Available online at www.sciencedirect.com

ScienceDirect

Mathematics and Computers in Simulation 102 (2014) 131–143

MATHEMATICS AND COMPUTERS IN SIMULATION

www.elsevier.com/locate/matcom

Original article

Inference in stochastic frontier analysis with dependent error terms

Rachida El Mehdi [a], Christian M. Hafner [b,*]

[a] Mohammed First University of Oujda and Institut de Statistique, Biostatistique et Sciences Actuarielles, Université Catholique de Louvain, Voie du Roman Pays, 20, B-1348 Louvain-la-Neuve, Belgium

[b] Institut de Statistique, Biostatistique et Sciences Actuarielles and CORE, Université Catholique de Louvain, Voie du Roman Pays, [illegible] B-1348 Louvain-la-Neuve, Belgium

Received 16 August 2011; received in revised form 10 June 2013; accepted [illegible] September 2013
Available online 28 October 2013

Abstract

Stochastic frontier analysis (SFA) is often used to estimate technical efficiency of en[illegible]ch as firms, countries or municipalities. A potential dependence between the two components of the error [illegible]n be taken into account by a copula function. While estimation of the model is straightforward using the Corrected Ordinary Least Squares (COLS) and Maximum Likelihood (ML) methods, an open issue concerns the inference of the technical efficiencies. We propose a parametric bootstrap algorithm which is suitable for the dependence case. This allows us to estimate the efficiency percentile confidence intervals. We apply the model to the estimation of technical efficiencies of Moroccan municipalities.

Keywords: Bootstrap; Copulas; Efficiency; Inference; Stochastic frontier an[illegible]

1. Introduction

Efficiency analysis has often [illegible]ed out using nonparametric frontier models such as the Data Envelopment Analysis (DEA) or the Free Disposal Hull (FDH). An alternative approach is to use Stochastic Frontier Analysis (SFA), which includes an error term so that deviations from the frontier can be purely random without necessarily indicating ineffice[illegible]. SFA can be formulated both in a parametric or nonparametric framework, but the parametric SFA has certainly been predominant in the literature and in applications. The basic idea of all approaches is the comparison between the Decision Making Unit (DMU, firms for example) in order to know how inputs are used to produce outputs and the c[illegible]is ba[illegible]e Technical Efficiency (*TE*) score achieved by each unit. By definition, technical efficie[illegible] reflects t[illegible]bility of the firm to obtain maximal output from a given set of inputs.

The [illegible]parametri[illegible]ontier approach using DEA or FDH requires minimal assumptions regarding the structure of the produc[illegible] and d[illegible] not impose restrictions on the functional form relating inputs and outputs. It does not account for noise in t[illegible]so it implicitly assumes that every deviation from the frontier is considered as inefficiency.

However, in the parametric SFA, assumptions have to be made both about the functional form and the distribution of the two types of error, namely, the symmetric stochastic error term and the divergence of observations from the efficient frontier. This stochastic frontier approach in the efficiency analysis was simultaneously and independently introduced

* Corresponding author.
E-mail addresses: rachida.elmehdi@uclouvain.be (R. El Mehdi), christian.hafner@uclouvain.be (C.M. Hafner).

0378-4754/$36.00 © 2013 IMACS. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.matcom.2013.09.011

by [2,13]. Later, several extensions have been proposed by, for example [1,8,10,18,19]. A FRONTIER software was developed by [6] in order to estimate the stochastic frontier production and the cost function in the case where the two components of the error term are independent. This software is now also available in the statistical computation environment R, see [16]. As a consequence of its increasing computational availability, stochastic frontier analysis has been widely applied in several areas.

Recently, [19] has proposed an SFA model allowing for dependence between the two error components. The dependence can be explicitly modelled using copula functions, while maintaining typical assumptions about the marginal distribution of the error terms. Estimation of the model using the Corrected Ordinary Least Squares (COLS) and Maximum Likelihood (ML) methods is straightforward but can be computationally challenging. Furthermore, inference about the technical efficiencies is not standard. In this paper, we propose a bootstrap procedure which is an extension of an algorithm proposed by [18] to the copula case. This allows to obtain not only point estimates, but also confidence intervals for the estimated technical efficiencies.

We apply the model to the estimation of technical efficiencies of Moroccan municipalities defining operating receipts as input and financial autonomy as output. The model is estimated with alternative distributions for the one-sided error term, as well as alternative copulas. The best model is selected using classical information criteria. The obtained bootstrap confidence intervals for the technical efficiency estimates are narrow, confirming the adequacy of our methodology and the interpretation of the results. We find that, contrary to common understanding, no municipality in the central regions of the country is close to the frontier.

The remainder of the paper is organized as follows. Section 2 gives an overview of parametric SFA and its history, Section 3 presents the model with dependent error terms and explains the estimation and inference using the bootstrap. Section 4 presents the application of the proposed methodology, and finally the conclusions will summarize the analysis.

2. Parametric stochastic frontier models

Classical parametric stochastic frontier models assume that [illegible] a production function f that converts $X \in \mathbb{R}_+^p$, a vector of inputs of dimension p, into a scalar output $Y \in \mathbb{R}_+$. Supposing that one has n observations of (X_i, Y_i) the model can be written for the ith DMU as

$$y_i = f(x_i, \beta) + \varepsilon_i, \quad i = 1, \ldots, n \tag{1}$$

where $y_i = \log(Y_i)$, $x_i = \log(X_i)$, β is a vector of parameters of dimension $l+1$ to be estimated, and ε_i is a stochastic error term. The function $f(x_i, \beta)$ is interpreted as the production frontier.

The stochastic term ε_i contains information about both the noise and the inefficiency. It can be decomposed into a technical inefficiency and a noise term which can be estimated. In particular, a typical specification is given by

$$\varepsilon_i = v_i - u_i, \tag{2}$$

where v_i is a Gaussian error term, ($v_i \sim N(0, \sigma_V^2)$), and u is a stochastic error term with non-negative support ($u_i \geq 0$, a.s.).

Note that the stochastic component v_i that describes random noise affecting the production process is not directly attributable to the producer or the underlying technology. The noise may come from weather changes, economic adversities, etc. The other component, u_i, measures technical inefficiency in the sense that it measures the shortfall of output y_i from its maximal possible value given by the stochastic frontier ($f(x_i, \beta) + v_i$) and it is equal to zero for a technically efficient decision unit. Then, the one-sided error term $u_i \geq 0$ allows the distinction between DMU (e.g. firms) that are on the frontier ($u_i = 0$) and others that are below the frontier ($u_i > 0$).

The stochastic model then permits to estimate β and its standard errors and, consequently, to make statistical tests of hypotheses. However, one of the criticisms of this model is that there is no *a priori* justification for the selection of the distributional form for u_i. Several choices have been made in the literature, see e.g. the overview of [12], for example, the exponential, the half-normal, the truncated normal or the Gamma distribution. Furthermore, in order to decompose the error term ε into its two components, one has to make assumptions on their dependence. Classical SFA assumes that they are independent. Let us first recall this approach, see e.g. [11].