Journal of Econometrics 7 (1978) 67–86. © North-Holland Publishing Company

STOCHASTIC SPECIFICATION OF PRODUCTION FUNCTIONS AND ECONOMIC IMPLICATIONS*

Richard E. JUST and Rulon D. POPE

University of California, Berkeley, CA 94720, USA

Received May 1976, final version received April 1977

The stochastic specification of input–output response is examined. Postulates are set forth which seem reasonable on the basis of *a priori* theorizing and observed behavior. It is found that commonly used formulations are restrictive and may lead to inefficient and biased results. A function which satisfies the postulates is suggested. Two- and four-stage procedures for estimation of the resulting function are then outlined. The estimators are shown to be consistent and, in the latter case, asymptotically efficient under normality.

1. Introduction

The intent of this paper is to explore the implications of risk for both theoretical and econometric production studies and to explore econometric possibilities for a production function specification with reasonable risk implications. Accordingly, this paper considers appropriate production function formulations under risk. It becomes apparent that popular formulations of stochastic production functions are very restrictive for many cases in which risk is important. For example, consider inputs like frost protection or, possibly, pesticide use which are generally believed to have a risk-reducing effect on output at least in some ranges of use. If one investigates the effects of these inputs using the familiar Cobb–Douglas production function with log-linear disturbances, then one incorrectly imposes a risk-increasing effect on output. After introducing several other reasonable risk considerations, other common stochastic production functions also appear inadequate. It is argued in this paper that a useful production function should possess sufficient flexibility so that the effects of inputs on the deterministic component of production is different than on the stochastic component. Such an alternative is suggested, and the related econometric possibilities are investigated. The econometric implications are also of apparent importance for production studies in which risk is not a primary interest.

*Giannini Foundation Paper no. 488.

2. Needed flexibility in empirical problems

The literature abounds with efforts to generalize production functions on the basis of *a priori* theorizing about technical production response, empirically observed behavior, and, perhaps, suitability and convenience for estimation and empirical use.[1] In both theoretical and empirical work with these functions, however, the stochastic component of production has usually been incorporated in one of three ways,[2]

$$y = F_1(X) \equiv f(X)\,\mathrm{e}^{\varepsilon}, \qquad E(\varepsilon) = 0, \tag{1}$$

$$y = F_2(X) \equiv f(X)\,\varepsilon, \qquad E(\varepsilon) = 1, \tag{2}$$

$$y = F_3(X) \equiv f(X) + \varepsilon, \qquad E(\varepsilon) = 0, \tag{3}$$

where y is production, $f(\cdot)$ is the deterministic component of production (possibly following any of the common forms), F_j is the stochastic version of the production function, X is a vector of inputs, and ε is a stochastic disturbance.

Typically, these specifications have been made without regard for reasonable risk considerations. Analogous to the way in which generalizations were introduced in the deterministic specification of production functions, this section attempts to incorporate risk considerations of similar economic relevance in order to improve specification of stochastic components of production. Results demonstrate that estimates based on common specifications are uninformative with respect to risk considerations. Accordingly, any ensuing evaluation of policy which affects risk or inputs related to risk may be inadequate.

Consider the following postulates which seem reasonable and, perhaps, necessary to reflect stochastic, technical input–output relationships. For simplicity, the discussion of risk will be confined to arguments relating to variance rather than higher moments.[3] Also, many of the arguments will be couched in an agricultural context because the wide range of input characteristics in that industry easily facilitates the discussion.

[1]For example, the Cobb–Douglas production function was developed as an improvement over the polynomial functional form to allow homogeneity and diminishing marginal productivity with constant returns to scale [Cobb–Douglas (1928)], and the CES production function was subsequently developed to represent production with a constant elasticity of substitution different from unity [Arrow et al. (1961)]. Additional generalizations have since been incorporated in the transcendental function and the transcendental logarithmic function to represent other reasonable economic relationships [Halter, Carter, and Hocking (1957); Christensen, Jorgenson, and Lau (1973)].

[2]For convenience in notation, upper-case letters are used to denote the random production function and lower-case letters are used for the deterministic component of production.

[3]Of course, risk should be more broadly defined for distributions other than the normal [Hanoch and Levy (1969)], but most empirical production studies assume normality anyway.

Postulate 1. Positive production expectations $[E(y) > 0]$.

This is a rather trivial requirement and requires no explanation.

Postulate 2. Positive marginal product expectations $[\partial E(y)/\partial X_i > 0]$.

Consistent with deterministic production theory, it seems reasonable to consider only factors which make a positive marginal contribution to expected output in the region of use.[4]

Postulate 3. Diminishing marginal product expectations $[\partial^2 E(y)/\partial X_i^2 < 0]$.

This criterion corresponds to the usual concavity conditions in deterministic production theory. Arguments analogous to those made in the deterministic case seem to equally support the stochastic case. That is, if any firm in an industry is risk-neutral, then that firm would be induced to grow infinitely large when Postulate 3 is not satisfied [assuming Postulate 1 and Postulate 2 are satisfied]; this possibility seems implausible.

Postulate 4. A change in variance for random components in production should not necessarily imply a change in expected output when all production factors are held fixed $[\partial E(y)/\partial V(\varepsilon) = 0$ possible$]$.

Remote situations may exist in which it makes sense for variance to behave like a factor of production in output expectations; but the more reasonable case would be where it does not.

Postulate 5. Increasing, decreasing, or constant marginal risk should all be possibilities $[\partial V(y)/\partial X_i \gtreqless 0$ *possible where* $V(y) = E[y-E(y)]^2]$.

Based on real world observations and *a priori* reasoning, it seems that all three possibilities can exist depending on the input considered. For example, agricultural inputs such as land, fertilizer, and chemical thinning practices seem to make a positive contribution to variance of production in some cases. On the other hand, pesticides, irrigation, frost protection, disease-resistant seed varieties, and overcapitalization all possibly have a negative effect on the variance of production attributable to weather, insects, and crop diseases [Carlson (1969); Just and Pope (1977)].[5] Particularly in agriculture, increased machinery availability can help to expedite a harvest and shorten the period of vulnerability to adverse weather conditions.

[4]Although it is conceivable that some input may be used at a point where its expected marginal product is negative if it also were leading to substantial reduction in variance, no such examples seem apparent.

[5]Although there are perhaps regions of both increasing and decreasing marginal risk for each of these inputs, it seems that there is some region where an activity, such as use of smudge pots for frost protection, must have a negative contribution to variance.

Postulate 6. A change in risk should not necessarily lead to a change in factor use for a risk-neutral (profit-maximizing) producer [$\partial X_i^*/\partial V(\varepsilon) = 0$ *possible where* X_i^* *is the optimal input level*].

It is commonly held that a change in risk does not affect the behavior of a risk-indifferent producer [Arrow (1971)]. Some cases have been developed in which this is not necessarily true [Walters (1960); Just (1975)], but it appears desirable that the choice of production function should not exclude the possibility of Postulate 6 *a priori.*

Postulate 7. The change in the variance of marginal product with respect to a factor change should not be constrained in sign a priori *without regard to the nature of the input* [$\partial V(\partial y/\partial X_i)/\partial X_j \gtreqless 0$ *possible*].

Here, perhaps, intuition is somewhat less clear than in previous cases. It seems, however, that some cases exist where the marginal product variability of, say, land increases when other inputs are held fixed. For example, as one farms on a more extensive basis with the same inputs, his 'fire fighting' ability is reduced. That is, one is more subject to adverse weather conditions, etc., during critical operations such as harvesting and planting. On the other hand, an input such as smudge pots in frost protection is likely to have decreasing marginal product variability since survival of the crop becomes more probable, and marginal product presumably tends in probability toward zero with increased input use (in the relevant range). Similar arguments also make it seem unreasonable to constrain complementarity in marginal product variability among inputs without regard for the nature of these inputs.

Postulate 8. Constant stochastic returns to scale should be possible [$F(\theta X) = \theta F(X)$ *for scalar* θ].

For example, if one produces with, say, two production units rather than one where exactly the same inputs (including quality, management, etc.) are used with each, then exactly twice as much output results. This concept has been employed by both Arrow (1971) and Stiglitz (1974); and although it may not necessarily hold in all cases, it certainly appears that this possibility exists.

3. Implications of common production function specifications

On the basis of the above eight postulates, consider the popular specification in (1) which includes the usual stochastic formulation for Cobb–Douglas, CES, transcendental, and translog production functions. It can be verified that Postulates 1, 2, 3, and 8 may hold. However, in the case of Postulate 5, one can easily verify that

$$E(y) = f(X)E(\mathrm{e}^{\varepsilon}), \qquad V(y) = f^2(X)V(\mathrm{e}^{\varepsilon}),$$

$$\frac{\partial E(y)}{\partial X_i} = f_i E(\mathrm{e}^{\varepsilon}), \qquad \frac{\partial V(y)}{\partial X_i} = 2ff_i V(\mathrm{e}^{\varepsilon}).$$

Since $E(\mathrm{e}^{\varepsilon})$ must be positive, the imposition of Postulate 2 implies that $f_i > 0$. Additionally, using Postulate 1 requires that $f > 0$; hence, $\partial V(y)/\partial X_i \geqq 0$ since $V(\mathrm{e}^{\varepsilon})$ must always be non-negative. That is, the marginal risk effect must always be positive with the various popular production function specifications included in (1) when expected production and marginal products are positive. It thus appears that the risk effects of such (possibly) risk-reducing inputs associated with overcapitalization and use of pesticides, irrigation, and frost protection equipment in agriculture cannot be properly reflected.

Consider also Postulate 4. Here it is readily apparent that a change in $V(\varepsilon)$, *ceteris paribus*, causes a change in

$$E(y) = f(X)E(\mathrm{e}^{\varepsilon}),$$

for any distribution of ε satisfying (1); furthermore, $E(y)$ may also depend on higher moments if variance is fixed.[6] Indeed, in the case of normality, one finds that $E(\mathrm{e}^{\varepsilon}) = \mathrm{e}^{\sigma/2}$ where $\varepsilon \sim N(0, \sigma)$ [Zellner, Kmenta, and Dreze (1966)] and, hence, that

$$\frac{\partial E(y)}{\partial V(\varepsilon)} = \frac{1}{2\sigma} f(X)\,\mathrm{e}^{\sigma/2} > 0,$$

if $V(\varepsilon) = \sigma$ when Postulate 1 is also imposed.

Turning to Postulate 6, it is apparent that the above property of (1) leads to another *a priori* limitation. For simplicity, consider only the case with log-normality since a counterexample is sufficient to refute Postulate 6. With prices fixed and unaffected by firm decisions, first-order conditions for expected profit

[6]Note that, when $E(\varepsilon) = 0$, one obtains

$$E(\mathrm{e}^{\varepsilon}) = E\left[1+\frac{\varepsilon}{1!}+\frac{\varepsilon^2}{2!}+\ldots+\frac{\varepsilon^n}{n!}+\ldots\right]$$
$$= 1+\frac{\sigma}{2}+\sum_{k=3}^{\infty}\frac{m_k}{k!},$$

where m_k represents the kth moment about zero (the mean) for the ε distribution. It is thus apparent that a change in variance or any higher moment, *ceteris paribus*, leads to a change in $E(\mathrm{e}^{\varepsilon})$ and, hence, $E(y)$.

maximization with (1) are

$$f'(X) = \mathrm{e}^{-\sigma/2}\,\frac{\gamma}{P},$$

where f' is a vector of first derivatives f_i, P is output price, and γ is a vector of input prices. In this case one finds that profit maximization implies

$$\frac{\partial X^*}{\partial \sigma} = -\tfrac{1}{2}\,\mathrm{e}^{-\sigma/2} G^{-1}\,\frac{\gamma}{P} > 0, \tag{4}$$

assuming concavity of f and complementarity [Bear (1965)] where X^* is the optimal input level and G is the appropriate Hessian matrix.[7] Hence, Postulate 6 is also not satisfied.

Finally, consider Postulate 7. From (1), it can be determined that

$$\frac{\partial V(\partial y/\partial X_i)}{\partial X_j} = 2 f_i f_{ij} V(\mathrm{e}^{\varepsilon}).$$

But, if Postulate 3 holds, then

$$\frac{\partial^2 E(y)}{\partial X_i^2} = f_{ii} E(\mathrm{e}^{\varepsilon}) < 0,$$

which implies that $f_{ii} < 0$ so that $\partial V(\partial y/\partial X_i)/\partial X_i$ is always constrained to be negative with the formulation in (1). With complementarity of $f(X)$ or with all positive production elasticities in, say, a Cobb–Douglas production function, it can also be shown that $f_{ij} > 0$, $i \neq j$, so that complementarity is also arbitrarily imposed on all marginal product variability relationships.

The practical implications of using the popular stochastic formulation in (1) are several. Suppose one is interested in evaluating a policy that affects use of some input. If an input is one that makes a negative contribution to risk, then the change is expected utility will be understated for both policies which encourage or discourage use of that input (for a risk-averse producer). For example, if one evaluates a policy such as controlling, say, pesticide use and risk is thereby increased, then the true utility loss for a risk-averse producer will be greater than when the risk effect is incorrectly estimated as a reduction. Similar incorrect conclusions can also result in evaluating other policies when affected

[7]Note that the Cobb–Douglas function satisfies complementarity when partial production elasticities are positive.

inputs increase risk because of the lack of flexibility in the popular specification in (1).

Further, considering the implications of (1) relating to Postulate 6, it appears that serious reservations must be attached to both policy evaluation and forecasting with the relationship in (1). Suppose, for example, that $\varepsilon \sim N(0, \sigma)$ so that $E(e^{\varepsilon}) = e^{\sigma/2}$ and consider the effects of some new production-stabilizing technology (possibly induced by policy actions). Using the production function in (1) leads to the rather unintuitive result in (4) implying that input use and, hence, expected output are reduced in response to the reduction in risk. Normally, however, one would expect either no response or an increase in production (particularly in the risk-averse case) [Sandmo (1971)].[8]

Turning to other functional forms, it seems that problems also exist with most of the other simple stochastic specifications in common use. These investigations are made in the appendix, but the results are presented in table 1. All of the simple specifications in (1), (2), and (3) fail to satisfy at least two postulates; but the specification in (1), which is in most common use, fails to satisfy more of the postulates than any of the others which were investigated.

4. An alternative stochastic specification

Since production specifications in common use fail to satisfy many of the risk postulates, an investigation of several more general specifications is appropriate. In formulating such generalizations, it is apparently necessary to introduce some function of the inputs, $h(X)$, which perturbs the effects of the disturbance in such a way that relationships of inputs with risk are not determined solely by the relationships of inputs with expected output. Two simple classes of production functions of this type are possible – those with the perturbed disturbance $h(X)\varepsilon$ appearing multiplicatively and those where $h(X)\varepsilon$ appears additively.[9] The cases where $h(X)\varepsilon$ appears multiplicatively, however, also fail to satisfy several of the risk postulates (see table 1 and the appendix). In the general additive case,

$$y = F_4(X) = f(X)+h(X)\varepsilon, \qquad E(\varepsilon) = 0, \quad V(\varepsilon) = \sigma; \tag{5}$$

however, it can easily be verified that all of the Postulates 1 through 8 are satisfied (again, see table 1 and the appendix).

Of course, there are more general specifications which may also satisfy

[8]There are also significant econometric implications which must be considered when (1) is not appropriate, but these will be discussed later in the paper.

[9]Of course, some more general specifications are also possible, but a discussion of the associated problems will be delayed until later in the paper.

Table 1

Evaluation of stochastic production function specifications in terms of Postulates 1 through 8.[a]

Specification		Postulates 1	2	3	4	5	6	7	8
$y = f(X)e^{\varepsilon}$,	$E(\varepsilon) = 0$	x	x	x					x
$y = f(X)\varepsilon$,	$E(\varepsilon) = 1$	x	x	x	x		x		x
$y = f(X)+\varepsilon$,	$E(\varepsilon) = 0$	x	x	x	x		x		
$y = f(X)g(\varepsilon)$		x	x	x	x		x		x
$y = f(X)h(X)\varepsilon$		x	x	x	x		x		x
$y = f(X)e^{h(X)\varepsilon}$		?	?	?		?	?	?	
$y = f(X)+h(X)\varepsilon$,	$E(\varepsilon) = 0$	x	x	x	x	x	x	x	x
$y = f(X)h(X, \varepsilon)$		x	x	x	x	x	x	x	x
$y = f(X, \varepsilon)$		x	x	x	x	x	x	x	x

[a]An 'x' implies that the respective postulate can be satisfied, while a '?' indicates a postulate for which rejection is not definite but the range of flexibility appears overly restrictive (see appendix).

Postulates 1 through 8 – for example, special cases of the general function

$$y = f(X, \varepsilon), \tag{6}$$

or

$$y = f(X)h(X, \varepsilon), \tag{7}$$

where ε possibly becomes a vector of random disturbances [note that eq. (5) is a special case of eqs. (6) and (7)].[10] But such a function is too general to investigate risk properties and to provide useful information for empirical work.

5. Econometric procedures in production function estimation

Since the representations in (6) and (7) are too general to suggest useful econometric procedures, it appears that a reasonable choice of production function might be any special case of (5) for which f satisfies the usual requirements for a production function and h is possibly linearly homogeneous with sufficient flexibility so that the signs and magnitudes of hh_i and h_ih_{ij} are not determined *a priori*. Hence, most of the usual production function forms can

[10]Special cases of these specifications also include the broad class of random coefficient functions. But here, again, if random coefficients enter in such a way that $E[h(X, \varepsilon)]$ depends on the variance (covariance) of ε, then Postulates 4 and 6 will not be satisfied. For special functional forms like the Cobb–Douglas or translog functions, Postulates 5 and 7 can also fail [indeed, this case is quite similar to the case treated in the appendix with $y = f(X)e^{h(X)\varepsilon}$].

possibly be used for h as well as f. To facilitate parts of the discussion, it will be assumed that both f and h are log-linear in parameters. The Cobb–Douglas, transcendental, translog, and an approximation of the CES (Kmenta) production functions all have the property of being log-linear in parameters. Also, to facilitate the discussion and proofs, the functions f and h will be written as functions of parameter vectors α and β, respectively, so that the econometric relationship for observation t may be represented as[11]

$$y_t = f(Z_t, \alpha) + h(Z_t, \beta)\varepsilon_t,$$
$$E(\varepsilon_t) = 0, \quad E(\varepsilon_t^2) = 1, \quad E(\varepsilon_t \varepsilon_\tau) = 0, \quad \text{for} \quad t \neq \tau, \tag{8}$$

where

$$\ln f(Z_t, \alpha) \equiv (\ln Z_t)'\alpha \equiv z_t'\alpha, \tag{9}$$

$$\ln h(Z_t, \beta) \equiv (\ln Z_t)'\beta \equiv z_t'\beta, \tag{10}$$

$$Z_t = Z(X_t).$$

5.1. Consistent estimation

The estimation method considered here is as follows. First, define u_t by

$$u_t = h(Z_t, \beta)\varepsilon_t,$$

so that (8) can be written as

$$y_t = f(Z_t, \alpha) + u_t, \qquad E(u_t) = 0, \tag{11}$$

where disturbances are heteroscedastic.[12] Under appropriate assumptions, a nonlinear least-squares regression of (11) provides a consistent estimate of α, say $\hat{\alpha}$. Hence, consistent estimates of u_t are given by

$$\hat{u}_t = y_t - f(Z_t, \hat{\alpha}). \tag{12}$$

Consistent estimates of β can then be obtained in a second stage following methods suggested by Hildreth and Houck (1968) where the square of the residual is regressed on its asymptotic expectation. In this case one can regress $\hat{u}_t^2$ on $h^2(Z_t, \beta)$ since $\hat{u}_t$ is consistent for u_t and $E(u_t^2) = h^2(Z_t, \beta)$. This regression

[11]Note that no generality is lost in assuming $E(\varepsilon_t^2) = 1$ since, if $E(\varepsilon_t^2) = \sigma$, the h function could simply be modified by a multiplicative factor of σ.

[12]This implies that Z_t is either non-stochastic or stochastically independent of u_t; otherwise $E(u_t) = 0$ may not be the case.

can be accomplished very simply by linear means since, in logarithmic form, the right-hand side of the regression relationship is given by (10). These results are stated more precisely in the following theorems.

Theorem 1. Consider the regression equation in (11) *with underlying specifications in* (8), (9), *and* (10) *and suppose that* (i) *the matrix*

$$M_{ZZ} = T^{-1} \sum_{t=1}^{T} z_t z_t'$$

approaches a matrix $\bar{M}$ *which is non-singular as* T *increases,* (ii) $\alpha \in \theta = \{\alpha; |\alpha| \leqq \alpha^*\}$ *for finite* α^* *and the true parameter vector* α_0 *is an interior point of* θ, *and* (iii) Z_t *and the admissible parameter space for* β, *say* Ω, *are suitably bounded so that* $0 < |h(Z_t, \beta)| \leqq \bar{h} < \infty$ *for all* t *and* $\beta \in \Omega$, e.g., $0 < Z_t \leqq Z^*$, $t = 1, 2, \ldots,$ *and* $\Omega = \{\beta; |\beta| \leqq \beta^*\}$ *for finite* Z^* *and* β^*. *Then, the unweighted nonlinear least-squares estimator for* α *in* (11) *is consistent if* α *exists.*

Proof. The proof of this theorem relies heavily on the general lemma given by Malinvaud (1970) for consistency in nonlinear regression. The first condition of Malinvaud's lemma is satisfied when

$$S_T = \sum_{t=1}^{T} Q(\alpha_t^2) > 0,$$

for $\alpha \neq \alpha_0$ and T sufficiently large where

$$Q_t(\alpha) = f(Z_t, \alpha) - f(Z_t, \alpha_0).$$

In this case, however, $Q_t(\alpha) = 0$ only when $\ln f(Z_t, \alpha) - \ln f(Z_t, \alpha_0) = z_t'(\alpha - \alpha_0) = 0$ which cannot hold for all t as T gets large under the non-singularity condition in (i) if $\alpha \neq \alpha_0$. Hence, $[Q_t(\alpha)]^2 > 0$ for some t. The other condition required by Malinvaud's lemma is that

$$\Pr\left\{\sup_{\alpha \in \omega} \sum_{t=1} g_t(\alpha)\varepsilon_t \geqq \tfrac{1}{2}\right\}$$

tends to zero as T increases for all closed sets ω not containing α_0, the true parameter vector, where

$$g_t(\alpha) = \frac{Q_t(\alpha) h(Z_t, \beta)}{\sum_{t=1}^{T} Q_t^2(\alpha)}.$$

First, note that the conditional distribution of

$$\sum_{t=1}^{T} g_t(\alpha)\varepsilon_t,$$

given Z_t, $t = 1, 2, \ldots, T$, has mean zero and variance

$$\sigma_T = \sum_{t=1}^{T} g_t^2(\alpha),$$

and σ_T is positive for large T when $\alpha \in \omega$ by the arguments above. Hence, Chebyshev's inequality yields

$$\Pr\left\{\left|\sum_{t=1}^{T} g_t(\alpha)\varepsilon_t\right| \geqq \tfrac{1}{2}\right\} \leqq 4\sigma_T, \quad \text{for} \quad \alpha \in \omega.$$

If $\sigma_T \to 0$ for all $\alpha \in \omega$, the theorem is proved. But

$$\sigma_T = \sum_{t=1}^{T} g_t^2(\alpha) < h^{-2}\left\{\sum_{t=1}^{T} Q_t^2(\alpha)\right\}^{-1},$$

and

$$\sum_{t=1}^{T} Q_t^2(\alpha) \to \infty,$$

as T increases since condition (i) implies that

$$\sum_{t=1}^{T}\Big[\ln f(Z_t, \alpha) - \ln f(Z_t, \alpha_0)\Big]^2 = \sum_{t=1}^{T}\Big[z'(\alpha-\alpha_0)\Big]^2 \to \infty, \text{ for } \alpha \in \omega,$$

and the exponential function is monotonically increasing at an increasing rate.

Theorem 2. Suppose the conditions of Theorem 1 holds – that the ε_t are identically and independently distributed and fourth moments of the ε_t distribution exist. Then consistent estimates of β except for the constant term are obtained by ordinary least-squares (OLS) regression of $\ln|\hat{u}_t|$ *on z_t [where $\hat{u}_t$ is given by* (12) *and $\hat{\alpha}$ is the α estimate from Theorem* 1] *if the true parameter vector β_0 is an interior point of Ω. Furthermore, the OLS estimator $\hat{\beta}$ is asymptotically normal, and a consistent estimator of the covariance matrix of $\hat{\beta}$ is given by the usual OLS covariance estimator.*

F

Proof. Note that $\hat{u}_t - u_t$ converges to zero trivially by Theorem 1. Hence, asymptotically, a regression equation may be written as

$$|\hat{u}_t| \to |u_t| = h(Z_t, \beta)\,|\varepsilon_t|,$$

or, in logarithmic terms, using (10) and dropping a factor of 2,

$$\ln|u_t| = z_t'\beta + \ln|\varepsilon_t|.$$

Letting $\bar{\beta} = E(\ln|\varepsilon_t|)$ and $v_t = \ln|\varepsilon_t| - \bar{\beta}$, it is then evident that the regression equation,

$$\ln|\hat{u}_t| = \bar{\beta} + z_t'\beta + v_t, \tag{13}$$

applies asymptotically where the v_t are identically and independently distributed with zero mean. Hence, Theorems 6 and 7 of Jennrich (1969) are applicable (his other conditions can be verified trivially as in the usual linear case); hence, the conclusions of Theorem 2 follow aside from modifications in the constant term.

For purposes of later comparison, one can further determine for the regression equation investigated in Theorem 2,

$$\begin{aligned} \ln \hat{u}_t^2 &= \ln h^2(Z_t, \beta) + \ln \tilde{\varepsilon}_t^2 \\ &= 2z_t'\beta + \ln \tilde{\varepsilon}_t^2, \end{aligned} \tag{14}$$

[which is equivalent to (13)] that

$$\ln \tilde{\varepsilon}_t^2 = \ln \frac{\hat{u}_t^2}{h^2(Z_t, \beta)}.$$

But by Theorem 1, $\hat{u}_t - u_t \to 0$ as $T \to \infty$; hence, $\ln \tilde{\varepsilon}_t^2$ converges in distribution to an identically and independently distributed variable v_t^* defined by

$$v_t^* \equiv \ln \frac{u_t^2}{h^2(Z_t, \beta)} \equiv \ln \varepsilon_t^2,$$

(the logarithm of a chi-square variate with one degree of freedom) where, by the methods of Abramowitz and Stegum (1965, pp. 260, 943) as demonstrated by Harvey (1976, p. 462),

$$E(v_t^*) \equiv \bar{\beta} = -1.2704, \qquad V(v_t^*) = 4.9348. \tag{15}$$

Using (14) and (15), it is apparent that the asymptotic covariance matrix of $\hat{\beta}$

(defined by Theorem 2) is, more precisely,

$$V(\hat{\beta}) \to 4.9348\left[\sum_{t=1}^{T} 4z_t z_t'\right]^{-1} = 1.2337\left[\sum_{t=1}^{T} z_t z_t'\right]^{-1}. \tag{16}$$

It is also apparent that a consistent estimator for the constant term coefficient in h is obtained by adding 1.2704 to the one produced by the methods of Theorem 2.

5.2. *Efficient estimation*

The usual continuation of the above procedure in seeking asymptotic efficiency is to estimate the covariance matrix for the regression in (11) using the second-stage estimate of β and then compute a feasible nonlinear generalized least-squares estimator [Hildreth and Houck (1968), Zellner (1962), Gallant (1975)]. Since Jennrich's methods can be used to show asymptotic efficiency of nonlinear generalized least squares for α when β is known and the information matrix is block-diagonal with respect to α and β (as shown below), this usual procedure presumably also leads to asymptotically efficient estimation of α in the case of this paper. However, no improvements in estimation of β are obtained nor can they be made using a similar type of procedure. That is, since the disturbances in (13) are asymptotically independent and identically distributed, when the ε_t in (8) are identically and independently distributed, the second-stage estimator of β corresponds to Aitken's estimation; and no greater asymptotic efficiency can be gained by pursuing eq. (13). Nevertheless, this second-stage estimator of β is asymptotically inefficient.

To see this, note that asymptotic efficiency can be attained for both α and β in the case of normality with maximum likelihood (ML) procedures. In this case the log-likelihood function for (8) is

$$\ln L = -\frac{T}{2}\ln 2\pi - \sum_{t=1}^{T} z_t'\beta - \tfrac{1}{2}\sum_{t=1}^{T}(y_t - \exp(z_t'\alpha))^2 \exp(-2z_t'\beta).$$

Hence, the inverse of the information matrix or, equivalently, the asymptotic covariance matrix of the ML estimator $(\tilde{\alpha}'\tilde{\beta}')'$ of $(\alpha'\beta')'$ is

$$\left\{ E\left[\begin{array}{c:c} \dfrac{\partial^2 \ln L}{\partial\alpha\partial\alpha'} & \dfrac{\partial^2 \ln L}{\partial\alpha\partial\beta'} \\ \hdashline \dfrac{\partial^2 \ln L}{\partial\beta\partial\alpha'} & \dfrac{\partial^2 \ln L}{\partial\beta\partial\beta'} \end{array}\right]_{\substack{\alpha=\tilde{\alpha}\\ \beta=\tilde{\beta}}} \right\}^{-1}$$

$$= \left[\begin{array}{c:c} \left(\sum_{t=1}^{T} \exp\,[2z_t'(\alpha-\beta)]z_t z_t'\right)^{-1} & 0 \\ \hdashline 0 & \dfrac{1}{2}\left(\sum_{t=1}^{T} z_t z_t'\right)^{-1} \end{array}\right]. \tag{18}$$

Thus, comparing (16) and (18), ML estimation of β is more than twice as efficient as step estimation, asymptotically.[13]

Unfortunately, however, ML estimates of α and β cannot be easily and directly obtained with most computerized least-squares procedures because of the presence of the Jacobian [the second right-hand term of (17)] although the same general gradient and search principles are applicable. Alternatively, Fisher's method of scoring [Rao (1973, p. 366) and Theil (1971, p. 527)] suggests an iterative approach to ML estimation with

$$\tilde{\alpha}_{j+1} = \tilde{\alpha}_j - \left[\sum_{t=1}^{T} \exp\,[2z_t'(\tilde{\alpha}_j - \tilde{\beta}_j)] z_t z_t' \right]^{-1} \times \sum_{t=1} \exp\,[z_t'(\tilde{\alpha}_j - 2\tilde{\beta}_j)][y_t - \exp\,(z_t'\tilde{\alpha}_j)] z_t, \tag{19}$$

$$\tilde{\beta}_{j+1} = \tilde{\beta}_j + \left[\sum_{t=1}^{T} z_t z_t' \right]^{-1} \sum_{t=1}^{T} \left[1 - \left(y_t - \exp\,(z_t'\tilde{\alpha}_j) \right)^2 \exp\,(-z_t'\tilde{\beta}_j) \right] z_t. \tag{20}$$

Upon noting one additional fact, a rather simple four-step procedure for asymptotically efficient estimation of both α and β emerges. That is, only one iteration of (19) and (20) is required to produce estimators (α^*, β^*) with the same asymptotic distribution as ML estimators ($\tilde{\alpha}$, $\tilde{\beta}$) when consistent estimators are used to give starting values [Jorgenson (1961, p. 240)]. It is in this context that Theorems 1 and 2 take on significance; namely, consistency and asymptotic efficiency are attained by first applying the methods of Theorems 1 and 2 and then modifying those estimates using (19) and (20):

Step 1. Nonlinear least-squares regression of y_t *on* $f(Z_t, \alpha)$ *producing a consistent estimate* $\hat{\alpha}$.

Step 2. A linear least-squares regression of $\ln|\hat{u}_t|$ *on* z_t *producing* $\hat{\beta}$ *where* $\hat{u}_t = y_t - f(Z_t, \hat{\alpha})$.

Step 3. Modifying $\hat{\alpha}$ *as follows*:

$$\alpha^* = \hat{\alpha} - \left[\sum_{t=1}^{T} \exp\,[2z_t'(\hat{\alpha} - \hat{\beta})] z_t z_t' \right]^{-1} \sum_{t=1}^{T} \exp\,[z_t'(\hat{\alpha} - 2\hat{\beta})] \hat{u}_t z_t.$$

[13]Interestingly, the relative efficiency, 0.4052, es exactly the same as Harvey (1976) found in comparing step estimation with ML estimation for eq. (A.2) of the appendix with multiplicative heteroscedasticity.

Step 4. Modify $\hat{\beta}$ as follows:

$$\beta^* = \hat{\beta} + 1.2704 e_1 + \left[\sum_{t=1}^{T} z_t z_t'\right]^{-1} \sum_{t=1}^{T} \left[1 - \hat{u}_t^2 \exp(-z_t'\hat{\beta})\right] z_t,$$

where $e_1 = (1\ 0 \ldots 0)'$ and the first coefficient of β represents the constant factor in h [see Harvey (1976, p. 464) *for a similar procedure*].

6. Consequences of common estimation procedures

If the model in (8) which carries reasonable risk implications is applicable and one follows the usual procedure of log linear regression, the implications are several. First, it is well known that the covariance estimates and t ratios produced by linear regression techniques are not applicable in their usual sense. Hence, standard statistical procedures used to investigate model specification are not theoretically valid.

But second, note that, in logarithmic terms, (8) becomes

$$\ln y_t = \ln f(Z_t, \alpha) + \ln [1 + \varepsilon_t h(Z_t, \beta)/f(Z_t, \alpha)].$$

Hence, where

$$X = \begin{bmatrix} z_1' \\ \vdots \\ z_T' \end{bmatrix}, \qquad V = \begin{bmatrix} u_1 \\ \vdots \\ u_T \end{bmatrix},$$

and

$$u_t \equiv u_t(\varepsilon_t) \equiv \ln [1 + \varepsilon_t h(Z_t, \beta)/f(Z_t, \alpha)],$$

one finds that the OLS estimate α^* has expectation

$$E(\alpha^*) = \alpha + (X'X)^{-1} X' E(V) \neq \alpha,$$

since $E(V) < 0$ by Jensen's inequality and concavity of $u_t(\cdot)$,[14]

$$E[u_t(\varepsilon_t)] < u_t[E(\varepsilon_t)] = u_t(0) = 0.$$

OLS is, of course, also inconsistent since the magnitude of bias does not generally depend on T [see Kmenta (1971, p. 402) for a similar discussion]. Taking

[14]One can also note that the direction of bias may be in either direction. For example, if

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad E(V) = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix},$$

then α_1 is biased upwards and α_2 is biased downwards.

expectations of a Taylor series expansion of u_t

$$E(u_t) = -\frac{1}{2}\frac{E(\varepsilon_t^2)h^2(Z_t, \beta)}{f^2(Z_t, \alpha)} - \frac{1}{4}\frac{E(\varepsilon_t^4)h^4(Z_t, \beta)}{f^4(Z_t, \alpha)} - \frac{1}{6}\frac{E(\varepsilon_t^6)h^6(Z_t, \beta)}{f^6(Z_t, \alpha)} \dots,$$

it is also evident that the bias is necessarily small only when instability in production is small (relative to expected production). One must then question the use of log linear estimates of f if the propositions of this paper hold.

It is also interesting to consider the relationship of the estimation method presented in this paper and that proposed, for example, by Hildreth and Houck for the random coefficient regression problem. They have discussed their problem in the context of production function estimation where the production function could be represented by

$$\ln y_t = z_t'\alpha = z_t'\alpha_0 + z_t'\delta,$$

where α is random with $\alpha = \alpha_0 + \delta$, $E(\delta) = 0$. The procedure they suggest would then involve first regressing $\ln y_t$ on z_t disregarding heteroskedasticity much like the approach suggested by Theorem 1. Then, in a second stage, the squares of the estimated disturbances are regressed on z_t to estimate the diagonal of the covariance matrix for δ. This procedure is also the same as suggested by Theorem 2 of this paper aside from a logarithmic transformation. However, in the Hildreth and Houck case, all coefficient estimates in the second stage should be positive since they represent variances. In the case discussed in this paper, on the other hand, the second-stage coefficients may be of either sign depending on the relationship of the independent variable with variability. Hildreth and Houck (1968) present a lengthy discussion of the oft-encountered problem where some second-stage estimate is negative and how to artificially avoid such results. On the basis of the model in this paper, however, it appears that negative estimates in the second stage could well be indicative of the stochastic specification in this paper. Indeed, when (asymptotically) significant negative results are obtained in the second stage, it would appear that the empirical cases suggested by Hildreth and Houck lead to rejection of the random coefficient model in favor of the model in (6).

Appendix

The appendix investigates the extent to which Postulates 1 through 8 are satisfied by alternative stochastic specifications and substantiates much of table 1. Consider first the multiplicative specification in (2). Here it suffices to consider the more general case used by Stiglitz with

$$y = (FX_5) = f(X)h(\varepsilon), \tag{A.1}$$

where h is any function of ε. In this case one can verify that Postulates 1, 2, and 3 are satisfied when the corresponding deterministic conditions hold for $f(X)$. Postulates 4 and 6 also hold with (A.1) because $E(y) = f(X)$ does not depend on the distribution (variance) of ε. Finally, Stiglitz (1974) has shown that Postulate 8 holds when $f(X)$ is linearly homogeneous. In the case of Postulate 5, however, problems still remain. Obviously,

$$V(y) = f^2(X)V[h(\varepsilon)].$$

Hence,

$$\frac{\partial V(y)}{\partial X_i} = 2ff_iV[h(\varepsilon)] \geqq 0,$$

since Postulates 1 and 2 imply $ff_i > 0$; and strict equality occurs only when production is non-stochastic. So, again, many of the same counterintuitive arguments which apply to (1) are applicable. Regarding Postulate 7, arguments analogous to those for the specification in (1) are again applicable if one replaces $V(e^{\varepsilon})$ and $E(e^{\varepsilon})$ with $V[h(\varepsilon)]$ and $E[h(\varepsilon)]$, respectively. Hence, arbitrary complementarity relationships are also associated with (2).

Consider now the case in (3) with an additive disturbance which is, perhaps, indicative of production studies prior to the introduction of the Cobb–Douglas production function. Here, again, it is evident that Postulates 1, 2, and 3 are satisfied when the corresponding deterministic conditions hold for $f(X)$. Also, it is immediate that Postulates 4 and 6 hold since expected production does not depend on the distribution of ε. But it is also clear that the variance of production is independent of input use; hence, Postulate 5 is not satisfied. Also, all marginal products are necessarily non-stochastic contrary to Postulate 7. Finally, observe that

$$F_3(\theta X) = f(\theta X)+\varepsilon \neq \theta[f(X)+\varepsilon] = \theta F_3(X).$$

Hence, contrary to Postulate 8, constant stochastic returns to scale are also not possible.

In the multiplicative case where perturbations of the disturbance term are introduced, it is clear that the case with $y = f(X)h(X)\varepsilon$ is formally equivalent to (2). But the functional form with

$$y = F_6(X) \equiv f(X)\, e^{h(X)\varepsilon} \tag{A.2}$$

has not been treated. For simplicity, suppose that $\varepsilon \sim N(0, \sigma)$; hence,

$$E(y) = f(X)\, e^{h^2(X)\sigma/2},$$

$$\frac{\partial E(y)}{\partial X_i} = f_i\, e^{h^2\sigma/2} + f\sigma hh_i\, e^{h^2\sigma/2}, \tag{A.3}$$

$$\frac{\partial^2 E(y)}{\partial X_i^2} = (f_{ii}+2f_i hh_i\sigma + fh^2h_i^2\sigma^2 + fh_i^2\sigma^2 + fhh_{ii}\sigma)\,\mathrm{e}^{h^2\sigma/2}. \tag{A.4}$$

It is thus evident that Postulates 1, 2, and 3 are possibly satisfied with (A.2); however, (A.3) seems to indicate that $\partial E(y)/\partial X_i$ may not be positive if hh_i is somewhat large negatively since $f_i < f$ is normally the case in the relevant range of production. Furthermore, (A.4) implies that either h or h_i must be small if f_{ii} and f_i are small in absolute value by comparison with f. It is also disturbing that concavity of expected output depends on subjective variances. That is, suppose one's subjective variance increases. Since $fh^2h_i^2\sigma^2 > 0$ and $fh_i^2\sigma^2 > 0$ by Postulate 1, it is clear that $\partial^2E(y)/\partial X_i^2$ becomes positive for sufficiently large σ. It can further be determined from (A.2) that

$$\frac{\partial E(y)}{\partial V(\varepsilon)} = \tfrac{1}{2}f(X)h^2(X)\,\mathrm{e}^{h^2(X)\sigma/2}, \tag{A.5}$$

$$\frac{\partial V(y)}{\partial X_i} = 2[\mathrm{e}^{h^2(X)\sigma}-1]E(y)\,\frac{\partial E(y)}{\partial X_i} + E^2(y)\sigma hh_i\,\mathrm{e}^{h_2(X)\sigma}. \tag{A.6}$$

Since Postulate 1 implies $f(X) > 0$ in this case, it is obvious from (A.5) that $\partial E(y)/\partial V(\varepsilon) \geqq 0$ with strict equality occurring only when production is nonstochastic. Hence, Postulate 4 is definitely not satisfied. In the case of Postulate 5, no definite conclusion is possible. It is only evident that the first term in (A.6) must be positive when Postulates 1 and 2 hold; hence, the possibility of $\partial V(y)/\partial X_i < 0$ implies that hh_i must be sufficiently large negatively. But, as indicated above, this requirement may lead to a contradiction of Postulates 2 and 3. Postulates 6 and 7 are more tedious to investigate, and indeterminate conclusions are also reached. For example, one finds under expected profit maximization that

$$\frac{\partial X^*}{\partial V(\varepsilon)} = -\tau^{-1}\,\mathrm{e}^{h^2(X)\sigma/2}[\tfrac{1}{2}f'h^2 + fhh'(1+\tfrac{1}{2}h^2\sigma)],$$

where τ^{-1} is the inverse Hessian matrix (with all negative elements under Bear's conditions), f' is a vector of partial derivatives f_i, and h' is a vector of partial derivatives h_i. Since the sign of $\partial X^*/\partial V(\varepsilon)$ is indeterminate, Postulate 6 apparently can be satisfied. However, $\partial X_i^*/\partial V(\varepsilon)$ requires that some hh_j is sufficiently negative as in the case with Postulate 5. This tends to contradict Postulates 1 and 2. In the case of Postulate 8, it can be easily verified that

$$F_6(\theta X) = f(\theta X)\,\mathrm{e}^{h(\theta X)\varepsilon} \neq \theta f(X)\,\mathrm{e}^{h(X)\varepsilon} = \theta F_6(X),$$

unless $h(X)$ does not depend on X. Hence, Postulate 8 also fails except in the special case where F_6 reduces to F_5 in (A.1).

In the case of (5), however, one can easily verify that

$$E(y) = f(X), \quad \frac{\partial E(y)}{\partial X_i} = f_i, \quad \frac{\partial^2 E(y)}{\partial X_i^2} = f_{ii};$$

hence, Postulates 1, 2, and 3 are satisfied whenever f satisfies the corresponding deterministic criteria usually imposed on production functions. Also, one obtains

$$\frac{\partial E(y)}{\partial V(\varepsilon)} = 0, \quad V(y) = h^2(X)\sigma, \quad \frac{\partial V(y)}{\partial X_i} = 2hh_i\sigma.$$

Thus, Postulate 4 is trivially satisfied, and Postulate 5 is satisfied since the sign of $\partial V(y)/\partial X_i$ switches as the sign of h switches. It is also trivial that Postulate 6 is satisfied since expected output does not depend on the (subjective) variance $V(\varepsilon)$ above.

To verify Postulate 7, note that

$$\frac{\partial y}{\partial X_i} = f_i + h_i\varepsilon, \quad V\left(\frac{\partial y}{\partial X_i}\right) = h_i^2\sigma, \quad \frac{\partial V(\partial y/\partial X_i)}{\partial X_j} = 2h_ih_{ij}\sigma.$$

Obviously, negative, positive, and zero possibilities exist depending on the signs of h_{ij}. Finally, it is evident that Postulate 8 holds since

$$F_4(\theta X) = f(\theta X) + h(\theta X)\varepsilon = \theta[f(X) + h(X)\varepsilon] = \theta F_4(X),$$

whenever f and h are linearly homogeneous functions. In terms of the standards imposed by Postulates 1 through 8, the stochastic specification in (5) thus seems to hold reasonable promise. This is true regardless of whether the deterministic component of (5) follows a Cobb-Douglas, CES, translog, or any other common production function specification.

References

Abramowitz, M. and I.A. Stegum, eds., 1965, Handbook of mathematical functions (Dover Publications, New York).

Arrow, K.J., 1971, Essays in the theory of risk bearing (Markham, Chicago, IL).

Arrow, K., H. Chenery, B. Minhas and R. Solow, 1961, Capital–labor substitution and economic efficiency, Review of Economics and Statistics 43, 225–250.

Bear, D., 1965, Inferior inputs and the theory of the firm, Journal of Political Economy 73, 287–289.

Carlson, G., 1969, Bayesian analysis of pesticide use, 1969 Proceedings of the Business and Economic Statistics Section, American Statistical Association (Washington, DC).

Christensen, L.R., D.W. Jorgenson and L.J. Lau, 1973, Transcendental logarithmic production frontiers, Review of Economics and Statistics 55, 28–45.

Cobb, C. and P. Douglas, 1928, A theory of production, American Economic Review 18, 139–165.

Gallant, A.R., 1975, Seemingly unrelated nonlinear regressions, Journal of Econometrics 4, 35–50.

Halter, A., H. Carter and J. Hocking, 1957, A note on the transcendental production function, Journal of Farm Economics 39, 974–996.

Hanoch, G. and H. Levy, 1969, The efficiency analysis of choices involving risk, Review of Economic Studies 36, 335–346.

Harvey, A.C., 1976, Estimating regression models with multiplicative heteroscedasticity, Econometrica 44, 461–465.

Hildreth, C. and J.P. Houck, 1968, Some estimates for a linear model with random coefficients, Journal of the American Statistical Association 63, 584–595.

Jennrich, R.I., 1969, Asymptotic properties of nonlinear least squares estimators, Annals of Mathematical Statistics 40, 633–643.

Jorgenson, D.W., 1961, Multiple regression analysis of a Poisson process, Journal of the American Statistical Association 56, 235–245.

Just, R.E., 1975, Risk aversion under profit maximization, American Journal of Agricultural Economics 57, 347–352.

Just, R.E. and R.D. Pope, 1977, On the relationship of input decisions and risk, in: J. Roumasset, J. Boussard and I.J. Singh, eds., Risk and uncertainty in agricultural development (University of California Press, Berkeley, CA) forthcoming.

Kmenta, J., 1971, Elements of econometrics (Macmillan, New York).

Malinvaud, E., 1970, The consistency of nonlinear regressions, Annals of Mathematical Statistics 41, 956–969.

Rao, C.R., 1973, Linear statistical inference and its applications, 2nd ed. (Wiley, New York).

Sandmo, A., 1971, On the theory of the competitive firm under price uncertainty, American Economic Review 61, 65–73.

Stiglitz, J., 1974, Incentives and risk sharing in sharecropping, Review of Economic Studies 41, 219–255.

Theil, Henri, 1971, Principles of econometrics (Wiley, New York).

Walters, A., 1960, Marginal productivity and probability distributions of factor services, Economic Journal 70, 325–330.

Zellner, A., 1962, An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, Journal of the American Statistical Association 57, 348–368.

Zellner, A., J. Kmenta and J. Drèze, 1966, Specification and estimation of Cobb–Douglas production function models, Econometrica 34, 784–795.