Identifiability of the Stochastic Frontier Models

Bandyopadhyay, Debdas and Das, Arabinda
University of Kalyani

June 2007

Identifiability of the Stochastic Frontier Models

Debdas Bandyopadhyay[1] & Arabinda Das[2]
Department of Statistics
University of Kalyani
Kalyani-741 235
India

Abstract

This paper examines the identifiability of the standard single-equation stochastic frontier models with uncorrelated and correlated error components giving, *inter alia*, mathematical content to the notion of "near-identifiability" of a statistical model. It is seen that these models are at least locally identifiable but suffer from the "near-identifiability" problem. Our results also highlight the pivotal role played by the Signal to Noise Ratio in the "near-identifiablity" of the stochastic frontier models.

Keywords: Identification, Stochastic frontier model, Information Matrix, Signal to Noise Ratio

JEL Classification: C3, C10, C52

AMS Classification: 91B70, 62F10, 62B10

[1]Address for correspondence: Debdas Bandyopadhyay, Department of Statistics, University of Kalyani, Kalyani-741 235, West Bengal, India. E-Mail: debdas.banerjee@gmail.com
[2]Arabinda Das, Department of Statistics, University of Kalyani, Kalyani-741 235, West Bengal, India. E-mail: ara.das@gmail.com.
Research of the second author is supported by a fellowship from Council for Scientific and Industrial Research, Govt. of India.

Identifiability of the Stochastic Frontier Models

Debdas Bandyopadhyay[3] & Arabinda Das[4]
Department of Statistics
University of Kalyani
Kalyani-741 235
India

1. Introduction

Ever since its introduction, the stochastic frontier model (Aigner, Lovell and Schmidt, 1977; Meeusen and van den Broeck, 1977 and Battese and Corra, 1977), hereafter SFM, has been extensively used for likelihood-based statistical inference regarding the firm level productive inefficiency (See, among others, Kumbhakar and Lovell, 2000 for an excellent introduction to the stochastic frontier literature). Likelihood-based inference, however, is possible only if the model is identifiable (for definition and characterization of the identification problem associated with different statistical models, see, among others, Rao, 1992 and references therein). When the model is not identifiable there are at least two different models (probability structures) with *exactly* equal likelihood of generating the sample observations. This makes the likelihood-based inference of unidentifiable models logically invalid, as the models can no longer be discriminated using the likelihood function. A more frequently encountered problem in statistics and econometrics, however, is the problem of "near-identification". In this case there are two or more models with *approximately equal* likelihood of generating the sample. For a near- identifiable model, the likelihood-based inference is logically valid but the resulting estimates are imprecise and unstable as the information matrix of such a model, though non-singular, is near-singular. For example, though the maximum likelihood estimates of the parameters of a near-identifiable model have the usual optimal asymptotic properties, the estimates are imprecise (large asymptotic variance) and unstable (highly sensitive to small change in sample). Near-exact multicollinearity is a classic example of near-identifiability of a statistical model. We may, however, note that although the

[3]Address for correspondence: Debdas Bandyopadhyay, Department of Statistics, University of Kalyani, Kalyani-741 235, West Bengal, India. E-Mail: debdas.banerjee@gmail.com
[4]Arabinda Das, Department of Statistics, University of Kalyani, Kalyani-741 235, West Bengal, India. E-mail: ara.das@gmail.com.
Research of the second author is supported by a fellowship from Council for Scientific and Industrial Research, Govt. of India.

identification problem is a well-researched topic in both statistics and econometrics, the problem of near-identifiability has not received adequate attention in either literature (Rao, 1992, pp. 134).

Greene (1993, pp. 79) was perhaps first to mention this problem in the context of estimation of the truncated normal SFM. He observed "the cost (of non-zero μ) appears to be that the log-likelihood is relatively flat in the dimension of μ" (Greene 1993, pp 79). Subsequently, Ritter and Simar (1997) showed through simulation that, even with the sample of a few hundred observations, it is difficult to correctly identify the normal-gamma SFM when the sample is generated by one of its sub-model or limiting models. They also observed the classical symptoms associated with "near-identifiable" models viz. imprecise and unstable maximum likelihood estimates of the model parameters. Ritter and Simar (1997), however, neither analyzed the identification problem of the SFM analytically nor considered the near-identification as a problem distinct from the identification problem. As shown here the stochastic frontier models are in fact near-identifiable.

In this paper we carry forward the work of Ritter and Simar (1997) and examine analytically the identification status of the standard stochastic frontier models with uncorrelated and correlated error components. In doing so we give mathematical content to the notion of near–identifiability of a statistical model and show that all the single equation standard frontier models with uncorrelated error components viz. the exponential (Meeusen and van den Broeck, 1977), the half-normal (Aigner et al., 1977), the truncated-normal (Stevenson, 1980) and the gamma (Greene, 1990) frontiers are in fact either globally identifiable or at least locally identifiable but each of them suffer from near-identifiability problem. Secondly, the recently introduced truncated bivariate normal SFM (Pal and Sengupta, 1999; Bandyopadhyay and Das, 2006) is shown to be either unidentifiable or near-identifiable even in a restricted parameter space. Finally, we link the near-identifiability problem of an SFM with its signal to noise ratio (SNR) parameter and show that a disproportionately high or low SNR leads to the near-identifiability problem[5].

---

[5] SNR is defined as the ratio of the variances of the inefficiency and noise in a stochastic frontier model (See, Bandyopadhyay and Das (2006), pp. 174).

In the next section we briefly discuss the identification and near-identification problem in the context of parametric statistical model and state the different criteria for identification and near-identification. In section 3, we use these criteria to examine the identification status of the different stochastic frontier models. In section 4, we relate near-identifiability of an SFM with its SNR parameter. The final section sums up the findings of the study and scope for future work in the context of the problem of near-identifiability of the stochastic frontier models.

2. The Identification Problem

The problem of identification of a statistical model is concerned with proper specification of the theoretical structure of a model that generates the sample observations. The identification problem results from the inability of the sample to discriminate between the two probability structures. Likelihood-based inference, regarding the model, however, is possible only if each data generation process corresponds to one and only one probability structure. Thus when a model is not identifiable, there is no logical basis for likelihood-based inference regarding the model.

Consider *the parametric statistical model* given the family $\prod = \{F_\theta(x), \theta \in \Omega\}$ where $F_\theta(x)$ is the distribution function, indexed by the parameter vector $\theta \in \Omega \subset \mathbb{R}^m$. Let $f_\theta(x)$ be the associated density function that satisfies all the regularity conditions for validity of the Cramer-Rao inequality (See, Cramer 1946, pp. 479). Let $\mathbb{S}$ denote the sample space and $L(\theta \mid x)$ be the likelihood function of θ given the sample $x \in \mathbb{S}$. We use the following definitions (Rothenberg, 1971):

Definition 1: Two parameters $\theta_1, \theta_2 \in \Omega$ are said to be *observationally equivalent* if $L(\theta_1 \mid x) = L(\theta_2 \mid x)$ for some $x \in \mathbb{S}$.

Definition 2: A parameter $\theta_0 \in \Omega$ is said to be *globally identifiable* if there is no other parameter $\theta (\neq \theta_0) \in \Omega$, which is observationally equivalent to θ_0.

Definition 3: The statistical model Π is said to be globally identifiable when every parameter point θ in Ω is globally identifiable.

Definition 4: A parameter θ_0 is said to be *locally identifiable* if there exits an open neighborhood of θ_0, say, $N_\varepsilon(\theta_0)$ for some $\varepsilon > 0$, such that there is no $\theta \in N_\varepsilon(\theta_0)$, which is observationally equivalent to θ_0.

We may note that global identification implies local identification but the converse is not true. Also there may exist an identifiable re-parameterization of the model even when the model is unidentifiable on its natural parameter space.

2.1 Criteria for Global and Local Identification:

Necessary and sufficient condition for global identification of θ_0 when the support of $F_\theta(y)$ is independent of θ is given by the existence of unique solution of $H(\theta,\theta_0)$=0 at $\theta = \theta_0$, where $H(\theta,\theta_0)$ is the expected Kullback-Leibler information for discriminating θ and θ_0 (Rao, 1992, pp. 122). This condition, however, is difficult to check in practice as $H(\theta,\theta_0)$ will not generally have a closed form. For the distributions belonging to the exponential family this condition is equivalent to non-singularity of Fisher's information matrix. However, no such result exists for the distributions belonging to the non-exponential family (which is the case for all the stochastic frontier models) and the conditions of identification are derived in a problem specific manner (Rao, 1992). In this paper we shall use the following results to examine the global and local identifiability of the different SFM.

Result 1 ***(Rothenberg, 1971, pp. 584)*****:** If there exist m known functions $\phi_1(Y),.........,\phi_m(Y)$ such that, for all θ in Ω, $\theta_i = E[\phi_i(Y)]$ for i=1,2…m, where θ_i is the ith element of θ, then every θ in Ω is identifiable.

Result 2 ***(Rothenberg, 1971, pp. 579)*****:** Let θ_0 be a "regular point" of $I(\theta_0)$, the Fisher's information matrix at θ_0. Then θ_0 is locally identifiable if and only if $I(\theta_0)$ is nonsingular[6].

---

[6]A point $\theta_0 \in \Omega$ is said to be a regular point of a matrix $M(\theta)$, whose elements are continuous functions of θ, if there exist an open neighborhood of θ_0 in which $M(\theta)$ has constant rank (*Rothenberg, 1971, pp. 579).*

We may note that while the Result 1 provides sufficient condition for global identification, the condition provided in the Result 2 is both necessary and sufficient for local identification. Therefore, in order to see if θ is identifiable or not, we shall first check if Result 1 holds or not. When no conclusion regarding the global identifiability of a model could be drawn using the Result 1, only then we use the Result 2 to check if the model is locally identifiable or not.

2.2 "Near-identification" Problem:

The near-identification problem refers to the situation where two probability structures (parameters) are "nearly" observationally equivalent. Formally, we say a parameter $\theta_0 \in \Omega$ is "nearly observationally equivalent" if there is another parameter $\theta(\neq \theta_0) \in \Omega$ such that $L(\theta_0 \mid x) \approx L(\theta \mid x)$ for some $x \in \mathbb{S}$. When θ_0 is near-identifiable, the likelihood surface is 'nearly flat" around θ_0. Therefore, we define:

Definition 5: A parameter θ_0 is said to be "near-identifiable" if there exits a neighborhood around θ_0, say, $N_\varepsilon(\theta_0)$, and a very small positive η such that for all $\theta \in N_\varepsilon(\theta_0)$, $\left|L(\theta,x) - L(\theta_0,x)\right| < \eta$. Equivalently, θ_0 is locally near-identifiable if

$$\lim_{\theta \to \theta_0} \frac{\delta \log L}{\delta \theta} = 0 \tag{2.1}$$

Definition 6: A model $\prod = \left\{F_\theta(x), \theta \in \Omega\right\}$ is said to be near-identifiable if there exists $\theta \in \Omega$ such that θ is near- identifiable.

Clearly, when θ_0 is "regular", a sufficient condition for near-identifiability of θ_0 is that the Fisher's information matrix $I(\theta)$ becomes "nearly singular" as $\theta \to \theta_0$. In multi-parameter case one can similarly define near-identifiability of a subset of components of θ_0. Let us partition θ and θ_0 as $\theta = (\theta_1, \theta_2)'$ and $\theta_0 = (\theta_{01}, \theta_{02})$. Let $L(\theta_1 \mid \theta_2, x)$ be the conditional likelihood of θ_1 given θ_2 and *x*. Then

Definition 7: θ_{01} is said to be locally near-identifiable if there exits a neighborhood around θ_{01}, say, $N_\varepsilon(\theta_{01})$, and a very small positive η such that

$\left|L(\theta_1,\theta_2 \mid x)-L(\theta_{01},\theta_2 \mid x)\right|<\eta$ for every $\theta_1 \in N_\varepsilon(\theta_{01})$ and every θ_2. Equivalently, θ_{01} is locally near-identifiable if

$$\lim_{\theta_1\to\theta_{01}} \frac{\delta \log L(\theta_{01},\theta_2 \mid x)}{\delta\theta_1}=0 \tag{2.2}$$

As in case of full parameter case, here too a sufficient condition for near identifiability of θ_{01} is given by near singularity of Fisher's information matrix $I(\theta)$ as $\theta_1 \to \theta_{01}$. Moreover, a model is near-identifiable when at least one component of θ is near-identifiable.

3. Identification of the Stochastic Frontiers:

The stochastic production frontier model of the ith firm is given by

$$y_i = f_\beta(x_i).\exp(v_i - u_i),\ i=1,2,\cdots n \tag{3.1}$$

$$-\infty < v < \infty, 0 < u < \infty$$

where $f(x,\beta)$ is the deterministic production frontier representing the maximum possible output achievable from a bundle of inputs (x) and a given technology $f_\beta(.)$, indexed by the parameter vector β. It is assumed that the actual output (y) of a firm is affected by two random factors; one uncontrollable, called the statistical noise (v) and the other controllable, called the inefficiency (u). The deterministic frontier subject to statistical noise, $y^s = f_\beta(x).\exp(v)$, is called the stochastic frontier and gives the potential output of a firm for different input bundles. The amount by which the actual output falls short of the potential output viz. $\exp(-u)$ measures the technical inefficiency of the firm. We may note that the inefficiency u is a non-negative random variable as $y \le y^s$ for all x and β. While the probability distribution of the "noise" is assumed to be normal, the same for the firm level inefficiency has been modeled by exponential, half-normal, truncated normal or gamma distribution. The noise and the inefficiency are traditionally assumed to be statistically independent though recently, in a few studies, this assumption has been relaxed (Pal and Sengupta, 1999; Smith, 2004; Burns, 2004 and Bandyopadhyay and Das, 2006).

In this section we examine the identifiability of two stochastic frontier models viz. the normal-gamma frontier (Greene, 1990) and the truncated bivariate normal frontier (Pal and Sengupta, 1999; Bandyopadhyay and Das, 2006). We may note that these two models include all the standard stochastic frontier models either as a sub-model or as a limiting model. Before proceeding for identification, we should also note that i) the frontier $f\left(x_i, \beta\right)$ of equation (3.1) is linear in $\beta \in \mathbb{R}^k$ and ii) the c.d.f. of the composite error $\varepsilon = v - u$, $F_\theta(\varepsilon)$, is 'regular" in the sense of Cramer (1946, pp. 479) and does not involve β. Thus, when $F_\theta(\varepsilon)$ is globally identifiable, the probability model of y, $G_\eta(y)$ where $\eta = (\beta, \theta)^{'}$, is also identifiable as long as the matrix $X = (x_1, x_2 .... x_k)$ has full column rank. On the other hand, if $F_\theta(\varepsilon)$ is not identifiable, then $G_\eta(y)$ is also not identifiable even when $X^{'}$ has full column rank.

3.1 Identifiability of the Normal-Gamma SFM:

In the parametric set up, the normal-gamma stochastic frontier model (Greene, 1990) provides the most flexible description of the firm's inefficiency. However, the implementation of this model has been restricted as much because of the complicated nature of its likelihood function as its identification problem. Ritter and Simar (1997) was first to demonstrate through simulation that the model "is poorly conditioned for samples of up to several observations …" (pp. 2). They also observed that the estimates "suffer from substantial imprecision, are ambiguous or can not be calculated at all" and "the full model is hard to identify" (pp. 2). These observations indicate that the model is near-unidentifiable. In this section we show that the normal-gamma stochastic frontier model is globally identifiable but suffers from the near-identifiablity problem.

The normal-gamma SFM is given by the equation (3.1) along with the assumptions i) $v \sim N(0, \sigma_v^2)$ ii) $u \sim G(P, \theta)$ and iii) u and v are independently distributed. Let $\eta = (\beta, \theta, P, \sigma_v^2)^{'}$ be the parameter vector and $I(\eta)$ be the Fisher's information matrix of the normal-gamma SFM. Then the following result shows that the normal-gamma SFM is globally identifiable.

Theorem 1: Normal-Gamma SFM is globally identifiable.

Proof: See Appendix.

We may note that global identification of normal-gamma SFM also establishes the global identification of normal-exponential SFM, as the later is a sub-model of the former at $P = 1$. It also implies that the Fisher's information matrix $I(\eta)$ is non-singular at every point in the parameter space. However, in the next section we show that $|I(\eta)| \approx 0$ for some η in the parameter space. In other words, we show that the normal-gamma SFM is near-identifiable.

3.1.2 Near–identifiability of the Normal-Gamma SFM:

The near-identifiability problem of the normal-gamma SFM can be demonstrated considering the limiting behavior of the characteristic function of the composite error term ε. For example, it can be easily checked as $P \to \infty$, the inefficiency (gamma) distribution tends to the normal distribution and the parameters of the component distributions cannot be separately identified. In other words, the normal-gamma SFM tends to be near-identifiable for very large values of the shape parameter of the gamma distribution. Similarly, it can be shown that as $P \to 1$ and $\theta \to \infty$, the inefficiency distribution becomes degenerate at 0 and the gamma frontier model tends to the Gaussian least squares model. In the next theorem we show that as $P \to 1$ and $\theta \to \infty$, the log-likelihood function of the normal-gamma SFM becomes "nearly-flat".

Theorem 2: $\lim\limits_{\substack{P \to 1 \\ \theta \to \infty}} \dfrac{\partial \ln f(\gamma)}{\partial \theta} \to 0$.

Proof: See Appendix.

Thus the slope of the log-likelihood function becomes "nearly flat" and the normal-gamma SFM becomes near-identifiable in that region of the parameter space where the shape parameter is around unity and the scale parameter is very large.

3.2 Identifiability of the Truncated Bivariate Normal SFM:

The truncated bivariate normal SFM, hereafter BNSFM, is obtained from equation (3.1) under the assumption that the component errors v and u are jointly distributed as truncated bivariate normal; u being truncated at an unknown non-negative point, say, u_0. The p.d.f of y is given by

$$f(y;\eta)=\left[\sigma_*\Phi\left(-\frac{u_*\sqrt{1+\lambda^2-2\lambda\rho}}{\sigma_*\lambda}\right)\right]^{-1}\Phi\left(\frac{u_*}{\sqrt{1-\rho^2}\lambda\sigma_*}+\mu\frac{y-\xi}{\sigma_*}\right)\phi\left(\frac{y-\xi}{\sigma_*}\right) \qquad (3.2)$$

$$-\infty<y<\infty$$

where $\eta=(\beta,\mu_v,\mu_u,\sigma_*,\lambda,\rho,u_0)'$, $u_*=(u_0-\mu_u)$, $\xi=x'\beta+\mu_v-\mu_u$, $\lambda=\sigma_u/\sigma_v$, $\sigma_*^2=\sigma_u^2+\sigma_v^2-2\rho\sigma_u\sigma_v$ and $\mu=-(\lambda-\rho)/\sqrt{1-\rho^2}$.

From the expression of the p.d.f. of y given in (3.2), it is obvious that if both μ_u and μ_v are non-zero then the model is not identifiable as there are infinitely many combinations of μ_u and μ_v that yield the same value for the likelihood function of the model. By the same argument, one can also see that the model is not identifiable if both u_0 and μ_u are non-zero. Thus the model may be identifiable in the following two alternative situations namely i) $\mu_v=u_0=0$ (Pal and Sengupta, 1999) and ii) $\mu_u=\mu_v=0$ (Bandyopadhyay and Das, 2006). Let $\eta_1=(\beta,0,\mu_u,\sigma_*,\lambda,\rho,0)$, $\eta_2=(\beta,0,0,\sigma_*,\lambda,\rho,0)$ and $I(\eta)$ be the Fisher's information matrix for the model. Then the following theorem shows that while the Bandyopadhyay and Das (2006) model is unidentifiable, the Pal and Sengupta (1999) model is near-identifiable.

Theorem 3: $I(\eta_2)$ is singular. Also $\lim_{\lambda\to\infty}|I(\eta_1)|=0$.

Proof: See Appendix.

3.3 Identification of the Half-normal and the Truncated-Normal SFM:

The BNSFM includes the half-normal ($u_0=\mu_v=\mu_u=\rho=0$) and the truncated-normal ($u_0=\mu_v=\rho=0$) SFM as sub-models and tends to Gaussian least squares model as $\lambda\to 0$ (see section 3.3.1 below). The p.d.f. of y under the half-normal (Aigner et al., 1977) and the truncated-normal (Stevenson, 1980) stochastic frontier models can be obtained substituting in (3.2) $u_0=\mu_v=\mu_u=\rho=0$ and $u_0=\mu_v=\rho=0$ respectively.

It can be easily checked that the second and higher order moments of y under the truncated normal SFM are non-linear function of the parameters and the individual parameters of the model cannot be uniquely expressed as continuous function of the population moments. Thus the global identifiability of the model cannot be established using the Result 1. However, the following theorem establishes the global identifiability of the half-normal stochastic frontier.

Theorem 4: The normal-half-normal SFM is globally identifiable.

Proof: See Appendix

3.3.1 Near-identifiability of the Truncated Bivariate Normal SFM:

As in case of normal-gamma SFM, we can examine the near-identifiability of the BNSFM studying the limiting behavior of its characteristic function given by

$$\psi_y(t) = \frac{\exp\left(it\xi - \dfrac{t^2\sigma_*^2}{2}\right)\Phi\left(\dfrac{\alpha + it\sigma_*\mu}{\sqrt{1+\mu^2}}\right)}{\Phi\left(\dfrac{\alpha}{\sqrt{1+\mu^2}}\right)}$$

where $\alpha = -(u_0 - \mu_u)\big/(\sqrt{1-\rho^2}\sigma_*\lambda)$, $\mu = -(\lambda - \rho)\big/\sqrt{1-\rho^2}$, $\xi = x'\beta + \mu_v - \mu_u$. It can be checked that, as $\lambda \to 0$, the characteristic function of the BNSFM tends to that of the Gaussian least squares model. Therefore, when $\lambda \to 0$, normal-half-normal, normal-truncated-normal and truncated bivariate normal SFMs cannot be distinguished from each other. Moreover, it can be easily checked that under the following parametric transformation

$$\xi = x'\beta + \mu_v - \mu_u,$$

$$\sigma_*^2 = \sigma_v^2 + \sigma_u^2 - 2\rho\sigma_v\sigma_u$$

$$\alpha = -\frac{u_0 - \mu_u}{\sqrt{1-\rho^2}\sigma_*\lambda},$$

$$\mu = -\frac{\lambda - \rho}{\sqrt{1-\rho^2}},$$

y follows extended skew-normal distribution i.e. $y \sim ESN(\xi, \sigma_*, \mu, \alpha)$ which is known to be near-identifiable (Capitanio et al., 2003). Also using the characteristic function approach, one can show that the p.d.f. of y tends to be that of skew-normal distribution as $\mu_u \to 0$. In the next theorem we show that the model is near-identifiable. Since following theorem holds for all values of μ_u, it also shows that the normal-half-normal SFM, though globally identifiable, suffers from near-identifiability problem.

Theorem 5: Normal-Truncated-normal SFM is near-identifiable.

Proof: See Appendix.

4. Near–identifiability and the Signal to Noise Ratio:

Bandyopadhyay and Das (2006, pp. 174) defined the signal to noise ratio (SNR) of a model as the ratio of the variances of inefficiency (u) and the noise (v) and studied the relationship between SNR and the firm level inefficiency. In this section we discuss the relationship between SNR and the near-identifiability of a model. An interesting aspect of the above results on near-identifiability is that all the models considered in this paper tends to be near-identifiable as the ratio of $\sigma_u/\sigma_v = \lambda$ tends to 0 and / or ∞. Moreover, except in case of normal-gamma SFM, these results hold good irrespective of the values of the other parameters. Since the signal to noise ratio (SNR) for all the models tend to 0 or ∞ according as λ tends to 0 or ∞, the above results on the near-identifiabiliy shows that all the standard models tend to be near-identifiable as the SNR tends to 0 or ∞. Thus we see that as the variance of one of the component distributions becomes too large or too small *vis-a-vis* that of the other distribution the model becomes near-identifiable and it becomes difficult to identify the parameters of the component distributions. In so far as the expected information regarding the parameters of a distribution is inversely related to the variance of the distribution, the information content of the sample regarding the parameters of inefficiency distribution becomes negligible as the variance of the noise becomes extremely large in comparison with that of the noise. Consequently, it becomes difficult to separately identify the parameters of the inefficiency distribution on the basis of information provided in the sample. Exactly opposite happens as the variance of the inefficiency becomes extremely large in comparison with that of the noise. Thus our results analytically establish the empirical conclusion of Ritter and Simar (1997) viz. a disproportionately high or low value of the variance of one of the component distribution in SFM will make identification of the parameters of component distribution extremely difficult. In other words, disproportionately large or small SNR will make the SFM near-identifiable.

5. Conclusions:

Model identification is an essential prerequisite for the likelihood-based inference of a statistical model. In this article we have analytically examined the identifiability of the standard single equation SFMs with uncorrelated and correlated error components. Giving mathematical content to the notion of near-identifiability of a statistical model, we have shown that each of these SFMs suffers from near-identification problem although they are at least locally identifiable. In particular, we have determined for the different SFM the near-identifiable parameters around which the log-likelihood function becomes "nearly-flat". Our results also provide the analytical support to the empirical conclusion of Ritter and Simar (1997) viz. a too large or a too small variance of one of the component distribution make the identification of their parameters extremely difficult. This result also highlights the pivotal role played by the SNR of a SFM in rendering it near-identifiable. However, some of the important questions that arise in this context and that have been left unanswered here are: how to determine from the sample the extent of "near-identifiablity" of a near-identifiable parameter? Does the near-identifiablity problem of the single equation SFM carry over to the simultaneous equation set-up? These are some of the questions that are currently being explored by the present authors in the context of the near-identification problem of the stochastic frontier models.

References

Aigner, D., K. Lovell and P. Schmidt (1977). Formulation and Estimation of Stochastic Frontier Production Function Models. *Journal of Econometrics*, 6, 21-37.

Bandyopadhyay and Das (2006). On measures of Technical Inefficiency and Production Uncertainty in Stochastic Frontier Production Model with Correlated Error Components. *Journal of Productivity Analysis*, Vol. 26, No. 2, 165-180.

Battese, G. and G. Corra (1977). Estimation of a Production Frontier Model, with Application to the Pastoral Zone off Eastern Australia. *Australian Journal of Agricultural Economics*, 21, 169-179.

Burns, R. C. J. (2004). The Simulated Maximum Likelihood Estimation of Stochastic Frontier Models with Correlated Error Components. Unpublished dissertation, Department of Econometrics and Business Statistics, The University of Sydney, Australia.

Capitanio, A., A. Azzalini and E. Stanghellini (2003). Graphical Models for Skew-normal Variates. *Scandinavian Journal of Statistics,* 30, 129-144.

Cramer, H. (1946). *Mathematical Methods in Statistics.* Princeton University Press, Princeton.

Greene, W. H. (1990). A Gamma Distributed Stochastic Frontier Model. *Journal of Econometrics*, 1990, 46, 141-164.

Greene, W. H. (1993). The Econometric Approach to Efficiency Analysis, in H. O. Fried, C. A. K. Lovell and S. S. Schmidt, eds., *The Measurement of Productive Efficiency: Techniques and Applications*. New York: Oxford University Press.

Kumbhakar, S. and C. A. K. Lovell (2000). *Stochastic Frontier Analysis*. Cambridge University Press, Cambridge.

Meeusen, W. and J. van den Broeck (1977). Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error. *International Economic Review*, 18, 435-444.

Pal, M. and A. Sengupta (1999). A Model of FPF with Correlated Error Components: An Application to Indian Agriculture. *Sankhya,* Vol-61, Series B, 337-350.

Rao, B. L. S. Prakash (1992). *Identifiability in Stochastic Models*, Academic Press, New York.

Ritter, C. and L. Simar (1997). Pitfalls of Normal-Gamma Stochastic Frontier Models, *Journal of Productivity Analysis*, 8, pp. 167-182.

Rothenberg, T. J. (1971). Identification in Parametric Models. *Econometrica*, 39, No. 3, 577-591.

Smith, M. D. (2004). Stochastic Frontier Model with Correlated Error Components. Working paper, Department of Econometrics and Business Statistics, The University of Sydney, Australia.

Stevenson, R. (1980). Likelihood Functions for Generalized Stochastic Frontier Estimation. *Journal of Econometrics*, 13, 57-66.

Appendix

Proof of Theorem 1:

In order to establish the global identifiability of the normal-gamma SFM, we note that the first four central moments of y are $\mu_1(y) = x'\beta - P/\theta$, $\mu_2(y) = \sigma_v^2 + P/\theta^2$, $\mu_3(y) = 2P/\theta^3$ and $\mu_4(y) = 3\sigma_v^4 + 6\sigma_v^4 P/\theta^2 + 3P(P+2)/\theta^4$ (Greene, 1990, p. 152) from which one can obtain the following consistent estimates of β, θ, P and σ_v^2 as $\hat{\beta} = \left(x'x\right)^{-1} x'\left(y - \hat{P}/\hat{\theta}\right)$, $\hat{\theta} = -3m_3/\left(m_4 - 3m_2^2\right)$, $\hat{P} = -\hat{\theta}^3 m_3/2$, $\sigma_v^2 = m_2 - \hat{P}/\hat{\theta}^2$.

Therefore, by Result 1, the model is globally identified.

Proof of Theorem 2:

Let the parameter vector be $\eta = \left(\beta', \theta, P, \sigma_v^2\right)$

Then the log-likelihood function of the model is given by

$$l(\eta) = p\log\theta - \log\overline{|P} - \theta\left(y - x'\beta\right) + \sigma_v^2\theta^2/2 + \log\Phi\{(y - x'\beta + \sigma^2\theta\}/2 + \log h\left(P-1, y\right)$$

where $\eta = \left(\beta', \theta, P, \sigma_v^2\right)'$, $h\left(P-1, y\right) = E\left(u_1^{P-1}\right)$ and $u_1 \sim \frac{1}{2}N\left(y - x'\beta + \sigma_v^2\theta, \sigma_v^2\right)$.

Then, from Greene (1990, pp. 150),

$$\frac{\partial \ln f(\eta)}{\partial\theta} = \frac{P}{\theta} - \frac{h(P,\varepsilon)}{h(P-1,\varepsilon)} = E(u) - E(u \mid \varepsilon)$$

Also, from Greene (1990, pp. 157), we know, $\underset{P\to 1}{lt}\, E(u/\varepsilon) = \lambda + \sigma_u . h(\lambda/\sigma_u)$ where $\lambda = -(\varepsilon - \theta.\sigma_u^2)$ and h(.) is the hazard rate. Now, using the approximation, $h(z) \approx z$ we get $\underset{P\to 1}{lt}\, \partial \ln f(\eta)/\partial\theta = 1/\theta$. Therefore, $\underset{\substack{P\to 1\\ \theta\to\infty}}{lt}\, \partial \ln f(\eta)/\partial\theta = 0$.

Thus, normal-gamma SFM becomes near-identifiable as the scale parameter tends to unity and the shape parameter becomes infinitely large.

Proof of Theorem 3:

Substituting η_1 in (3.2), we get the log-likelihood function of the model, $l\left(\eta_1\right)$ as

$$l\left(\eta_1\right) = -\log\sigma_* - \log\Phi\left(\frac{\mu_u\sqrt{1+\lambda^2-2\lambda\rho}}{\sigma_*\lambda}\right) + \log\Phi\left(\frac{\mu_u}{\sqrt{1-\rho^2}\,\lambda\sigma_*} + \mu\,\frac{y - x'\beta + \mu_u}{\sigma_*}\right) - \frac{\left(y - x'\beta + \mu_u\right)^2}{2\sigma_*^2}$$

Taking limits of the individual elements of the Fisher's information matrix for this model, one gets,

$$\lim_{\lambda\to\infty} I(\eta_1) = \begin{pmatrix} \frac{1}{\sigma_*^2}\left(-\mu^2\Lambda_{10}+1\right)x^{'}x & \frac{1}{\sigma_*^2}\left(-\mu^2\Lambda_{10}+1\right)x & -\frac{1}{2\sigma_*^3}\left(\frac{\mu\Lambda_{11}}{\sigma_*}+\mu E\left[h(z)\right]+\frac{\Lambda_{01}}{\mu\sigma_*}\right)x & 0 & m(\lambda,\rho)\left[\frac{1}{\sigma_8}\Lambda_{11}+E\left[h(z)\right]\right]x \\ & \frac{1}{\sigma_*^2}\left(-\mu^2\Lambda_{10}+1\right) & \frac{1}{2\sigma_*^3}\left[\frac{\mu\Lambda_{11}}{\sigma_*}+\mu E\left[h(z)\right]+\frac{\Lambda_{01}}{\mu\sigma_*}\right] & 0 & m(\lambda,\rho)\left[\frac{1}{\sigma_*}\Lambda_{11}+E\left[h(z)\right]\right] \\ & & \frac{1}{\sigma_*^3}\left[\frac{2}{\sigma_*}+\frac{3}{4}\Lambda_{11}-4\frac{\Lambda_{02}}{\mu^2\sigma_*}\right] & 0 & \frac{1}{2\sigma_*^2}\left[\sqrt{k}\frac{\mu_u}{\sigma_*}-\frac{\Lambda_{12}}{\mu\rho_*^3}+m(\lambda,\rho)\frac{\Lambda_{11}}{\mu}\right] \\ & & & 0 & 0 \\ & & & & -k\frac{\mu_u^2}{\sigma_*^2}-\left[\frac{m^2(\lambda,\rho)}{\rho_*^3}\frac{\Lambda_{11}}{\mu}\right] \end{pmatrix}$$

where $\Lambda_{ij} = E\left[\Psi^i(z_*)z_*^j\right]$ and $z_* = \mu\frac{y-x^{'}\beta-\mu_v}{\sigma}$

Thus the determinant of the Fisher's information matrix tends to zero as the ratio of the variances of inefficiency and noise increases indefinitely and hence the single equation version of the Pal and Sengupta (1999) model, though globally identifiable, becomes "near-identifiable" as the signal to noise ratio tends to zero.

In order to show global unidentifiability of the Bandyopadhyay and Das (2006) model, let us substitute η_2 in (3.2) to get the log-likelihood function of the model as

$$l(\eta_2) = -\log\sigma_* + \log\Phi\left(\mu\frac{y-x^{'}\beta}{\sigma_*}\right) - \frac{\left(y-x^{'}\beta\right)^2}{2\sigma_*^2}$$

From the above log-likelihood function, it can be checked that the different elements of the observed information matrix satisfy the following relations:

$$\frac{\partial^2 l(\eta_2)}{\partial\lambda\partial\beta^{'}} = c\frac{\partial^2 l(\eta_2)}{\partial\rho\partial\beta},\quad \frac{\partial^2 l(\eta_2)}{\partial\lambda\partial\sigma_*} = c\frac{\partial^2 l(\eta_2)}{\partial\rho\partial\sigma_*},\quad \frac{\partial^2 l(\eta_2)}{\partial\lambda^2} = c\frac{\partial^2 l(\eta_2)}{\partial\rho\partial\lambda},\quad \frac{\partial^2 l(\eta_2)}{\partial\lambda\partial\rho} = c\frac{\partial^2 l(\eta_2)}{\partial\rho^2}$$

where the vector $c = (1+\rho\lambda)/(1-\rho^2)$.

Taking expectation of the both sides of the above relations, one gets the different elements of the information matrix $I(\eta_2)$. It can be checked that the resulting information matrix is singular as the last two rows of $I(\eta_2)$ are identical. Thus the Bandyopadhyay and Das (2006) model is globally unidentifiable.

Proof of Theorem 4:

The p.d.f. of *y* is given by

$$f(y;\gamma)=\frac{1}{\sigma}\Phi\left(\lambda\frac{y-x^{'}\beta}{\sigma}\right)\phi\left(\frac{y-x^{'}\beta}{\sigma}\right), \qquad -\infty<y<\infty$$

where $\gamma=\left(\beta^{'},\lambda,\sigma\right)$.

It can be checked that the first three central moments of y are respectively given by $\mu_1(y)=x^{'}\beta+b\delta$, $\mu_2(y)=\sigma^2\left(1-b^2\delta^2\right)$ and $\mu_3(y)=\delta^3 b\left(2b^2-1\right)$ where $b=\sqrt{2/\pi}$, and $\delta=\lambda\big/\sqrt{1+\lambda^2}$. Replacing the population moments, $\mu_r(y)$, by their sample counterparts $m_r(y)$, we get the consistent estimates of β, λ and σ as $\hat{\beta}=\left(x^{'}x\right)^{-1}x^{'}\left(y-b\hat{\delta}\right)$, $\hat{\delta}=\left(m_3\big/b\left(2b^2-1\right)\right)^{1/3}$ and $\hat{\sigma}^2=m_2-\left(1-b^2\hat{\delta}^2\right)^7$. Therefore, by Result 1, the model is globally identified.

Proof of Theorem 5:

The log-likelihood function based on single observation of the model given in (3.2) is

$$l(\gamma)=\log L(\gamma)=-\log\sigma-\log\Phi\left(\frac{\mu_u\sqrt{1+\lambda^2}}{\sigma\lambda}\right)+\log\Phi\left(\frac{\mu_u}{\sigma\lambda}+\lambda\frac{y-x^{'}\beta-\mu_u}{\sigma}\right)-\frac{1}{2\sigma^2}\left(y-x^{'}\beta-\mu_u\right)^2$$

Let the observed information matrix be $\tilde{I}(\gamma)=\dfrac{\partial^2 ln(\gamma)}{\partial\gamma_i\partial\gamma_j}$ where $\gamma=\left(\beta^{'},\mu_u,\sigma,\lambda\right)^{'}$ and

$$\frac{\partial^2 l(\gamma)}{\partial\beta\partial\beta^{'}}=\left[\Psi(z_2)\frac{\lambda^2}{\sigma^2}-\frac{1}{\sigma^2}\right]x^{'}x$$

$$\frac{\partial^2 l(\gamma)}{\partial\beta\partial\mu_u}=\left[\Psi(z_2)\frac{1}{\sigma^2}\left\{1-\lambda^2\right\}-\frac{1}{\sigma^2}\right]x$$

$$\frac{\partial^2 l(\gamma)}{\partial\beta\partial\sigma}=\frac{1}{\sigma^2}\left(-\Psi(z_2)\frac{\mu_u+\lambda^2\varepsilon_1}{\sigma}+h(z_2)\lambda-\frac{2\varepsilon_1}{\sigma}\right)x$$

$$\frac{\partial^2 l(\gamma)}{\partial\beta\partial\lambda}=\left[\Psi(z_2)\frac{\lambda}{\sigma^2}\left\{\mu_u\frac{1}{\lambda^2}+\varepsilon_1\right\}-h(z_2)\frac{1}{\sigma}\right]x$$

$$\frac{\partial^2 l(\gamma)}{\partial\mu_u^2}=\frac{1}{\sigma^2}\left[\Psi(z_1)\frac{\lambda_1^2}{\lambda^2}-\Psi(z_2)\frac{1}{\lambda^2}-1\right]$$

$$\frac{\partial^2 l(\gamma)}{\partial\mu\partial\sigma}=\frac{\lambda_1}{\sigma^2\lambda}\left[-\Psi(z_1)\frac{\lambda_1}{\sigma\lambda}+h(z_1)\right]+\frac{1}{\sigma\lambda^2}\left[\Psi(z_2)\frac{\mu_u}{\sigma^2}+\varepsilon_1\frac{\lambda^2}{\sigma^2}-h(z_2)\right]-\frac{2\varepsilon_1}{\sigma^3}$$

$$\frac{\partial^2 l(\gamma)}{\partial\mu_u\partial\lambda}=\frac{1}{\sigma\lambda^2}\left[-\frac{\Psi(z_1)}{\lambda}+\frac{h(z_1)}{\lambda_1}\right]+\frac{1}{\sigma\lambda}\left[\Psi(z_2)\left\{\varepsilon_1-\frac{\mu_u}{\lambda^2}\right\}+\frac{h(z_2)}{\lambda}\right]$$

$$\frac{\partial^2 l(\gamma)}{\partial\sigma^2}=-\frac{1}{\sigma^2}-\frac{\lambda_1}{\lambda\sigma^3}\left[\Psi(z_1)\frac{\mu_u^2\lambda_1}{\lambda\sigma}+2h(z_1)\right]+\frac{1}{\sigma^3}\left(\frac{\mu_u}{\lambda}+\varepsilon_1\lambda\right)\left\{\left(\frac{\mu_u}{\lambda}+\varepsilon_1\lambda\right)\frac{\Psi(z_2)}{\sigma}+2h(z_2)\right\}-\frac{3\varepsilon_1^2}{\sigma^4}$$

$$\frac{\partial^2 l(\gamma)}{\partial\sigma\partial\lambda}=\Psi(z_1)\frac{\mu_u}{\lambda^3}-\frac{\mu_u h(z_1)}{\sigma^2\lambda^2\lambda_1}+\left(\frac{\mu_u}{\lambda^2\sigma}+\frac{\varepsilon_1}{\sigma}\right)\left(\frac{\Psi(z_2)}{\sigma^3\lambda}(\mu_u+\varepsilon_1\lambda^2)-\frac{h(z_2)}{\sigma}\right)$$

$$\frac{\partial^2 l(\gamma)}{\partial\lambda^2}=\frac{\mu_u}{\sigma\lambda_1}\left[\Psi(z_1)\frac{1}{\lambda^4\lambda_1}-h(z_1)\left\{\frac{2}{\lambda^3}+\frac{1}{\lambda_1^2}\right\}\right]-\Psi(z_2)\frac{1}{\sigma^2}\left\{\frac{\mu_u}{\lambda^2}+\varepsilon_1\right\}^2+2h(z_2)\frac{\mu_u}{\sigma\lambda^3}$$

where

$$\Psi(z_1)=\left[\frac{z_1\phi(z_1)}{\Phi(z_1)}+\frac{\phi^2(z_1)}{\Phi^2(z_1)}\right],\ \Psi(z_2)=\left[\frac{z_2\phi(z_2)}{\Phi(z_2)}+\frac{\phi^2(z_2)}{\Phi^2(z_2)}\right],\ z_1=\frac{\mu_u\lambda_1}{\sigma\lambda},$$

$z_2=\frac{\mu_u}{\sigma\lambda}+\lambda\frac{\varepsilon_1}{\sigma}$, $\lambda_1=\sqrt{\lambda^2+1}$, $\varepsilon_1=y-x'\beta-\mu_u$. Taking expectations of the above terms one gets the Fisher's information matrix $I(\beta,\mu_u,\sigma,\lambda)$ from which it can be checked that

$$\lim_{\lambda\to 0} I(\beta,\mu_u,\sigma,\lambda)=\begin{pmatrix}\frac{1}{\sigma^2}x'x & \frac{1}{\sigma^2}x & \frac{2E(\varepsilon_1)}{\sigma^3}x & 0\\ & \frac{1}{\sigma^2} & \frac{2E(\varepsilon_1)}{\sigma^3} & 0\\ & & \frac{1}{\sigma^2}+\frac{3E(\varepsilon_1^2)}{\sigma^4} & 0\\ & & & 0\end{pmatrix}$$

Therefore, as the variance of the noise increases infinitely vis-à-vis the variance of inefficiency i.e. as signal to noise ratio goes to zero, the determinant of the Fisher's information matrix tends to zero and the normal-truncated-normal SFM becomes near-identifiable.

[7] These estimates are consistent as they can be expressed as continuous functions of sample moments.