Economics Letters 118 (2013) 233–236

Contents lists available at SciVerse ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

GMM estimation of stochastic frontier model with endogenous regressors

Kien C. Tran[a,*,1], Efthymios G. Tsionas[b]

[a] Department of Economics, University of Lethbridge, 4401 University Drive S, Lethbridge, Alberta, T1K 3M4, Canada
[b] Department of Economics, Athens University of Economics and Business, 76 Patission Street, 10434 Athens, Greece

ARTICLE INFO

Article history:
Received 18 June 2012
Received in revised form
28 August 2012
Accepted 23 October 2012
Available online 31 October 2012

JEL classification:
C13

Keywords:
Endogeneity
Generalized method of moments
Maximum likelihood
Technical efficiency

ABSTRACT

A convenient and simple GMM procedure for estimating stochastic frontier models in the presence of endogenous regressors is proposed. Monte Carlo simulations show that the proposed estimator works very well in finite samples. We apply the proposed method to panel data of Norwegian dairy farms to illustrate the usefulness of the proposed approach.

1. Introduction

Recently, Kutlu (2010) proposed a modification of the Battese and Coelli (1992) estimator to account for endogenous regressors via one-step and two-step ML procedures. The main advantage of a two-step estimator is that it is easy to implement but the standard errors of the estimated parameters are incorrect and inconsistent, hence a bootstrap procedure is needed in order to obtain the correct standard errors. In contrast, the one-step estimation is more complicated to compute but it provides the correct standard errors of the estimated parameters.

In this note we suggest a different one-stage estimation approach by looking at the first order conditions of the correct likelihood function, when endogeneity is taken into account, and propose a simple GMM estimator that is consistent and asymptotically efficient.

Section 2 gives the model specification and derivation of the GMM estimator. Section 3 presents the results of Monte Carlo experiments to examine the finite sample performance of the GMM estimator. Section 4 applies the proposed estimator to a large unbalanced panel data from Norwegian dairy farms. Concluding remarks are given in Section 5.

2. The model and GMM estimation procedure

Consider the following stochastic production frontier model with endogenous regressors:

$$y_{it} = z'_{1,it}\alpha + x'_{it}\beta + v_{it} - u_{it} \tag{1}$$

$$x_{it} = z_{2,it}\gamma + \varepsilon_{it} \tag{2}$$

where 'prime' denotes transpose of a vector or a matrix, y_{it} is a scalar dependent variable, x_{it} is a $(1 \times p)$ vector of endogenous regressors, $z_{1,it}$ is a $(1 \times q_1)$ vector of exogenous regressors, $z_{2,it} = I_p \otimes \tilde{z}_{2,it}$, $\tilde{z}_{2,it}$ is of dimension q_2 $(q_2 \geq p)$, and $\tilde{z}_{2,it}$ is assumed to be strictly exogenous in the sense that $E(\varepsilon_{it}|z_{2,it}) = 0$ and $E(\xi_{it}|z_{it}, \varepsilon_{it}) = E(\xi_{it}|\varepsilon_{it})$ where $\xi_{it} = v_{it} - u_{it}$, and $z_{it} = (z_{1,it}, z_{2,it})$. We assume the error terms ε_{it} and v_{it} satisfy the following:

$$\begin{pmatrix} \tilde{\varepsilon}_{it} \\ v_{it} \end{pmatrix} \equiv \begin{pmatrix} \Omega_{\varepsilon}^{-1/2}\varepsilon_{it} \\ v_{it} \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} I_p & \rho\sigma_v \\ \rho'\sigma_v & \sigma_v^2 \end{bmatrix} \right) \tag{3}$$

and Ω is a $(p \times p)$ variance–covariance matrix of ε_{it} and ρ is a $(p \times 1)$ correlation vector between v_{it} and ε_{it}. We assume that $u_{it} \sim$ i.i.d. $|N(0, \sigma_u^2)|$ and independent of x_{it}, z_{it}, v_{it} and ε_{it}. Other distributional assumptions of u_{it} such as exponential, gamma or truncated normal with exogenous variables dependent mean (e.g., Battese and Coelli, 1992, 1995) can be used without affecting the proposed methodology given below.

By Cholesky decomposition of the variance–covariance matrix of $(\tilde{\varepsilon}_{it}\ v_{it})'$, we can write (3) as:

$$\begin{pmatrix} \tilde{\varepsilon}_{it} \\ v_{it} \end{pmatrix} = \begin{bmatrix} I_p & 0 \\ \sigma_v\rho' & \sigma_v\sqrt{1-\rho'\rho} \end{bmatrix} \begin{pmatrix} \tilde{\varepsilon}_{it} \\ \tilde{\omega}_{it} \end{pmatrix} \tag{4}$$

* Corresponding author. Tel.: +1 403 329 2511; fax: +1 403 329 2519.
E-mail addresses: kien.tran@uleth.ca (K.C. Tran), tsionas@aueb.gr (E.G. Tsionas).

1 This paper was written while the first author was a visiting research fellow at the Department of Economics, Athens University of Economics and Business and the University of Macedonia, Greece.

0165-1765/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.econlet.2012.10.028

where $\tilde{\omega}_{it} \sim N(0, 1)$ and independent of $\tilde{\varepsilon}_{it}$. Using (4), we can write (1) and (2) as

$$y_{it} = z'_{1,it}\alpha + x'_{it}\beta + \sigma_v \rho' \Omega^{-1/2}(x_{it} - z_{2,it}\gamma) + \omega_{it} - u_{it} \quad (5)$$

where $\omega_{it} \sim N(0, (1-\rho'\rho)\sigma_v^2)$. Let $\tilde{\xi}_{it} = \omega_{it} - u_{it}$, $\sigma_s^2 = (1-\rho'\rho)\sigma_v^2 + \sigma_u^2$ and $\lambda = \sigma_u/(\sigma_v\sqrt{1-\rho'\rho})$, then the probability density function of $\tilde{\xi}_{it}$ is given by

$$f(\tilde{\xi}_{it}) = \frac{2}{\sigma_s}\phi\left(\frac{\tilde{\xi}_{it}}{\sigma_s}\right)\Phi\left(\frac{-\lambda\tilde{\xi}_{it}}{\sigma_s}\right), \quad -\infty < \tilde{\xi}_{it} < \infty \quad (6)$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal PDF and CDF respectively. Let $y_i = (y_{i1}, \ldots, y_{iT})'$, $x_i = (x'_{i1}, \ldots, x'_{iT})'$, $z_i = (z'_{i1}, \ldots, z'_{iT})'$ and which denotes the $(m \times 1)$ vector parameter θ as $\theta = (\alpha, \beta, \gamma, \eta, \lambda, \sigma_s^2)'$ where $\eta = -\left[\frac{\sigma_s^2}{(1+\lambda^2)(1-\rho'\rho)}\right]^{1/2}\Omega_\varepsilon^{-1/2}\rho$, then for the sample observations (y_i, x_i, z_i), the joint log-likelihood function of y_i and x_i is given by:

$$\ln L(\theta) = \ln L_{y|x}(\theta) + \ln L_x \quad (7)$$

where

$$\ln L_{y|x}(\theta) = \sum_{i=1}^{n}\ln f(y_i|x_i, z_i, \theta) \propto -\frac{nT}{2}\ln\sigma_s^2 + \frac{1}{\sigma_s}\sum_{i=1}^{n}\ln\Phi\left[-\lambda\left(y_i - z'_{1,i}\alpha - x'_i\beta - \eta(x_i - z_i\gamma)\right)\right] - \frac{1}{2\sigma_s^2}\sum_{i=1}^{n}\left(y_i - z'_{1,i}\alpha - x'_i\beta - \eta(x_i - z_i\gamma)\right)^2 \quad (8a)$$

and

$$\ln L_x = \sum_{i=1}^{n}\ln f(x_i) \propto -\frac{nT}{2}\ln(|\Omega_\varepsilon|) - \frac{1}{2}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon'_{it}\Omega_\varepsilon^{-1}\varepsilon_{it} \quad (8b)$$

where in (8a) we have used all the exogenous variables in the model $z_i = (z_{1,i}, z_{2,i})$ instead of $z_{2,i}$ in the definition of ε_i from (2). Let $\partial \ln L(x_i, z_i, \theta)/\partial\theta$ denote the $(m \times 1)$ corresponding likelihood score vector, where m is the total number of parameters in (7). Then we have

$$n^{-1}\sum_{i=1}^{n}\partial\ln f(x_i, z_i, \theta)/\partial\theta = n^{-1}\sum_{i=1}^{n}g_1(x_i, z_i, \theta) = 0. \quad (9a)$$

From (2), the p-first order conditions for OLS are given by:

$$n^{-1}\sum_{i=1}^{n}z'_i(x_i - z_i\gamma) = n^{-1}\sum_{i=1}^{n}g_2(x_i, z_i, \gamma) = 0 \quad (9b)$$

where $z_i = (z_{1,i}, z_{2,i})$. Eqs. (9a) and (9b) constitute a set of $(p+m)$ moment conditions that form the basis for our GMM estimator. Define the moments vector $G(x_i, z_i, \theta, \gamma) = (g_1(x_i, z_i, \theta)', g_2(x_i, z_i, \gamma)')'$, then the joint GMM estimator takes the form

$$\left(\hat{\gamma}, \hat{\theta}\right) = \arg\min_{\gamma,\theta}\left(n^{-1}\sum_{i=1}^{n}G(x_i, z_i, \theta, \gamma)'\right) \times W_n\left(n^{-1}\sum_{i=1}^{n}G(x_i, z_i, \theta, \gamma)\right) \quad (10)$$

where W_n is a symmetric positive definite weighting matrix. To analyze the asymptotic property of the GMM estimator in (9), it is useful to define the following partition matrices:

$$D = \begin{bmatrix} G_{1,\gamma} & 0 \\ G_{2,\gamma} & G_{2,\theta} \end{bmatrix}, \qquad \Omega = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}$$

where $G_{1,\gamma} = E[\partial g_1(\cdot)/\partial\gamma]$, $G_{2,\gamma} = E[\partial g_2(\cdot)/\partial\gamma]$, $G_{2,\theta} = E[\partial g_2(\cdot)/\partial\theta]$ and $V_{ij} = E[g_i(.)g_j(\cdot)']$ for $i, j = 1, 2$. Then, under suitable regularity conditions, see, for example, Hansen (1982) or Newey (1984), it can be easily shown that the GMM estimator given in (10) is consistent and

$$\sqrt{n}\left(\hat{\gamma}, \hat{\theta}\right) \to N\left(0, \left(D^{-1}\Omega D^{-1'}\right)\right).$$

A consistent estimator of the asymptotic covariance matrix of $\left(\hat{\gamma}, \hat{\theta}\right)$ can be easily obtained by replacing the unknown parameters in D and Ω respectively, with their consistent estimates from (10). Note that the asymptotic variance of the GMM estimator is unaffected by the choice of the weighting matrix W_n because in this case, we have an exact identification where the number of moment conditions are exactly the same as the number of parameters to be estimated. Thus in practice one can set $W_n = I$.

The proposed GMM estimator in (10) is an extension of the Newey (1984) GMM estimator for the nonlinear regression case. It provides consistent and correct standard errors of the estimated parameters, and it is fairly simple to compute given the current existing computing power and readily automated GMM estimation program. The asymptotic efficiency of the proposed GMM estimator can be obtained with just one iteration so that numerical searches can be minimized or avoided, albeit in practice iterating to convergence would be more preferable.

3. Monte Carlo simulations

To examine the finite sample performance of our proposed GMM estimator, we conduct some Monte Carlo experiments. To this end, we consider the following data generating process (see Kutlu, 2010):

$$y_{it} = z_{1,it}\alpha + x_{it}\beta + v_{it} - u_{it}$$
$$x_{it} = z_{2,it}\gamma + \varepsilon_{it}, \quad \varepsilon_{it} \sim N\left(0, \sigma_\varepsilon^2\right)$$

where u_{it} is generated as i.i.d. $\left|N\left(0, \sigma_u^2\right)\right|$ and the random variable $z_{it} = (z_{1,it}, z_{2,it})'$ is generated by a VAR (1) process: $\binom{z_{1,it}}{z_{2,it}} = R_2\binom{z_{1,it-1}}{z_{2,it-1}} + e_{it}$, $e_{it} \sim N(0_2, I_2)$, and $\binom{z_{1,i1}}{z_{2,i1}} \sim N\left(0_2, (I_2 - R_2)^{-1}\right)$ where 0_2 is a null vector of dimension 2 and $R_2 = \begin{pmatrix} 0.4 & 0.05 \\ 0.05 & 0.4 \end{pmatrix}$. The vector of random errors $(v_{it}, \varepsilon_{it})'$ is generated by $\binom{v_{it}}{\varepsilon_{it}} \sim N\left(0_2, \begin{pmatrix} \sigma_v^2 & \rho\sigma_v\sigma_\varepsilon \\ \rho\sigma_v\sigma_\varepsilon & \sigma_\varepsilon^2 \end{pmatrix}\right)$. In our experiment, we fix $\alpha = \beta = 0.5$, $\sigma_\varepsilon^2 = 1$, $\sigma_u = \sigma_v = 1$ and $\gamma = 1$. We set the values of $\rho = \{0.0, 0.4, 0.8\}$ and consider two sample sizes: $(n, T) = (50, 15)$ and $(100, 15)$. The simulations are replicated 1000 times, and all the computations are done using the GAUSS program. For the purpose of comparison, we also compute the standard MLE estimators. Simulation results of the parameter estimates' MSE are displayed in Tables 1 and 2.

Our simulations show that when there is no correlation (i.e., no endogeneity in the regressor), the GMM estimator performs almost as well as the standard MLE for all ranges of the parameters considered. However, when there is correlation and as the correlation increases, the MLE deteriorates quickly and becomes severely biased, while the proposed GMM estimator remains unbiased. In addition, for a fixed $T = 15$, as the sample size n doubles, the estimated MSE