Production economics

1982

UCD Department of Agricultural Economics

WORKING PAPER SERIES

University of California, Davis
Department of Agricultural Economics

TESTING THE STOCHASTIC STRUCTURE OF PRODUCTION: A FLEXIBLE MOMENT-BASED APPROACH

by

John M. Antle

Working Paper No. 82-2 (Revised)

TESTING THE STOCHASTIC STRUCTURE OF PRODUCTION:
A FLEXIBLE MOMENT-BASED APPROACH

The aim of this paper is to develop a flexible moment-based approach to specifying, estimating, and testing stochastic production models. This approach provides a statistical methodology for estimating not only "mean" output as a function of inputs, as is done in conventional production function models, but also the variance, third moment, and higher moments can be specified and estimated as functions of inputs. The moment-based approach to the study of production economics is motivated by a number of facts.

First, the probability distribution of output is a unique function of its moments and, therefore, the behavior of the firm under stochastic production can be defined in terms of the relationships between inputs and these moments. Any characteristic of a firm's stochastic technology can be measured and tested using the moment functions.

Second, conventional econometric production models are based on the *ad hoc* appending of additive or multiplicative random error terms to deterministic production functions. This is true for both the "mean" and "frontier" production models in the literature. It is shown in this paper that these models are not generally adequate representations of the probability distribution of output, because they impose arbitrary restrictions on the moments of output which result in arbitrary restrictions on the behavior of the firm. Other studies have also found that the error specification has important economic implications (Just and Pope, 1978, Pope and Kramer 1979, Newberry and Stiglitz, 1982). *These restrictions can not be tested using conventional production function models.*

A third motivation for the moment based approach comes from empirical evidence (Day 1965, Anderson 1974, Roumasset 1975, Just and Pope 1979, Nikiphoroff 1981, Antle and Goodger 1982) which shows that second, third, and fourth moments of output may be functions of inputs. The theory of decision making under uncertainty shows that empirical production models need to account for these relationships.

The above facts suggest that an econometric production model is needed which provies a general representation of the probability distribution of output and which does not impose arbitrary restrictions on the moments. The moment-based approach developed in this paper is "flexible" in the sense that it imposes relatively few restrictions, or maintained hypotheses, on the probability distribution of output and thus provides a useful framework for testing the stochastic structure of production.

Estimators for the moment-based approach are developed within a conventional linear regression framework. The estimation methodology can be viewed as a generalization of the heteroscedastic regression models devised by Goldfeld and Quandt (1972) and Amemiya (1978). The generalized least squares (GLS) technique is used to obtain estimators for the moment functions which are consistent and asymptotically normal under the conditions of the Lindberg-Feller central limit theorem so they can be used to test hypotheses. As Judge et al. (1982) have noted, one limitation of the Goldfeld-Quandt and Amemiya models is that negative variance estimates can be obtained for the GLS model; in this paper it is shown that using standard nonlinear programming techniques, and recently developed software, variances for the GLS models can be estimated under the constraint that they are nonnegative. To test cross-moment restrictions implied by conventional production models, a joint GLS estimator is developed which is a heteroscedastic version of Zellner's (1962) seemingly unrelated regression estimator.

The paper begins with a discussion of the restrictions imposed on the stochastic structure of production and the behavior of the firm by the production function models in the literature. In the second section the theoretical foundations for the flexible moment-based approach are developed, moment-based approximations are discussed, and a quadratic moment model is used to illustrate how the stochastic structure of production can be tested. The third section develops the econometric model and the associated parameter estimators. The final section presents an application of the moment-based approach to milk production data. The results strongly support the hypothesis that both the variance and third moment of the output distribution are statistically significant functions of inputs. The restrictions implied by models in the literature are formally tested. The results show that both the multiplicative error model and the heteroscedastic, additive error model are poor approximations to the output distribution and are rejected by the data.

I. MAINTAINED HYPOTHESES OF STOCHASTIC SPECIFICATIONS.

Before introducing the flexible moment-based approach, I shall discuss the restrictions, or maintained hypotheses, embodied in conventional stochastic production models, and the economic implications of these restrictions.

First consider the maintained hypotheses of the multiplicative error model

(1) $Q = m(x, \beta)e^{u}$

where Q is output, $x = (x_1,...,x_n)$ is a vector of inputs, β is a conformable parameter vector, and u is a random error. This specification is convenient because it can be expressed as an additive error model by taking logarithms and it satisfies the constraint that output is non-negative. However, there is little theoretical justification for this particular model and it can be shown to impose a number of arbitrary restrictions on the stochastic structure of the production process. Letting $E[\cdot]$ denote the mathematical expectation operator, the mean of output is

$$\mu_1 = E[Q] = m(x, \beta)\, E[e^u]$$

the variance is

$$\mu_2 = E[Q - E(Q)]^2 = m(x, \beta)^2\, E[e^u - E(e^u)]^2$$

and in general the ith moment about μ_1 is

$$\mu_i = m(x, \beta)^i\, E[e^u - E(e^u)]^i.$$

The multiplicative error model implies that the mean and the higher moments of the probability distribution of output are functions of inputs through the function $m(x,\beta)$. The set of restrictions, or maintained hypotheses, implied by this model can be expressed in terms of the elasticities of moments with respect to inputs. From above we have, for $\mu_i \neq 0$,

$$(2)\quad \eta_{ik} \equiv \frac{\partial \mu_i}{\partial x_k} \cdot \frac{x_k}{\mu_i} = i\, \frac{\partial m(x,\beta)}{\partial x_k} \cdot \frac{x_k}{m(x,\beta)} = i\eta_{1k},\ i \geq 2.$$

Thus the elasticity of the ith moment with respect to the kth input, η_{ik}, is proportional to the mean production elasticity η_{1k}. These are the cross-moment restrictions implied by model (1).[1]

Another frequently used specification is the additive error model

$$(3)\quad Q = m(x, \beta) + u.$$

Typically u is assumed to be independently and identically distributed across all observations, and the distribution of u is assumed not to depend on x. Under this specification <u>only</u> the mean of the output distribution is assumed to be a function of inputs; all other moments are independent of x. Thus model (3) implies that $\eta_{ik} = 0$ for all $i \geq 2$.

Just and Pope (1978) have suggested that some of the restrictions embodied in models (1) and (2) can be circumvented by utilizing the heteroscedastic model proposed by Harvey (1976). Model (3) is specified with the heteroscedastic error structure

$$u = h(x, \gamma)\varepsilon$$

where ε is an independently and identically distributed error term. This allows the inputs to have different effects on the mean $m(x, \beta)$ and the variance of output

$$h(x, \gamma)^2 \, E(\varepsilon^2).$$

While this model does represent a generalization of models (1) and (3) because it does not restrict the effects of inputs on the variance to be related to the mean, it can easily be shown that it does restrict the effects of x across the second and higher moments in exactly the way model (1) does across all moments. To see this we simply note that, from above,

$$E(u^i) = h(x, \gamma)^i \, E(\varepsilon^i) = \mu_i.$$

For $i > 2$ and $E(\varepsilon^i) \neq 0$ the parameters of the ith moment are directly related to the parameters of the second moment; in particular,

$$(4) \quad \eta_{ik} \equiv \frac{\partial \mu_i}{\partial x_k} \cdot \frac{x_k}{\mu_i} = i \, \frac{\partial h(x,\gamma)}{\partial x_k} \cdot \frac{x_k}{h(x,\gamma)} = \frac{i}{2} \, \eta_{2k}, \; i > 2.$$

Therefore, the elasticity of each higher nonzero moment with respect to an input is directly proportional to the elasticity of the second moment with respect to that input. The restrictions (2) imposed on moment functions by model (1) are identical to the restrictions (4) of the Just-Pope model <u>except</u> that the former apply to <u>all</u> moments and the latter apply to <u>second</u> and <u>higher</u> moments. The restrictions in (4) are valid if output follows a "two parameter" distribution, such as the normal distribution, otherwise they generally are not valid.[2]

Restrictions (2) and (4) are economically important because they constrain the firm's behavior under uncertainty. To illustrate, consider the negative exponential utility function $U(\pi) = a-be^{-c\pi}$, where a, b, and c are positive parameters. For simplicity assume prices are nonstochastic and define "normalized" profit as $\pi = Q - \sum_{i=1}^{n} r_i x_i$, where r_i is the ith input price divided by the output price. An mth order Taylor series expansion of $U(\pi)$ about expected profit $\bar{\pi}$ gives

$$E\left[U(\pi)\right] = a-be^{-c\bar{\pi}} - be^{-c\bar{\pi}} \sum_{i=2}^{m} \frac{(-c)^i}{i!} \mu_i.$$

To further simplify the discussion consider a third-order expansion of the utility function. The first-order condition for maximization of expected utility can then be written as

$$(5) \quad \frac{\partial \mu_1}{\partial x_k} + \delta^{-1} \frac{(-c)}{2} \frac{\partial \mu_2}{\partial x_k} + \delta^{-1} \frac{(-c)^2}{6} \frac{\partial \mu_3}{\partial x_k} = r_k, \ k=1,\ldots, n,$$

where

$$\delta = 1 + \frac{(-c)^2}{2} \mu_2 + \frac{(-c)^3}{6} \mu_3.$$

Equation (5) can be rewritten as

$$(6) \quad \eta_{1k} + \delta^{-1} \frac{(-c)}{2} \frac{\mu_2}{\mu_1} \eta_{2k} + \delta^{-1} \frac{(-c)^2}{6} \frac{\mu_3}{\mu_1} \eta_{3k} = \frac{r_k x_k}{\mu_1}, \ k=1,\ldots, n,$$

which shows that the firm's behavior can be expressed in terms of the elasticities of moments with respect to inputs. In equilibrium the k^{th} factor share $r_k x_k/\mu_1$ equals a linear combination of the η_{ik}. Equation (6) shows that as the Arrow-Pratt risk aversion parameter c approaches zero, inputs are chosen such that the mean production elasticity η_{1k} equals the mean factor share $r_k x_k/\mu_1$, as would be the case for a risk-neutral firm. For large positive values of c, equation (6) shows that the equilibrium condition of the risk-neutral firm generally is not satisfied.[3]

Equation (6) can be used to show how restrictions (2) and (4) affect the behavior of the firm. Substituting (2) into (6) and simplifying we have

$$(7)\quad \eta_{1k}\left[1 + \delta^{-1}(-c)\,\frac{\mu_2}{\mu_1} + \delta^{-1}\frac{(-c)^2}{2}\frac{\mu_3}{\mu_1}\right] = r_k x_k/\mu_1$$

Equation (7) implies that $\eta_{1k} \gtreqless r_k x_k/\mu_1$ as $c\mu_3/2 \gtreqless \mu_2$. Thus restrictions (2) imply that if the output distribution is negatively skewed the firm chooses inputs such that $\eta_{1k} > r_k x_k/\mu_1$; and if $c\mu_3/2 = \mu_2$ the risk-averse firm behaves exactly like a risk-neutral firm! There does not appear to be any theoretical explanation why such behavior should in fact be observed. Equation (6) shows input choices could satisfy $\eta_{1k} > r_k x_k/\mu_1$, $\eta_{1k} = r_k x_k/\mu_1$, or $\eta_{1k} < r_k x_k/\mu_1$, regardless of the value of μ_3.

It is thus clear that models which impose arbitrary restrictions on the moment functions also impose arbitrary restrictions on the firm's behavior.

II. THEORETICAL FOUNDATIONS OF THE FLEXIBLE MOMENT-BASED APPROACH.

In this section I develop the theoretical foundations for a moment-based approach to the study of production economics. This approach is motivated by the fact that the conventional production function approach to modeling stochastic production processes imposes arbitrary restrictions on the relationship between inputs and the probability distribution of output. Instead of prameterizing a deterministic production function and appending an error term to it, the moment-based approach begins with a general parameterization of the moments of the probability distribution of output. In this way more flexible representations of output distributions are obtained.

The problem of uniquely characterizing the probability distribution of output is solved by utilizing the results of the "Stieltjes moment problem" (Rao 1973, p. 106). These results show that the probability distribution of output is a unique function of its moments.[4] Therefore, all economically relevant characteristics of the technology must be embodied in the relationships between inputs and moments. The behavior of the firm under production uncertainty can, therefore, always be defined in terms of the moments of the probability distribution of output. The moment-based approach to production economics, therefore, begins with a general representation of the moment functions which describe a stochastic technology. Let the probability distribution of output Q for a given input set x be $f(Q|x)$. The moment functions are written generally as

$$\mu_1(x, \gamma_1) = \int Q\, f(Q|x)\, dQ \tag{8}$$
$$\mu_i(x, \gamma_i) = \int (Q - \mu_i)\, f(Q|x)\, dQ, \quad i \geq 2.$$

where the γ_i are parameters relating x to μ_i. With this approach the production model may exhibit not only heteroscedasticity (μ_2 a function of x) but also "heteroskewness" as a function of inputs (μ_3 a function of x), and generally any moment of the distribution may be a function of inputs.[5]

The general representation of the moments of the probability distribution of output in equation (8) is "flexible" in the sense that each moment function depends on a distinct parameter vector. Thus, within or cross-moment restrictions are not imposed on the model, in contrast to the production function models discussed above which restrict the relationship between inputs and moments because all moment functions depend on the same parameters. The need for a flexible representation of the stochastic production model is analogous to the need for "flexible functional forms" in production and demand models (Diewert 1974, Fuss, McFadden, and Mundlak 1978).

While the flexible moment-based approach relaxes restrictions on the moment functions, it creates an "incidental parameter problem" in the sense that there are as many different parameter vectors as moments. Obviously, a useful representation of stochastic technologies cannot be based on a very large number of parameters. In the flexible moment-based approach this problem is resolved by the principle that research should strive to obtain a good approximation to the true underlying relationships identified by theory. Since we seek a good approximation to the true distribution of output, we must ask how many moments one need know to adequately represent the behaviorally relevant characteristics of the distribution. Kendall and Stuart (1976) provide one means of solving this problem. They show that a probability distribution can be approximated to the nth degree by an nth degree polynomial whose coefficients are functions of the first n moments of the distribution. They conclude from this result (p. 90):

> Thus distributions which have a finite number of the lower moments in common will, in a sense, be approximations one to another. We shall encounter many cases where, although we cannot determine a distribution explicitly, we may ascertain its moments at least up to some order; and hence we shall be able to approximate to the distribution by finding another distribution of known form which has the same lower moments. In practice, approximations of this kind often turn out to be remarkably good, even when only the first three of four moments are equated.

These observations by Kendall and Stuart are consistent with the usual practice of characterizing distributions in terms of the main "shape" characteristics of distributions, namely location (mean), dispersion (variance), skewness (third moment), and possibly also kurtosis (fourth moment). The fact that many distributions can be adequately represented in terms of four or fewer moments is also demonstrated by the Pearson system of distributions (Kendall and Stuart, 1976, Ch. 6). Members of the Pearson system, such as the Beta and Gamma distributions, are known to be functions of not more than the first four moments.[6] Another argument in favor of moment-based approximations has been put forward by Anderson et al. (1980, pp. 97-98) for the analysis of firm behavior in the expected utility framework. They note that when expected utility is approximated by a Taylor series, as in the previous section, terms beyond the third or forth moment usually add insignificantly to the precision of the approximation. These considerations all suggest that a useful approach would be to utilize three or four moments to represent a stochastic production process.

The principle that empirical research should strive for good approximations also suggests that it would be desirable to choose a functional form for the moment functions which may be a reasonable approximation to the true functions. It is also desirable to choose a flexible functional form because theory provides little information about the relationship between inputs and moments. As will become clear in the following section, a good approximation to the "mean" function is especially important because the residuals from the mean function play an important role in the model. A poor approximation to the mean function could introduce substantial bias into the parameter estimates of the other moment functions. Polynomial expansions appear to be likely candidates for tractable and flexible linear representations of moment functions. In addition to the Taylor series approximation, other polynomials such as the Laurent series may provide good approximations (Barnett et al., 1982).

To illustrate the specification of a flexible moment model, consider the following quadratic moment model. Letting x_k be the kth element of the input vector x, let

$$(9)\quad \mu_i = \beta_{io} + \sum_{k=1}^{n} \beta_{ik}\, x_k + \frac{1}{2} \sum_{k}^{n} \sum_{\ell}^{n} \gamma_{ik\ell}\, x_k x_\ell,\ i=1,\ldots,m$$

where the model is specified with m moments. Then

$$\frac{\partial \mu_i}{\partial x_k} = \beta_{ik} + \sum_{\ell=1}^{n} \gamma_{ik\ell}\, x_\ell .$$

The restrictions discussed in Section I can be tested using this model. For example, the restrictions in (4), that $\eta_{ik} = \frac{i}{2}\, \eta_{1k}$, are expressible as

$$(10)\quad \beta_{ik} + \sum_{\ell=1}^{n} \gamma_{ik\ell} x_\ell = \frac{i}{2} \frac{\mu_i}{\mu_2} [\beta_{2k} + \sum_{\ell=1}^{n} \gamma_{2k\ell} x_\ell].$$

With parameter estimates of the moment functions these parameter restrictions can be tested at any data point. Other hypotheses, such as the sign of the marginal effects $\partial\mu_i/\partial x_k$, can also be tested at any data point. Standard statistical procedures such as analysis of variance can be used to test for parameter differences across groups of firms. For example, to test the hypothesis that moment functions differ according to firm size, the sample can be stratified by firm size and the analysis of variance and covariance can be applied. With such tests one can determine the qualitative and quantitative structure of the moment functions.

III. SPECIFYING, ESTIMATING, AND TESTING MOMENT FUNCTIONS

A. The Linear Moment Model

In this section estimators are devised for the parameters of the output distribution moment functions. A linear moment model (LMM) is specified with the moments assumed to be linear (in the parameters) functions of the inputs or other exogenous variables. Feasible GLS estimators are derived for the LMM and are shown to be asymptotically equivalent to the true GLS estimators. Under standard statistical assumptions these estimators are consistent and asymptotically normal, and thus provide a statistical foundation for hypothesis testing.

The LMM is defined as follows. Q_j is output of the jth firm, $x_j = (x_{j1},\ldots,x_{jn})$ is the input vector with $x_{j1} = 1$, β is conformable to x_j, u_j is a random error with mean zero, and the "mean" function is

$$(11)\quad \begin{aligned} Q_j &= x_j\gamma_1 + u_j, \quad j=1,\ldots,N \\ \mu_{1j} &\equiv E(Q_j) = x_j\gamma_1. \end{aligned}$$

Equation (8) shows that higher moments of Q_j also may be functions of x_j. Define $\mu_{ij} \equiv E(u_j^i)$, $i > 2$, as the ith moment of Q_j about its mean μ_{1j}. Then let the "ith moment function" be

$$(12)\quad u_j^i = x_j\gamma_i + v_{ij}, \quad E(v_{ij}) = 0, \; i > 2, \; j=1,\ldots,N$$

so that $\mu_{ij} = x_j\gamma_i$ for all i. The LMM, represented by equations (11) and (12), contains a different parameter vector γ_i for each moment function and thus does not impose restrictions on the γ_i either within or across moments. Therefore, the LMM is indeed a more general representation of the output distribution than the models described in Section 1 and is sufficiently general for testing restrictions on the moment function parameters within and across equations.

The following assumptions will be maintained in the derivation of estimators for the LMM parameters:

(i) $E(u_j u_{j'}) = 0$ for $j \neq j'$.

(ii) the x_j are bounded, and the (Nxn) matrix X of the x_j is such that

$$\lim \frac{X'X}{N} = M_x$$

is a positive definite matrix.

(iii) letting u and v_i be the (Nx1) vectors of the u_j and v_{ij},

$$\text{plim}\,\frac{X'u}{N} = \text{plim}\,\frac{X'v_i}{N} = 0 \quad \text{for all i,}$$

and $X'u/\sqrt{N}$ and $X'v_i/\sqrt{N}$ converge in distribution to a well-defined limiting distribution.

Under the above assumptions a least squares regression of Q_j on x_j produces a consistent estimate $\hat{\gamma}_1$ of γ_1. The residuals of this regression are

$Q_j = x_j \gamma_i + u_j$

(13) $\hat{u}_j = u_j + x_j(\gamma_1 - \hat{\gamma}_1)$ → $\hat{u}_j = Q_j - X_j \hat{\gamma}$ $= (X_j \gamma_i + u_j) - X_j \hat{\gamma}_i = u_j + X_j(\gamma_1 - \hat{\gamma}_1)$

By Slutsky's theorem it follows that

(14) $\text{plim } \hat{u}_j^i = [u_j + x_j\gamma_1 - \text{plim } x_j\hat{\gamma}_1]^i = u_j^i$, for all i.

For notational consistency define y_1 as the (Nx1) vector of the Q_j, and y_i as the (Nx1) vector of the u_j^i. Also define $\hat{y}_i$, $i > 2$, as the (Nx1) vector of the $\hat{u}_j^i$, and let $\hat{v}_i = \hat{y}_i - x\gamma_i$. Regression of $\hat{y}_i$ on x gives

$$\hat{\gamma}_i = (X'X)^{-1}X'\hat{y}_i, \quad i > 2.$$

It follows from assumptions (ii) and (iii) and equation (14) that

(15) $\text{plim } \hat{\gamma}_i = \gamma_i, \quad i > 2.$

In addition, assumption (i) and equation (12) imply

(16) $E(u_j^i) = \mu_{ij} = x_j\gamma_i$ ith moment

(17) $E(v_{ij}^2) = \mu_{2i,j} - \mu_{ij}^2.$

Equations (12), (13), and (14) show that

(18) $E(u_j^2) = \text{plim } x_j\hat{\gamma}_2 = \mu_{i2}$

(19) $E(v_{ij}^2) = \text{plim } [x_j\hat{\gamma}_{2i} - (x_j\hat{\gamma}_i)^2] = \mu_{2i,j} - \mu_{ij}^2.$

These results suggest that one can obtain a consistent estimator $\hat{\gamma}_i$ of γ_i, for any i, using least squares regressions (11) and (12), with u_j^i replaced by $\hat{u}_j^i$ in (12). Equations (16) and (17) show that error terms u_j and v_{ij} are heteroscedastic and their variances are functions of the γ_i. We can construct "feasible" GLS estimators for the γ_i as follows: define Ω_1 as the (NxN) diagonal matrix of the μ_{2j} and Ω_i, $i > 2$, as the diagonal matrix of the $[\mu_{2i,j} - \mu_{ij}^2]$; also define $\hat{\Omega}_1$ and $\hat{\Omega}_i$, $i > 2$, in terms of $x_j\hat{\gamma}_2$ and $[x_j\hat{\gamma}_{2i} - (x_j\hat{\gamma}_i)^2]$. The feasible GLS estimators for the γ_i are then:

$$(20)\quad \tilde{\gamma}_1 = (X'\hat{\Omega}_1^{-1}X)^{-1} X'\hat{\Omega}_1^{-1} y_1$$

$$(21)\quad \tilde{\gamma}_i = (X'\hat{\Omega}_i^{-1}X)^{-1} X'\hat{\Omega}_i^{-1}\hat{y}_i, \quad i > 2.$$

We need to determine the statistical properties of the $\tilde{\gamma}_i$. It has been shown elsewhere for the conventional regression model that under conditions (i), (ii) and (iii), feasible GLS estimators converge in distribution to the true GLS estimators (White 1980). However, (21) differs from the usual equations for the feasible GLS estimator because $\tilde{\gamma}_i$, $i > 2$, depends on $\hat{y}_i$ rather than y_i. Therefore it is necessary to prove that the asymptotic convergence holds for (21).

<u>Theorem 1</u>. The $\tilde{\gamma}_i$ defined in equation (21) converge in distribution to

$$\bar{\gamma}_i = (X'\Omega_i^{-1}X)^{-1} X'\Omega_i^{-1} y_i, \ i > 2.$$

<u>Proof</u>: Define the ith diagonal element of Ω_i as ω_{ii} and similarily define $\hat{\omega}_{ii}$. Further define

$$(22)\quad \Delta\omega = \sup\{|\hat{\omega}_{ii}^{-1} - \omega_{ii}^{-1}|\}.$$

By Slutsky's theorem and equations (15)-(19), plim $\Delta\omega = 0$. Sufficient conditions for convergence of $\tilde{\gamma}_i$, $i > 2$, are

$$\text{plim } X'(\hat{\Omega}_i^{-1} - \Omega_i^{-1})X/N = 0$$

$$(23)\quad \text{plim } X'\hat{\Omega}_i^{-1}\hat{v}_i/\sqrt{N} = \text{plim } X'\Omega_i^{-1}v_i/\sqrt{N}, \ i > 2.$$

Using (22) and (ii) and (iii),

$$\text{plim } X'(\hat{\Omega}_i^{-1} - \Omega_i^{-1})\, X/N < \text{plim } X'X\Delta\omega/N = M_x \text{ plim } \Delta\omega = 0.$$

$$(24)\quad \text{plim } X'(\hat{\Omega}_i^{-1} - \Omega_i^{-1})\hat{v}_i/\sqrt{N} < \text{plim } \frac{X'\hat{v}_i}{\sqrt{N}} \text{ plim } \Delta\omega = 0.$$

From (14) plim $\hat{v}_i = v_i$, so

$$(25)\quad \text{plim } X'\Omega_i^{-1}\hat{v}_i/\sqrt{N} = \text{plim } X'\Omega_i^{-1} v_i/\sqrt{N}\,.$$

Together equations (24) and (25) show that (23) is satisfied and hence the theorem is proved. <u>Q.E.D.</u>

Using the results of Theorem 1 a large sample estimation algorithm for the γ_i may proceed as follows:

(a) estimate the "mean" function (11) and compute the residuals $\hat{u}_j$.

(b) estimate the regression

$$\hat{u}_j^i = x_j\gamma_i + \hat{v}_{ij}$$

for all moments deemed relevant to the analysis to obtain consistent estimates of the γ_i.

(c) compute the feasible GLS estimators (20) and (21).

Since the covariance matrices are diagonal, the GLS regressions can be computed as weighted least squares regressions with weights for (11) given by $(x_j\hat{\gamma}_2)^{-.5}$ and weights for (12) given by

$$[x_j\hat{\gamma}_{2i} - (x_j\hat{\gamma}_i)^2]^{-.5}.$$

One practical difficulty with the above estimation procedure is that the estimated variances used in the GLS regressions may be negative. This problem has arisen in the literature on the estimation of heteroscedastic regression models; in fact the model developed above can be interpreted as a generalization of heteroscedastic regression models proposed by Goldfeld and

Quandt (1972) and Amemiya (1977). Several other models have been proposed to overcome the negative variance problem (see Judge et al., 1982, Ch. 14), although it can be shown that those models are similar to the Just-Pope model and impose the restrictions on the moment functions described above (Antle 1981). To overcome the negative variance problem for the LMM I shall show that standard nonlinear programming methods provide a means of consistently estimating the parameters of the even moments under the non-negativity restriction. This approach is motivated by the fact that negative variance estimates are due to either small sample bias or sampling error in the estimates of the γ_i. The non-negativity constraint on even moments holds with probability one, so the constrained estimator has smaller bias than the unconstrained one. Since restricted estimators generally are more efficient than unrestricted estimators, the mean squared error of the inequality-constrained estimator should be less than the mean squared error of the unconstrained estimator in small samples. In large samples the consistency of the $\hat{\gamma}_i$ assures that the constraint holds and the problem disappears. Hence, use of inequality-constrained estimators for the γ_i can improve small sample properties of the estimates and does not affect the large sample consistency properties.

The inequality-constrained estimator for γ_2 is obtained by choosing γ_2 to solve

$$\min_{\gamma_2} \sum_{j=1}^{N} [\hat{u}_j^2 - x_j\gamma_2]^2 \quad \text{subject to} \quad x_j\gamma_2 \geq 0.$$

To estimate the γ_{2i} for $i \geq 2$, choose γ_{2i} to solve

$$\min_{\gamma_{2i}} \sum_{j=1}^{N} [\hat{u}_j^{2i} - x_j\gamma_{2i}]^2 \quad \text{subject to} \quad [x_j\gamma_{2i} - (x_j\hat{\gamma}_i)^2] \geq 0,$$

where $\hat{\gamma}_i$ has been obtained from a previous regression. The latter inequality constraint simultaneously forces γ_{2i} to satisfy the requirement that $\mu_{2i,j} > 0$ as well as the requirement that the variance of u_j^i be non-negative. Such inequality-constrained minimization problems can be solved with software such as the MINOS program (see Murtaugh and Saunders, 1977, 1978).

B. Hypothesis Testing and the Joint GLS Estimator

Testing hypotheses on the parameters of the LMM requires knowledge of the estimators' distributions. Under the conditions of the Lindberg-Feller Theorem (Gnedenko and Kolmogorov, 1954), the GLS estimators $\tilde{\gamma}_i$ defined in (20) and (21) have a normal limiting distribution with a covariance matrix $(X'\hat{\Omega}_i^{-1}X)^{-1}$. Therefore, in large samples, tests of hypotheses on individual parameters can be based on single-equation estimates using this limiting distribution. However, single equation GLS estimates are generally less efficient than joint GLS estimates because the error terms are correlated across equations. These cross-equation correlations must be taken into account when calculating test statistics for cross-equation parameter restrictions such as those discussed in section 1.[7] The structure of the cross-equation covariances is seen as follows: for the first and ith moments the error terms u_j and v_{ij} give

$$E(u_j v_{ij}) = E(u^{i+1}{}_j - u_j\mu_{ij}) = \mu_{i+1,j}$$

and for the ith and kth moments

$$E(v_{ij}v_{kj}) = E(u_j^i - \mu_{ij})(u_j^k - \mu_{kj}) = \mu_{i+k,j} - \mu_{ij}\mu_{kj},\ i,\ k > 1.$$

By the assumption of independence of the u_j,

$$E(u_j\ v_{ij}{}') = 0\ ,\ j\neq j'$$

$$E(v_{ij}\ v_{\ell j}{}') = 0\ ,\ j\neq j'.$$

To illustrate the joint GLS estimator consider a three moment model, $y = W\delta + \varepsilon$, where

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix},\ W = \begin{bmatrix} X_1 & o & o \\ o & X_2 & o \\ o & o & X_3 \end{bmatrix},\ \delta = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix},\ \varepsilon = \begin{bmatrix} u \\ v_2 \\ v_3 \end{bmatrix},$$

$$\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} & \Omega_{13} \\ \Omega_{21} & \Omega_{22} & \Omega_{23} \\ \Omega_{31} & \Omega_{32} & \Omega_{33} \end{bmatrix},$$

where Ω_{ii}, i=2,3, is the diagonal matrix of $\mu_{2i,j} - \mu_{ij}^2$; $\Omega_{i1} = \Omega_{1i}$, i=1,2,3 is the diagonal matrix of $\mu_{1+i,j}$; and $\Omega_{32} = \Omega_{23}$ is the diagonal matrix of $\mu_{5j} - \mu_{2j}\,\mu_{3j}$. The feasible joint GLS estimator of δ is

$$(26)\quad \tilde{\delta} = (W'\,\hat{\Omega}^{-1}W)^{-1}\,W'\,\hat{\Omega}^{-1}\,\hat{y}$$

where $\hat{\Omega}$ and $\hat{y}$ are as defined analogously to Ω and y. We now prove the following generalization of Theorem 1.

<u>Theorem 2</u>: $\tilde{\delta}$ <u>defined in</u> (26) <u>converges in distribution to</u>

$$(27)\qquad \bar{\delta} = (W'\,\Omega^{-1}\,W)^{-1}\,W'\,\Omega^{-1}\,y.$$

<u>Proof</u>: To utilize the results of Theorem 1 we need to show that Ω^{-1} satisfies Slutsky's Theorem, i.e., that the elements of Ω^{-1} can be defined as continuous functions of the elements of Ω. This fact can be shown by noting that the Ω_{ii} are diagonal matrices and by applying the partitioned matrix inverse rule to Ω. It follows that

$$\Omega^{-1} = \begin{bmatrix} \Omega^{11} & \Omega^{12} & \Omega^{13} \\ \Omega^{12} & \Omega^{22} & \Omega^{23} \\ \Omega^{13} & \Omega^{23} & \Omega^{33} \end{bmatrix}$$

where the Ω^{ij} are diagonal matrices whose diagonal elements are continuous functions of the elements of Ω and independent of N, the sample size. Therefore, for all N we can define the elements ω^{ij} of Ω^{-1} and claim that plim $\hat{\omega}^{ij} = \omega^{ij}$ where $\hat{\omega}^{ij}$ is the (i,j) element of $\hat{\Omega}^{-1}$. Proceeding along the lines of Theorem 1, define

$$\Delta\omega = \sup\,\{|\hat{\omega}^{ij} - \omega^{ij}|\}$$

for which plim $\Delta\omega = 0$. Now define R as the 3N x 3N matrix obtained by replacing the Ω_{ij} in Ω with identity matrices. Then, as in Theorem 1, it follows that

$$\text{plim } W'(\hat{\Omega}^{-1} - \Omega^{-1})W/N \leq \text{plim } \frac{W'RW}{N} \Delta\omega = 0$$

$$\text{plim } W'(\hat{\Omega}^{-1} - \Omega^{-1})\hat{\varepsilon}/\sqrt{N} \leq \text{plim } \frac{W'R\hat{\varepsilon}}{\sqrt{N}} \Delta\omega = 0$$

which is sufficient to show $\tilde{\delta}$ converges in distribution to $\bar{\delta}$. Q.E.D.

To compute the feasible joint GLS estimator defined in (27), $\hat{\Omega}$ must be a positive definite matrix. The consistent moment estimates, however, do not necessarily satisfy this requirement. Using the methodology discussed above for imposing inequality constraints it is possible to restrict the consistent moment estimates so that $\hat{\Omega}$ is positive definite. For example, the necessary and sufficient conditions for Ω defined above to be positive definite are $|C_1| > 0$, $|C_2| > 0$, $|C_3| > 0$, where C_i is the i^{th} principal minor of Ω. These conditions amount to inequality restrictions on the moments. Since $|C_1| = |\Omega_{11}|$ and $|C_2| = (\Omega_{11}\Omega_{22}) - \Omega_{21}^2$, the first two conditions can be imposed using the linear inequality constraints $\Omega_{11} > 0$ and $\Omega_{22} > \hat{\Omega}_{21}^2/\hat{\Omega}_{11}$. However, for C_i, $i > 2$, the constraints become nonlinear and require that several moment functions be jointly estimated. Such estimation procedures are possible using the MINOS/AUGMENTED nonlinear optimization program but are likely to be very costly. Therefore, joint GLS estimation may be extremely costly when more than two moment functions are jointly estimated. This means that it is advisable to test cross-moment parameter restrictions with pairs of jointly estimated moment functions as a first step in an analysis. Whether more than pair-wise estimation is attempted should depend on the need for improved efficiency relative to the increased estimation cost.

IV. AN APPLICATION TO MILK PRODUCTION

I now apply the flexible moment approach to milk production data. Milk production is an attractive process to study because it is a true single-product process. In addition, inputs are chosen prior to the production period and are therefore exogenous to output. The model is not subject to simultaneous equation bias as would be the case if inputs were chosen sequentially (Antle 1983).

The monthly data represent nine Tulare County, California, dairies over a 30 month period. These high quality data were obtained from a computerized data collection and processing system. For a detailed discription of the data, see Goodger (1981). The results reported here are based on an in-depth analysis of milk production by Antle and Goodger (1982).

I use the quadratic form of the LMM given in equation (9). The data are monthly time series so it is necessary to account for autocorrelation. Because the nine dairies are subject to similar weather and climatic shocks, and only 30 (or fewer) observations are available per dairy, the same first-order autocorrelation coefficient is assumed for all dairies. The mean function is assumed to be

$$Q_{jt} = \mu_{1jt} + u_{jt}, \; j=1, \ldots, 9, \; t=1, \ldots, 30,$$

where

$$u_{jt} = \rho u_{jt-1} + \varepsilon_{jt}, \; |\rho|<1,$$

and

$$E(\varepsilon_{jt}^{k}\varepsilon_{j't'}^{\ell}) = 0, \; j\neq j', \; t\neq t'$$

$$= \mu_{k+\ell,jt}, \; j=j', \; t=t'.$$

Therefore, by applying the transformation

$$Q_{jt} - \rho Q_{jt-1} = \mu_{1jt} - \rho\mu_{1jt-1} + \varepsilon_{jt}$$

to the mean function, the higher moment functions can be estimated by hypothesizing that the moments of ε_{jt} are functions of inputs.

Based on this model, the following estimation procedure was used: First the mean function was estimated to obtain a consistent estimate of ρ. The data were transformed with ρ and the mean equation was re-estimated to obtain residuals $\hat{\varepsilon}_{jt}$. Second, the $\hat{\varepsilon}_{jt}$ were used in inequality-constrained regressions to consistently estimate the parameters of μ_{2jt}, μ_{3jt}, μ_{4jt}, μ_{5jt}, and μ_{6jt}. Third, these estimated moments were used to compute the feasible GLS estimators for μ_{1jt}, μ_{2jt}, μ_{3jt}. To test the cross-moment restrictions given in equations (2) and (4), the parameters of μ_{2jt} and μ_{3jt} were jointly estimated.[8]

The inputs in the model are feed; animal capital measured as herd size adjusted for breed, age, and health; physical capital measured as milking capacity of the dairy; and management measured by an index computed from a survey of the dairy managers. Environmental and herd health variables were also included in the model. The mean production elasticities for the first three moments with respect to each input are presented in Table 1 with their standard errors. The χ^2 statistic (as in Theil, 1971, Ch. 8) for the null hypothesis of zero slope coefficients of each moment function are also presented in Table 1. These statistics show that all three moments are clearly significant functions of inputs. The fundamental hypothesis of the moment-based approach, that moments are functions of inputs, is supported by the data.

The elasticities in Table 1 suggest that the restrictions implied by the multiplicative error model, given in equation (2), are rejected, since the elasticities of higher moments are not all positive. The negative elasticities of μ_2 with respect to physical capital and μ_3 with respect to feed show that the multiplicative error model is a very poor approximation to the output distribution. The elasticities in Table 1 also indicate that the restrictions implied by the Just-Pope model (equation 4) are rejected, because the elasticities of the third moment are not equal to 1.5 times the elasticities of the variance. Indeed, except for the animal capital variable, the elasticities of μ_2 have the opposite sign from the elasticities of μ_3. Thus it would clearly be inappropriate to constrain all the elasticities of μ_2 and μ_3 to have the same sign.

A formal test of the restrictions of the multiplicative error model and the Just-Pope model is obtained by noting that both (2) and (4) imply $\eta_{3k} = 3\eta_{2k}/2$. Using equation (10) these restrictions can be expressed as:

$$\beta_{3k} = (3\mu_{3jt}/2\mu_{2jt})\,\beta_{2k}$$
(28)
$$\gamma_{3k\ell} = (3\mu_{3jt}/2\mu_{2jt})\gamma_{2k\ell}$$

for all k and ℓ. These restrictions were tested at the sample means of the moments by computing

$$\chi^2/R = \frac{SSE_r - SSE_u/R}{SSE_u/DF}$$

where SSE_r is the sum of squared residuals under the restriction, SSE_u is the sum of squared residuals from the unrestricted model, R is the number of restrictions, and DF is the degrees of freedom. Asymptotically, χ^2 has the Chi-square distribution with R degrees of freedom. Under restrictions (28) we obtain $\chi^2(18) = 76.84$ which far exceeds the 5 percent critical value of 28.26. Thus, as examination of the moment elasticities suggests, both the multiplicative error model and the Just-Pope model are rejected by the data in favor of a more flexible specification.

IV. CONCLUSIONS AND FUTURE RESEARCH

The aim of this paper is to develop a flexible representation of a firm's stochastic production process which can be estimated and subjected to statistical test. This study is motivated by the fact that conventional production function specifications impose restrictions on the probability distribution of output which cannot be tested with the conventional models. These restrictions have important implications for the behavior of the firm. Because output distributions are unique functions of their moments the stochastic structure of any production process can be inferred by measuring its moments. A linear moment model (LMM) is developed which is sufficiently flexible to test restrictions within and across moments such as those implied by the conventional production function models discussed in section I. A quadratic representation of the moment functions is suggested in section II as a flexible linear-in-parameters moment model.

The LMM restricts moment functions to be linear in the parameters. Although polynomial functions such as the quadratic can provide flexible representations of moment functions, statistical problems may arise because of a large number of parameters. Using Malinvaud's (1970) results on consistency of nonlinear least squares regressions, it would appear that the statistical results obtained in this paper for linear moment functions could be extended to the case of nonlinear functions, thus permitting models to be specified with fewer parameters.

Due to previous methodological limitations, little is currently known about the stochastic structure of production. The approach developed in this paper opens the way for testing the characteristics of the stochastic structure of production processes which are known to have important behavioral implications but heretofore could not be rigorously tested. The estimates presented here showed that, in the case of milk production, the multiplicative error model and the Just-Pope model are not sufficiently flexible to represent the output distribution. Therefore, inferences of firm behavior under uncertainty based on those models could be misleading. We now need further measurement and testing of the stochastic structure of production processes to discover what empirical regularities exist and to further explore their implications for our understanding of firm behavior.

cg 1/6/83 P32

Footnotes

1. The reader should note that most recent specifications of frontier production functions (see Forsund, Lovell, and Schmidt 1980) have multiplicative errors and therefore impose restrictions of the form (2) across moments.

2. When u is normal it follows that

$$\mu_{2k} = \frac{\mu_2^{2k}\,(2k!)}{2^k\,k!}.$$

3. Higher moments have been introduced in the finance literature in numerous studies. For a survey of the literature and analysis of a 3 moment portfolio model see Francis and Archer, 1979, Ch. 16.

4. The solution of the "Stieltjes moment problem" shows that a sufficient condition for a set of moments to define a unique distribution is that the range of the random variable is finite. This condition clearly holds for output, which also shows that all moments of output exist.

5. An alternative derivation of these moment functions can be based on the general theory of regression (Kendall and Stuart, 1979, Ch. 28). Assume the joint distribution function of Q and x is f(Q,x); then for any value x_o of x we can define

$$\mu_1(x_o) = \int Qf(Q,x_o)dQ/\int f(Q,x_o)dQ$$

$$\mu_i(x_o) = \int [Q-\mu_1(x_o)]^i\, f(Q,x_o)dQ/\int f(Q,x_o)dQ.$$

and thus obtain the moments as functions of x as in equation (5). The inputs can be treated as predetermined variables in the model under the following assumptions: (a) the firm chooses inputs to maximize the mathematical expectation of a function of output in a single period, or

(b) the firm chooses inputs over time using an open-loop control algorithm. However, if inputs are chosen sequentially and there is information feedback, inputs become endogenous variables in the model. See Antle (1983) for a detailed analysis of this issue.

6. Day's (1965) pioneering study of yield distributions in agriculture used the Beta distribution and the method of moments to show how the shape of yield distributions changes with the level of nitrogen fertilizer inputs.

7. Note that production models based on cross-section data usually have the same regressors in each equation. However, there is still an efficiency gain from the joint GLS estimator relative to single-equation GLS because the covariance matrix is different from the seemingly unrelated regression case.

8. As noted above, joint GLS estimation poses the problem of obtaining a positive definite covariance matrix for the equation system. It is sufficient to jointly estimate the second and third moment functions to test the restrictions implied by the multiplicative error model and the Just-Pope model.

Table 1

Mean Moment Elasticities Based on GLS Estimates of Quadratic Moment Functions

Elasticity with respect to:	Moment		
	First	Second	Third
Feed	.059 (.027)	.448 (.779)	-33.276 (12.238)
Physical Capital	.678 (.311)	-4.051 (1.580)	153.102 (44.584)
Animal Capital	.877 (.065)	1.057 (.992)	8.414 (16.727)
Management	.230 (.291)	-1.557 (2.241)	75.269 (32.726)
$\chi^2(18)$	7130	855.5	67.68

Note: Standard errors in parentheses. Moment elasticities computed at sample means of the data.

References

Amemiya, T. "A Note on a Heteroscedastic Model." Journal of Econometrics 6(1977): 365-370.

Anderson, J. R. "Sparse Data, Climatic Variability, and Yield Uncertainty in Response Analysis." American Journal of Agricultural Economics 56(1973): 77-83.

Antle, J. M. "Estimation of Heteroscedastic Regression Models Whose Moments are Functions of Exogenous Variables." Department of Agricultural Economics, University of California, Davis, Working Paper No. 81-5, 1981.

__________. "Sequential Decision Making in Production Models." American Journal of Agricultural Economics, forthcoming, May 1983.

__________, and W. J. Goodger. "Measuring Stochastic Technology: The Case of Tulare Milk Production." Department of Agricultural Economics, University of California, Davis, Working Paper, 1982.

Barnett, W. A., P. A. Spindt, and E. K. Offenbacher. "The Empirical Properties of Divisia Monetary Indexes." Paper presented at A.E.A. meetings, Washington, D.C., December 1981.

Day, R. H. "Probability Distributions of Field Crops." Journal of Farm Economics 47(August 1965): 713-741.

Diewert, W. E. "Applications of Duality Theory." In Frontiers of Quantitative Economics, Vol. III, edited by M. D. Intrilligator and D. A. Kendricks, North-Holland, 1974.

Forsund, F. R., C.A.K. Lovell, and P. Schmidt. "A Survey of Frontier Production Functions and of Their Relationship to Efficiency Measurement." Journal of Econometrics 13(1980): 5-25.

Francis, J. C. and S. H. Archer. Portfolio Analysis. Prentice Hall, Inc. Englewood Cliffs, 1979.

Fuss, M., D. McFadden and Y. Mundlak. "A Survey of Functional Forms in the Economic Analysis of Production." In Production Economics: A Dual Approach to Theory and Applications, edited by M. Fuss and D. McFadden. Amsterdam: North-Holland, 1978.

Gnedenko, B. V. and A. N. Kolmogorov. Limit Distributions for Sums of Independent Random Variables. Addison-Wesley, 1954.

Goldfeld, S. M. and R. E. Quandt. Nonlinear Methods in Econometrics. Amsterdam: North-Holland, 1972.

Goodger, W. A. "The Description, Measurement, and Economic Assessment of the Contribution of Veterinary Services to Large Scale Dairy Operations in California." Ph.D. dissertation, University of California, Davis, 1981.

Harvey, A. C. "Estimating Regression Models with Multiplicative Heteroscedasticity." Econometrica 44(1976): 461-465.

Judge, G. G., R. C. Hill, W. E. Griffiths, H. Lutkepohl, and T. C. Lee. Introduction to the Theory and Practice of Econometrics. New York: Wiley, 1982.

Just, R. E. and R. D. Pope. "Stochastic Specification of Production Functions and Economic Implications." Journal of Econometrics 7(1978): 67-86.

__________________________. "Production Function Estimation and Related Risk Considerations." American Journal of Agricultural Economics 61(May 1979): 277-284.

Kendall, M. and A. Stuart. The Advanced Theory of Statistics, Volume 1, 4th Edition. New York: Macmillan, 1976; Volume 2, 4th Edition, 1978.

Malinvand, E. "The Consistency of Nonlinear Regressions." Annals of Mathematical Statistics 41(1970): 956-969.

Murtaugh, B. A. and M. A. Saunders. "Large-Scale Linearly Constrained Optimization." Mathematical Programming 14(1978): 41-72.

________________. MINOS User's Guide. Technical Report 77-9 (February 1977), Department of Operations Research, Stanford University, 1977.

Newberry, D.M.G. and J. E. Stiglitz. "The Choice of Techniques and the Optimality of Market Equilibrium with Rational Expectations." Journal of Political Economy 90(April 1982): 223-246.

Nikiphoroff, B. "An Economic Analysis of Resource Allocation in Small Farms Under Uncertainty: A Case Study of Paraguay." Ph.D. dissertation, University of California, Davis, 1981.

Pope, R. D. and R. A. Kramer. "Production Uncertainty and Factor Demands for the Competitive Firm." Southern Economic Journal 46(October 1979): 489-501.

Rao, C. R. Linear Statistical Inference and Its Application. New York: Wiley, 1973.

Roumasset, J. A. Rice and Risk: Decision Making Among Low Income Farmers. Amsterdam: North-Holland, 1976.

Theil, H. Principles of Econometrics. New York: John Wiley and Sons, 1971.

White, H. "A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity." Econometrica 48(1980): 817-838.