Econometric Estimates of Scale Economies in Electric Power Generation: The Role of Excess Capacity*

By

Mamiko V. Howorka

September 1981

*Material in this working paper came directly from the dissertation of the same title presented to the Graduate Council of the University of Florida in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The author acknowledges the help of Sandra H. Kirks and the Public Utilities Research Center in the preparation of this report.

The views expressed in this paper do not necessarily represent those of PURC or its sponsors.

In the United States free enterprise economy, competition is generally relied on as a means of achieving good performance, such as keeping prices low and service qualities reasonable. However, in the public utility sector, important economic decisions are guided and controlled, not by competitive pressure, but by government regulation. Regulation, in many ways, restricts rather than supplements competition by imposing restrictions on entry, through licensing (as with utility franchises). Regulation may also intervene in decision-making with respect to price, output, investment, and financing; these decisions would have been made through the competitive market mechanism, without regulation.

The electric power industry contends with substantial regulation by public authorities. The main economic justification for regulating the electric power industry is scale economies, associated with natural monopoly. The technological characteristics of this industry are such that the average costs decrease as output expands, and such scale economies would not be exhausted by a few large firms serving the entire market. Therefore, due to economies of scale, only a single or a limited number of large firm(s) can provide service at least cost; unrestricted entry could merely bring about more expensive and poorer service through wasteful duplication of heavy fixed (or capacity costs) and right-of-ways. Such duplication by potential competitors may also incur higher social costs as their facilities affect adversely the environment. Under these circumstances, unrestrained competition is both undesirable and infeasible. The electric firms are, therefore, under regulatory control in the hope that institutionalization of the firm's monopoly status would enable the firm to increase its level of output and thereby take advantage of its decreasing cost technology. The above argument represents an economic case for granting the electric utilities monopoly status in their local markets.

At the same time, the government controls the firm's pricing policies through rate-of-return regulation to insure that institutionalized monopoly power will not be abused to earn excessive monopoly profits by charging customers high prices. Rate-of-return regulation restricts prices to the level which would just enable the regulated firm to earn a "fair" rate of return on its invested capital base. The firms are also obliged to meet the demand in their franchised areas at the prices consistent with rate-of-return regulation. In brief, the electric utilities are subject to a complicated combination of privileges, public duties and responsibilities placed on them by society.

Yet, even where effective regulation is called for, very few would dispute that reliance on competition can assist in assuring good performance. On the margin, natural monopolies might compete for fringe service areas or for specific customers, since electricity generation scale economies may outweigh distribution and transmission costs for such service area extensions. The implication of such an argument is that proper policies should be directed at preserving some competition. At the same time, regulators would have to balance off efficiency gained by allowing firms to coordinate or integrate enough so as to exploit full potential economies of scale with the advantages of preserving competitive opportunity of viable firms. Economies of scale, in this sense, require that regulators seek the economically optimum combination of competition and regulation as well as the effective form of competition itself.

Weiss (1975) has made an interesting proposal for reorganization of the U.S. electric power industry. He proposed that firms be vertically disintegrated by separating the generation of electricity from its transmission and distribution. The policies based on Weiss' proposal would lead to increased levels of competition at the generation stage. However, whether (as Weiss claims) such an action would result in lowering rates charged for electric power depends largely on the extent of scale economies at the generating stage. If scale economies in electricity generation have been exhausted, then

> . . . . a small number of extremely large firms are not required for efficient production and . . . policies designed to promote competition in electric power generation cannot be faulted in terms of sacrificing economies of scale. (Christensen and Greene, 1976)

If not, the question becomes whether the potential benefits of increased competition offset the inefficiency resulting from forgoing substantial scale economies existing at the generation stage. Therefore, policies encouraging integration or coordination at the generation stage need to be judged on the basis of trade-offs between the additional benefits from coordination and the disadvantages of suppressed competition.

In light of the rising real cost of electric power, econometric estimates of scale economies in electric power generation bear important public policy implications. The purpose of this study is to provide more accurate estimates by modeling the production structure with the translog cost function in which firm's excess capacity is explicitly specified.

Almost all previous studies in this literature have consistently indicated the existence of scale economies throughout the range of observations, thereby supporting the natural monopoly concept as applied to the electric power industry. For the first time, however, scale exhaustion at the firm level was reported in the study of Christensen and Greene. Christensen and Greene (C-G) found no significant scale economies for firms producing more than 19 billion KWH per year. Unfortunately, this C-G study as well as their subsequent study, which has introduced coordination and regional dummy variables into their previous model, has somewhat limited public policy implications for the following reasons:

1. Based on the C-G model, it is not appropriate to fit only one cost function for all observations. C-G have used cross-section data to distinguish scale economies from cost reductions resulting from technological improvement. Their cross-section data, however, contain firms whose plants were built in different years, ranging from the early 1920's to 1970. The difference is almost 50 years. During this period of time, the electric power industry had gone through rapid technical changes. It is not therefore proper to assume the embodiment of the same level of technology in their cross-section data just because the observations were made in the same year. The question is empirical and needs to be tested.

In our study, by defining the "vintage of firm" as the weighted average of the years its plants were built, the equality of each pair of regression equations in various vintage groups will be tested, first with the C-G single-output model and later with our expanded model. The results indicate that, based on the C-G model, the fitted cost functions in the newest and the oldest vintage groups are significantly different from each other. According to the modeling approach that C-G have adopted, this finding necessarily implies that their estimated cost function embodies different levels of technology. Therefore, more than one cost function should have been fitted for these observations. This criticism alone is sufficiently serious to cast doubt on the results obtained by C-G.

2. C-G's results rest on the restrictive assumption of full capacity utilization. This assumption would not be so binding if all firms had the same capacity utilization rate, but this is obviously not the case. Indeed, the capacity utilization rates among electric utilities are considerably different. In this situation, the use of output as a measure of firm size might have distorted their estimates of scale economies.

To avoid this type of distortion, some authors have introduced capacity-utilization variable into their models. Our approach is to introduce firm's unused (or excess) capacity for this purpose.

The use of excess capacity offers an advantage by providing some consideration to the peak load aspect of output, which has been neglected totally in the previous firm-level analyses. Firm's excess capacity may be considered as an explanatory variable of both fixed and variable costs by taking into account the effects of variations in fuel input due to differing capacity utilization rates, since fuel input per KWH decreases as capacity operation approaches the optimal level. Assuming the U-shaped relationship between fuel input and capacity utilization rate, and also between capital charges and installed capacity, we could represent firm's total cost by function which is quadratic in the logarithms of output, excess capacity, and input prices.

Therefore, the functional form we used to describe the production structure of the U.S. electric power generation is the translog cost function in which firm's excess capacity is explicitly specified. Applying the principle of duality, we directly modeled the structure of cost without imposing arbitrary a priori restrictions on the structure of production. Like Christensen and Greene, we introduced several dummy variables, indicating differences among firms in degree of coordination and in location of generating facilities.

In addition, we controlled for plant structure and reliance on nuclear and/or hydroelectric plants, two characteristics not explored by these authors. Special tests were conducted to see whether the level of technology varied across time.

Our results do not support the recent findings of the scale exhaustion at the generation stage, within the range of observations. In sharp contrast with the Christensen and Greene results, we found significant scale economies for almost all firms in their sample. Those 15 firms operating in the region of no significant scale economies are small firms producing less than 8 billion KWH annually.

Reliable estimates of scale economies, however, were obtained at the cost of foregoing the simple method of estimating cost savings associated with various degrees of coordination. The explicit consideration of firm's excess capacity divided cost savings from coordination into two parts: those in the form of reductions in excess capacity, and those resulted from the benefits of coordination other than the reductions in excess capacity.

The Southern and Western regions were found to have cost advantages over North Central firms. The firms in South enjoy roughly a 4 percent cost advantage, and those in West about a 6 to 7 percent cost advantage over the North Central firms. These regions may face favorable fuel mix possibilities, relatively favorable weather affecting load patterns, and a generally lower price level (especially in the South). The firms producing more than half of electricity by outdoor or semi-outdoor had approximately a 3 percent cost advantage over those relying on indoor plants. A similar magnitude of cost advantage was obtained for firms relying on nuclear and/or hydroelectric plants.

Because of our findings about the existing substantial scale economies at the generation stage, the potential benefits from competition (through vertical disintegration) might be offset by higher costs. Our findings support the requirement of large firms for efficient production. Public policy initiatives ought to be guided by more than regression analysis, given the complex cost structures under investigation. Yet, this study does not support radical restructuring of the industry.

The empirical results do not find substantial gains from power pooling. Power pooling without common ownership may have provided less cost savings to the member firms than informal coordination has done to the unaffiliated firms.

Yet, the evaluation of power pooling should be based on the full benefits of pooling. We found that the firms in the South and the West have cost advantage over those in other regions, so power pooling interconnecting the utilities in different regions might yield cost savings. Alternatively, the cost savings for members of formal pools may already be captured by the inclusion of an excess capacity variable, since pools also yield savings through reductions in reserve requirements.

The underlying production technology is very complex, with different regulatory constraints and environmental regulations affecting the various firms. Sorting out all the influences is a very difficult task. We have succeeded in introducing some new variables into the cost structure, using a very general functional form. The advantages of this estimation framework are many, although unanswered questions remain. By focusing on scale economies, we obtained results which stand in stark contrast with those of Christensen and Greene, who found scale economies stopping at 20 billion KWH. Thus, the natural monopoly justification of electric utilities cannot be readily tossed aside.

Power Pooling in Electric Power
Generation*

By

Mamiko V. Howorka

September 1981

*The essence of this report is taken from the author's doctoral dissertation "Econometric Estimates of Scale Economies in Electric Power Generation: The Role of Excess Capacity". The aid of Sandra H. Kirks and the Public Utilities Research Center in the preparation of this report is acknowledged.

The technological and institutional characteristics of the electric power industry create an inherent tendency to decreasing average costs over much of the market. Electric utilities are obliged to meet widely fluctuating demand in their service areas, with the product except for KWH based on pumped storage, produced instantaneously (on demand). Because of this obligation, each individual utility must make a heavy investment in capacity large enough to meet its peak demands as well as to maintain idle capacity, called reserve capacity, in excess of peak requirements. Reserve capacity is maintained for the purpose of assuring a reliable electric service in case of a breakdown in some of the generating units, possible load forecasting error, or maintenance needs. This presence of stochastic supply and demand components is a principal cause of the presence of excess capacity in the electric power industry. Therefore, electric utilities in general reach considerably short of the designed level of capacity operation at which short-run average cost attains its minimum value.

Electric utility operation is considered as a natural monopoly not because of the existence of short-run decreasing costs but of long-run decreasing costs over the entire extent of the market. What it amounts to is that the average costs of larger capacities would be less than those of the smaller if operated at optimum rates.

Technological progress which brought about substantial increases in the size of efficient generators and long-distance transmission at high voltages is the kind that increases the extent of economies of scale. As a result, larger plants will have lower average costs than smaller plants. Such economies of scale imply that monopoly may be the most efficient form of organization for particular areas.

Even though neither heavy fixed costs nor potentially duplicative investments (e.g., right-of-way) necessarily make for natural monopoly, it is these fixed costs that might be wastefully duplicated by competing firms if economies of scale exist in the utilization of these facilities. Duplication therefore becomes inefficient only in the presence of economies of scale.

Kahn identified another source of these potential economies of scale on the demand side. The requirement on electric utilities to maintain the capacity large enough to meet peak demands tends to make it more efficient to supply many customers over a wider region than a few small areas. Coordination across firms serving customers with different peaks can create what one called demand-diversity cost savings. The resulting economies of diversity of demand do not necessarily set a case for monopoly but for geographical integration of operations.

The essence of intercompany coordination in the electric power industry is to exploit these scale economies generally available only in larger systems. The degree of coordination which two or more utilities undertake for the purpose of increased economy and reliability in power system planning or operations can range from very loose agreements for energy transfers to coordinated planning and operations, to completely integrated operations. There are no standardized coordination arrangements as each individual utility has different needs and system design. Although nearly all electric utilities engage in some form of coordination, many coordination arrangements are informal without any contractual obligation. In this study, we define "a power pool" as a coordination arrangement with a contractual obligation.

Breyer and MacAvoy have identified six categories in which cost savings may be attained through operation of a group of generating units and a network of interconnecting transmission lines for services to multiple population centers as if the parts were one system. One of them is central dispatch through which operating costs can be held to a minimum. Central dispatch is a program to select for dispatch the energy from the generators with the lowest marginal operating costs of those available in a power pooled system. Most firms use central dispatch for their own operations, but not for all the member utilities viewed as a single system.

The second saving is a reduction on costs for meeting peak demands by taking advantage of demand diversities over wider geographic areas. If daily or seasonal peaks in two regions do not coincide, the electric utilities in these regions can save generating capacity by exchanging power or using the peaking capacity in one to supply part of the peak demand in the other. That is, neither needs to maintain the full capacity necessary to meet its own peak requirements. At the same time, both can benefit from a higher average utilization of capacity as well as higher load factors. The third saving is a reduction in reserve costs. Reserve generating capacity is maintained in case of a breakdown in some of the generating units, possible underestimation of peak demand and maintenance needs. Coordinating the operations of several systems reduces the need for breakdown reserves by allowing each member firm to call upon the reserves of other members if a generator outage occurs. The capacity needed for maintenance needs is also reduced by staggering their maintenance schedules to allow each member to use the same spare generators to substitute for the risk of underforecasting demand since unexpected changes in demand faced by one firm can be absorbed by reserve capacity of other members.

A reduction on generating costs is the fourth saving. Here, utilities take advantage of economies of scale in generator size. Even though an individual firm may not face sufficient demand to justify an installation of a large generator, several firms combined into a single system may plan jointly and take turns in building very large units and sell the surplus among the members.

There is also a possible reduction in transmission reliability costs through coordinated planning across a wide geographical area. Finally, social costs may be reduced by coordination. Joint planning can lower environmental costs by reducing total construction and locating the plants in the areas involving lower social costs.

However, voluntary coordination is likely to be incomplete and therefore to fall considerably short of exploiting fully these possible advantages of optimal coordination or those of complete integration where financial consolidation promises additional efficiencies. There is no straightforward explanation for the failure of electric utility companies to cooperate to the fullest extent for optimum performance. The following may be among the most important obstacles.

The resistance to thoroughly integrated power pooling may be partially explained by the difficulty in dividing the potential gains from coordination among the member companies which are, in some degree, competitive and at the same time highly regulated. Although direct competition among electric utilities is quite limited, there are certain areas and kinds of competition between electricity companies to induce industrial customers to locate in their territories, or competition at the edges of service areas to determine which company is to take over an overlapping or adjacent market. There is also competition in wholesale power market and so-called yardstick competition between private and non-private utility systems. The latter generally involves the comparison of performances of private and non-private systems which may be used to make certain political decisions, such as which type of utility system is to be certificated for future service or to serve present or expanding needs. Comparisons of this sort, however, are not fair because compared with private utilities, the public and cooperatives pay less taxes and obtain their capital on more favorable terms; on the other hand, many of them face offsetting disadvantages from the smaller scale of their generating units as well as their dependence on private companies for much of their energy.

Power pools are arrangements made by two or more utility companies in search of greater economy and reliability, each with something to contribute. Certain complications arise from the fact that participants benefit to different degrees from pooling. The smaller companies would benefit most since they are the farthest from achieving the full economies of scale by themselves. Larger companies which have already achieved substantial economies than to pool with small firms, which may thereby become their potential rivals. The large firms could lose markets through increased competition, the cost of which may well exceed the cost savings the large firms might obtain from pooling.

As in the case of sharing the gains, there is also great difficulty in apportioning certain costs. One such example may be the costs of strengthening the transmission system from which all members benefit, but in different and indeterminate amounts. Since the amount of benefit each company derives cannot be easily determined, it would be difficult for them to agree on apportionment. State regulatory authorities would certainly be reluctant to allow companies to contribute the cost of building a transmission wire in some other states unless it will obviously benefit the people in its own state.

Management problems in pooling constitute another type of obstacle. While there is a difficulty in decision making without superior authority with the power to resolve disagreements, establishment of such power by reducing the representation of the small firms in committee would make a pool of separate companies more like a single holding company.

Pools also face more difficulty in planning and operation than a single company does. Since pools are composed by independent firms, member companies are unlikely to operate the whole system most efficiently against their own objectives. For planning it could matter whose rate base was expanded, since a larger rate base means greater total profits allowed to earn.

This obstacle as well as those created by potential competition of members and managerial conflicts might be removed by financial integration. If individual companies could financially integrate each would find it more consistent with its own interest to perform each function "with the least-cost combination of media." We therefore expect that the available economies of scale have been more successfully realized by financially integrated firms than by a pool of independent firms.