Econometric Benchmarking of Cost Performance: The Case of U.S. Power Distributors

Mark Newton Lowry[*], *Lullit Getachew*[**], *and David Hovde*[**]

Benchmarking of cost efficiency has growing use in energy utility regulation. The state of the art has been limited in many countries by the small size of available national data sets and poor data on capital cost. Data available in the United States place fewer constraints on benchmarking methods. This paper develops an econometric cost benchmarking model for power distribution that is based on U.S. data. The model can address total cost and its major components. Numerous cost drivers are identified. Statistical tests of efficiency hypotheses are performed. The cost performances of utilities are compared to the industry norm. The suitability of the alternative frontier standard in regulatory applications is discussed.

1. INTRODUCTION

Benchmarking has in recent years become a widely used tool in the assessment of energy utility performance. Managers use benchmarking studies to assess how well their companies are doing. Benchmarking is also used in the regulatory arena to help establish utility rates. In recent years, benchmarking has played an important role in ratemaking in Australia, Canada, Europe (e.g. Britain and Norway), and Latin America (e.g. Bolivia and Panama). Studies have been presented in U.S. rate proceedings but rarely at the initiative of regulators.

Benchmarking of utility performance for regulation requires accurate cost evaluations and such appraisals are challenging. There are important differences among companies in business conditions that influence cost. It is difficult to establish benchmarks that properly control for such conditions even with abundant and high quality data. The data sets available for utility benchmarking

e Energy Journal, Vol. 26, No. 3.

* Corresponding author, Partner, Pacific Economics Group, 22 E. Mifflin Street, Suite 302, Madison, WI 53705, USA. Email: mnlowry@earthlink.net.

** Senior Economists, Pacific Economics Group, 22 E. Mifflin Street, Suite 302, Madison, WI 53705, USA.

are, however, often quite limited. In most countries, data are available for only a few years and for less than thirty utilities in a given industry. Few countries have gathered the data needed for accurate capital cost benchmarking. Yet capital accounts for more than half of the cost of most "wire" and "pipe" utilities. International benchmarking can produce more sizable data sets but has been hampered by a lack of data standardization and by differences in the activities that utilities perform.

These data challenges have had important consequences for the benchmarking methods used in regulation. In most countries, benchmarking has focused on operation and maintenance (O&M) expenses rather than total cost performance. Benchmarking techniques have been favored that require few data. For example, econometric cost models employed have generally consisted of O&M expenditure functions with simple functional forms and few explanatory variables.[1] The studies rarely consider the statistical significance of benchmarking results. A finding that a utility's cost performance is significantly inferior to a given benchmark is, after all, less likely the smaller is the sample.

The United States is one country where data challenges have not greatly restricted utility benchmarking methods. For several utility industries there, good data are publicly available for many companies for periods of more than ten years. The data include those needed for rigorous benchmarking of capital costs. The favorable data environment has encouraged the estimation of more complex econometric cost benchmarking models and statistical tests of efficiency hypotheses. Studies frequently address total cost as well as O&M expenses.

Another noteworthy feature of many U.S. benchmarking studies has been their use of an average industry efficiency standard.[2] Studies employing this standard measure the extent to which subject utilities have cost performance above or below the apparent industry norm. A frontier performance standard has been more commonly used in benchmarking studies prepared for European regulators.[3] Studies employing this standard measure the extent to which subject utilities have failed to reach the apparent cost efficiency frontier. Benchmarking methods available for making frontier comparisons include corrected ordinary least squares (COLS), stochastic frontier analysis (SFA), and data envelopment analysis (DEA).[4]

1. Examples include the 1997/98 and 2004/05 updates of the price controls for distribution network operators in Britain. Both reviews employed O&M expenditure functions featuring only a composite output variable.

2. Benchmarking studies based on an average industry efficiency standard are also encountered in Australian and Canadian regulation.

3. Jamasb and Pollitt (2001) provide a useful survey of the use of frontier benchmarking studies in regulation.

4. A good introductory discussion of frontier benchmarking methods is Coelli et al (1998). Seminal early works in the frontier benchmarking literature include Meeusen and van den Broek (1977), Aigner, Lovell and Schmidt (1977) and Schmidt and Sickles (1984). More recent studies include Hattori (2002), Hattori, Jamasb and Pollitt (2002), Huiebert (2002) and Burns and Weyman-Jones (1996).

This paper presents an econometric benchmarking study using U.S. data on the cost of power distribution. The study employs a sample of data for sixty-six U.S. distributors spanning twelve years. The model can evaluate O&M expenses, capital cost, and total cost.[5] An average cost standard is employed and statistical tests of efficiency hypotheses are undertaken.

Model specification was aided by previous research on the cost structure of power distribution. The seminal article in the field is Neuberg (1977). Noteworthy recent contributions include Hjalmarsson and Veiderpass (1992), Salvanes and Tjotta (1994) and Yatchew (2000).

The balance of the paper is divided into six sections. We first discuss the use of average and frontier efficiency standards in benchmarking in Section 2. We present the data used in Section 3 and the benchmarking methodology in Section 4. The results are discussed in Section 5. Concluding comments appear in Section 6.

2. EFFICIENCY STANDARDS FOR BENCHMARKING

In considering an appropriate efficiency standard for use in benchmarking it is useful to start by enunciating some basic criteria for selecting benchmarking methods used in regulation. Two such criteria are accuracy and fairness. Considerations of both suggest that an average industry efficiency standard is a worthy alternative to a frontier standard.

With regard to accuracy, consider first that there is currently no effective way to identify the sustainable minimum cost of utility service. At each point in time several utilities in a sample used for benchmarking will likely incur costs that are below the sustainable minimum. A power distributor may, for example, postpone tree trimming costs that are ultimately quite essential to the maintenance of satisfactory service quality. Existing frontier benchmarking methods estimate the distance from the unsustainable cost frontier and are therefore inherently biased in measurement of the distance from the more relevant long run sustainable frontier. This problem is not encountered with an average industry standard.[6]

The accuracy of frontier methods is also limited by the current state of the art. For example, rigorous econometric research on total cost and its major components, capital and O&M expenses, is commonly undertaken using multiple equation systems that are estimated by methods that control for cross equation correlation. It is also desirable to use econometric methods in cost research that correct for heteroscedasticity. SFA estimation procedures that can estimate the parameters of multiple equation cost models, control for cross-equation correlation

5. Assessments of O&M expenses have a long-run character since they do not fully consider how much capital the distributors utilize.

6. Yatchew (2001), in discussing how best to implement benchmarking in regulation, points to similar difficulties in obtaining estimates of "best practices" since they are variable. In addition, he notes that methods that estimate best practices suffer the most from outliers whereas those that estimate the average are less susceptible to them.

and heteroscedasticity are not yet, to our knowledge, readily available. However, these procedures have been developed by the authors for benchmarking using the alternative industry norm standard.

The fairness of a benchmarking method can be defined as its consistency with generally accepted standards for the distribution of the benefits of market activities. The competitive market standard is compelling in this regard. In competitive markets, firms with superior performance earn above average returns. This is true even in the long run.[7] If regulation is to emulate the operation and outcomes of a competitive market, companies with markedly superior performance must therefore be allowed to earn rates of return above the competitive norm. If the industry's best-observed practice is imposed on all firms, any firm that fails to achieve this standard will earn below average returns. This would be true even for superior performers that nevertheless fall short of the industry's best performance.

Data from more competitive industries can shed light on these issues. For example, the authors have surveyed frontier benchmarking studies in agriculture and banking and have found that the typical firm in such industries is about 20% below frontier efficiency. These distance estimates may reflect both the inefficiency of typical firms and the difference between sustainable and unsustainable minimum cost.

3. DATA

Our study is based on a sample of data for U.S. power distributors spanning the period 1991 to 2002. The primary source of the cost and quantity data used in the study is the Federal Energy Regulatory Commission (FERC) Form 1. Major investor-owned electric utilities (IOUs) in the United States are required by law to file this form annually. Data reported on Form 1 must conform to the FERC's Uniform System of Accounts. Many respondents are vertically integrated utilities that also own generation and transmission facilities. However, all are required to separately itemize their distribution costs.

All major U.S. electric IOUs were included in the sample that filed the Form 1 electronically in 2002; have reported, together with any important predecessor companies, the necessary data continuously since they achieved a

7. There are both short-run and long-run equilibria in competitive markets. In the short run, equilibrium occurs whenever quantity supplied equals quantity demanded. But the industry will not be in long-run equilibrium if average returns in the industry are not equal to the competitive rate of return, defined to be the opportunity cost of capital. For example, if average industry returns exceed the competitive rate of return, long-run equilibrium is established as new firms enter the industry and existing firms expand their production, thereby increasing supply and driving down prices and average returns. This process continues until the industry's average return equals the competitive rate of return. For evidence that superior performers continue to earn above-average returns even in the long run, see L. Schwalbach, U. Grabhoff, and T. Mahmood, "The Dynamics of Corporate Profits," *European Economic Review*, October 1989, 1625-1639.

"major" designation; and submitted plausible data in the periods required. Data from sixty-six companies met all of these standards and were used in the study.[8]

Power distribution services are defined to include the local delivery, customer account, sales, and information services provided by distributors. We do not address the costs that they incur for power procurement. The total cost of distribution services thus defined comprises the costs of plant ownership, operation and maintenance.

Our benchmarking method involves the decomposition of cost into three input categories: capital services, labor services, and non-labor O&M inputs. The cost of labor is defined as the sum of O&M salaries and wages, pensions and other employee benefits. The cost of other O&M inputs is defined as assigned O&M expenses net of these labor costs. This input category includes the services of contract workers, insurance, real estate rents, equipment leases, and miscellaneous materials.

The study uses a service price approach to measuring the cost of plant ownership that is based on the economic value of utility plant.[9] Under this approach, the cost of capital is the product of capital price and quantity indexes. This method controls for differences between utilities in the age of their investments.

The capital price index ($WKS_{h,t}$) that we employ is one appropriate for capital services in a competitive rental market. Its formula is:

$$WKS_{h,t} = \mathrm{d}*WKA_{h,t} + WKA_{h,t-1}\,[r_t - (WKA_{h,t} - WKA_{h,t-1})/WKA_{h,t-1}] \quad (1)$$

where for each firm h in year t, $WKA_{h,t}$ is the capital asset price index[10], r_t is the cost of funds, and d is the depreciation rate, which is assumed constant.

The first term in this expression corresponds to the cost of depreciation. A geometric decay treatment of depreciation is used. The second term corresponds to the difference between opportunity cost and capital gains. The term in brackets is the real rate of return on capital. This term is smoothed to reduce capital cost volatility. The cost of capital normally includes certain tax expenses. However, we have chosen to exclude taxes due in part to their volatility.

The capital quantity index that we employ ($XK_{h,t}$) is based on the following perpetual inventory method:

$$XK_{h,t} = (1-d) * XK_{h,t-1} + \frac{VI_{h,t}}{WKA_{h,t-1}}. \quad (2)$$

8. Other sources of data are also used primarily to measure input prices. The sources are Whitman, Requardt & Associates; R.S. Means and Associates; the Bureau of Economic Analysis (BEA) of the U.S. Department of Commerce; the Bureau of Labor Statistics (BLS) of the Department of Labor; and DRI-WEFA.

9. See Hall and Jorgensen (1967) for a seminal discussion of the service price method of capital cost measurement.

10. These data are reported in the *Handy-Whitman Index of Public Utility Construction Costs*, a publication of Whitman, Requardt and Associates.

Here, $VI_{h,t}$ is the value of gross additions to utility plant.

The explanatory variables used in the cost model comprise three measures of output, three input prices, and seven variables that represent miscellaneous other business conditions. The latter group of variables may usefully be called "Z variables". The three output quantity variables are the number of retail customers, the volume of power deliveries to such customers and the line miles in a utility's service territory[11]. The input prices are for labor, non-labor O&M input, and capital.

The Z variables included in the model are the number of gas customers served, the percentage of line miles overhead, average precipitation, a measure of system age, the value of transmission and generation plant, the percentage of deliveries that are made to residential and commercial customers, and average temperature. These variables are discussed further in the results section. The model also contains a trend variable. It permits predicted cost to shift over time for reasons other than changes in the specified business conditions. Table 1 presents the average values of cost model business conditions over the 2000-2002 period.

Table 1. Average Values of Variables in the Benchmarking Study

Variable	Units	U.S. Sample Average
Power Delivery Cost	Dollars	388,187,880
Number of Customers	Count	755,347
Retail Deliveries	MWh	19,683,942
Line Miles	Miles	20,332
Price of Capital Services	Index Number	14.10
Price of Labor Services	Dollars / Year	37,870
Price of Materials	Index Number	102.11
Number of Gas Customers Served	Count	179,743
% of Line Miles Overhead	Percent	77%
Average Precipitation	Inches/Year	36.35
Ten Year Customer Growth	Percent	13%
Transmission and Generation Plant	Dollars	2,169,649,000
% of Deliveries Residential and Commercial	Percent	69%
Temperature	°F	54.16

11. Line miles may also reasonably be viewed as a network variable. With either interpretation, the intention is to capture the cost impact of system extensiveness.

4. BENCHMARKING METHODOLOGY

4.1 Cost Model

We estimate performance relative to the average by using the dual representation of production technology. A simplified version of the dual cost function for a panel data set is:

$$C_{ht} = \beta_0 = X_{ht}\beta + \varepsilon_{ht} \qquad h=1\ldots n,\ t = 1\ldots T. \tag{3}$$

Here for each firm h in year t, C_{ht} is total cost, X_{ht} is a vector of explanatory variables, the β term represents model parameters and ε_{ht} is an error term. The term ε_{ht} embodies a firm specific measure of inefficiency, α_h, and random noise, η_{ht}:

$$\varepsilon_{ht} = \alpha_h + \eta_{ht} \qquad h=1\ldots n,\ t = 1\ldots T. \tag{4}$$

The model that is estimated is:

$$\begin{aligned} C_{ht} &= \beta_0 + \overline{\alpha} + X_{ht}\beta + \alpha_h - \overline{\alpha} + \eta_{ht} \qquad h=1\ldots n,\ t = 1\ldots T. \\ &= \beta_0^* + X_{ht}\beta + \varepsilon_{ht}^* \\ \varepsilon_{ht}^* &= \alpha_h - \overline{\alpha} + \eta_{ht} \end{aligned} \tag{5}$$

We assume a well behaved random noise with $E(\eta_{ht}) = 0$ and $E(\varepsilon_{ht}^* \mid X_{ht}) = 0$ Thus, the expected difference between the predicted and actual cost of the average firm equals zero. Using parameter estimates of the model, we get an estimate of each firm's efficiency as follows[12]:

$$C_{ht} - \hat{C}_{ht} = \hat{\varepsilon}_{ht}^* = \hat{\alpha}_h - \hat{\overline{\alpha}} + \hat{\eta}_{ht} \qquad h=1\ldots n,\ t = 1\ldots T \tag{6}$$

$$E(C_{ht} - \hat{C}_{ht}) \approx \frac{1}{T}\sum_{t=1}^{T}(C_{ht} - \hat{C}_{ht}) = \hat{\alpha}_h - \hat{\overline{\alpha}} \qquad h=1\ldots n,\ t = 1\ldots T$$

4.2 Model Specification

The functional form selected for this study is the translog. This form has considerable flexibility and is widely used in econometric cost research

12. This formulation is inspired by Afriat (1972) and Richmond (1974) who suggested a modified version of COLS. Their model is based on a production function for cross sectional data, but can be applied to a cost function as follows. One can estimate the parameters $C_i = \beta_0 + X_i\beta + \varepsilon_i$ by OLS. Then, we obtain $\hat{\beta}_0^* = \hat{\beta}_0 + E(\hat{\varepsilon}_i)$ and $\hat{\varepsilon}_i^* = \hat{\varepsilon}_i - E(\hat{\varepsilon}_i)$. Unlike the COLS case, we can think of this version of inefficiency estimation as measuring performance relative to the average rather than the frontier.

when a sample of adequate size is available.[13] The general form of the translog econometric cost model, where time and firm subscripts have been suppressed for simplicity, is:

$$\begin{aligned}\text{In } C = \alpha_0 + \sum_i \alpha_i \ 1\text{n } Y_i + \sum_j \alpha_j \ 1\text{n } W_j \\ + \ \frac{1}{2}\left[\sum_i \sum_k \gamma_{ik} \ 1\text{n } Y_i \ 1\text{n } Y_k + \sum_j \sum_n \gamma_{jn} \ 1\text{n } W_j \ 1\text{n } W_n\right] \\ \sum_i \sum_j \gamma_{jh} \ 1\text{n } W_j \ 1\text{n } Y_h + \sum_l \ \alpha_i \ 1\text{n } Z_l + \alpha_i T + \varepsilon.\end{aligned} \tag{7}$$

Here Y_i denotes one of M variables that quantify output and W_j denotes one of N input prices. In addition, Z_l denotes one of the L additional business conditions, T is a trend variable, and ε denotes the error term. Notice that, to preserve degrees of freedom, we have not interacted the Z variables with the other variables.[14]

Yatchew (2000), in his study of power distribution cost, finds that the measure of efficiency is not affected by homothetic and log linear representations, with nonparametric scale effect, of the technology. We will apply our measure of inefficiency to these and the homogeneous representations of production technology to determine the effect functional representation has on efficiency.

Benchmarks for capital costs and O&M expenses can be obtained by augmenting the cost equation with the set of cost share equations implied by Shepard's Lemma. Simultaneous estimation of these equations and the cost function can, furthermore, enhance the efficiency of parameter estimates. The general form of a cost share equation for a representative input price category, *j*, can be written as:

$$S_j = \alpha_j + \sum_i \gamma_{jh} \ 1\text{n } Y_i + \sum_n \gamma_{jn} \ 1\text{n } W_n. \tag{8}$$

4.2 Estimation Methods

It is well known that if there exists contemporaneous correlation between the errors in the system of regression equations, more efficient estimates can be obtained by using a Feasible Generalized Least Squares (FGLS) approach. [15] It is also known that more efficient estimators can be obtained by iterating this procedure to convergence.[16] Since we estimate the unknown disturbance matrices

13. For further discussion of the translog functional form see Guilkey et. al. (1983) and Gagne and Ouellette (1998).

14. Interaction of the Z variables with the input prices is a sensible enhancement when O&M expenses are the focus of research.

15. See Zellner, A. (1962).

16. That is, we iterate the procedure until the determinant of the difference between any two consecutive estimated disturbance matrices are approximately zero.

consistently, the estimators we eventually compute are equivalent to Maximum Likelihood Estimates (MLE).[17]

The firms in our panel dataset are characterized by varying scales of operation, which suggests that we relax the assumption of constant error variance across the N groups. The presence of group-wise heteroscedasticity leads to biased standard errors and, hence, incorrect inference. Our estimation procedure addresses this problem and thereby produces parameter and efficiency estimates that are consistent and efficient.

We also undertake statistical tests of efficiency hypothesis. In assessing performance relative to the average, it is desirable to test the hypothesis that a utility is not an average performer. A conclusion of superior or inferior cost performance, relative to the average, can be reached if this hypothesis is rejected at a designated level of confidence.

To form confidence intervals for our statistical test of efficiency hypothesis, we drop the target utility when estimating the model used to do the predictions. This allows us to use standard errors for out of sample predictions to test whether a utility's efficiency score is statistically significantly different from the average. The confidence interval is wider for utilities that are further away from the average in the sample. The standard errors can also be used to compute t-statistics for tests of efficiency hypotheses.

5. EMPIRICAL RESULTS

5.1 Parameter Estimates

We estimate four specifications of our benchmarking cost model. These are the translog (model 1), the homothetic (model 2), the homogeneous (model 3), and the log-linear (model 4) specifications. The production technology is restricted to be homothetic if the cost function can be written as a separable function of factor prices and output.

The homothetic restriction implies that $\gamma_{jh} = 0$, $\forall j$ and h, so that all factor price and output interaction terms drop out of the cost function. In this case, the slopes of the isoquants are preserved along every ray from the origin and returns to scale are unaffected by factor prices.

The production technology is made homogeneous by imposing the restriction that $\gamma_{ik} = 0$, $\forall i$ and k. In this case, returns to scale are unaffected by increases in output and the unit cost cannot take a u-shaped form. The log-linear model requires that all second order terms drop out of the cost model.

Table 2 presents parameter estimates from all four models. Since estimation was done on data that were mean scaled and logged, the parameter estimates for the first order terms are estimates of the overall elasticity of the variables at the sample mean. These estimates are generally plausible as to sign

17. See Dhrymes (1971), Oberhofer and Kmenta (1974), Magnus (1978).

and magnitude. In models 1 and 2, all parameter estimates of first order terms are, additionally, significant at the 10% level.

Parameter estimates of the prices reflect their shares in total cost. For example, the estimate of the price of capital, α_k, is about 0.60 in all models reflecting the capital intensive nature of the power distribution business. In all four models, the parameter estimates on the number of customers, α_{yn}, show this variable to be the dominant output-related cost driver. The effect on cost of retail deliveries and line miles is in each case about one-third that of that of customer numbers for the average firm. The use of line miles as a variable, together with the number of customers, allows us to account for the impact of customer density on cost. It is generally more costly to serve a less dense service territory. The positive parameter estimate for line miles reflects this since it indicates that holding the number of customers constant, the higher the line miles the lower the density of the service territory and the higher the cost. The parameter estimate on the trend variable, α_t, reflects a modest average downward shift in cost of 0.6% to 0.8% over time.

The coefficients on the additional business conditions are also sensible in all four models. The parameter estimate on the number of customers receiving gas distribution service, α_{z1}, reflects the cost impact of diversification into gas distribution. The negative estimate reflects cost reduction from scope economies.

The second Z variable is the percentage of line miles that is not underground. Underground lines provide a higher quality of service than overhead lines but are also more costly. The negative parameter estimate on the percentage of line miles not underground, α_{z2}, reflects that cost is higher the greater the undergrounding.

The third Z variable is the average precipitation in the service territory. This serves as a proxy for forestation, which raises O&M cost due to tree trimming and other special maintenance activities. The positive parameter estimate on α_{z3} reflects this.

Accurate benchmarking of the total cost of power delivery, which is a capital intensive business, requires consideration of the age of the distribution system.[18] We generally expect a younger system to have higher capital cost but lower O&M expenses. The effect on total cost is unclear. We proxy system age by computing the share of the total number of customers served that have been added in the last ten years, $\frac{N_t - N_{t-10}}{N_t}$, for each year of the sample period. The negative parameter estimate on α_{z4}, suggests that the newer the distribution system, the lower is total cost.

The fifth Z variable is the value of transmission and generation plant. The negative parameter estimate of this variable indicates that there is a systematic difference in the distribution cost of specialized distributors and of vertically integrated utilities. Utilities that engage in generation and transmission have lower distribution cost.

18. Yatchew (2001) broke ground in this area by accounting for the impact of system age on total distribution cost per customer.

Table 2. Parameter Estimates

	Parameter Estimates							
coefficients	**Model 1**		**Model 2**		**Model 3**		**Model 4**	
α_0	15.074	(0.013)	15.045	(0.013)	15.039	(0.013)	15.054	(0.012)
α_{yn}	0.549	(0.024)	0.570	(0.023)	0.602	(0.023)	0.574	(0.022)
α_{yv}	0.214	(0.025)	0.194	(0.026)	0.179	(0.025)	0.221	(0.024)
α_{yl}	0.225	(0.017)	0.220	(0.016)	0.218	(0.015)	0.226	(0.014)
α_l	0.171	(0.002)	0.179	(0.002)	0.179	(0.002)	0.189	(0.002)
α_k	0.598	(0.002)	0.594	(0.002)	0.594	(0.002)	0.590	(0.002)
α_m	0.231	(0.002)	0.227	(0.002)	0.227	(0.002)	0.221	(0.002)
γ_{ynyn}	0.848	(0.100)	0.812	(0.100)				
γ_{yvyv}	0.908	(0.096)	1.058	(0.099)				
γ_{ylyl}	0.038	(0.063)	-0.043	(0.061)				
γ_{ynyv}	-0.800	(0.088)	-0.935	(0.088)				
γ_{ynyl}	0.028	(0.056)	0.166	(0.055)				
γ_{yvyl}	-0.159	(0.057)	-0.176	(0.057)				
γ_{ll}	-0.027	(0.022)	-0.080	(0.021)	-0.082	(0.021)		
γ_{kk}	-0.016	(0.024)	-0.065	(0.025)	-0.074	(0.026)		
γ_{mm}	0.043	(0.036)	0.146	(0.037)	0.155	(0.038)		
γ_{lk}	0.008	(0.017)	0.012	(0.018)	0.020	(0.019)		
γ_{lm}	0.019	(0.018)	0.068	(0.016)	0.062	(0.016)		
γ_{km}	0.008	(0.019)	0.053	(0.016)	0.054	(0.019)		
γ_{lyn}	0.011	(0.007)						
γ_{lyv}	-0.035	(0.006)						
γ_{lyl}	0.004	(0.005)						
γ_{kyn}	-0.065	(0.007)						
γ_{kyv}	0.077	(0.007)						
γ_{kyl}	-0.001	(0.005)						
γ_{myn}	0.054	(0.008)						
γ_{myv}	-0.042	(0.007)						
γ_{myl}	-0.003	(0.006)						
α_{z1}	-0.005	(0.001)	-0.008	(0.001)	-0.011	(0.001)	-0.010	(0.001)
α_{z2}	-0.138	(0.025)	-0.096	(0.025)	-0.285	(0.023)	-0.247	(0.023)
α_{z3}	0.145	(0.013)	0.122	(0.013)	0.134	(0.013)	0.153	(0.012)
α_{z4}	-0.023	(0.007)	-0.021	(0.007)	-0.018	(0.008)	-0.007	(0.008)
α_{z5}	-0.019	(0.009)	-0.016	(0.008)	-0.012	(0.008)	-0.013	(0.008)
α_{z6}	0.433	(0.034)	0.442	(0.035)	0.194	(0.031)	0.242	(0.032)
α_{z7}	-0.103	(0.047)	-0.111	(0.049)	-0.175	(0.053)	-0.191	(0.052)
α_t	-0.008	(0.002)	-0.008	(0.002)	-0.006	(0.002)	-0.008	(0.002)
system-R^2	0.985		0.982		0.976		0.975	

Standard errors are in parentheses

The sixth Z variable accounts for the special impact of delivering power to residential and commercial customers. The positive parameter estimate indicates that it is more costly to serve these customers. This finding reflects, in part, the fact that industrial customers often obtain fewer distribution services from distributors.

The seventh Z variable is the average temperature over the twelve year period that prevailed in each distributor's service territory. The negative parameter estimate of this variable suggests that, all else equal, utilities in colder regions bear higher cost. Ice storms are likely to be one contributing factor.

5.2. Rankings

Our estimates are based on a panel covering the period 1991-2002. Results from these estimates can be used to determine efficiency of the firms in any given year or over a period of years. Since utilities plan their systems for expected business conditions over a series of years and some explanatory variables are volatile, cost benchmarking should be undertaken over a multiyear timeframe. We choose a three-year period as one providing a reasonable balance between the need for multiyear analysis and contemporary relevance. Thus, we use estimated residuals from the 2000-2002 period to determine efficiency measures.

In addition, we are interested to know whether different functional forms affect efficiency measures and the ranking of firms by efficiency. We compare the translog cost function, which requires extensive data for accurate parameter estimation, and the other three specifications of our cost model discussed earlier. We present efficiency scores and firm rankings from the four models in Table 3.

The homothetic specification produces rankings quite similar to the full translog model, while the homogeneous and log-linear specifications produce rankings that are substantially different. For the first firm, for example, the efficiency score from model 1 shows that its cost is 11.1% above the average, while that from model 2 shows its cost to be 10.9% above the average. Models 3 and 4 show it to have cost that is 3.5% above and 2.8% below the average, respectively.

Table 4 presents Spearman rank correlation coefficients for the rankings of the four models. It can be seen that the correlation coefficients between model 1 and models 3 and 4 are quite a bit lower than that between models 1 and 2.

To select one of the models for benchmarking purposes, we test the full model against the homothetic, the homogeneous and the log-linear models. For this purpose we use the likelihood ratio statistic, which has a χ^2_J distribution with J degrees of freedom, where J is the number of restrictions. The test results in Table 5 show that we can reject all three models and use the results of the full translog model for benchmarking. The results in tables 3-5 suggest that it is desirable to use benchmarking models with flexible functional forms.

Table 3. Scores and Rankings of the Models

	Model 1		Model 2		Model 3		Model 4	
Utility_id	**Score**	**Rank**	**Score**	**Rank**	**Score**	**Rank**	**Score**	**Rank**
2	0.111	54	0.109	50	0.035	38	-0.028	33
9	-0.031	33	-0.035	29	-0.060	26	-0.100	22
12	0.084	48	0.099	49	0.131	51	0.113	52
13	0.052	44	0.155	57	0.105	47	0.072	43
14	0.031	40	-0.035	30	0.075	43	0.108	50
17	0.070	45	0.075	44	0.106	48	0.079	45
21	-0.187	11	-0.179	10	-0.100	22	-0.096	23
22	0.106	52	0.136	54	0.122	50	0.111	51
23	-0.036	29	-0.059	23	0.060	41	0.070	42
25	-0.083	22	-0.057	24	0.142	54	0.168	57
27	-0.035	30	0.004	36	0.038	39	0.007	38
29	0.046	43	0.083	45	0.119	49	0.106	49
30	-0.193	10	-0.193	8	-0.283	4	-0.269	7
31	-0.310	4	-0.310	4	-0.312	1	-0.337	2
36	-0.015	35	-0.018	33	0.004	35	-0.022	34
47	0.349	64	0.332	65	0.197	60	0.137	53
50	-0.176	12	-0.093	19	-0.273	7	-0.216	11
53	0.026	39	0.033	40	0.045	40	0.030	39
62	-0.327	2	-0.337	2	-0.208	12	-0.298	4
63	-0.066	24	-0.042	26	-0.019	34	-0.069	29
67	0.094	50	0.087	47	-0.046	29	-0.017	36
73	-0.059	26	-0.067	22	-0.052	28	-0.071	27
89	-0.005	37	-0.016	34	0.004	36	-0.021	35
91	-0.006	36	0.031	39	0.082	44	0.069	41
92	-0.322	3	-0.340	1	-0.274	6	-0.313	3
93	-0.064	25	-0.090	20	-0.109	19	-0.094	24
98	-0.295	5	-0.274	5	-0.286	2	-0.275	6
99	-0.081	23	-0.038	28	-0.224	10	-0.201	13
101	0.042	42	0.047	43	0.034	37	0.097	48
109	0.107	53	0.091	48	0.292	62	0.285	62
110	0.187	61	0.157	58	0.330	63	0.322	64
119	-0.001	38	0.003	35	-0.262	8	-0.224	10
130	0.198	62	0.226	63	0.194	59	0.202	61
131	-0.035	31	0.010	37	-0.101	21	-0.116	20
133	-0.209	8	-0.201	7	-0.078	24	-0.062	32
135	-0.171	13	-0.153	12	-0.213	11	-0.228	9
136	-0.214	7	-0.221	6	-0.194	14	-0.199	14
138	-0.040	27	-0.041	27	-0.037	33	-0.066	31
140	0.144	57	0.159	59	0.169	57	0.182	60
141	0.103	51	0.132	53	0.095	45	0.093	46
142	0.365	65	0.399	66	0.347	64	0.303	63
149	-0.040	28	-0.025	32	-0.039	30	-0.066	30

continued

Table 3. Scores and Rankings of the Models (continued)

	Model 1		Model 2		Model 3		Model 4	
Utility_id	Score	Rank	Score	Rank	Score	Rank	Score	Rank
150	0.369	66	0.332	64	0.372	66	0.390	66
152	-0.022	34	-0.035	31	-0.038	31	-0.070	28
153	-0.032	32	0.014	38	0.067	42	0.074	44
154	-0.086	19	-0.147	15	-0.087	23	-0.085	26
156	-0.149	15	-0.162	11	-0.107	20	-0.132	19
157	-0.209	9	-0.148	14	-0.282	5	-0.282	5
159	0.131	56	0.163	60	0.150	55	0.161	56
163	0.185	60	0.208	61	0.360	65	0.373	65
167	0.093	49	0.120	52	0.207	61	0.174	58
169	-0.093	17	-0.113	17	-0.285	3	-0.342	1
171	-0.374	1	-0.318	3	-0.243	9	-0.239	8
172	0.119	55	0.115	51	0.167	56	0.137	54
178	0.079	46	0.086	46	0.132	52	0.095	47
180	0.081	47	0.039	42	-0.057	27	-0.107	21
181	0.163	58	0.141	55	0.170	58	0.182	59
182	-0.101	16	-0.121	16	-0.163	16	-0.152	15
183	-0.158	14	-0.150	13	-0.204	13	-0.151	16
185	-0.083	21	-0.087	21	-0.152	17	-0.150	17
186	0.229	63	0.218	62	0.142	53	0.158	55
196	-0.090	18	-0.095	18	-0.120	18	-0.138	18
198	0.037	41	0.039	41	-0.037	32	-0.009	37
201	0.164	59	0.144	56	0.101	46	0.068	40
202	-0.083	20	-0.043	25	-0.074	25	-0.092	25
203	-0.242	6	-0.187	9	-0.192	15	-0.212	12

Table 4. Spearman Rank Correlation Matrix, 1999-2001 Period

Model	Model 1	Model 2	Model 3	Model 4
Model 1	1.00000			
Model 2	0.98146	1.00000		
Model 3	0.87930	0.87696	1.00000	
Model 4	0.86607	0.86106	0.98297	1.00000

5.3 Use of Results in Regulation

To illustrate the application of benchmarking results in regulation, we assume that they are being used in the context of North American-style price cap regulation. Under this system, the growth of utility rates is limited to the growth in a price cap index (PCI). The growth rate of the PCI equals the growth rate of an inflation measure less an X factor. The X factor reflects industry productivity growth plus a stretch factor that is intended to reflect a company's potential for accelerated productivity growth.

Table 5. Statistical Results for Alternative Cost Model Specifications

Null hypothesis	J	LR-test stat	Critical value (5% level)
Homotheticity	12	162.91	21.03
Homogeneity	18	374.23	28.87
Log-linearity	25	371.51	37.65

Our benchmarking results are clearly relevant to the establishment of the stretch factor. One sensible approach is to establish different stretch factors for utilities with costs that are significantly superior to the industry norm, not statistically distinguishable from the norm, and significantly inferior to the norm. For these three categories, suggested stretch factors are 0%, 0.5%, and 1%. While other stretch factors may be considered, it bears noting that customers ultimately receive benefits from the externalization of ratemaking that benchmarking achieves as well as from the stretch factors assigned.

Table 6 presents efficiency scores, their t-stats, and the implied stretch factors from the translog benchmark model. We use a critical value of 1.65 to determine the cutoff points for superior and inferior cost performance. About 1/3 of utilities are thus deemed superior and inferior cost performers. The majority, 2/3 of utilities, are average cost performers.

6. CONCLUSION

Our study shows that econometric cost benchmarking models of considerable sophistication can be developed for power distributors with a quality dataset of adequate size. Both total cost and its major components can be benchmarked. Statistical methods can be used in model specification and application including, most notably, tests of efficiency hypotheses. These methods are also applicable in the cost appraisal and regulation of other utility businesses. For instance, the authors have developed similar models for power transmission, bundled power service, gas distribution, and water distribution.

The results also suggest several ways to improve the contribution of statistical benchmarking to utility regulation. One is to improve the size and quality of datasets. This can be achieved by greater use of multinational data and, for individual countries, the gathering of better capital cost data and the gradual accumulation of panel datasets. Further development of econometric methods is also desirable. Needed extensions include the refinement and diffusion through standard econometric packages of SFA methods for multi-equation cost models. These methods would, ideally, make corrections for heteroscedasticity and for cross-equation correlation of error terms.

Whether or not these advances are made, we believe that statistical tests of efficiency hypotheses should play a greater role in regulation where benchmarking is used. Generally speaking, conclusions about efficiency should be more difficult to make the smaller and less varied is a sample, the more atypical a

Table 6. Appraisals of Total Cost Performance and Suggested Stretch Factors, 2000-2002

Utility_id	Score	t-stat	p-value	Indicated stretch factor
Significantly Superior Performers				
171	-0.374	-5.705	0.000	0.0
136	-0.214	-5.545	0.000	0.0
31	-0.310	-5.357	0.000	0.0
98	-0.295	-4.051	0.000	0.0
62	-0.327	-3.369	0.001	0.0
92	-0.322	-3.048	0.002	0.0
21	-0.187	-2.972	0.003	0.0
203	-0.242	-2.961	0.003	0.0
30	-0.193	-2.339	0.020	0.0
183	-0.158	-2.255	0.024	0.0
156	-0.149	-1.681	0.093	0.0
Average Performers				
50	-0.176	-1.512	0.131	0.5
157	-0.209	-1.465	0.143	0.5
169	-0.093	-1.393	0.164	0.5
135	-0.171	-1.287	0.198	0.5
202	-0.083	-1.250	0.212	0.5
196	-0.090	-1.200	0.231	0.5
99	-0.081	-1.177	0.240	0.5
154	-0.086	-1.058	0.290	0.5
25	-0.083	-0.893	0.372	0.5
133	-0.209	-0.824	0.410	0.5
182	-0.101	-0.815	0.415	0.5
63	-0.066	-0.758	0.448	0.5
73	-0.059	-0.679	0.497	0.5
185	-0.083	-0.626	0.531	0.5
153	-0.032	-0.611	0.541	0.5
23	-0.036	-0.589	0.556	0.5
93	-0.064	-0.567	0.571	0.5
138	-0.040	-0.465	0.642	0.5
149	-0.040	-0.403	0.687	0.5
131	-0.035	-0.301	0.764	0.5
27	-0.035	-0.290	0.772	0.5
9	-0.031	-0.274	0.784	0.5
36	-0.015	-0.215	0.830	0.5
152	-0.022	-0.177	0.859	0.5
89	-0.005	-0.088	0.930	0.5
91	-0.006	-0.087	0.931	0.5
119	-0.001	-0.005	0.996	0.5
53	0.026	0.249	0.803	0.5
198	0.037	0.294	0.769	0.5
14	0.031	0.305	0.760	0.5
101	0.042	0.371	0.711	0.5

continued

Table 6. Appraisals of Total Cost Performance and Suggested Stretch Factors, 2000-2002 (continued)

Utility_id	Score	t-stat	p-value	Indicated stretch factor
13	0.052	0.518	0.605	0.5
17	0.070	0.661	0.509	0.5
109	0.107	0.733	0.464	0.5
12	0.084	0.818	0.414	0.5
2	0.111	0.891	0.373	0.5
29	0.046	0.953	0.341	0.5
167	0.093	1.005	0.315	0.5
180	0.081	1.028	0.304	0.5
178	0.079	1.046	0.296	0.5
67	0.094	1.118	0.264	0.5
140	0.144	1.391	0.165	0.5
172	0.119	1.535	0.125	0.5
159	0.131	1.641	0.101	0.5
Significantly Inferior Performers				
110	0.187	1.745	0.081	1.0
141	0.103	1.846	0.065	1.0
163	0.185	1.940	0.053	1.0
186	0.229	1.974	0.049	1.0
181	0.163	2.214	0.027	1.0
201	0.164	2.339	0.020	1.0
130	0.198	2.960	0.003	1.0
22	0.106	3.054	0.002	1.0
47	0.349	4.497	0.000	1.0
142	0.365	4.593	0.000	1.0
150	0.369	6.080	0.000	1.0

subject utility is from the sample used to appraise it, and the more poorly the cost model explains the data on which it is based. Active use of hypotheses tests will thus encourage regulators to employ better benchmarking methods. The ability of a benchmarking method to facilitate hypotheses tests on efficiency should be an important consideration in method selection.

REFERENCES

Afriat, S.N. (1972). "Efficiency Estimation of Production Functions." *International Economic Review* 13: 568-598.

Aigner, D.J., C.A.K. Lovell and P. Schmidt (1977). "Formulation and Estimation of Stochastic Frontier Production Function Models." *Journal of Econometrics* 6(2): 21-38.

Burns, P. and T.G. Weyman-Jones (1996). "Cost Functions and Cost Efficiency in Electricity Distribution: A Stochastic Frontier Approach." *Bulletin of Economic Research* 48(1): 41-64.

Coelli, T., D.S.P. Rao and G.E. Battese (1998). *An Introduction to Efficiency and Productivity Analysis*: Norwell, MA. Kluwer Academic Publishers.

Dhrymes, P. J. (1971). "Equivalence of Iterative Aitkin and Maximum Likelihood Estimators for a System of Regression Equations." *Australian Economic Papers* 10: 20-4.

DRI-WEFA (various years). *Utility Cost Information Service.*

Gagne, R. and P. Ouellette (1998). "On the Choice of Functional Forms: Summary of a Monte Carlo Experiment." *Journal of Business and Economic Statistics* 16(1): 118-124.

Guilkey, D. K., C.A.K. Lovell and R.C. Sickles (1983). "A Comparison of the Performance of Three Flexible Functional Forms." *International Economic Review* 24: 591-616.

Hattori, T. (2002). "Relative Performance of U.S. and Japanese Electricity Distribution: An Application of Stochastic Frontier Analysis." *Journal of Productivity Analysis* 18(3): 269-284.

Hattori, T., T. Jamasb and M.G. Pollitt (2002). "Relative Performance of UK and Japanese Electricity Distribution Systems 1985-1998: Lessons for Incentive Regulation." DAE Working Paper: 0212(CMI03). University of Cambridge.

Hiebert, L.D. (2002). "The Determinants of the Cost Efficiency of Electric Generating Plants: A Stochastic Frontier Approach." *Southern Economic Journal* 64(4): 935-946.

Hjalmarsson, L. and A. Veiderpass (1992). "Efficiency and Ownership in Swedish Electricity Retail Distribution." *Journal of Productivity Analysis* 3: 7-23.

Hall, R. and D. W. Jorgensen (1967). "Tax Policy and Investment Behavior." *American Economic Review* 57: 391-4140.

Hulten, C. and F. Wykoff (1981). "The Measurement of Economic Depreciation" in C. Hulten, ed., *Depreciation, Inflation, and the Taxation of Income From Capital.* Washington, D.C.: Urban Institute.

Jamasb, T. and M. Pollitt (2001). "Benchmarking and Regulation: International Electricity Experience," *Utilities Policy* 9: 107-130.

Magnus, J. R. (1978). "Maximum Likelihood Estimation of the GLS Model with Unknown Parameters in the Disturbance Covariance Matrix." *Journal of Econometrics* 7: 281-312.

Means, R.S. (1999). *Heavy Construction Cost Data.* Kingston, MA.

Meeusen, W. and J. van den Broeck (1977). "Efficiency Estimation from Cobb-Douglas Production Functions With Composed Error," *International Economic Review* 18: 435-444.

Neuberg, L. (1977). "Two Issue in Municipal Ownership of Electric Power Distribution Systems." *Bell Journal of Economics* 8: 302-323.

Oberhofer, W. and Kmenta, J. (1974). "A General Procedure for Obtaining Maximum Likelihood Estimates in Generalized Regression Models." *Econometrica* 42: 579-90.

Richmond, J. (1974). "Estimating the Efficiency of Production." *International Economic Review* 15: 515-521.

Salvanes, K. and S. Tjotta (1998). "Productivity Differences in Multiple Output Industries," *Journal of Productivity Analysis* 4: 23-43.

Schmidt, P. and R.C. Sickles (1984). "Production Frontiers and Panel Data." *Journal of Business and Economic Statistics* 2: 367-374.

U.S. Department of Commerce (1994). *Statistical Abstract of the United States.* Washington, D.C.

U.S. Department of Commerce (various issues). *Survey of Current Business.* Washington, D.C.

U.S. Department of Commerce, unpublished data on the stocks and service lives of the capital of Local Distribution Companies.

U.S. Department of Labor (1998). *National Compensation Survey.* Washington, D.C.

U.S. Federal Energy Regulatory Commission (various years). *FERC Form 1.* Washington, D.C.

Varian, H. (1984). *Microeconomic Analysis.* New York: Norton and Company.

Whitman, Requardt and Associates (1993). *Handy-Whitman Index of Public Utility Construction Costs.* Baltimore, MD.

Yatchew, A. (2000). "Scale Economies in Electricity Distribution: A Semiparametric Analysis." *Journal of Applied Econometrics* 15: 187-210.

Yatchew, A. (2001). "Incentive Regulation of Distributing Utilities Using Yardstick Competition." *The Electricity Journal*, January/February: 56-60.

Zellner, A. (1962). "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests of Aggregation Bias." *Journal of the American Statistical Association* 57: 348-68.