Working Paper 93-06
Statistics and Econometrics Series 06
March 1993

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

BAYESIAN EFFICIENCY ANALYSIS WITH A FLEXIBLE COST FUNCTION

Gary Koop, Jacek Osiewalski and Mark F.J. Steel*

Abstract

In this paper we describe the use of Gibbs sampling methods for drawing posterior inferences in a model with an asymptotically ideal price aggregator, non-constant returns to scale and composed error. An empirical example illustrates the sensitivity of efficiency measures to assumptions made about the functional form of the frontier.

Key Words
Stochastic Frontiers; Asymptotically Ideal Model; Composed Error; Gibbs Sampler; Metropolis Chain.

*Koop, Department of Economics, University of Toronto and Faculty of Economics and Politics, University of Cambridge; Osiewalski, Department of Econometrics, Academy of Economics, Kraków; Steel, Department of Statistics and Econometrics, Universidad Carlos III de Madrid. We would like to thank John Geweke for providing the computer code for sampling from the truncated multivariate Student-t distribution. The second author would like to acknowledge the hospitality of Tilburg University.

Cost or production functions with composed error are commonly used by microeconomists in the measurement of firm inefficiency. At the same time, the development of seminonparametric methods has allowed researchers to work with very flexible cost functions without composed error. In this paper, we unite these two strands of the literature and develop Bayesian techniques for analyzing flexible functional form cost functions with composed error. We argue that such techniques allow for much more accurate understanding of firm efficiency than do traditional methods.

The paper is organized as follows. Section 1 discusses composed error models. Section 2 describes the asymptotically ideal model (AIM) which forms the basis of the seminonparametric approach we use in the paper. Section 3 introduces the AIM cost function with composed error and develops the Gibbs sampler. Section 4 applies our method to an empirical example, and Section 5 concludes.

Section 1: Composed Error Models

Composed error models were first introduced by Meeusen and van den Broeck (1977) and Aigner, Lovell and Schmidt (1977). Bauer (1990) provides a survey of the literature. The basic model is given by:

$$\ln(C_i)=h(S_i,\gamma)+z_i+v_i. \tag{1}$$

This model decomposes the log of observed costs for firm i (C_i) into three parts: i) The log of the actual frontier which depends on S_i, a vector of exogenous variables, and which represents the minimum possible cost of producing a given level of output with certain input prices. ii) A non-negative disturbance, z_i, which captures the level of firm inefficiency. iii) A symmetric disturbance, v_i, which captures other effects due, for instance, to measurement error.

In an empirical exercise assumptions are commonly made about these three components. Usually one takes the v_is to be i.i.d. $N(0,\sigma^2)$ and independent of the z_is, an assumption we maintain throughout this paper. Assumptions, which are not so innocuous, must also be made about h(.,.) and z_i. Typically, z_i is taken to be i.i.d. D(.), where D(.) is some one-sided distribution on $\Re^+$. Common choices for D(.) are truncated Normal, exponential or Gamma. Since interest usually centers on firm inefficiency, accurate estimation of the z_is is essential, and choosing an inappropriate D(.) may have harmful consequences. In a previous paper (van den Broeck, Koop, Osiewalski and Steel (1993), hereafter BKOS), four different choices for D(.) were used: truncated Normal, and Gamma with shape parameters 1, 2 and 3.[1] We were able to take weighted averages across our four choices for D(.) by using posterior model probabilities as weights, and argued that such an approach was preferable to choosing one particular distribution for D(.).

Although BKOS addressed the issue of uncertainty about D(.), it assumed that h(.,.) was linear in γ, an assumption we propose to relax in this paper. The exact functional form used is described in Section 2. For present purposes it is sufficient to note that estimates of z_i can be sensitive to choice of h(.,.), and that most of the existing literature assumes that h(.,.) takes a simple form (eg. Greene (1990) and BKOS use a variant of the Cobb-Douglas cost function). Accordingly, we intend to examine to what extent inferences on firm efficiencies, usually the prime objective of composed error models, can depend on the functional form of the frontier.[2] By using seminonparametric methods, we intend to let the data reveal what h(.,.) should be.

[1]The Gamma with shape parameter 1 is the exponential distribution.

[2]The use of panel data can eliminate the need for distributional assumptions to be made for z_i. However, even with panel data it is important to specify h(.,.) correctly. Thus, the techniques of this paper are relevant even without the composed error framework. Indeed, it is worth stressing that the Gibbs sampling techniques developed here are innovative even for the analysis of standard cost functions. By eliminating the z_i term, our techniques are able to provide an exact Bayesian analysis of the standard Asymptotically Ideal Model with nonconstant returns to scale. In addition, they can be extended quite easily to other nonlinear models such as the generalized translog.

Section 2: The Asymptotically Ideal Model[3]

The large amount of research that has gone into finding flexible functional forms testifies to the great importance of avoiding gross specification error. This is especially so in the case of composed error models since measures of inefficiency can be very misleading if an inappropriate choice for h(.,.) is made. The separation of the two error terms is the main challenge in these models, and generally proves to be least robust to arbitrary assumptions made by the user. In view of this problem, we propose using seminonparametric techniques to approximate the underlying cost function.

Such techniques involve taking an expansion of a parametric form for the cost function. If properly chosen, the resulting seminonparametric cost function can, as the order of expansion increases, approach any possible function. The seminonparametric approximation we use in this paper is based on the Muntz-Szatz expansion and results in the Asymptotically Ideal Model (AIM) discussed in Barnett, Geweke and Wolfe (1991b). To motivate the advantages of the AIM, let us consider two criteria for judging a cost function: regularity and flexibility. If a cost function is regular, it satisfies the restrictions implied by economic theory; if it is flexible, it includes a wide variety of functional forms. A cost function such as the translog, which involves taking a second order Taylor Series expansion about a point, is locally flexible but not regular. The translog may be made regular at a particular data point by imposing restrictions. However, since such restrictions involve both parameters and the data, they can only be imposed at a point. In contrast, the AIM model uses the Muntz-Szatz expansion, which is globally flexible. Global regularity is imposed on the AIM model through parametric restrictions alone.[4]

Seminonparametric methods are useful in that they allow for

[3]Much of the material in this section is drawn from Koop and Carey (1992).

[4]The importance of being able to impose global regularity has been emphasized by many authors. See Barnett, Geweke and Wolfe (1991b), p. 15, for an extensive bibliography.

the data to determine what the key properties of the cost function should be. A danger associated with some seminonparametric methods is the possibility of overfitting. Early seminonparametric models (Gallant (1981)) used Fourier expansions such that economic functions of interest were approximated using sines and cosines. Since cost functions are concave, many terms in the Fourier expansion are typically necessary, increasing greatly the risk of overfitting. It is for this reason that we favor the Muntz-Szatz expansion over the Fourier expansion; it allows for the approximation of globally regular cost functions with an expansion globally regular at every degree. The AIM model fits only that part of the data that is globally regular, thereby eliminating the risk of overfitting. For a more detailed discussion of overfitting see Barnett, Geweke and Wolfe (1991a) pp. 433-434 or Barnett, Geweke and Wolfe (1991b) p. 12.

More specifically, consider the model with one output (Q) and three input prices ($p=(p_1,p_2,p_3)'$). If constant returns to scale hold, then there exists a price aggregator, f(p), such that the frontier cost function takes the form:

$$C(Q,p)=Qf(p)\,. \tag{2}$$

A seminonparametric approach to this simple case would involve choosing an expansion to model f(p). Barnett, Geweke, and Wolfe (1991b) use the Muntz-Szatz expansion[5] and call the resulting model the AIM model. Note that, while the model given above is globally flexible, in the absence of restrictions it is not globally regular, since f(p) can be any function, including non-concave or non-homogeneous (or even negative outside the range of the data). Barnett, Geweke and Wolfe describe how linear homogeneity can be imposed on f(p) in a simple way. In addition, the authors ensure that C(Q,p) is quasi-concave and non-decreasing in input prices by requiring that all the coefficients

[5]We define the Muntz-Szatz expansion later. For present purposes it is sufficient to note that this expansion has the property that it can reach any continuous function if enough terms in the expansion are taken.

of f(p) be non-negative. Their resulting AIM model is globally regular in that any order expansion will satisfy the restrictions implied by economic theory.

It is worth noting that the restriction that the coefficients of f(p) be non-negative is a sufficient but not necessary condition for monotonocity and quasi-concavity. Hence the AIM model with this restriction is not globally flexible; that is, there exist some regular functions which cannot be approximated by C(Q,p) once the coefficients are constrained to be non-negative. Despite this drawback, Barnett, Geweke and Wolfe argue that the restricted AIM specification is extremely flexible and adequate for their purposes (Barnett, Geweke and Wolfe (1991b), p. 18).[6]

Non-constant returns to scale that vary with output can be incorporated by specifying the frontier cost function as:

$$C(Q,p)=Q^{\beta_1+\beta_2 \ln Q}f(p)\,. \qquad (3)$$

The above cost frontier will be adopted here with an AIM(q) form for f(p), where q is the order of the expansion. The Muntz-Szatz expansion has yet to be explicitly defined, as its general formulation is complicated. Once we impose linear homogeneity, however, it can be greatly simplified. Below we give the linear homogenous Muntz-Szatz expansion for f(p) for q=1 and 2, for the case where there are three inputs. (Since our empirical example involves only 123 firms, setting q=3 or higher would probably provide too rich a parameterization. In any case, the evidence we obtain clearly suggests that q=2 is sufficient for our application).

---

[6]There are other ways of imposing regularity. For example, Gallant and Golub (1984) develop computational methods for imposing curvature restrictions at any arbitrary set of points. Their methods can be used to ensure that a cost function is quasi-concave at each data point. Terrell (1992) develops a Bayesian method for imposing regularity conditions over regions of prices, rather than simply at particular data points. Since it is very simple to impose, we use the non-negativity restriction given in Barnett, Geweke and Wolfe (1991b) here.

f(p) for use in AIM(1)

$$f(p)=\alpha_1p_1+\alpha_2p_2+\alpha_3p_3+\alpha_4p_1^{\frac{1}{2}}p_2^{\frac{1}{2}}+\alpha_5p_1^{\frac{1}{2}}p_3^{\frac{1}{2}}+\alpha_6p_2^{\frac{1}{2}}p_3^{\frac{1}{2}}$$

f(p) for use in AIM(2)

$$\begin{aligned}f(p)=&\alpha_1p_1+\alpha_2p_2+\alpha_3p_3+\alpha_4p_1^{\frac{1}{2}}p_2^{\frac{1}{2}}+\alpha_5p_1^{\frac{1}{2}}p_3^{\frac{1}{2}}+\alpha_6p_2^{\frac{1}{2}}p_3^{\frac{1}{2}}\\&+\alpha_7p_1^{\frac{3}{4}}p_2^{\frac{1}{4}}+\alpha_8p_1^{\frac{3}{4}}p_3^{\frac{1}{4}}+\alpha_9p_2^{\frac{3}{4}}p_3^{\frac{1}{4}}+\alpha_{10}p_1^{\frac{1}{4}}p_2^{\frac{3}{4}}+\alpha_{11}p_1^{\frac{1}{4}}p_3^{\frac{3}{4}}\\&+\alpha_{12}p_2^{\frac{1}{4}}p_3^{\frac{3}{4}}+\alpha_{13}p_1^{\frac{1}{2}}p_2^{\frac{1}{4}}p_3^{\frac{1}{4}}+\alpha_{14}p_1^{\frac{1}{4}}p_2^{\frac{1}{2}}p_3^{\frac{1}{4}}+\alpha_{15}p_1^{\frac{1}{4}}p_2^{\frac{1}{4}}p_3^{\frac{1}{2}}\end{aligned}$$

Linear homogeneity in input prices is assured, since the exponents in each term sum to one. If each element of $\alpha=(\alpha_1,..,\alpha_k)'$, (k=6 for AIM(1) and k=15 for AIM(2)) is non-negative, then f(p) is non-negative, non-decreasing and quasi-concave for all positive input prices. It is worth noting that the first degree expansion yields a cost function identical to the commonly used generalized Leontief model.

Section 3: The Gibbs Sampler[7]

If we use (3) to model the log of the cost frontier, $h(S_i,\gamma)$ in (1), we obtain the model used in this paper:

$$\ln(C_i)=\beta_1\ln(Q_i)+\beta_2\ln^2(Q_i)+\ln(f(p_i))+z_i+v_i, \quad \textbf{(4)}$$

where v_i is i.i.d. $N(0,\sigma^2)$, z_i is i.i.d. exponential[8] with parameter λ and $p_i=(p_{1i},p_{2i},p_{3i})'$. The likelihood function for the model based on a sample of size N can be easily derived as in BKOS (Section 3), but for the sake of brevity we do not present it here. We assume a reference prior density which is flat on

[7]The Gibbs sampler is a technique for obtaining a random sample from a joint distribution by taking random draws only from the full conditional distributions. A detailed description of the technique can be found in Casella and George (1992) and Gelfand and Smith (1990).

[8]It would be trivial to allow z_i to take other forms. However, we do not do so here in order to focus analysis on the modelling of the frontier. Note that the posterior odds analysis in BKOS favored, for the same application considered here (albeit with a Cobb-Douglas price aggregator), the exponential assumption for z_i.

$\beta=(\beta_1,\beta_2)'$, $\ln(\sigma^2)$ and $\ln(\lambda)$, but where the elements of α are restricted to be nonnegative. That is,

$$p(\alpha,\sigma^2,\beta,\lambda)\propto\sigma^{-2}\lambda^{-1}p(\alpha),$$

where $p(\alpha)=1$ if all the elements of α are nonnegative and $=0$ otherwise. These assumptions define our Bayesian model from which posterior inferences on the parameters or the z_is can be made.

In BKOS we carried out a Bayesian analysis of a similar model using Monte Carlo integration with importance sampling. In a subsequent paper (Koop, Steel and Osiewalski (1992), hereafter KSO), we argued that the computational difficulties surrounding Monte Carlo integration with importance sampling were truly daunting and recommended the use of Gibbs sampling methods instead. The Gibbs sampler derived in KSO was found to work very well and yielded very accurate results with a relatively light computational burden. However, the Gibbs sampler in KSO was derived for a Cobb-Douglas price aggregator in (3), implying that $h(.,.)$ was linear in γ. Because the log of the AIM cost function used in this paper is not linear in α, a subvector of $\gamma'=(\alpha',\beta')$, the Gibbs sampler is different from that developed in KSO. It is worth emphasizing that, although the Gibbs sampler derived here is for the extension to the AIM cost function given in (3), similar methods can be used to carry out a Bayesian analysis of other nonlinear cost functions.

To develop our Gibbs sampler we begin with some notation: let $y_i=-\ln(C_i)$, $x_i=(-\ln(Q_i)\ -\ln^2(Q_i))'$, and $w_i'\alpha=f(p_i)$ (ie. let w_i contain the decreasing fractional powers of p_i given in Section 2). Furthermore, let X, z, y and w_α indicate the vectors or matrices containing data on all firms for x_i, z_i, y_i and $\ln(w_i'\alpha)$.

Since, conditionally on α, the frontier is linear, we can draw on results from KSO to state:

$$\begin{aligned} p(\beta,\sigma^{-2}\,|\,Data,z,\alpha,\lambda^{-1}) &= p(\beta,\sigma^{-2}\,|\,Data,z,\alpha)= \\ f_G(\sigma^{-2}\,|\,\frac{N-2}{2},\frac{1}{2}(y+z+w_\alpha-X\hat{\beta})'(y+z+w_\alpha-X\hat{\beta}))\,&f_N(\beta\,|\,\hat{\beta},\sigma^2(X'X)^{-1}), \end{aligned} \quad (5)$$

where

$$\hat{\beta}=(X'X)^{-1}X'(y+z+w_{\alpha}),$$

and $f_N(.|a,A)$ and $f_G(.|b,c)$ denote the Normal density with mean vector a and covariance matrix A and the Gamma density with mean b/c and variance b/c^2, respectively. Furthermore, given z, λ is independent of all the data and the other parameters such that:

$$p(\lambda^{-1}|\,Data,z,\alpha,\beta,\sigma^{-2})=p(\lambda^{-1}|\,z)=f_G(\lambda^{-1}|\,N,z'\iota), \qquad (6)$$

where ι is an Nx1 vector of ones. The conditional posterior for z takes the form:

$$p(z|\,Data,\alpha,\beta,\lambda^{-1},\sigma^{-2})\propto f_N(z|\,X\beta-w_{\alpha}-y-\frac{\sigma^2}{\lambda}\iota,\sigma^2 I_N)\prod_{i=1}^{N}I(z_i\geq 0) \qquad (7)$$

where I(.) is the indicator function and I_N is the NxN identity matrix.

Given β, σ^{-2} and z we have a nonlinear regression model in α, which leads to:

$$\begin{aligned}&p(\alpha|\,Data,z,\beta,\lambda^{-1},\sigma^{-2})=p(\alpha|\,Data,z,\beta,\sigma^{-2})\\&\propto p(\alpha)\exp[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i-x_i'\beta+z_i+\ln(w_i'\alpha))^2].\end{aligned} \qquad (8)$$

A Gibbs sampler can be set up in terms of conditional densities (5), (6), (7) and (8).[9] Note that, with the exception of (8), it is easy to take random draws from all of these densities. To take draws from (7) we use the truncated Normal random number generator suggested in Geweke (1991), while (5) and (6) involve Normal and Gamma distributions only.

It remains to discuss random number generation from (8), which does not take the form of any standard density. To this end, we set up an independence Metropolis algorithm (see Tierney (1991) for a theoretical discussion and Chib and Greenberg (1992) for an application). Like the Gibbs sampler, the Metropolis algorithm, originally proposed by Metropolis *et al.* (1953), is based on a Markov chain. We use a special case of the Metropolis implementation in Hastings (1970). A Markovian transition kernel

[9] In practice we draw in the order (8), (5), (6) and (7).

drives the chain by generating candidate values for the next draw. These candidates are then either accepted with a certain probability, or rejected, in which case the chain remains at the current value. The independence Metropolis chain draws candidates independently and always from the same density, $\theta(.)$. So, on the ith pass, this algorithm generates a candidate, α^*, from $\theta(\alpha)$. The random draw from (8), α^i, is then either α^* or α^{i-1} with a certain probability. If the procedure stays at the same value for α over several passes, the value acquires more and more weight. As a consequence, the algorithm will generate a serially correlated sample from (8). Tierney stresses that this method works best if $\theta(\alpha)$ is a good approximation to the actual posterior.

Since equation (8) takes the form of a nonlinear regression model in α, we let $\theta(\alpha)$ be a multivariate Student-t distribution truncated to the nonnegative orthant.[10] The mean, α_0, and covariance matrix, Ω, of the underlying Student-t distribution are calculated using an approximate generalized least squares procedure.[11] Thus, we take candidate draws from

$$\theta(\alpha) \propto p(\alpha)\, f_S^k(\alpha \mid \nu, \alpha_0, A),$$

where $p(\alpha)$ is our prior for α which ensures nonnegativity and $f_S^k(.|\upsilon,\alpha_0,A)$ is the k-variate Student-t density with υ degrees of freedom, location vector α_0 and precision matrix A (ie. covariance matrix $\Omega=(\upsilon/\upsilon-2)A^{-1}$). Moreover, we denote by $\tau(\alpha)$ the ratio of the kernels of $p(\alpha|Data,z,\beta,\sigma^{-2})$ and $\theta(\alpha)$:

[10] We use the algorithm described in Geweke (1991) to draw from this truncated distribution.

[11] The approximate GLS procedure is based on the fact that $y^*=\exp(X\beta-y-z)$ is approximately linear in α. In particular,

$$y^* = w'\alpha \exp(v) \cong w'\alpha\,(1+v) = w'\alpha + v^*,$$

where v^* is Normal with standard deviation $w'\alpha\sigma$. Hence, given starting values for β and z, we can use two- or three- step GLS to obtain an estimate for α as well as a covariance matrix. α_0 and Ω are then set equal to these values.

$$\tau(\alpha)=\frac{\exp[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i-x_i'\beta+z_i+\ln(w_i'\alpha))^2]}{[1+\frac{1}{v}(\alpha-\alpha_0)'A(\alpha-\alpha_0)]^{-\frac{v-k}{2}}}.$$

The degrees of freedom parameter, υ, is chosen on the basis of computational considerations. It is important that $\theta(\alpha)$ have tails at least as fat as (8) to avoid the algorithm getting stuck at a tail value for α with very high $\tau(\alpha)$. In practice, we set $\upsilon=3$. Our independence Metropolis algorithm for drawing α^i can then be defined as follows:

Step 1: Take a draw, α^*, from $\theta(\alpha)$.

Step 2: Calculate $K=\tau(\alpha^*)/\tau(\alpha^{i-1})$.

Step 3: Take a draw, u, from the uniform (0,1) distribution.

Step 4: If K>u then $\alpha^i=\alpha^*$, else $\alpha^i=\alpha^{i-1}$.

In practice, this algorithm works quite well, provided α_0, A and υ are chosen according to the strategy outlined above. The Gibbs sampler seems much less sensitive to choice of initial values for β, σ^{-2} and λ^{-2}. Tierney (1991) mentions the combination of Gibbs and Metropolis steps in a single Markov chain strategy, as used here, and notes that such a hybrid chain is uniformly ergodic provided $\tau(\alpha)$ is bounded, which leads to the strongest form of convergence.

Section 4: Empirical Results

The application discussed in this section is the same as that analyzed in BKOS, KSO and Greene (1990), who provides a complete listing of the data. The data set contains observations from N=123 electric utility companies in the United States in 1970. In addition to output and cost figures, the data set contains information on three input prices: labor price (p_1), capital price (p_2) and fuel price (p_3).

The version of the Gibbs sampler we adopt is that described in Gelfand and Smith (1990). That is, instead of starting the Gibbs sampler and then taking one long run, we take several shorter runs each starting at the same initial values. We carry out M runs, each containing L passes, and keep only the Lth pass out of these runs.

The issue of whether to use one long run from the Gibbs sampler (sequential Gibbs sampling) or to restart every Lth pass (parallel Gibbs sampling) has been discussed in the literature (see, for example, Tanner (1991), Carlin *et al.* (1992), Casella and George (1992), Gelman and Rubin (1992) and Raftery and Lewis (1992)). The question of which variant is preferable is no doubt a problem-specific one, but in our application, the restarting method was found to work best. As in KSO, restarting is required to prevent the path from becoming "temporarily trapped in a nonoptimal subspace" (Tanner (1991), p. 91; see also Zeger and Karim (1991) and Gelman and Rubin (1992)).

We set starting values for α in AIM(1) to α_0 based on the GLS procedure explained in footnote 11 and choose posterior means from KSO as starting values for the other parameters and $r_i=\exp(-z_i)$, $i=1,..,N$. For AIM(2) we start α at the posterior means from the AIM(1) model for $\alpha_1,..,\alpha_6$ and at zero for the other elements of α, retaining KSO starting values for the rest of the parameters.

It is also important to evaluate the accuracy of our Gibbs sampling methods. To this end, we present numerical standard errors (NSEs) calculated using the formula given in Geweke (1992). This formula involves the use of spectral methods for which we use a Parzen window with truncation point $2M^{1/2}$.[12]

Results with M=2500 and L=50 are given in Tables 1 and 2 and Figure 1. Table 1 contains posterior means and standard deviations of all the parameters along with NSEs corresponding to the means. In Table 2, results from KSO (based on 10,000 draws with L=5) are added for comparison. Since our focus is on the efficiency measures, Table 1 can be dealt with quickly. For present purposes it its sufficient to note that: i) NSEs are very small; hence our estimates are quite accurate and RNEs (not

[12]For a more thorough discussion of the practical details required to implement our Gibbs sampler, the reader is referred to KSO.

presented here) indicate that the restarting has partly broken the positive serial correlation of the draws, leading to even higher efficiencies than i.i.d. sampling from the posterior for all the parameters except α. Due to the positive correlation inherent in the independence Metropolis chain for α, its RNEs are somewhat lower. ii) The results of AIM(1) and AIM(2) are very similar, indicating that we need not proceed with higher order approximations (ie. $q>2$). iii) Posterior moments of common parameters clearly indicate substantial differences between the AIM specification used here and that used in KSO.

Note that returns to scale (RTS) can vary across firms. The posterior means of RTS for the minimum, median and maximum output firms are 2.86, 1.04 and .85, respectively, for the AIM(1) specification. The corresponding numbers for AIM(2) are 2.74, 1.04 and .87. These results indicate that average sized firms tend to exhibit roughly constant RTS while small (large) firms tend to exhibit increasing (decreasing) RTS.

Table 2 and Figure 1 present evidence on the efficiency measures. We define $r_i=\exp(-z_i)$ as our measure of firm specific efficiency and r_f as the efficiency of a hypothetical average unobserved firm (see BKOS for details). Table 2 presents posterior moments of r_f and r_i ($i=1,..,5$) for the AIM(1) and AIM(2) expansions. Results from KSO are reproduced at the bottom of the table for comparative purposes. Two major findings are immediately apparent: using the AIM model, efficiency measures are much closer to full efficiency than in KSO as well as showing much less variation over firms. These results are reinforced in Figure 1 which plots r_f for all three models considered. Finally, we note that AIM(1) and AIM(2) lead to virtually identical inference on efficiencies.

It cannot be overemphasized that the results presented in this section are not caused by overfitting, since the restrictions imposed on our cost frontier ensure that it is globally regular. Hence our frontier only fits that part of the data that is globally regular.[13] In fact, the measure of fit described in Appendix A indicates that the cost function used in KSO and the AIM(2) here fit the data equally well, but give very different inferences about the relative importance of symmetric and asymmetric error components.[14]

Insofar as they hold in other data sets, the findings of this section convey a serious warning to empirical researchers working with stochastic frontier models. Estimated efficiency measures are found to be quite sensitive to the choice of functional form for the frontier. Furthermore, it is found that, for our AIM model, which is very flexible but globally regular, inefficiencies are quite small. This suggests that it is not unlikely that previous measures of efficiency reported in empirical studies do not reflect efficiency at all; but rather, merely misspecification of the frontier.

Section 5: Conclusions

This paper carries out a Bayesian analysis of the AIM cost function with composed error. Two important contributions to the existing literature are made: 1) On a theoretical level, the paper develops a Gibbs sampler for analyzing the AIM cost function with non-constant returns to scale and with composed error. It emphasizes that the techniques developed can be easily extended to other nonlinear models as well as to models without composed error. 2) Empirical results presented here indicate that measured efficiencies can be very sensitive to the choice of functional form for the frontier. In fact, the cost function

[13]That is, we assume that the part of the data that is globally regular is the frontier, while any remainder is allocated to the composed error. It is always necessary to make such an assumption in composed error models.

[14]This measure of fit is .0302 for AIM(1), .0301 for AIM(2) and .0301 for the KSO specification. The proportion of this measure that comes from measurement error (v_i) is .9049 for AIM(1), .8972 for AIM(2), but only .4426 for the KSO specification. Thus, although our different frontiers fit the data equally well, the KSO specification allocates much more of the residual to inefficiency (and much less to measurement error) than do the AIM specifications.

based on the AIM aggregator converges to a very different frontier from that based on the Cobb-Douglas aggregator. This latter finding should be a warning to researchers working with stochastic frontier models.

References

Aigner, D., C.A.K. Lovell, and P. Schmidt (1977). "Formulation and estimation of stochastic frontier productions function models," **Journal of Econometrics**, 6, 21-37.

Barnett, W., J. Geweke, and M. Wolfe (1991a). "Seminonparametric Bayesian estimation of consumer and factor demand models," in **Equilibrium Theory and Applications: Proceedings of the Sixth International Symposium in Economic Theory and Econometrics**, W. A. Barnett, B. Cornet, C. d'Aspermont, J. Gabszewicz and A. Mas-Colell (eds.), Cambridge University Press.

Barnett, W., J. Geweke, and M. Wolfe (1991b). "Seminonparametric Bayesian estimation of the Asymptotically Ideal Production Model," **Journal of Econometrics**, 49, 5-50.

Bauer, P.W. (1990). "Recent developments in the econometric estimation of frontiers," **Journal of Econometrics**, 46, 39-56.

van den Broeck, J., G. Koop, J. Osiewalski, and M. Steel (1993). "Stochastic frontier models: A Bayesian perspective," **Journal of Econometrics**, forthcoming.

Carlin, B., N. Polson and D. Stoffer (1992). "A Monte Carlo approach to nonnormal and nonlinear state-space modelling," **Journal of the American Statistical Association**, 87, 493-500.

Casella, G. and E. George (1992). "Explaining the Gibbs sampler," **The American Statistician**, 46, 167-174.

Chib, S. and E. Greenberg (1992). "Bayesian inference in regression models with AR(p) and MA(q) errors," manuscript.

Gallant, A. R. (1981). "On the bias in flexible functional forms and an essentially unbiased form: The Fourier flexible form," **Journal of Econometrics**, 15, 211-245.

Gallant, A. R. and G. Gollub (1984). "Imposing curvature restrictions on flexible functional forms." **Journal of Econometrics**, 26, 295-321.

Gelfand, A.E. and A.F.M. Smith (1990). "Sampling-based approaches to calculating marginal densities," **Journal of the American Statistical Association**, 85, 398-409.

Gelman, A. and D. Rubin (1992). "A Single series from the Gibbs sampler provides a false sense of security," in **Bayesian Statistics 4**, J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith (eds.), Oxford University Press, Oxford.

Geweke, J. (1991). "Efficient simulation from the multivariate Normal and Student-t distributions subject to linear constraints," in **Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface**, E.M. Keramidas and S.M. Kaufman (eds.), Interface Foundation of North America.

Geweke, J. (1992). "Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments," in **Bayesian Statistics 4**, J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith (eds.), Oxford University Press, Oxford.

Greene, W. H. (1990). "A gamma-distributed stochastic frontier model," **Journal of Econometrics**, 46, 141-163.

Hastings, W.K. (1970). "Monte Carlo sampling methods using Markov chains and their applications," **Biometrika**, 57, 97-109.

Koop, G. and K. Carey (1992). "Using seminonparametric methods to model hospital cost functions: The multi-product Asymptotically Ideal Model," manuscript.

Koop, G., M.F.J. Steel and J. Osiewalski (1992). "Posterior analysis of stochastic frontier models using Gibbs sampling," Universidad Carlos III working paper 9245.

Meeusen, W. and J. van den Broeck (1977). "Efficiency estimation from Cobb-Douglas production functions with composed error," **International Economic Review**, 8, 435-444.

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller (1953). "Equations of state calculations by fast computing machines," **Journal of Chemical Physics**, 21, 1087-1091.

Raftery, A. and S. Lewis (1992). "How many iterations in the Gibbs sampler?" in **Bayesian Statistics 4**, J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith (eds.), Oxford University Press, Oxford.

Smith, A.F.M. and A.E. Gelfand (1992). "Bayesian statistics without tears: A sampling-resampling perspective," **The American Statistician**, 46, 84-88.

Tanner, M. (1991). **Tools for Statistical Inference**, Springer-Verlag.

Terrell, D. (1992). "Incorporating regularity conditions in flexible functional forms," **Econometric Reviews**, forthcoming.

Tierney, L. (1991). "Exploring posterior distributions using Markov chains," in **Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface**, E.M.

Keramides and S.M. Kaufman (eds.), Interface Foundation of North America.

Zeger, S. and M. Karim (1991). "Generalized linear models with random effects: A Gibbs sampling approach," **Journal of the American Statistical Association**, 86, 79-86.

Appendix A: Measuring Fit in Composed Error Models

The actual deviation from the theoretical frontier is $\varepsilon_i = v_i + z_i$, where v_i and z_i have the properties described in the body of the paper. Thus, a natural sampling characteristic of fit is

$$E(\varepsilon_i^2 \mid \sigma^2, \lambda) = E(v_i^2 \mid \sigma^2) + E(z_i^2 \mid \lambda)$$
$$= \sigma^2 + var(z_i \mid \lambda) + E^2(z_i \mid \lambda) = \sigma^2 + 2\lambda^2 .$$

From a Bayesian standpoint, this is a random variable, so that we use as a measure of fit the posterior expectation:

$$E(\sigma^2 + 2\lambda^2 \mid Data) = E(\sigma^2 \mid Data) + 2[var(\lambda \mid Data) + E^2(\lambda \mid Data)] .$$

All the quantities in the equation above can easily be calculated in our Gibbs sampling procedure.

The same posterior measure of fit can be obtained in a predictive Bayesian fashion. Consider an unobserved (forecasted) firm, for which the deviation is $\varepsilon_f = v_f + z_f$. Assuming independent sampling, the posterior expected squared deviation for this unobserved firm is

$$E(\varepsilon_f^2 \mid Data) = \int_0^{\infty}\int_0^{\infty} E(\varepsilon_f^2 \mid \sigma^2, \lambda) p(\sigma^2, \lambda \mid Data) \, d\sigma^2 d\lambda$$
$$= E(\sigma^2 + 2\lambda^2 \mid Data) .$$

It should be stressed that this is not the same measure of fit as that used in BKOS. All the models in that paper used the same theoretical frontier and, hence, ignored the systematic part of the deviation. The measure used in BKOS is given by

$$TV_f = var(\varepsilon_f \mid Data) = E(\varepsilon_f^2 \mid Data) - E^2(z_f \mid Data) .$$

Table 1: Posterior Moments of Parameters

		AIM(1)			AIM(2)	
	Mean	NSE	St Dev	Mean	NSE	St Dev
β_1	.241	.001	.052	.257	8.16E-4	.043
β_2	.041	7.18E-5	.004	.040	5.80E-5	.003
σ^{-2}	37.471	.095	5.813	37.864	.112	5.721
λ^{-1}	41.462	.507	30.069	39.891	.602	33.080
σ^2	.027	----	.004	.027	----	.004
λ	.034	----	.018	.035	----	.018
α_1	1.93E-6	1.32E-7	1.39E-6	1.39E-6	1.48E-8	7.60E-7
α_2	1.68E-4	8.26E-6	1.19E-4	1.21E-4	1.47E-6	6.69E-5
α_3	7.83E-4	3.87E-5	4.63E-4	6.78E-4	5.81E-6	3.23E-4
α_4	2.08E-5	1.43E-6	1.75E-5	1.43E-5	1.63E-7	8.24E-6
α_5	1.31E-4	4.65E-6	5.14E-5	7.89E-5	1.35E-6	5.55E-5
α_6	5.54E-4	3.20E-5	3.56E-4	3.77E-4	5.16E-6	2.18E-4
α_7	----	----	----	1.25E-6	4.29E-8	2.11E-6
α_8	----	----	----	2.08E-6	6.68E-8	3.58E-6
α_9	----	----	----	5.60E-5	2.12E-6	9.49E-5
α_{10}	----	----	----	1.10E-5	3.97E-7	1.90E-5
α_{11}	----	----	----	5.85E-5	2.11E-6	9.57E-5
α_{12}	----	----	----	1.61E-4	6.11E-6	2.67E-4
α_{13}	----	----	----	6.47E-6	2.24E-7	1.12E-5
α_{14}	----	----	----	1.94E-5	6.05E-7	3.28E-5
α_{15}	----	----	----	3.45E-5	1.16E-6	5.77E-5

Table 2: Posterior Moments of Efficiency Measures

	r_f	r_1	r_2	r_3	r_4	r_5
AIM(1)						
Mean	.968	.972	.979	.964	.972	.976
NSE	2.99E-4	----	----	----	----	----
St Dev	.038	.033	.022	.041	.032	.026
AIM(2)						
Mean	.967	.970	.979	.962	.970	.973
NSE	3.06E-4	----	----	----	----	----
St Dev	.039	.033	.022	.043	.032	.028
KSO*						
Mean	0.917	0.730	0.973	0.943	0.926	0.963
NSE	7.4E-4	----	----	----	----	----
St Dev	0.077	0.111	0.026	0.047	0.058	0.034

*KSO indicates results from the exponential model with linear frontier analyzed in KSO.

FIGURE 1: POSTERIOR EFFICIENCIES

2500 passes through the Gibbs sampler