Journal of Econometrics 190 (2016) 360–373

Contents lists available at ScienceDirect

Journal of Econometrics

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

Unobserved heterogeneity and endogeneity in nonparametric frontier estimation

Léopold Simar [a,*], Anne Vanhems [b], Ingrid Van Keilegom [a]

[a] Institut de Statistique, Biostatistique et Sciences Actuarielles, Université catholique de Louvain, Voie du Roman Pays 20, B1348 Louvain-la-Neuve, Belgium

[b] University of Toulouse, Toulouse Business School and Toulouse School of Economics, France

ARTICLE INFO

Article history:
Available online 23 June 2015

JEL classification:
C13
C14
C49
D24

Keywords:
Unobserved heterogeneity
Endogeneity
Nonparametric frontiers
Robust estimation of frontiers
Conditional efficiency

ABSTRACT

In production theory, firm efficiencies are measured by their distances to a production frontier. In the presence of heterogeneous conditions (like environmental factors) that may influence the shape and the position of the frontier, traditional measures of efficiency obtained in the space of inputs/outputs are difficult to interpret, since they mix managerial inefficiency and shift of the frontier. This can be corrected by using nonparametric conditional efficiencies. In this paper we extend these concepts in the case where the heterogeneity is not observed. We propose a model where the heterogeneity variable is linked to a particular input (or output). It is defined as the part of the input (or the output), independent from some instrumental variable through a nonseparable nonparametric model. We discuss endogeneity issues involved in this model. We show that the model is identified and analyze the asymptotic properties of proposed nonparametric estimators. When using FDH estimators we achieve a limiting Weibull distribution, whereas when using the robust order-m estimators we obtain the asymptotic normality. The method is illustrated with some simulated and real data examples. A Monte-Carlo experiment shows how the procedure works for finite samples.

1. Introduction

In production theory and efficiency analysis, the technical efficiency of a production unit (a firm) is measured by an appropriate distance of this unit to the production frontier, which is the geometrical locus of optimal combinations of the inputs and outputs in the set of attainable production plans (the production set). The economic theory underlying this analysis dates back to the works of Koopmans (1951), Debreu (1951) and Farrell (1957). The first empirical analysis where the production set, its boundary and the resulting efficiency are estimated from a sample of observed units is due to Farrell (1957).

Parametric models have been used in the econometric literature starting from the works of Aigner and Chu (1968) or Greene (1980) for parametric deterministic frontier models and of Aigner et al. (1977), Meeusen and Van den Broek (1977) using stochastic frontier models (see Kumbhakar and Lovell, 2000 and the references therein for a nice overview). Nonparametric approaches have been developed after the pioneering work of Farrell (1957) with the DEA estimator, popularized by Charnes et al. (1978) and the FDH estimator of Deprins et al. (1984). These nonparametric approaches are based on envelopment techniques in the space of inputs and outputs and so are sensitive to outliers and extreme points in the cloud of observed points. Robust alternatives have been proposed in Cazals et al. (2002), Aragon et al. (2005) and Daouia and Simar (2007) using partial frontiers (order-m and order-α quantile frontiers). Today, these nonparametric estimators have been analyzed from a statistical point of view and inference on efficiency estimates is available, mainly by using bootstrap techniques. These nonparametric techniques have been adapted to stochastic frontier models, such as the semi parametric model of Fan et al. (1996) and recently Kneip et al. (2015) (see the recent survey Simar and Wilson, in press for details, and the references therein).

It is now recognized that in the presence of heterogeneous conditions (environmental factors, …) that are not under the control of the producer but that may affect the frontier level, the traditional measures based only on inputs and outputs are difficult to interpret, because the units are benchmarked against a frontier level that may not be attainable under their environmental conditions

* Corresponding author.
E-mail addresses: leopold.simar@uclouvain.be (L. Simar), a.vanhems@tbs-education.fr (A. Vanhems), ingrid.vankeilegom@uclouvain.be (I. Van Keilegom).

http://dx.doi.org/10.1016/j.jeconom.2015.06.015
0304-4076/

(see Simar and Wilson, 2007, 2011, for a detailed discussion).[1] In a sense, these traditional measures combine information on the managerial inefficiency and on the shift of the frontier. A solution to this problem is to use attainable sets and frontier levels that may depend on these heterogeneous conditions: this is the idea of defining conditional efficiency scores, initiated by Cazals et al. (2002) and extended in Daraio and Simar (2005). Here too the statistical properties of these estimators and their robust versions, have been established (in Cazals et al., 2002; Daouia and Simar, 2007; Jeong et al., 2010). As explained in detail in e.g. Daraio and Simar (2005, 2007) and Bădin et al. (2012), the comparison of conditional and unconditional efficiency measures allows to capture the effect of the environmental variable on the shift of the frontier.[2]

So far, these conditional approaches are based on the assumption that these heterogeneous conditions are known and observed. However, this may not be the case. We may infer that some latent factors influence the production process and in particular the set of attainable combinations of inputs and outputs. The objective of our paper is to propose a model where we may have unobserved heterogeneity. In this paper we propose one approach that allows to identify and estimate this latent variable. This will be achieved through a model where the heterogeneity variable is linked to a particular input (or an output). It is defined as the part of the input (or the output), independent from some instrumental variable through a nonseparable nonparametric model. This model involves endogeneity issues that will be discussed. Under usual regularity assumptions, we show that the model is identified, we propose a nonparametric estimator and analyze its asymptotic properties. To the best of our knowledge we are not aware of existing results for handling latent or unobserved heterogeneity in nonparametric frontier models.

The paper is organized as follows. Section 2 summarizes the basic notations and concepts of conditional efficiency measures, and gives the basic model for introducing unobserved heterogeneity in the production process. It also gives the natural nonparametric estimators of the elements of the model and investigates the links with the endogeneity issue. Then, Section 3 establishes the asymptotic properties of our estimators. Section 4 gives a few simple illustrative examples with simulated and real data, and Section 5 indicates how the procedure works for finite samples through a limited Monte-Carlo experiment. Finally, Section 6 contains some conclusions and ideas for future research, Appendix A gives an original way to derive optimal bandwidths for estimating conditional efficiencies sharing monotonicity properties, and Appendix B contains the proofs of the asymptotic results.

2. The model

2.1. Conditional efficiency scores

We first summarize the existing tools for handling the presence of observable heterogeneous (or environmental) factors $Z \in \mathcal{Z} \subseteq \mathbb{R}^d$ in a production process where inputs $X \in \mathbb{R}_+^p$ are used to produce the output $Y \in \mathbb{R}_+$.[3] The production set is the set of technically possible combinations of inputs and output. In the presence of environmental factors, when $Z = z$ this is defined as

$$\Psi(z) = \{(x, y) \in \mathbb{R}_+^p \times \mathbb{R}_+ \mid x \text{ can produce } y, \text{ when } Z = z\}. \quad (2.1)$$

This set is the support of the joint random variable (X, Y), conditionally on $Z = z$. The marginal support of the variables (X, Y) is given by

$$\Psi = \{(x, y) \in \mathbb{R}_+^p \times \mathbb{R}_+ \mid x \text{ can produce } y\} = \bigcup_{z \in \mathcal{Z}} \Psi(z). \quad (2.2)$$

It is the support of the joint marginal distribution of (X, Y). The "separability condition" described in Simar and Wilson (2007, 2011) is the assumption that $\Psi(z) = \Psi$ for all $z \in \mathcal{Z}$. The traditional Farrell–Debreu efficiency score for a firm operating at level (x, y) is defined by the distance in the output direction to the upper boundary of Ψ, and it is given by

$$\lambda(x, y) = \sup\{\lambda \mid (x, \lambda y) \in \Psi\}. \quad (2.3)$$

The separability assumption is a strong assumption, and if it is not fulfilled, the efficiency score $\lambda(x, y)$ is more difficult to interpret since the frontier may not be reachable for the firm facing the environmental conditions z. The same is obviously true for their nonparametric DEA or FDH estimators. Ignoring this factor creates in addition endogeneity issues that are discussed below; we will see that the "separability condition" can be interpreted as partial exogeneity having less severe consequences.

To overcome these difficulties, Cazals et al. (2002) and Daraio and Simar (2005) introduce the concept of conditional frontier and of conditional efficiency score for a unit operating at the level (x, y) and facing conditions z. It is defined as

$$\lambda(x, y|z) = \sup\{\lambda > 0 \mid (x, \lambda y) \in \Psi(z)\}, \quad (2.4)$$

where the production process is described by the conditional distribution of (X, Y) given $Z = z$. It is convenient to characterize this distribution by the probability of being dominated[4]:

$$\begin{aligned} H_{X,Y|Z}(x, y|Z = z) &= P(X \le x, Y \ge y|Z = z) \\ &= S_{Y|X,Z}(y|X \le x, Z = z) F_{X|Z}(x|Z = z), \quad (2.5) \end{aligned}$$

where $S_{Y|X,Z}$ denotes a survival function and $F_{X|Z}$ a cumulative distribution function; we note also the nonstandard conditioning on X $(X \le x)$ and the usual conditioning $Z = z$ for the environmental variables. Under the free disposability assumption,[5] the output conditional score (see e.g. Daraio and Simar, 2005), can also be defined for all x such that $F_{X|Z}(x|Z = z) > 0$ as

$$\begin{aligned} \lambda(x, y|z) &= \sup\{\lambda > 0 \mid H_{XY|Z}(x, \lambda y|Z = z) > 0\} \quad (2.6) \\ &= \sup\{\lambda > 0 \mid S_{Y|X,Z}(\lambda y|X \le x, Z = z) > 0\}. \quad (2.7) \end{aligned}$$

Nonparametric estimators are obtained by plugging-in empirical versions of the probabilities appearing on the right hand side of these equations. The conditioning on $Z = z$ requires the use of smoothing techniques and the derivation of the optimal bandwidth for Z.

In our particular setup where Y is univariate, the frontier can be described by a conditional production function:

$$\phi(x, z) = \sup\{y \mid F_{Y|X,Z}(y|X \le x, Z = z) < 1\}, \quad (2.8)$$

since for univariate y, $F_{Y|X,Z}(y|X \le x, Z = z) = 1 - S_{Y|X,Z}(y|X \le x, Z = z)$. Similarly, the output conditional efficiency score can be defined as

$$\lambda(x, y|z) = \sup\{\lambda > 0 \mid F_{Y|X,Z}(\lambda y|X \le x, Z = z) < 1\}, \quad (2.9)$$

where for univariate Y, it is obvious that $\phi(x, z) = \lambda(x, y|z)\, y$.

---

[1] We do not consider situations where the firms choose their environmental conditions or control them, which is a different story.

[2] It should be noted that Banker and Morey (1986) and Ruggiero (1996) have earlier considered environmental factors in DEA, modeling them as non-discretionary inputs. The use of this appealing approach is limited by some restrictions: (i) the direction of the effect has to be known in advance (favorable or not favorable to efficiency), (ii) the direction of the effect has to be monotone, (iii) these variables must satisfy free disposability and convexity assumptions, like the other inputs–outputs

[3] We will do the presentation in an output orientation where firms try to reach the maximal possible output for a given level of inputs. The same could be done in an input orientation where the firms try to reduce their input (cost) $X \in \mathbb{R}_+$ for a given level of their outputs $Y \in \mathbb{R}_+^q$.

[4] Here and in the sequel, inequalities on vectors should be understood component by component.

[5] The free disposability assumption means that if (x, y) is achievable, then $(\tilde{x}, \tilde{y})$ is also achievable once $(\tilde{x} - x, y - \tilde{y}) \ge 0$. It is technically possible to waste resources.