The University of Sydney Business School
The University of Sydney

BUSINESS ANALYTICS WORKING PAPER SERIES

Endogenous Environmental Variables

In Stochastic Frontier Models

Christine Amsler, Artem Prokhorov and Peter Schmidt

Abstract

This paper considers a stochastic frontier model that contains environmental variables that affect the level of inefficiency but not the frontier. The model contains statistical noise, potentially endogenous regressors, and technical inefficiency that follows the scaling property, in the sense that it is the product of a basic (half-normal) inefficiency term and a parametric function of the environmental variables. The environmental variables may be endogenous because they are correlated with the statistical noise or with the basic inefficiency term. Several previous papers have considered the case of inputs that are endogenous because they are correlated with statistical noise, and if they contain environmental variables these are exogenous. One recent paper allows the environmental variables to be correlated with statistical noise. Our paper is the first to allow both the inputs and the environmental variables to be endogenous in the sense that they are correlated either with statistical noise or with the basic inefficiency term. Correlation of inputs or environmental variables with the basic inefficiency term raises non-trivial conceptual issues about the meaning of exogeneity, and technical issues of estimation of the model.

Keywords: endogeneity, stochastic frontier, environmental variables
.

May 2017

BA Working Paper No: BAWP-2017-02
http://sydney.edu.au/business/business_analytics/research/working_papers

Endogenous Environmental Variables

In Stochastic Frontier Models

Christine Amsler
Michigan State University

Artem Prokhorov
University of Sydney

Peter Schmidt
Michigan State University

April 9, 2017

Abstract

This paper considers a stochastic frontier model that contains environmental variables that affect the level of inefficiency but not the frontier. The model contains statistical noise, potentially endogenous regressors, and technical inefficiency that follows the scaling property, in the sense that it is the product of a basic (half-normal) inefficiency term and a parametric function of the environmental variables. The environmental variables may be endogenous because they are correlated with the statistical noise or with the basic inefficiency term.

Several previous papers have considered the case of inputs that are endogenous because they are correlated with statistical noise, and if they contain environmental variables these are exogenous. One recent paper allows the environmental variables to be correlated with statistical noise. Our paper is the first to allow both the inputs and the environmental variables to be endogenous in the sense that they are correlated either with statistical noise or with the basic inefficiency term. Correlation of inputs or environmental variables with the basic inefficiency term raises non-trivial conceptual issues about the meaning of exogeneity, and technical issues of estimation of the model.

JEL Classification Codes: C10, C26, C36

Keywords: endogeneity, stochastic frontier, environmental variables

Corresponding author: Peter Schmidt, Department of Economics, Michigan State University, East Lansing, MI 48824, USA. Email: schmidtp@msu.edu. Telephone: 517-355-8381.

1. INTRODUCTION

This paper considers a stochastic frontier model that contains environmental variables that affect the level of inefficiency but not the frontier. The model contains statistical noise, potentially endogenous inputs, and technical inefficiency that follows the scaling property, in the sense that it is the product of a basic (half-normal) inefficiency term and a parametric function of the environmental variables. Both the inputs and the environmental variables may be endogenous because they are correlated with the statistical noise component of the error or with the basic inefficiency term.

The first stochastic frontier papers with endogeneity appear to be Kutlu (2010) and Tran and Tsionas (2013). Inputs may be endogenous because they are correlated with the statistical noise component of the error, but they are not correlated with inefficiency. There are no environmental variables. This model was extended by Karakaplan and Kutlu (2013) to include environmental variables, but they are exogenous.

Tran and Tsionas (2015) use a copula to allow dependence between the inputs and the composed error (the sum of statistical noise and inefficiency). There are no environmental variables.

Amsler, Prokhorov and Schmidt (2016), hereafter APS, was the first paper to allow endogeneity of the inputs with respect to statistical noise and inefficiency separately. There were no environmental variables. The present paper is essentially an extension of Amsler, Prokhorov and Schmidt to allow environmental variables. Both the inputs and the environmental variables can be endogenous because they are correlated with statistical noise and/or they are correlated with the basic inefficiency term. The novelty of the paper is allowing the endogenous variables to be correlated with the basic inefficiency term. This raises non-trivial issues of the meaning of exogeneity and difficult and novel technical issues of estimation.

Kutlu (2016), which was written independently of and roughly contemporaneously with this paper, allows for endogenous inputs and endogenous environmental variables, but only in the sense of correlation with the statistical noise.

Another somewhat similar paper is Griffiths and Hajargasht (2016). They assume a panel data setting. They do not have environmental variables in the same sense that we do, but the distribution of inefficiency differs across firms because it depends on the firm means of the endogenous inputs, so that they do have correlations between noise, inputs and inefficiency. This paper and the current paper are similar in intent but they actually do not have much in common analytically because the models are different.

Our paper makes three contributions to the literature. First, it provides a systematic treatment of endogeneity in stochastic frontier models generally and models with endogenous environmental variables more specifically. Second, it discusses instrumental variables estimation as well as maximum likelihood. Third, it allows environmental variables to be endogenous because they are correlated with either the statistical noise or the basic inefficiency term or both.

The plan of the paper is as follows. In Section 2 we outline the model and define some basic notation. Section 3 considers IV estimation in the case that the endogenous environmental variables are correlated with noise but not with the basic inefficiency error. Section 4 considers MLE, in the same setting. Section 5 considers the case that the endogenous environmental variables may be correlated with basic inefficiency as well as statistical noise. Section 6 gives the results of some simulations, and Section 7 contains our concluding remarks.

2. NOTATION AND BASICS OF THE MODEL

Consider the stochastic frontier model

$$(1) \qquad y_i = \alpha + x_i'\beta + v_i - u_i\ ,\ i = 1, \dots, N\ .$$

In the basic stochastic frontier model of Aigner, Lovell and Schmidt (1977), it was assumed that v_i is normal and u_i is half-normal, and x_i, v_i and u_i are independent. In this paper we will consider the case (commonly assumed in the stochastic frontier literature) that u_i depends on some "environmental variables" q_i that do not influence the frontier output, but which do influence the level of inefficiency u_i. We will consider the case that the distribution of u_i satisfies the "scaling property":

$$(2) \qquad u_i = u_i^o \exp(q_i'\delta)\ .$$

Here the "basic inefficiency term" u_i^o is distributed as $N^+(0, \sigma_u^2)$, i.e. half normal, while $\exp(q_i'\delta)$ is the "scaling function." This is the so-called RSCFG model of Reifschneider and Stevenson (1991), Caudill and Ford (1993) and Caudill, Ford and Gropper (1995).

We want to allow some or all of the inputs and environmental variables to be endogenous, so we will partition x_i and q_i:

$$(3) \qquad x_i = \begin{bmatrix} x_{1i} \\ x_{2i} \end{bmatrix}\ ,\ q_i = \begin{bmatrix} q_{1i} \\ q_{2i} \end{bmatrix},$$

where x_{1i} and q_{1i} are exogenous, and x_{2i} and q_{2i} are endogenous. The full set of exogenous instruments is denoted as z_i:

$$(4) \qquad z_i = \begin{bmatrix} 1 \\ x_{1i} \\ q_{1i} \\ w_i \end{bmatrix}$$

where w_i = any "outside instruments."

For some of the estimators we consider we need to assume reduced form equations for the endogenous inputs and environmental variables:

(5a) $x_{2i} = \Pi_x' z_i + \eta_i$

(5b) $q_{2i} = \Pi_q' z_i + \tau_i$

In the most general version of our model, v_i, u_i^o, η_i and τ_i may all be correlated with each other.

Finally, we define $\mu^o = E(u_i^o) = \sqrt{\frac{2}{\pi}}\sigma_u$, and $u_i^* = u_i^o - \mu^o$.

All of the papers referred to in the Introduction, except for Tran and Tsionas (2015) and Griffiths and Hajargasht (2016), are special cases of this model. Kutlu (2010) and Tran and Tsionas (2013) have no environmental variables and do not allow correlation between the endogenous inputs and inefficiency. Karakaplan and Kutlu (2013) have environmental variables but they are exogenous. Amsler, Prokhorov and Schmidt (2016) do not have environmental variables but they do allow correlation between the endogenous inputs and inefficiency. Kutlu (2016) has environmental variables that can be endogenous, but does not allow correlation of the endogenous inputs or environmental variables with the basic inefficiency term.

Allowing the environmental variables to be correlated with the basic inefficiency term is a decidedly non-trivial extension of the previous literature, in large part because there are subtle issues of what is a proper (or useful) definition of "exogenous" and "endogenous." For example, in an IV setting we want exogenous variables to be valid instruments, and in an MLE setting we want exogenous variables to be things that we can condition on. We give careful definitions of exogeneity that make these things true.

3. IV ESTIMATION WITH ENVIRONMENTAL VARIABLES CORRELATED WITH NOISE BUT NOT WITH BASIC INEFFICIENCY

In this Section we allow endogenous environmental variables, but they are endogenous only in the sense of being correlated with statistical noise, not with the basic inefficiency term. This case was also considered by Kutlu (2016). However, here we consider estimation by instrumental variables (IV) rather than MLE.

In APS, one of the methods of estimation considered was corrected 2SLS (C2SLS). Here the model was estimated by 2SLS (IV), and then the intercept was corrected. The intercept needed correction because, with no environmental variables, $\mu^o = E(u_i) = \sqrt{\frac{2}{\pi}}\sigma_u$ and 2SLS implicitly estimates the model $y_i = (\alpha - \mu^o) + x_i'\beta + (v_i - u_i^*)$. The correction is based on an estimate of σ_u^2, where σ_u^2 and σ_v^2 are estimated using the second and third moments of the 2SLS residuals. This is a straightforward generalization of the corrected OLS estimator of Aigner, Lovell and Schmidt (1977), which was for the case that all of x_i is exogenous.

With environmental variables, things are different because $E(u_i)$ is no longer constant. So now we rewrite the model as follows:

(6) $$y_i = \alpha + x_i'\beta - \mu^o \exp(q_i'\delta) + [v_i - u_i^* \exp(q_i'\delta)]$$

where as before $u_i^* = u_i^o - \mu^o$. A detail is that, assuming $\delta \neq 0$, the term $\mu^o \exp(q_i'\delta)$ will not be confounded with the intercept and so no correction of the intercept will be necessary.

For IV to be consistent, we will need to have

(7) $$E[(y_i - \alpha - x_i'\beta + \mu^o \exp(q_i'\delta))\,|z_i] = 0,$$

or at least the weaker condition that $E[z_i(y_i - \alpha - x_i'\beta + \mu^o \exp(q_i'\delta))] = 0$. The condition in (7) is equivalent to

(8) $$E[(v_i - u_i^* \exp(q_i'\delta))|z_i] = 0.$$

We will have $E[v_i|z_i] = 0$ under all of the definitions of exogeneity that we consider. So fundamentally we will need to make assumptions that guarantee that

(9) $$E[(u_i^* \exp(q_i'\delta))|z_i] = 0 \qquad \text{(IV moment condition)}$$

or at least $E[u_i^* \exp(q_i'\delta) \cdot z_i] = 0$ (weaker IV moment condition).

A trivial observation is that $E[(u_i^* \exp(q_i'\delta))\,|z_i]$ would equal zero under any reasonable definition if q_i is contained in z_i and $E[u_i^o|z_i] = E(u_i^o)$. That is the case that q_i is exogenous. But if $q_i = \begin{bmatrix} q_{1i} \\ q_{2i} \end{bmatrix}$ and only q_{1i} is part of z_i, this is not helpful.

This leads us to the following definition of exogeneity of z_i.

DEFINITION 1: z_i is exogenous if (i) $[v_i|z_i] = 0$; (ii) $E[u_i^o|z_i, q_i] = E[u_i^o|q_i] = E(u_i^o)$.

THEOREM 1: If z_i is exogenous in the sense of Definition 1, the IV moment condition (8) holds.

Proof: If Definition 1 applies, then

$$E[u_i^*|z_i, q_i] = E[u_i^o|z_i, q_i] - E[\mu^o|z_i, q_i] = \mu^o - \mu^o = 0. \tag{10}$$

and then

$$\begin{aligned} E[(u_i^* \exp(q_i'\delta))\,|z_i] &= E_{q|z} E[(u_i^* \exp(q_i'\delta))\,|z_i, q_i] \\ &= E_{q|z} \exp(q_i'\delta)\, E[u_i^*|z_i, q_i] = 0. \end{aligned} \tag{11}$$

Given that the IV moment condition holds, the IV estimator will be consistent if standard regularity conditions (including the rank and order conditions) hold. These are familiar conditions that we will not seek to refine.

The force of Definition 1 is that endogeneity of q_{2i} is with respect to v_i only. That is, endogeneity of q_{2i} means correlation with v_i, not with u_i^o. Although we are not yet speaking in terms of independence, Definition 1 translated into statements of independence would be: z_i is exogenous if z_i is independent of v_i and (z_i, q_i) is independent of u_i^o. Fundamentally the point of view is that the representation $u_i = u_i^o \exp(q_i'\delta)$ is as a product of independent parts. So, for example, if $u_i^o \sim N^+(0, \sigma_u^2)$, then $u_i | q_i \sim N^+(0, \sigma_u^2 \exp(2q_i'\delta))$. That fits with the usual discussion of the RSCFG model, but that is not surprising since the usual discussion is in terms of exogenous q_i (and x_i).

To understand why this strong assumption is necessary, consider the following alternative (too weak) definition of exogeneity.

DEFINITION 2 (Too Weak): $\boldsymbol{z_i}$ is exogenous if (i) $[\boldsymbol{v_i} | \boldsymbol{z_i}] = \mathbf{0}$; (ii) $\boldsymbol{E}[\boldsymbol{u_i^o} | \boldsymbol{z_i}, \boldsymbol{q_i}] = \boldsymbol{E}[\boldsymbol{u_i^o} | \boldsymbol{q_i}]$.

Definition 2 is appealing because it is about the relationship of z_i to the other variables. It does not restrict the relationship between q_{2i} and u_i^o. However, unfortunately, it does not imply that the IV moment condition (8) holds. To see this, we calculate

(12) $$E[(u_i^* \exp(\mathrm{q}_i'\delta)) |\, z_i] = E[(u_i^o \exp(q_i'\delta)) | z_i] - E[(\mu^o \exp(q_i'\delta)) | z_i].$$

The second term on the r.h.s. of (12) simply equals $\mu^o E[\exp(q_i'\delta)\, | z_i]$. The first term is

(13) $$E[(u_i^o \exp(q_i'\delta))\, | z_i] = E_{q|z} \exp(q_i'\delta) E[u_i^o | z_i, q_i] = E_{q|z} \exp(q_i'\delta)\, E[u_i^o | z_i, q_i]$$

where the last equality is by Definition 2. However, these two terms (first and second terms on the right hand side of (12)) are not equal unless $E(u_i^o | q_i) = E(u_i^o) \equiv \mu^o$. This is not implied by Definition 2.

4. MLE WITH ENVIRONMENTAL VARIABLES CORRELATED WITH NOISE BUT NOT WITH INEFFICIENCY

We continue to assume the basic stochastic frontier model (1) with environmental variables entering the specification of u_i as in (2). Also as before $u_i^o \sim N^+(0, \sigma_u^2)$, $v_i \sim N(0, \sigma_v^2)$, and we partition x_i as $x_i = \begin{bmatrix} x_{1i} \\ x_{2i} \end{bmatrix}$.

For completeness, we first mention briefly the case in which all of q_i is exogenous, and x_{2i} is endogenous because it is correlated with v_i. This model was considered by Karakaplan and Kutlu (2013). It is a relatively straightforward extension of Kutlu (2010), Tran and Tsionas (2013) and APS, which did not contain environmental variables. Some algebraic details for this case are given in Appendix 1.

4.1 The Case That x_{2i} and q_{2i} Are Endogenous (Correlated with v_i)

We now consider the main case of interest, in which x_{2i} and q_{2i} are endogenous in the sense that they are correlated with v_i. This case was also considered independently by Kutlu (2016). The instruments are $z_i = \begin{bmatrix} 1 \\ x_{1i} \\ q_{1i} \\ w_i \end{bmatrix}$ as in equation (4) above. We have the SF model of equations (1) and (2), plus we now assume reduced form equations for the endogenous variables:

(14a) $$x_{2i} = \Pi_x' z_i + \eta_i$$

(14b) $$q_{2i} = \Pi_q' z_i + \tau_i$$

We can write these as $p_i = \Pi' z_i + \xi_i$ where $p_i = \begin{bmatrix} x_{2i} \\ q_{2i} \end{bmatrix}$, $\Pi = [\Pi_x, \Pi_q]$, $\xi_i = \begin{bmatrix} \eta_i \\ \tau_i \end{bmatrix}$.

We define ψ_i and Ω similarly to the case considered in Appendix 1:

(15) $$\psi_i = \begin{bmatrix} v_i \\ \xi_i \end{bmatrix} = \begin{bmatrix} v_i \\ \eta_i \\ \tau_i \end{bmatrix}, \ \Omega = \begin{bmatrix} \sigma_v^2 & \Sigma_{v\xi} \\ \Sigma_{\xi v} & \Sigma_{\xi\xi} \end{bmatrix} = \begin{bmatrix} \sigma_v^2 & \Sigma_{v\eta} & \Sigma_{v\tau} \\ \Sigma_{\eta v} & \Sigma_{\eta\eta} & \Sigma_{\eta\tau} \\ \Sigma_{\tau v} & \Sigma_{\tau\eta} & \Sigma_{\tau\tau} \end{bmatrix}.$$

We then make the following Assumption.

ASSUMPTION 1.

(i) $\boldsymbol{\psi_i | z_i \sim N(0, \Omega)}$

(ii) $\boldsymbol{u_i^o | v_i, x_i, q_i, w_i \sim N^{+}(0, \sigma_u^2)}$

So once again u_i^o is independent of everything else, and the errors v_i, η_i and τ_i are independent of the instruments z_i.

We note that $u_i|(v_i, \xi_i, z_i) = u_i|(\xi_i, z_i) \ \sim \ N^{+}(0, \sigma_{u,i}^2)$, where $\sigma_{u,i}^2 = \sigma_u^2 \exp(2q_i'\delta)$ as in Appendix 1.

We can factor the joint density of $(u_i, v_i, \xi_i)|z_i$ as follows:

(16) $$f_{u,v,\xi}((u_i, v_i, \xi_i)|z_i) = f_u(u_i|v_i, \xi_i, z_i) \cdot f_v(v_i|\xi_i, z_i) \cdot f_\xi(\xi_i|z_i)$$
$$= f_u(u_i|\xi_i, z_i) \cdot f_v(v_i|\xi_i, z_i) \cdot f_\xi(\xi_i|z_i)$$

From (16) we can obtain

(17) $$f_{\varepsilon,\xi}\big((\varepsilon_i, \xi_i)\big|z_i\big) = \{\textstyle\int_0^\infty f_u(u|\xi_i, z_i) \cdot f_v[(\varepsilon_i + u)|\xi_i, z_i]du\} \cdot f_\xi(\xi_i|z_i)$$

The term in brackets is the convolution of $N^{+}(0, \sigma_{u,i}^2)$ and $N(\mu_{c,i}, \sigma_c^2)$, where $\sigma_{u,i}^2$ is defined above and

(18) $$\mu_{c,i} = \Sigma_{v\xi}\Sigma_{\xi\xi}^{-1}\xi_i \ , \ \sigma_c^2 = \sigma_v^2 - \Sigma_{v\xi}\Sigma_{\xi\xi}^{-1}\Sigma_{\xi v}.$$

This convolution yields the skew-normal density:

(19) $$f_\varepsilon(\varepsilon_i|\xi_i, z_i) = \left(\frac{2}{\sigma_i}\right) \cdot \varphi\left(\frac{\varepsilon_i - \mu_{c,i}}{\sigma_i}\right) \cdot \Phi\left(\frac{-\lambda_i(\varepsilon_i - \mu_{c,i})}{\sigma_i}\right).$$

In the expression in (19), $\sigma_i^2 = \sigma_{u,i}^2 + \sigma_c^2$; $\lambda_i = \sigma_{u,i}/\sigma_c$; φ is the standard normal density function; and Φ is the standard normal cdf.

The term outside the brackets in (17) is the multivariate normal density:

(20) $$f_\xi(\xi_i|z_i) = \text{constant} \cdot \ |\Sigma_{\xi\xi}|^{-1/2}\exp\left(-\frac{1}{2}\ \xi_i'\Sigma_{\xi\xi}^{-1}\xi_i\right)$$

Combining (19) and (20), we obtain

$$(21) \qquad f_{\varepsilon,\xi}\big((\varepsilon_i,\xi_i)\big|z_i\big) = \left(\frac{2}{\sigma_i}\right)\cdot\varphi\left(\frac{\varepsilon_i-\mu_{c,i}}{\sigma_i}\right)\cdot\Phi\left(\frac{-\lambda_i(\varepsilon_i-\mu_{c,i})}{\sigma_i}\right)\cdot\ |\Sigma_{\xi\xi}|^{-1/2}\exp\left(-\frac{1}{2}\ \xi_i'\Sigma_{\xi\xi}^{-1}\xi_i\right)$$

To obtain the likelihood, we substitute $\varepsilon_i = y_i - \alpha - x_i'\beta$ and $\xi_i = p_i - \Pi' z_i$. Then we take logarithms and sum over i. This yields the result we seek.

THEOREM 2: $\boldsymbol{\ln L = \text{constant} + \ln L_1 + \ln L_2}$ **where**

$$\textbf{(22A)} \qquad \boldsymbol{\ln L_1 = -\frac{1}{2}\sum_i \ln\sigma_i^2 - \frac{1}{2}\sum_i[(y_i - \alpha - x_i'\beta - \mu_{c,i})/\sigma_i]^2}$$

$$\boldsymbol{+\sum_i \ln\left[\Phi\left(\frac{-\lambda_i(y_i-\alpha-x_i'\beta-\mu_{c,i})}{\sigma_i}\right)\right]}$$

and

$$\textbf{(22B)} \qquad \boldsymbol{\ln L_2 = -\frac{N}{2}\ln\left|\Sigma_{\xi\xi}\right| - \frac{1}{2}\sum_i\left(p_i{}' - z_i'\Pi\right)\Sigma_{\xi\xi}^{-1}(p_i - \Pi' z_i)\ .}$$

This likelihood is similar to the likelihood given in equations (A3) - (A5) in Appendix 1, for the case that all of q_i is exogenous, which in turn is similar to the likelihood given in equations (13a) – (13c) of APS, for the case that there is no q_i.

We obtain the MLE by maximizing the likelihood with respect to the parameters $(\alpha,\beta,\sigma_v^2,\sigma_u^2,\delta,\Sigma_{v\xi},\Sigma_{\xi\xi},\Pi)$. Alternatively, as suggested by Kutlu (2010), Karakaplan and Kutlu (2013) and APS, we can use a two-step procedure. Step 1 is to estimate the parameters $\Sigma_{\xi\xi}$ and Π from the reduced form equations, that is by maximizing $\ln L_2$. This yields $\widehat{\Pi}$ = OLS of p_i on z_i and $\widehat{\Sigma}_{\xi\xi} = \frac{1}{N}\sum_i\big(p_i - \widehat{\Pi}' z_i\big)\,(p_i - \widehat{\Pi}' z_i)'$. Step 2 is to estimate the rest of the parameters by maximizing $\ln L_1$ taking the estimates of $\Sigma_{\xi\xi}$ and Π as given. This is essentially a control function approach where the control function in the SF model equation is $\Sigma_{\xi\xi}^{-1}(p_i - \Pi' z_i)$ and the coefficients are $\Sigma_{v\xi}$. The two-step procedure is generally different from the MLE because it ignores the information about $\Sigma_{\xi\xi}$ and Π contained in $\ln L_1$. A practical implication is that the conventionally-calculated standard errors from the Step 2 estimation need to be adjusted to reflect the fact that $\Sigma_{\xi\xi}$ and Π have been estimated, except when $\Sigma_{v\xi} = 0$ (there is no endogeneity). See APS, Section 4.3, for more detail.

4.2 Prediction of $\boldsymbol{u_i}$

The usual predictor of u_i is $\hat{u}_i = E[u_i|\varepsilon_i]$, as suggested by Jondrow et al. (1982). However, given the model of Section 4.1 we can define a better predictor of u_i, namely $\tilde{u}_i = E[u_i|\varepsilon_i, \xi_i]$. Even though u_i is independent of ξ_i, ξ_i is correlated with, and therefore informative about, v_i. Therefore, conditional on $\varepsilon_i = v_i - u_i$, ξ_i is informative about u_i.

This point was first made by APS, Section 4.4, in the context of the simpler model with no q_i. The results that follow here are logically the same as in APS though they are algebraically more complex.

Suppose that we transform $(u_i, \varepsilon_i, \xi_i)$ into $(u_i, \tilde{\varepsilon}_i, \xi_i)$ where

(23) $$\tilde{\varepsilon}_i = \varepsilon_i - \mu_{c,i} \text{ where } \mu_{c,i} = \Sigma_{v\eta}\Sigma_{\eta\eta}^{-1}\xi_i \, .$$

Then ξ_i is independent of $(u_i, \tilde{\varepsilon}_i)$, so that $E[u_i|\varepsilon_i, \xi_i] = E[u_i|\tilde{\varepsilon}_i, \xi_i] = E[u_i|\tilde{\varepsilon}_i]$. Define $\sigma_{u,i}^2$, σ_c^2 and $\mu_{c,i}$ as above. Also define $\lambda_i = \sigma_{u,i}/\sigma_c$, $\sigma_i^2 = \sigma_{u,i}^2 + \sigma_c^2$, $\mu_{*,i} = -\frac{\sigma_{u,i}^2}{\sigma_i^2}\tilde{\varepsilon}_i$ and $\sigma_{*,i}^2 = \frac{\sigma_{u,i}^2\sigma_c^2}{\sigma_i^2}$.

Then the same argument as in Appendix B of APS implies that the distribution of u_i conditional on $\tilde{\varepsilon}_i$ (or conditional on ε_i and ξ_i) is $N^+(\mu_{*,i}, \sigma_{*,i}^2)$. This leads to the explicit expression:

(24) $$\tilde{u}_i = E[u_i|\varepsilon_i, \xi_i] = E[u_i|\tilde{\varepsilon}_i] = \sigma_{*,i}[\Lambda(h_i) - h_i]$$

where $h_i = \frac{\lambda_i}{\sigma_i}\tilde{\varepsilon}_i$ and where $\Lambda(h) = \varphi(h)/[1 - \Phi(h)]$ is the standard normal hazard function. This is somewhat similar to equation (3) of Jondrow et al. (1982).

This is a better predictor than the former predictor, $\hat{u}_i = E[u_i|\varepsilon_i]$, because $\sigma_c^2 < \sigma_v^2$. So, paradoxically, while endogeneity complicates parameter estimation, it makes it possible (subject to the assumptions of the model) to improve the precision of prediction of u_i.

The obvious disadvantage of the new estimator of u_i is that now the reduced form must in fact be a correctly specified model with normal errors. Unlike in the standard linear model this is now a substantive assumption.

5. ESTIMATION WITH ENDOGENOUS VARIABLES CORRELATED WITH INEFFICIENCY

5.1 Notation and Assumptions

We will now consider the case that x_{2i} and q_{2i} may be correlated with u_i^o as well as v_i. APS considered the case that x_{2i} may be correlated with u_i^o, but correlation of q_{2i} with u_i^o is novel.

We will continue to assume the stochastic frontier model as in equation (1), the error specification of equation (2), and the reduced form equations for x_{2i} and q_{2i} as in equation (14). As in the previous Section, we write the reduced form equations as $p_i = \Pi' z_i + \xi_i$ where $p_i = \begin{bmatrix} x_{2i} \\ q_{2i} \end{bmatrix}$, $\Pi = [\Pi_x, \Pi_q]$, $\xi_i = \begin{bmatrix} \eta_i \\ \tau_i \end{bmatrix}$. Also, as in equation (14) above, we will use the notation that $\psi_i = \begin{bmatrix} v_i \\ \xi_i \end{bmatrix} = \begin{bmatrix} v_i \\ \eta_i \\ \tau_i \end{bmatrix}$, $\Omega = \begin{bmatrix} \sigma_v^2 & \Sigma_{v\xi} \\ \Sigma_{\xi v} & \Sigma_{\xi\xi} \end{bmatrix} = \begin{bmatrix} \sigma_v^2 & \Sigma_{v\eta} & \Sigma_{v\tau} \\ \Sigma_{\eta v} & \Sigma_{\eta\eta} & \Sigma_{\eta\tau} \\ \Sigma_{\tau v} & \Sigma_{\tau\eta} & \Sigma_{\tau\tau} \end{bmatrix}$.

Then we make the following Assumption.

ASSUMPTION 2.

(i) $\boldsymbol{\psi_{ij}|z_i \sim N(0, \Omega_{jj})}$

(ii) $\boldsymbol{u_i^o | z_i \sim N^+(0, \sigma_u^2)}$

(iii) The joint distribution of $\begin{bmatrix} \boldsymbol{\psi_i} \\ \boldsymbol{u_i^o} \end{bmatrix}$ conditional on $\boldsymbol{z_i}$ is characterized by the marginal distributions in (i) and (ii), plus the Gaussian (multivariate normal) copula.

In Assumption 2, ψ_{ij} is the j^{th} element of the vector ψ_i and Ω_{jj} is the j^{th} diagonal element of the matrix Ω.

A copula is a joint distribution whose marginals are uniform. It captures the dependence in a joint distribution. Sklar's Theorem says that the specification of a joint distribution determines the marginal distributions and the copula; conversely, if we specify the marginal distributions and a copula, this determines a joint distribution which has the specified marginals. Similarly to what was done in APS, Section 4.5.2, we have therefore specified normal marginal distributions for the elements of ψ_i and a half-normal marginal distribution for u_i^o. The normal marginal distributions for the elements of ψ_i plus the Gaussian copula imply that $\psi_i | z_i$ is distributed as $N(0, \Omega)$, as in Assumption 1. However, in Assumption 2 u_i^o is independent only of z_i, so that now u_i^o can be correlated with the noise v_i and with the reduced form errors η_i and τ_i. Therefore now it is possible for x_{2i} or q_{2i} to be endogenous because they are correlated either with v_i or with u_i^o.

The joint distribution of $\begin{bmatrix} \psi_i \\ u_i^o \end{bmatrix}$ will depend on the variances Ω_{jj} (which include σ_v^2 and the variances of the reduced form errors), σ_u^2, and the correlation parameters in the copula. As a matter of generic notation we will let θ represent this set of parameters. Note that the distribution of $\begin{bmatrix} \psi_i \\ u_i^o \end{bmatrix}$ does not depend on z_i; we are assuming that the random elements of the model are independent of the instruments.

Also note that, if we should wish to assume that u_i^o and v_i are independent, we just need to set the relevant correlation parameter in the copula equal to zero. Smith (2008) has suggested using a copula to allow dependence between u_i and v_i in the basic stochastic frontier model, but the concept of endogeneity does not require such dependence.

The distribution of $\begin{bmatrix}\psi_i \\ u_i^o\end{bmatrix}$ cannot be written in closed form, but (as noted by APS) it is easy to make simulation draws from it. This will be the basis of our estimation methods.

5.2 IV Estimation

We will now consider IV estimation of the model. The basis of IV estimation is the moment condition:

(25) $$E[(y_i - \alpha - x_i'\beta + \omega_i)|z_i] = 0$$

where

(26) $$\omega_i = \omega(z_i, \delta, \Pi_q, \theta_\tau) = E[u_i|z_i] = E[u_i^o \exp(q_i'\delta)|z_i] \,.$$

Here the notation θ_τ is used to represent the parameters that are needed to characterize the joint distribution of u_i^o and τ_i. The parameters that are needed to characterize the distribution of η_i or its correlations with u_i^o and τ_i are not relevant at this point.

We can calculate $\omega(z_i, \delta, \Pi_q, \theta_\tau)$ to any desired degree of precision by the following simulation. (i) Given θ_τ, draw a value of $\begin{bmatrix}\tau_i \\ u_i^o\end{bmatrix}$ from the joint distribution of u_i^o and τ_i. (ii) Given z_i, Π_q and τ_i, calculate a value of $q_{2i} = \Pi_q' z_i + \tau_i$. (iii) Given u_i^o, q_{1i}, q_{2i} and δ, calculate a value of $u_i = u_i^o \exp(q_i'\delta)$. (iv) Repeat this procedure many times and average the value of u_i (over draws of the simulation, not over *i*) to obtain $\omega_i = \omega(z_i, \delta, \Pi_q, \theta_\tau)$.

Note that q_{1i} is part of z_i and so it is taken as given; it is not drawn as part of the simulation.

The moment conditions (25) would lead naturally to the following GMM estimator. We would minimize the IV criterion function

(27) $$\sum_i[\left(y_i - \alpha - x_i'\beta - \omega(z_i, \delta, \Pi_q, \theta_\tau)\right)' z_i]C^{-1} \sum_i[z_i' \left(y_i - \alpha - x_i'\beta - \omega(z_i, \delta, \Pi_q, \theta_\tau)\right)].$$

Here C is a positive definite weighting matrix, and the optimal C is $V[z_i' \left(y_i - \alpha - x_i'\beta - \omega(z_i, \delta, \Pi_q, \theta_\tau)\right)]$, which is estimable in the usual way from the results of a first-step estimation with arbitrary C.

As it stands, the minimization would be with respect to the parameters $\alpha, \beta, \delta, \Pi_q$ and θ_τ. That is a very large set of parameters. We can simplify the IV minimization by observing that some of the parameters can be estimated from the reduced form. For example, least squares applied to the reduced form equation for q_{2i} provides a consistent estimate of Π_q, say $\widehat{\Pi}_q$. Also, if we define k_q to be the dimension of q_{2i}, we can consistently estimate the k_q reduced form error variances and the $\frac{1}{2}k_q(k_q - 1)$ reduced form error covariances from the sums of squares and cross products of the reduced form residuals. These are the parameters in the matrix $\Sigma_{\tau\tau}$ defined in equation (15) above. From these covariances we can calculate the implied copula parameters for the distribution of τ_i.

With that as motivation, we partition $\theta_\tau = \begin{bmatrix}\theta_{\tau 1} \\ \theta_{\tau 2}\end{bmatrix}$, where $\theta_{\tau 1}$ contains the reduced form variances and the copula parameters that determine $\Sigma_{\tau\tau}$, and where $\theta_{\tau 2}$ contains σ_u^2 and $c_{\tau u}$, where $c_{\tau u}$ is generic notation for the k_q copula parameters that determine the dependence between u_i^o and τ_i. We can calculate a consistent estimate, say $\hat{\theta}_{\tau 1}$, of $\theta_{\tau 1}$, based on the results of the estimation of the reduced form equation for q_{2i}. Then we can minimize the modified IV criterion function

(28) $\sum_i[\left(y_i - \alpha - x_i'\beta - \omega\left(z_i, \delta, \widehat{\Pi}_q, \hat{\theta}_{1\tau}, \theta_{\tau 2}\right)\right) z_i']C^{-1} \sum_i[z_i \left(y_i - \alpha - x_i'\beta - \omega\left(z_i, \delta, \widehat{\Pi}_q, \hat{\theta}_{1\tau}, \theta_{\tau 2}\right)\right)]$

with respect to α, β, δ and $\theta_{\tau 2}$.

5.3 MLE

We will now consider maximum likelihood estimation of the model. Omitting the subscript "i" for typographical simplicity, we begin with the joint density of the random elements of the model, namely

(29) $f_{\eta,\tau,v,u^o}(\eta, \tau, v, u^o).$

This is the density implied by the assumptions of Section 5.2, namely, marginal normality of v and of the elements of η and τ, half-normality of u^o, and the normal copula.

Then we want to transform this set of variables to the set $\eta, \tau, \varepsilon, u^o$ where

(30) $\varepsilon = v - u^o \exp(q'\delta) = v - u^o\gamma \exp(\tau'\delta_2)$

and where $\gamma = \exp\ \left(q_1'\delta_1 + z'\Pi_q\delta_2\right)$. Note that $\mathrm{v} = \varepsilon + u^o\gamma\exp(\tau'\delta_2)$, and the Jacobian of the transformation from η, τ, v, u^o to $\eta, \tau, \varepsilon, u^o$ is unity. Therefore the joint density of $\eta, \tau, \varepsilon, u^o$ is

(31) $f_{\eta,\tau,\varepsilon,u^o}(\eta, \tau, \varepsilon, u^o) = f_{\eta,\tau,v,u^o}(\eta, \tau, \varepsilon + u^o\gamma\exp(\tau'\delta_2), u^o).$

Next, we integrate out u^o:

(32) $f_{\eta,\tau,\varepsilon}(\eta, \tau, \varepsilon) = \int f_{\eta,\tau,\varepsilon,u^o}(\eta, \tau, \varepsilon, u^o)du^o$

$= \int f_{\eta,\tau,v,u^o}(\eta, \tau, \varepsilon + u^o\gamma\exp(\tau'\delta_2), u^o)du^o$.

The range of the integral is from zero to infinity.

Finally, we substitute $\eta = x_2 - \Pi_x'z, \tau = q_2 - \Pi_q'z$ and $\varepsilon = y - \alpha - x'\beta$ to obtain the joint density of x_2, q_2 and y:

(33) $f_{x_2,q_2,y}(x_2, q_2, y) = f_{\eta,\tau,\varepsilon}\left(x_2 - \Pi_x'z, q_2 - \Pi_q'z, y - \alpha - x'\beta\ \right)$

This leads to the (log) likelihood

(34) $$\ln L = \sum_{i=1}^{N} \ln(f_{x_2,q_2,y}(x_{2i}, q_{2i}, y_i))$$

The joint density of $\eta, \tau, v, \varepsilon$ can be written in closed form. See Appendix 2. However, this is not a simple expression since it includes the term $\Phi^{-1}[F_{u^o}(u^o)]$. There is no explicit expression for the integral in equation (32). It can be evaluated numerically (by drawing from the distribution of u^o). See Appendix 2 for the details.

6. SIMULATIONS

In this Section, we report the results of simulations designed to investigate the performance of some of the estimators proposed in this paper. Specifically, we will investigate the performance of the MLE's of Section 4.1 and Section 5.3. We do not attempt to perform a detailed investigation of the statistical properties of the estimators, but we want to make sure there are no obvious serious problems (e.g. large biases or variances). Also, and perhaps more importantly, we just want to make sure that our models are estimable in a practical sense. This paper has not presented any formal identification results, and so we are interested in making sure that we have identification in the numerical sense, that is, that the various likelihoods have sharp maxima.

6.1 Data Generation

The DGP for the simulations is described in terms of the model given by equations (1), (2), (5a) and (5b), with additional notation defined in equations (3) and (4). Either Assumption 1 or Assumption 2 will hold depending on whether endogenous variables are correlated only with statistical noise v, or also with the basic inefficiency term u^o.

We will first define a "base case" and then consider some changes from this case. The base case is defined as follows.

1. N = 200.

2. One each of x_1, x_2, q_1, q_2.

3. Two outside instruments, w_1 and w_2.

4. The instruments x_1, q_1, w_1, w_2 all N(0,1), and equicorrelated with correlation 0.5.

5. The errors η, τ, v (components of ψ) all N(0,1), equicorrelated with correlation 0.5, and independent of the instruments.

6. σ_u^2 = $\pi/(\pi - 2)$ so that var(u^o) = 1.

7. α = 0.

8. β chosen such that the two coefficients are equal and var($x'\beta$) = var($v - u$).

9. Π_x and Π_q chosen such that in Π_x the coefficients are equal and var($\Pi_x{}'z$) = var(η), and similarly for Π_q.

10. δ = 0.

11. u^o independent of the other errors (though this will only sometimes be imposed in estimation).

6.2 The Model of Section 4.1

We first consider the model of Section 4.1. In this model η, τ and v are correlated, but u^o is independent of η, τ and v. Therefore x_2 and q_2 are endogenous because they are correlated with v, not u^o. For this case the likelihood can be written in closed form (equations (22A) and (22B)).

Table 1 gives the results of estimation for a single simulated data set, for each of three values of N. The point of this is simply to verify that the model can be estimated (e.g. there is no apparent failure of identification) and yields sensible results. By focusing on a single data set, we can do the usual numerical sensitivity checks (e.g. verifying that the same estimates result from different choices of starting values for the maximization algorithm) that are hard to implement in a many-replication simulation.

As a broad statement, the results in Table 1 do seem sensible. We did not encounter numerical difficulties and there are no obvious anomalies. As N increases, the estimates generally get closer to the true values, and their (OPG) standard errors decrease at approximately the correct $N^{-1/2}$ rate. The estimates of σ_u^2 are imprecise, but that is a common feature in all stochastic frontier models.

We present the 2-step estimates only for N = 200, because in that case, and the other cases we consider, they are very similar to the 1-step estimates.

Table 2 gives the results for the same data generating process, but for a simulation with 1000 replications. Here we give both the simulation standard deviations of the estimates and the average across replications of the standard errors. As in Table 1, the results are unexceptional. There are no serious biases, the standard errors are reliable (the average of the standard errors is close to the simulation standard deviations), and both standard errors and standard deviations decrease at the correct $N^{-1/2}$ rate. As in Table 1, σ_u^2 is estimated imprecisely, but the standard errors accurately reflect this.

6.3 The Model of Section 5.3

Next we consider the model of Section 5.3, which allows u^o, η, τ and v to be correlated. Therefore x_2 and q_2 can be endogenous because they are correlated with v or with u^o or both. However, we will simplify this model a little by making u^o and v independent in data generation and imposing independence of u^o and v in estimation. This restriction seems sensible in empirical applications, and correlation between u^o and v is not part of the endogeneity issue.

We also simplify the data generation process by making u^o independent of η and τ in the DGP. This enables us to generate the data without using a copula. However, we do not impose the restriction that u^o is independent of η and τ in estimation, so a copula is necessary for estimation of the model. We use the Gaussian copula, which is consistent with the DGP. Note that for this case we need to simulate the likelihood. We do so using S draws from the distribution of u^o to evaluate (estimate) the expectation in equation (A10), where the values of S will be reported below.

The results for a single simulated data set, for each of three values of N, are given in Table 3. For this set of results we use S = 100 draws to evaluate the likelihood. Note that the correlation parameters at the bottom of the Table (e.g. $\rho_{v\eta}$) are the correlation parameters in the Gaussian copula, but because v, η and τ are marginally normal they are also the simple correlations of the indicated variables.

Table 3 is similar in format and intent to Table 1. The results are also qualitatively similar. We did not encounter numerical difficulties and the only anomaly is that the parameters $\rho_{\eta u}$ and $\rho_{\tau u}$ (the parameters in the Gaussian copula that determine the correlations between $u^o$and η and τ, respectively) and σ_u^2 are estimated quite imprecisely, especially when N = 200. The estimates of the other parameters are generally a little less precise than in Table 1, but that is not surprising given that we are now estimating a more complicated model.

The imprecision of the estimation of $\rho_{\eta u}$ and $\rho_{\tau u}$disappears if we use large enough values of N and S. For example, when we constructed one data set with N = 10,000 and used S = 1000, the estimates of $\rho_{\eta u}$ and $\rho_{\tau u}$ were 0.005 and 0.034, with standard errors of 0.056 and 0.042. Of course, N = 10,000 is not empirically relevant for these kinds of models, but at least it seems that there is no fundamental (e.g. identification) issue. For σ_u^2, we obtained an estimate of 2.453 (versus true value of 2.752), with a standard error of 0.163, so the imprecision is more resistant to increases in sample size.

Table 4 is similar in format and intent to Table 2, though for a different model. It gives the results for the same data generating process as for Table 3, but for a simulation with 100 replications. Once again we use $S = 100$. Here as in Table 2 we give both the simulation standard deviations of the estimates and the average across replications of the standard errors.

The results for the parameters other than σ_u^2, $\rho_{\eta u}$ and $\rho_{\tau u}$ are mostly unexceptional. There are no serious biases, the standard errors are generally reliable (the average of the standard errors is close to the simulation standard deviations), and both standard errors and standard deviations decrease at the correct $N^{-1/2}$ rate. As in Table 3, σ_u^2, $\rho_{\eta u}$ and $\rho_{\tau u}$ are estimated imprecisely. The standard errors for these parameters reflect this imprecision but they are not as reliable as for the other parameters. For example, for $\rho_{\eta u}$ for $N = 500$, compare the simulation standard deviation of 0.252 to the average standard error of 0.172; and for σ_u^2 for $N = 500$, compare the simulation standard deviation of 1.048 to the average standard error of 0.686. In both cases the standard deviation is about 50% larger than the average standard error. Also for σ_u^2 there appears to be a non-trivial bias. However, given the large standard deviation of the estimate and the small number (100) of replications, this bias is not statistically significantly different from zero.

It is possible that both the bias (for σ_u^3) and the unreliability of the standard errors (for all three parameters) would disappear for larger values of R (number of replications) and S (number of draws to evaluate the expectation in the simulated likelihood). Table 5 reports some results for larger values of R and S. All of these results are for $N = 500$. The first three columns of results simply repeat the information in the middle columns of Table 4, for ease of comparison, and correspond to $S = R = 100$. The second three columns are for $S = 500, R = 192$ and the last three columns are for $S = 500, R = 384$. The number of replications is not as important as the number of draws used to evaluate the likelihood. Increasing S from 100 to 500 essentially removes the bias in the estimate of σ_u^2. However, it does not much improve the accuracy of the standard errors. This issue deserves more attention in future work.

7. CONCLUDING REMARKS

In this paper, we consider the case that the inefficiency term u_i in a stochastic frontier model depends on some "environmental variables" q_i that do not influence the frontier output, but which do influence the level of inefficiency. Specifically we assume a "scaling" model in which $u_i = u_i^o \exp(q_i'\delta)$ where u_i^o is the "basic inefficiency term" (e.g. half-normal). When both q_i and the inputs x_i are exogenous this is a familiar model in the stochastic frontier literature.

We first consider the case that some components of x_i and/or q_i are endogenous because they are correlated with v_i, though not with u_i^o. This case, which is also considered by Kutlu (2016), is relatively easily handled. We then turn to the more novel and difficult case that some components of x_i and/or q_i are endogenous because they are correlated with u_i^o as well as possibly with v_i. We show how to estimate the model by IV and also by MLE. Neither method is simple, because a specific copula must be assumed to model the correlation of u_i^o with the endogenous variables, and because simulation methods are necessary to form the IV criterion function or the likelihood.

The paper also makes the potentially important point that, although endogeneity complicates estimation of the model, it also enables more precise prediction of the inefficiencies.

APPENDIX 1

We define the instrument set $z_i = \begin{bmatrix} 1 \\ x_{1i} \\ q_i \\ w_i \end{bmatrix}$ and we write down a set of reduced form equations:

$x_{2i} = \Pi_x' z_i + \eta_i$. Let $\psi_i \equiv \begin{bmatrix} v_i \\ \eta_i \end{bmatrix}$ and $\Omega = \begin{bmatrix} \sigma_v^2 & \Sigma_{v\eta} \\ \Sigma_{\eta v} & \Sigma_{\eta\eta} \end{bmatrix}$. We make the following assumption.

ASSUMPTION 1(A).

(i) $\boldsymbol{\psi_i | z_i \sim N(0, \Omega)}$

(ii) $\boldsymbol{u_i^o | v_i, z_i, x_{2i} \sim N^+(0, \sigma_u^2)}$; **equivalently,** $\boldsymbol{u_i^o | v_i, x_i, q_i, w_i \sim N^+(0, \sigma_u^2)}$

Assumption 1(A) says that u_i^o is independent of everything else in the model as well as any outside instruments w_i. Note that q_i is exogenous in the current case because it is independent of ψ_i and u_i^o . The only source of endogeneity in the model is that v_i is correlated with x_{2i} if $\Sigma_{v\eta} \neq 0$.

Now we have $u_i = u_i^o \exp(q_i' \delta)$ and therefore

(A1) $$u_i | v_i, z_i, x_{2i} \sim N^+(0, \sigma_{u,i}^2) \ , \ \sigma_{u,i}^2 = \sigma_u^2 \exp(2 q_i' \delta) \ .$$

This is still a half-normal distribution but with the (pre-truncation) variance varying over i. The variation over i is exogenous because it depends on q_i (and δ) and q_i is exogenous.

This is essentially the same as the model of Kutlu (2010) and Karakaplan and Kutlu (2013). We just have to change σ_u^2 to $\sigma_{u,i}^2 = \sigma_u^2 \exp(2 q_i' \delta)$. Explicitly, define

(A2) $$\mu_{c,i} = \Sigma_{v\eta} \Sigma_{\eta\eta}^{-1} (x_{2i} - \Pi_x' z_i) \ , \sigma_i^2 = \sigma_{u,i}^2 + \sigma_c^2 \ , \lambda_i = \frac{\sigma_{u,i}}{\sigma_c} \ , \sigma_c^2 = \sigma_v^2 - \Sigma_{v\eta} \Sigma_{\eta\eta}^{-1} \Sigma_{\eta v} \ .$$

Also let Φ be the standard normal cdf. Then the likelihood is

(A3) $$\ln L = \ln L_1 + \ln L_2$$

(A4) $$\ln L_1 = -\tfrac{1}{2} \textstyle\sum_i \ln \sigma_i^2 - \frac{1}{2\sigma_i^2} \sum_i (y_i - \alpha - x_i' \beta - \mu_{c,i})^2$$

$$+\sum_i \ln\left[\Phi\left(\frac{-\lambda_i(y_i-\alpha-x_i'\beta-\mu_{c,i})}{\sigma_i}\right)\right]$$

(A5) $$\ln L_2 = -\frac{N}{2}\ln\left|\Sigma_{\eta\eta}\right| - \frac{1}{2}\sum_i(x_{2i}' - z_i'\Pi_x)\Sigma_{\eta\eta}^{-1}(x_{2i} - \Pi_x' z_i)\ .$$

We obtain the MLE by maximizing the likelihood with respect to the parameters $(\alpha, \beta, \sigma_v^2, \sigma_u^2, \delta, \Sigma_{v\eta}, \Sigma_{\eta\eta}, \Pi)$.

APPENDIX 2

As in Section 5.3, we assume marginal normality of v and of the elements of η and τ, half-normality of u^o, and the normal copula. Thus, as explained in Section 5.1, $\psi_i = \begin{bmatrix} v_i \\ \eta_i \\ \tau_i \end{bmatrix}$ is distributed as $N(0, \Omega)$. Let the dimensions of x_2 and q_2 be k_x and k_q, and denote the error variances as $\sigma_v^2 = \Omega_{11}$, $\sigma_{\eta,j}^2 = \Omega_{j+1,j+1}$ for $j = 1, \ldots, k_x$, and $\sigma_{\tau,j}^2 = \Omega_{j+k_x+1,j+k_x+1}$ for $j = 1, \ldots, k_q$. Then the joint density of η, τ, v, u^o is

(A6) $$f_{\eta,\tau,v,u^o}(\eta, \tau, v, u^o) = c(F_{\eta_1}, \ldots, F_{\eta_{k_x}}, F_{\tau_1}, \ldots, F_{\tau_{k_q}}, F_v, F_u) \cdot$$
$$[\textstyle\prod_{j=1}^{k_x} f_{\eta_j}] \cdot [\prod_{j=1}^{k_q} f_{\tau_j}] \cdot f_v \cdot f_u\ .$$

In equation (A6), "F" represents the cdf of the random variable indicated by the subscript, evaluated at the point indicated on the left hand side of the equation. For example, if Φ is the standard normal cdf, then $F_{\eta_j} = \Phi\left(\frac{\eta_j}{\sigma_{\eta,j}}\right)$, $F_{\tau_j} = \Phi\left(\frac{\tau_j}{\sigma_{\tau,j}}\right)$, $F_v = \Phi\left(\frac{v}{\sigma_v}\right)$, and F_u is the cdf of $N^+(0, \sigma_u^2)$ evaluated at u^o (or, equivalently, the cdf of $N^+(0,1)$ evaluated at (u^o/σ_u)). Similarly, "f" represents the corresponding pdf; for example, $f_v = \frac{1}{\sigma_v}\varphi\left(\frac{v}{\sigma_v}\right)$, where φ is the standard normal pdf. Also in (A6) "c" is the Gaussian (normal) copula given by

(A7) $$c(s_1, \ldots, s_n) = |R|^{-1/2} \exp\{-\tfrac{1}{2}t'(R^{-1} - I)t\}$$

with $n = k_x + k_g + 2$ and $t = [\Phi^{-1}(s_1), \dots, \Phi^{-1}(s_n)]'$, and where R is the correlation matrix of the copula.

Note that the typical elements of t are as follows. (i) $\Phi^{-1}\left(F_{\eta_j}\right) = \Phi^{-1}\left(\Phi\left(\frac{\eta_j}{\sigma_{\eta,j}}\right)\right) = \frac{\eta_j}{\sigma_{\eta,j}} \equiv \eta_j^{(s)}$; (ii) $\Phi^{-1}\left(F_{\tau_j}\right) = \frac{\tau_j}{\sigma_{\tau,j}} \equiv \tau_j^{(s)}$; (iii) $\Phi^{-1}(F_v) = \sigma_v^{-1}v = v^{(s)}$; and (iv) $\Phi^{-1}\left(F_u(u^o)\right)$. We will let $\eta^{(s)}$ and $\tau^{(s)}$ be the vectors with typical elements $\eta_j^{(s)}$ and $\tau_j^{(s)}$ respectively.

The density given in (A6) and (A7) corresponds to the expression given in equation (29) of the text. Next we want the joint density of $\eta, \tau, \varepsilon, u^o$ as in equation (31) of the text. This is given by

(A8)
$$f_{\eta,\tau,\varepsilon,u^o}(\eta,\tau,\varepsilon,u^o) = |R|^{-\frac{1}{2}}\exp\left\{-\tfrac{1}{2}t_u{}'(R^{-1}-I)t_u\right\}\cdot \left[\textstyle\prod_{j=1}^{k_x} f_{\eta_j}\right]\cdot\left[\textstyle\prod_{j=1}^{k_q} f_{\tau_j}\right]\cdot f_v(\varepsilon+u^o\gamma\exp(\tau'\delta_2)\cdot f_u(u^o)$$

where $\gamma = \exp(q_1'\delta_1 + z'\Pi_q\delta_2)$ and where

(A9)
$$t_u = \{\eta^{(s)\prime}, \tau^{(s)\prime}, \sigma_v^{-1}(\varepsilon + u^o\gamma\exp(\tau'\delta_2)), \Phi^{-1}(F_u(u^o))\}'$$

Finally, as in equation (32) of the text, we integrate out u^o:

(A10)
$$f_{\eta,\tau,\varepsilon}(\eta,\tau,\varepsilon) = \int f_{\eta,\tau,\varepsilon,u^o}(\eta,\tau,\varepsilon,u^o)du^o = |R|^{-\frac{1}{2}}\left[\textstyle\prod_{j=1}^{k_x} f_{\eta_j}\right]\cdot\left[\textstyle\prod_{j=1}^{k_q} f_{\tau_j}\right]\cdot E_{u^o}\left[\exp\left\{-\tfrac{1}{2}t_u{}'(R^{-1}-I)t_u\right\}f_v(\varepsilon+u^o\gamma\exp(\tau'\delta_2))\right],$$

where "E_{u^o}" indicates the expectation over the distribution of u^o. We cannot evaluate this expectation analytically, but we can evaluate it numerically by averaging the indicated expression over many repeated draws from the distribution of u^o (for given values of all of the parameters, and of η, τ and ε).

TABLE 1

Case of Section 4.1

$u^o$is independent of η, τ, v in data generation and independence is imposed in estimation
One replication only

		N=200		N=200	N=500		N=2000	
		1 Step		2 Step	1 Step		1 Step	
	TRUE	Estimate	s.e.	Estimate	Estimate	s.e.	Estimate	s.e.
α	0	-0.009	(0.243)	-0.015	0.266	(0.104)	-0.068	(0.068)
β_1	0.661	0.875	(0.187)	0.872	0.560	(0.106)	0.630	(0.048)
β_2	0.661	0.640	(0.209)	0.640	0.641	(0.108)	0.650	(0.058)
δ_1	0	0.034	(0.131)	0.027	-0.013	(0.054)	0.028	(0.030)
δ_2	0	0.071	(0.115)	0.073	-0.038	(0.042)	-0.031	(0.023)
$\Pi_{x,0}$	0	-0.001	(0.072)	-0.003	0.043	(0.046)	-0.001	(0.022)
$\Pi_{x,1}$	0.316	0.373	(0.095)	0.367	0.343	(0.060)	0.317	(0.029)
$\Pi_{x,2}$	0.316	0.270	(0.088)	0.294	0.302	(0.059)	0.262	(0.028)
$\Pi_{x,3}$	0.316	0.091	(0.084)	0.084	0.284	(0.059)	0.312	(0.028)
$\Pi_{x,4}$	0.316	0.456	(0.092)	0.444	0.372	(0.057)	0.348	(0.027)
$\Pi_{q,0}$	0	0.044	(0.077)	0.042	-0.027	(0.044)	0.002	(0.023)
$\Pi_{q,1}$	0.316	0.382	(0.097)	0.376	0.359	(0.058)	0.311	(0.030)
$\Pi_{q,2}$	0.316	0.259	(0.088)	0.280	0.243	(0.055)	0.277	(0.029)
$\Pi_{q,3}$	0.316	0.183	(0.094)	0.177	0.250	(0.058)	0.347	(0.028)
$\Pi_{q,4}$	0.316	0.361	(0.099)	0.350	0.354	(0.055)	0.304	(0.028)
σ_u^2	2.752	2.076	(0.866)	2.057	3.631	(0.457)	2.537	(0.256)
σ_v^2	1	1.099	(0.362)	1.106	0.721	(0.161)	1.023	(0.106)
σ_η^2	1	0.877	(0.086)	0.877	0.983	(0.068)	0.955	(0.032)
σ_τ^2	1	0.972	(0.095)	0.972	0.909	(0.061)	1.015	(0.033)
$\sigma_{v\eta}$	0.5	0.525	(0.210)	0.524	0.423	(0.130)	0.465	(0.063)
$\sigma_{v\tau}$	0.5	0.469	(0.174)	0.467	0.451	(0.082)	0.439	(0.047)
$\sigma_{\eta\tau}$	0.5	0.487	(0.072)	0.486	0.450	(0.044)	0.481	(0.026)

TABLE 2

Case of Section 4.1

$u^o$is independent of η, τ, v in data generation and independence is imposed in estimation
Simulation with 1000 replications

		N=200			N=500			N=2000		
		Estimate		Avg.	Estimate		Avg.	Estimate		Avg.
	TRUE	(mean)	Std.Dev.	Std.Err.	(mean)	Std.Dev.	Std.Err.	(mean)	Std.Dev.	Std.Err
α	0	-0.042	(0.292)	(0.233)	-0.013	(0.137)	(0.133)	-0.005	(0.067)	(0.063)
β_1	0.661	0.670	(0.164)	(0.176)	0.662	(0.101)	(0.106)	0.663	(0.053)	(0.051)
β_2	0.661	0.655	(0.185)	(0.201)	0.660	(0.120)	(0.120)	0.657	(0.059)	(0.058)
δ_1	0	-0.024	(0.537)	(0.214)	-0.003	(0.062)	(0.065)	-0.002	(0.029)	(0.030)
δ_2	0	0.007	(0.181)	(0.164)	0.000	(0.059)	(0.054)	0.001	(0.023)	(0.024)
$\Pi_{x,0}$	0	0.005	(0.072)	(0.076)	0.001	(0.044)	(0.046)	0.001	(0.021)	(0.023)
$\Pi_{x,1}$	0.316	0.312	(0.089)	(0.098)	0.316	(0.054)	(0.059)	0.317	(0.028)	(0.029)
$\Pi_{x,2}$	0.316	0.317	(0.087)	(0.094)	0.316	(0.056)	(0.056)	0.317	(0.028)	(0.027)
$\Pi_{x,3}$	0.316	0.319	(0.089)	(0.092)	0.317	(0.052)	(0.055)	0.316	(0.027)	(0.027)
$\Pi_{x,4}$	0.316	0.316	(0.084)	(0.092)	0.317	(0.054)	(0.055)	0.316	(0.027)	(0.027)
$\Pi_{q,0}$	0	0.004	(0.071)	(0.076)	0.000	(0.044)	(0.046)	0.000	(0.022)	(0.023)
$\Pi_{q,1}$	0.316	0.313	(0.091)	(0.098)	0.318	(0.056)	(0.059)	0.316	(0.028)	(0.029)
$\Pi_{q,2}$	0.316	0.321	(0.089)	(0.094)	0.314	(0.055)	(0.056)	0.316	(0.027)	(0.027)
$\Pi_{q,3}$	0.316	0.318	(0.089)	(0.093)	0.318	(0.055)	(0.055)	0.316	(0.027)	(0.027)
$\Pi_{q,4}$	0.316	0.314	(0.085)	(0.093)	0.317	(0.055)	(0.056)	0.314	(0.027)	(0.027)
σ_u^2	2.752	2.624	(0.954)	(0.859)	2.707	(0.509)	(0.516)	2.735	(0.257)	(0.245)
σ_v^2	1	1.040	(0.366)	(0.356)	1.020	(0.199)	(0.208)	1.009	(0.101)	(0.098)
σ_η^2	1	0.978	(0.100)	(0.109)	0.992	(0.063)	(0.065)	0.998	(0.031)	(0.032)
σ_τ^2	1	0.980	(0.097)	(0.108)	0.993	(0.061)	(0.066)	0.998	(0.031)	(0.032)
$\sigma_{v\eta}$	0.5	0.497	(0.214)	(0.225)	0.501	(0.137)	(0.136)	0.504	(0.066)	(0.066)
$\sigma_{v\tau}$	0.5	0.499	(0.152)	(0.165)	0.504	(0.094)	(0.099)	0.502	(0.046)	(0.048)
$\sigma_{\eta\tau}$	0.5	0.489	(0.075)	(0.085)	0.499	(0.050)	(0.052)	0.499	(0.024)	(0.025)

TABLE 3

Case of Section 5.3

u^oindependent of η, τ, v in data generation but only independence of $u^o$and v imposed in estimaton
Gaussian copula correctly assumed in estimation
One replication only

		N=200		N=200	N=500		N=2000	
	TRUE	1 Step Estimate	s.e.	2 Step Estimate	1 Step Estimate	s.e.	1 Step Estimate	s.e.
α	0	-0.194	(0.310)	-0.148	0.302	(0.125)	-0.220	(0.067)
β_1	0.661	0.806	(0.183)	0.809	0.538	(0.105)	0.628	(0.049)
β_2	0.661	0.680	(0.198)	0.674	0.665	(0.107)	0.640	(0.057)
δ_1	0	0.175	(0.155)	0.122	0.003	(0.054)	0.014	(0.030)
δ_2	0	-0.028	(0.118)	0.008	-0.042	(0.057)	-0.032	(0.026)
$\Pi_{x,0}$	0	0.072	(0.096)	-0.003	0.019	(0.047)	-0.020	(0.024)
$\Pi_{x,1}$	0.316	0.385	(0.094)	0.367	0.341	(0.061)	0.315	(0.029)
$\Pi_{x,2}$	0.316	0.255	(0.090)	0.294	0.294	(0.060)	0.267	(0.029)
$\Pi_{x,3}$	0.316	0.098	(0.086)	0.084	0.287	(0.060)	0.311	(0.028)
$\Pi_{x,4}$	0.316	0.468	(0.093)	0.444	0.374	(0.058)	0.347	(0.027)
$\Pi_{q,0}$	0	0.144	(0.101)	0.042	-0.026	(0.046)	0.003	(0.026)
$\Pi_{q,1}$	0.316	0.395	(0.096)	0.376	0.361	(0.058)	0.311	(0.030)
$\Pi_{q,2}$	0.316	0.249	(0.093)	0.280	0.232	(0.055)	0.279	(0.029)
$\Pi_{q,3}$	0.316	0.192	(0.099)	0.177	0.251	(0.057)	0.345	(0.028)
$\Pi_{q,4}$	0.316	0.372	(0.100)	0.350	0.360	(0.054)	0.304	(0.028)
σ_u^2	2.752	2.131	(1.332)	2.508	4.389	(0.625)	2.342	(0.250)
σ_v^2	1	1.156	(0.368)	1.115	0.647	(0.148)	1.046	(0.105)
σ_η^2	1	0.895	(0.092)	0.877	0.987	(0.069)	0.952	(0.032)
σ_τ^2	1	1.005	(0.115)	0.972	0.909	(0.061)	1.015	(0.033)
$\rho_{v\eta}$	0.5	0.637	(0.134)	0.635	0.190	(0.204)	0.364	(0.073)
$\rho_{v\tau}$	0.5	0.597	(0.120)	0.585	0.600	(0.115)	0.441	(0.061)
$\rho_{\eta\tau}$	0.5	0.538	(0.057)	0.527	0.477	(0.034)	0.490	(0.018)
$\rho_{\eta u}$	0	0.359	(0.310)	0.277	-0.260	(0.119)	-0.131	(0.069)
$\rho_{\tau u}$	0	0.471	(0.274)	0.346	0.016	(0.142)	0.010	(0.081)

TABLE 4

Case of Section 5.3

u^oindependent of η, τ, v in data generation but only independence of $u^o$and v imposed in estimaton
Gaussian copula correctly assumed in estimation
Simulation with 100 replications

		N=200			N=500			N=2000		
		Estimate		Avg.	Estimate		Avg.	Estimate		Avg.
	TRUE	(mean)	Std.Dev.	Std.Err.	(mean)	Std.Dev.	Std.Err.	(mean)	Std.Dev.	Std.Err
α	0	-0.015	(0.347)	(0.245)	0.005	(0.223)	(0.166)	0.008	(0.191)	(0.084)
β_1	0.661	0.675	(0.192)	(0.171)	0.684	(0.096)	(0.107)	0.663	(0.049)	(0.052)
β_2	0.661	0.648	(0.236)	(0.192)	0.655	(0.115)	(0.120)	0.659	(0.064)	(0.059)
δ_1	0	-0.019	(0.162)	(0.111)	-0.005	(0.078)	(0.065)	0.002	(0.034)	(0.030)
δ_2	0	0.030	(0.189)	(0.103)	0.003	(0.081)	(0.061)	0.002	(0.039)	(0.029)
$\Pi_{x,0}$	0	-0.007	(0.086)	(0.077)	0.008	(0.050)	(0.049)	-0.004	(0.029)	(0.024)
$\Pi_{x,1}$	0.316	0.323	(0.088)	(0.095)	0.320	(0.055)	(0.059)	0.315	(0.027)	(0.029)
$\Pi_{x,2}$	0.316	0.295	(0.076)	(0.087)	0.314	(0.055)	(0.056)	0.312	(0.029)	(0.028)
$\Pi_{x,3}$	0.316	0.312	(0.092)	(0.086)	0.312	(0.049)	(0.056)	0.319	(0.027)	(0.027)
$\Pi_{x,4}$	0.316	0.316	(0.098)	(0.085)	0.323	(0.051)	(0.056)	0.317	(0.029)	(0.027)
$\Pi_{q,0}$	0	-0.006	(0.077)	(0.077)	-0.002	(0.050)	(0.049)	0.000	(0.029)	(0.025)
$\Pi_{q,1}$	0.316	0.315	(0.100)	(0.096)	0.323	(0.057)	(0.059)	0.318	(0.027)	(0.029)
$\Pi_{q,2}$	0.316	0.312	(0.082)	(0.090)	0.314	(0.050)	(0.057)	0.314	(0.025)	(0.028)
$\Pi_{q,3}$	0.316	0.317	(0.096)	(0.088)	0.315	(0.053)	(0.055)	0.314	(0.029)	(0.027)
$\Pi_{q,4}$	0.316	0.311	(0.090)	(0.088)	0.315	(0.055)	(0.055)	0.315	(0.027)	(0.027)
σ_u^2	2.752	2.954	(1.470)	(1.015)	2.947	(1.048)	(0.686)	2.968	(0.811)	(0.351)
σ_v^2	1	1.066	(0.489)	(0.348)	1.020	(0.214)	(0.226)	1.020	(0.159)	(0.112)
σ_η^2	1	0.991	(0.104)	(0.113)	0.994	(0.062)	(0.067)	1.001	(0.034)	(0.032)
σ_τ^2	1	0.988	(0.093)	(0.110)	0.987	(0.067)	(0.066)	0.996	(0.028)	(0.032)
$\rho_{v\eta}$	0.5	0.460	(0.317)	(0.195)	0.466	(0.185)	(0.147)	0.457	(0.087)	(0.079)
$\rho_{v\tau}$	0.5	0.455	(0.227)	(0.161)	0.463	(0.171)	(0.125)	0.481	(0.085)	(0.066)
$\rho_{\eta\tau}$	0.5	0.513	(0.054)	(0.060)	0.500	(0.033)	(0.035)	0.500	(0.015)	(0.017)
$\rho_{\eta u}$	0	-0.032	(0.367)	(0.216)	-0.019	(0.252)	(0.172)	-0.020	(0.127)	(0.096)
$\rho_{\tau u}$	0	-0.005	(0.344)	(0.192)	-0.017	(0.206)	(0.151)	-0.044	(0.091)	(0.083)

TABLE 5

Case of Section 5.3

u^oindependent of η, τ, v in data generation but only independence of $u^o$and v imposed in estimaton
Gaussian copula correctly assumed in estimation
Simulation with N = 500

		S = 100, R = 100			S = 500, R = 192			S = 500, R = 384		
	TRUE	Estimate (mean)	Std.Dev.	Avg. Std.Err.	Estimate (mean)	Std.Dev.	Avg. Std.Err.	Estimate (mean)	Std.Dev.	Avg. Std.Err.
α	0	0.005	(0.223)	(0.166)	-0.048	(0.248)	(0.173)	-0.033	(0.236)	(0.173)
β_1	0.661	0.684	(0.096)	(0.107)	0.663	(0.112)	(0.106)	0.665	(0.109)	(0.107)
β_2	0.661	0.655	(0.115)	(0.120)	0.664	(0.133)	(0.120)	0.656	(0.127)	(0.120)
δ_1	0	-0.005	(0.078)	(0.065)	0.011	(0.082)	(0.070)	0.001	(0.081)	(0.069)
δ_2	0	0.003	(0.081)	(0.061)	-0.009	(0.097)	(0.071)	-0.004	(0.093)	(0.070)
$\Pi_{x,0}$	0	0.008	(0.050)	(0.049)	0.004	(0.047)	(0.046)	0.002	(0.045)	(0.046)
$\Pi_{x,1}$	0.316	0.320	(0.055)	(0.059)	0.323	(0.060)	(0.058)	0.320	(0.059)	(0.058)
$\Pi_{x,2}$	0.316	0.314	(0.055)	(0.056)	0.309	(0.057)	(0.056)	0.314	(0.057)	(0.056)
$\Pi_{x,3}$	0.316	0.312	(0.049)	(0.056)	0.319	(0.055)	(0.054)	0.314	(0.056)	(0.055)
$\Pi_{x,4}$	0.316	0.323	(0.051)	(0.056)	0.314	(0.055)	(0.055)	0.316	(0.056)	(0.055)
$\Pi_{q,0}$	0	-0.002	(0.050)	(0.049)	0.009	(0.051)	(0.046)	0.006	(0.048)	(0.046)
$\Pi_{q,1}$	0.316	0.323	(0.057)	(0.059)	0.317	(0.059)	(0.058)	0.318	(0.058)	(0.058)
$\Pi_{q,2}$	0.316	0.314	(0.050)	(0.057)	0.306	(0.060)	(0.055)	0.313	(0.058)	(0.055)
$\Pi_{q,3}$	0.316	0.315	(0.053)	(0.055)	0.318	(0.057)	(0.055)	0.316	(0.058)	(0.054)
$\Pi_{q,4}$	0.316	0.315	(0.055)	(0.055)	0.322	(0.054)	(0.055)	0.316	(0.056)	(0.054)
σ_u^2	2.752	2.947	(1.048)	(0.686)	2.701	(0.915)	(0.685)	2.753	(0.909)	(0.697)
σ_v^2	1	1.020	(0.214)	(0.226)	1.048	(0.284)	(0.226)	1.046	(0.277)	(0.228)
σ_η^2	1	0.994	(0.062)	(0.067)	0.988	(0.066)	(0.065)	0.988	(0.061)	(0.065)
σ_τ^2	1	0.987	(0.067)	(0.066)	0.993	(0.062)	(0.066)	0.991	(0.061)	(0.065)
$\rho_{v\eta}$	0.5	0.466	(0.185)	(0.147)	0.469	(0.189)	(0.145)	0.479	(0.179)	(0.147)
$\rho_{v\tau}$	0.5	0.463	(0.171)	(0.125)	0.469	(0.183)	(0.126)	0.476	(0.176)	(0.127)
$\rho_{\eta\tau}$	0.5	0.500	(0.033)	(0.035)	0.502	(0.034)	(0.035)	0.500	(0.035)	(0.035)
$\rho_{\eta u}$	0	-0.019	(0.252)	(0.172)	0.019	(0.292)	(0.194)	0.018	(0.290)	(0.191)
$\rho_{\tau u}$	0	-0.017	(0.206)	(0.151)	0.014	(0.226)	(0.160)	0.011	(0.217)	(0.159)

REFERENCES

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

Amsler, C., A. Prokhorov and P. Schmidt (2016), "Endogeneity in Stochastic Frontier Models," *Journal of Econometrics*, 190, 280-288.

Caudill, S.B. and J.M. Ford (1993), "Biases in Frontier Estimation Due to Heteroskedasticity," *Economics Letters*, 41, 17-20.

Caudill, S.B., J.M. Ford and D.M. Gropper (1995), "Frontier Estimation and Firm-Specific Inefficiency Measures in the Presence of Heteroskedasticity," *Journal of Business and Economic Statistics*, 13, 105-111.

Griffiths, W.E. and G. Hajargasht (2016), "Some Models for Stochastic Frontiers with Endogeneity," *Journal of Econometrics*, 190, 341-348.

Jondrow, J., C.A.K. Lovell, I.S. Materov and P. Schmidt (1982), "On Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model," *Journal of Econometrics*, 19, 233-238.

Karakaplan, M.U. and L. Kutlu (2013), "Handling Endogeneity in Stochastic Frontier Analysis," unpublished manuscript.

Kutlu, L. (2010), "Battese-Coelli Estimator with Endogenous Regressors," *Economics Letters*, 109, 79-81.

Kutlu, L. (2016), "A Time-Varying True Individual Effects Model with Endogenous Regressors," unpublished manuscript.

Reifschneider, D. and R. Stevenson (1991), "Systematic Departures from the Frontier: A Framework for the Analysis of Firm Inefficiency," *International Economic Review*, 32, 715-723.

Smith, M.D. (2008), "Stochastic Frontier Models with Dependent Error Components," *Econometrics Journal*, 11, 172-192.

Tran, K.C. and E.G. Tsionas (2013), "GMM Estimation of Stochastic Frontier Models with Endogenous Regressors," *Economics Letters*, 118, 233-236.

Tran, K.C. and E.G. Tsionas (2015), "Endogeneity in Stochastic Frontier Models: Copula Approach without External Instruments," *Economics Letters,* 133, 85-88.