variable costs such that the plant produces enough revenue to cover baseload capacity.[15] That price, however, is less than the variable cost of a peaking plant, which recovers all of its capital (nonvariable) costs during the peak period.

---

[14] Studies of whether there is market power in the sale of electricity at peak periods typically neglect the capital component, measuring instead the margin between price and only the average variable cost of the most expensive plant necessary to supply peak power. Brennan, Timothy, "Mismeasuring Electricity Market Power," *Regulation* (Spring 2003): 60–65.

[15] This suggests that off-peak prices will also exceed variable costs and include a component for capital recovery, casting additional doubt on studies that attempt to ascertain market power by comparing prices and variable costs.

Capacity-based reserve programs

A natural interpretation of a reserve margin would be that for each unit of electricity one plans to produce, one has to have Z% more capacity. We can embed such a requirement in our present model. Assume further that if demand is at peak, the imposer of the capacity requirement—the regulator, grid manager, or independent system operator—allows generators to sell energy from this added capacity at a "strike price," P_S, following our characterization of reserve margins as a call option. Following the benchmark model, we abstract from uncertainty to display more clearly the expected effects of capacity-based reserve policies. With no uncertainty, this "call option" becomes nothing more than a partial payment by the grid operator for the capacity. We presume that $P_P \geq P_S \geq 0$, and that the grid operator covers the cost of this power through sales, perhaps at the market price, perhaps at cost.

First, focus on the peak-period sellers. Assume that for each unit of capacity they use, they have to install Z (presumably less than one) additional units of peaking capacity, and that the regulator imposing the requirement can ensure that the reserve capacity is suitable for providing power at peak periods. We assume that peak plants are in service a fraction $R < 1$ of the time. The seller gets P_P for sales from the unit it installs on its own, and P_S for sales from the Z units it has to install in addition to comply with the margin requirement.[16] The marginal operating cost of planned and reserve capacity remains V_P.

Let P_P' be the equilibrium peak price with the reserve requirement. In equilibrium, peak sellers will install capacity until the point at which the revenues cover its costs:

$$R[P_P' + ZP_S] = [K_P + RV_P][1 + Z], \tag{4}$$

implying

$$P_P' = K_P[1 + Z]/R + V_P[1 + Z] - ZP_S.$$

The net change in the peak price from the benchmark is

$$P_P' - P_P = Z\left[\frac{K_P}{R} + V_P - P_S\right] = Z[P_P - P_S] > 0. \tag{5}$$

[16] We consider below the possibility that the capacity is installed but not used. Again, we assume that regulators can verify that reserve capacity is suitable for peak use.

The option requirement implicitly taxes some peak purchases to subsidize those at the strike price. If the strike price equals operating cost, then the peak price charged in the market overall increases as if capacity were taxed at the rate K_PZ/R.[17] If the strike price P_S is the same as the equilibrium price P_P', $P_P' = P_P$. The reserve requirement has no effect on peak-period prices and overall capacity.[18]

Turning now to the baseload plants, we assume that under the reserve requirement, the baseload generators install Z units of *peak* capacity for every unit of *baseload* capacity they install. This assumption has three justifications. First, peak capacity costs less than baseload capacity (see equation (1)) and thus would be chosen by baseload plants if they were free to choose.[19] Second, the nominal purpose of this policy is to provide extra capacity when electricity demand is greater, not when it is relatively low. Third, absent a strike price below the presumably low baseload price, the argument above for peak power would mean that the baseload plants would end up charging the same baseload energy price and installing the same amount of capacity, making a reserve margin meaningless.

For every unit of baseload capacity installed and operated full-time, the baseload generator earns P_P' R amount of the time and P_B', the new baseload price, 1 – R of the time. We also assume that the baseload plant sells energy from its reserve capacity at the strike price P_S, operating the fraction R of the time demand is at peak. Accordingly, in equilibrium, revenues from the last unit of baseload installed must cover associated costs:

$$R[P_P' + ZP_S] + [1 - R]P_B' = K_B + V_B + ZK_P + RZV_P.$$

Using equation (4) to substitute for the first term,

$$[K_P + RV_P][1 + Z] + [1 - R]P_B' = K_B + V_B + ZK_P + RZV_P.$$

All of the terms involving Z cancel out, leaving

$$P_B' = V_B + \frac{K_B + RV_B - [K_P + RV_P]}{1 - R}.$$

[17] This is the same effect as if the extra Z units of capacity were not used.

[18] This is why the strike price is important. If energy from the reserve capacity can be sold at the market price, the margin requirement has no effect.

[19] This explains why peaking plants would add peaking capacity to meet their reserve margins as well.

From equation (3), $P_B' = P_B$. *The reserve capacity requirement has no effect on the baseload price.* The costs of the margin requirement are just covered by the increase in rents the baseload plant earns during the peak period as a result of imposing the margin requirement on peaking plants.

Combining these results for peak and baseload plants shows that reserve requirements implemented as a fraction of capacity will increase, not decrease, expected volatility. Customers who cannot obtain peak capacity at the lower strike price will pay higher peak prices, covering the surplus captured by those able to obtain electricity on-peak at the lower strike price. Those customers continue to pay the same baseload price they would pay without the requirement. Hence, mandating reserves increases volatility. Whether that increase would be even worse absent the requirement, once uncertainty (particularly regarding unit outages) is introduced is a separate question. But this increase in volatility is a clear strike against reserve requirements instituted in this manner.

Reviewing the case in which capacity is built but not used shows that retaining unused capacity makes matters worse, as one would expect. For this unused capacity, the strike price P_S would be zero, but no operating costs V_P would be incurred in running that additional capacity. Hence, from equation (10),

$$P_P' - P_P = K_P Z/R.$$

The requirement raises peak prices by the cost of the added capacity, divided by the fraction of time that capacity is being used.[20] For baseload plants, the higher peak price they can charge the fraction R of the time demand is at peak will cover their reserve requirements as well, so baseload prices do not change. Hence, a requirement to maintain excess capacity will simply raise peak-load prices.

Imposing requirements on energy production rather than capacity

The inability of option-like capacity-based requirements to reduce peak prices on more than a subset of sales or to a subset of customers suggests taking a different approach. As before,

[20] If the generators or regulators cannot commit to the excess capacity remaining unused, then the analysis reverts to the case above, in which energy from the nominally reserved excess capacity can also be sold at the prevailing peak price. In that case, as we saw, the reserve requirement is meaningless, with no effect on either peak or baseload prices.

beginning with the benchmark model, we assume that peak and off-peak demand is known with certainty—that is, that there is no excess capacity in either period. Hence, we can equate output and capacity in both periods. This allows us to treat an extra capacity or reserve margin requirement as a tax on energy production used to subsidize capital.

Let $Q_P(P_P)$ be the amount of electricity demanded on-peak, and let $Q_B(P_B)$ be the baseload quantity demanded at any given time within the relevant period (peak or baseload). With certainty in equilibrium, we have $N_B = Q_B$ units of baseload capacity and $N_P = Q_P - Q_B$ units of peaking capacity. The total amount of energy supplied over a given unit of time is $[1 - R]Q_B + RQ_P = N_B + RN_P$.

Let T be a tax on energy supplied and S be a subsidy of capacity. Because output (Q) equals capacity (N) during both peak and baseload periods, the condition that the tax revenues equal the capacity subsidy is

$$TN_B + TRN_P = SN_B + SN_P,$$

implying that

$$S = T\frac{N_B + RN_P}{N_B + N_P}. \tag{6}$$

Letting an asterisk indicate the presence of the tax and subsidy plan, the effect of the tax and subsidy on peak prices, from (2), will be

$$P_P{}^* = [K_P - S]/R + V_P + T. \tag{7}$$

From (6), the net change in the peak-period price is

$$P_P{}^* - P_P = T - S/R = T\ \frac{N_B}{N_B + N_P}\ \frac{R-1}{R} < 0. \tag{8}$$

The tax and subsidy scheme causes peak-period power prices to fall (unless R = 1, i.e., there is no baseload period and everything is peak), as expected, since the capacity-output ratio is greater for peak units than for baseload plants.

Consider the effect of this tax and subsidy plan on baseload prices P_B. From (3), reducing capacity costs by S and increasing variable energy costs by T, we obtain

$$P_B = [V_B + T] + \frac{[[K_B - S] + R[V_B + T]] - [[K_P - S] + R[V_P + T]]}{1 - R}. \tag{9}$$

The added S and T terms in the numerator of (9) cancel. From (3) and (9), accordingly,

$$P_B^* - P_B = T. \qquad (10)$$

The baseload price increases by the energy tax. From the first two terms on the left side of (3), the subsidy on capacity has no effect on price. Although it reduces the cost of baseload capacity (subtracting S from K_B), it reduces the rents earned from on-peak sales by the same amount (subtracting S from K_P).

Equations (8) and (10) lay out the story. The tax and subsidy scheme raises baseload prices and generates a net subsidy to peak-period power sales. The size of the net subsidy depends on R, the fraction of time at which demand is at peak, and the quantities of peak and baseload capacity installed. The precise relationship between the tax and the subsidy requires knowing the demand functions $Q_B(.)$ (= N_B) and $Q_P(.)$ (= N_P). In this setting, the effect of the requirement is to reduce volatility, raising baseload prices and reducing peak prices. This matches, at least qualitatively, the pattern we would expect risk-averse or expenditure-smoothing consumers to prefer—paying a premium when times are good to reduce costs when times are bad.

In the model designed here, the policymaker who knows the demand functions could, within the limits alluded to just above, set a target peak-period price reduction and then use the tax to achieve it.[21] However, knowing the peak-period demand curve $Q_P(.)$, the policymaker could just as easily set an output target equal to what would be demanded at the target price. Moreover, because of certainty and, thus, the absence of unused capacity in equilibrium, the policy could be translated yet again as a demand by the policymaker for an increase in peak capacity above that which one would see absent policy intervention.[22] Letting ΔN_P be that increase in capacity, the relationship between it and the price target P_P^* is

$$\Delta N_P = Q_P(P_P^*) - Q_P(P_P).$$

One might be able to characterize the policy as equivalent to imposing a "reserve margin" of $\Delta N_P/[N_P + N_B]$. However, as we see below, one cannot entirely translate this policy into a tax or margin requirement on capacity itself.

---

[21] Because N_B goes to zero as T gets arbitrarily large, a policymaker has only limited ability to reduce peak-period power prices by this method.

[22] Again, see Jaffe and Felder, n. 4 *supra*, for a general discussion of uncertainty's effect on using prices or quantities to implement capacity policy.

Were one to impose a tax on capacity rather than on output, the subsidy on capital would equal the tax, with no effect on electricity markets. Tracing a capacity tax through the above analysis, one would first observe that the numerator in (4) equals the denominator, implying that S = T. Because the tax on is now on capacity instead of output, equation (5), specifying the new peak-period price, becomes

$$P_P{*} = V_P + \frac{K_P + S - T}{R} = V_P + \frac{K_P}{R} = P_P,$$

implying no change in the peak-period price, hence no change in peak capacity. A similar recharacterization of equation (9) above shows that the baseload price P_B would also not change.

Capacity requirements as an extreme-peak price control

As noted at the outset, a common if dubious method for ascertaining market power and setting an appropriate wholesale price cap is to assume that the competitive price would be the average variable cost of the marginal unit dispatched. The seemingly obvious problem with such pricing is that it would ensure that the marginal unit would never recover its capital costs. Competitive markets would never set prices in this way.

A market in which firms were regulated to provide such capacity could do so if the requirements led to sufficiently large capacity to lower price to the average variable cost of those units.[23] To examine such a possibility, we retain baseload and peak plants but also include units to meet what we will call "extreme peak" demand. Following Dansby,[24] we assume that the capital cost per unit of output falls and variable costs rise as one moves from baseload to peak to extreme. Letting subscripts B, P, and E indicate baseload, peak, and extreme units, respectively, with K as per unit capital cost and V as per unit variable cost, we have

$$V_B < V_P < V_E$$

and

[23] A different approach, not modeled here, would be to require sufficient capacity such that the price of electricity does not go above the cost of a blackout were the electricity unavailable, also known as the "value of lost load" (VoLL). National Electricity Code Administrator, Reliability Panel Final Report, *Review of VoLL in the National Electricity Market: Report and Recommendations*, July 1999; see also Joskow, n. 1 *supra* at 37; Jaffe and Felder, n. 4 *supra*.

[24] Dansby, n. 13 *supra*.

$K_B > K_P > K_E$.

We amend our notation slightly to deal with this case. Hoping redundant notation will be more helpful than confusing, we suppose that for a fraction B of the time, only baseload plants operate. For an additional fraction P of the time, peak plants come online as well, and for fraction E of the time, all three types of plants operate, where $B + P + E = 1$. We also will assume that the efficient plant for each level of demand is unique—that is, that the same technology would not be used to build units designed to serve two of the three levels.[25]

If the regulator establishes capacity requirements to ensure that the extreme-demand price is V_E, no firm will supply capacity just to meet that demand. Its expected revenue from each unit of capacity, EV_E, would be less than the cost of supplying and operating that capacity, $K_E + EV_E$. Hence, all capacity to serve extreme demand has to come from peak and baseload suppliers. Under this requirement, each baseload and peak supplier has to install an extra Z units of extreme capacity (the least-cost extreme capacity, by assumption) as a reserve requirement for each unit of capacity it installs. Z is set such that enough capacity is available during extreme-demand periods to keep price at V_E.

A peak supplier does not operate when only baseload plants are needed. It operates its peak units a fraction P of the time, selling at price P_P. At extreme-demand periods occurring a fraction E of the time, it operates its peak units and its extreme units, selling electricity at price V_E. For each unit of peak capacity it installs, its revenues are PP_P and $E[1+Z]V_E$, where Z represents the amount of extreme capacity it is required to install for each unit of peak capacity.[26] Its per capacity unit operating costs are $[P + E]V_P$ for its peak units and EZV_E for its extreme units. Its costs include capital costs $K_P + ZK_E$. In equilibrium, to ensure cost recovery for each unit of peak capacity installed to meet reserve requirements,

---

[25] This assumption requires that six inequalities hold. Baseload plants run all the time, peak plants a fraction P + E of the time, and extreme plants the fraction E. The inequalities are as follows:

For baseload: $K_B + V_B < K_P + V_P$, $K_B + V_B < K_E + V_E$.

For peak units: $K_P + [P+E]V_P < K_B + [P+E]V_B$, $K_P + [P+E]V_P < K_E + [P+E]V_E$.

For extreme units, $K_E + EV_E < K_B + EV_B$, $K_E + EV_E < K_P + EV_P$.

The inequalities relating baseload and extreme plants are redundant. If a peak plant is uneconomical as a baseload plant and an extreme plant uneconomical as a peak plant, then the extreme plant is uneconomical as a baseload plant. Similarly, if a baseload plant is uneconomical as a peak plant and a peak plant uneconomical as an extreme plant, then a baseload plant is uneconomical as an extreme plant.

[26] At extreme-demand periods, the peak unit would sell power from its peak plants at V_E as well.

$$PP_P + E[1 + Z]V_E = K_P + ZK_E + [P + E]V_P + EZV_E. \qquad (11)$$

The EZV_E terms on both sides of the equation cancel out because extreme price equals variable costs. Hence,

$$P_P = V_P + \frac{[K_P + EV_P] + ZK_E - EV_E}{P}. \qquad (12)$$

The conditions ensuring that a peak plant not be economical as an extreme plant and vice versa ensure that $P_P > V_P$.[27]

We look at the similar per unit revenue and cost calculation for baseload plants. During a fraction B of the time, these plants obtain revenue per unit P_B. At the fraction P where demand is at peak, the per unit revenue is P_P. During the fraction of time E of extreme demand, the baseload unit gets V_E from its baseload unit and ZV_E from the extreme-capacity unit it was required to install. Accordingly, the per baseload unit revenue is $BP_B + PP_P + E[1+Z]V_E$. Per baseload unit capacity costs are $K_B + ZK_E$. Its per baseload unit operating costs are V_B, running that baseload unit all the time, plus EZV_E, when it runs its extreme-demand unit. Hence, in equilibrium,

$$BP_B + PP_P + E[1+Z]V_E = K_B + ZK_E + V_B + EZV_E.$$

We can substitute the right-hand side of (11) for the last two terms of the left-hand side in the above expression, cancel out the EZV_E terms, and obtain

$$P_B = V_B + \frac{K_B - K_P - [P + E][V_P - V_B]}{B}. \qquad (13)$$

This is the same as equation (3) above, where P+E is R, the time of above-baseload demand, and B = 1 – R, the time of low demand. As in that case, *the reserve requirement has no effect on the baseload price*. The baseload plant recovers the added cost of the capacity requirement during the peak period. Hence, baseload capacity is unchanged under the requirement as well, as we posit that only enough such capacity is installed to meet demand at P_B.

The requirement does affect the price during peak periods, and hence the amount of peak and extreme capacity installed. To see this, we can calculate what the extreme and peak prices

---

[27] This follows from the equations in n. 25 *supra*, specifically that $K_E + EV_E < K_P + EV_P$. We can also assume that $V_E > P_P$ so that the capacity requirement does not lead peak prices to exceed extreme prices, but that is not necessary for the results.

would be absent the capacity requirement. Without such a requirement, extreme plants would have to cover their costs at price P_E during the fraction E of the time when they operate. Accordingly,

$$P_E = V_E + K_E/E. \tag{14}$$

To compare peak-period prices with and without the reserve requirement, let P_P^o be the peak-period price without the requirement, leaving P_P as the price under the regulation. Peak plants thus enter up to the point at which the per capacity unit revenues from sales at P_P during fraction P of the time and P_E at fraction E of the time cover the costs of constructing that capacity and operating it the fraction P + E of the time.

$$PP_P^o + EP_E = K_P + [P + E]V_P.$$

From (14) and rearranging terms,

$$P_P^o = V_P + \frac{[K_P + EV_P] - [K_E + EV_E]}{P}. \tag{15}$$

Again, because peak plants are uneconomical to run only at extreme times, the numerator of the fraction on the left-hand side is positive, implying that the peak-period price is above peak-period variable cost V_P.

Comparing (12) and (15),

$$P_P - P_P^o = \frac{[1 + Z]K_E}{P}.$$

A policy to set Z large enough that extreme-period prices fall to variable cost raises the peak-period price. It increases enough not merely to cover the added capital cost (ZK_E/P), but also to cover the lost rents (K_E/P) that would have been earned during the extreme period had the price been P_E as calculated in (14) rather than V_E.

As noted above, baseload prices do not change, since baseload plants follow the price of peak plants during the fraction of time P + E that the peak plants are operating. Consequently, the effect of a requirement to construct enough capacity to ensure that prices remain at variable cost during extreme periods will be to raise prices during peak periods, imposing an effective tax of $\frac{[1 + Z]K_E}{P}$. These higher peak prices imply less electricity demanded during peak periods, and hence less peak capacity will be built. Prices and baseload capacity remain unaffected.

Since the requirement reduces the prices during extreme periods, with less peak capacity and no additional baseload capacity, more high-cost, extreme-capacity plants will be built under the requirement, increasing the size of the requirement (Z) and the effective tax on peak energy necessary to meet the policy objective. It is difficult to imagine a scenario in which the desired price path would be to have baseload prices remain unchanged while just peak prices rise to ameliorate extreme-peak situations. However, if both peak and extreme demands occur within the same seasons or times of the day, and other seasons or times have demand at baseload levels, seasonal or time-of-use pricing will replicate the changes in price imposed by reserve requirements implemented in this fashion.

Summary observations

That capacity requirements are desirable to counter the blackout externality arising from grid interconnectedness and vulnerability to load imbalances is conventional wisdom. Less clear are how capacity requirements are designed and how their costs are covered. We examined three scenarios:

- If requirements exist as a right for the grid operator to summon peak power at a price below the spot price, they leave baseload prices unchanged. Those unable to get peak power at the strike price subsidize those who can. For those customers, reserve requirements increase price volatility, counter to the presumed intention of the policy.
- If requirements are effectively an output tax to subsidize capacity, baseload prices increase and peak prices fall. Wealth is redistributed, as might happen if customers wanted to smooth payments over time or insure against high peak prices but were unable to do so.
- If requirements mandate the installation of sufficient extreme-demand capacity to ensure pricing at variable cost, baseload prices are unchanged. Peak prices rise to cover the cost of the capacity, reducing the amount of peak capacity that would have been installed and increasing the construction of extreme-demand capacity. Such a result does not seem to follow from any expected preference for insurance against supply risk or payment smoothing, but it could match the outcomes when electricity prices are seasonally or time-of-use based.

Recognizing that the burden of paying for capacity requirements is highly sensitive to their design could lead to a more careful assessment of their extent and how they should be implemented, to mitigate uncertainty and the blackout externality at least cost.