Choosing among Alternative Cost Function Specifications: An Application to Italian Multi-utilities

Massimiliano PIACENZA, Davide VANNONI

Working Paper 4, 2003

© Hermes
Real Collegio Carlo Alberto
Via Real Collegio, 30
10024 Moncalieri (To)
011 640 27 13 - 642 39 70
info@hermesricerche.it
http://www.hermesricerche.it

Choosing among Alternative Cost Function Specifications: An Application to Italian Multi-utilities

Massimiliano Piacenza
(*Ceris-CNR*, *HERMES*)

Davide Vannoni
(*University of Turin*, *Ceris-CNR*, *HERMES*)

May 2003

Abstract

The paper analyses the cost properties of a sample of Italian utilities providing in combination gas, water and electricity services. The estimates from a multi-product *Composite* cost function (Pulley and Braunstein, 1992) are compared with the ones coming from other traditional functional forms such as the Standard Translog, the Generalized Translog, and the Separable Quadratic. The results show that the composite model provides a better description of data and highlight the presence of global *scope* and *scale* economies for the 'median' firm of the sample.

Keywords: Multi-Utilities, Scope and Scale Economies, Composite Cost Function.
JEL Code: L97, L5, L21, C3.

1. The Composite Cost Function Model

In a seminal contribution, Pulley and Braunstein (1992) proposed a novel functional form - the *Composite* cost function - which is well suited for examining cost properties of multi-product firms. Such a model, as well as other widely used alternative cost functions, are nested into the following *General* specification (PB_G):

$$C^{(\phi)} = \left\{ \exp\left[\left(\alpha_0 + \sum_i \alpha_i Y_i^{(\pi)} + \frac{1}{2}\sum_i \sum_j \alpha_{ij} Y_i^{(\pi)} Y_j^{(\pi)} + \sum_i \sum_r \delta_{ir} Y_i^{(\pi)} \ln P_r \right)^{(\tau)} \right] \right.$$
$$\left. \cdot \exp\left[\sum_r \beta_r \ln P_r + \frac{1}{2} \sum_r \sum_l \beta_{rl} \ln P_r \ln P_l \right] \right\}^{(\phi)} \qquad [1]$$

where the superscripts in parentheses π, ϕ and τ represent Box-Cox transformations (for example $Y_i^{(\pi)} = (Y_i^{\pi} - 1)/\pi$ for $\pi \neq 0$ and $Y_i^{(\pi)} \to \ln Y_i$ for $\pi \to 0$). C is the long-run cost of production, Y_i refers to outputs, and P_r indicates factor prices.

By applying the *Shephard's Lemma*, the associated input cost-share equations are:

$$S_r = \left(\sum_i \delta_{ir} Y_i^{(\pi)} \right) \cdot \left[\alpha_0 + \sum_i \alpha_i Y_i^{(\pi)} + \frac{1}{2} \sum_i \sum_j \alpha_{ij} Y_i^{(\pi)} Y_j^{(\pi)} + \sum_i \sum_r \delta_{ir} Y_i^{(\pi)} \ln P_r \right]^{\tau - 1} + \beta_r + \sum_l \beta_{rl} \ln P_l \qquad [2]$$

The composite specification (PB_C) is obtained by setting $\pi = 1$ and $\tau = 0$. In a similar vein, the well-known *Generalized Translog* (GT) and *Standard Translog* (ST) models, as well as a *Separable Quadratic* (SQ) functional form can be estimated by imposing simple restrictions on the system [1]-[2].[1]

The PB cost functions originate from the combination of the log-quadratic input price structure of the ST and GT specifications with a quadratic structure for outputs. The latter is appropriate to model cost behavior in the range of zero output levels and gives the PB specifications an advantage over the ST and GT forms as far as the measurement of both economies of scope and product-specific economies of scale are concerned. In addition, the log-quadratic input price structure can be easily constrained to be linearly homogeneous.[2]

The relatively few studies which employed the PB specification when examining economies of scale and scope referred to the banking (e.g., Pulley and Braunstein, 1992; Pulley and Humphrey, 1993; McKillop et al., 1996) and telecommunications sectors (e.g., Braunstein and Pulley, 1998; Bloch et al., 2001; McKenzie and Small, 1997). Overall, the composite model proved to be successful in obtaining more stable and reliable estimates than the alternative functional forms. However, in many instances the methodology was employed without devoting particular attention to some relevant aspects, i.e. those concerning model selection and the estimation of the Box-Cox parameters ϕ and τ.

The PB_G model proposes to transform both sides of the cost function - from $C = C(Y, P)$ to $C^{(\phi)} = [C(Y, P)]^{(\phi)}$ - in order to enlarge the set of plausible empirical specifications[3]. The optimal value of ϕ can be found either *i*) by searching over a grid of given ϕ values and judging on the basis of the sum of squared errors (SSE) or *ii*) by direct estimation, resorting to standard non-linear least squares routines. Pulley and Braunstein (1992) were able to use approach *ii*) only for the PB_C and SQ specifications, while they were forced to fix a ϕ value to estimate the general PB_G model. On the other hand, McKillop et al. (1996) relied entirely on the grid-search approach. All

[1] More precisely, the GT model is obtained by setting $\phi = 0$ and $\tau = 1$, while the ST model requires the further restriction $\pi = 0$. The SQ model is obtained from the PB_C specification by adding the restrictions $\delta_{ir} = 0$ for all *i* and *r*.

[2] To be consistent with cost minimization, [1] must satisfy symmetry ($\alpha_{ij} = \alpha_{ji}$ and $\beta_{rl} = \beta_{lr}$ for all couples *i,j* and *r, l*) as well as the following properties: *a*) non-negative fitted costs; *b*) non-negative fitted marginal costs with respect to outputs; *c*) homogeneity of degree one of the cost function in input prices ($\Sigma_r \beta_r = 1$ and $\Sigma_l \beta_{rl} = 0$ for all *r*, and $\Sigma_r \delta_{ir} = 0$ for all *i*); *d*) non-decreasing fitted costs in input prices; *e*) concavity of the cost function in input prices.

[3] *Y* is the vector of outputs and *P* is the vector of input prices.

the other empirical works appeared in the literature did not attempt to estimate the general model, but opted for simplified versions of equation [1][4]. Moreover, the few studies which devoted effort to compare the different alternatives and/or to search for the optimal ϕ did not recur to summary statistics for the system of equations [1]-[2] but looked at the log-likelihoods of the cost function only. Finally, the choice between *non-nested* specifications (i.e. PB_C versus GT) was mainly made by balancing statistical fit with the satisfaction of regularity conditions.

In this paper we estimate the general model PB_G in order to test for the presence of scope and scale economies in a sample of multi-utilities providing gas, water and electricity services. Particular attention is devoted to methodological issues, in that:

- we are able to estimate the general model [1] jointly with the input cost-share equations in [2];
- we use the log-likelihoods for the system [1]-[2] to choose among different *nested* models;
- we recur to an adjusted LR statistics (Vuong, 1989) to select among *non-nested* specifications.

2. Methodology

All the specifications of the multi-product cost function are estimated jointly with their associated input cost-share equations. In our two-inputs case, to avoid singularity of the covariance matrix only the labor equation (S_L) was retained and included in each system. Before the estimation, all variables were standardized on their respective sample medians. Parameter estimates were obtained via a non-linear GLS estimation (NLSUR), which ensures estimated coefficients to be invariant with respect to the omitted share equation.

Assuming the error terms are normally distributed, the concentrated *log-likelihood* for the estimated *cost function* and related *labor-share equation* can be respectively computed via (Greene, 1997)

$$\ln L_C = (\phi - 1)\sum_{t=1}^{T} \ln C_t - \frac{T}{2}[1 + \ln(2\pi)] - \frac{T}{2}\ln\left[\frac{1}{T}\sum_{t=1}^{T}\hat{\psi}_{C\,t}^{2}\right] \qquad [3a]$$

$$\ln L_{S_L} = -\frac{T}{2}[1 + \ln(2\pi)] - \frac{T}{2}\ln\left[\frac{1}{T}\sum_{t=1}^{T}\hat{\psi}_{L\,t}^{2}\right] \qquad [3b]$$

where t is the single observation, $\hat{\psi}_C$ and $\hat{\psi}_L$ are the estimated residuals of the two regressions, and $[(\phi\text{-}1)\ \Sigma_t \ln C_t)]$ is the logarithm of the Jacobian of the transformation of the dependent variable from C_t to $C_t^{(\phi)}$ ($J = \prod_{t=1}^{T} J_t$ with $J_t = \left|\ \partial\psi_{C\,t} / \partial C_t\ \right| = C_t^{\phi - 1}$).

[4] Braunstein and Pulley (1998) and Bloch et al. (2001) set ϕ equal to 0, while McKenzie and Small (1997) imposed the restriction ϕ=1.

Similarly, the concentrated *system log-likelihood* is defined by:

$$\ln L_{(C,S_L)} = \ln J - \frac{T}{2}\left[2(1+\ln(2\pi)) + \ln|\Omega|\right] \qquad [4]$$

where J is the Jacobian of the transformation of (C_t, S_{Lt}) to $(C_t^{(\phi)}, S_{Lt})$, and Ω is the (2×2) matrix of residual sum of squares and cross products for the system, with the *pq*th element of Ω, Ω_{pq}, equal to $\frac{1}{T}\sum_{t=1}^{T}\hat{\psi}_{p_t}\hat{\psi}_{q_t}$ and $p, q = C, S_L$.

Differently from previous studies, we use the *system log-likelihoods* in order to select the preferred specification of the cost function. In the case of *nested models* (i.e PB_C versus SQ and GT versus ST), the standard LR statistic can be used, while for strictly *non-nested* or overlapping models (i.e. GT versus PB_C) we adopt the general procedure developed in Vuong (1989). According to the latter, the standard LR statistic is normalized by a factor $\hat{\omega}_T$, so that the final likelihood ratio statistic is:

$$\mathrm{VLR} = \frac{\mathrm{LR}}{\hat{\omega}_T} = \frac{2(\ln L_{\mathrm{PB}} - \ln L_{\mathrm{GT}})}{\left[\sum_{t=1}^{T}(\hat{\psi}'_{\mathrm{PB}t}\Omega_{\mathrm{PB}}^{-1}\hat{\psi}_{\mathrm{PB}t} - \hat{\psi}'_{\mathrm{GT}t}\Omega_{\mathrm{GT}}^{-1}\hat{\psi}_{\mathrm{GT}t})^2\right]^{\frac{1}{2}}} \qquad [5]$$

where $\hat{\psi}_t$ is for each observation the (2 × 1) column vector of the estimated residuals from the cost function and labor-share equation $\begin{pmatrix}\hat{\psi}_{C_t}\\ \hat{\psi}_{S_{Lt}}\end{pmatrix}$, and Ω is the estimated covariance matrix.

3. An Application to Italian Multi-Utilities

Our sample refers to 90 Italian public utilities which were providers, in combination or as specialised units, of gas, water and electricity services in the years 1994, 1995 and 1996, for a total of 270 pooled observations. Given the presence in the sample of 39 specialised units (19 for gas, 16 for water and 4 for electricity), 37 two-output firms (31 gas-water, 1 gas-electricity and 5 water-electricity combinations) and 14 three-output utilities, we can investigate the presence of economies of scope.

Data on costs, output quantities and input prices are obtained by integrating the information available in the annual reports of each company with additional information drawn from questionnaires sent to managers. Long-run cost (C) is the sum of labor cost and other factors costs, including energy, materials, services and depreciation. The three output categories are: cubic meters of gas (Y_G); cubic meters of water (Y_W); and kilowatt hours of electricity (Y_E). Productive factors are labor (L) and other inputs (O). The price of labor in each utility (P_L) is given by the ratio of total salary expenses to the number of employees. The price of other factors (P_O) is obtained by dividing the costs for energy, materials, services and for the use of physical capital by the length of the network. For multi-utilities inputs prices are obtained by computing a weighted average of prices in each sector.

The summary results of the NLSUR estimations for the general model and the four nested specifications ST, GT, SQ, and PB_C are presented in Table 1. The highest value of the system log-likelihood is 109,75 and is reached by the general model PB_G. The LR tests allow to reject the GT, ST, and SQ specifications, while the restricted composite model PB_C cannot be rejected (critical ${}_{0.01}\chi^2_{(2)}$ = 9.21; computed $\chi^2_{(2)}$ = 2.84).[5] Moreover, pair-wise comparisons show that the PB_C specification adjusts better the observed data than the SQ alternative, and that the GT specification is preferred over the ST one. In order to choose among the *non-nested* specifications which exhibit the highest log-likelihood statistics, we recur to the Vuong's statistics. The value of VLR=6.16 indicates that the composite cost function provides a better description of the technology of multi-utilities than the GT alternative. [6] Finally, the last rows of table 1 show that all the specifications exhibit a good degree of satisfaction of the output and price regularity conditions, with the exception of the SQ model, where 33 cases of negative marginal costs were ascertained.

Table 2 presents the estimates of cost elasticities with respect to outputs ($\varepsilon_{CY_i} = \partial \ln C(Y) / \partial \ln Y_i$, i = G, W, E) and labor price ($\varepsilon_{C_{S_L}} = S_L$), as well as the results for scale and scope economies for the 'median' firm[7]. While in the five estimated models the labor price elasticity ranges from 0.22 to 0.28, the output elasticities show a greater variability, with ST and GT specifications according more weight to gas and SQ, PB_C and PB_G models according more weight to electricity. The relative advantages of the composite specification can be appreciated also by comparing the measures of global scale (*SL*) and scope (*SC*) economies.[8] For the median firm, the former are 0.94 and 1.30 according to the ST and GT models, and 1.10, 1.12, and 1.07 for the SQ , PB_C and PB_G specifications, while economies of scope range from -75% (ST model) to 63.5% (GT model). The PB_C, SQ and PB_G models provide comparable and more plausible values of scope economies of the order of 16%-22%. Finally, the results for *product-specific scope economies* SC_i[9] highlight that the cost advantages of multi-product firms with respect to

[5] In fact in the general model τ is not significantly different from zero while π is very close to one.

[6] The R^2 and log-likelihoods of the cost and labor-share equations, as well McElroy's (1977) $R_*{}^2$ for the NLSUR system, which are reported in table 1, provide further support for the PB models.

[7] The 'median' firm is an hypothetical unit providing about 71 million m^3 of gas, 11 million m^3 of water and 221 million kwh of electricity, and facing median values of input prices.

[8] These are respectively $SL = 1/\sum_i \varepsilon_{CY_i}$ and $SC = \left[C(Y_G,0,0) + C(0,Y_W,0) + C(0,0,Y_E) - C(Y_G,Y_W,Y_E)\right] / C(Y_G,Y_W,Y_E)$.

[9] $SC_i = \left[C(Y_i) + C(Y_{-i}) - C(Y)\right] / C(Y)$, where $C(Y_i)$ and $C(Y_{-i})$ are the costs of producing only output i and the outputs different from i, respectively. $SC_i > 0$ (< 0) indicates a cost disadvantage (advantage) in the "stand-alone" production of output i.

specialised utilities is higher in the case of gas ($SC_G = 0.14$ and $SC_G = 0.11$ according to the PB_G and PB_C respectively)[10].

7. Conclusions

This paper analyses the cost structure of a sample of firms operating, as specialised units or as combination utilities, in gas, water and electricity sectors. The empirical strategy focuses on the Composite cost function model (PB_C) introduced by Pulley and Braunstein (1992). The latter, by combining a log-quadratic input price structure with a quadratic structure for multiple outputs, is more suitable to investigate the presence of economies of scope as compared to the Standard Translog (ST) and Generalised Translog (GT) specifications. After having set the above alternative functional forms within a general specification (PB_G), we carry out LR-type tests in order to select among nested and non-nested models. The results confirm the merits of the PB-type cost functions and show the existence of global and product-specific economies of scope, as well as of global returns to scale. From a policy standpoint, our findings suggest that specialised firms could reduce their costs by evolving into multi-utilities providing network services such as gas, water and electricity.

References

BLOCK H., MADDEN G. AND SAVAGE S. J. (2001), "Economies of Scale and Scope in Australian Telecommunications", *Review of Industrial Organisation*, 18, 219-227.

BRAUNSTEIN Y. M. AND PULLEY L. B. (1998), "Economies of Scale and Scope in Telephony: Applying the Composite Cost Function to Bell System Data", in *Communication and Trade: Essays in Honor of Merhoo Jussawalla*, New Jersey: Hampton Press, 181-192.

FRAQUELLI G., PIACENZA M., AND VANNONI D. (2002), "Scope and Scale Economies in Multi-Utilities: Evidence form Gas, Water and Electricity Combinations", *working paper 'G. Prato',* 64, September, Torino.

GREENE W.H. (1997), "*Econometric Analysis*", third edition, Prentice Hall, New Jersey.

MCELROY M. (1977), "Goodness of Fit for Seemingly Unrelated Regressions: Glahn's R^2 and Hooper's $\overline{R}^2$", *Journal of Econometrics*, 6, 381-87.

MCKENZIE D. J. AND SMALL J. P. (1997), "Econometric Cost Structure Estimates for Cellular Telephony in the United States", *Journal of Regulatory Economics,* 12, 147-157.

MCKILLOP D. G., GLASS C. J. AND MORIKAWA Y. (1996), "The Composite Cost Function and Efficiency in Giant Japanese Banks", *Journal of Banking and Finance*, 20, 1651-1671.

[10] For a more comprehensive analysis of cost properties of Italian multi-utilities, including estimates of aggregate and product specific scale and scope economies for firms of different size-classes, see Fraquelli et al. (2002), where a simplified version of the PB_C model (with ϕ set equal to zero) was estimated.

PULLEY L. B. AND BRAUNSTEIN Y. M. (1992), "A Composite Cost Function for Multiproduct Firms with an Application to Economies of Scope in Banking", *Review of Economics and Statistics*, 74, 221-230.

PULLEY L. B. AND HUMPHREY D. B. (1993), "The role of fixed costs and cost complementarities in determining scope economies and the cost of narrow banking proposals", *The Journal of Business*, 66, 3, 437-462.

VUONG Q. H. (1989), " Likelihood Ratio Tests for Model Selection and non-nested Hypotheses", *Econometrica*, 57, 2, 307-333.

Table 1. NLSUR estimation: General (PB_G), Composite (PB_C), Separable Quadratic (SQ), Generalized Translog (GT) and Standard Translog (ST) cost function models [a]

	PB_G MODEL	PB_C MODEL	SQ MODEL	GT MODEL	ST MODEL
Box-Cox Parameters					
ϕ	0.3182 (0.039)	0.3158 (0.038)	0.2369 (0.041)	0	0
π	1.1105 (0.087)	1	1	0.1276 (0.028)	0
τ	-0.0854 (0.075)	0	0	1	1
Cost function					
Log-likelihood	-149.65	-150.33	-151.46	-227.76	-234.42
R^2	0.9146	0.9137	0.9010	0.7850	0.7742
Labor-share equation					
Log-likelihood	246.22	243.89	198.27	233.68	234.86
R^2	0.5171	0.5089	0.3113	0.4703	0.4748
System log-likelihood	109.75	108.33	60.14	13.71	5.45
Goodness of fit [b]	0.8677	0.8679	0.8436	0.7114	0.6959
LR test statistics					
PB_G versus other models	--	2.84	99.22 [c]	192.08 [c]	208.61 [c]
PB_C versus SQ	--	--	96.38 [c]	--	--
GT versus ST	--	--	--	--	16.53 [c]
VLR test statistic [d]					
PB_C versus GT	--	--	--	6.16 [c]	--
Satisfaction of regularity conditions					
Output regularity violations	1.1%	7.2%	11.9%	1.4%	6.1%
Price regularity violations	0	0	4.3%	0	0

[a] The estimated asymptotic standard errors presented in parentheses have been computed using the 'delta' method (see Greene, 1997, pages 278-280).
[b] The goodness-of-fit measure for the NLSUR systems is McElroy's (1977) R_*^2.
[c] The null hypothesis that the two models fit equally well the data is rejected at the 1% level of significance.
[d] See Vuong (1989). The VLR statistic is distributed as a $N(0,1)$.

Table 2. Cost elasticites with respect to outputs and factor prices, global economies of scale, and global and product-specific economies of scope: estimates for PB_G, PB_C, SQ, GT and ST models (at median values of outputs and factor prices)*

	PB_G	PB_C	SQ	GT	ST
Output elasticities					
ε_{CYG}	0.281	0.273	0.208	0.363	0.449
	(0.041)	(0.032)	(0.027)	(0.050)	(0.042)
ε_{CYW}	0.164	0.167	0.245	0.201	0.298
	(0.036)	(0.037)	(0.029)	(0.045)	(0.045)
ε_{CYE}	0.492	0.456	0.457	0.208	0.320
	(0.043)	(0.028)	(0.026)	(0.058)	(0.068)
Factor price elasticities					
S_L	0.284	0.284	0.223	0.241	0.232
	(0.015)	(0.014)	(0.011)	(0.026)	(0.025)
S_O	0.716	0.716	0.777	0.759	0.768
	(0.015)	(0.014)	(0.011)	(0.026)	(0.025)
Global scale economies					
SL	1.067	1.116	1.099	1.296	0.938
	(0.060)	(0.042)	(0.032)	(0.156)	(0.071)
Global scope economies					
SC	0.222	0.181	0.157	0.639	-0.753
	(0.049)	(0.041)	(0.038)	(0.246)	(0.062)
Product-specific scope economies					
SC_G	0.137	0.114	0.120	0.926	-0.606
	(0.040)	(0.035)	(0.028)	(0.363)	(0.083)
SC_W	0.107	0.086	0.065	0.795	-0.161
	(0.032)	(0.032)	(0.025)	(0.417)	(0.223)
SC_E	0.122	0.098	0.098	0.543	-0.650
	(0.039)	(0.035)	(0.030)	(0.230)	(0.085)

* Estimated asymptotic standard errors in parentheses. The subscripts are G = gas, W = water, E = electricity, L = labor, O = other inputs.