

# Additive General Error Models for Production, Cost, and Derived Demand or Share Systems

---

Marjorie B. McElroy

*Duke University*

Many empirical studies of production specify a deterministic model of the firm, derive the implied behavioral equations (input demand or share system), and then “embed this system in a stochastic framework” by tacking on linear error terms. In contrast, this paper proposes general error models (GEMs) in which the error specification is an integral part of the optimization model. These models are the statistical embodiment of Stigler’s view that apparent observed inefficiencies reflect the investigator’s ignorance of the true optimization problem. Additive GEMs are proposed and interpreted. Specification tests indicate that a translog additive GEM is superior to the standard translog specification.

## I. Introduction and Summary

Error specifications play a fundamental role in the estimation of producers’ demand, cost, and production systems. Many researchers have given careful thought to the specification of error terms for production, cost, and profit functions and for derived input demand and share systems. These include Theil (1965, 1974), Zellner, Kmenta, and Drèze (1966), Zellner and Revankar (1969), Fuss, McFadden, and Mundlak (1978), Simmons and Weiserbs (1979), Woodland (1979), Duncan (1980), and Rossi (1983), as well as the stochastic frontier literature authoritatively surveyed by Førsund,

Earlier versions of this paper were presented at the Fifth World Congress of the Econometric Society, Cambridge, Mass., August 1985, and at the Winter Meeting of the Econometric Society, New York, December 1985. I would like to thank Ariel Pakes, John Strauss, and John Whitaker for helpful comments and Lisa Ouye and Fallaw Sowell for research assistance. This work was supported by NSF grant SEC-8409262.

[*Journal of Political Economy*, 1987, vol. 95, no. 4]

© 1987 by The University of Chicago. All rights reserved. 0022-3808/87/9504-0005\$01.50

Lovell, and Schmidt (1980). Apart from share errors summing to zero, most of this work proceeds as if the static neoclassical model of a profit-maximizing firm has no implications for the specification of the stochastic parts of the implied behavioral equations (cost function, input demand, and share functions), and vice versa. To date, the sources of error have been seen as outside of the theory (e.g., random environment or errors in measurement), violations of the theory (errors in optimization, both technical and allocative), or lack of generality of the theory (decision making under risk).

In contrast, this paper uses the bare-bones, static, profit-maximizing model of the firm to derive a rich set of implications for the stochastic specification of behavioral equations. It is neither necessary nor desirable to follow a common procedure in which the first step is to derive the deterministic cost and derived demand or share functions and the second step is to “embed” these equations in a stochastic framework by tacking error terms onto these functions. Many issues treated here are at least touched on by Fuss et al. (1978).<sup>1</sup> They explicitly recognize that the “stochastic specification is an intrinsic part of the specification of the production model” (p. 249) and urge that this specification be guided by “visualization of the *true* process” as determined by nature and not by econometric convenience.

I call such models that treat the stochastic specification of the firm’s behavioral equations (cost, derived input demand, and share equations) as an integral part of the underlying model of the firm “general error models” (GEMs). A GEM specification of the firm’s (primal or dual) optimization problem includes the error structure. Consequently, the estimated GEM behavioral equations (e.g., cost and derived demand equations) inherit their stochastic specification from the underlying maximization problem. Errors in GEMs are transmitted between the primal and the dual or the demand and share systems via standard duality relationships, thereby endowing all these representations of technology and behavior with mutually consistent error specifications. The specification problems automatically solved by this mutual consistency of GEM errors include the following: (i) Standard comparisons of the functional form of a demand system with that of a share system typically tack linear error terms onto each system, resulting in inconsistent error specifications; in contrast, the use of a GEM

<sup>1</sup> For this reference I am indebted to an anonymous referee. Fuss et al. specify an exemplary Cobb-Douglas general error model with errors stemming from a firm effect, a random effect (unknown at the time of the production decision), and errors in observation (on the part of both the firm and the observer) and discuss the consequences for the properties of various estimators.

would guarantee a consistent error specification. (ii) Tests such as Appelbaum's (1978) have resulted in questioning the validity and internal consistency of neoclassical production theory. Such results can arise solely on account of inconsistent error specifications between the primal and the dual. A GEM would automatically provide a consistent error specification and preclude conflict between the primal and the dual. (iii) Unlike conventional error specifications, for a GEM any given full-information estimator (such as full-information maximum likelihood [FIML]) is identical for all representations of that GEM. (iv) With some exceptions such as Fuss et al. (1978), Berndt and Khaled (1979), and the stochastic frontier literature, previous research has assumed that there is no functional relationship between input demand (or share) errors and cost function error terms. In contrast, GEMs reveal exact functional relationships between demand (or share) errors and the error term in the cost function. (v) General error models also reveal the precise way in which, relative to a stochastic demand system, the stochastic cost function is redundant; but, relative to a stochastic share system, the stochastic cost function is required for full-information estimation. (vi) Previous research either ignores the fact that estimated shares must lie on the unit simplex or implements this restriction by resorting to special distributions such as Woodland's (1979) use of the Dirichlet distribution. Since they are derived theoretically, GEM shares automatically lie on the unit simplex. Special distributional assumptions or estimation procedures are not required to restrict predicted GEM shares to the unit simplex.

One way to motivate GEMs is to start with the core static model of a profit-maximizing firm with perfect information in a perfectly competitive and riskless world and to assume that firms differ from each other according to parameters that are known by the decision maker but not by the outside observer. To the observer these firm effects manifest themselves as random parameters in the production function and, in the additive GEM of this paper, as additive error terms in the firm's dual cost and input demand (or share) functions. This random parameters interpretation is, however, inessential. For additive GEMs, permissible interpretations of the errors include measurement error in the factor inputs as well as a stochastic frontier with "technical errors." Additional information would be required to distinguish among these interpretations.

Stigler (1976) argued that observed inefficiencies in behavior are only apparent inefficiencies: that if one only knew the complete and correct criterion function of the maximizer (inclusive of the costs associated with contract enforcement, the costs of information pertinent to the choice of technology, and so forth), then these apparent

inefficiencies would disappear. General error models may be viewed as a mathematical embodiment of Stigler's view that apparent inefficiencies arise because we are ignorant of the true optimizing problem: inefficiency is in the eye of the beholder. Hence GEMs are appropriate when data or other resource constraints do not permit the researcher to model these unknown costs and objectives.

Motivated both by tractability and by comparability with widely used error specifications, Section II introduces the neoclassical "additive general error model" (additive GEM or AGEM). Neoclassical additive GEMs are defined as static neoclassical models that give rise to cost, demand, and share equations with additive errors. Two isomorphic representations of the empirical implications of the neoclassical additive GEM are given. These are (in Sec. IIA) an  $n$ -input demand system and (in Sec. IIB) a cost-cum-share system. The isomorphism between these two representations is inclusive of the error structure. Section IIC shows that the traditional share model with additive errors is inconsistent with the additive GEM. For empirical work, the isomorphism between the cost-cum-share and demand systems guarantees that the entire menu of extant deterministic input demand and share systems can be incorporated into additive GEMs. This menu includes all the flexible functional forms (e.g., the translog) as well as the exact functional forms (e.g., the constant elasticity of substitution [CES] production function). Section IID gives alternative interpretations of additive GEMs while Section IIE specifies a translog additive GEM.

In Section III nonlinear (iterated) seemingly unrelated regression (NLITSUR) techniques are shown to yield consistent (maximum likelihood) estimators for additive GEMs. Also, an encompassing strategy for specification tests is shown to yield Breusch-Pagan heteroscedasticity tests for the AGEM model versus the traditional share model with additive errors. In Section IV, NLITSUR methods are used to estimate the translog additive GEM using the Berndt and Wood (1975) data on U.S. manufacturing. These results are compared with those for the traditional share model with additive errors. Generalized Breusch-Pagan specification tests indicate that this translog additive GEM specification is superior to the traditional specification of translog shares with additive errors. The final section discusses the use of GEMs as an empirical research strategy as well as extensions of this work to other (nonadditive) functional forms and to GEMs for consumption theory.<sup>2</sup>

<sup>2</sup> I am currently working on other functional forms for GEMs in production as well as GEMs for consumer demand systems (see McElroy 1981, 1986).

## II. General Error Models with Additive Errors in Production, Cost, Input Demand, and Share Systems

This section presents the additive general error model of cost, input demand, and input share systems. For maximum comparability with the empirical literature, this additive GEM is for a cost-minimizing firm with an exogenously determined level of output. Although the essential insights are available from this AGEM, it is important to realize that the AGEM specification applies to a wider class of models. By replacing the duality results for production, cost, and input demands (Shephard's lemma) by duality results for production, profit, and input demand and output supply (Hotelling's lemma), the AGEM is readily extended to the case in which the firm chooses its output level(s).

Consider the deterministic model of a profit-maximizing firm that purchases inputs  $\mathbf{x} = (x_1, \dots, x_n)$  at prices  $\mathbf{w} = (w_1, \dots, w_n)$  to produce output  $y$  according to the additive GEM production function

$$\begin{aligned} y &= q(x_1 - \epsilon_1, \dots, x_n - \epsilon_n; \theta) \\ &\equiv q(\mathbf{x} - \boldsymbol{\epsilon}; \theta). \end{aligned} \quad (1)$$

Here  $q(\cdot)$  gives the maximum level of output for given inputs  $\mathbf{x}$  with *known* parameters  $(\theta, \boldsymbol{\epsilon})$ , where  $\boldsymbol{\epsilon} \equiv (\epsilon_1, \dots, \epsilon_n)'$ . The function  $q(\cdot)$  is assumed to be twice continuously differentiable, concave, and non-decreasing in its arguments.<sup>3</sup> The parameters  $\theta$  are the same for all firms and, henceforth, are suppressed.

To motivate the model the parameter vector  $\boldsymbol{\epsilon}$  is assumed known to the firm but unknown to the researcher and to vary across firms according to a distribution that is free of  $y$  and  $\mathbf{w}$ , has a zero mean vector, and an  $n \times n$  positive definite covariance matrix,  $\boldsymbol{\Sigma}$ . Further, in the neighborhood of any observed input vector  $\mathbf{x}$  the distribution of  $\boldsymbol{\epsilon}$  is assumed to be such that any realized  $\epsilon_i$  is almost surely sufficiently small that  $x_i - \epsilon_i > 0$  for all  $i$ .<sup>4</sup>

Define the deterministic cost function as  $c(y, \mathbf{w})$ , which gives the minimum cost of producing output level  $y$  under production function (1) given input prices  $\mathbf{w}$  when the firm-specific parameters are zero, that is, when  $\boldsymbol{\epsilon} = \mathbf{0}_n$ . Also define the firm-specific cost function,  $C(y, \mathbf{w}, \boldsymbol{\epsilon})$ , as the minimum cost of producing  $y$  under prices  $\mathbf{w}$  and param-

<sup>3</sup> It is also required that  $q(\mathbf{x} - \mathbf{0}, \theta) > 0$  for all  $x \gg 0$ , i.e., that  $q(\cdot)$  evaluated at strictly positive  $\mathbf{x}$  and  $\boldsymbol{\epsilon} = \mathbf{0}$  be positive. Strict concavity can be weakened. For a summary of the menu of alternative regularity conditions of production functions, see Diewert (1974), Lau (1974), and Shephard (1974).

<sup>4</sup> In effect the AGEM production function (1) is assumed to hold in the neighborhood of the observed input and output levels  $\mathbf{x}$  and  $y$ .

eters  $\epsilon$ , where in general  $\epsilon \neq \mathbf{0}_n$ . It is straightforward to show that a firm has the AGEM production function (1) iff its firm-specific cost function,  $C(y, \mathbf{w}, \epsilon)$ , is given by the AGEM cost function

$$C = C(y, \mathbf{w}, \epsilon) = c(y, \mathbf{w}) + \sum_{j=1}^n w_j \epsilon_j, \quad (2)$$

which is the sum of the deterministic cost function  $c(y, \mathbf{w})$  and a price-weighted sum of the  $\epsilon_j$ 's. Ceteris paribus, the larger a firm's parameter  $\epsilon_i$ , the larger its cost of producing output  $y$  for given prices  $\mathbf{w}$ . Note that the "deterministic" cost function is also the mean cost function, or  $E[C(y, \mathbf{w}, \epsilon)] = c(y, \mathbf{w})$ . Both  $C(\cdot)$  and  $c(\cdot)$  are concave and linear homogeneous in  $\mathbf{w}$ , and dual to  $q(\mathbf{x} - \epsilon)$  and  $q(\mathbf{x})$ , respectively.

The remainder of this section establishes two isomorphic representations of the empirical implications of the AGEM cost-minimizing model of the firm: an  $n$ -input demand system (subsection A) and a cost-cum-share system consisting of  $(n - 1)$  share functions and the cost function (subsection B).

#### A. Representation of Empirical Implications of the AGEM as a Derived Input Demand System

Shephard's lemma yields the conditional firm-specific input demand functions

$$x_i = C_i(y, \mathbf{w}, \epsilon) = c_i(y, \mathbf{w}) + \epsilon_i, \quad i = 1, \dots, n, \quad (3)$$

where  $C_i(\cdot)$  and  $c_i(\cdot)$  are homogeneous of degree zero in  $\mathbf{w}$ . Further,  $E[C_i(y, \mathbf{w}, \epsilon)] = c_i(y, \mathbf{w})$  is the deterministic or average demand function of the population of firms. If  $\epsilon_i > 0$  ( $\epsilon_i < 0$ ), the firm uses  $\epsilon_i$  more (less) of  $x_i$  to produce  $y$  than does the average firm; this in turn costs the firm  $w_i \epsilon_i$  more (less) than the average firm; see equation (2).<sup>5</sup> These demand functions satisfy the first-order conditions

$$\frac{q_i(\mathbf{x} - \epsilon)}{q_j(\mathbf{x} - \epsilon)} = \frac{w_i}{w_j}, \quad i, j = 1, \dots, n, i \neq j, \quad (4)$$

where  $q_i$  is the  $i$ th partial derivative of the production function (1).

Weighting (3) by  $w_i$ , summing over the  $n$  inputs, and using homogeneity retrieves the cost function

$$\sum w_i x_i = \sum w_i c_i(y, \mathbf{w}) + \sum w_i \epsilon_i = c(y, \mathbf{w}) + \sum w_i \epsilon_i \equiv C(y, \mathbf{w}, \epsilon). \quad (5)$$

<sup>5</sup> Fuss et al. (1978) also exhibit a precise relationship among cost, demand, and share errors. Berndt and Khaled (1979) obtained a similar relationship between the errors in their input-output equations and the unit cost function (see their eq. [29] in particular). Since they assumed that errors are errors in cost-minimizing behavior, they speculated that sufficient conditions to ensure that the unit cost function error is positive would be "very messy" or "intractable." For GEMs this is not a problem.

Consequently, the input demands (3) identically satisfy adding up for every  $y$  and  $w$  and realization of the vector  $\epsilon$ . The  $n$ -equation demand system embodies all the empirical implications of the model; the cost function is redundant.

*B. Representation of Empirical Implications of the AGEM as a Cost-cum-Share System*

A cost-minimizing firm produces a given level of output subject to the AGEM production function (1) iff it has the AGEM cost function (2) iff it has demand system (5). Further, (1), (2), or (5) implies that the firm has the AGEM share demand system<sup>6</sup>

$$S_i = s_i(y, \mathbf{w}) + v_i, \quad i = 1, \dots, n, \tag{6}$$

where the  $i$ th observed share is defined as  $S_i \equiv w_i x_i / \sum_j w_j x_j$  (implying that  $\sum S_i \equiv 1$ ), where the *deterministic* part of the  $i$ th share function is given by

$$s_i(y, \mathbf{w}) \equiv \frac{\partial \ln c(y, \mathbf{w})}{\partial \ln w_i} \equiv \frac{w_i c_i(y, \mathbf{w})}{c(y, \mathbf{w})}, \quad i = 1, \dots, n, \tag{7}$$

and the *stochastic* part is given by

$$v_i = v_i(y, \mathbf{w}, \epsilon) = \frac{1}{c(y, \mathbf{w}) + \sum w_j \epsilon_j} [w_i \epsilon_i - s_i(y, \mathbf{w}) \sum w_j \epsilon_j], \tag{8}$$

$$i = 1, \dots, n.$$

Linear homogeneity of the  $c(\cdot)$  in  $\mathbf{w}$  implies that

$$\sum w_i c_i(y, \mathbf{w}) = c(y, \mathbf{w}) \tag{9}$$

and hence that the deterministic shares sum to one ( $\sum_j s_j(y, \mathbf{w}) = 1$ ), that the errors sum to zero ( $\sum_j v_j = 0$ ), and that both sides of the share equations (6) sum identically to one.

Because of this linear dependence, unlike the demand system (5), one cannot recover the cost function from the share system (6). The missing information on the level of production can be added, however, in the form of the cost equation.<sup>7</sup> Then any arbitrarily chosen

<sup>6</sup> A constructive derivation of (6) is to multiply (3) by  $w_i$  and divide by total cost,  $C(y, \mathbf{w}, \epsilon) \equiv \sum w_j x_j$ , to get

$$\frac{w_i x_i}{\sum w_j x_j} = w_i \cdot \frac{c_i(y, \mathbf{w})}{c(y, \mathbf{w}) + \sum w_j \epsilon_j} + \frac{w_i \epsilon_i}{c(y, \mathbf{w}) + \sum w_j \epsilon_j}, \quad i = 1, \dots, n,$$

multiply the first term on the right-hand side by  $1 \equiv (C - \sum w_j \epsilon_j)/c$ , and then collect deterministic and stochastic terms. Alternatively, one could partially differentiate the natural logarithm of the cost function (2).

<sup>7</sup> Alternatively, one could add (albeit asymmetrically) an arbitrarily selected single demand function to determine the level. For the problem of maximizing output for a given fixed cost, one would add the production function instead of the cost function.

$n - 1$  share equations from (6) plus the cost function (2) embody all the empirical implications of cost-minimizing behavior. This  $n$ -equation cost-cum-share system is isomorphic to the  $n$ -equation demand system. Without loss of generality we will take the redundant share equation to be the  $n$ th and write the cost-cum-share system as

$$\begin{aligned} C &= c(y, \mathbf{w}) + v_c, \\ S_i &= s_i(y, \mathbf{w}) + v_i, \quad i = 1, \dots, n - 1, \end{aligned} \tag{10}$$

where the  $v_i$ 's are defined in (8) and  $v_c$  is defined as  $\sum_{j=1}^n w_j \epsilon_j$ .

*C. A Restatement of the Isomorphism and the Implied Restrictions on the Error Structure*

The building blocks of the demand system (3) are the  $n$  deterministic demand functions  $d(y, \mathbf{w}) = (c_1(y, \mathbf{w}), \dots, c_n(y, \mathbf{w}))$  plus the  $n$  jointly distributed random errors,  $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_n)$ . It is obvious that any given demand system (3)—one in which the deterministic demands  $d(y, \mathbf{w})$  are specified and the joint distribution of  $\boldsymbol{\epsilon}$  is given—can be written as a cost-cum-share system in which the share functions are

$$\begin{aligned} S_i &= s_i^*(d(y, \mathbf{w}), \mathbf{w}; \boldsymbol{\epsilon}) \\ &= \frac{w_i c_i(y, \mathbf{w})}{\sum_j w_j c_j(y, \mathbf{w})} + \frac{1}{\sum_j w_j c_j(y, \mathbf{w}) + \sum_j w_j \epsilon_j} \\ &\quad \times \left[ w_i \epsilon_i - \frac{w_i c_i(y, \mathbf{w})}{\sum_j w_j c_j(y, \mathbf{w})} \sum_j w_j \epsilon_j \right], \quad i = 1, \dots, n - 1. \end{aligned} \tag{11}$$

In this way one could use the deterministic Diewert-Leontief demand functions (Diewert 1974, p. 113) to specify a cost-cum-share system. Note that if the jointly distributed  $\boldsymbol{\epsilon}$  are independently and identically distributed (i.i.d.) across observations, then the share errors in (11) are precluded from being i.i.d.: the error term (the last term in [11]) varies across observations with  $y$  and  $\mathbf{w}$ .

Provided a regularity condition on the cost-cum-share errors is satisfied, the converse of the previous paragraph also holds. The building blocks of the cost-cum-share system (10) are the deterministic cost function and  $n - 1$  deterministic share functions,  $\bar{s}(y, \mathbf{w}) = (c(y, \mathbf{w}), s_1(y, \mathbf{w}), \dots, s_{n-1}(y, \mathbf{w}))$ , plus the jointly distributed cost error and  $n - 1$  share errors  $\tilde{\mathbf{v}} = (v_c, v_1, \dots, v_{n-1})$ . And, provided the regularity condition on  $\tilde{\mathbf{v}}$  is met, any given cost-cum-share system (10)—one in which the deterministic cost and shares  $s(y, \mathbf{w})$  are speci-



fied and the joint distribution of  $\tilde{\mathbf{v}}$  is given—can be written as an AGEM demand system:

$$x_i = x_i^*(\bar{s}(y, \mathbf{w}), \mathbf{w}; \tilde{\mathbf{v}}) = \frac{1}{w_i} s_i(y, \mathbf{w})c(y, \mathbf{w}) + \frac{1}{w_i} \{[c(y, \mathbf{w}) + v_c]v_i + s_i(y, \mathbf{w})v_c\}, \quad i = 1, \dots, n. \tag{12}$$

In subsection *E* below, (12) is used to specify a translog AGEM demand system. In (12),  $\tilde{\mathbf{v}}$  must meet the regularity condition that

$$e_i(\bar{s}(y, \mathbf{w}), \mathbf{w}; \tilde{\mathbf{v}}) = \frac{1}{w_i} \{[c(y, \mathbf{w}) + v_c]v_i + s_i(y, \mathbf{w})v_c\} \tag{13}$$

is free of  $y$  and  $\mathbf{w}$ . Note that when  $\bar{s}(y, \mathbf{w})$  and  $\tilde{\mathbf{v}}$  are regarded as a function of  $d(y, \mathbf{w}), \mathbf{w}$ , and  $\boldsymbol{\epsilon}$ , then

$$e_i(\bar{s}(y, \mathbf{w}), \mathbf{w}; \tilde{\mathbf{v}}) \equiv \boldsymbol{\epsilon}_i, \quad i = 1, \dots, n. \tag{14}$$

The regularity condition (13) rules out the possibility that the share errors  $\tilde{\mathbf{v}}$  are i.i.d. across observations.<sup>8</sup>

*D. Interpretations of the Errors in Additive GEMs*

To reiterate, the interpretation of additive GEM errors as random parameters is straightforward. Each firm minimizes cost subject to its known firm-specific production function (1) with parameters  $(\boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_n)$  that differ across firms according to a joint distribution that is independent of input prices and the output level. These parameters are known to the firm but unknown to the econometrician. They appear in every primal, or dual, representation of production.

Additional interpretations of additive GEMs are available. The  $\boldsymbol{\epsilon}$ 's may be regarded as classic errors in measurement of the inputs,

<sup>8</sup> Suppose that the  $\tilde{\mathbf{v}}$  are i.i.d. across observations. Then the obvious candidate for the underlying production function is  $y = q^*(\mathbf{x}, \mathbf{w}, \boldsymbol{\epsilon})$ , which is the (perhaps only locally defined) explicit function found by solving for  $y$  from  $y = q(\mathbf{x} - \boldsymbol{\epsilon}(y, \mathbf{w}, \tilde{\mathbf{v}}))$ , where the  $i$ th element of  $\boldsymbol{\epsilon}(y, \mathbf{w}, \tilde{\mathbf{v}})$  is the error term from (13). At best,  $q^*(\cdot)$  has the right curvature and gives the maximum  $y$  for given  $\mathbf{x}, \mathbf{w}$ , and  $\boldsymbol{\epsilon}$ . But  $q^*(\cdot)$  is price dependent and is a different function from  $q(\cdot)$ . Thus  $q^*(\cdot)$  is certainly not the underlying production function that is dual to the cost function (2) and the shares (6). In other work, I am investigating nonadditive GEMs in which i.i.d. share errors are permitted, e.g., the log-additive GEM, with cost-cum-share system

$$\ln C = \theta(y, \mathbf{w}) + \sum \eta_i \ln w_i + \eta_0, \\ S_i = \frac{\partial \theta(y, \mathbf{w})}{\partial \ln w_i} + \eta_i, \quad i = 1, \dots, n, \quad \sum_{i=1}^n \eta_i = 0.$$

Here  $\exp[\theta(y, \mathbf{w})]$  is the deterministic cost function and  $\eta_1, \dots, \eta_{n-1}, \eta_0$  are the GEM errors with  $\eta_1, \dots, \eta_{n-1}, \eta_0$  i.i.d. across observations.

where each measured  $x_i$  equals the true  $x_i^*$  plus measurement error on the part of the observer,  $\epsilon_i$ . Thus the true production function is  $q(\mathbf{x}^*) \equiv q(\mathbf{x} - \boldsymbol{\epsilon})$ , where  $\mathbf{x}^*$  is the true input vector. These measurement errors additively enter the demand system (3) as  $\epsilon_i$  and the cost function (2) as the weighted sum  $\sum w_j \epsilon_j$ .

Finally, by allowing  $\boldsymbol{\epsilon}$  to have a positive mean vector and resolving the  $\epsilon_i$ 's into a symmetrically distributed random variable plus a one-sided random error, the additive GEM may be interpreted as a stochastic frontier model with technical errors. This interpretation is discussed elsewhere (McElroy 1986), and adjusted SUR estimates are proposed. That paper also shows that, when the true model is an additive GEM but the usual restrictions found in the technical and allocative inefficiency literature are imposed, the researcher will tend to "find" inefficiencies even though none exists.

Although panel data may permit one to distinguish among these interpretations of AGEM errors, a single cross section or time series will not. Regardless of the interpretation, however, the relationship among the AGEM errors in the production, demand, cost, and share equations is intuitively pleasing. If for a particular firm  $\epsilon_i > 0$  ( $\epsilon_i < 0$ ), then the additive demand error in (3) shows that the firm uses  $\epsilon_i$  more (less) of  $x_i$  in the production of a given level of output  $y$  than does the average firm. Via (2), this in turn costs the firm  $w_i \epsilon_i$  more than the average firm. Although share errors (8) depend on the output level, all input prices, and the entire vector of AGEM errors  $\boldsymbol{\epsilon}$ , nonetheless

$$\frac{\partial v_i}{\partial \epsilon_i} = \frac{w_i [1 - S_i(y, \mathbf{w}, \boldsymbol{\epsilon})]}{C(y, \mathbf{w}, \boldsymbol{\epsilon})} > 0, \quad i = 1, \dots, n, \quad (15)$$

$$\frac{\partial v_i}{\partial \epsilon_k} = \frac{-w_i S_i(y, \mathbf{w}, \boldsymbol{\epsilon})}{C(y, \mathbf{w}, \boldsymbol{\epsilon})} < 0, \quad i, k = 1, \dots, n, k \neq i, \quad (16)$$

so that via  $v_i$  a larger  $\epsilon_i$  increases the cost share of input  $x_i$  and decreases the remaining cost shares.<sup>9</sup>

Thus the AGEM share error specification stands in contrast to the customary share error assumptions. The customary assumption has been that share errors are i.i.d. across observations and bear no particular relationship to the error in the cost equation (n. 5 gives two exceptions). Christensen and Greene (1976, p. 662), for example, clearly state the reasoning behind this customary assumption: "Since

<sup>9</sup> Note that the elasticity of share  $i$  with respect to  $\epsilon_i$  is

$$\frac{\partial S_i}{\partial \epsilon_i} \frac{\epsilon_i}{S_i} = \frac{w_i \epsilon_i}{C(y, \mathbf{w}, \boldsymbol{\epsilon})} \frac{1 - S_i(y, \mathbf{w}, \boldsymbol{\epsilon})}{S_i(y, \mathbf{w}, \boldsymbol{\epsilon})}$$

so that the percentage increase in the share is larger the larger the share of total cost accounted for by  $w_i \epsilon_i$  and the smaller the expenditure share of  $x_i$  in the total.

the cost share equations are derived by differentiation, they do not contain the disturbance term from the cost function." By contrast, the additive GEM has the advantage of an explicit, interpretable, and empirically exploitable functional relationship between cost and share errors.

E. *Flexible Functional Form GEMs:*

*A Translog Example*

This subsection presents a translog additive GEM. Assume that the deterministic production function  $q(\mathbf{x})$  is such that the dual mean cost function is translog. Thus the firm-specific AGEM cost function (2) takes the form

$$C(y, \mathbf{w}, \boldsymbol{\epsilon}) = \exp h(y, \mathbf{w}) + \sum w_j \epsilon_j, \tag{17}$$

where

$$\begin{aligned} h(y, \mathbf{w}) = & \ln \alpha_0 + \sum_i \alpha_i \ln w_i + \sum_i \sum_j \frac{1}{2} \gamma_{ij} \ln w_i \ln w_j \\ & + \sum_i \mu_i \ln w_i \ln y + \mu \ln y + \frac{\theta}{2} (\ln y)^2, \end{aligned}$$

and where the  $\alpha$ 's,  $\gamma$ 's, and  $\mu$ 's are parameters restricted to satisfy linear homogeneity in prices and symmetry, or

$$\sum_i \alpha_i = 1, \gamma_{ij} = \gamma_{ji}, \sum_i \gamma_{ij} = \sum_j \gamma_{ji} = 0, \sum_i \mu_i = 0. \tag{18}$$

The corresponding firm-specific share equations are the special case of (6),

$$S_i = \alpha_i + \sum_j \gamma_{ij} \ln w_j + \mu_i \ln y + v_i(y, \mathbf{w}, \boldsymbol{\epsilon}), \quad i = 1, \dots, n, \tag{19}$$

where  $v_i(\cdot)$  is given in (8) above. Via (12) and (14) the empirical implications of this translog additive GEM are embodied in the  $n$ -equation demand system

$$x_i = \frac{1}{w_i} [\exp h(y, \mathbf{w})] \left( \alpha_i + \sum_j \gamma_{ij} \ln w_j + \mu_i \ln y \right) + \epsilon_i, \quad i = 1, \dots, n, \tag{20}$$

subject to (18). This translog additive GEM is nonlinear in the parameters of the system but linear in the error terms and, as shown in Section III below, readily estimated.

Equivalently, the empirical implications are embodied in the  $n$ -

equation cost-cum-share system, (17) along with (19). Since the endogenous variable  $C(y, \mathbf{w}, \boldsymbol{\epsilon})$  enters the denominator of every  $v_i$  in the share system, it is a highly inconvenient, nonlinear system. Nonetheless, given distributional assumptions on the errors,  $\boldsymbol{\epsilon}$ , maximum likelihood estimators for both this cost-cum-share representation and the demand representation are identical: both are representations of the same underlying GEM. One can “solve” all the problems associated with estimating the share representation by estimating the isomorphic demand representation instead. Note that the estimated shares,  $\hat{s}_i(y, \mathbf{w}) = w_i \hat{c}_i(y, \mathbf{w}) / \sum w_j \hat{c}_j(y, \mathbf{w})$ , will automatically lie on the unit simplex.

### III. Estimation and Testing of Additive GEMs

By definition, the demand system for an additive GEM has additive errors that are i.i.d. across observations. Further, since every representation of a given GEM embodies the same stochastic specification, one need only estimate the most convenient representation, in this case the demand system. The demand system may be linear or nonlinear in parameters. For example, the generalized Diewert-Leontief AGEM demand system of the form (3) would be linear in parameters, whereas the translog AGEM of the form (20) would be nonlinear. Either way, across observations the AGEM error vector is assumed to be i.i.d. as  $\boldsymbol{\epsilon}(t) \sim (\mathbf{0}_n, \boldsymbol{\Sigma})$ , where  $\mathbf{0}_n$  is an  $n$  vector of zeros and  $\boldsymbol{\Sigma}$  is an  $n \times n$  positive definite contemporaneous covariance matrix.<sup>10</sup> Thus one may regard the demand system (3) or (12) as  $n$  nonlinear SUR regressions with cross-equations restrictions (as in [18] for the translog) and obtain strongly consistent and asymptotically normal least-squares estimators following Gallant (1986). Alternately, one may make the additional assumption that the errors are jointly normally distributed and obtain maximum likelihood estimators via iterated SUR nonlinear regressions with cross-equations restrictions following Gallant (1986, chap. 6).

It is important to test the additive GEM error specification against conventional error specifications, in particular, against the customary share error specification with additive i.i.d. errors.<sup>11</sup> On the one hand, in the AGEM the demand system errors,  $\boldsymbol{\epsilon}$ , have a joint distribution that is i.i.d. across observations and is free of input prices and the level of output; this implies that the share errors  $\bar{v}$  have a nasty ratio form,

<sup>10</sup> It would be straightforward to modify this structure by letting the  $\boldsymbol{\epsilon}$ 's be correlated over time in the manner of Parks (1967).

<sup>11</sup> To cite a few of the many examples, Berndt and Christensen (1973, 1974), Christensen and Greene (1976), and Appelbaum (1978) all use this specification. More sophisticated specifications such as Fuss (1977) also stem from this setup.

are heteroscedastic (see eq. [8]), and therefore cannot be i.i.d. On the other hand, in the conventional specification of a share system, the errors have a joint distribution that is i.i.d. across observations and is free of input prices and the level of output, which implies that the corresponding demand system errors have a nasty multiplicative form, are heteroscedastic, and therefore cannot be i.i.d. (see [13]).

The foregoing suggests the following encompassing strategy for testing. Suppose that the demand cost and share systems both have two-part error terms, each being the sum of a nice error term and a heteroscedastic one. Hence in the demand system the error term for the  $i$ th demand equation would be

$$d_i = \epsilon_i + e_i(y, \mathbf{w}, \tilde{\mathbf{v}}), \quad i = 1, \dots, n, \tag{21}$$

where  $e_i(\cdot)$  is defined in (13) and where  $\epsilon$  and  $\tilde{\mathbf{v}}$  are i.i.d. across observations as  $\epsilon \sim N(\mathbf{0}_n, \Sigma)$  and  $\tilde{\mathbf{v}} \sim N(\mathbf{0}_n, \Omega)$ ;  $\Sigma$  and  $\Omega$  are both positive definite covariance matrices.<sup>12</sup> Thus  $\epsilon_i$  is the nice component of the demand error and  $e_i(y, \mathbf{w}, \tilde{\mathbf{v}})$  is the heteroscedastic one (dual to the well-behaved share error,  $\tilde{\mathbf{v}}$ ). The specification (21) forms the basis of the alternative hypothesis to be pitted against the null hypothesis that the  $i$ th demand function error is simply  $d_i = \epsilon_i$ . Thus a readily testable implication of the alternative hypothesis is that the  $n \times n$  covariance matrix for  $(d_1, \dots, d_n)$  is heteroscedastic via dependence on  $y$  and  $\mathbf{w}$ . The corresponding null is that  $(d_1, \dots, d_n)$  has the fixed covariance matrix  $\Sigma$ . More precisely, letting  $t$  denote the  $t$ th observation, the alternative is

$$H_A: E[d_i(t)d_j(t)] = \sigma_{ij} + z'_{ij}(t)\alpha_{ij}(t), \quad i, j = 1, \dots, n, t = 1, \dots, T, \tag{22}$$

where  $\sigma_{ij}$  is the  $ij$ th element of the contemporaneous covariance of the  $\epsilon$ 's,  $\Sigma$ . Here, in accordance with (13), and noting that  $v_{ct}$  is the error for the  $t$ th observation on the cost function in (10), we get

$$\begin{aligned} E[e_i(y_t, w_t, \tilde{v}_t)e_j(y_t, w_t, \tilde{v}_t)] &= E\left\{\left[\frac{c(y_t, w_t)v_{it} + v_{ct}v_{it} + s_i(y_t, w_t)v_{ct}}{w_{it}}\right]\right. \\ &\quad \times \left.\left[\frac{c(y_t, w_t)v_{jt} + v_{ct}v_{jt} + s_j(y_t, w_t)v_{ct}}{w_{jt}}\right]\right\} \\ &= z'_{ij}(y_t, w_t)\alpha_{ij}(t), \end{aligned}$$

where  $z_{ij}(y_t, w_t)$  is the following  $8 \times 1$  vector whose elements depend on  $(y_t, w_t)$ :

<sup>12</sup> Correspondingly, the share functions have error terms  $v_i + v_i(y, \mathbf{w}, \epsilon)$ . The symmetric tests for the share functions involve ratios of stochastically dependent random variables and are not pursued here.

$$z'_{ij}(y_t, w_t) = [c^2(y_t, w_t), 2c(y_t, w_t), c(y_t, w_t)s_j(y_t, w_t), c(y_t, w_t)s_i(y_t, w_t), \\ 1, s_j(y_t, w_t), s_j(y_t, w_t)s_i(y_t, w_t), s_i(y_t, w_t)]/w_{it}w_{jt}, \quad (23)$$

and where  $\alpha_{ij}(t)$  is the following  $8 \times 1$  vector of constants whose elements are functions of the elements of  $\Omega$ :

$$\alpha'_{ij}(t) = E(v_{it}v_{jt}, v_{ct}v_{it}v_{jt}, v_{ct}v_{it}, v_{ct}v_{jt}, v_{ct}^2v_{it}v_{jt}, v_{ct}^2v_{it}, v_{ct}^2v_{jt}).$$

Under the null hypothesis, the covariance terms,  $E[e_i(y_t, w_t, \bar{v}_t)e_j(y_t, w_t, \bar{v}_t)]$ , corresponding to the nice share errors  $\bar{v}$  vanish.<sup>13</sup> More specifically, the null hypothesis is then that each  $\alpha_{ij}$  is a vector of zeros, or

$$H_0: E[d_i(t)d_j(t)] = \sigma_{ij}, \quad i, j = 1, \dots, n, t = 1, \dots, T.$$

A generalization of Breusch-Pagan Lagrange multiplier (LM) heteroscedasticity tests (1979) to a system of equations is that if the  $\epsilon(t)$  are jointly normal and i.i.d. across observations, then under the  $(n^2 + n)/2$  null hypotheses, the statistics

$$LM_{ij} = \frac{\sum_{t=1}^T [z'_{ij}(y_t, w_t)\hat{\alpha}_{ij}]^2}{2}, \quad i, j = 1, \dots, n, i \leq j, \quad (24)$$

are asymptotically distributed as  $\chi^2$  random variates with eight degrees of freedom (the number of parameters in  $\alpha_{ij}$ ). Here  $\hat{\alpha}_{ij}$  is the ordinary least squares estimator of  $\alpha_{ij}$  from the regression of  $(x_i - \hat{x}_i)(x_j - \hat{x}_j)/\hat{\sigma}_{ij}$  on  $z_{ij}$ , and in turn,  $\hat{x}_i$ ,  $\hat{x}_j$ , and  $\hat{\sigma}_{ij}$  are the NLITSUR estimators of the corresponding unhatted quantities. These tests can be approximated asymptotically by replacing  $z_{ij}(y_t, w_t)$  with  $\hat{z}_{ij}(y_t, w_t)$ , the vector whose elements are determined by the predicted total cost and shares,  $\hat{c}(y_t, w_t)$  and  $\hat{s}_i(y_t, w_t)$ , for the  $t$ th observation from NLITSUR in accordance with (23). Rejections of the null hypotheses would be evidence in favor of i.i.d. errors for share equations ( $\bar{v}$ ) with their corresponding heteroscedastic covariance in (22). Nonrejections would support the AGEM specification with i.i.d. demand errors,  $\bar{\epsilon}$ .<sup>14</sup>

Note that since  $i$  and  $j$  run from 1 to  $n$  commodities and since covariance matrices are symmetric, there are  $(n^2 + n)/2$  hypotheses to test. By stacking the equations, one could obtain one grand joint-generalized Breusch-Pagan test to jointly test all  $(n^2 + n)/2$  hypotheses. The  $(n^2 + n)/2$  tests proposed here, however, are more likely to uncover some evidence for rejection than is one grand joint test.

<sup>13</sup> Note that the null also shifts the intercept terms. Under the null,  $E[d_i(t)] = 0$ , whereas under the alternative,  $E[d_i(t)] = \text{cov}(v_{it}\epsilon_{it}) = w_{it}^{-1}E\{[c(y, \mathbf{w}) + v_c]v_{it} + s_i(y, \mathbf{w})v_{it}\} = \Omega_{ni}/w_i$ , where  $\Omega_{ni} \equiv E(v_c v_i)$ .

<sup>14</sup> A sharper alternative could be used that would account for the restrictions on  $\alpha$  under the alternative.

Therefore, since the multiple tests proposed here stack the cards against the AGEM, these multiple tests are indicated.

#### IV. An Empirical Example and Results from a Specification Test

For comparability with other studies, the translog additive GEM model was estimated with constant returns to scale ( $\mu_i = \theta = 0$ ,  $\mu = 1$ ) maintained. These data were annual aggregates for U.S. manufacturing, 1947–71, from Berndt and Wood (1975), who describe them in detail. These data contain total input cost as well as price and quantity indices for  $n = 4$  inputs (capital, labor, energy, and materials) and  $T = 25$  observations. As pointed out by Berndt and Wood, at the aggregate industry level the exogeneity of prices is in doubt. This raises a serious question with no obvious solution. Their solution was to use a set of instrumental variables “considered exogenous to the United States manufacturing sector” (p. 263) to obtain fitted values for the prices by employing iterated three-stage least squares. In contrast, others have argued that in this situation the use of instruments is likely to yield the results that “may be very sensitive to the choice of instruments [Appelbaum 1978]” (p. 94) and therefore avoided instruments. In my judgment, especially when the goal is, as here, to compare models, it is better to have results that are insensitive to the choice of instruments. Therefore, I eschewed their use.

Table 1 presents maximum likelihood estimates of the coefficients (and their asymptotic  $t$ -ratios) for two constant-returns-to-scale translog models. Both models have translog deterministic share functions with constant returns to scale. The models differ solely by their error assumptions. Under their respective error assumptions, FIML estimates are obtained for each. The first column gives the FIML estimates for the translog additive GEM model (eqq. [17]–[20]) when the demand errors ( $\epsilon_1, \dots, \epsilon_n$ ) are assumed i.i.d.  $N(\mathbf{0}_n, \mathbf{\Sigma})$ . These were obtained via NLITSUR regressions for the demand system (20). The second column gives the ITSUR regression parameter estimates for the translog cost-cum-share model:

$$\ln C = h(y, \mathbf{w}) + v_c, \quad (25)$$

$$S_i = \alpha_i + \sum_j \gamma_{ij} \ln w_j + v_i, \quad i = 1, \dots, n - 1,$$

where  $h(y, \mathbf{w})$  is the standard deterministic translog cost function given above, and  $(v_1, \dots, v_{n-1}, v_c)$  are assumed i.i.d.  $N(\mathbf{0}_n, \mathbf{\Omega})$ . As pointed out in Section IID above, this latter error assumption embodies the standard error assumptions in the literature but is inconsis-

TABLE 1  
 COMPARISON OF NLITSUR ESTIMATES OF A TRANSLOG  
 FORM AGEM ERROR SPECIFICATION WITH THE  
 TRADITIONAL ERROR SPECIFICATION

Parameter	AGEM (1)	Traditional Error Specification (2)
$\alpha_0$	5.086 (480.39)	5.077 (526.27)
$\alpha_K$	.054 (37.71)	.051 (42.07)
$\alpha_L$	.257 (106.93)	.252 (113.55)
$\alpha_E$	.040 (35.67)	.041 (50.96)
$\alpha_M$	.648 (184.25)	.656 (198.75)
$\gamma_{KK}$	.038 (8.11)	.037 (7.51)
$\gamma_{KL}$	.008 (2.69)	.018 (4.73)
$\gamma_{KE}$	-.007 (-2.33)	-.006 (-2.26)
$\gamma_{KM}$	-.038 (-4.72)	-.048 (-6.00)
$\gamma_{LL}$	.068 (10.35)	.081 (11.49)
$\gamma_{LE}$	.006 (2.33)	.007 (3.05)
$\gamma_{LM}$	-.082 (-8.86)	-.106 (-9.81)
$\gamma_{EE}$	.016 (2.26)	.019 (3.80)
$\gamma_{EM}$	-.014 (-1.37)	-.021 (-2.67)
$\gamma_{MM}$	.134 (6.25)	.176 (8.34)

NOTE.—*t*-values are in parentheses.

tent with the AGEM. Thus, despite a common deterministic share structure, (20) and (25) are not nested.

As table 1 shows, the estimated coefficients are similar for the two models; the  $\alpha_i$ 's differ by at most 0.008 and the  $\gamma_{ij}$ 's differ by at most 0.042. Thus the estimated coefficients appear to be robust to the differences in the error assumptions. This robustness is comforting in that violent swings in parameter estimates due solely to changes in error assumptions would call into question the specification of the deterministic parts of the model (the translog in this case).



Although the parameter estimates are quite similar for the two models, in 13 of 15 cases the  $t$ -ratios for the AGEM model (col. 1) are smaller than for the traditional i.i.d. share errors model (col. 2). These results—robustness of coefficient estimates but different  $t$ -ratios (estimated standard errors)—partially parallel those of Woodland (1979). He compared maximum likelihood estimates for a statistically acceptable share error structure (a Dirichlet distribution that confines the estimated shares to the unit simplex) with estimates for a standard model with normal i.i.d. share errors. He too found the estimated parameter values to be robust with respect to these specifications of the error distributions. He also found that if the customary error specification is incorrectly assumed to be the correct model, then in his study the estimated  $t$ -ratios (calculated via the wrong formula) are on average too small. The thrust of the results presented here differs from Woodland's. On the basis of the limited evidence of this study, if the AGEM is the true model but the traditional error specification is used, then the reported  $t$ -ratios will be calculated via the wrong formula and tend to be bigger than the true  $t$ -ratios. This is a potentially serious problem. There are many possible reasons why these results differ from Woodland's. An interesting candidate is that his use of the Dirichlet distribution restricts the covariances of the share errors to be negative—a restriction not present in the AGEM.

It is worth noting that the problem is especially serious since the parameters of interest are often not the coefficients of the cost and share functions, but elasticities that are not at all robust with respect to the underlying parameter estimates, particularly for inputs with small shares.

Table 2 permits a comparison of the estimated Allen-Uzawa partial elasticities of substitution between the translog GEM and traditional translog specification. These elasticities were calculated at the sample means shares ( $\bar{y}_i$ 's) in accordance with the well-known formulas

$$\begin{aligned}\sigma_{ii} &= \frac{\gamma_{ii} + \bar{y}_i^2 - \bar{y}_i}{\bar{y}_i^2}, \\ \sigma_{ij} &= \frac{\gamma_{ij} + \bar{y}_i \bar{y}_j}{\bar{y}_i \bar{y}_j}, \quad i \neq j.\end{aligned}\tag{26}$$

As these formulas show, elasticities involving inputs that account for only small shares in total cost are quite sensitive to the size of the estimated  $\gamma_{ij}$ 's. Small differences in estimated parameters can and do make large differences in estimates of partial elasticities. Here, for example, for seven out of 10 estimated partial elasticities, the larger (in absolute value) estimate is more than 25 percent greater than the

TABLE 2

ESTIMATED ALLEN-UZAWA PARTIAL ELASTICITIES OF SUBSTITUTION  
FOR THE TRANSLOG AGEM AND THE TRADITIONAL  
TRANSLOG ERROR SPECIFICATIONS CALCULATED AT THE  
SAMPLE MEAN SHARES ( $\bar{y}_i$ 's)

$\sigma_{ij}$	AGEM	Traditional Share Errors
$\sigma_{KK}$	-4.58	-4.77
$\sigma_{KL}$	1.56	2.20
$\sigma_{KE}$	-2.06	-1.51
$\sigma_{KM}$	-.14	-.44
$\sigma_{LL}$	-1.17	-1.56
$\sigma_{LE}$	1.46	1.59
$\sigma_{LM}$	.52	.38
$\sigma_{EE}$	-13.51	-11.64
$\sigma_{EM}$	.50	.27
$\sigma_{MM}$	-.25	-.15

smaller estimate. An interesting case is  $\sigma_{KE}$ , the partial elasticity for capital and energy. Although the AGEM and traditional specifications result in estimates of  $\gamma_{KE}$  that differ by only .001, the corresponding estimated partial elasticity of substitution between capital and labor is 37 percent higher for the AGEM than for the traditional error specification. This extreme sensitivity of estimated partial elasticities to the error assumptions is due to the small cost shares of capital and energy.<sup>15</sup>

The results above dramatize the need for efficient parameter estimates and the use of correct formulas in calculating standard errors and *t*-ratios. We therefore turn to the specification tests. Table 3 presents LM statistics (24) for the Breusch-Pagan tests outlined in Section III above. These are based on a model that encompasses both the AGEM and the traditional share specification. In each of these  $(4^2 + 4)/2 = 10$  cases, the null hypothesis is that the AGEM covariance specification holds and that the traditional share error specification adds nothing to the covariance; the alternative hypothesis is that the errors exhibit the heteroscedasticity implied by the traditional i.i.d. share error specification. In every one of these 10 tests, the null hypothesis that the GEM specification holds cannot be rejected. As pointed out in Section III, these tests really stack the deck in favor of

<sup>15</sup> Over this sample period the mean cost shares for capital and energy are approximately .05 and .04, respectively. Thus, although the two estimates of  $\gamma_{KE}$  differ (in absolute value) by only  $\Delta\hat{\gamma}_{KE} = .001$ , according to (26) the implied estimates of  $\sigma_{KE}$  will differ by .5:  $\Delta\hat{\sigma}_{KE} = (\bar{y}_K\bar{y}_E)^{-1}\Delta\hat{\gamma}_{KE} = (500)(.001) = .5$ .

TABLE 3  
 BREUSCH-PAGAN TESTS OF HETEROSCEDASTICITY  
 IN THE NESTED MODEL

$LM_{ij}^*$	$K$	$L$	$E$	$M$
$K$	4.58	...	...	...
$L$	4.50	7.82	...	...
$E$	2.15	4.74	5.40	...
$M$	3.46	.47	5.70	5.99

\* The subscripts  $i$  and  $j$  are from  $\hat{\epsilon}_i\hat{\epsilon}_j/2(\hat{\sigma}_{ij})^2$ , where  $\hat{\epsilon}_i$  and  $\hat{\epsilon}_j$  are the ITSUR residual vectors from the  $i$ th and  $j$ th input demand equations and  $\hat{\sigma}_{ij}$  is the corresponding estimated variance or covariance. In each case there are eight degrees of freedom with .05 and .01 cutoffs of 15.5 and 20.1, respectively.

rejecting the AGEM model.<sup>16</sup> Thus for these data there is strong evidence in favor of the AGEM over the traditional share error specification.

## V. GEMs as a Research Strategy

Error specifications are fundamental to empirical work. The general error model approach of this paper is consistent with the fundamental and historically fruitful theory-based research strategy: push the static neoclassical model as far as possible. After this core theory is exhausted, then introduce generalizations such as true dynamics or risky decisions. To reiterate, GEMs are a mathematical embodiment of Stigler's (1976) view that apparent inefficiencies arise because we are ignorant of the true optimization problem. Thus GEMs are appropriate when the data or resource constraints do not permit modeling these unknown costs and objectives.

The additive GEM permits the best of both worlds. First, one can choose the deterministic demand or share system from the entire extant menu of deterministic systems derived from production theory or approximation theory. This includes all the flexible functional forms such as the translog and the generalized Diewert-Leontief, as well as exact functional forms such as the constant and variable elasticity of substitution. Second, one can use the simple and theoretically justifiable AGEM error structure.

The GEMs proposed and estimated in this paper extend the neoclassical theory of the firm to embrace additive error specifications in

<sup>16</sup> The LM test may have low power. Also we know that individual tests on the 10 hypotheses such as  $E[\epsilon_i(t)\epsilon_j(t)]$  for  $i, j = 1, 2, 3, 4$  and  $i \neq j$  are more likely to uncover evidence of rejection than one grand joint test that all 10 null hypotheses hold. Nonrejection of the individual tests implies nonrejection for a joint test.

input demand and cost-cum-share systems. These additive GEMs have many theoretical advantages over conventional error specifications; they are readily estimated. Further, the empirical example given here indicates that the translog AGEM model is superior to the standard translog share model with additive i.i.d. errors.

The importance of any error specification resides in its impact on empirical results. Conventional error specifications for behavioral equations, although inconsistent with the specified underlying optimization problem, may, nevertheless, yield acceptable empirical results. Research on GEMs consonant with economic theory will either discover that conventional estimators are robust or alter existing empirical conclusions in important ways. Much work is needed before an empirical verdict on the additive general error model of production emerges, and many obvious extensions require investigation.<sup>17</sup>

## References

- Appelbaum, Elie. "Testing Neoclassical Production Theory." *J. Econometrics* 7 (February 1978): 87–102.
- Berndt, Ernst R., and Christensen, Laurits R. "The Translog Function and the Substitution of Equipment, Structures, and Labor in U.S. Manufacturing, 1929–68." *J. Econometrics* 1 (March 1973): 81–113.
- . "Testing for the Existence of a Consistent Aggregate Index of Labor Inputs." *A.E.R.* 64 (June 1974): 391–404.
- Berndt, Ernst R., and Khaled, Mohammed S. "Parametric Productivity Measurement and Choice among Flexible Functional Forms." *J.P.E.* 87 (December 1979): 1220–45.
- Berndt, Ernst R., and Wood, David O. "Technology, Prices, and the Derived Demand for Energy." *Rev. Econ. and Statis.* 57 (August 1975): 259–68.
- Breusch, T. S., and Pagan, Adrian R. "A Simple Test for Heteroscedasticity and Random Coefficient Variation." *Econometrica* 47 (September 1979): 1287–94.
- Christensen, Laurits R., and Greene, William H. "Economies of Scale in U.S. Electric Power Generation." *J.P.E.* 84, no. 4, pt. 1 (August 1976): 655–76.
- Diewert, W. E. "Applications of Duality Theory." In *Frontiers of Quantitative Economics*, vol. 2, *Contributions to Economic Analysis*, edited by Michael D. Intriligator and David A. Kendrick. Amsterdam: North-Holland, 1974.
- Duncan, Gregory M. "Formulation and Statistical Analysis of the Mixed, Continuous/Discrete Variable Model in Classical Production Theory." *Econometrica* 48 (May 1980): 839–52.
- Førsund, Finn R.; Lovell, C. A. Knox; and Schmidt, Peter. "A Survey of Frontier Production Functions and of Their Relationship to Efficiency Measurement." *J. Econometrics* 13 (May 1980): 5–25.
- Fuss, Melvyn A. "The Demand for Energy in Canadian Manufacturing: An

<sup>17</sup> Among these extensions are the relationship between GEMs and stochastic frontier production functions with both technical and allocative errors as well as GEMs for neoclassical consumer theory. Research is under way in all these areas (see McElroy 1981, 1986).

- Example of the Estimation of Production Structures with Many Inputs." *J. Econometrics* 5 (January 1977): 89–116.
- Fuss, Melvyn A.; McFadden, Daniel; and Mundlak, Yair. "A Survey of Functional Forms in the Economic Analysis of Production." In *Production Economics: A Dual Approach to Theory and Applications*, vol. 1, *The Theory of Production*, edited by Melvyn A. Fuss and Daniel McFadden. Amsterdam: North-Holland, 1978.
- Gallant, A. Ronald. *Nonlinear Statistical Models*. New York: Wiley, 1986.
- Lau, Lawrence J. "Applications of Duality Theory: Comment." In *Frontiers of Quantitative Economics*, vol. 2, *Contributions to Economic Analysis*, edited by Michael D. Intriligator and David A. Kendrick. Amsterdam: North-Holland, 1974.
- McElroy, Marjorie B. "Duality and the Error Structure of Consumer Demand and Producer Demand and Cost Systems." Discussion Paper no. 80-10. Chicago: Nat. Opinion Res. Center, Econ. Res. Center, 1981.
- . "General Error Models for Stochastic Production Frontiers with Both Technical and Allocative Errors." Durham, N.C.: Duke Univ., 1986.
- Parks, Richard W. "Efficient Estimation of a System of Regression Equations When Disturbances Are Both Serially and Contemporaneously Correlated." *J. American Statist. Assoc.* 62 (June 1967): 500–509.
- Rossi, Peter. "Stochastic Specification of Cost and Production Relationships." Mimeographed. Evanston, Ill.: Northwestern Univ., 1983.
- Shephard, Ronald W. "Applications of Duality Theory: Comment." In *Frontiers of Quantitative Economics*, vol. 2, *Contributions to Economic Analysis*, edited by Michael D. Intriligator and David A. Kendrick. Amsterdam: North-Holland, 1974.
- Simmons, Peter, and Weiserbs, Daniel. "Translog Flexible Functional Forms and Associated Demand Systems." *A.E.R.* 69 (December 1979): 892–901.
- Stigler, George J. "The Existence of X-Efficiency." *A.E.R.* 66 (March 1976): 213–16.
- Theil, Henri. "The Information Approach to Demand Analysis." *Econometrica* 33 (January 1965): 67–87.
- . "A Theory of Rational Random Behavior." *J. American Statist. Assoc.* 69 (June 1974): 310–14.
- Woodland, A. D. "Stochastic Specification and the Estimation of Share Equations." *J. Econometrics* 10 (August 1979): 361–83.
- Zellner, Arnold; Kmenta, Jan; and Drèze, Jacques. "Specification and Estimation of Cobb-Douglas Production Function Models." *Econometrica* 34 (October 1966): 784–95.
- Zellner, Arnold, and Revankar, N. S. "Generalized Production Functions." *Rev. Econ. Studies* 36 (April 1969): 241–50.