DISCUSSION PAPER

September 2008 ■ RFF DP 08-33

Is the Benefit of Reserve Requirements in the "Reserve" or the "Requirement"?

Timothy J. Brennan

1616 P St. NW
Washington, DC 20036
202-328-5000 www.rff.org

RESOURCES
FOR THE FUTURE

Is the Benefit of Reserve Requirements in the "Reserve" or the "Requirement"?

Timothy J. Brennan

Abstract

Reliability in electricity markets is, in many respects, a public good, in that one supplier's failure to meet its customers' demands can cause failure throughout the grid. This creates a blackout externality. One of the remedies for a blackout externality is a reserve requirement, where load-serving entities have capacity on hand to meet demand in the case of unexpected surges in demand or unit failures. Modeling the magnitude of the externality as a positive function of use and negative function of capacity reveals that a benefit of capacity requirements is that covering their costs imposes a tax on usage. After illustrating this possibility, a model addressing the sector as a whole, where spot markets can resolve individual but not overall shortfalls, illustrates that capacity requirements should be increased or decreased to exploit this usage tax effect.

Key Words: electricity, reliability, reserve requirements, capacity

JEL Classification Numbers: L94, H23, D24

Contents

Is the Benefit of Reserve Requirements in the "Reserve" or the "Requirement"?

Timothy J. Brennan*

Introduction

Considerable attention has been given to reserve requirements in electricity. The aspects that have received the most attention are on the "reserve" side of the equation. These include separating capacity payments from energy prices and designing capacity markets to set that price. This discussion is typically divorced from the fact that capacity payments will be passed on to consumers as energy prices. Brennan (2003) considered how capacity payments would affect end user prices—an issue that had been neglected and that appears to remain so.[1]

Much of the literature on capacity markets treats them solely as a regulatory requirement imposed on sellers, with little consideration of effects on end users. Because those users do not pay separately for capacity, these payments will be translated into prices based on energy use. How this happens depends on the nature of the reserve requirement itself, particularly the "strike price" paid for exercising the option to use that capacity, and the basis for imposing that reserve requirement. Brennan (2003) examined different structures of reserve requirements to see if the effects on off-peak and peak prices reflect plausible benefits to consumers, such as ways in which they might want to purchase insurance against blackouts by adding reserves.[2]

* Professor, Public Policy and Economics, University of Maryland, Baltimore County; Senior Fellow, Resources for the Future. Email: brennan@umbc.edu. I thank Larry Blank, Seth Blumsack, Mark Friese, Kiwan Lee, Catherine McDonough, and other participants in the Rutgers Advanced Workshop in Regulation and Competition, 27th Annual Eastern Conference, for comments and suggestions. All errors and shortcomings are the author's sole responsibility.

[1] At least within the economics literature, the question of optimal reserve requirements has received little, if any, attention. In the main, if not the only, recent article addressing this issue, Joskow and Tirole (2007) focus more on general system planning assuming inelastic demand rather than how one might optimally tax use to mitigate a blackout externality.

[2] McDonough and Kraus (2008) provide very useful illustrations of the effects of capacity charges on real-time electricity prices.

This analysis builds upon that prior work to examine how the effects on usage prices could assist in the design of reserve requirements themselves. The driver of policy in this area is the *blackout externality* that occurs when the failure of one supplier to meet its customers' demands increases the likelihood of an outage across the wider grid.[3] This externality depends essentially on the likelihood that a supplier would not be able to meet demand from its customers.[4] Undoubtedly, this depends on a wide variety of factors, including the suppliers' decisions regarding unit maintenance, grid management technology that can anticipate and thus restrict the spread of blackouts, and the ability of the supplier to put units back into service.

To garner some insight into how reserve requirements should be designed in light of this blackout externality and the effects on use, we can simplify the analysis by assuming that reserve requirements affect only the likelihood of a blackout, not its magnitude. This assumption, although probably a significant simplification, is common in theoretical analyses of liability rules designed to induce potential tortfeasors to take optimal care. In contrast to settings in which optimal care is induced by *ex post* rules that require those responsible for accidents to cover their costs (Cooter and Ulen 2004), reserve requirements are *ex ante* rules that set care levels in advance. Implicit in assuming a need for reserve requirements is that *ex post* bearing of costs will not work. This is likely to be justified, particularly if the costs of a blackout are highly variable and difficult to measure, assigning blame is problematic, and the damages are likely to exceed the assets of the individual generator (Brennan and Palmer 1998).[5]

---

[3] As Larry Blank has observed, if outages can be restricted to the customers of generators who fail to meet demands, no blackout externality exists. Without that externality, generators can optimally compete on the basis of reliability, and policy interventions, such as reserve requirements to maintain reliability of the grid as a whole, are not needed. A similar argument applies to policies to mandate real-time prices, which are unnecessary unless one generator's failure to meet peak demands causes outages elsewhere in the grid (Brennan 2004).

[4] As just noted, this externality can also rationalize the imposition of real-time meters, even though the market should already incorporate the efficiency benefits from time-based pricing when doing so exceeds the cost of the meters (Brennan 2004).

[5] See Boyd and Ingberman (1994) for an analysis of tort liability when bankruptcy can render tortfeasors potentially judgment-proof.

Because the likelihood of a blackout depends on use relative to capacity, the virtue of reserve requirements may be not only in the reserves, but also in imposing on users the cost of meeting the requirement. To mitigate a blackout externality, one could want to reduce use as much as increase capacity. Imposing the payment for reserves on the basis of use, particularly during peak usage periods when the risk of a blackout is greatest, would be beneficial on its own as an efficient response to the externality.

The models presented below illuminate this possible benefit. The first is a simple model for illustrative purposes, in which energy demand does not vary over time. Even this simple model is instructive, in showing why incorporating the blackout externality probably requires that reserves be just that—not just more capacity, but more *unused* capacity. Thinking about requirements in this way has implications for the magnitude of requirements and how they might be imposed on individual load-serving entities (LSEs). The following models begin by looking at individual firms first without demand varying across time. The next model exploits the analytical convenience of using spot markets to justify considering the blackout as a function of grid-wide use and capacity rather than individual use and capacity. This also allows incorporating time-varying probabilities of blackouts into the design of the optimal policy. This demonstrates how reserve markets should be designed to incorporate the merits of an implicit tax on use to account for the blackout externality.

The Basics

An unrealistically simple benchmark case, with no variation in use over time and perfect information regarding use and cost, nevertheless illuminates the role reserve requirements might play in inducing appropriate reductions in electricity use. On the supply side, the generation sector produces X megawatt-hours (mWh) of energy with capacity K. Capacity cannot be exceeded; $X \leq K$. The marginal cost of producing the X^{th} unit of energy at any particular time, given capacity at least as great as X, is $c(X)$; the capacity that can produce at least cost is always used first. The cost of capacity used to provide the X^{th} unit is $h(X)$. The total cost of producing X units with K units of capacity, $C(X, K)$ is

$$C(X,K) = \int_o^X c(x)dx + \int_0^K h(k)dk \ . \tag{1}$$

The marginal cost, including capacity, of producing the X^{th} unit is also increasing, that is, $c' + h' \geq 0$. However, capacity used only a fraction of the time will probably cost less than capacity used all of the time, implying $h' < 0$ (Crew and Kleindorfer 1986). To ensure that capacity that would be used more often—where the fraction of time used t is between 0 and 1—has a lower marginal cost of production, it suffices to assume that c'' and h'' are both greater than zero, that is, the increase in marginal cost is increasing in production and the reduction in the cost of capacity used to produce additional electricity falls as more is added.[6]

To understand the effect of the blackout externality on use and capacity, the first model presents a base case, assuming that demand does not change over time. Neglecting time allows the demand curve to be given by a fixed $p(X)$. Overall benefit $W(X)$ gross of payments by consumers for electricity is given by

$$W(X) = \int_o^X p(x)dx \ ; \tag{2}$$

hence $W'(X) = p(X)$. With demand constant over time, capacity would always be used, so $K = X$. The total cost of production is then $C(X, X)$ as defined by (1). Maximizing net welfare involves choosing total production X (with matching capacity) to maximize the difference between $W(X)$ and $C(X, X)$, giving the amount of electricity X^* such that price equals marginal cost, which here is given by

$$p(X^*) = c(X^*) + h(X^*). \tag{3}$$

In a competitive environment, any particular plant i would take $p(X^*)$ as given, and choose its own output X_i and capacity $K_i = X_i$ to the point where price equals marginal cost, locally

[6] See the appendix. We make this assumption to cover the possibility, but do not look at the time-varying implementation of capacity here.

satisfying (3).[7] Requiring that supply equals demand ensures that (3) holds globally. Two aspects of this simple relationship are noteworthy. First, efficient pricing will include capacity rents, with price above short-run marginal cost of the marginal unit ($c(X^*)$); this contrasts with views that efficiency requires bids and prices to equal short-run marginal cost at all times.[8] Second, excess capacity would not arise if demand does not vary over time, either systematically (e.g., summer days vs. spring evenings) or stochastically.

Incorporating the Blackout "Tort"

In some cases, externalities can be modeled without looking at specific generators. For example, climate change policies are based on total carbon dioxide (CO_2) emissions or CO_2-equivalent emissions; the greenhouse effect attributable to all emitters is identical. With respect to blackouts, however, the risk of a grid-wide blackout is not simply based on the probability that aggregate demand exceeds available capacity. Rather, it depends on the degree to which each individual supplier's decisions affect the likelihood of a blackout. Accordingly, further analysis of the effect of the blackout externality requires decomposing the aggregate cost and competitive supply function described above into those of individual generators.

To do this, a model of the blackout externality with slightly more detail is needed. In the interest of simplification, treat the cost imposed by a blackout as B. In this way, the effect of an individual generator i's actions on a blackout become purely a matter of the degree to which they change the probability ρ of a blackout. The expected harm of the blackout is just ρB. As noted above, the determinants of the probability of a blackout will be the differences across providers between the capacity each generator has in place and the energy each one supplies. Let X_i and K_i

[7] To allow this condition to hold for all generators, we assume away indivisibilities in terms of the size or characteristics of units available to any individual generator.

[8] Such views have led to exaggerated estimates of market power and inefficient price ceilings (Brennan 2006). One of the scenarios in which one may need a reserve requirement and separate capacity market is when regulators or market mitigation agencies prevent suppliers from bidding prices that exceed average variable costs, so as to recover capacity costs.

be generator i's energy supply and capacity, respectively, indexing i from 1 to N to reflect N suppliers in place. This yields

$$\rho = \rho(X_1, K_1, X_2, K_2, \ldots X_N, K_N), \tag{4}$$

where $\partial\rho/\partial X_i \geq 0$ and $\partial\rho/\partial K_i \leq 0$.[9]

Because of the interconnectedness of the grid, a blackout is akin to a joint tort, in which the probability of harm is the result of the choices made by individual actors. In the tort case, the variable reflects the level of care taken by potential tortfeasors (Landes and Posner 1987). In this context, "care" is equivalent to "excess capacity;" in other words, $K_i - X_i$. At the efficient outcome, each generator would act as if it is internalizing the expected cost of its actions, making choices simultaneously with other generators rather than sequentially.[10] Were firm i to face those costs, assuming it takes price P as given, it will choose its output X_i and capacity K_i to maximize profits Π_i given by

$$\Pi_i(X_i, K_i) = PX_i - C(X_i, K_i) - \rho(\ldots, X_i, K_i, \ldots)B, \tag{5}$$

where

$$C(X_i, K_i) = \int_o^{X_i} c_i(x)dx + \int_0^K h_i(k)dk \tag{6}$$

is the cost to firm i of producing X_i and K_i given the generation units it has, which have marginal costs given by c_i and h_i. The following conditions hold for the individual firm: $X_i \leq K_i$, c_i, $c_i' \geq 0$, $h_i > 0$, $h_i' \leq 0$, $c_i' + h_i' > 0$. To ensure that added capacity is generally worthwhile when less frequently invoked (although time is still neglected), c_i'', $h_i'' > 0$.

[9] Of course, other aspects of the grid outside of generation—such as transmission capacity and maintenance—can affect the likelihood of a blackout.

[10] Sequential liability introduces optimal sequential rules (Wittman 1981) that are probably impractical in the blackout context.

From (5), and letting λ be the shadow cost of meeting the condition that $X_i \leq K_i$, firm i chooses X_i and K_i to maximize

$$PX_i - C(X_i, K_i) - \rho(X_i, K_i)B - \lambda[X_i - K_i], \tag{7}$$

which gives the conditions

$$P - c_i(X_i) - [\partial\rho/\partial X_i]B - \lambda = 0 \tag{8}$$

and

$$- h_i(K_i) - [\partial\rho/\partial K_i]B + \lambda = 0.^{11} \tag{9}$$

The constraint may be binding or nonbinding. If the constraint is binding, $X_i = K_i$. If so, no reserve capacity exists, despite the blackout externality. Combining (8) and (9) gives

$$P = c(X_i) + h(X_i) + [\partial\rho/\partial X_i + \partial\rho/\partial K_i]B. \tag{10}$$

This implicitly defines the supply function $X_i(P)$; P then is set by equating supply from the N firms and demand, defined by

$$W'^{-1}(P) = \sum_{1}^{N} X_i(P), \tag{11}$$

where $W'^{-1}(P)$ is the inverse of the marginal willingness to pay W' derived from (2).

If the constraint is binding, the effect of internalizing the cost is to change the scale of the generator. The effect depends on the sign of $\partial\rho/\partial X_i + \partial\rho/\partial K_i$, which one can think of as the marginal probability that the firm will not be able to meet its demand, and thus cause a blackout, as a function of scale. Because the first term is positive and the second negative—increasing production relative to capacity increases the likelihood of a blackout, and vice versa—$\partial\rho/\partial X_i$ +

[11] All second-order conditions necessary for values satisfying these conditions to be a global maximum are assumed.

$\partial\rho/\partial K_i$ can be either positive (a greater likelihood of a blackout as the firm increases in size) or negative.

One might speculate, *a priori*, that larger firms have more units. As all units are assumed to operate at capacity, the chance that all units will *not* fail will fall with the number of units, suggesting that this term is positive. From (9), if $c' + h' > 0$ as assumed, this implies that X_i would be smaller, with optimal internalization of the costs of a blackout. This reduction in the quantity of energy supplied at a given price shifts back the competitive supply curve, causing the price to increase. Hence, we have:

> **Finding 1:** *If output remains at capacity, optimal internalization of the blackout externality probably causes the price of electricity to rise because, holding prices constant, the optimal size of a generation company falls.*

In this sense, the optimal blackout policy is equivalent to a tax on production equal to the increase in the probability of a blackout with supplier size multiplied by the cost of a blackout.

If the constraint is nonbinding, $K_i > X_i$. In this case, one would observe unused capacity when no unit failures occur. In this setting, the optimal X_i and K_i are given by (8) and (9) with $\lambda = 0$:

$$P = \mathrm{c_i}(\mathrm{X_i}) + [\partial\rho/\partial X_i]B \qquad (12)$$

and

$$h_i(K_i) = -[\partial\rho/\partial K_i]B. \qquad (13)$$

The first condition, (12), implies that the price of energy should equal its marginal cost of production, neglecting capacity costs, along with a tax $[\partial\rho/\partial X_i]B$ to reflect the fact that increasing output, holding capacity fixed, increases the likelihood of a blackout. The second condition, (13), implies that capacity should be installed up to the point where the cost of added capacity ($h(K_i)$) equals the marginal benefit from reducing the likelihood of a blackout ($-[\partial\rho/\partial K_i]B$).

These results, however, have two shortcomings. First, because $h' < 0$—capacity used less often costs less—second-order conditions for a profit maximum need not hold. We can assume that the second-order conditions for an optimum from the planner's perspective hold, so that (12) and (13) define a local maximum. In particular, we would require, taking the derivative of the first-order condition (9) with respect to K_i, that

$$-\,h_i'(K_i) < [\partial^2\rho/(\partial K_i)^2]B, \tag{14}$$

that is, the benefits of adding capacity in terms of reducing the likelihood of a blackout fall faster than do the costs of the added capacity. However, this does indicate that a system in which the firm receives a subsidy of $-\,[\partial\rho/\partial K_i]B$ for each unit of capacity installed would result in an infinite installation of capacity. Thus, we have

Finding 2: *Optimal policies with excess capacity require specifying the target capacity rather than providing a subsidy to install capacity.*

A subsidy as such is unlikely to be forthcoming in any event. This leads to the second shortcoming that cannot be easily assumed away. If the capacity constraint does not hold, the firm will act as if the price covers only the variable costs of producing energy in addition to the blackout externality tax $[\partial\rho/\partial X_i]B$. Unlike the base case, in which capacity is fully used, the price will not be affected by the marginal cost of capacity $h(K_i)$. The revenues taken in by the firm, PX_i, thus need not cover the total costs $C(X_i, K_i)$ given in (6).

The constraint that revenues cover costs need not be binding. The variable costs of production ($c(X_i)$) and the costs including capacity ($c(X_i) + h(X_i)$) are increasing in X. As a consequence, a price equal to the marginal cost of generation would produce inframarginal rents that could cover the costs of a requirement to purchase extra capacity. With no binding constraints—that is, with the cost of excess capacity covered by inframarginal rents—internalizing the blackout externality leads the firm to optimize

$$PX_i - C(X_i, K_i) - \rho(X_i, K_i)B, \tag{15}$$

leading to the first-order conditions (12) and (13) without caveats. With no constraints, one would optimally want to tax usage and use the revenues to cover the cost of meeting the ideal capacity level.

On the other hand, if electricity rates need to be sufficiently high to generate revenues to cover the cost of extra capacity, as one would expect, then the optimization problem (7) requires a constraint that revenues cover costs,

$$PX_i = C(X_i, K_i). \tag{16}$$

We then choose X_i and K_i to maximize welfare from firm i, subject to the constraints that revenues cover costs, giving

$$PX_i - C(X_i, K_i) - \rho(X_i, K_i)B - \lambda[X_i - K_i] - \theta[PX_i - C(X_i, K_i)]. \tag{17}$$

If the constraint that output equals capacity is not binding, but the cost coverage constraint is binding, then the constrained welfare objective becomes simply

$$-\rho(X_i, K_i)B - \theta[PX_i - C(X_i, K_i)]. \tag{18}$$

The objective becomes the minimization of the blackout externality subject to the constraint that revenues from selling power just cover the costs. When the revenues are constrained to equal costs, the net welfare of operating the plant is zero. The first-order conditions are

$$-[\partial\rho/\partial X_i]B - \theta[P - c(X_i)] = 0 \tag{19}$$

and

$$-[\partial\rho/\partial K_i]B - \theta[h(K_i)] = 0. \tag{20}$$

Using (19) to calculate θ, substituting for θ in (19), and rearranging terms gives

$$P = c(X_i) - \frac{\partial\rho/\partial X_i}{\partial\rho/\partial K_i} h(K_i)\,. \tag{21}$$

Because the numerator in the fraction on the right-hand side of (21), $\partial\rho/\partial X_i$, is positive, and the denominator, $\partial\rho/\partial K_i$, is negative, (21) implies that at the optimum, price exceeds marginal cost. This yields

> **Finding 3:** *If, at the excess capacity optimum, revenues just cover costs, the optimal energy price is above marginal cost, to cover the cost of optimum capacity. The virtues of the capacity requirement include its tendency to induce a tax on usage, which has the side benefit of mitigating the blackout externality.*

Note from (21) that the optimum tax is independent of the size of the blackout externality.

Relying on a Spot Market and Introducing Time

The above analyses, and the policies they imply, presume that reserve requirements should be imposed on individual utilities, as the chance of a blackout depends on the individual utility's supply and capacity choices. However, suppose that individual generators can purchase energy on the spot market for resale if their individual units go down. If so, the chance of a blackout could be thought of as a function not so much of individual relationships between capacity and supply, but the entire grid's balance of the two.

The assumption of an effective spot market in which suppliers can trade with each other to meet demand allows for one major simplification—ρ, the probability of a blackout, can be treated as a function of the aggregate output X and capacity K, with $\rho_X > 0$, $\rho_K < 0$, and with the constraint that $X \leq K$. This allows us to consider welfare in the sector as a whole, rather than taking price as given, although the resulting conditions are the same. From a social perspective, we want to optimize

$$W(X) - C(X, K) - \rho(X, K)B - \lambda(X - K), \tag{22}$$

with $C(X, K)$ defined by (1). Because $W'(X) = \mathrm{p}(X)$, $C_X = c(X)$, and $C_K = h(K)$, the first-order conditions are

$$p(X) = c(X) + \rho_X B + \lambda, \tag{23}$$

$$h(K) = -\rho_K B + \lambda, \tag{24}$$

and

$$\lambda(X - K) = 0. \tag{25}$$

These conditions are essentially those for individual firms, and the analysis above applies to the cases when the constraint is and is not binding.

The simplification of looking at the sector as a whole allows the incorporation of time into the model, which requires a slight modification of the notation. Let $W(X, t)$ be the consumer benefit of using x units of electricity at time t, which runs from 0 to T. Capacity remains K, fixed over t. The chance of a blackout at any given time, $\rho(X, K)$, is taken to be independent of t, other than how X varies over time. At any time t, $X \leq K$, which is represented as a constraint λ that can also vary over time t. Because energy is generated over time, whereas capacity is fixed, the cost in (1) can be decomposed into two separate components. With sufficient capacity in place, the total cost of generating X units of electricity at any given time, $C(X)$, is

$$C(X) = \int_o^X c(x)dx\,, \tag{26}$$

and the total cost of the capacity K used to generate electricity, $H(K)$, is

$$H(K) = \int_0^K h(k)dk\,. \tag{27}$$

Aggregate welfare over time T is then given by

$$\int_0^T [W(X(t),t) - C(X(t)) - \rho(X(t),K)B - \lambda(t)(X(t) - K)]dt - H(K)\,. \tag{28}$$

At any time t, $X(t)$ and $\lambda(t)$ are given by

$$p(X(t), t) = c(X(t)) + \rho_X(X(t), K)B + \lambda(t) \tag{29}$$

and

$$\lambda(t)[X(t) - K] = 0. \tag{30}$$

From (29), price at any time should equal marginal production cost, plus the external blackout cost, plus the shadow price of additional output that would be available with one more unit of capacity. Absent policy intervention, the *X* chosen in the market at any time *t* will exceed the optimal level, as the blackout externality cost, $\rho_X(X(t), K)B$, would not be internalized.

The optimal level of capacity is defined by the first-order condition

$$\int_0^T -\rho_K(X(t), K)dt + \int_0^T \lambda(t)dt = h(K). \tag{31}$$

Capacity should be installed up to the point where the marginal benefit, the sum of the reduced blackout risk aggregated over time (the first term on the left in (31)). The shadow price of additional output aggregated over time (the second term on the left) just equals the marginal cost of additional capacity. This will be more than the capacity the market would produce on its own. The second term, the shadow price, represents a benefit of additional capacity that individual generators can capture by making capacity available and selling energy from it in the spot market. The first term, however, represents a benefit to the grid as a whole that energy sellers would not internalize.

Note that optimal capacity as defined by (31) need not be more than what the market would provide. Although that condition includes an uninternalized externality, which increases the level of capacity, it is also based on the level of use $X(t)$ as defined in (29). That condition does reflect the blackout externality associated with use. Thus, that level of capacity is lower than what the market would provide, reducing the capacity that would be set under (31). The optimal combination of use and capacity could result in less use, but no more capacity, than what one would obtain in the market. It remains the case that the level of capacity one would get with *only* the usage-based negative blackout externality ρ_X internalized would be less than that if the capacity-based positive externality ρ_K were internalized as well. Let ΔK be the difference between these two.

Conditions (29) and (31) illuminate the potential double benefit of a reserve requirement. Suppose that one were to impose a tax on electricity use at any time (presupposing real-time pricing, of course) of $\rho_X(X(t), K)B$. The total tax revenue, R, would then be

$$R = \int_0^T \rho_X(X(t), K)BX(t)dt \, . \tag{32}$$

One could obtain this tax in reverse by requiring added capacity over and above what the sector would construct with supply at any time equal to $X(t)$, the cost of which requires an increase in revenue equal to R as defined in (32) . To translate best such a capacity-funding tax into the effect on use, one would want to assign the obligation so that use at any given time necessitates paying a premium of $\rho_X(X(t), K)B$ on each unit of electricity used. The per-unit tax would be greater as $X(t)$ increases and quite likely would be zero for a sufficiently small $X(t)$.

One virtue of looking at reserve requirements in this way is that an explicit tax on electricity may be politically infeasible. The optimal tax revenue defined in (32) could pay for more, the same as, or less than ΔK of added capacity. Ideally, the payment should be separated from the capacity requirement. But if the only way to impose a per-unit tax is as a contribution to capacity, the capacity requirement should be distorted to reflect the benefits of the tax.[12] In particular, setting the capacity requirement as optimal, as defined by (31), will lead to too little use reduction if, at the optimal tax, one would buy more capacity if the revenues were used to do so. We thus have

Finding 4: *If the optimal tax would pay for more than the optimal amount of reserves, reserve requirements should be increased to induce optimal use reductions. If the optimal tax would pay for less than the optimal amount of reserves, one should reduce reserve requirements below the optimum to prevent undue reductions in use.*

[12] Similar benefits may result from using electricity surcharges to pay for investments in energy efficiency (Brennan 2008, at 19).

Conclusion and Next Steps

Working through models of use, capacity, and blackout externalities illustrates the optimal policies to induce generators to internalize the blackout costs they impose when use is close to capacity. This leads to various insights, depending on whether internalizing such externalities would cause capacity to exceed use, in a simple model without time-varying demand. Such a model suggests that optimal policies would reduce supply on the theory that, in larger-scale firms, a unit is more likely to go down and thus impose an externality. The declining cost of capacity, as it is less likely to be used, requires specifying a capacity requirement rather than a subsidy if excess capacity in reserve is desired. Finally, and most germane to this paper, if revenues from energy sales must cover excess capacity costs, the optimal energy price is above marginal cost, which mitigates the blackout externality as well.

Adapting this framework to allow a spot market at any time implies that the likelihood of a blackout is a function of aggregate supply and capacity, not individual supply and capacity. This makes it relatively easy to allow demand—and thus the blackout externality and whether capacity constraints bind—to vary over time. Incorporating those aspects into the model results a key finding—that reserve requirements should be increased or decreased to exploit the benefits of, in effect, taxing use to internalize the blackout externality.

This examination so far is simplistic and cursory. It neglects an explicit calculation of the probability that capacity would be used, and what price the energy produced by that capacity would receive, if called upon by an LSE or system operator to prevent a blackout. This would affect the expense of capacity held specifically in reserve, indicating what sorts of technologies would and perhaps should be specified.[13] Making explicit the stochastic aspects of blackouts, in terms of both likelihood and severity, would be an important next step.

[13] One presumably would not want to allow capacity requirements to be nominally satisfied by by installing cheap but ineffective technologies.

Appendix

To provide a continuous representation of the relationship between the variable cost $c(X)$ of the marginal unit, $h(X)$ of capacity needed for that marginal unit, and t—here the fraction of time that unit of capacity is in operation—let $t(X)$ be the time at which one would be just indifferent between using the technology to produce X and using that to produce $X + dX$; in other words, the point where

$$t(X)[c(X)] + h(X) = t(X)c(X + dX) + h(X + dX). \tag{A1}$$

Collecting terms and dividing by dX gives

$$t(X)c'(X) + h'(X) = 0. \tag{A2}$$

This can be differentiated to obtain

$$c''(X)t(X) + t'(X)c'(X) + h''(X) = 0, \tag{A3}$$

and rearranging and suppressing the argument of the functions gives

$$t' = -\frac{tc'' + h''}{c'}. \tag{A4}$$

This defines the condition for which $t'(X) < 0$; in other words, the condition for which it is profitable to use the technology to produce X a smaller fraction of the time than the technology to produce Y if X is greater than Y. This will hold if c', c'', and h'' are all positive, as indicated in the text.

References

Boyd, James, and Daniel Ingberman. 1994. Noncompensatory Damages and Potential Insolvency. *Journal of Legal Studies* 23: 895–910.

Brennan, Timothy. 2003. Electricity Capacity Requirements: Who Pays? *Electricity Journal* 16(8): 11–22.

Brennan, Timothy. 2004. Market Failures in Real-Time Metering. *Journal of Regulatory Economics* 26: 119–39.

Brennan, Timothy. 2006. Preventing Monopoly or Discouraging Competition? The Perils of Price–Cost Tests for Market Power in Electricity. In *Electric Choices: Deregulation and the Future of Electric Power*, edited by Andrew N. Kleit. Lanham, MD: Rowman and Littlefield, 163–79.

Brennan, Timothy. 2008. "Night of the Living Dead" or "Back to the Future"? Electric Utility Decoupling, Reviving Rate-of-Return Regulation, and Energy Efficiency. Discussion Paper 08-27. Washington, DC: Resources for the Future.

Brennan, Timothy, and Karen Palmer. 1998. Not Enough Anarchy in the U.K. *Regulation* 21(2): 2–3.

Cooter, Robert, and Thomas Ulen. 2004. *Law and Economics*. Boston, MA: Pearson Addison Wesley.

Crew, Michael, and Paul Kleindorfer. 1986. *The Economics of Public Utility Regulation*. Cambridge, MA: MIT Press.

Joskow, Paul, and Jean Tirole. 2007. Reliability and Competitive Electricity Markets. *RAND Journal of Economics* 38: 60–84.

Landes, William, and Richard Posner. 1987. *The Economic Structure of Tort Law*. Cambridge, MA: Harvard University Press.

McDonough, Catherine, and Robert Kraus. 2008. Making Hourly Pricing Programs Work Without Scarcity Pricing in the Hourly Day-Ahead Market. Paper presented at the Center for

Research in Regulated Industries, Rutgers University, Advanced Workshop in Regulation and Competition, 27th Annual Eastern Conference. May 15, 2008, Skytop, PA.

Wittman, Donald. 1981. Optimal Pricing of Sequential Inputs: Last Clear Chance, Mitigation of Damages and Related Doctrines in the Law. *Journal of Legal Studies* 10: 65–91.