NBER WORKING PAPERS SERIES

QUALITY ADJUSTED COST FUNCTIONS

Paul J. Gertler

Donald M. Waldman

Working Paper No. 3567

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 1990

We thank Steve Garber, Gary Hunt, Jacob Klerman, Bridger Mitchell, V. K. Smith, and participants of seminars at University of Colorado, Johns Hopkins University, and the RAND Corporation for helpful discussions. Gertler gratefully acknowledges the financial support of the National Institute of Aging, Grant #1PO1AG08291-01. This paper is part of NBER's research program in Health Economics. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #3567
December 1990

QUALITY ADJUSTED COST FUNCTIONS

ABSTRACT

We propose a simple method for estimating cost functions in the presence of endogenous and unobserved quality. The theory of production, the equilibrium conditions implied by optimizing behavior, and exogenous influences on product demand are used to identify the model. An important advantage of the method is that the data requirements, above those necessary for standard cost function estimation, are minimal and the data are usually readily available. Specifically, exogenous information that influences the demand for the firm's product is required.

We apply this method to estimate quality-adjusted cost functions in the nursing home industry. Estimation of a translog cost function that ignores quality yields seriously misleading estimates of marginal cost and economies of scale. In particular, while estimation of a quality-exogenous cost function reports economies of scale, estimation of a quality adjusted cost function reveals diseconomies of scale for high quality homes, constant returns to scale for average quality homes, and economies of scale for low quality homes.

Paul J. Gertler
The Rand Corporation
1700 Main Street
P.O. Box 2138
Santa Monica, CA 90406

Donald M. Waldman
Department of Economics
University of Colorado
Campus Box 256
Boulder, CO 80309-0256

1. *INTRODUCTION*

The structure of production is a cornerstone of economic research and crucial to the implementation of regulatory policy. Indeed, economists regularly estimate cost functions to learn about marginal cost, elasticities of substitution, and economies of scale which are then used to make inferences about firm behavior and economic efficiency. The vast majority of these studies, however, have treated product quality as an unobserved factor that is assumed to be either constant or uncorrelated with included variables.[1] If this assumption is correct, standard analyses produce consistent estimates of important parameters. In many cases, however, this assumption is almost surely incorrect, implying that conventional cost function estimates do not provide accurate representations of the structure of production. In this paper, we propose a novel method for estimating cost functions in the presence of unobserved and endogenous quality, and use that method to estimate quality-adjusted cost functions for nursing homes.

The fact that economists routinely ignore the possibility of endogenous quality in cost function estimation is surprising. The notion that firms product-differentiate and engage in non-price competition has been an active area of research since Chamberlain (1933). Empirically, implicit markets for endogenous product quality characteristics have been studied extensively in the hedonic pricing literature since Rosen (1974). Moreover, a significant amount of attention has been paid to the performance and welfare implications of endogenous quality responses to regulatory policy.[2] Indeed, in the presence of price regulation, one should expect even greater endogenous variation in product quality, since quality characteristics are the only strategic variables that these firms can use to pursue profit maximization. In particular, endogenous quality responses to price regulation in the nursing home industry have been explored in Dusansky (1989), Nyman (1985), and Gertler (1989).

Braeutigam and Pauly (1986) is one of the few articles that explores the implications of endogenous product quality in the estimation of cost functions. They argue that one should consider endogenous quality to be the standard case in cost function estimation,

[1] A few cost function studies have included observed or hedonic quality measures in estimation as exogenous variables. These include Friedlander and Spady (1980), Braeutigam, Daughety and Trunquist (1982, 1984), Fuss and Waverman (1981), and Christensen, Cummings and Schoech (1980). Friedlander and Spady (1980) raise the issue of the endogeneity of quality. They note, therefore, that their parameter estimates represent a "behavioral" rather than structural cost function.

[2] Examples of studies that focus on endogenous quality responses to price regulation include Posner (1971), Stigler (1971), White (1972), Rosse and Panzar (1974), Douglas and Miller (1974), Panzar (1975), Spence (1975), Schmalensee (1977), Joskow and Noll (1981), Nyman (1985), Gertler (1989), Allen and Gertler (1990).

especially in regulated industries. Further, they demonstrate that, as a result of the firm's optimization process, product quality must be correlated with included right-hand side variables in any empirical cost function specification. Failure to account for quality will result in biased cost function estimates and, therefore, invalid characterizations of the firm's structure of production. They provide a very simple and useful specification test to determine if cost function estimates suffer from bias due to endogenous and unobserved quality. The value of their test is that it does not require one to directly observe quality characteristics. Braeutigam and Pauly use this test to demonstrate that cost functions in the regulated automobile insurance industry suffer from such bias.

If quality is so important, why is it often ignored in cost function estimation? Perhaps the major reason is data problems. Quality is usually defined in terms of the characteristics of goods other than the physical units in which the good is priced (Lancaster, 1976; Leffler, 1982). In many industries, and especially in service sectors, the number of quality characteristics can be quite large and difficult to measure. Typically, the data requirements for such endeavors are onerous and data collection itself can be quite costly.

The quality-adjusted cost function proposed below treats quality as endogenous and unobserved, and does not require large amounts of additional data. The theory of production, the equilibrium conditions implied by optimizing behavior, and exogenous influences on product demand identify the model. An important advantage of the quality-adjusted cost function is that the data requirements, above those necessary for standard cost function estimation, are minimal and the data are usually readily available. Specifically, exogenous information that influences the demand for the firm's product is required. Thus, as in hedonic pricing models, cross-market variation in the exogenous determinants of product demand is needed for identification (see Brown and Rosen, 1982). The same type of information is employed by Deaton (1988), who uses demographic differences across geographic locations to correct prices for quality differences.

This method is used to estimate quality-adjusted cost functions in the nursing home industry. We find that the estimation of a quality-exogenous translog cost function yields seriously misleading estimates of marginal cost and economies of scale. When quality is ignored, the estimates are consistent with recent nursing home cost function studies (Dor, 1989; McKay, 1988; Nyman, 1989) in suggesting large economies of scale. However, estimates of a quality-adjusted cost function show diseconomies of scale for high quality homes, constant returns to scale for average quality homes, and economies of scale for low quality homes.

2. *THE BEHAVIORAL MODEL*

This section describes the behavioral model assumed to produce the observed data on firm output and costs, and the unobserved (to the econometrician) differences in product quality. Consider a firm selling a multidimensional product. The product is sold in units measured by one of the product's attributes. This attribute defines the quantity units in which the product is sold and the other attributes characterize the product's quality. For example, two attributes of milk are volume and percent butterfat. Milk is sold in quarts, implying that volume is quantity and percent butterfat is quality. For the distinction to have nontrivial implications for cost function estimation, the firm must spend resources to produce higher quality, and consumers must be willing to pay more for a higher quality product.

The firm's demand function is defined in terms of its price, quantity, and quality. Changes in quality attributes shift the demand curve. This notion of quality is consistent with all market structures where there is product differentiation, including monopoly, oligopoly, monopolistic competition, and the competitive specifications of the hedonic pricing literature. As in the hedonic literature, consumers are assumed to have a range of tastes for quality, implying that firms producing products with differing quality levels can coexist in the same market. In addition to price and quality, the quantity demanded from the firm is influenced by cross-market differences in characteristics of consumers (for example, population size, income, and education) and the degree of competition.

Suppose that the product has only two attributes (quantity and quality) or that all of the quality attributes can be aggregated into a single index of quality. Let Q be the quality level per unit output. Then, the firms demand function is given by:

$$Y = Y(Q, P, Z) \quad , \tag{1}$$

where P is product price (per unit quantity), and Z is a vector of variables exogenous to the firm that shift the demand curve. Demand is assumed to be falling in price and rising in quality, while the effect of changes in Z depend upon the definitions of these variables.[3]

[3] This demand function applies to monopolistic and monopolistically competitive market structures. In oligopolistic markets, demand will also depend upon the price and quality of other firms in the market. The application of our methods to those markets does not present new problems, as indicated below.

Given a transformation (production) function and input prices, the cost function is derived by choosing inputs to minimize the costs of producing quantity Y at quality Q. Let W be a vector of exogenous factor input prices. The cost function is:

$$C = C(Y, Q, W) \quad , \tag{2}$$

where cost is assumed to be increasing in all its arguments.

The firm chooses price and quality to maximize profits, Π:

$$\Pi = P\,Y(Q, P, Z) - C[\,Y(Q, P, Z), Q, W\,] \;, \tag{3}$$

The levels of P and Q that maximize (3) are assumed to satisfy the first order conditions:

$$\frac{\partial \Pi}{\partial P} = 0 \quad \rightarrow \quad P\frac{\partial Y}{\partial P} + Y = \frac{\partial C}{\partial Y}\frac{\partial Y}{\partial P} \;, \tag{4}$$

$$\frac{\partial \Pi}{\partial Q} = 0 \quad \rightarrow \quad P\frac{\partial Y}{\partial Q} = \frac{\partial C}{\partial Y}\frac{\partial Y}{\partial Q} + \frac{\partial C}{\partial Q} \;. \tag{5}$$

The price condition (4) equates marginal revenue from a price change to marginal cost, while the quality condition (5) equates the marginal revenue from increasing quality to the marginal cost of quality. Equations 1, 2, 4, and 5 can be solved for the endogenous variables P and Q in terms of the exogenous variables W and Z:

$$P = P(W, Z) \;, \tag{6}$$

$$Q = Q(W, Z) \;. \tag{7}$$

Equations 1, 2, 6, and 7 determine an equilibrium output, price, quality, and generate observed costs given by equation 2. The parameters of the cost function are the objects of estimation.

3. *EMPIRICAL SPECIFICATION*

In this section, the behavioral model is used to derive the empirical specification and discuss identification. With observed quality, the empirical model consists of the structural demand and cost functions (1 and 2), and the equilibrium price and quality equations, (6 and 7). The behavioral theory implies a set of identifying restrictions on the parameters of these equations. Namely, the Z are excluded from the cost function (2). Variation in Z shifts the product demand function and the equilibrium price and quality equations. Therefore, (2) is identified and may be estimated once functional form assumptions are made.

The difficulty in identifying the parameters of the cost function arises not only because both quantity and quality are endogenous, but also because quality is unobserved. We show that the cost function is identified in a system that includes the cost function (2), the equilibrium quality equation (7), and an expression for quantity in terms of quality and other exogenous variables. Identification will be achieved through a set of parametric restrictions implied by the first order conditions, information that affects product demand but not costs, and nonlinearities in the cost function.

We begin by deriving the empirical specification. To obtain an expression for Y conditional on Q, substitute (6) into (1) to give[4]

$$Y = f(Q, W, Z) \quad . \tag{8}$$

Since Q is unobserved, it is necessary to substitute (7) into (2) and (8) to produce the empirical model:

$$C = C[Q(W, Z), Y, W] \quad , \tag{9}$$

$$Y = Y[Q(W, Z), W, Z] \quad . \tag{10}$$

[4]This conditional demand function (8) is consistent with all market structures (see previous footnote). In oligopolistic markets, prices and qualities of other firms enter the demand function. The equilibrium values of these prices and qualities will be functions of the W's and Z's faced by these firms. Since firms in the same market face the same W's and Z's, the prices and qualities of the other firms have already been implicitly substituted for in the conditional demand function.

Since the Q(·) function appears in both (9) and (10), the parameters on the determinants of quality can only differ across (9) and (10) by a factor of proportionality. This factor is the ratio of the marginal cost of quality to the marginal change in demand with respect to quality. These proportionality restrictions apply regardless of the functional form of the cost and product demand functions.

The proportionality restrictions and the availability of variables exogenous to the firm that shift the product demand function are necessary to identify the model. The first order conditions imply that the optimal level of quality depends upon both exogenous supply and demand factors, whereas the cost function depends only on input prices and quality. Thus, conditional on quantity and input prices, the variation in costs across firms that is *correlated* with exogenous determinants of product demand reflects quality variation. That is, it reflects firms' quality (resource) responses to different demand structures. [5,6]

Although the identifying information Z also enters the right hand side of (9) via the Q(·) function, its coefficients are restricted by the cross-equation proportionality constraints. These restrictions are the same as those that characterize the "multiple indicator multiple cause" (MIMIC) latent variable model (see Joreskog and Goldberger, 1975). In this MIMIC specification the indicators of the latent variable quality are (1) costs filtered for variation due to quantity and input prices and (2) demand filtered for variation due to exogenous determinants. The remaining variation in filtered costs and demand are functions solely of quality and random disturbances. Thus, the indicators of quality, filtered costs and demand, have the same covariation with the causes (determinants) of quality but are measured in different units. The causes of the latent variable quality are exogenous supply and demand factors.

An additional problem arises in the latent variable framework when causes of the latent variable also directly affect the observable outcomes. This is the case in our model because the latent variable Q is determined by a reduced form equation which necessarily contains all exogenous variables. Therefore, input prices are common to the cost function (2) and the

[5] This is basically the same idea put forth in Braeutigam and Pauly (1986) to develop their test for the endogeneity of unobserved quality in cost function estimation. They noted that if quality is endogenous and omitted from the cost function, then, by substituting the reduced-form quality function into the cost function, price and exogenous determinants of demand will affect costs.

[6] Variations in cost or input demands after controlling for quantity and input prices are generally attributed to technical inefficiency. However, while this source of variation may be related to input usage, it should not be related in any way to demand variables.

equilibrium quality function (7). The result of this is that there will be fewer reduced form estimates than structural parameters, a violation of the order condition for identification. This can be seen by specifying linear cost, demand, and equilibrium quality equations and solving for the reduced form. The conditions for identification can be satisfied in these cases by the introduction of nonlinearities into the cost function. Nonlinearities, such as the square of output or the interaction of quantity and quality, are natural in the cost function, where, in fact, a linear specification would be unusual.[7] These nonlinearities add higher order moments of Z and interactions of Z and W to the reduced form which provide the additional information necessary for identification.[8] The model is identified for all of the flexible functional forms since they include higher order moments.

Overidentifying information is available from the input demands. By Shephard's Lemma the demand for input i is

$$X_i = X_i\,[Q(W, Z), Y, W] \quad , i = 1, \ldots, n \quad , \tag{11}$$

where n is the number of inputs. Since the $Q(\cdot)$ function appears in equations (11) as well as in (9) and (10), the parameters on the determinants of quality can only differ across these equations by a factor of proportionality. The parameters of the share equations are a subset of the parameters of the cost function. Since the input demand functions add no new structural parameters, the addition of (11) adds overidentifying information and efficiency in estimation.

As in all latent variable models, the model is identified only up to an arbitrary factor of proportionality (Aigner et al, 1984). This is because the latent variable quality has no units of measurement. Without loss of generality, an arbitrary normalization defines a metric for quality and completes identification of the model. The estimates of the structure of production are invariant to the normalization.

[7] This identification problem and solution is very similar to the identification issue in hedonic models as first noted by Brown and Rosen (1982).

[8] The model specified above is identified for a wide class of technologies, including all of the flexible functional forms. Excluded from this list, however, are linear cost functions. Linear cost functions are, of course, very restrictive technologies, where the marginal costs of quantity and quality are constant, independent of one another, and independent of the input prices. This imposes constant returns to scale and constant marginal costs. Linear cost functions also rule out, *a priori*, the likely property that the marginal cost of producing another physical unit of output, though constant, is different for firms producing products of different quality levels.

4. *ECONOMETRIC METHODS*

The estimation of a quality-adjusted cost function requires specification of functional forms for the cost, conditional demand, and equilibrium quality equations as well as stochastic assumptions. We estimate a quality-adjusted translog cost function. The firm is assumed to observe a series of shocks (ϵ_i) that influence its decision making (and hence its costs), but these shocks are unobserved to the investigator (McElroy 1986). Treating quantity and quality as the two attributes of output, the translog cost function with n inputs is:

$$\log C = \alpha_0 + \alpha_y \log Y + \alpha_q \log Q + \sum_{i=1}^{n} \alpha_{wi} \log W_i + \frac{1}{2} \beta_{yy} (\log Y)^2$$
$$+ \beta_{yq} \log Y \log Q + \frac{1}{2} \beta_{qq} (\log Q)^2 + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} \log W_i \log W_j$$
$$+ \sum_{i=1}^{n} \beta_{yi} \log Y \log W_i + \sum_{i=1}^{n} \beta_{qi} \log Q \log W_i + \sum_{i=1}^{n} \epsilon_i \log W_i + \tau \qquad (12)$$

where C is cost, Y is output, Q is quality, the W_i are input prices, and τ is white noise. The corresponding share equations are:

$$S_i = \alpha_{wi} + \sum_{j=1}^{n} \beta_{ij} \log W_j + \beta_{yi} \log Y + \beta_{qi} \log Q + \epsilon_i \quad , \quad i = 1,\ldots,n. \qquad (13)$$

As a result of the McElroy error structure, the share equation disturbances are additive and logically consistent with the cost function error specification. The variance-covariance matrix of the cost function and the share errors is unrestricted. However, errors are assumed to be uncorrelated across firms.

The parameters of the cost and share equations are constrained by the usual symmetry restrictions ($\beta_{ij}=\beta_{ji}$), and homogeneity of degree one in input prices implies:

$$\sum_{i=1}^{n} \alpha_{wi} = 0; \quad \sum_{i=1}^{n} \beta_{ij} = \sum_{j=1}^{n} \beta_{ij} = 0; \quad \sum_{i=1}^{n} \beta_{yi} = 0; \text{ and } \sum_{i=1}^{n} \beta_{qi} = 0 \quad .$$

These constraints do not place restrictions on measures of scale economies or input price elasticities.

We also specify a logarithmic conditional product demand function:

$$\log Y = \gamma_0 + \gamma_q \log Q + \sum_{i=1}^{L} \gamma_{wi} \log W_i + \sum_{i=1}^{K} \gamma_{zi} \log Z_i + \upsilon \qquad (14)$$

where Z_i (i = .,...,K) are exogenous factors that shift the demand function and υ is a random disturbance.

To complete the functional form assumptions, we specify a logarithmic equilibrium quality equation:

$$\log Q = \sum_{i=1}^{K} \lambda_{zi} \log Z_i + \sum_{i=1}^{n} \lambda_{wi} \log W_i + \epsilon \quad , \qquad (15)$$

where ϵ is a random disturbance. As in most latent variable models (see Aigner et al., 1984; Joreskog and Goldberger, 1975) the intercept is not identified.

While the first-order term of the unobservable, log Q, adds information to the reduced-form cost, share and product demand functions, the second-order term $(\log Q)^2$ does not. That is, its inclusion in the cost function does not add to the number of reduced-form exogenous variables in any of the equations. Its presence merely duplicates all terms quadratic in log W and log Z and all terms in log W·log Z, due to the presence of log Y·log W, $(\log Y)^2$, and log Y·log Q in the cost function. The addition of this term does, however, add a substantial number of nonlinear constraints on the parameter space, which serve mostly to complicate optimization in an already overidentified model. This greatly increases the computational burden of estimation. It is not clear that the additional flexibility of the cost function due to the presence of a term in $(\log Q)^2$ is worth the cost imposed upon estimation. While it is possible to identify its coefficient, β_{qq}, our intuition is that we are asking a lot of the data to distinguish both first and second order terms in unobservables. Therefore, we simplify the model by setting β_{qq} equal to zero.

Since quality is unobserved, it must be eliminated from the cost, share and product demand equations by substitution, producing:

$$\log C = \alpha_0 + \alpha_y \log Y + \sum_{i=1}^{n} (\alpha_{wi} + \alpha_q \lambda_{wi}) \log W_i$$

$$+ \frac{1}{2} \beta_{yy} (\log Y)^2 + \sum_{i=1}^{n} (\beta_{yi} + \beta_{yq} \lambda_{wi}) \log Y \log W_i$$

$$+ \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (\beta_{ij} + 2\beta_{qi} \lambda_{wj}) \log W_i \log W_j + \alpha_q \sum_{i=1}^{K} \lambda_{zi} \log Z_i$$

$$+ \beta_{yq} \sum_{i=1}^{K} \lambda_{zi} \log Z_i \log Y + \sum_{i=1}^{n} \beta_{qi} \sum_{j=1}^{K} \lambda_{zj} \log Z_j \log W_i + \zeta_c \qquad (16)$$

$$S_\ell = \alpha_{w\ell} + \sum_{j=1}^{n} (\beta_{\ell j} + \beta_{q\ell} \lambda_{wj}) \log W_j + \beta_{y\ell} \log Y$$

$$+ \beta_{q\ell} \sum_{j=1}^{K} \lambda_{zj} \log Z_j + \zeta_{s\ell}, \qquad \ell = 1, \ldots, n, \qquad (17)$$

$$\log Y = \gamma_0 + \sum_{i=1}^{n} (\gamma_{wi} + \gamma_q \lambda_{wi}) \log W_i + + \sum_{i=1}^{K} (\gamma_{zi} + \gamma_q \lambda_{zi}) \log Z_i + \zeta_y \qquad (18)$$

where ζ_c, $\zeta_{s\ell}$, and ζ_y are functions of ϵ, υ, and ξ, and the parameters, and are heteroskedastic.

The model is estimated with a minimum distance estimator (see Chamberlain, 1982). This estimator is especially well-suited for a system of equations with nonlinear, cross-equation restrictions and general (unknown) forms of heteroskedasticity. Chamberlain shows that the estimator is consistent and asymptotically normal. Further, he proves that it is efficient for this type of model. Implementation of the estimator involves minimizing a function of the distance between the vectors of the structural parameters and reduced form coefficients. Thus, expressions are needed for the reduced form coefficients in terms of the structural parameters. The reduced form cost and share functions are obtained by substituting equation 18 into equations 16 and 17. These equations along with the reduced form product demand equation (18) produce a vector of reduced-form coefficients on all first- and second- order moments of the exogenous variables in the vectors log W and log Z, and a constant. In obvious notation these moments and the solution for the reduced-form coefficients in terms of the structural parameters are displayed in Table 1. In Appendix A we provide the distance function that is minimized and the properties of the resulting estimator.

5. *EMPIRICAL EXAMPLE: THE CASE OF NURSING HOMES*

To demonstrate the importance of the endogenous quality, we estimate a quality-adjusted cost function for nursing homes. The nursing home industry is one in which endogenous quality is of real concern and knowledge of the structure of production is of policy interest. Indeed, recent studies have focused on nursing homes quality responses to changes in price regulation and competition (Dusansky, 1989; Gertler, 1989: Nyman 1985). Since the government pays for approximately 50% of all nursing home care through the Medicaid program, understanding the structure of nursing home production is crucial for setting Medicaid reimbursement rates.

The specification is a three-input, quality-adjusted cost function: where the inputs are skilled labor, unskilled labor, and supplies.[9] Quantity is measured in annual patient days (divided by 365). The exogenous demand information includes the population over age 65 in the market, per capita income in the market, the Herfindahl index of market concentration (competitiveness), the Medicaid reimbursement rate, and a "casemix" index of health status. The data come from a 1980 survey of New York State nursing homes (New York State, 1980). The sample includes 279 proprietary nursing homes. Descriptive statistics for these variables are provided in Table 2. The selection of the sample and the construction of the variables are described in Appendix B.

Total output, Y, is exogenous in the nursing home industry (see Scanlon, 1980; Vogel and Palmer, 1983; Nyman, 1985; Dusansky, 1989; Gertler, 1989). Government regulation imposes a capacity constraint on nursing homes that is binding. Nursing homes can fill their capacity with two types of patients: those who finance their care privately (private pay patients) and those whose care is financed through a government entitlement program (primarily Medicaid). Homes can charge private-pay patients what the market will bear and receive the Medicaid reimbursement rate for Medicaid patients. However, homes cannot legally discriminate by method of payment in their provision of services, so that the same quality is supplied to both private-pay and Medicaid patients. Whereas private-pay demand depends on both price and quality, Medicaid demand is perfectly elastic at the Medicaid reimbursement

[9] In this application we have ignored capital stock, which is usually treated as a quasi-fixed factor. A nursing home's capital stock mainly consists of the building and there are very limited substitution possibilities between labor and capital. We experimented with using the building's area in square feet per bed as a measure of a nursing home's capital stock. None of the additional coefficients were statistically significant at conventional levels nor were there any substantive changes in existing coefficient estimates.

rate. Homes use price and quality to compete for private-pay patients knowing that they can always fill excess capacity with Medicaid patients at the Medicaid reimbursement rate. Thus, price and quality choices determine the mix of private and Medicaid patients, while the total number of patients is determined exogenously by regulation. In this case, the cost function can be identified without the conditional demand equation. Therefore, equation (18) is not estimated – the empirical model reduces to equations (16) and (17).

Estimates are presented for both quality-adjusted and quality-exogenous cost functions. The quality-exogenous model treats quality as an unobserved factor that is uncorrelated with the right-hand side variables in the cost and share equations, and therefore is subsumed in the error terms. The parameter β_{q1} has been normalized to one in the quality-adjusted cost function. The coefficients and asymptotic t-statistics for the quality-adjusted cost function and equilibrium quality equation are presented in the first and third columns of Table 3, and the estimates of the quality-exogenous cost function are presented in the second column.

The quality-adjusted estimation results agree nicely with economic theory and are generally precisely estimated as indicated by the t-statistics. The economic interpretation of the results is discussed in detail below. Braeutigam and Pauly (1986) show that there is unobserved and endogenous quality bias if one can reject the null hypothesis that the coefficients of the demand variables in the cost function are zero. This null hypothesis is overwhelmingly rejected, leading us to prefer the quality-adjusted estimates.

The estimated equilibrium quality equation (Table 3, third column) has great intuitive appeal. The first two coefficients, both negative and significantly different from zero, measure the effect of input prices on quality. Increases in input prices raise the marginal cost of quality. Therefore, the results are consistent with the notion that as the marginal cost of quality rises, firms reduce quality. The next six coefficients represent demand-side variables. The coefficients on population, income and their interaction are jointly significant. They indicate that homes in markets where the aged population is larger and incomes are higher supply greater quality. The negative sign of the interaction indicates that while the total effect of increasing income on quality is positive, the size of the effect is smaller in markets with larger elderly populations. Similarly, while the total effect of population is positive, the size of the effect is smaller in markets with greater per capita income. The coefficient on case-mix indicates that homes supply sicker patients with more quality. The results also show that quality is lower in more concentrated markets, indicating that more competition results in higher quality. Finally, the coefficient on the Medicaid reimbursement rate is negative and significantly different from zero. This result is consistent with the theoretical and empirical literature on nursing home behavior (Nyman 1985; Dusansky 1989; Gertler 1989).

There are substantive differences between the parameter estimates for the quality-exogenous and the quality-adjusted cost functions. First, the quality-adjusted model has five additional parameters corresponding to quality. One of those is the parameter β_{ql}, which has been normalized to one. The estimates of three of the remaining four parameters are significantly different from zero. In addition, inspection of the first two columns of Table 3 reveals that there are important differences in sign and magnitude between coefficient estimates on variables common to both models.

As for all flexible functional forms, the interpretation of the coefficients is difficult. A more meaningful comparison of the bias from ignoring quality is in terms of estimates of the structure of production: economies of scale and average and marginal cost curves.

Consider economies of scale, which are measured by $1-\partial \ln C/\partial \ln Y$ (Christensen and Greene, 1976). In the quality-exogenous case, scale economies are estimated to be 0.061 at the mean of the data. This implies that a doubling of patients will reduce average cost by approximately 6%. These results are very similar to those found by Dor (1989) using national data, McKay (1988) using Texas data, and Nyman (1989) using 1983 New York data.

Economies of scale measures in the quality-adjusted model tell quite a different story. Since economies of scale depend upon the level of quality, measures were calculated for homes that supply different quality. Low, average, and high quality were obtained by predicting quality for each of the homes using the estimated reduced-form quality equation. The minimum, mean, and maximum of these predictions were used as the measures of low, average, and high quality. While we find economies of scale for low-quality homes, the cost function exhibits constant returns to scale at average quality and diseconomies of scale at high quality. Specifically, scale economies for the average size home are .065 for low quality, -.001 for average quality, and -.044 for high quality. The presence of diseconomies of scale in high quality homes is consistent with the conventional wisdom for nursing homes. Higher quality is produced through labor intensive activities such as personal contact with patients by employees, and highly personalized physical and psychological therapy. These activities become more difficult to manage as the home becomes large.

To further illustrate the differences, Figure 1 is a graph of the average cost per patient day for both the quality-adjusted and quality-exogenous models. The quality-exogenous average cost curve is downward sloping indicating economies of scale. Three quality-adjusted average cost functions are drawn; one for a low quality home, a second for an average quality home, and third for a high quality home. The average cost curve for the low quality home is downward sloping, indicating economies of scale. For average quality, the curve is flat, indicating constant returns to scale, while for high quality the curve is rising, indicating diseconomies of scale. Note that as quality rises, the slope of the average cost curve increases. This fanning of the average cost curves in the quality-adjusted case is missed when quality is taken to be exogenous.

In the empirical application, the results demonstrate that a quality-exogenous translog cost function yields seriously misleading estimates of the structure of production. There are three pieces of evidence that support this conclusion. First, consistent with the quality endogenous theoretical model, exogenous demand influences were found to be significant determinants of cost, after controlling for quantity and input prices. Second, the parameter estimates of the quality-endogenous model differed substantively from those of the quality-exogenous model. Finally, while the quality-exogenous model reported economies of scale, the quality-endogenous model revealed that economies of scale differ by quality level.

Appendix A

This Appendix shows how the minimum distance estimator is applied to the model. The presentation follows Chamberlain (1982) closely; reference is made to that paper for proofs. The discussion of the minimum distance estimator is facilitated by simplifying the notation. Define $\mathbf{Y}_i = (\log C_i, S_{1i}, \ldots, S_{ni}, \log Y_i)$, and let $\mathbf{X}_i$ be the vector of the explanatory variables (the first column of Table 1), where i indexes observations. Define $\Pi' = (\pi_1', \ldots, \pi_{n+2}')$ and $\Delta = (\pi_1, \ldots, \pi_{n+2})$, where π_ℓ is the vector of parameters in reduced form equation ℓ, for $\ell = 1, \ldots, n+2$. The system of reduced form equations can be written

$$\mathbf{Y}_i' = \Delta'\mathbf{X}_i + \eta_i' \qquad (19)$$

where $\eta_i = (\eta_{1i}, \ldots, \eta_{n+2,i})'$, and $\eta_{\ell i}$ is the error term in the ℓ^{th} reduced-form equation for individual i. Arrange the structural parameters in a vector θ. Then the constraints expressed in Table 1 can be represented by the functions $\Pi = f(\theta)$.

The minimum distance estimator chooses θ to minimize the length of $\hat{\Pi} - f(\theta)$ in the metric of $V(\eta_i)$, the covariance matrix of the vector of reduced-form errors. In order to derive the exact distance function to be minimized, let $\hat{\Pi}$ be the least squares estimator of Π, and let $\hat{\Delta}$ be analogously defined. Then

$$\sqrt{T}\,(\hat{\Pi} - \Pi) \xrightarrow{a} N(0,\Omega)$$

where $\Omega = E[V(\eta_i \mid \mathbf{X}_i) \otimes \Phi^{-1}(\mathbf{X}_i\mathbf{X}_i')\Phi^{-1}]$, $V(\eta_i \mid \mathbf{X}_i)$ is the conditional variance of η_i given $\mathbf{X}_i$, $\Phi = E(\mathbf{X}_i\mathbf{X}_i')$, and there are T observations. A consistent estimator of Ω is

$$\hat{\Omega} = T^{-1}\sum_{i=1}^{T} [(\mathbf{Y}_i - \hat{\Delta}'\mathbf{X}_i)(\mathbf{Y}_i - \hat{\Delta}'\mathbf{X}_i)' \otimes \hat{\Phi}^{-1}(\mathbf{X}_i\mathbf{X}_i')\,\hat{\Phi}^{-1}] \quad , \qquad (20)$$

where $\hat{\Phi}^{-1} = T^{-1}\sum_{i=1}^{T} \mathbf{X}_i\mathbf{X}_i'$.

The minimum distance estimator of θ is $\hat{\theta}$ chosen to minimize

$$g(\theta) = [\hat{\Pi} - f(\theta)]'\hat{\Omega}^{-1}[\hat{\Pi} - f(\theta)] \quad . \tag{21}$$

The minimum distance estimator $\hat{\theta}$ is consistent with limiting distribution

$$\sqrt{T}(\hat{\theta} - \theta) \overset{a}{\rightarrow} N(0, F'\hat{\Omega}^{-1}F) \quad ,$$

where $F = \dfrac{\partial f(\theta)}{\partial \theta'}$.

In addition, the criterion function for the minimum distance estimator serves as a test of the model restrictions. Given the number of structural and reduced form parameters discussed above, under the null hypothesis that the restrictions are valid we have

$$g(\hat{\theta}) \overset{a}{\rightarrow} \chi^2_{(J)} \quad ,$$

where $J = n^2 + nK + K(K+1)/2 - 2n - 3$.

Appendix B

The data are constructed from New York State's 1980 survey of Long Term Care Facilities. The sample consists of 279 nursing homes chosen from 798 possible cases. Excluded were private not-for-profit homes, government homes, hospital attached homes, and non-reporting homes. Unless otherwise specified, the variables are daily averages, with the unit of observation being the nursing home. Descriptive statistics are presented in Table 2.

We specify a three input translog cost function where the inputs are skilled labor, unskilled labor, and supplies. The skilled wage is taken to be the average wage of a registered nurse in the home's market area and the wage of unskilled labor is taken to be the average wage of an aide in the home's market area. A home's market area is defined later. The price of supplies is assumed to be the same across homes. Supplies are measured in terms of expenditures so that its price is one for all homes. We assume that all nursing homes face the same cost of capital. Capital expenditures are small in nursing homes since the majority of capital is the building. The size of buildings is roughly proportional to the total number of patients within the nursing home. In our data, there is little variation in square feet per bed across homes. Therefore, if the price of capital is not the same across New York State then our estimation results can be interpreted as a short run cost function, where capital is proportional to capacity.

The exogenous demand variables are the per capita income of the people living in the nursing home's market area, the population over age 65 in the nursing home's market area, a casemix index of health status, an index of market concentration, and the Medicaid reimbursement rate. The income and population data are from the 1980 census. The concentration index is a measure which is negatively related to the competitiveness of the market.

Defining a home's market requires some work. Since homes do not compete for Medicaid patients, the appropriate market to analyze is the private-pay patient market. The common assumption is that a home's geographic market is the county in which the home is located (e.g. Nyman, 1985), but patient origin data indicate that most homes care for a substantial number of patients whose last residence was not the county in which the home is located. Instead, separate market areas are defined for each home based on patient origin data. Homes are assumed to participate in several county markets. A home's participation in a county market is given by the proportion of the home's private-pay patients from that county. Thus, a home's market area is defined as the counties in which its private-pay patients last resided, and the proportion of its private-pay patients from each county.

This market definition guides the construction of the demand variables. Each home's market population is computed as a weighted sum of the number of persons over age 65 in each county, using the home's proportion of private-pay patients from the counties as weights. Similarly, the per capita income of the population in a home's market area is computed as the weighted sum of the county's per capita incomes.

The concentration level of a home's private-pay patient market is computed as a weighted sum of the county market concentration levels. The notion is that counties comprise separate markets and nursing homes compete for private-pay patients in several counties. The competitiveness of a home's market is a weighted average of the competitiveness of the county markets. The decomposition into county markets is artificial, but is necessary since the data come aggregated at the county level.

The concentration of each county market is computed using the Herfindahl-Hirschmann index, which is the sum of squares of each home's share (proportion) of a county's private pay patients. Let y_{ij} be the number of patients from county j in home i. If s_j is equal to the total number of patients in county j, then the concentration of county market j is:

$$M_j = \sum_{i=1}^{n} (y_{ij}/s_j)^2$$

The concentration of a home's market is the weighted average of the M_j's. Specifically, the concentration index for home i is:

$$H_i = \sum_{j=1}^{J} s_{ij}^2 \, M_j$$

where s_{ij} is the share of home i's patients from county j.

The casemix index of health status is based on the Katz Activities of Daily Living (ADL) index. The ADL index is a measure of a patient's functional level. Katz (1963) developed the index to explicitly measure function levels among the chronically ill and aging population, and it has proven a valid and reliable measure. The ADL index is computed from disability scores assigned to patients in eight functional areas. For each home, patients' ADL scores are summed and divided by the number of patients in the home. The result is an index of the average ill-health of the patients in a facility.

In 1980, New York reimbursed nursing homes using a cost-plus method. New York computed a home's 'plus' factor based on an owner's equity, debt structure, the size of the facility, and the value of assets. Care was taken to ensure that homes could not manipulate this formula by constantly reselling the home so as to increase the value of its assets. Also, the size of the home and assets were controlled by regulation. Thus, some of the factors in the plus formula were exogenous to the home. Alternatively, homes could to some extent manipulate their equity and debt structure so as to maximize their plus factor net of taxes, but the equity and debt structure decisions are independent of variable input and patient-mix choices. The plus factor is used as the Medicaid reimbursement rate.

References

Aigner, D., Hsiao, C., Kapteykn, A., and Wansbeek, T. 1984. "Latent Variable Models in Econometrics," in *Handbook of Econometrics* 2, eds. Z. Griliches and M. Intrilligator, Amsterdam: North-Holland, 1321-1393.

Allen, R. and P. Gertler. 1990. "Regulation and the Provision of Quality to Heterogeneous Consumers," *Journal of Regulatory Economics*, forthcoming.

Braeutigam, R., A. Daughety, and M. Turnquist. 1982. "The Estimation of a Hybrid Cost Function for a Railroad Firm." *Review of Economics and Statistics*, 64, 394-404.

Braeutigam, R., A. Daughety, and M. Turnquist. 1984. "A Firm-Specific Analysis of Economies of Density in the U.S. Railroad Industry," *Journal of Industrial Economics*, 33, 3 - 20.

Braeutigam, R. and M. Pauly. 1986. "Cost Function Estimation and Quality Bias: The Regulated Automobile Insurance Industry," *Rand Journal of Economics*, 17, 606-617.

Brown, J. and H. Rosen. 1982. "On the Estimation of Structural Hedonic Price Models," *Econometrica*, 50, 765-768.

Chamberlain, Edward H. 1933. *The Theory of Monopolistic Competition*. Cambridge, Mass.: Harvard University Press.

Chamberlain, G. 1982. "Multivariate Models for Panel Data," *Journal of Econometrics*, 18, 5-46.

Christensen, L., D. Cummings, and P. Schoech. 1979. "Econometric Estimation of Scale Economies in Telecommunications," University of Wisconsin-Madison, Social Systems Research Institute Paper 8013.

Christensen, L and W. Greene. 1976. "Economics of Scale in U.S. Electricity Power Generation," *Journal of Political Economy*, 84, 655-674

Deaton, A. 1988. "Quantity, Quality and Spatial Price Variation," *American Economic Review*, 78, 418-430.

Dor, Avi. 1989. "The Costs of Medicare Patients in Nursing Homes in the United States: A Multiple Output Analysis," *Journal of Health Economics*, 8, 253-270.

Douglas, G. and J. Miller. 1974. *Economic Regulation of Domestic Air Transport*, Washington, D.C.: Brookings Institution.

Dusansky, R. 1989. "On the Economics of Institutional Care of the Elderly in the U.S.: The Effects of a Change in Government Reimbursement," *Review of Economic Studies*, 56, 141-150.

Friedlander, A. and R. Spady. 1980. *Freight Transport Regulation: Equity, Efficiency and Competition in the Rail and Trucking Industries*, Cambridge: MIT Press.

Fuss, M. and Waverman, L. 1981. "Regulation and the Multiproduct Firm: The Case of Telecommunications in Canada," in G. From, ed. *Studies in Public Regulation*, Cambridge: MIT Press.

Gertler, P. J. 1988. "A Latent Variable Model of Quality Determination," *Journal of Business and Economic Statistics*, 6, 97-104.

Gertler, P. J. 1989. "Subsidies, Quality, and the Regulation of Nursing Homes," *Journal of Public Economics*, 39, 33-53.

Harrington, C. and J. Swan. 1984. "Medicaid Nursing Home Reimbursement Policies Rates and Expenditures," *Health Care Financing Review*, 7, 1-38.

Joreskog, K., and Goldberger, A. 1975, "Structural Equations in the Social Sciences," *Econometrica*, 40, 979-1001.

Joskow, P. 1980. "The Effects of Competition and Regulation on Hospital Bed Supply and the Reservation Quality of the Hospital," *The Bell Journal of Economics*, 11.

Katz, S., 1963, "Studies of Illness in the Aged: The Index of ADL," *Journal of the American Medical Association*, 195, 914-919.

Lancaster, K. 1976. "A New Approach to Consumer Theory," *Journal of Political Economy*, 74, 132-157.

Leffler, K. 1982. "Ambiguous Changes in Product Quality," *American Economic Review*, 72,956-967.

McElroy, M. 1986. "Generalized Error Models for Cost Function," *Journal of Political Economy*, 94.

McKay, N., 1988. "Costs and Scale Economies in the Nursing Home Industry," *Journal of Human Resources*, 23, 57-75.

New York State Department of Health and Human Services, Office of Health Systems Management, 1980, Title XIX, Nursing Home Cost Survey.

Nyman, J., 1985. "Prospective and 'Cost Plus' Medicaid Reimbursement, Excess Medicaid Demand, and the Quality of Nursing Home Care," *Journal of Health Economics*, 4, 237-259.

Nyman, J. 1989. "The Marginal Cost of Nursing Home Care," *Journal of Health Economics*, 8, 193-208.

Panzar, J. 1975. "Regulation, Service Quality, and Market Performance: A Model of Airline Rivalry." Center for Research in Economic Growth, Stanford University, Memorandum 184.

Posner, R. 1971. "Taxation of Regulation," *Bell Journal of Economics and Management Science*, 2, 22-50.

Rosen, Sherwin. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 82, 34-55.

Rosse, J. and J. Panzar. 1974. "Models of Regulated Monopoly with Service Quality and Averch-Johnson Effects: Preempirical Comparative Statics." Center for Research in Economic Growth, Stanford University, Memorandum 176.

Scanlon, W., 1980. "The Market for Nursing Home Care: A Case of an Equilibrium with Excess Demand as a Result of Public Policy, unpublished Ph.D. dissertation (University of Wisconsin-Madison).

Schmalensee, R. 1977. "Comparative Static Properties of Regulated Airline Oligopolies," *Bell Journal of Economics*, 8, 565-576.

Spence, A.M. 1975. "Monopoly, Quality, and Regulation," *Bell Journal of Economics*, 6, 417-429.

Stigler, G. 1971. "The Theory of Economic Regulation," *Bell Journal of Economics and Management Science*, 2, 3-21.

Vogel, R. and H. Palmer, eds. 1983. *Long Term Care: Perspectives from Research and Demonstrations.* HCFA, HHS, Washington, D.C.

White, L. 1972. "Quality Variation When Prices are Regulated," *Bell Journal of Economics and Management Science*, 3, 425-436.

</ref_section>

Table 1
Reduced Form Parameters as Functions of
Structural Parameters

Moment	a. Cost Function
1	$\alpha_0 + \alpha_y \gamma_0$
log W_i	$\alpha_y(\gamma_{wi} + \gamma_q \lambda_{wi}) + \alpha_{wi} + \alpha_q \lambda_{wi} + \beta_{yy}\gamma_0(\gamma_{wi} + \gamma_q \lambda_{wi})$
log Z_i	$(\alpha_y + \beta_{yy})c_i + \alpha_q \lambda_{zi} + \beta_{yq}\lambda_{zi}\gamma_0$
log W_i·log W_j	$(\beta_{ij} + 2\beta_{qi}\lambda_{wj}) + b_i(\gamma_{wi} + \gamma_q \lambda_{wj}) + \beta_{yy}(\gamma_{wi} + \gamma_q \lambda_{wi})(\gamma_{wj} + \gamma_q \lambda_{wj})$
log W_i·log Z_j	$\beta_{qi}\lambda_{zj} + b_i c_i + \beta_{yq}\lambda_{zj}(\gamma_{wi} + \gamma_q \lambda_{wi}) + \beta_{yy} c_i(\gamma_{wi} + \gamma_q \lambda_{wi})$
log Z_i·log Z_j	$\beta_{yy} c_i c_j + \beta_{yq}\lambda_{zi} c_j$

	b. i-th Share Equation
1	$\alpha_{wi} + \beta_{yi}\gamma_0$
log W_j	$\beta_{ij} + \beta_{qi}\lambda_{wj} + \beta_{yi}(\gamma_{wi} + \gamma_q \lambda_{wi})$
log Z_j	$\beta_{qi}\lambda_{zj} + \beta_{yi} c_i$

	c. Conditional Product Demand Equation
1	γ_0
log W_j	$\gamma_{wi} + \gamma_q \lambda_{wi}$
log Z_j	$\gamma_{zi} + \gamma_q \lambda_{zi}$

Definitions: $b_i = \beta_{yi} + \beta_{yq}\lambda_{wi}$; $c_i = \gamma_{zi} + (\gamma_{wi} + \gamma_q \lambda_{wi})$.

Table 2

MEANS AND STANDARD DEVIATIONS
(N = 279)

Variable	Mean	Standard Deviation
Total Cost ($/day)	5319.74	4287.0
Output (total patients)	121.227	74.230
Cost Shares		
Skilled labor	0.206	0.045
Unskilled labor	0.537	0.053
Supplies expenditures	0.257	0.065
Wages		
Skilled wage ($/hour)	9.752	3.384
Unskilled wage ($/hour)	7.118	2.451
Demand shifters		
Casemix index	0.407	0.154
Medicaid plus factor	6.839	2.714
Population 65+ (10000's)	0.964	0.863
Per capita income ($1000's)	0.706	0.142
Herfindahl index	0.124	0.118
Population*Herfindahl	0.744	0.701

Table 3

PARAMETER ESTIMATES
(ASYMPTOTIC T-STATISTICS)

	Cost Function		Quality Function	
	Quality Adjusted	Quality Exogenous		
ALPHA0	2.466	0.726	LAMBDAW1	-0.014
	(12.69)	(1.75)		(3.87)
ALPHAW1	0.105	0.113	LAMBDAW2	-0.022
	(11.35)	(7.95)		(6.15)
ALPHAW2	0.455	0.448		
	(43.30)	(31.45)		
ALPHAW3	0.438	0.439		
	(2.38)	(1.38)	CASEMIX	0.160
ALPHAY	0.993	1.769	(LAMBDAZ1)	(19.49)
	(11.56)	(9.66)		
ALPHQ	0.221	---	MEDPLUS	-0.418
	(0.22)		(LAMBDAZ1)	(4.44)
BETAW1W1	0.101	0.143		
	(13.38)	(13.59)	POPULATION	0.045
BETAW1W2	-0.064	-0.116	(LAMBDAZ1)	(1.74)
	(-7.51)	(11.81)		
BETAW2W2	0.153	0.162	INCOME	0.069
	(14.62)	(14.23)	(LAMBDAZ1)	(1.37)
BETAW1W3	-0.036	-0.027		
	(11.80)	(0.41)	HERF	-2.601
BETAW2W3	-0.090	-0.046	(LAMBDAZ1)	(4.97)
	(15.87)	(0.48)		
BETAW3W3	0.125	-0.074	POP*INC	-0.008
	(6.42)	(0.73)	(LAMBDAZ1)	(2.28)
BETAYY	0.004	-0.088		
	(0.43)	(4.43)		
BETAYQ	0.979	---		
	(4.75)			
BETAY1	0.004	0.009		
	(2.31)	(0.03)		
BETAY2	-0.012	0.993		
	(4.47)	(2.73)		
BETAY3	0.008	-0.010		
	(3.14)	(0.14)		
BETAQ1	1.0*	---		
	(---)			
BETAQ2	0.384	---		
	(4.86)			
BETAQ3	-1.384	---		
	(18.52)			

* Normalized to 1.

Figure 1

Average Cost Functions

Quality-Adjusted/Quality-Exogenous