

Working Paper 92-45
Statistics and Econometrics Series 31
December 1992

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

POSTERIOR ANALYSIS OF STOCHASTIC FRONTIER MODELS USING GIBBS SAMPLING

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

Abstract

In this paper we describe the use of Gibbs sampling methods for making posterior inferences in stochastic frontier models with composed error. We show how the Gibbs sampler can greatly reduce the computational difficulties involved in analyzing such models. Our findings are illustrated in an empirical example.

Key Words

Composed Error Models; Bayesian Econometrics; Gibbs Sampling; Rejection Sampling.

*Koop, Department of Economics, Boston University; Steel, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid, Osiewalski, Department of Econometrics, Academy of Economics, Krakow. We are grateful to John Geweke for providing us with a Fortran implementation of his mixed rejection algorithm for drawing from a truncated Normal distribution.

Introduction

Stochastic frontier models have a long history in empirical economics. They developed out of the reasoning that production functions derived from economic theory should measure the greatest amount of output that can be obtained from a given set of inputs. In practice, firms rarely, if ever, are perfectly efficient, so that actual output produced should lie below the production frontier. It is now standard practice to work with models where the deviation from the theoretical production function is composed of two parts, one of which is negative and interpreted as inefficiency; the other of which is symmetric and interpreted as measurement error. This composed error framework was first introduced in Meeusen and van den Broeck (1977) and Aigner, Lovell and Schmidt (1977). The reader is referred to Bauer (1990) for a survey of the literature.

In a previous paper (van den Broeck, Koop, Osiewalski and Steel (1993), hereafter BKOS), we used Bayesian methods to analyze stochastic frontier models and argued that such methods had several advantages over their classical counterparts in the treatment of these models. Most importantly, the Bayesian methods we outlined enabled us to provide exact finite sample results for any feature of interest and to take fully into account parameter uncertainty. In addition, they made it easy to treat uncertainty about which model to use since they recommended taking weighted averages over all models, where the weights were posterior model probabilities. We illustrated the Bayesian approach using an empirical example described in Greene (1990), and posterior properties were evaluated using Monte Carlo integration with importance sampling. Unfortunately, the empirical approach was computationally very demanding and we were led to conclude that, at least for larger problems, Monte Carlo integration was not a feasible method. In this paper, we describe how Gibbs sampling can be used to substantially reduce the computational demands encountered in our Bayesian approach.¹

¹A third alternative to Gibbs sampling or Monte Carlo integration with importance sampling would be simple Monte Carlo integration. Ley and Steel (1992) use rejection methods to draw

The paper is organized as follows. Section 1 introduces the Bayesian model we consider in the paper and Section 2 develops the Gibbs sampler. Section 3 applies the Gibbs sampler to the same illustrative example we used in BKOS. We find that an iteratively calibrated Gibbs sampler with restarting performs very well in such a numerically complicated problem.

Section 1: The Model

The basic stochastic frontier model which we use is given by:

$$y_i = X_i' \beta + v_i - z_i, \quad (1)$$

where y_i is the log of output (or the negative of log cost) for firm i ($i=1, \dots, N$); X_i is a vector of k exogenous variables; v_i is a symmetric disturbance capturing measurement error in the stochastic frontier; z_i is a nonnegative disturbance capturing the level of firm inefficiency; and z_i and v_i are independent of each other and across firms. We use the notation y , z , and v to indicate vectors containing all the y_i s, z_i s, and v_i s; and X to indicate the matrix of observations on the explanatory variables. Moreover, we assume that v_i is $N(0, \sigma^2)$. Four different specifications for z_i are then considered, three corresponding to different Gamma distributions and one to a truncated Normal. We allow for z_i to have a Gamma distribution with integer shape parameter $j=1, 2, 3$, and refer to these models as M_j ($j=1, 2, 3$). This implies

$$p_j(z_i | \theta_j) = f_G(z_i | j, \lambda^{-1}) = \frac{z_i^{j-1}}{\lambda^j \Gamma(j)} \exp(-\lambda^{-1} z_i) I(z_i \geq 0), \quad (2)$$

where $\theta_j = (\beta', \sigma^{-2}, \lambda^{-1})'$; $I(\cdot)$ is the indicator function; $\Gamma(\cdot)$ is

directly from the posterior for a version of our model with no measurement error. However, when measurement error is allowed for, these rejection methods can no longer be used.

²We parameterize in terms of σ^{-2} and λ^{-1} rather than σ^2 and λ . This choice has no substantive implications; it merely ensures that we can draw from Gamma densities instead of inverted Gamma densities.

the gamma function; and $f_G(\cdot|j, \lambda^{-1})$ indicates the Gamma density with parameters j and λ^{-1} ($j=1,2,3$). Note that the $j=1$ case corresponds to the exponential distribution. The fourth model, M_4 , assumes z_i has a truncated Normal density. That is,

$$p_4(z_1|\theta_4) = p_4(z_1|\psi, \omega^{-2}) = \frac{f_N^1(z_1|\psi\omega, \omega^2)}{\Phi(\psi)} I(z_1 \geq 0), \quad (3)$$

where $\theta_4 = (\beta', \sigma^{-2}, \psi, \omega^{-2})'$; $\Phi(\cdot)$ is the cumulative Normal; and $f_N^M(\cdot|\cdot)$ is the M -variate Normal density. In our case this underlying Normal distribution has mean $\psi\omega$ and variance ω^2 .³

The focus of interest in our empirical illustration is in making posterior inferences about the parameters and z . The derivation of the posterior under M_j ($j=1,2,3,4$) is given in BKOS. For the Gamma models (M_1, M_2 and M_3), we work with the diffuse prior, $p(\theta_j) \propto \sigma^2 \lambda$ ($j=1,2,3$), but the techniques described are also valid for the more general classes of priors considered in BKOS. Although the posteriors for θ_j and z are very difficult to work with, conditional posterior densities are much simpler. The next section describes how a Gibbs sampler can be set up in terms of the following conditional densities:

$$p_j(\beta, \sigma^{-2} | \lambda^{-1}, y, z, X) = p_j(\beta, \sigma^{-2} | y, z, X) = f_N^k(\beta | \beta, \sigma^2 (X'X)^{-1}) f_G(\sigma^{-2} | \frac{N-k}{2}, \frac{(y+z-X\beta)'(y+z-X\beta)}{2}), \quad (4)$$

where

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

$$p_j(\lambda^{-1} | \beta, \sigma^{-2}, y, z, X) = p_j(\lambda^{-1} | z) = f_G(Nj, z'1), \quad (5)$$

where 1 is an $N \times 1$ vector of ones; and

³We find this parameterization of the Normal to be useful in that the posterior of ψ is better behaved than that of the mean of the Normal.

$$p_j(z|\beta, \sigma^{-2}, \lambda^{-1}, y, X) \propto f_N^N(z|X\beta - y - \frac{\sigma^2}{\lambda}, \sigma^2 I_N) \prod_{i=1}^N z_i^{j-1} I(z_i \geq 0), \quad (6)$$

where I_N is the $N \times N$ identity matrix.

A Gibbs sampler for M_4 can be derived in a similar way. However, we use an informative prior for M_4 ⁴ which is described in more detail in BKOS. It is sufficient merely to note here that this prior is a modified version of the usual natural conjugate prior, and is given by:

$$p_4(\psi|\omega^{-2}) = p_4(\psi) = 2f_N^1(\psi|0, a)\Phi(\psi), \quad (7)$$

and

$$p_4(\omega^{-2}) = f_G(\omega^{-2}|\frac{v_0}{2}, \frac{b}{2}), \quad (8)$$

where a , b and v_0 are prior hyperparameters. In this paper we set $a=1$, $v_0=10$ and $b=.1783$. A detailed justification of these choices, which imply the prior median of our measure of efficiency to be .875, is given in BKOS. For the remaining parameters, we use a standard noninformative prior: $p(\beta, \sigma^{-2}) \propto \sigma^2$, independent of ψ and ω^{-2} .

With this prior, conditional densities for M_4 can now be created in a relatively straightforward fashion. Note first that, in equation (4), the posterior for β and σ^{-2} given z does not depend on λ^{-1} , and hence, that it is easy to show that the posterior takes exactly the same form for M_4 as for M_j ($j=1,2,3$). It can also be shown that:

$$p_4(z|\beta, \sigma^{-2}, \psi, \omega^{-2}, y, X) \propto f_N^N(z|\frac{\sigma^2\psi\omega^{-1} - \omega^2(y - X\beta)}{\omega^2 + \sigma^2}, \frac{\sigma^2\omega^2}{\omega^2 + \sigma^2} I_N) I(z_i \geq 0, i=1, \dots, N), \quad (9)$$

and

⁴The posterior under M_4 behaves very badly when a noninformative prior is used. For this reason, we do not present results for this case either here or in BKOS.

$$\begin{aligned}
p_4(\psi, \omega^{-2} | \beta, \sigma^{-2}, y, z, X) = \\
p_4(\psi, \omega^{-2} | z) = f_N^1(\psi | \frac{a}{1+Na} \frac{\sum z_i}{\omega}, \frac{a}{1+Na}) \quad (10) \\
f_G(\omega^{-2} | \frac{N+v_0}{2}, \frac{\sum z_i^2 - \frac{a}{1+Na} (\sum z_i)^2 + b}{2}).
\end{aligned}$$

In this paper, we are interested in making marginal posterior inferences about z and θ and, for this reason, have no direct interest in the conditional densities listed above. Nonetheless, the next section describes how such marginal inferences can be made on the basis of (4), (5), (6), (9) and (10).

Section 2: The Gibbs Sampler

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

A glance at equations (6), (4) and (5) (and (9), (4) and (10) for M_4) confirms that the stochastic frontier model considered in this paper can be analyzed using Gibbs sampling methods. That is, even though the posteriors of θ_j and z are unwieldy, the conditionals for a suitable partition of the set of parameters are easy to work with. By taking successive draws from (6), (4) and (5) (or (9), (4) and (10)), each conditional on previous draws from the other conditional densities, we can create a sample which can be treated as random from the joint posterior.

Note that, in the application in Section 3, $N=123$ and $k=5$, so that the dimension of the Gibbs sampler is 130 for M_1 , M_2 and M_3 and 131 for M_4 . However, the partition we choose implies that only three steps are required for each pass through the sampler. Any arbitrary function of interest, $g(\theta_j, z)$, can be approximated by its sample mean, g^* , based on S passes. Moreover, numerical standard errors (NSEs) can also be calculated in order to quantify the approximation error. Geweke (1992) recommends estimating the NSE of g^* by $[S^{-1}S_g(0)]^{1/2}$, where $S_g(0)$ is an estimate

of the spectral density of the series $g(\theta_j^{(i)}, z^{(i)})$ ($i=1, \dots, S$) at frequency zero, and $\theta_j^{(i)}$ and $z^{(i)}$ are the sampled values. Another useful diagnostic is relative numerical efficiency (RNE). Geweke (1992) shows that RNE can be approximated by the ratio $\text{var}^*(g(\theta_j, z))/S_g(0)$, where $\text{var}^*(g(\theta_j, z))$ is the Gibbs sampling estimate of $\text{var}(g(\theta_j, z))$. RNE constitutes a good measure of computational efficiency in that the number of drawings necessary to achieve a given degree of numerical accuracy is inversely related to the RNE. In fact, RNE is the efficiency relative to i.i.d. drawings from the posterior. We use a Parzen window with truncation point at $2S^*$ in our calculation of $S_g(0)$ in Section 3.

To implement the Gibbs sampler for all the models, we must take pseudo-random draws from conditional densities (4), (5), (6), (9) and (10). It is trivial to take draws from the densities given in (4), (10) and (5) since the first two of these are Normal-Gamma; and the last, Gamma. Taking random draws from the densities given in (6), however, is somewhat more difficult. Due to our assumption of independence across firms, it is only necessary that we take draws from

$$p_j(z_i | y_i, X_i, \beta, \lambda^{-1}, \sigma^{-2}) \propto z_i^{j-1} f_N^1(z_i | X_i \beta - y_i - \frac{\sigma^2}{\lambda}, \sigma^2 I_N) I(z_i \geq 0). \quad (11)$$

We consider two separate strategies for drawing from (11), one of which works only for M_1 and M_4 ; the other, for M_j ($j=1, 2, 3, 4$).

If $j=1$ then (11) is a univariate Normal density truncated to be nonnegative. Simulation from this density is simple using the algorithm described in Geweke (1991). Note that (9) is also truncated Normal so that Geweke's algorithm can be used for M_4 as well. Note also, however, that (11) does not take the form of any standard density for either M_2 or M_3 . Nevertheless, we can sample from (11) for M_2 or M_3 using rejection methods⁵ found in Smith and Gelfand (1992). To obtain a random sample from a posterior,

⁵It is, of course, possible to use rejection methods for M_1 and M_4 as well as for M_2 and M_3 . For M_1 and M_4 , such methods were not used for the final runs given in Section 3 since direct simulation from the truncated Normal was about four times as fast.

the authors recommend taking a random sample from the prior and including each draw with a certain probability. In our context, if $\theta^{(l-1)}$ is the drawn value for θ_j from the $(l-1)$ th pass through the Gibbs sampler, the Smith-Gelfand algorithm is as follows:

Step 1. Take a random draw, $z_i^{(1)}$ from $p(z_i | \theta^{(l-1)})$, which is Gamma (see equation (2)).

Step 2. Take a random draw, w , from the uniform on $(0,1)$.

Step 3. Calculate $K = p(y_i | z_i^{(1)}, X_i, \theta^{(l-1)}) / p(y_i | \hat{z}_i, X_i, \theta^{(l-1)})$, where \hat{z}_i maximizes $p(y_i | z_i, X_i, \theta^{(l-1)})$.

Step 4. If $w < K$, then accept $z_i^{(1)}$; else return to Step 1.

This simple algorithm can be used to generate random numbers from (11) and, thus, (6) as well. The only minor drawback is that K must be evaluated for each firm ($i=1, \dots, N$) at each pass ($l=1, \dots, S$). However, this is not a major problem, as

$$p_j(y_i | z_i, X_i, \theta^{(l-1)}) = f_N^1(y_i | X_i' \beta^{(l-1)} - z_i, \sigma^{2(l-1)}).$$

Since z_i is restricted to be nonnegative, \hat{z}_i is either 0 or $X_i' \beta^{(l-1)} - y_i$, depending on the sign of $X_i' \beta^{(l-1)} - y_i$. Hence K can easily be calculated and is large enough in practice to ensure that the algorithm does not reject too many draws.⁶

In this section, we have argued that random draws can easily be taken from the conditional densities given in (4), (5), (6), (9) and (10). It follows that a Gibbs sampler can be set up to

⁶Zeger and Karim (1991) and Carlin et al. (1992) also use rejection methods within the Gibbs sampler. Rather than draw from the prior density, however, they match Gaussian kernels to the conditionals in question. In our case, sampling with rejection from the truncated Normal will not work since K in Step 3 then becomes

$$\frac{\prod_{i=1}^n [z_i^{(1)}]^{j-1}}{\max \prod_{i=1}^n z_i^{j-1}}$$

which is zero as z_i^{j-1} is unbounded for $j=2$ or 3 . Hence, the probability of accepting a draw would be zero.

evaluate the posterior properties of z and θ_j in our stochastic frontier model. It is worth stressing at this stage that our Gibbs sampler, unlike the Monte Carlo approach we outlined in BKOS, yields a draw of z at each pass, and that for this reason, the inefficiency measures for all N firms are produced as a by-product of our Gibbs sampling methodology.

Section 3: The Application

The application discussed in this section is the same as that analyzed in BKOS and in 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. Greene (1990) and BKOS fit a cost function based on the Cobb-Douglas cost function, but which is generalized to allow returns to scale to vary with output. If Q_i , C_i , P_{li} , P_{ki} , and P_{fi} are output, cost, labor price, capital price, and fuel price of firm i , respectively, then the model we work with here is given by:

$$-\ln\left(\frac{C_i}{P_{fi}}\right) = -\beta_0 - \beta_1 \ln Q_i - \beta_2 \ln^2 Q_i - \beta_3 \ln\left(\frac{P_{li}}{P_{fi}}\right) - \beta_4 \ln\left(\frac{P_{ki}}{P_{fi}}\right) + v_i - z_i.$$

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 of S passes, we take several shorter runs each starting at the same initial value (ie. at the same $\theta^{(0)}$). We carry out M runs, each containing L passes, and keep only the L th pass out of these runs. Moreover, we use this implementation in an iterative fashion, starting from an arbitrary initial value and using small Gibbs runs, typically with $L=10$ and $M=500$ or $1,000$, to calibrate the starting value $\theta^{(0)}$.⁷ As soon as convergence is achieved, a final Gibbs run is

⁷That is, after each calibration run we take the posterior means as the starting value for the next run and convergence is achieved if resulting posterior means are close enough (relative difference of no more than 20%) to the starting values of that run. These runs are computationally cheap since NSEs (which require spectral analysis) do not need to be evaluated at this stage.

carried out. $M=10,000$ was used and some experimentation revealed that $L=5$ was more than sufficient.

The issue of whether to use one long run from the Gibbs sampler (often called sequential Gibbs sampling) or to restart every L th pass (often called 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)). Although the theoretical properties of these two variants of the Gibbs sampler are the same, they can, in practice, be quite different. Advocates for taking one long run note that the restarting method typically wastes a lot of information (eg. our methods produce 50,000 passes, but we only use 10,000 of them when actually calculating posterior moments). Advocates for restarting the Gibbs sampler note that this technique can be very effective in breaking serial correlation and prevents 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)). 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. Using one long run without introducing prior information in models M_j ($j=1,2,3$) very often led to a circling in the region of full efficiency and very small λ . For M_4 , undesirable behavior (with the precision blowing up) occurs even with the informative prior in (7) and (8). In any case, the strong serial correlation induced by a single long Gibbs run⁸ produces very small numerical efficiencies: For M_4 the RNEs, without restarting, range from .027 for σ^{-2} to .18 for β_1 . Tables 1 and 2 indicate that RNEs with our restarting method are typically around 1, suggesting that Gibbs sampling is roughly as efficient as is sampling directly from the posterior.

Results using the Gibbs sampling algorithm described above are given in Tables 1, 2 and 3. In Table 1 we present posterior

⁸An alternative way to reduce the serial correlation is to use only every K th pass in a long sequence. However, this would not solve the problem of the path possibly getting stuck. Furthermore, evidence from Raftery and Lewis (1992) and Geweke (1992) supports taking $K=1$ in most cases.

means and standard deviations as well as NSEs and RNEs for the posterior means for all parameters for M_1 , M_2 and M_3 . Table 2 presents the same numbers for M_4 . In Table 3, we focus on measures of efficiency. Since our data is in logs, it makes sense to define the firm-specific efficiency of firm i as $r_i = \exp(-z_i)$. Both posterior means and standard deviations of r_i for $i=1, \dots, 5$,⁹ and of the measure of average efficiency, r_f , are presented.¹⁰

The numbers reported in Tables 1, 2 and 3 are very close to those of BKOS, with some exceptions. The most important of these occur for M_4 (eg. in BKOS we found the posterior means of ψ and ω^{-2} to be 0.41 and 62.80; here these numbers are 1.98 and 145.78, respectively). Even with an informative prior, it was very difficult to find a good importance function for M_4 in BKOS. In contrast, results using the restarted Gibbs sampler are very stable indeed once the prior information in (7) and (8) is introduced. Calibration was fast for all models (2 iterations) and results very insensitive to changes in L and M . As further evidence of the advantages of Gibbs sampling, consider Figures 1 and 2. These figures contain plots of the posterior of r_f for all four models using Gibbs sampling and Monte Carlo integration with importance sampling, respectively. Clearly, the plots in Figures 1 and 2 indicate that Gibbs sampling has vastly superior numerical properties, especially for M_4 . Figure 1 is calculated using only 10,000 drawings (ie. $M=10,000$ and $L=5$), while Figure 2 is calculated on the basis of 100,000 drawings from an importance function. Furthermore, for both methods we initially used maximum likelihood estimates, but for importance sampling we needed many more iterations in order to fine-tune the importance function. Note that the iterations in our Gibbs sampler are only required to fix the location of $\theta^{(0)}$, whereas location, spread and functional form of an importance function have to be chosen for the Monte Carlo approach.

The runs included in this paper involved 50,000 passes

⁹We could easily present posterior moments of r_i for $i=6, \dots, 123$ as well but do not do so for space considerations.

¹⁰Here we define average efficiency as the efficiency of an unobserved firm. See BKOS for details.

through the Gibbs sampler. Execution times for M_1 , M_2 , M_3 and M_4 were 893 minutes, 2388 minutes, 2283 minutes and 960 minutes, respectively, on a 386-33 PC using GAUSS-386 VM version 2.0. Note that M_1 and M_4 were much faster since they did not involve the use of rejection methods.¹¹ By way of comparison with Monte Carlo techniques, note that Figure 2 is based on 100,000 antithetic draws from the importance functions selected for our four models. This final run alone took 616 minutes, 866 minutes, 1166 minutes and 700 minutes for M_1 , M_2 , M_3 and M_4 , respectively, where efficiencies for the first five firms only were calculated.

These execution times do not represent the entire computational burden since, for Monte Carlo, the final importance functions we used were selected on the basis of numerous runs. In particular, we used a strategy where we iteratively updated the importance function on the basis of calculated posterior means of the parameters. This search for an adequate importance function was invariably a long one (for example, a run of 100,000 draws from the final importance function usually followed five or more preliminary runs, each of 50,000 draws). As the iterative calibration of the Gibbs sampler typically converged after 2 iterations with $L=10$ and $M=1,000$, the initial costs of the Gibbs sampler were much less. This makes the total computing cost (and, in particular, the effort of the researcher) required for Figure 2, higher than for Figure 1. In addition, and more importantly, the Gibbs sampling results are much less erratic¹² than are the Monte Carlo results.

So as to really put this comparison into perspective, we

¹¹For M_2 and M_3 rejection methods were used for the 123 elements of z . The average number of drawings per pass were 2198 and 1973 for M_2 and M_3 , respectively. Accordingly, roughly 16 draws for each firm were rejected on average before one was accepted. Acceptance rates tend to go up as L becomes smaller and $\theta^{(0)}$ moves towards convergence in our iterative calibration.

¹²The many modes in Figure 2 are caused by the very high weights being attributed to a small number of drawings. Even with conscientious fine-tuning of the importance function, this numerical instability could not be avoided. For this reason, the binwidth in BKOS was chosen to be 0.04 whereas here it is 0.01 in all figures for critical comparison. For a clear exposition of Monte Carlo importance sampling methods, see Geweke (1989).

also computed NSEs for the Monte Carlo runs. For the truncated Normal model, M_4 , the NSEs obtained with 100,000 drawings from a carefully selected importance function are roughly 10 to 20 times as large as are those reported in Table 2 (based on only 10,000 drawings)! This implies that the efficiency of a Gibbs run is at least 100 times better, even if the number of drawings is much smaller. Indeed, RNEs of the Monte Carlo estimates of posterior means range from .0003 for σ^{-2} to .001 for ω^{-2} . In other words, whereas Gibbs sampling with restarting is as efficient as drawing from the posterior, the Monte Carlo approach using importance sampling requires 1,000 or more draws for every draw directly from the posterior in order to obtain the same numerical accuracy.

We based the comparison of Gibbs and Monte Carlo results on plots of actual drawings. However, based on the Rao-Blackwell theorem, direct use of the knowledge of conditional densities can lead to substantial improvements in our estimates of posterior characteristics (see, eg., Gelfand and Smith (1990), Gelfand, Smith and Lee (1992) and Casella and George (1992)). In obtaining Table 3 we have averaged the conditional posterior moments of r_f over drawn values of θ_j to arrive at marginal moments.¹³ For Figure 3 we have followed the same strategy with the conditional posterior densities of r_f . The plot given in Figure 3 is based on only 1,000 drawings ($M=1,000$, $L=5$) and is extremely stable to changes in M or L . A plot taking a mere hundred drawings (not presented here) is indistinguishable from Figure 3, at a computational cost which ranges from less than 5 minutes for M_1 to less than 20 minutes for M_2 !

Conclusion

In this paper, we have shown how Gibbs sampling methods can be used to greatly reduce the computational burdens which complicate the Bayesian analysis of stochastic frontier models with composed error. We show how the posterior conditional

¹³Results on the basis of the drawn values alone coincide in the first three decimals, indicating exemplary numerical stability.

densities can be used to set up a Gibbs sampler, and use the rejection method suggested in Smith and Gelfand (1992) to take random draws from the conditional of z for M_2 and M_3 . All other conditionals are either truncated Normal, Normal-Gamma or Gamma. An application illustrates the enormous computational gains which can be achieved with the use of Gibbs sampling methods.

Bibliography

- 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.
- 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**, forthcoming.
- 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.
- Gelfand, A., A. Smith and T. Lee (1992). "Bayesian analysis of constrained parameter and truncated data problems using Gibbs sampling," **Journal of the American Statistical Association**, 87, 523-532.
- 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. (1989). "Bayesian inference in econometric models using Monte Carlo integration," **Econometrica**, 57, 1317-1339.
- 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.

- Ley, E. and M. F. J. Steel (1992). "Bayesian econometrics: Conjugate analysis and rejection sampling using *Mathematica*," Chapter 15 in **Economic and Financial Modeling Using Mathematica**, H. Varian (ed.), Springer-Verlag, forthcoming.
- Meeusen, W. and J. van den Broeck (1977). "Efficiency estimation from Cobb-Douglas production functions with composed error," **International Economic Review**, 8, 435-444.
- 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.
- 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.

Table 1: Posterior Moments of Parameters for M_1, M_2, M_3

	β_0	β_1	β_2	β_3	β_4	σ^{-2}	λ^{-1}
M_1							
Mean	-7.48	0.43	0.03	0.25	0.05	78.98	11.55
NSE	2.9E-3	4.0E-4	2.7E-5	6.0E-4	6.0E-4	0.18	0.03
RNE	1.25	1.02	1.02	1.13	1.03	0.97	0.87
S.D.	0.33	0.04	2.7E-3	0.06	0.06	17.97	2.36
M_2							
Mean	-7.37	0.40	0.03	0.25	0.07	62.03	22.41
NSE	3.2E-3	3.9E-4	2.7E-5	6.4E-4	6.3E-4	0.12	0.04
RNE	1.08	1.04	1.05	1.07	0.92	0.98	1.04
S.D.	0.34	0.04	2.7E-3	0.07	0.06	11.85	4.12
M_3							
Mean	-7.33	0.39	0.03	0.25	0.07	56.03	33.51
NSE	3.4E-3	4.1E-4	2.9E-5	6.4E-4	6.0E-4	0.09	0.05
RNE	1.04	0.91	0.91	1.12	1.08	1.12	1.12
S.D.	0.34	0.04	2.7E-3	0.07	0.06	9.03	5.23

Table 2: Posterior Moments of Parameters for M_4

	Mean	NSE	RNE	S.D.
β_0	-7.37	3.3E-3	1.08	0.34
β_1	0.39	3.9E-4	0.94	0.04
β_2	0.03	2.7E-5	0.95	2.7E-3
β_3	0.24	7.0E-4	0.91	0.07
β_4	0.08	6.4E-4	0.94	0.06
σ^{-2}	67.21	0.14	0.89	13.42
ψ	1.98	1.9E-3	1.08	0.19
ω^{-2}	145.78	0.29	1.32	33.76

Table 3: Posterior Moments of Efficiency Measures for all Models

	r_f	r_1	r_2	r_3	r_4	r_5
M_1						
Mean	0.917	0.730	0.973	0.943	0.926	0.963
S.D.	0.077	0.111	0.026	0.047	0.058	0.034
NSE	7.4E-4	---	---	---	---	---
RNE	1.076	---	---	---	---	---
M_2						
Mean	0.913	0.836	0.955	0.927	0.921	0.945
S.D.	0.060	0.088	0.030	0.046	0.049	0.036
NSE	5.8E-4	---	---	---	---	---
RNE	1.047	---	---	---	---	---
M_3						
Mean	0.914	0.870	0.944	0.920	0.918	0.936
S.D.	0.049	0.066	0.031	0.042	0.043	0.034
NSE	4.4E-4	---	---	---	---	---
RNE	1.230	---	---	---	---	---
M_4						
Mean	0.843	0.782	0.925	0.859	0.856	0.903
S.D.	0.070	0.063	0.048	0.060	0.060	0.053
NSE	6.8E-4	---	---	---	---	---
RNE	1.048	---	---	---	---	---

Notes to all Tables

1) Mean, NSE and RNE indicate the posterior mean, corresponding numerical standard error and relative numerical efficiency; while S.D. indicates standard deviation.

2) Results for M_1 , M_2 and M_3 are derived under the noninformative prior described in Section 1; and for M_4 , under the informative prior described in Section 1 (ie. with $a=1$, $\nu_0=10$ and $b=.1783$).

FIGURE 1: POSTERIOR EFFICIENCIES
Gibbs sampling with 10,000 drawings

..... j=1 - - - - j=2 - - - - j=3 ——— j=4 ——— inf.

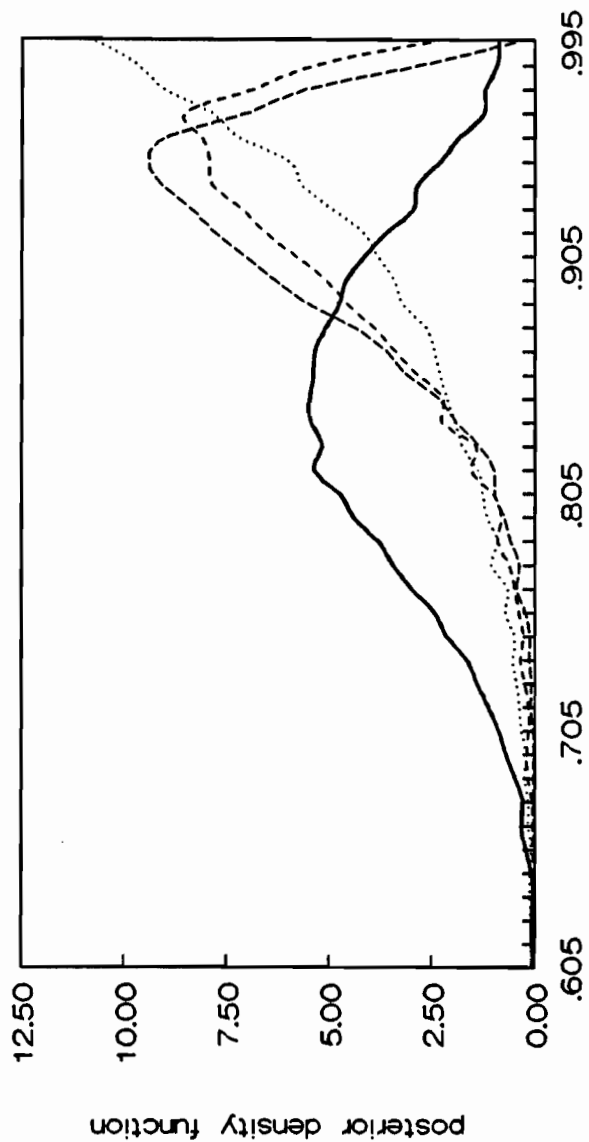


FIGURE 2: POSTERIOR EFFICIENCIES

Monte Carlo with 100,000 drawings

..... j=1 - - - - j=2 - - - - j=3 ——— j=4 ——— inf.

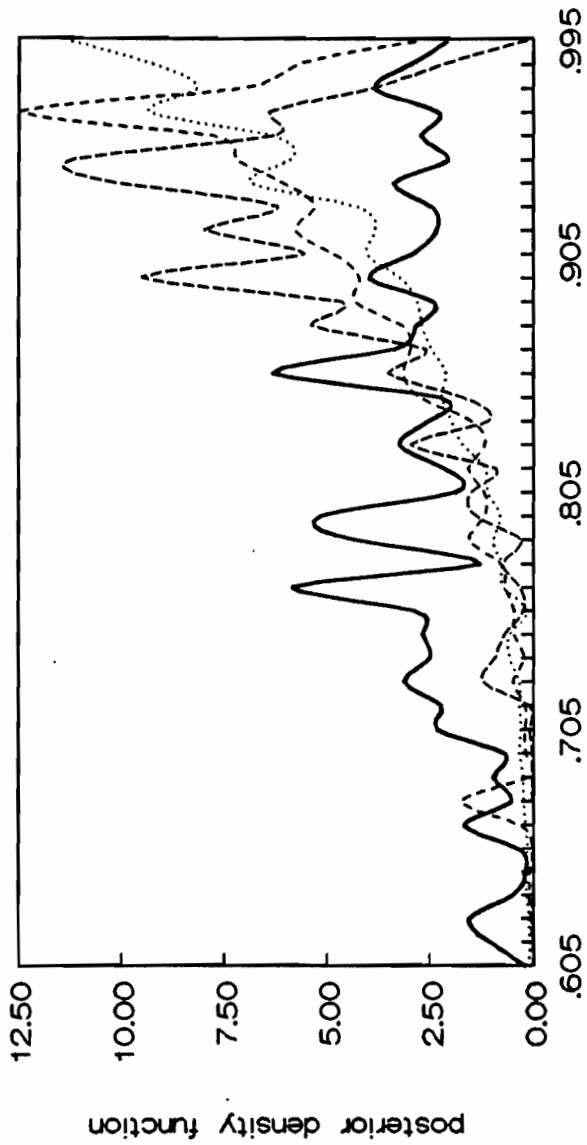


FIGURE 3: POSTERIOR EFFICIENCIES

Gibbs using conditionals; 1000 drawings

..... j=1 - - - - j=2 - - - - j=3 — j=4 inf.

