Journal of Econometrics 13 (1980) 57–66. © North-Holland Publishing Company

LIKELIHOOD FUNCTIONS FOR GENERALIZED STOCHASTIC FRONTIER ESTIMATION*

Rodney E. STEVENSON

University of Wisconsin, Madison, WI 53706, USA

1. Introduction

Economic relationships based on optimization behavior define efficient frontiers of minimum (e.g. cost) or maximum (e.g. production) attainment for any set of relevant conditions. Traditional econometric methods for estimating stochastic economic relationships have implicitly assumed that all economic agents are successful in reaching the efficient frontier. If, however, the economic agents are not equally efficient, then the average relationships estimated by ordinary least squares methods might not reflect the frontier relationships. Our purpose here is to develop a specification for a stochastic frontier model. We seek to create a model which is reasonably general without regard to the pattern of efficiency distribution throughout the sample — a model which subsumes both the special cases of zero and non-zero modes for the distribution of efficiency levels among the economic agents.

Early efforts at specifying frontiers [e.g. Aigner and Chu (1968), Timmer (1971), Afriat (1972), Richmond (1974), or Schmidt (1975)] were either non-stochastic, of questionable economic and statistical justification, or contrary to the usual maximum likelihood regularity conditions. More recently Aigner, Amemiya, and Poirier (1976), Aigner, Lovell, and Schmidt (1977) (ALS), and Meensen and van den Broeck (1977) have suggested stochastic error specifications which are consistent with a stochastic frontier (or random measurement error on the dependent variable) and variable efficiency among the economic agents. While the existence of efficiency variation can be inferred from the Aigner, Amemiya, and Poirier specification, ALS and others explicitly set forth a joint destiny function based on an error form $\varepsilon = u + v$ where u is the error associated with interagent efficiency differences

*This research was supported by the National Science Foundation (Grant no. DAR77-16084-A01). I am indebted to Lane Bishop, Thomas Cowing, William Greene, David Reifschneider, and Donald Waldman for their valuable advice at various stages of the research. However all errors are of my own making.

and v is the error reflecting the stochastic characteristic of the frontier or the measurement error with respect to the dependent variable. ALS assume v to be distributed normally with zero mean and variance of σ_v^2 while assuming that u is distributed either with a 'half normal' density or with an exponential density. Thus for the production case, ALS have (see fig. 1)

$$
\begin{aligned}
\mathrm{ALS}_1:\quad & f(u)=\frac{\sqrt{2}}{\sqrt{\pi}\sigma_u}\exp\left[1/2(u/\sigma_u)^2\right] && \text{for}\quad u<0,\\
& =0 && \text{otherwise;}\\
\mathrm{ALS}_2:\quad & f(u)=\theta\exp(\theta u) && \text{for}\quad u<0,\\
& =0 && \text{otherwise.}
\end{aligned}
\tag{1}
$$

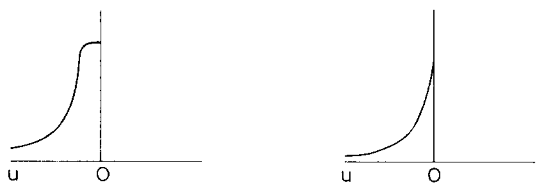

Fig. 1. Left: ALS_1 $f(u)$, and right: ALS_2 $f(u)$.

Both of these specifications assume a distribution of u which has a mode at $u=0$. It is not clear, however, why the mode of u should be expected to occur at $u=0$. If the error term u represents the level of inefficiency, the ALS specifications are based on an implicit assumption that the 'likelihood' of inefficient behavior monotonically decreases for increasing levels of inefficiency. However, characteristics such as degree of educational training, intelligence, persuasiveness, etc. (factors which relate to managerial efficiency) are not likely distributed with such a monotonically declining density function over the population. Since the economic agents are humans or human institutions, the possibility of a non-zero mode for the density function of u would seem a more tenable presumption. Both of the ALS specifications can be generalised to permit a non-zero mode for the density function of u as well as to enable the testing of the special case of a zero mode, as follows.

2. Stochastic specification for cost functions

2.1. Normal-truncated normal

Let us assume that the frontier relationship we seek to estimate is the dual cost function. We assume the error of the cost function is

$$\varepsilon = u + v, \tag{2}$$

where u and v are independently distributed. Given cost minimization behavior, u will be non-negative. Let us further assume that u and v are distributed as

$$\begin{aligned} k(u) &= \frac{1}{(1-F^*(-\mu/\sigma_u))\sqrt{2\pi}\,\sigma_u}\exp\left[-\frac{1}{2}\left(\frac{u-\mu}{\sigma_u}\right)^2\right] && \text{for} \quad u>0, \\ &= 0 && \text{otherwise;} \end{aligned} \tag{3}$$

and

$$g(v) = \frac{1}{\sqrt{2\pi}\,\sigma_v}\exp\left[-\frac{1}{2}\left(\frac{v}{\sigma_v}\right)^2\right] \qquad \text{for all } v,$$

where $F^*(\cdot)$ is the distribution function for a standard normal random variable. Simply stated, u is assumed to be distributed as a truncated normal with mode μ, and v is assumed to be distributed as a normal with zero mean and variance σ_v^2. The joint density function for $\varepsilon = u + v$ is given as

$$\begin{aligned} h(\varepsilon) = \int_0^\infty & \frac{1}{(1-F^*(-\mu/\sigma_u))2\pi\sigma_u\sigma_v} \\ & \times \exp\left[-\frac{1}{2}\left(\left(\frac{u-\mu}{\sigma_u}\right)^2 + \left(\frac{\varepsilon-u}{\sigma_v}\right)^2\right)\right]\mathrm{d}u, \end{aligned} \tag{4}$$

which integrates to

$$h(\varepsilon) = \sigma^{-1} f^*\left(\frac{\varepsilon-\mu}{\sigma}\right)\left[1-F^*\left(-\frac{\mu}{\sigma\lambda}-\frac{\varepsilon\lambda}{\sigma}\right)\right]\left[1-F^*\left(-\frac{\mu}{\sigma_u}\right)\right]^{-1}, \tag{5}$$

where $\sigma = (\sigma_u^2 + \sigma_v^2)^{1/2}$, $\lambda = \sigma_u/\sigma_v$ and f^* is the standard normal density evaluated at $((\varepsilon-\mu)/\sigma)$. Note that at $\mu=0$, $h(\varepsilon)$ becomes

$$h(\varepsilon)\big|_{\mu=0}=\frac{2}{\sigma}f^*\left(\frac{\varepsilon}{\sigma}\right)\left[1-F^*\left(-\frac{\varepsilon\lambda}{\sigma}\right)\right], \tag{6}$$

which is the cost function analog to the ALS formulation.

The mean and variance of ε are

$$E(\varepsilon)=E(u)=\frac{\mu a}{2}+\frac{\sigma_u a}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\mu}{\sigma_u}\right)^2\right], \tag{7}$$

$$V(\varepsilon)=V(u)+V(v)$$
$$=\mu^2\frac{a}{2}\left(1-\frac{a}{2}\right)+\sigma_u^2\frac{a}{2}\left(\frac{\pi-a}{\pi}\right)+\sigma_v^2,$$

where $a=(1-F^*(-\mu/\sigma_u))^{-1}$. At $\mu=0$, the mean and variance of ε becomes

$$E(\varepsilon)\big|_{\mu=0}=\frac{\sqrt{2}}{\sqrt{\pi}}\sigma_u,$$

$$V(\varepsilon)\big|_{\mu=0}=\sigma_u^2\left(\frac{\pi-2}{\pi}\right)+\sigma_v^2. \tag{8}$$

Noting that σ_u can be written as $\sigma(\lambda^{-2}+1)^{-\frac{1}{2}}$, the logged likelihood function for the cost function frontier model is

$$\ln L=\ln(\boldsymbol{Y}|\boldsymbol{\beta},\lambda,\sigma^2,\mu)$$
$$=-\frac{n}{2}\ln\sigma^2-\frac{n}{2}\ln 2\pi-\frac{1}{2\sigma^2}\sum_{i=1}^{n}((Y_i-\boldsymbol{\beta}'\boldsymbol{X}_i)-\mu)^2$$
$$+\sum_{i=1}^{n}\ln\left[1-F^*\left(\sigma^{-1}\left(-\frac{\mu}{\lambda}-(Y_i-\boldsymbol{\beta}'\boldsymbol{X}_i)\lambda\right)\right)\right]$$
$$-n\ln\left[1-F^*\left(-\frac{\mu}{\sigma}(\lambda^{-2}+1)^{\frac{1}{2}}\right)\right], \tag{9}$$

where $Y_i=\boldsymbol{\beta}'\boldsymbol{X}_i+\varepsilon_i$ with $\boldsymbol{\beta}$ and $\boldsymbol{X}_i$ being $[1\times K]$ vectors.

Taking derivatives, we have

$$\frac{\partial \ln L}{\partial \beta}=\frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i-\boldsymbol{\beta}'\boldsymbol{X}_i-\mu)\boldsymbol{X}_i-\frac{\lambda}{\sigma}\sum_{i=1}^{n}\frac{f_{2i}^*}{(1-F_{2i}^*)}\boldsymbol{X}_i, \tag{10}$$

$$\begin{aligned}\frac{\partial \ln L}{\partial \lambda}&=-\sum_{i=1}^{n}\frac{f_{2i}^*}{(1-F_{2i}^*)}\left(\frac{\mu}{\lambda^2}-(Y_i-\boldsymbol{\beta}'\boldsymbol{X}_i)\right)\frac{1}{\sigma}\\&\quad+\frac{n\mu}{\sigma\lambda^3}(\lambda^{-2}+1)^{-\frac{1}{2}}\frac{f_1^*}{(1-F_1^*)},\end{aligned} \tag{11}$$

$$\begin{aligned}\frac{\partial \ln L}{\partial \sigma^2}&=-\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^{n}(Y_i-\boldsymbol{\beta}'\boldsymbol{X}_i-\mu)^2\\&\quad+\frac{1}{2\sigma^3}\sum_{i=1}^{n}\frac{f_{2i}^*}{(1-F_{2i}^*)}\left(-\frac{\mu}{\lambda}-(Y_i-\boldsymbol{\beta}'\boldsymbol{X}_i)\lambda\right)\\&\quad+\frac{n\mu(\lambda^{-2}+1)^{\frac{1}{2}}}{2\sigma^3}\frac{f_1^*}{(1-F_1^*)},\end{aligned} \tag{12}$$

and

$$\begin{aligned}\frac{\partial \ln L}{\partial \mu}&=\frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i-\boldsymbol{\beta}'\boldsymbol{X}_i-\mu)+\frac{1}{\lambda\sigma}\sum_{i=1}^{n}\frac{f_{2i}^*}{(1-F_{2i}^*)}\\&\quad-\frac{n(\lambda^{-2}+1)^{\frac{1}{2}}}{\sigma}\frac{f_1^*}{(1-F_1^*)},\end{aligned} \tag{13}$$

where f_1^* and F_1^* are the standard normal density and distribution functions, respectively, evaluated at $(-(\mu/\sigma)(\lambda^{-2}+1)^{\frac{1}{2}})$ and f_{2i}^* and F_{2i}^* are the standard normal density and distribution functions evaluated at $(\sigma^{-1}(-\mu/\lambda-(Y_i-\boldsymbol{\beta}'\boldsymbol{X}_i)\lambda))$.

The first-order derivatives can be used in a nonlinear optimization algorithm to derive the MLE estimates of β, λ, σ^2, μ, and the associated variances.

The model described above can be compared to the OLS model predicated on a single error $(\sim N(0,\sigma_v^2))$ using the standard likelihood ratio test where the restricted model is computed with $\mu=\lambda=0$. Alternatively, the adequacy of the ALS specification can be determined by computing the t-statistic for μ.

2.2. *Normal-gamma*

An alternative specification of the stochastic frontier model, and a generalization of the second model suggested by ALS, can be derived by assuming that $v \sim N(0,\sigma_v^2)$ and that the density function for $f(u)$ is the gamma density,

$$f(u)=\frac{1}{\Gamma(m+1)}\theta^{m+1}u^m\exp(-\theta u) \quad \text{for} \quad u>0, \quad \theta>0, \quad m>-1,$$
$$=0 \qquad \text{otherwise.} \tag{16}$$

Given $g(v)=g(\varepsilon-u)$ from above, the joint density function for $\varepsilon=u+v$ is given by

$$h(\varepsilon)=\int_0^\infty \frac{1}{\Gamma(m+1)\sqrt{2\pi}\,\sigma_v}\theta^{m+1}u^m\exp\left[-\frac{1}{2\sigma_v^2}\left(\frac{\theta\sigma^2 u}{2}+(\varepsilon-u)^2\right)\right]\mathrm{d}u, \tag{17}$$

which simplifies to

$$h(\varepsilon)=\frac{\sigma_v^m}{\Gamma(m+1)\sqrt{2\pi}}\exp\left(-\varepsilon\theta+\frac{\theta^2\sigma_v^2}{2}\right)\theta^{m+1}\int_w^\infty (t-w)^m\exp\left[-\frac{t^2}{2}\right]\mathrm{d}t, \tag{18}$$

where $w=(-\varepsilon/\sigma_v+\theta\sigma_v)$. The mean and variance of ε are given by

$$E(\varepsilon)=E(u)=\frac{m+1}{\theta},$$

$$V(\varepsilon)=V(u)+V(v)=\frac{m+1}{\theta^2}+\sigma_v^2. \tag{19}$$

Assuming that m takes on only integer values, we derive the following:

For $\underline{m=0}$,

$$h(\varepsilon)=\theta\exp\left(-\varepsilon\theta+\frac{\theta^2\sigma_v^2}{2}\right)\left[1-F^*\left(-\frac{\varepsilon}{\sigma_v}+\theta\sigma_v\right)\right]. \tag{20}$$

For $\underline{m=1}$,

$$h(\varepsilon)=\sigma_v\frac{\theta^2}{2}\exp\left(-\varepsilon\theta+\frac{\theta^2\sigma_v^2}{2}\right)\left[\frac{1}{\sqrt{2\pi}}\left(1-\exp\left(\frac{1}{2}\left(-\frac{\varepsilon}{\sigma_v}+\theta\sigma_v\right)^2\right)\right)\right.$$
$$\left.-\left(-\frac{\varepsilon}{\sigma_v}+\theta\sigma_v\right)\left(1-F^*\left(-\frac{\varepsilon}{\sigma_v}+\theta\sigma_v\right)\right)\right]. \tag{21}$$

For $m=2$,

$$h(\varepsilon)=\sigma_v^2\frac{\theta^3}{6}\exp\left(-\varepsilon\theta+\frac{\theta^2\sigma_v^2}{2}\right)\left[\frac{1}{2}\left(1-G^*\left(\left(-\frac{\varepsilon}{\sigma_v}+\theta\sigma_v\right)^2\right)\right)\right.$$
$$-\frac{\sqrt{2}}{\sqrt{\pi}}\left(-\frac{\varepsilon}{\sigma_v}+\theta\sigma_v\right)\left(1-\exp\left(\frac{1}{2}\left(-\frac{\varepsilon}{\sigma_v}+\theta\sigma_v\right)^2\right)\right)$$
$$\left.+\left(-\frac{\varepsilon}{\sigma_v}+\theta\sigma_v\right)^2\left[1-F^*\left(-\frac{\varepsilon}{\sigma_v}\right)\right]\right],\qquad \text{etc.} \tag{22}$$

$F^*(\cdot)$ is the distribution function for a standard normal, and $G^*(\cdot)$ is the chi-square distribution with three degrees of freedom. We note that at $m=0$, the joint density function is the cost function analog of the ALS formulation. For each integer value of m, the 'conditional' logged likelihood function can be formed, and MLE estimates of $\boldsymbol{\beta}$, σ_v^2, and θ can be derived along with the associated variances. M can be selected by comparing the value of the likelihood functions evaluated at the various optimal values of $\boldsymbol{\beta}$, σ_v^2, and θ. Thus, it is possible to test the adequacy of an exponential specification for the efficiency error term.

3. Empirical analysis

To test the empirical significance of the generalized stochastic frontier model, estimation routines were run on two separate data sets.[1] With each data set OLS, ALS, and generalized frontier parameter estimates are derived. For the frontier models, we restrict our empirical analysis to the estimation of the normal-truncated normal cases.

The first example uses U.S. primary metals industry (SIC33) 1975–1958 data. This data has been previously analyzed by Hildebrand and Liu (1965), Aigner and Chu (1968), Aigner, Lovell and Schmidt (1977).[2] We estimate two models. First we estimate a simple Cobb–Douglas function,

$$\ln Y=\beta_0+\beta_1\ln L+\beta_2\ln K+\varepsilon, \tag{23}$$

[1]The frontier estimation routines used in this paper are based on a computer program developed by Donald Waldman. Adaptation of the Waldman program to the generalized ALS case and estimation of the specific equations reported here were done by David Reifschneider.

[2]Data for the estimation of eqs. (21) and (22) were kindly provided by William Greene. As Greene noted in his correspondence to the author, ambiguities in the original description of the Hildebrand–Liu data have made an exact duplication of the original data set quite difficult. As noted in the estimates of eq. (22), we have not been able to replicate the original Hildebrand and Liu results.

where Y is value added per establishment, L is labor input per establishment, and K is the ratio of the gross book value of plant and equipment per establishment.

The second model is that employed by Hildebrand and Liu (1965), i.e.,

$$\ln Y=\beta_0+\beta_1 \ln L+\beta_2(\ln R \cdot \ln K)+\varepsilon, \tag{24}$$

where R is the ratio of net book value to gross book value of plant and equipment.

Parameter estimates for eqs. (23) and (24) are given in tables 1 and 2, respectively. As indicated in the tables, μ is estimated to be negative and, by an asymptotic t-test, significant.

Table 1

Estimates of eq. (23).

	OLS	ALS	Generalized ALS
β_0	0.234[a]	−0.0002	0.503[a]
β_1	0.694[a]	0.694[a]	0.687[a]
β_2	0.328[a]	0.328[a]	0.333[a]
σ		0.174[a]	0.175[a]
$\lambda^{\frac{1}{2}}$		0.009	456.5
μ			−0.502[a]

[a]Significant at 99% level.

Table 2

Estimates of eq. (24).

	OLS	ALS	Generalized ALS
β_0	1.212[a]	1.030	1.380[a]
β_1	0.862[a]	0.862[a]	0.805[a]
β_2	0.052[a]	0.052[a]	0.063[a]
σ		0.168[a]	0.189[a]
$\lambda^{\frac{1}{2}}$		0.014	66.5
μ			−0.318[a]

[a]Significant at 99% level.

For our second example, we estimated a cost function premised on the cost minimizing behavior of electrical utility firms. Utilizing a trans-log flexible form specification and assuming linear homogeneity in factor prices, our estimating equation for a firm employing capital, labor, and fuel for electricity generation is

$$\begin{aligned}\ln C = {} & \alpha_0 + \alpha_1 \ln P_L + \alpha_2 \ln P_F + 1/2\alpha_{11}(\ln P_L)^2 + 1/2\alpha_{22}(\ln P_F)^2 \\ & + \alpha_{12} \ln P_L \ln P_F + \beta_1 \ln Q + 1/2\beta_2(\ln Q)^2 \\ & + \beta_{11} \ln Q \ln P_L + \beta_{12} \ln Q \ln P_F, \end{aligned} \tag{25}$$

where C = normalized costs, P_L = normalized price of labor, P_F = normalized price of fuel, and Q = kilowatt hours of electricity generated. In accordance with the assumed constraint of linear homogeneity in prices, C, P_L, and P_F are normalized by the price of capital.[3] The sample is based upon firm level observations for 81 electrical utilities in 1970.

Table 3

Estimates of eq. (25).

	OLS	ALS	Generalized ALS
α_0	18.250[a]	17.930[a]	17.850[a]
α_1	0.768[a]	0.788[a]	0.760[a]
α_2	0.300[a]	0.227[a]	0.218[a]
α_{11}	−0.493	−0.217	0.206
α_{22}	0.104	0.158	0.276
α_{12}	0.377	0.146	0.019
β_1	0.915[a]	0.903[a]	0.908[a]
β_2	0.032	0.014[a]	0.048[a]
β_{11}	0.174	0.131[a]	0.123[a]
β_{12}	0.081	0.115[a]	0.130[a]
σ		0.233[a]	0.191[a]
$\lambda^{\frac{1}{2}}$		3.030	2.103[a]
μ			0.187[a]

[a]Significant at 99% level.

Comparative estimates of eq. (25) are contained in table 3. As indicated in that table the estimate of μ is positive and, as indicated by an asymptotic t-test, significant.

[3]A description of the data is provided in Stevension (1980).

4. Conclusions

Our purpose has been to present a more general error specification for a stochastic frontier estimation model. Our model is more general than those previously presented in the literature in that the specifications presented here are appropriate both for efficiency distributions (i.e., the distribution of u) which have modes at $u=0$ (the ALS cases) and for those with non-zero modes. While the normal-gamma distribution case which we presented here utilized an integer specification of one of the parameters (m), no such restrictions were imposed for the normal-truncated normal case. The costs for the generalized specifications presented here are slightly more complicated first order derivatives of the likelihood function and one additional parameter to be estimated.

Finally, it is worth noting that μ in the normal-truncated normal case is not constrained to be positive (e.g., for a cost function) or negative (e.g., for a production). Thus estimates generated from the normal-truncated normal case (e.g., with a satisfactory negative μ in the cost function case) might reveal a distribution pattern for u which would be quite similar to an exponential distribution.

References

Afriat, S.N., 1972, Efficiency estimation of a production function, International Economic Review 13, Oct., 568–598.

Aigner, D.J. and S.F. Chu, 1968, On estimating the industry production function, American Economic Review 58, 826–839.

Aigner, D.J., T. Ameniya and D.J. Poirier, 1976, On the estimation of production frontiers, International Economic Review 17, June, 377–396.

Aigner, D.J., C.A.K. Lovell and P. Schmidt, 1977, Formulation and estimation of stochastic frontier production function models, Journal of Econometrics 6, July, 21–37.

Meensen, W. and J. van den Broeck, 1977, Efficiency estimation from Cobb-Douglas production functions with composed error, International Economic Review 18, June, 435–444.

Richmond, J., 1974, Estimating the efficiency of production, International Economic Review 15, June, 415–421.

Schmidt, P., 1975, On the statistical estimation of parametric frontier production functions, Review of Economics and Statistics 58, May, 238–239.

Stevenson, R.E., 1980, Measuring technological bias, American Economic Review, forthcoming.

Timmer, C.P., 1971, Using a probabilistic frontier production function to measure technical efficiency, Journal of Political Economy 79, 776–794.