Economics Letters 155 (2017) 5–8

Contents lists available at ScienceDirect

Economics Letters

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

A note on the likelihood ratio test on the equality of group frontiers

Cliff J. Huang [a], Hung-pin Lai [b,*]

[a] Vanderbilt University, United States
[b] National Chung Cheng University, Taiwan

HIGHLIGHTS

- A likelihood ratio test on the equality of group production frontiers is proposed.
- The heterogeneity in the error distribution may exist and should be considered in the LR test.
- We reproduce the results of Rao et al. (2004) and also demonstrate our LR test for comparison.

ARTICLE INFO

Article history:
Received 12 January 2017
Received in revised form
21 February 2017
Accepted 25 February 2017
Available online 7 March 2017

Keywords:
Metafrontier
Likelihood-ratio test
Heterogeneity

ABSTRACT

The empirical applications of the metafrontier analysis have largely been popularized by the publications of Battese et al. (2004), and O'Donnell et al. (2008). Prior to the estimation of the metafrontier function, it is generally suggested to conduct a likelihood-ratio (LR) test to test if all group frontiers are the same. In this short note, we point out that although the null hypothesis of the suggested LR test restricts all frontiers to be the same for all groups, the heterogeneity in the error distribution may exist and should be considered. We propose a restricted LR test as an alternative procedure.

1. Introduction

Central to the metafrontier production analysis is that, while firms in different regions or groups may choose a particular production technology depending on the production environment encountered, there exists a "meta" technology that envelops all the heterogeneous ones. Consider the data generating process (DGP) of the output y_{it}^j of the ith firm in the tth period in the jth group using the input x,

$$y_{it}^j = f^j\left(x;\beta^j\right) + \upsilon_{it}^j - u_{it}^j,\quad i = 1\ldots N_j,$$
$$t = 1\ldots T_i,\ j = 1\ldots J, \tag{1}$$

where $f^j\left(x;\beta^j\right)$ denotes the jth group production technology, β^j is a parameter and x is the input level.[1] A common assumption imposed on the two random components in the stochastic frontier studies is $\upsilon_{it}^j \sim N\left(0,\sigma_{\upsilon j}^2\right)$, $u_{it}^j \sim N^+(0,\sigma_{uj}^2)$, and υ_{it}^j and u_{it}^j are independent to each other. The frontier function f^j and the parameter β^j differentiate the group production technologies.

The metafrontier of the heterogeneous groups is defined as a convex hull, $f^m\left(x;\beta^m\right) \equiv conv\left\{\bigcup_j f^j\left(x;\beta^j\right)\right\}$. Had the heterogeneity in technology not existed, there would be no need for metafrontier production analysis, i.e., $f^m\left(x;\beta^m\right) = f^j\left(x;\beta^j\right)$ for all j.

The empirical applications of the metafrontier analysis have largely been popularized by the publications of Battese et al. (2004), and O'Donnell et al. (2008). Prior to the estimation of the metafrontier function $f^m\left(x;\beta^m\right)$, it is generally suggested that a likelihood-ratio (LR) test of the null hypothesis that all group frontiers are the same be performed by computing the statistic $\lambda = -2\left\{\ln\left(L_{H_0}/L_{H_1}\right)\right\}$, where $\ln L_{H_0}$ is the value of the log-likelihood function for the stochastic frontier estimated by pooling the data for all groups under the homoscedastic variance component assumption ($\sigma_{\upsilon j}^2 = \sigma_\upsilon^2$, $\sigma_{uj}^2 = \sigma_u^2$ for all j), and $\ln L_{H_1}$ is the sum of the values of the log-likelihood functions for the separate group frontiers (e.g., Rao et al., 2004; Battese et al., 2004; O'Donnell et al., 2008).

Since the specification $f^j(x;\beta^j)$ in (1) defines the jth group production technology and $\left(\sigma_{\upsilon j}^2,\ \sigma_{uj}^2\right)$ defines the distributions

* Corresponding author.
E-mail addresses: cliff.huang@vanderbilt.edu (C.J. Huang), ecdhpl@ccu.edu.tw (H.-p. Lai).

[1] A production technology or production function is a function $f(x) = \max\{y|y \in P(x)\} = \max\{y|x \in L(y)\}$ where $P(x)$ is the output set and $L(y)$ is the input set (see p. 26 in Kumbhakar and Lovell, 2000). The defined jth group production function $f^j\left(x;\beta^j\right)$ is a parametric representation of $f(x)$.

http://dx.doi.org/10.1016/j.econlet.2017.02.037
0165-1765/

of the group's random noise and inefficiency, empirically it is inconceivable that the group's variance components would be homoscedastic. In this short note, we point out that the above-suggested LR test of pooling the data under the hypothesis H_0 that all group frontiers are identical, i.e., $f^m(x;\beta^m) = f^j(x;\beta^j)$ for all j, is inappropriate when the group's variance components are heteroscedastic, and propose a restricted LR test as an alternative procedure.[2]

2. Likelihood ratio test on the equality of group production frontiers

Define the composite error $\varepsilon_{it}^j = \upsilon_{it}^j - u_{it}^j$. It can be shown that the composite error of the model given by (1) has the following density function,

$$h\left(\varepsilon_{it}^j\right) = \frac{2}{\sqrt{\sigma_{\upsilon j}^2+\sigma_{uj}^2}}\phi\left(\frac{\varepsilon_{it}^j}{\sqrt{\sigma_{\upsilon j}^2+\sigma_{uj}^2}}\right)\Phi\left(\frac{-\varepsilon_{it}^j\sigma_{uj}/\sigma_{\upsilon j}}{\sqrt{\sigma_{\upsilon j}^2+\sigma_{uj}^2}}\right), \tag{2}$$

where $\phi(.)$ and $\Phi(.)$ are the probability density function (pdf) and cumulative distribution function (cdf) of the standard normal distribution, respectively. The log-likelihood function of the N_j firms in the jth group is

$$\begin{aligned}\ln L^j &= \sum_{i=1}^{N_j}\sum_{t=1}^{T_i}\ln h\left(\varepsilon_{it}^j\right)\\ &= \sum_{i=1}^{N_i}\sum_{t=1}^{T_i}\left\{\ln 2-\frac{1}{2}\ln\left(\sigma_{\upsilon j}^2+\sigma_{uj}^2\right)\right.\\ &\quad\left.+\ln\phi\left(\frac{\varepsilon_{it}^j}{\sqrt{\sigma_{\upsilon j}^2+\sigma_{uj}^2}}\right)+\ln\Phi\left(\frac{-\varepsilon_{it}^j\sigma_{uj}/\sigma_{\upsilon j}}{\sqrt{\sigma_{\upsilon j}^2+\sigma_{uj}^2}}\right)\right\}.\end{aligned} \tag{3}$$

The within group one-sided errors u_{it}^j in the above log-likelihood function (3) are assumed to be independently, identically, distributed (i.i.d.). However, in the panel model of Battese and Coelli (1992) with the time-varying decaying inefficiency, the one-sided error is generally specified as a function of time t,

$$u_{it}^j = e^{-\eta(t-T)}u_i^j, \tag{4}$$

where $u_i^j \sim N^+(0,\sigma_{uj}^2)$ is time-invariant but varies with groups. It can then be shown that the composite error has the following alternative log-likelihood function,

$$\begin{aligned}\ln L^j &= -\frac{1}{2}\left(\sum_{i=1}^{N_j}T_i\right)\left(\ln(2\pi)+\ln\sigma_{sj}^2\right)\\ &\quad-\frac{1}{2}\sum_{i=1}^{N_j}(T_i-1)\ln\left(1-\gamma_j\right)\\ &\quad-\frac{1}{2}\sum_{i=1}^{N_j}\ln\left\{1+\left(\sum_{t=1}^{T_i}\eta_{it}^{j2}-1\right)\gamma_j\right\}+N_j\ln 2\\ &\quad+\sum_{i=1}^{N_j}\ln\left\{1-\Phi\left(-z_i^*\right)\right\}+\frac{1}{2}\sum_{i=1}^{N_j}z_i^{*2}\\ &\quad-\frac{1}{2}\sum_{i=1}^{N_j}\sum_{t=1}^{T_i}\frac{\varepsilon_{it}^{j2}}{\left(1-\gamma_j\right)\sigma_{sj}^2},\end{aligned} \tag{5}$$

where $\sigma_{sj}^2 = \left(\sigma_{\upsilon j}^2+\sigma_{uj}^2\right)$, $\gamma_j = \sigma_{uj}^2/\sigma_{sj}^2$, $\eta_{it}^j = e^{-\eta_j(t-T_i)}$, and

$$z_i^* = \frac{-\gamma_j\sum_{t=1}^{T_i}\eta_{it}^j\varepsilon_{it}^j}{\left[\gamma_j\left(1-\gamma_j\right)\sigma_{sj}^2\left\{1+\left(\sum_{t=1}^{T_i}\eta_{it}^{j2}-1\right)\gamma_j\right\}\right]^{1/2}}.$$

For the empirical metafrontier analysis, one often needs to conduct the likelihood-ratio test to examine whether all group frontiers are the same, i.e., to test the hypothesis H_0,

$$H_0: f^1\left(x;\beta^1\right) = f^2\left(x;\beta^2\right) = \cdots = f^J\left(x;\beta^J\right). \tag{6}$$

Under the hypothesis H_0, denote $f^m(x;\beta^m) = f^j\left(x;\beta^j\right)$ for all j. Then the group's stochastic frontier regression (1) becomes

$$y_{it}^j = f^m\left(x;\beta^m\right)+\upsilon_{it}^j-u_{it}^j,\quad i=1\ldots N_j,\ t=1\ldots T_i,\quad j=1\ldots J, \tag{7}$$

with the composite error,

$$\varepsilon_{it}^j = y_{it}^j - f^m\left(x;\beta^m\right). \tag{8}$$

The log-likelihood value, $\ln L_{H_0}$, of (7) under the hypothesis that all group frontiers are identical is obtained by the maximization

$$\ln L_{H_0} = \max_{\left\{\beta^m,\sigma_{\upsilon 1}^2,\sigma_{u1}^2,\ldots,\sigma_{\upsilon J}^2,\sigma_{uJ}^2\right\}}\sum_{j=1}^{J}\ln L^j, \tag{9}$$

where $\ln L^j = \sum_{i=1}^{N_j}\ln f(\varepsilon_{i1}^j,\ldots,\varepsilon_{iT_i}^j;\beta^m,\ \sigma_{\upsilon j}^2,\sigma_{uj}^2)$ depends on the distribution assumptions of υ_{it}^j and u_{it}^j, and ε_{it}^j is defined in (8). Alternatively, if the group production technologies are heterogeneous, then the group frontiers are estimated separately without imposing the constraint that $f^m(x;\beta^m) = f^j\left(x;\beta^j\right)$ for all j. The unconstrained log-likelihood value, $\ln L_{H_1}$, of (7) is obtained by the maximization

$$\ln L_{H_1} = \sum_{j=1}^{J}\max_{\left\{\beta^1,\ldots,\beta^J,\sigma_{\upsilon 1}^2,\sigma_{u1}^2,\ldots,\sigma_{\upsilon J}^2,\sigma_{uJ}^2\right\}}\ln L^j, \tag{10}$$

where $\ln L^j = \sum_{i=1}^{N_j}\ln f(\varepsilon_{i1}^j,\ldots,\varepsilon_{iT_i}^j;\beta^j,\sigma_{\upsilon j}^2,\sigma_{uj}^2)$ and the composite error is,

$$\varepsilon_{it}^j = y_{it}^j - f^j\left(x;\beta^j\right). \tag{11}$$

The likelihood-ratio test of hypothesis (6) involves computing the statistic

$$\lambda = -2\left\{\ln L_{H_0}-\ln L_{H_1}\right\}, \tag{12}$$

which has a Chi-square (χ^2) distribution with degrees of freedom $d = \sum_{j=1}^{J}dim\left(\beta^j\right) - dim\left(\beta^m\right)$, where $dim\,(.)$ is the dimension of the parameter.

We emphasize that the maximization in (9) under the hypothesis that all group frontiers are the same is not equivalent to the maximum likelihood estimation of a single stochastic frontier regression (7) by pooling the data of all groups to compute the log-likelihood value $\ln L_{H_0}$, which is commonly suggested in empirical studies (e.g., Rao et al., 2004; Battese et al., 2004; O'Donnell

[2] If the truth is indeed homoscedastic, i.e., $\sigma_{\upsilon j}^2 = \sigma_\upsilon^2$, $\sigma_{uj}^2 = \sigma_u^2$ for all j, the LR test of identical group frontiers by assuming full heterogeneous variance components, i.e., $\sigma_{\upsilon j}^2 \neq \sigma_{\upsilon j'}^2$ and $\sigma_{uj}^2 \neq \sigma_{uj'}^2$ would still be valid, but less inefficient since more variance parameters are estimated. On the contrary, if the variances were heterogeneous, the LR test under the assumption of homogeneous variances would be invalid. In either case, however, the LR test of identical group frontiers under the heterogeneous variance assumption is valid. We appreciate an anonymous referee's insightful comment.