Exhibit KEDNE/LRK-3

THE COST PERFORMANCE OF BOSTON GAS

Pacific Economics Group

Economic and Litigation Consulting

THE COST PERFORMANCE OF BOSTON GAS

January 28, 2003

Larry Kaufmann, Ph.D.
Partner

David Hovde, M.S.
Senior Economist

John Kalfayan, M.S.
Senior Economist

Lullit Getachew, Ph.D.
Senior Economist

PACIFIC ECONOMICS GROUP
22 East Mifflin, Suite 302
Madison, Wisconsin USA 53705
608.257.1522 608.257.1540 Fax

TABLE OF CONTENTS

1. INTRODUCTION AND SUMMARY **1**

1.1 INTRODUCTION 1

1.2 SUMMARY OF RESEARCH 2

1.2.1 Definition of Cost *2*

1.2.2 The Sample *2*

1.2.3 Econometric Research *2*

2. DATA ISSUES **3**

2.1 DATA 3

2.2 DEFINITION OF COST 5

2.2.1 Applicable Total Cost *5*

2.2.2 Cost Decomposition *5*

3. ECONOMETRIC RESEARCH **6**

3.1 AN OVERVIEW OF THE METHOD 6

3.2 BUSINESS CONDITION VARIABLES 9

3.2.1 Output Quantity Variables *9*

3.2.2 Input Prices *10*

3.3 OTHER BUSINESS CONDITIONS 10

3.4 BUSINESS CONDITIONS OF BOGAS 12

3.5 ECONOMETRIC RESULTS 12

3.5.1 Estimation Results *12*

3.5.2 Econometric Benchmarking Results *16*

APPENDIX: FURTHER DETAILS OF THE BENCHMARKING RESEARCH **18**

A.1 CAPITAL COST 18

A.2 ECONOMETRIC RESEARCH 20

A.2.1 Form of the Cost Model *20*

A.2.2 Estimation Procedure *22*

A.2.3 Predicting Cost *23*

REFERENCES **24**

Pacific Economics Group
Economic and Litigation Consulting

1. INTRODUCTION AND SUMMARY

1.1 Introduction

Statistical benchmarking has in recent years become a widely used tool in the assessment of utility performance. Managers look to benchmarking studies for indications of how well their companies are doing. Benchmarking also plays a growing role in regulation. Such studies can, for example, be used to assess the reasonableness of costs and consumer dividends at the start of multiyear rate plans.

Appraisals of utility performance are facilitated by the extensive data that utilities report to regulators and industry associations. However, accurate appraisals are still challenging. There are important differences between companies in the character of services provided, the overall scale of operations, the prices of production inputs, and other business conditions that influence their cost. Data are unavailable for many companies and do not cover all relevant business conditions where they are available.

Pacific Economics Group LLC ("PEG") personnel have been active for several years in statistical benchmarking research for utilities. We pioneered the use of scientific benchmarking in U.S. regulation and have testified on our work in several proceedings. Boston Gas (BoGas) is preparing a multiyear rate plan for its gas distribution services. It has commissioned PEG to measure its overall cost efficiency. We appraised its efficiency using econometric cost modeling.

This paper is a report on our benchmarking work for gas distribution. Following a brief summary of the work, Section 2 discusses the database used in the study and our calculation of distribution cost. Our econometric work is discussed in Section 3. Additional, more technical details of the research are presented in the Appendix.

1.2 Summary of Research

1.2.1 Definition of Cost

Our research addressed the efficiency of local gas distribution companies (LDCs) in managing the total cost of their distribution operations. Gas distribution services are defined to include the gas delivery, customer account and customer information services provided by LDCs. We do not address the cost of gas procurement services. The total cost of distribution services comprises the cost of plant ownership, operation, and maintenance.

1.2.2 The Sample

The econometric research was based on a sample of data for 43 distributors. The focus of benchmarking was the 1993-2000 period.

1.2.3 Econometric Research

The gas distribution cost performance of BoGas was appraised using an econometric cost model. Guided by economic theory, we developed a mathematical model in which the cost of gas distribution is a function of some quantifiable business conditions. The parameters of the model were estimated statistically using data on the historical costs of LDCs and the business conditions they faced. All key parameters were plausibly signed and highly significant.

We used the model to predict the average total cost of gas distribution services for BoGas given the business conditions it faced. The Company was found to face some challenging conditions in its efforts to contain gas distribution cost. For example, BoGas is not a combined gas and electric utility. The Company faces high prices for labor services and plant construction, and our results also show that there are special cost pressures from operating in the Northeast. It also has relatively more cast-iron main than any gas distributor in our sample.

BoGas's gas distribution cost was about 27% below the value predicted by the cost model, on average, from 1993 to 2000. This difference was statistically significant. We therefore conclude that BoGas is a significantly superior cost performer.

2. DATA ISSUES

2.1 Data

The primary source of the data used in our gas distribution cost research changed over the full sample period used in our benchmarking work. The *Uniform Statistical Report* (USR) was the primary source for the earliest years. Gas utilities are asked to file these reports annually with the American Gas Association (AGA). USR data for some variables are aggregated and published annually by the AGA in *Gas Facts.*

USRs are unavailable for many distributors. Many do not file complete USRs. Some LDCs that do file them do not release them to the public. The development of a satisfactory sample therefore required us to obtain basic cost and quantity data from alternative sources including, most notably, reports to state regulators. These reports often use as templates the Form 2 report that interstate gas transmission companies are required to file with the Federal Energy Regulatory Commission. Other sources of data were also used in the research. These included DRI/McGraw Hill; Whitman, Requardt & Associates; the Bureau of Economic Analysis ("BEA") of the U.S. Department of Commerce; and the Bureau of Labor Statistics ("BLS") of the U.S. Department of Labor.

Our econometric cost model is based on quality data for 43 gas distributors over the 1993 to 2000 period. The sample includes most of the nation's larger distributors.

The sampled distributors grouped by region are listed in Table 1. It can be seen that the regional distribution of sampled LDCs is somewhat uneven. In particular, the northeast accounts for 40% of the sampled companies, but accounts for only 23% of U.S. gas end users. Texas accounts for only 2% of the sample, but for 7% of gas end users. The table also indicates that the sampled LDCs served about 52% of all gas end users in the United States.

Table 1

SAMPLE FOR THE GAS DISTRIBUTION ECONOMETRIC RESEARCH

Region	Company	Number of Customers (2000)
Northeast		
	Boston Gas	542,792
	Brooklyn Union Gas	1,191,679
	Central Hudson Gas & Electric	63,851
	Commonwealth Gas	243,853
	Connecticut Energy	164,012
	Connecticut Natural Gas	155,641
	Consolidated Edison	1,048,357
	New Jersey Natural Gas	414,620
	Niagara Mohawk	548,075
	Orange & Rockland Utilities	118,718
	PECO	430,842
	People's Natural Gas	353,715
	PG Energy	155,992
	Providence Energy	172,965
	Public Service Electric & Gas	1,621,128
	Rochester Gas & Electric	285,944
South Atlantic		
	Atlanta Gas Light	1,530,000
	Baltimore Gas & Electric	595,239
	Public Service Company of North Carolina	357,736
	Washington Gas Light	868,362
North Central		
	Citizens Gas & Coke	265,450
	Consumers Power	1,594,484
	East Ohio Gas	1,234,854
	Illinois Power	399,361
	Interstate Power	50,270
	Madison G & E	113,781
	North Shore Gas	149,781
	Northern Illinois Gas	1,962,228
	Peoples Gas Light & Coke	840,560
	Wisconsin Gas	540,676
	Wisconsin Power & Light	157,077
South Central		
	Alabama Gas	465,656
	Louisville Gas & Electric	297,717
	Oklahoma Natural Gas	757,688
Southwest		
	Enserch	1,415,296
	Mountain Fuel Supply	705,878
	Southwest Gas	1,289,046
Northwest		
	Cascade Natural Gas	193,160
	Northwest Natural Gas	510,686
	Washington Natural Gas	580,283
California		
	Pacific Gas & Electric	3,818,679
	San Diego Gas & Electric	756,053
	Southern California Gas	5,008,579
Total for Sample		33,970,764
Industry Total *		64,804,630
Percentage of U.S. Total		52.4%

*Source For US Total: U.S. Energy Information Administration, *Natural Gas Annual 2000*

Preliminary, Privileged, and Confidential

2.2 Definition of Cost

2.2.1 Applicable Total Cost

Cost figures play an important role in our benchmarking methods. The applicable total cost of gas distribution was calculated as gas operation and maintenance ("O&M") expenses less gas production and procurement expenses plus total gas plant capital cost and a share of any common costs. The operations corresponding to this definition of cost include gas delivery, customer account, and customer information and other customer services of LDCs.

2.2.2 Cost Decomposition

Our benchmarking methods involve the decomposition of cost into three major input categories: capital services, labor services, and non-labor O&M inputs. The cost of gas delivery labor was defined as the sum of O&M salaries and wages and pensions and other employee benefits. The cost of other O&M inputs was defined to be O&M expenses net of these labor costs and of gas production and procurement expenses. This category includes the services of contract workers, insurance, real estate rents, equipment leases, and miscellaneous materials.

The study used a service price approach to measuring the cost of plant ownership that is based on the economic value of utility plant. Under this approach, the cost of capital is the product of a capital quantity index and the price of capital services. The cost of capital thus calculated includes depreciation, tax expenses, the opportunity cost of plant ownership, and capital gains. This method has a solid basis in economic theory and is well established in the scholarly literature. It controls in a precise and standardized way for differences between utilities in the age of their plants. Further details of our capital cost calculations are provided in Section A.1 of the Appendix.

3. ECONOMETRIC RESEARCH

3.1 An Overview of the Method

This section provides a substantially non-technical account of the econometric approach to benchmarking employed in this study. Additional, more technical details of the work are reported in the Appendix.

A mathematical model called a cost function was specified. Cost functions represent the relationship between the cost of a utility and quantifiable business conditions in its service territory. Business conditions are defined as aspects of a company's operating environment that influence its activities but cannot be controlled.

Economic theory was used to guide cost model development. We posited that the actual total cost (C_i) incurred by company, i, in service provision is the product of minimum achievable cost (C_i^*) and an efficiency factor ($efficiency_i$). This assumption can be expressed logarithmically as

$$\ln C_i = \ln C_i^* + \ln efficiency_i.^{1} \qquad [1]$$

The term ln indicates the natural log of a variable.

According to theory, the minimum total cost of an enterprise is a function of the amount of work it performs and the prices it pays for capital and labor services and other inputs to its production process. Theory also provides some guidance regarding the nature of the relationship between these business conditions and cost. For example, cost is apt to be higher the higher are input prices and the greater is the amount of work performed.

Here is a simple example of a minimum total cost function for gas distribution that conforms to cost theory.

$$\ln C_{i,t}^* = a_0 + a_1 \cdot \ln N_{i,t} + a_2 \cdot \ln W_{i,t} + u_{i,t}. \qquad [2]$$

[1] The logarithm of the product of two variables is the sum of their individual logarithms.

For each firm i in year t, the variable $N_{i,t}$ is the number of customers that the company serves. It quantifies one dimension of the work that it performs. The variable $W_{i,t}$ is the wage rate that the company pays. The wage rate and delivery volume are the measured business conditions in this cost function.

The term $u_{i,t}$ is the error term of the cost function. This term reflects errors in the specification of the model, including problems in the measurement of output and other business condition variables and the exclusion from the model of relevant business conditions. It is customary to assume a specific probability distribution for the error term that is determined by additional parameters, such as mean and variance.

Combining the results of Equations [1] and [2] we obtain the following model of cost:[2]

$$\ln C_{i,t} = \alpha_0 + \alpha_i \ln N_{i,t} + \alpha_2 \ln W_{i,t} + e_{i,t}. \qquad [3]$$

Here the *actual* (not minimum) total cost of a utility is a function of the two measured business conditions. The terms α_0, α_1, and α_2 are model parameters. Their values are assumed to be constant across companies and over some period of time. The α_0 parameter captures the efficiency factor for the average firm in the sample as well as the value of a_0 from Equation [3], the minimum total cost function. The values of α_1 and α_2 determine the effect of the two measured business conditions on cost. If the value of α_2 is positive, for instance, an increase in wage rates will raise cost.

---

[2] Here is the full logic behind this result:

$$\begin{aligned}
\ln C_{i,t} &= \ln C^*_{i,t} + \ln efficiency_i \\
&= \left(a_0 + a_1 \ln N_{i,t} + a_2 \ln W_{i,t} + u_{i,t}\right) + \ln efficiency_i \\
&= \left(a_0 + \ln efficiency^{average}\right) + a_1 \ln N_{i,t} + a_z \ln W_{i,t} \\
&\quad + \left[u_i + \left(\ln efficiency_i - \ln efficiency^{average}\right)\right] \\
&= \alpha_0 + \alpha_i \ln N_{i,t} + \alpha_2 \ln W_{i,t} + e_{i,t}
\end{aligned}$$

The term $e_{i,t}$ is the error term for equation [3]. We assume that it is a random variable. It includes the error term from the minimum total cost function. It also reflects the extent to which the Company's efficiency factor differs from the sample norm.

A branch of statistics called econometrics has developed procedures for estimating parameters of economic models. Cost model parameters can be estimated econometrically using historical data on the costs incurred by utilities and the business conditions that they faced. For example, a positive estimate for α_2 would reflect the fact that the cost reported by sampled companies was typically higher when higher wages were paid to employees.

Numerous statistical methods have been established in the econometrics literature for estimating parameters of economic models. In choosing among these, we have been guided by the desire to obtain the best possible model for cost benchmarking. Econometric methods are also useful in selecting business conditions for the model. Tests are available for the hypothesis that the parameter for a business condition variable equals zero. Variables were excluded from the model when such hypotheses could not be rejected.

A cost function fitted with econometric parameter estimates may be called an <u>econometric cost benchmark model</u>. We can use such a model to predict a company's cost given values for the variables that represent the business conditions that the company faced. Returning to our simple example, we might predict the (logged) cost of BoGas in period t as follows:[3]

$$\ln \hat{C}_{BoGas,t} = \hat{\alpha}_0 + \hat{\alpha}_1 \cdot \ln N_{BoGas,t} + \hat{\alpha}_2 \cdot \ln W_{BoGas,t}. \qquad [4]$$

Here $\hat{C}_{BoGas,t}$ denotes the predicted cost of the Company in period t, $N_{BoGas,t}$ is the number of customers it served, and $W_{BoGas,t}$ is the wage rate that it paid. The $\hat{\alpha}_0$, $\hat{\alpha}_1$, and $\hat{\alpha}_2$ terms are parameter estimates. Notice that in this model the cost benchmark reflects, through the estimate of parameter α_0, the *average* efficiency of the sampled utilities.

[3] Since this is a predicted equation using estimated parameters there is no error term.

Consider, now, that if the parameter estimates are unbiased and the expected value of $u_{i,t}$ is zero, the expected value of the percentage difference between the company's actual cost and that predicted by the model is the percentage difference between the efficiency factor of BoGas and that of the sample mean firm.

$$\ln\left(C_{BoGas,t} \Big/ \hat{C}_{BoGas,t} \right) = \ln\left(\text{efficiency}_{BoGas} \Big/ efficiency^{average} \right). \qquad [5]$$

This percentage difference is a measure of the company's cost performance.

A number like that generated by the cost benchmark model in [5] constitutes our best estimate of the company's cost given the business conditions that it faces. This is an example of a point prediction. An important characteristic of the econometric approach to benchmarking is that the statistical results provide information about the *precision* of such point predictions. According to econometric theory, precision is greater the lower is the variance of the model's prediction error. The variance of the prediction error can be estimated using a well-established formula. The formula shows that the precision of cost model predictions is greater to the extent that:

1) The model is more successful in explaining the variation in cost in the sample
2) The size of the sample is larger
3) The number of business condition variables included in the model is smaller
4) The business conditions of sample companies are more varied
5) The business conditions of the subject company are closer to those of the typical firm in the sample

3.2 Business Condition Variables

3.2.1 Output Quantity Variables

As noted above, economic theory suggests that quantities of work performed by utilities should be included in our cost model as business condition variables. There are two output quantity variables in our model: the number of retail customers and total throughput. We expect cost to be higher for higher values of each of these workload measures.

3.2.2 Input Prices

Cost theory also suggests that the prices paid for production inputs are relevant business condition variables. In this model, we have specified input price variables for capital, labor, and other O&M inputs.[4] We expect cost to be higher as the values of these price variables increase.

The labor price variable used in this study was constructed by PEG using data from the BLS. National Compensation Survey ("NCS") data for 1998 were used to construct average wage rates that correspond to each LDC's service territory. The wage levels were calculated as a weighted average of the NCS pay level for each job category using weights that correspond to the Electric, Gas, and Sanitary (EGS) sector for the U.S. as a whole. Values for other years were calculated by adjusting the 1998 level for changes in the Employment Cost Index for the EGS sector over the 1993-2000 period.

Prices for other O&M inputs are assumed to be the same in a given year for all companies. They are escalated by growth in the GDP-PI. Our general approach to the computation of a price index for capital services is described in Section 2.2. Further details of this calculation are found in the Appendix.

3.3 Other Business Conditions

Five additional business condition variables are included in the cost model. One is the percentage of distribution main not made of cast iron, calculated from American Gas Association data. Cast iron pipes were common in gas system construction in the early days of the industry. It is more heavily used in the older distribution systems, which tend to be in the eastern U.S. Greater use of cast iron typically involves both higher maintenance and replacement costs. A higher value for this variable means that a company owns fewer cast iron mains and has lower expected costs. Hence, we would expect the sign for this coefficient to be negative.

---

[4] The price index for other O&M inputs doesn't appear in the estimated parameter tables due to the imposition of the linear homogeneity restriction predicted by economic theory.

A second additional business condition variable is the number of power distribution customers served by the utility. This variable is intended to capture the extent to which the company has diversified into power distribution. Such diversification will typically lower cost due to the ability to share inputs (e.g., personnel, computer systems, meter readers) between the two services. Higher values for this variable indicate greater levels of diversification. We would therefore expect the value of this coefficient to be negative.

A third business condition was a dummy variable for distributors that operate in territories that are subject to frequent earthquakes. Systems in these territories may have to be designed differently to withstand earthquakes. Because these design differences are likely to entail additional costs, the coefficient on this variable is expected to be positive.

The model also included a dummy variable for distributors operating in the northeastern U.S. Previous econometric studies for BoGas have found that Northeast operations are associated with higher costs, even after controlling for factors like higher input prices. This coefficient is therefore expected to be positive. The northeast dummy takes a value of 1 for every distributor headquartered in the New England, New York, Pennsylvania or New Jersey and zero for all other companies.

Finally, the model included a PBR dummy variable for BoGas. This variable took a value of one for BoGas during the years when it operated under PBR (1997-2000) and zero for other years and for every other company. Because PBR is expected to lower costs by strengthening performance incentives, this coefficient was expected to be negative.

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. A trend variable captures the net effect on cost of diverse conditions, including technological change. It may also reflect the failure of the included business condition variables to properly measure the trends in relevant cost drivers. The model may, for instance, exclude an important cost driver or do a poor job of measuring such a driver. The trend variable might then capture the impact on cost of the trend in the driver.

3.4 Business Conditions of BoGas

Table 2 compares the average values over the 1993-2000 period of cost model business conditions for BoGas to the sample mean values of these variables. It can be seen that the average total cost of BoGas was just over 80% of the sample mean. Meanwhile, the number of customers served by BoGas was about 70% of the mean and its throughput was just below 80% of the mean.

Turning next to input prices, the table shows that BoGas had labor prices 13% above the sample mean. Its capital service price was about 9% above the mean.

Prices for other inputs were assumed to be the same across the sampled companies. This simplifying assumption may well distort results for BoGas. After all, it is quite possible that a region with high labor and construction costs also has higher average prices for other production inputs, especially those that are intensive in the use of local labor.

Regarding the other business conditions, note first that BoGas's percentage of gas distribution main that is not made of cast iron was well below the sample mean. In fact, BoGas had the most cast-iron intensive system in our sample, representing about 44% of main. This was more than twice the average share of cast iron main for the sample (17%) and reflects the age of the BoGas distribution network.

Note, finally, that BoGas has no power distribution customers. This has limited its opportunity to realize potential scope economies by sharing inputs with other utility services.

3.5 Econometric Results

3.5.1 Estimation Results

Estimation results for the cost model are reported in Table 3. The parameter values for the five additional business conditions and for the first order terms of the translogged variables are elasticities of the cost of the sample mean firm with respect to the basic variable. The first order terms are the terms that do not involve squared values of business condition variables or interactions between different variables. The table shades the results for these terms for reader convenience.

Table 2

Average Values of Variables in the Benchmarking Study: Gas Delivery

Variable	Units	U.S. Sample Average	Boston Gas	Boston Gas/ Sample Mean
Gas Delivery Cost	1,000 U.S. Dollars	380,027	311,651	0.82
Number of Customers	Customers	742,764	522,947	0.70
Total Throughput	mdkth	181,144	141,966	0.78
Price of Capital Services	Index Number	16.25	17.67	1.09
Price of Labor Services	Dollars per Employee	35,132	39,818	1.13
Price of Materials	Index Number	1.13	1.13	1.00
Number of Electric Customers	Customers	432,511	0	0.00
Percent of Main not Cast Iron	Percent	83.02%	56.14%	0.68
Earthquake Dummy	Binary	0.186	0.000	0.00
NE Dummy	Binary	0.372	1.000	2.69

Table 3

Translog Cost Function Regression Results: Gas Delivery

VARIABLE KEY*

L= Labor Price
K= Capital Price
N= Number of Customers
YV= Total Throughput
EC= Number of Electric Customers
NI= % of Main that is Non-cast Iron
EQ= Earthquake Dummy Variable
NE= Northeast Dummy
BG= Boston Gas PBR Dummy

EXPLANATORY VARIABLE	ESTIMATED COEFFICIENT	T-STATISTIC	EXPLANATORY VARIABLE	ESTIMATED COEFFICIENT	T-STATISTIC
L	0.202	60.15	EC	-0.010	-7.50
LL	-0.101	-2.19			
LK	0.010	0.41	NI	-0.210	-4.23
LN	0.014	1.58			
LYV	-0.023	-2.60	EQ	0.016	3.75
K	0.648	141.21	NE	0.059	9.11
KK	0.090	3.38			
KN	-0.028	-2.61	BG	-0.003	-3.92
KYV	0.038	3.54			
N	0.658	20.82	Constant	8.015	356.38
NN	-0.439	-6.41			
NYV	0.452	6.44	Trend	-0.005	-1.46
YV	0.210	6.32	System Rbar-Squared	0.975	
YVYV	-0.512	-6.83			

* Data for all variables were logged and mean-scaled prior to model estimation

The tables also report the values for the corresponding asymptotic t ratios. These were also generated by the estimation program and were used to assess the range of possible values for parameters that are consistent with the data. A parameter estimate is deemed statistically significant if the hypothesis that the true parameter value equals zero is rejected. This statistical test requires the selection of a critical value for the asymptotic *t* ratio. In this study, we employed critical values that are appropriate for a 90% confidence level given a large sample. The critical value was 1.645.

Examining the results in Table 3, it can be seen that the cost function parameter estimates were plausible as to sign and magnitude. With regard to the first order terms of the translogged variables, cost was found to be higher the higher were input prices and output quantities. At the sample mean, a 1% increase in the number of customers raised cost by 0.66%. A 1% hike in throughput raised cost by about 0.21%. The sum of those elasticities was about 0.87%. The number of customers served was clearly the dominant output-related cost driver.

The sum of the output elasticities is a common indicator of economies of scale. A sum less than one is an indicator that scale economies can be realized from balanced output growth for a firm of sample mean size. Since TFP growth depends in part on scale economies, this also means that it can differ from region to region of the U.S. due in part to differences in the pace of output growth.

Turning to results for the input prices, it can be seen that the elasticity of cost with respect to the price of capital services was 0.65%. This was more than three times the estimated elasticity of the price of labor. This reflects the capital intensiveness of the gas distribution business.

The coefficients on the additional business condition variables were also sensible and, with the exception of the trend variable, were statistically significant.

- Cost was lower as the percentage of distribution mains not made with cast iron increased.

- Cost was lower as the number of electric customers served by a distributor increased.
- Cost was higher for distributors operating in the Northeast.
- Cost was higher for distributors operating in territories with frequent earthquakes.
- The coefficient on the PBR dummy was –0.3%; this implies that, after controlling for each of the other business conditions in the model, BoGas's costs declined by 0.3% during the years when PBR was in effect.
- The estimate of the trend variable parameter was –0.005 and was not significant.

3.5.2 Econometric Benchmarking Results

Table 4 presents the results of our appraisals of BoGas's cost using the econometric model. The Company's average cost during the sample period was found to be about 27% below its predicted value. The hypothesis that the company was an average (or inferior) cost performer was rejected at the 99% confidence level. BoGas was therefore a significantly superior cost performer.

Table 4

Actual and Predicted Comprehensive Cost For Gas Distribution: 1993-2000 Boston Gas (U.S. $)

Actual Cost $1,000	Predicted Cost $1,000	Difference (%)	t-statistic
311,651	427,898	-27.2%	-5.59

APPENDIX:

FURTHER DETAILS OF THE BENCHMARKING RESEARCH

This section provides additional and more technical details of our benchmarking work. We first consider our method for computing capital cost. There follow treatments of our indexing and econometric work.

A.1 Capital Cost

A service price approach was chosen to measure the cost of plant ownership. This approach has a solid basis in economic theory and is widely used in scholarly empirical work.[5] In the application of the general methodology used in this study, capital cost in a given year t, CK_t, is the product of a capital service price index, WKS_t and a capital quantity index, XK_{t-1}.

$$CK_t = WKS_t \cdot XK_{t-1}. \qquad [6]$$

The service price index may be thought of as the annual cost (including the opportunity cost) of owning a unit of plant.

Each capital quantity index is constructed using inflation-adjusted data on the value of utility plant. Each service price index measures the trend in the hypothetical price of capital services from the assets in a competitive rental market. The price and quantity indexes require a consistent mathematical characterization of the process of plant deterioration.

In constructing the indexes we took 1983 as the benchmark or starting year for our gas distribution cost research. The values for these indexes in the benchmark year are based on the net value of plant as reported on the USR. We estimated the benchmark year (inflation adjusted) value of net plant by dividing the aggregate appropriate base year value by a "triangularized" weighted average of the values of an index of utility asset prices for a period ending in the

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

benchmark year equal to the lifetime of plant. A triangularized weighting gives greater weight to more recent values of this index, reflecting the notion that more recent plant additions have a disproportionate impact on the book value of plant.[6] The value of the asset-price index, WKA_t, is the applicable regional Handy-Whitman index of utility construction costs for the relevant asset category.[7]

The following formula was used to compute subsequent values of the capital quantity index:

$$XK_t = (1 - d) \cdot XK_{t-1} + \frac{VI_t}{WKA_t}. \qquad [7]$$

Here, the parameter, d, is the economic depreciation rate, VI_t is the value of gross additions to the utility plant and WKA_t is the index of utility plant asset prices.

The economic depreciation rate, d, was calculated as a weighted average of the depreciation rates for the structures and equipment used in the applicable industry. The depreciation rate for each structure and equipment category was obtained from the Bureau of Economic Analysis (BEA) of the U.S. Department of Commerce. The weights were based on net stock value data drawn from the same source.

The formula for the capital service price index, WKS_t, is:

$$WKS_t = \left(CK_t^{taxes} / XK_{t-1}\right) + r_t \cdot WKA_{t-1} + d \cdot WKA_t - \left(WKA_t - WKA_{t-1}\right). \qquad [8]$$

The four terms in this formula correspond to the four components of capital cost. These are: taxes, the opportunity cost of capital, depreciation, and capital gains.[8] Here, CK_t^{taxes} is the sum of total tax payments and franchise fees attributed to the LDC.[9] The term, r_t, is the user cost of capital for the U.S. economy.[10] PEG calculates this using data in the National Income and Product Accounts

---

[6] For example, in a triangularized weighting of 20 years of index values, the oldest index value has a weight of 1/210, the next oldest index has a value of 2/210, and so on. 210 is the sum of the numbers from 1 to 20. A discussion of triangularized weighting of asset price indexes is found in Stevenson (1980).

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

[8] The opportunity cost of capital is sometimes called the cost of funds.

[9] Franchise fees are a part of O&M expenses in our TFP trend indexes.

[10] The U.S. economy user cost of capital is not directly observable, but it can be measured by applying two economic relationships. The first economic pertains to the National Income and Products Accounts (NIPA)

(NIPA). The accounts are published by the Department of Commerce in its Survey of Current Business series. Capital gains are smoothed using a three-year moving average.

A.2 Econometric Research

A.2.1 Form of the Cost Model

The functional form selected for this study was the translog.[11] This very flexible function is the most frequently used in econometric cost research, and by some account the most reliable of several available alternatives.[12] The general form of the translog cost function is:

---

definitions of Gross Domestic Product (GDP) and the cost of inputs used by the U.S. economy. In the NIPA, the total cost of the U.S. economy inputs is equal to GDP. At the economy-wide level there are two inputs: labor and capital. Therefore the total cost of capital is equal to GDP less Labor Compensation (CL), or:

$$CK = GDP - CL \tag{1}$$

where CK represents the total cost of capital. The second relationship is between the total cost of capital and the components of the capital price equation. The total cost of capital is equal to the product of the quantity of capital input and the price of capital input, or:

$$CK = P_k \cdot K \tag{2}$$

where P_k represents the price and K the quantity of capital input. The price of capital can be decomposed into the price index for new plant and equipment (J), the opportunity cost of capital (r), the rate of depreciation (d), the inflation rate for new plant and equipment (l), and the rate of taxation on capital (t):

$$P_k = J \cdot (r + d - l + t) \tag{3}$$

Combining (2) and (3) one obtains the relationship:

$$\begin{aligned} CK &= J \cdot (r + d - l + t) \cdot K \\ &= r \cdot J \cdot K + d \cdot J \cdot K - l \cdot J \cdot K + t \cdot J \cdot K \\ &= r \cdot VK + D - l \cdot VK + T \end{aligned} \tag{4}$$

where D represents the total cost of depreciation, T total indirect business taxes and corporate profits taxes, and VK the current cost of plant and equipment net stock. Combining (1) and (4), one can derive the following equation for the opportunity cost of capital:

$$r = (GDP - CL - D - T + l \cdot VK) / (VK) \tag{5}$$

GDP, labor compensation, depreciation, and taxes are reported annually in the NIPA. The current cost of plant and equipment net stock and the inflation rate for plant and equipment are not reported in the NIPA, but are reported in Fixed Reproducible Tangible Wealth in the United States.

[11] The transcendental logarithmic (or translog) cost function can be derived mathematically as a second order Taylor series expansion of the logarithmic value of an arbitrary cost function around a vector of input prices and output quantities.

[12] See Guilkey (1983), et. al.

$$\ln C = \alpha_0 + \sum_h \alpha_h \ln Y_h + \sum_j \alpha_j \ln W_j$$
$$+ \frac{1}{2}\left(\sum_h \sum_k \gamma_{h,k} \ln Y_h \ln Y_k + \sum_j \sum_n \gamma_{j,n} \ln W_j \ln W_n \right) \qquad [9]$$
$$+ \sum_h \sum_j \gamma_{i,j} \ln Y_i \ln W_j$$

where Y_h denotes one of K variables that quantify output and the W_j denotes one of N input prices.

One aspect of the flexibility of this function is its ability to allow the elasticity of cost with respect to each business condition variable to vary with the value of that variable. The elasticity of cost with respect to an output quantity, for instance, may be greater at smaller values of the variable than at larger variables. This type of relationship between cost and quantity is often found in cost research.

Business conditions other than input prices and output quantities can contribute to differences in the costs of LDCs. To help control for other business conditions the logged values of some additional explanatory variables were added to the model in Equation [9] above.

The econometric model of cost we wish to estimate can then be written as:

$$\ln C = \alpha_o + \sum_h \alpha_h \ln Y_h + \sum_j \alpha_j \ln W_j$$
$$+ \frac{1}{2}\left[\sum_h \sum_k \gamma_{hk} \ln Y_h \ln Y_k + \sum_j \sum_n \gamma_{jn} \ln W_j \ln W_n \right] \qquad [10]$$
$$+ \sum_h \sum_j \gamma_{ij} \ln Y_h \ln W_j + \sum_h \alpha_h \ln Z_h + \alpha_t T + \varepsilon$$

Here the Z_h's denote the additional business conditions, T is a trend variable, and ε denotes the error term of the regression.

Cost theory requires a well-behaved cost function to be homogeneous in input prices. This implies the following three sets of restrictions:

$$\sum_{h=1}^{N} \frac{\partial \ln C}{\partial \ln W_h} = 1 \qquad [11]$$

$$\sum_{h=1}^{N} \frac{\partial^2 \ln C}{\partial \ln W_h \partial \ln W_j} = 0 \qquad \forall j = 1,\ldots,N \qquad [12]$$

$$\sum_{h}^{N} \frac{\partial^2 \ln C}{\partial \ln Y_h \partial \ln Y_j} = 0 \qquad \forall j = 1,\ldots,K \qquad [13]$$

Imposing the above $(1+N+K)$ restrictions implied by Equations [21-23] allow us to reduce the number of parameters that need be estimated by the same amount.

Estimation of the parameters in Equation [20] is now possible but this approach does not utilize all information available in helping to explain the factors that determine cost. More efficient estimates can be obtained by augmenting the cost equation with the set of cost share equations implied by Shepard's Lemma. The general form of a cost share equation for a representative input price category, j, can be written as:

$$S_j = \alpha_j + \sum_i \gamma_{h,j} \ln Y_h + \sum_n \gamma_{jn} \ln W_n \qquad [14]$$

We note that the parameters in this equation also appear in the cost model. Since the share equations for each input price are derived from the first derivative of the translog cost function with respect to that input price, this should come as no surprise. Furthermore, because of these cross-equation restrictions, the total number of coefficients in this system of equations will be no larger than the number of coefficients required to be estimated in the cost equation itself.

A.2.2 Estimation Procedure

We estimated this system of equations using a procedure first proposed by Zellner (1962).[13] It is well known that if there exists contemporaneous correlation between the errors in the system of regressions, more efficient estimates can be obtained by using a Feasible Generalized Least Squares (FGLS) approach. To achieve even a better estimator, PEG iterates this procedure to convergence.[14] Since we estimate these unknown disturbance matrices consistently, the estimators we eventually compute are equivalent to Maximum Likelihood Estimation (MLE).[15] Our estimates would thus possess all the highly desirable properties of MLE's.

---

[13] See Zellner, A. (1962).

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

[15] See Dhrymes (1971), Oberhofer and Kmenta (1974), Magnus (1978).

Before proceeding with estimation, there is one complication that needs to be addressed. Since the cost share equations by definition must sum to one at every observation, one cost share equation is redundant and must be dropped.[16] This does not pose a problem since another property of the MLE procedure is that it is invariant to any such reparameterization. Hence, the choice of which equation to drop will not affect the resulting estimates.

A.2.3 Predicting Cost

We now turn our attention to the topic of predicting the level of a utility's cost given its specific values for the explanatory variables. Fitting our cost model with the econometric parameter estimates, we obtain an econometric model of distributor cost. This can then be used to predict the historical cost of an LDC given its values for the specified business controls. It is well known that the ability of the model to make accurate predictions depends, in part, on the characteristics of the data reported for the utility as compared to the sample averages. The closer the firm's data are to the sample averages, the more accurate is the model's prediction. Alternatively, the more the characteristics of the utility's data lie outside those of the sample means, the less reliable is its predicted cost.

It should be noted that the model specification was determined using the data for all sampled companies, including BoGas. However, to compute the model parameters and standard errors for the prediction required that the utility of interest be dropped from the sample when we estimated the coefficients in the predicting equation.[17] The standard error based on this "out-of-sample" prediction was then used to construct the hypothesis tests for cost efficiency.

---

[16] This equation can be estimated indirectly from the estimates of the parameters left remaining in the model.

[17] This implies that the estimates used in constructing the predicting equation will vary slightly from those reported in the study.

REFERENCES

American Gas Association, *Gas Facts*, Arlington, VA, various issues.

Breusch, T. and A. R. Pagan (1980), "The LaGrange Multiplier Test and Its Applications to Model Specification in Econometrics," *Review of Economic Studies*, 47 pages 239-54.

Buse, A. (1982), "The Likelihood Ratio, Wald and LaGrange Multiplier Tests: An Expository Note," *The American Statistician*, 62, pages 153-7.

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

Guilkey, et. al. (1983), "A Comparison of Three Flexible Functional Forms," *International Economic Review*, 24, pages 591-616.

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

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

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

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

Mundlak, Y. (1978), "On the Pooling of Time Series and Cross Section Data," *Econometrica*, 46, pages 69-85.

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

U.S. Department of Commerce, *Statistical Abstract of the United States*, 1994.

U.S. Department of Commerce, *Survey of Current Business*, various issues.

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

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

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