Economics Letters 118 (2013) 151–154

Contents lists available at SciVerse ScienceDirect

Economics Letters

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

Misspecification in allocative inefficiency: A simulation study

Levent Kutlu *

School of Economics, Georgia Institute of Technology, Atlanta, GA, 30332-0615, USA

ARTICLE INFO

Article history:
Received 18 August 2012
Received in revised form
26 September 2012
Accepted 2 October 2012
Available online 11 October 2012

JEL classification:
C2
C3

Keywords:
Technical inefficiency
Allocative inefficiency
Stochastic frontier
Panel data

ABSTRACT

We make an extensive simulation analysis in order to investigate the consequences of ignoring the potentially complex and data dependent effects of allocative inefficiency on the estimation of stochastic frontier panel data models. Generally system estimators perform worse than single equation estimators. This result holds even when we approximate the allocative inefficiency.

1. Introduction

In the stochastic frontier framework, a cost–share system is rarely estimated as the inefficiency term in the cost equation and the deviations from the optimal shares from the observed shares are complicated functions of allocative inefficiency. Kumbhakar (1997) provides an exact solution but it is very difficult to estimate this model. Hence, Kumbhakar and Tsionas (2005) and Brissimis et al. (2009) provide first-order approximations to Kumbhakar's (1997) model. In the cross sectional data and single equation framework, Kumbhakar and Wang (2006) show (by simulations) that ignoring the allocative inefficiency can lead to non-negligible biases in the parameter estimates. Hence, ignoring allocative inefficiency can be problematic even for the single equation models. In contrast to Kumbhakar and Wang (2006), who only use a single equation maximum likelihood estimator, we use a variety of estimators in our study. This enables us to compare the performances of different estimators. One promising candidate for this setting is Kumbhakar's (1997) exact model. We make a Taylor series approximation to Kumbhakar's (1997) model.[1] In addition to this estimator we include two regression based (fixed effects (FE) and Cornwell et al. (1990) within (CSSW)) and two maximum likelihood based (single equation and system) estimators.

In Section 2, we explain the models that will be examined in our simulations. Section 3 gives more details about the simulations and the results. Finally, in Section 4, we make our conclusions.

2. Models

We consider five different models for our experiments. The first and second models are FE and CSSW, respectively. Since these estimators are well known, we omit the details. The third and fourth models are single equation maximum likelihood estimator (MLSE) and the system maximum likelihood estimator (MLSYS), respectively. We assume that

$$\begin{aligned}
\ln\left(\frac{C_{it}}{W_{it}^M}\right) &= \ln\left(\frac{\tilde{C}_{it}}{W_{it}^M}\right) + u_{it} + v_{it} \\
&= \beta_0 + \sum_k \beta_{Y_k} \ln(Y_{it}^k) \\
&\quad + \sum_m \beta_{W_m} \ln\left(\frac{W_{it}^m}{W_{it}^M}\right) \\
&\quad + \frac{1}{2}\sum_{k,l} \beta_{Y_{kl}} \ln(Y_{it}^k)\ln(Y_{it}^l) \\
&\quad + \frac{1}{2}\sum_{m,n} \beta_{W_{mn}} \ln\left(\frac{W_{it}^m}{W_{it}^M}\right)\ln\left(\frac{W_{it}^n}{W_{it}^M}\right) \\
&\quad + \sum_{k,m} \beta_{YW_{km}} \ln(Y_{it}^k)\ln\left(\frac{W_{it}^m}{W_{it}^M}\right) + u_{it} + v_{it} \qquad (1)
\end{aligned}$$

$$S_{it} = \tilde{S}_{it} + e_{it}, \qquad (2)$$

* Tel.: +1 404 894 4453; fax: +1 404 894 1890.
E-mail address: levent.kutlu@econ.gatech.edu.

[1] We follow Kumbhakar and Tsionas (2005) and Brissimis et al. (2009).

0165-1765/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.econlet.2012.10.004

where C_{it} is the total cost for firm i, Y_{it}^{k} is the kth output for firm i, W_{it}^{m} is the price index for the mth input for firm i, $u_{it} \geq 0$ is a random variable representing the total cost of inefficiency, $v_{it} \sim N(0, \sigma_v^2)$ is the traditional error term, S_{it} is the vector of observed input shares,[2] $\tilde{S}_{it}$ is the optimal input share implied by the Shephard's lemma, and $e_{it} \sim N(0, \Sigma_e)$ is the error term for the share equations.[3] We assume that u_{it}, v_{it}, and e_{it} are mutually independent. MLSE estimates Eq. (1) and MLSYS estimates the system. For MLSE and MLSYS, we assume that $u_{it} = u_i \sim \mathbf{N}^{+}(0, \sigma_u^2)$. MLSYS naively assumes that the error term for the share equations is independent from the inefficiency term in the cost equation. This assumption is inconsistent, because the error term in the share equations cannot be independent from the cost of allocative inefficiency.

The final estimator, MLKT, is a system estimator that approximates Kumbhakar's (1997) exact model.[4] This is done by using a first-order Taylor series approximation of the cost of allocative inefficiency and the allocative inefficiency error term around zero allocative inefficiency. Kumbhakar (1997) assumes a translog functional form for the cost function as given in Eq. (1). Shephard's lemma implies the following input share equations:

$$\tilde{S}_{it}^{m} = \frac{\partial \ln(\tilde{C}_{it})}{\partial \ln(W_{it}^{m})} = \beta_{W_m} + \sum_n \beta_{W_{mn}} \ln\left(\frac{W_{it}^{n}}{W_{it}^{M}}\right) + \sum_k \beta_{YW_{km}} \ln(Y_{it}^{k}). \tag{3}$$

The observed input share equations are given by

$$S_{it}^{m} = \tilde{S}_{it}^{m} + \xi_{it}^{m}. \tag{4}$$

Let $u_{it} = u_{it}^{T} + u_{it}^{A}$, where $u_{it}^{T}, u_{it}^{A} \geq 0$, with $u_{it}^{T} = B_t^{T} u_i^{T}$, where B_t^{T} is a function of time. Here, u_{it}^{T} and u_{it}^{A} represent the cost of technical inefficiency and the cost of allocative inefficiency, respectively. The cost of allocative inefficiency is modelled by utilizing the following relationship:

$$\frac{Pf_m(x_{it};\beta)}{Pf_M(x_{it};\beta)} = \frac{W_{it}^{m}}{W_{it}^{M}} \exp(\eta_{it}^{m}) = \frac{W_{it}^{m*}}{W_{it}^{M}}, \quad m = 1, 2, \ldots, M-1, \tag{5}$$

where Pf is the production function frontier dual to $\tilde{C}$ and $\eta_{it}^{m} = B_t^{\eta} \eta_i^{m}$, where B_t^{η} is a function of time. In what follows, we set $B_t^{T} = B_t^{\eta} = 1$. It can be shown that[5]

$$u_{it}^{A} = \ln(\tilde{C}_{it}^{*}) - \ln(\tilde{C}_{it}) + \ln(G_{it}), \tag{6}$$

where $\tilde{C}_{it}^{*} = \tilde{C}(Y_{it}, W_{it}^{*}; \beta)$, $W_{it}^{*} = (W_{it}^{1*}, W_{it}^{2*}, \ldots, W_{it}^{M-1*}, W_{it}^{M})$, $G_{it} = \sum_m [\frac{\partial \ln \tilde{C}_{it}^{*}}{\partial \ln W_{it}^{m*}}] \exp(-\eta_{it}^{m}) = \sum_m \tilde{S}_{it}^{m*} \exp(-\eta_{it}^{m})$, and $\tilde{S}_{it}^{m*} = \tilde{S}(Y_{it}, W_{it}^{*}; \beta)$. The input share equations are given by

$$S_{it}^{m} = \frac{\tilde{S}_{it}^{m*}}{G_{it}\exp(\eta_m)} + \xi_{it}^{m} + e_{it}^{m} = \tilde{S}_{it}^{m} + \xi_{it}^{m} + e_{it}^{m}, \tag{7}$$

where $\xi_{it}^{m} = \frac{\tilde{S}_{it}^{m*}}{G_{it}\exp(\eta_{it}^{m})} - \tilde{S}_{it}^{m}$ represents deviations from the optimal input shares due to input-specific allocative inefficiency and e_{it}^{m} is a term added to account for measurement errors. For the translog cost function,

$$u_{it}^{A} = \ln G_{it} + \sum_m \beta_{W_m} \eta_{it}^{m} + \sum_{m,n} \beta_{YW_{mn}} \ln\left(\frac{W_{it}^{m}}{W_{it}^{M}}\right) \eta_{it}^{n} + \frac{1}{2} \sum_{m,n} \beta_{W_{mn}} \eta_{it}^{m} \eta_{it}^{n} + \sum_{k,m} \beta_{YW_{km}} \ln(Y_{it}^{k}) \eta_{it}^{m} \tag{8}$$

$$\xi_{it}^{m} = \frac{\tilde{S}_{it}^{m}(1 - G_{it}\exp(\eta_{it}^{m})) + \sum_n \beta_{W_{mn}} \eta_{it}^{n}}{G_{it}\exp(\eta_{it}^{m})}. \tag{9}$$

As mentioned earlier, in the MLKT model we approximate u_{it}^{A} and ξ_{it} terms by a Taylor series expansion around $\eta_{it} = (\eta_{it}^{1}, \eta_{it}^{2}, \ldots, \eta_{it}^{M-1})' = 0$ to get a closed form expression for the log-likelihood function. Kumbhakar and Tsionas (2005) showed that $u_{it}^{A} \simeq 0$ and $\xi_{it}^{m} \simeq \sum_n H_{it}^{mn} \eta_{it}^{n}$, where $H_{it}^{mn} = \beta_{Wmm} - \tilde{S}_{it}^{m}(1 - \tilde{S}_{it}^{m})$ if $m = n$ and $H_{it}^{mn} = \beta_{Wmn} + \tilde{S}_{it}^{m}\tilde{S}_{it}^{n}$ if $m \neq n$. The distributions for our model are given by $v_{it} \sim \mathbf{N}(0, \sigma_v^2)$, $u_{it}^{T} = u_i^{T} \sim \mathbf{N}^{+}(0, \sigma_u^2)$, $\eta_{it} = \eta_i \sim \mathbf{N}(0, \Sigma_\eta)$, and $e_{it} \sim \mathbf{N}(0, \Sigma_e)$, where v_{it}, u_i^{T}, η_i, and e_{it} are mutually independent.

3. Simulations

We assume that Kumbhakar's (1997) exact model is the true model and that there are only two inputs. The distributions of random variables are given by $v_{it} \sim \mathbf{N}(0, \sigma_v^2)$, $u_{it}^{T} = u_i^{T} \sim \mathbf{N}^{+}(0, \sigma_u^2)$, $\eta_{it} = \eta_i \sim \mathbf{N}(0, \sigma_\eta^2)$, and $e_{it} \sim \mathbf{N}(0, \sigma_e^2)$, where v_{it}, u_i^{T}, η_i, and e_{it} are mutually independent. Moreover, u_{it}^{A} and ξ_{it}^{m} are generated according to Eqs. (8) and (9). The regressors, $X_{it} = [w_{it} y_{it}]'$, are generated by a bivariate VAR model[6]: $X_{it} = RX_{i,t-1} + \delta_{it}$, where $\delta_{it} \sim \mathbf{N}(0, \sigma_\delta^2 I_2)$ and $X_{i1} \sim N(0, \sigma_\delta^2 (I_2 - R^2)^{-1})$.[7] We make sure that the regularity conditions hold: (1) monotonicity with respect to w and y, (2) concavity with respect to w, (3) homogeneity of the degree 1 in w, (4) well-defined input shares, and (5) $u^{A} \geq 0$. The values of y_{it} and w_{it} are shifted by μ_{y_i} and μ_{w_i}, respectively. We used $\mu_{y_i} \sim N(\mu_y, \sigma_{\mu_y}^2)$ and $\mu_{w_i} \sim N(\mu_w, \sigma_{\mu_w}^2)$. At each simulation, we drew a larger number of firms than N and discarded those firms that violated the

Table 1
Simulation results.

Base					
PARAM	FE	CSSW	MLSE	MLSYS	MLKT
β_0	0.4852	0.4551	0.5206	0.5172	0.5173
β_w	0.4996	0.4996	0.4997	0.4989	0.4989
β_y	0.9003	0.9005	0.9002	0.9053	0.9051
β_{ww}	−0.1018	−0.1017	−0.1016	−0.0918	−0.0923
β_{wy}	0.1017	0.1018	0.1016	0.0920	0.0924
β_{yy}	0.0982	0.0980	0.0983	0.0979	0.0979
SIZE	FE	CSSW	MLSE	MLSYS	MLKT
β_0	–	–	0.2510	0.2530	0.2545
β_w	0.0535	0.0640	0.0645	0.6670	0.6650
β_y	0.0590	0.0730	0.0665	0.1010	0.1010
β_{ww}	0.0475	0.0540	0.0540	0.6495	0.6530
β_{wy}	0.0580	0.0590	0.0675	0.6570	0.6635
β_{yy}	0.0555	0.0635	0.0665	0.0695	0.0700
RMSE	0.1253	0.1339	0.1235	0.1483	0.1480
MAE	0.0850	0.0909	0.0838	0.1019	0.1018
RMSEeff	0.0392	0.0771	0.0373	0.0373	0.0488
MAEeff	0.0312	0.0593	0.0298	0.0297	0.0368
Mean u	0.2994	0.3294	0.2639	0.2650	0.2897
True u	0.2837	True Eff	0.7667		

[2] One of the input shares is omitted for obvious reasons.

[3] The linear homogeneity restrictions for the input prices is imposed by normalizing the cost and the prices by the price index for materials.

[4] KT stands for Kumbhakar and Tsionas (2005) as it uses the approximation proposed by them. Indeed, MLKT is a variation of the estimator proposed by Brissimis et al. (2009).

[5] See Kumbhakar (1997).

[6] We followed Park et al. (2003, 2007) and Kutlu (2010).

[7] In contrast to our simulations, Kumbhakar and Wang (2006) considers cross sectional data.