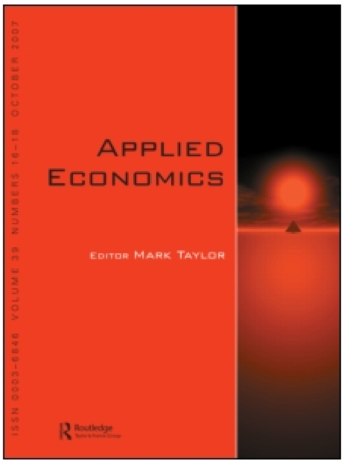

Applied Economics
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713684000

Imposing monotonicity on outputs in parametric distance function estimations

Sergio Perelman[a]; Daniel Santin[b]
[a] CREPP, Université de Liège, B-4000 Liège, Belgium [b] Department of Applied Economics VI, Universidad Complutense de Madrid, Campus de Somosaguas 28223 Pozuelo de Alarcón, Madrid, Spain

First published on: 28 January 2011

To cite this Article Perelman, Sergio and Santin, Daniel(2011) 'Imposing monotonicity on outputs in parametric distance function estimations', Applied Economics,, First published on: 28 January 2011 (iFirst)
To link to this Article: DOI: 10.1080/00036846.2010.493141
URL: http://dx.doi.org/10.1080/00036846.2010.493141

Applied Economics, 2011, 1–11, iFirst

Routledge
Taylor & Francis Group

Imposing monotonicity on outputs in parametric distance function estimations

Sergio Perelman[a] and Daniel Santin[b,*]

[a]*CREPP, Université de Liège, Bd Du Rectorat 7 (B31), B-4000 Liège, Belgium*

[b]*Department of Applied Economics VI, Universidad Complutense de Madrid, Campus de Somosaguas 28223 Pozuelo de Alarcón, Madrid, Spain*

The technology set involved in the estimation of a multi-output production frontier theoretically implies monotonicity on outputs. This is because an efficient firm cannot reduce the vector of outputs holding the vector of inputs fixed while it still belongs to the frontier. In empirical studies dealing with the estimation of parametric distance functions, this hypothesis is often violated by observations with far from average characteristics. To overcome this limitation, we propose a new approach for allowing the easy imposition of monotonicity on outputs in this context. This methodology is tested in the educational sector using Spanish student level data from the Programme for International Student Assessment (PISA) database. The results indicate that a nonnegligible 8.33% of the production units break the monotonicity assumption. Furthermore, although there is no statistically significant difference in efficiency distribution by school ownership, our methodology helps to detect a slight worse mathematical performance for students attending public schools.

I. Introduction

In the past decade, the parametric approach for measuring technical efficiency has gained growing consideration in applied economics. The main reason to explain this fact is that the parametric distance function allows modelling multi-input multi-output production problems without aggregation as it is done in a nonparametric Data Envelopment Analysis (DEA). However, the parametric tool is especially appealing for applied researchers because it allows to easily calculate production elasticities to help policy-makers and private managers in their decisions. Multi-output multi-input production technologies are frequently used in public services (education, health social services, etc.) as well as in other service activities generally operated by private companies (transportation, banking or insurance companies). Education is one of those multi-input multi-output economics fields in which developing consistent indexes of school performance and promoting yardstick competition in the sector can lead to improved human capital and reduced school failure without increasing budget efforts.

Due to specialization, it often happens that some Decision-Making Units (DMUs) produce proportionally more in one output than in others. For example, we consider the transportation of passengers and tons of freight by railways companies, other than firms only devoted to passenger or to

*Corresponding author. E-mail: dsantin@ccee.ucm.es

Applied Economics ISSN 0003–6846 print/ISSN 1466–4283 online © 2011 Taylor & Francis
http://www.informaworld.com
DOI: 10.1080/00036846.2010.493141

freight transportation. However, we are unlikely to find companies with extremely high, or extremely low, passenger transportation proportions. Likewise, in education, it is quite usual to observe students better prepared or motivated in some of the subjects with respect to others; as well as in the transportation example named above it is unusual to find students with, say, outstanding results in mathematics (reading) together with a very low reading (mathematics) level. As a consequence of the lack of extremely specialized DMUs, the econometric estimation of the corresponding parametric output distance function (e.g. Coelli and Perelman, 1999, 2000) will probably indicate a violation of microeconomic regularity conditions, mainly monotonicity, for some of the evaluated DMUs. The violation of monotonicity on outputs can damage the reliability of efficiency measures. This is because an efficient firm cannot reduce the vector of outputs holding the vector of inputs fixed while it still belongs to the frontier.

O'Donnell and Coelli (2005) proposed a Bayesian approach allowing the imposition of regularity conditions, among them monotonicity on outputs. In this article, we propose a new alternative approach which has the advantage of easy computation. To simplify, we only show the imposition of monotonicity on outputs for an output distance function.[1] This approach consists of the deterministic computation of an extra-distance radial component together with a set of output slacks for firms breaking the monotonicity assumption. The same approach could be used in such cases in which the monotonicity assumption would be considered as economically plausible, e.g. to measure congestion inefficiency or in the presence of bad output production restrictions for considering regulatory inefficiency (Färe *et al.*, 1984).

In order to illustrate the potentialities of the approach proposed here, we provide an application to Spanish educational data from the Programme for International Student Assessment (PISA), implemented in 2000 by the Organization for Economic Co-operation and Development (OECD). In this article, we employ the pupil as a DMU because it allows two main advantages. First, in efficiency educational research, student results are typically aggregated at school (Cordero-Ferrera *et al.*, 2008), district level (Banker *et al.*, 2004) or countries (Afonso and St. Aubyn, 2006), imposing a considerable limitation to simultaneously include the effect of a student's own background and the peer-group effect as different variables (Santín, 2006). Second, Hanushek *et al.* (1996) showed how aggregation can dramatically influence the statistical significance of inputs in the educational process upwards. We also investigate differences in student performance across Spanish public, private government dependent and private government independent schools and conclude that, once school inputs, student background and peer-group characteristics are taken into account, there is no statistically significant difference in efficiency distribution of the school regarding public–private ownership. However, our methodology is able to detect a slight worse mathematic performance of students attending public school.

This article is organized as follows. Section II presents the main properties and characteristics of parametric output distance functions. In Section III, we describe the procedure for imposing monotonicity on the output distance function. Section IV shows the Spanish educational data from the PISA database employed in the empirical application. Section V presents estimation results and illustrates the steps to impose monotonicity on outputs in order to obtain the corrected measurements of technical inefficiency. Section VI focusses on the main conclusions.

II. Measuring Efficiency Through Distance Functions

In defining a vector of inputs $x=(x_1,\ldots,x_K)\in\Re^{K+}$ and a vector of outputs $y=(y_1,\ldots,y_M)\in\Re^{M+}$, a feasible multi-input multi-output production technology can be defined using the output possibility set $P(x)$, which can be produced using the input vector x:

$P(x)=\{y\colon x \text{ can produce } y\}$, which is assumed to satisfy the set of axioms described by Färe and Primont (1995). This technology can also be defined as the output distance function proposed by Shephard (1970)

$$D_O(x,y)=\inf\{\theta:\theta>0,(x,y/\theta)\in P(x)\}$$

If $D_O(x,y)\leq 1$, then (x,y) belongs to the production set $P(x)$. In addition, $D_O(x,y)=1$, if y is located on the outer boundary of the output possibility set. In order to estimate the distance function in a parametric setting, a translog functional form is assumed. According to Coelli and Perelman (2000), this specification fulfils a set of desirable characteristics for its empirical estimation: flexible, easy to derive and allowing the imposition of homogeneity.

[1] The procedure can be easily extended to impose monotonicity on inputs also in an output distance function and monotonicity on outputs and inputs in an input distance function.

The translog output distance function specification herein adopted for the case of K inputs and M outputs is given as

$$\ln D_{Oi}(x,y) = \alpha_0 + \sum_{m=1}^{M}\alpha_m \ln y_{mi} + \frac{1}{2}\sum_{m=1}^{M}\sum_{n=1}^{M}\alpha_{mn} \times \ln y_{mi}\ln y_{ni} + \sum_{k=1}^{K}\beta_k \ln x_{ki} + \frac{1}{2}\sum_{k=1}^{K}\sum_{l=1}^{K}\beta_{kl}\ln x_{ki}\ln x_{li} + \sum_{k=1}^{K}\sum_{m=1}^{M}\delta_{km}\ln x_{ki}\ln y_{mi}, \quad i = 1,2,\ldots,N \tag{1}$$

where i denotes the ith unit (DMU) in the sample. In order to obtain the production frontier surface, we set $D_O(x,y) = 1$, which implies that $\ln D_O(x,y) = 0$.

The parameters of the above-mentioned distance function must satisfy a number of restrictions, among them symmetry and homogeneity of degree +1 in outputs. This latter restriction indicates that distances with respect to the boundary of the production set are measured by radial expansions.

According to Lovell *et al*. (1994), normalizing the output distance function by one of the outputs is equivalent to imposing homogeneity of a degree +1. Therefore, Equation 1 can be represented as follows:

$$\ln(D_{Oi}(x,y)/y_{Mi}) = \mathrm{TL}(x_i, y_i/y_{Mi}, \alpha, \beta, \delta), \quad i = 1,2,\ldots,N$$

where

$$\mathrm{TL}(x_i, y_i/y_{Mi}, \alpha, \beta, \delta) = \alpha_0 + \sum_{m=1}^{M-1}\alpha_m \ln(y_{mi}/y_{Mi}) + \frac{1}{2}\sum_{m=1}^{M-1}\sum_{n=1}^{M-1}\alpha_{mn} \times \ln(y_{mi}/y_{Mi})\ln(y_{ni}/y_{Mi}) + \sum_{k=1}^{K}\beta_k \ln x_{ki} + \frac{1}{2}\sum_{k=1}^{K}\sum_{l=1}^{K}\beta_{kl}\ln x_{ki}\ln x_{li} + \sum_{k=1}^{K}\sum_{m=1}^{M-1}\delta_{km}\ln x_{ki}\ln(y_{mi}/y_{Mi})$$

Rearranging the terms, the above function can be rewritten as follows:

$$-\ln(y_{Mi}) = \mathrm{TL}(x_i, y_i/y_{Mi}, \alpha, \beta, \delta) - \ln D_{Oi}(x,y), \quad i = 1,2,\ldots,N$$

where $-\ln D_{Oi}(x,y)$ corresponds to the radial distance from each point to the boundary. This deterministic framework can be estimated using the Corrected Ordinary Least Squares (COLS) method used by Lovell *et al*. (1994), the Parametric Linear Programming (PLP) method proposed for translog output distance functions by Färe *et al*. (1993) and the stochastic frontier analysis provided by Aigner *et al*. (1977).

On the one hand, the flexibility of the translog function is very useful for capturing possible second-order nonlinear relationships among the variables. However, on the other hand this specification can break the microeconomic assumption of monotonicity on outputs for some of the firms in empirical estimations. In this article, we provide a simple procedure to overcome this drawback.

III. Imposing Monotonicity on the Output Distance Function

According to O'Donnell and Coelli (2005), monotonicity on outputs implies the imposition of a condition on output distance function partial derivatives with respect to output defined by

$$r_m = \frac{\partial \ln D}{\partial \ln y_m} = \alpha_m + \sum_{n=1}^{M}\alpha_{mn}\ln y_n + \sum_{k=1}^{K}\delta_{km}\ln x_k$$

For D to be nondecreasing in y, it is required that

$$h_m = \frac{\partial D}{\partial y_m} = \frac{\partial \ln D}{\partial \ln y_m}\frac{D}{y_m} = r_m\frac{D}{y_m} \geq 0 \Leftrightarrow r_m \geq 0$$

The slope of the distance function between the two outputs, i.e. the Marginal Rate of Transformation (MRT), can be denoted as follows:

$$\mathrm{MRT}_{y_m y_n} = -\frac{\partial \ln y_m}{\partial \ln y_n} = -\frac{r_n}{r_m}$$

This expression fulfils monotonicity on outputs when $\mathrm{MRT}_{y_m y_n} \leq 0$. In the case that this condition was violated for some DMUs, as very often occurs in empirical studies, it can be imposed in a simple way, as illustrated in Fig. 1 for a two-output setting.

Theoretically, when a DMU such as A and B in Fig. 1 exhibits a positive slope on the estimated deterministic production frontier (FF') projection points A' and B', we can re-compute the distance to a new frontier after drawing *a strict production frontier* (GG' in Fig. 1), which fulfils the monotonicity assumption. This implies the adding up of a new extra distance component, $-\ln D_{Oi}^{\text{extra}}(x,y)$, $A'A''$ and $B'B''$

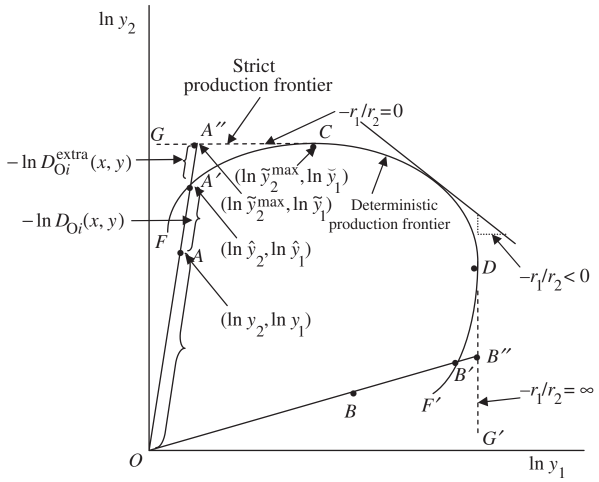

Fig. 1. Imposing monotonicity on outputs in a two outputs distance function

for A and B DMUs, respectively. In practice, we proceed in five steps as follows.

Step 1: This consists of the computation of the predicted efficient output vector on the estimated deterministic production frontier, hat denoted $\ln(\hat{y}_{ni})$, on behalf of the estimated production frontier parameters (points F to F' in Fig. 1).

$-\ln(\hat{y}_{Mi}) = \mathrm{TL}(x_i, y_i/y_{Mi}, \hat{\alpha}, \hat{\beta}, \hat{\delta})$, for the normalization output and

$-\ln(\hat{y}_{ni}) = -\ln(\hat{y}_{Mi}) - \ln(y_{ni}/y_{Mi})$, for the other outputs, using the output ratio relationships.

Step 2: Following the estimated output distance function parameters, we calculate $\mathrm{MRT}_{y_m y_n}$ for all DMUs focussing our interest only in points breaking the monotonicity on outputs $\mathrm{MRT}_{y_m y_n} > 0$ requirement before continuing with Step 3.

Step 3: This comprises of the computation of the output projection vector corresponding to the strict frontier, tilde denoted $\ln \tilde{y}_{ni}$, those are points A'' and B'' according to Fig. 1. To do this, we proceed as follows. First, we calculate output distance function partial derivatives with respect to all the outputs in order to detect DMUs where r_n is less than zero. Let us assume that we start with output M and DMU A, $r_{MA} < 0$. Once we know a DMU as A breaks monotonicity on M, our aim is to search for the maximum values $\ln \tilde{y}_{ni}^{\max}$ of the other outputs in the estimated distance frontier with giving A inputs endowment to remaining DMUs whatever ratio relationships they have. These maximum observed values are assigned to DMU A, projecting the M output holding the exogenous output ratios of DMU A constant.

$$\ln \tilde{y}_{ni}^{\max} = \max\left[\mathrm{TL}\left(x_A, y_i/y_{Mi}, \hat{\alpha}, \hat{\beta}, \hat{\delta}\right)\right]$$

$$\ln \tilde{y}_{MA} = \ln \tilde{y}_{ni}^{\max} - \ln\left(\frac{y_{nA}}{y_{MA}}\right)$$

Step 4: Finally, the new efficiency scores for each DMU are computed by adding up the *extra distance* term $-\ln D_{\mathrm{O}i}^{\mathrm{extra}}(x, y)$ to the estimated distance $\ln D_{\mathrm{O}i}(x, y)$. Here, we can separate the computed production frontier output vector, $(\ln(\hat{y}_{Mi}), \ln(\hat{y}_{ni}))$, from the strict production frontier output vector, $(\ln(\tilde{y}_{Mi}), \ln(\tilde{y}_{ni}^{\max}))$. The corresponding extra distance for DMUs A and B are therefore graphically measured in Fig. 1 by the Euclidean distances between OA' and OA'' and OB' and OB'', respectively. For DMU i, we obtain

$$\begin{aligned} &\ln D_{\mathrm{O}i}^{\mathrm{extra}}(x, y) \\ &\quad = d(A'A'') \\ &\quad = \sqrt{(\ln \hat{y}_{MA} - \ln \tilde{y}_{Mi})^2 + \sum_{n=1}^{M-1} \left(\ln \hat{y}_{ni} - \ln \tilde{y}_{ni}^{\max}\right)^2} \end{aligned}$$

Step 5: The radial expansion of a DMU breaking monotonicity on outputs to the strict production function originates a production target that presents an *output slack*.[2] As it is shown in Fig. 1, the A'' (B'') radial projection points are inefficient because DMU A (B) could produce more on output 1 (output 2) holding output 2 (output 1) constant, achieving point C (point D). The movement from A'' to C implies that the DMU could change their output ratio values. Sometimes, this could not be possible if these ratios are exogenously imposed (for a regulator, a politician, preferences, prices, etc.). For this reason, we will only apply this fifth step if the change is feasible in the analysed sector. In Point C $\ln \tilde{y}_{ni}^{\max}$ has the same value than in Step 3. The new target $\ln \breve{y}_{Mi}$ for DMU i to hold monotonicity on output M will be

$$\begin{aligned} r_{Mi} = \frac{\partial \ln D}{\partial \ln y_{Mi}} &= \hat{\alpha}_M + \hat{\alpha}_{Mm} \ln \breve{y}_{Mi} + \sum_{n=1}^{M-1} \hat{\alpha}_{Mn} \ln \tilde{y}_{ni}^{\max} \\ &\quad + \sum_{k=1}^{K} \hat{\delta}_{kM} \ln x_{ki} = 0 \end{aligned}$$

[2] The term output slack is new in the parametric frontier analysis but its meaning and interpretation is exactly the same than the output slack calculated with DEA.

Table 1. Descriptive statistics: outputs and inputs at pupil level in Spain by school type

Outputs and inputs \ School type	Variable	Public Mean (SD)	Private government dependent Mean (SD)	Private government independent Mean (SD)
Outputs				
Mathematics score	y_1	498.14 (83.56)	513.07 (81.20)	519.98 (81.32)
Reading score	y_2	519.26 (74.05)	526.39 (73.84)	543.71 (74.39)
Inputs				
School				
Computers/100 students	x_1	6.85 (3.52)	5.75 (5.08)	5.68 (2.84)
Teachers/100 students	x_2	9.02 (1.94)	5.75 (1.20)	5.73 (1.79)
Background				
Mother's level of education	x_3	2.69 (0.77)	2.88 (0.78)	3.04 (0.73)
Father's level of education	x_4	2.78 (0.83)	3.00 (0.80)	3.10 (0.77)
Cultural activities	x_5	2.45 (1.16)	2.62 (1.14)	2.74 (1.21)
Cultural possessions	x_6	2.98 (1.00)	3.17 (0.97)	3.33 (0.90)
Time spent on homework	x_7	3.33 (0.84)	3.40 (0.79)	3.49 (0.75)
Peer group				
Average mother's level of education	x_8	2.78 (0.38)	3.00 (0.46)	3.10 (0.41)
N	2449	1383	829	237

where rearranging terms

$$\ln \breve{y}_{Mi} = \frac{-\hat{\alpha}_M - \sum_{n=1}^{M-1} \hat{\alpha}_{Mn} \ln \tilde{y}_{ni}^{\max} - \sum_{k=1}^{K} \hat{\delta}_{kM} \ln x_{ki}}{\hat{\alpha}_{Mm}}$$

Finally, as is usual in empirical applications, we measure the output slack of output M in absolute magnitudes as follows:

$$\mathrm{OS}_{Mi} = \breve{y}_{Mi} - \tilde{y}_{ni}^{\max}$$

IV. Educational Data

In our empirical analysis, we use data from the PISA, implemented in 2000 by the OECD. PISA tests students in the subjects of reading, mathematics and sciences. Because the home, school and national contexts can play an important role in how students learn, PISA also collects extensive information about such background factors. The entire database comprises 32 countries, but this illustrative study is limited to the Spanish case. Given that the target 15-year-old population tends to be enrolled in different grades, we only selected upper 10th grade students for this study in order to obtain a more homogeneous sample to perform the efficiency analysis. To sum up, the analysis is based on a homogenous population composed of 2449 Spanish students attending 10th grade at 185 different schools, which, in the year 2000, completed the mathematics and reading PISA tests.

It is worth noting that PISA is methodologically highly complex and it exceeds the aims of this empirical application to present a complete explanation of the procedures followed in the sampling design. Nevertheless, for a complete review, OECD (2001, 2002) may be consulted. Table 1 displays descriptive information on the output and input measures used in the analysis by school ownership.

We consider two outputs: the students' scores obtained in the international mathematics and reading tests. As reported in Table 1, average reading scores were higher and at the same time, less widely distributed than mathematics scores. On the other hand, private independent schools and public financed private schools show better mean scores than public schools.

Two school inputs were selected: on the one hand, the *computer/student* ratio (corresponding to the total number of computers in the school divided by the total enrollment) and, on the other hand, the *teacher/student* ratio corresponding to the total teaching staff divided by the total school enrollment (full-time and part-time teachers are accounted for by 1.0 and 0.5, respectively). We think that both inputs are plausible indicators for the level of physical and human capital inside each school. As most students in Spain spend their entire secondary education in the same school, we argue that specific school ratios are better input indicators than those obtained at the (10th grade) classroom level. The *computer/student* ratio as well as the *teacher/student* ratio is higher for public schools. These ratios are very similar in both types of private schools.

We consider five student background inputs. All of these variables are represented by indexes that summarize the answers given by students to a series of related questions. *Mother* and *father's level of education* corresponds to the International Standard Classification of Education (ISCED) (OECD, 1999). The original categories contained in ISCED were redefined as four major possibilities: 1 = did not go to school; 2 = primary school completed; 3 = secondary school completed; and 4 = tertiary education completed. The *cultural activities* index was derived from how often students had participated in the following activities during the preceding year: visiting a museum or art gallery, attending the opera, ballet, a classical symphony or a concert or watching live theatre. The *cultural possessions* index was derived from student reports on the availability of the following items in their home: classical literature, poetry books and works of art. *Time spent on homework* was also derived from student reports on the amount of time they devoted to homework *per* week in reading, mathematics and science. Together with these variables, and taking advantage of using student level data, we introduce a variable to control for potential *peer-group* effects. The variable considered here is the *average mother's level of education* of the peers measured at class level (Rosenzweig and Wolpin, 1994). Given the nature and the treatment applied to the construction of these variables, their variation across the sample is limited. Even so, one can see in Table 1 that for all variables related with the background, the means are higher and the SDs lower in private government independent schools while pupils at public schools have the lower means and the higher variances.

V. Results and Discussion

A parametric output distance function was estimated assuming a stochastic translog technology, as indicated in Section I. Homogeneity of degree +1 was imposed by selecting one of the outputs, the students' scores in mathematics y_1 as the dependent variable, and the ratio y_2/y_1 as the explanatory variable, instead of y_2. However, for presentation purposes, in Table 2 the parameters corresponding to y_1 are reported, as calculated by application of the homogeneity condition.

Two different specifications were estimated in order to test the nonseparability hypothesis among outputs and inputs. For this purpose, following Coelli *et al*. (2005), we conducted a generalized Likelihood Ratio (LR) test, which allows contrasting whether or not input–output cross effect parameters are statistically significant. The null hypothesis was retained on the basis of this test; therefore, the results presented in Table 2 are those corresponding to the separable output distance function. In this case, the null hypothesis is rejected if the LR test exceeds $\chi^2_8(\alpha)$. For $\alpha = 0.05$, the critical value is 15.5, and we obtained LR = 10.74.

Parameter estimates

As is usual for the estimation of translog functions, the original variables y_m $(m = 1, 2)$ and x_k $(k = 1, \ldots, 8)$ were transformed in deviations to mean values. Therefore, the first-order parameters in Table 2 must be interpreted as distance function partial elasticities at mean values. For instance, those corresponding to the reading and mathematics scores are positive and indicate that student performance or efficiency increase (distance functions increase) when, *ceteris paribus*, their reading and mathematics scores increase. The opposite effect is observed for the scores in all first-order coefficients on inputs that are negative. This indicates that, at least at mean values and regardless of second-order effects, student performance decreases (distance functions decrease) when inputs increase. All these first-order coefficients are significant, with the sole exception of both school inputs: *computer/student* and *teacher/student* ratios.

Some general conclusions can, however, be drawn from these results without taking into account the second-order coefficients affecting school inputs. Several of them are statistically significant, e.g. β_{22}, β_{12} and β_{23}, which correspond to the *teacher/student* ratio in its quadratic form and in interaction with the *computer/student* ratio and the *mother's level of education* index, respectively.

In our case, a simpler Cobb–Douglas production function estimation would certainly be unable to discover cross effects between school inputs themselves or when combined with student background and peer-group inputs, and the conclusion would be that school does not matter. Therefore, one of the major advantages of parametric translog output distance function analysis at student level is that it can provide additional insights into the educational production process, overcoming model misspecification problems at the same time.

Table 2. Parametric output distance function estimations

Variables and parameters			*t*-ratio	Variables and parameters			*t*-ratio
Intercept	α_0	−0.1429	19.52	Inputs (Cont.)			
Outputs				$(\ln x_1)(\ln x_5)$	β_{15}	0.0188	1.98
$\ln y_1$ (mathematics score)	α_1	$\underline{0.3757}$		$(\ln x_1)(\ln x_6)$	β_{16}	−0.0152	1.28
$\ln y_2$ (reading score)	α_2	0.6243	41.45	$(\ln x_1)(\ln x_7)$	β_{17}	−0.0166	1.01
$(\ln y_1)^2$	α_{11}	$\underline{1.5089}$		$(\ln x_1)(\ln x_8)$	β_{18}	−0.0857	2.26
$(\ln y_2)^2$	α_{22}	1.5089	17.38	$(\ln x_2)(\ln x_3)$	β_{23}	−0.0601	1.69
$(\ln y_1)(\ln y_2)$	α_{12}	$-\underline{1.5089}$		$(\ln x_2)(\ln x_4)$	β_{24}	0.0616	1.69
Inputs				$(\ln x_2)(\ln x_5)$	β_{25}	−0.0073	0.42
$\ln x_1$ (computers/students)	β_1	−0.0002	0.05	$(\ln x_2)(\ln x_6)$	β_{26}	−0.0159	0.75
$\ln x_2$ (teachers/students)	β_2	−0.0046	0.54	$(\ln x_2)(\ln x_7)$	β_{27}	0.0017	0.06
$\ln x_3$ (mother's level of education)	β_3	−0.0357	3.35	$(\ln x_2)(\ln x_8)$	β_{28}	0.1638	2.42
$\ln x_4$ (father's level of education)	β_4	−0.0214	1.90	$(\ln x_3)(\ln x_4)$	β_{34}	−0.0570	1.96
$\ln x_5$ (cultural activities)	β_5	−0.0414	7.79	$(\ln x_3)(\ln x_5)$	β_{35}	0.0005	0.03
$\ln x_6$ (cultural possessions)	β_6	−0.0288	2.94	$(\ln x_3)(\ln x_6)$	β_{36}	0.0185	0.75
$\ln x_7$ (homework)	β_7	−0.0209	1.77	$(\ln x_3)(\ln x_7)$	β_{37}	−0.0063	0.22
$\ln x_8$ (peer−group)	β_8	−0.1497	7.81	$(\ln x_3)(\ln x_8)$	β_{38}	0.0240	0.30
$(\ln x_1)^2$	β_{11}	0.0124	1.17	$(\ln x_4)(\ln x_5)$	β_{45}	−0.0074	0.40
$(\ln x_2)^2$	β_{22}	0.1620	3.11	$(\ln x_4)(\ln x_6)$	β_{46}	−0.0162	0.70
$(\ln x_3)^2$	β_{33}	0.0930	2.01	$(\ln x_4)(\ln x_7)$	β_{47}	0.0121	0.43
$(\ln x_4)^2$	β_{44}	0.0250	0.59	$(\ln x_4)(\ln x_8)$	β_{48}	0.0879	1.15
$(\ln x_5)^2$	β_{55}	−0.0576	2.72	$(\ln x_5)(\ln x_6)$	β_{56}	0.0066	0.54
$(\ln x_6)^2$	β_{66}	−0.0189	0.70	$(\ln x_5)(\ln x_7)$	β_{57}	0.0288	1.82
$(\ln x_7)^2$	β_{77}	0.0015	0.04	$(\ln x_5)(\ln x_8)$	β_{58}	−0.0293	0.79
$(\ln x_8)^2$	β_{88}	0.0204	0.09	$(\ln x_6)(\ln x_7)$	β_{67}	0.0322	1.86
$(\ln x_1)(\ln x_2)$	β_{12}	−0.0656	3.70	$(\ln x_6)(\ln x_8)$	β_{68}	−0.0322	0.68
$(\ln x_1)(\ln x_3)$	β_{13}	−0.0079	0.43	$(\ln x_7)(\ln x_8)$	β_{78}	−0.0323	2.86
$(\ln x_1)(\ln x_4)$	β_{14}	0.0106	0.58				
Other ML parameters	γ	0.8067	30.84	Expected mean efficiency		0.8869	
	σ^2	0.0286	19.17				

Note: Underlined parameters are calculated by applying imposed homogeneity conditions.

Imposing curvature on the output distance function

Furthermore, the estimation of a parametric output distance function can violate monotonicity for some of the evaluated units. For this reason, it is worth evaluating the results. In educational production theory it is inconsistent that with the same quantities of inputs a student could reduce both scores remaining on the production frontier. The lack of theoretical sense of this result in education and in most of the economics fields leads us to evaluate the estimations obtained at each observation.[3] We proceed following the steps depicted in Section III.

Step 1: This consists of the computation of the predicted efficient output vector on the estimated production frontier, hat denoted $\ln(\hat{y}_{1i})$ and $\ln(\hat{y}_{2i})$, using the outputs transformed in deviations to mean values used in the estimation. In this application, the curvature of the deterministic production frontier is independent of input values because we assume inputs–outputs separability. For this reason and for simplicity in equations, we present the procedure assuming that all DMUs are centred around the mean value (zero in the deviations to mean estimation). Holding this in mind, the outputs in the deterministic production frontier are given as

$$-\ln(\hat{y}_{1i}) = -0.1429 + 0.6243\ln\left(\frac{y_{2i}}{y_{1i}}\right) + 1.5089\frac{1}{2}\left[\ln\left(\frac{y_{2i}}{y_{1i}}\right)\right]^2$$

$$-\ln(\hat{y}_{2i}) = -\ln(\hat{y}_{1i}) - \ln(y_{2i}/y_{1i})$$

[3] The monotonicity on inputs (the output distance function is nondecreasing in x) would imply that additional units of an input will not reduce the output vector. This assumption is closely related with the existence of *input congestion* which sometimes can be found in empirical and theoretical economics. To examine the theory behind the 'uneconomic region' of the production function, see Borts and Mishan (1962). For recent examples in education and health, see Flegg *et al.* (2004) and Ferrier *et al.* (2006), respectively.

For the sake of simplicity and interpretation, we undo the deviations to mean in outputs in order to follow the analysis with the original positive logs of each output working with $\ln(\hat{y}_1)$ and $\ln(\hat{y}_2)$. In Fig. 1, this corresponds to points from F to F'.

Step 2: We calculate $\mathrm{MRT}_{y_2y_1}$ for all DMUs focussing our attention only in those points breaking the monotonicity on outputs ($\mathrm{MRT}_{y_2y_1} > 0$). This stage also implies computation of the partial derivatives of the estimated distance function with respect to each output to know if DMU i fulfils with the assumption of monotonicity on output 1 and on output 2.

$$r_{\hat{y}_{1i}} = \frac{\partial \ln \hat{D}_i}{\partial \ln \hat{y}_{1i}} = 0.3757 + 1.5089 \ln \hat{y}_{1i} - 1.5089 \ln \hat{y}_{2i}$$

$$r_{\hat{y}_{2i}} = \frac{\partial \ln \hat{D}_i}{\partial \ln \hat{y}_{2i}} = 0.6243 + 1.5089 \ln \hat{y}_{2i} - 1.5089 \ln \hat{y}_{1i}$$

$$\mathrm{MRT}_{y_2y_1} = -\frac{\partial \ln y_2}{\partial \ln y_1} = -\frac{r_{\hat{y}_1}}{r_{\hat{y}_2}}$$

As we can see in Table 3, there are a number of pupils (204 cases; i.e. 8.33% of total) where monotonicity on outputs does not hold and the slope of the distance function becomes positive. This is probably due to the fact that, in real life, with very few exceptions, there are no pupils with outstanding results in reading (mathematics) and extremely bad results in mathematics (reading) and as a consequence the production frontier curves back in the wrong direction. If we fail to take this fact into account, we can underestimate inefficiency levels for those students projected at the stretches of the production frontier, which are breaking the monotonicity assumption in outputs.

Step 3: We compute the output projection vector corresponding to the strict frontier, tilde denoted $\ln \tilde{y}_{1A}$, that is points A'' and B'' according to Fig. 1. Once we know a DMU A breaks monotonicity on an output, our aim is to seek the maximum value $\ln \tilde{y}_{2i}^{\max}$ of the other output (points C and D in Fig. 1) in the distance frontier providing to all DMU the A inputs endowment. The maximum value found is assigned to DMU A projecting the other output, holding the exogenous output ratio of A constant.

$$-\ln(\hat{y}_{1i}) = -0.1429 + 0.6243 \ln\left(\frac{y_{2i}}{y_{1i}}\right) + 1.5089 \frac{1}{2}\left[\ln\left(\frac{y_{2i}}{y_{1i}}\right)\right]^2$$

$$-\ln(\hat{y}_{2i}) = -\ln(\hat{y}_{1i}) - \ln(y_{2i}/y_{1i}) \rightarrow \ln \tilde{y}_{2i}^{\max}$$

$$\ln \tilde{y}_{1A} = \ln \tilde{y}_{2i}^{\max} - \ln\left(\frac{y_{2A}}{y_{1A}}\right)$$

Note that $\ln \tilde{y}_{1A}$ will be always greater than $\ln(\hat{y}_{1A})$.

Step 4: New efficiency scores for each DMU are computed to adding up to the estimated distance $\ln \hat{D}_{\mathrm{O}i}(x, y)$ the *extra distance* term $-\ln \hat{D}_{\mathrm{O}i}^{\mathrm{extra}}(x, y)$, which separates the computed deterministic production frontier output vector, $(\ln(\hat{y}_{1i}), \ln(\hat{y}_{2i}))$, from the strict production frontier output vector, $(\ln(\tilde{y}_{Mi}), \ln(\tilde{y}_{ni}^{\max}))$. The corresponding extra distance for DMUs A and B are therefore graphically measured in Fig. 1 by the Euclidean distances between OA' and OA'' and OB' and OB'', respectively. For DMU A, we obtain the following equation:

$$\ln D_{\mathrm{O}A}^{\mathrm{extra}}(x, y) = d\left[(\ln \hat{y}_{1A}, \ln \hat{y}_{2A}); \left(\ln \tilde{y}_{1A}, \ln \tilde{y}_{2A}^{\max}\right)\right] = \sqrt{(\ln \hat{y}_{1A} - \ln \tilde{y}_{1A})^2 + \left(\ln \hat{y}_{2A} - \ln \tilde{y}_{2A}^{\max}\right)^2}$$

In our example for DMUs breaking monotonicity the new inefficiency values slightly decrease in mean from 0.877 to 0.863.

Step 5: As described in Section III, the radial expansion of a DMU to the strict production function originates a target that present an output slack. DMU A could produce more on one output holding constant the other. The new target $\ln \breve{y}_{1i}$ for DMU A to hold monotonicity on output 1 will be as follows:

$$r_{1A} = \frac{\partial \ln \hat{D}}{\partial \ln \hat{y}_{1A}} = 0.3757 - 1.5089 \ln \tilde{y}_{2i}^{\max} + 1.5089 \ln \breve{y}_{1A}$$

where rearranging terms

$$\ln \breve{y}_{1A} = \frac{-0.3757 + 1.5089 \ln \tilde{y}_{2i}^{\max}}{1.5089}$$

We measure output slack for output 1 in absolute values as follows:

$$\mathrm{OS}_{1A} = \breve{y}_{1A} - \tilde{y}_{1A}^{\max}$$

Table 3. Descriptive statistics for DMUs breaking the monotonicity on outputs

Distance slack	N
Rupture in mathematics ($r_1 < 0$)	194
Rupture in reading ($r_2 < 0$)	10
Total	204

For DMU B and output 2, we have

$$r_{2B} = \frac{\partial \ln \hat{D}}{\partial \ln \hat{y}_{2B}} = 0.6243 - 1.5089 \ln \tilde{y}_{1i}^{\max} + 1.5089 \ln \breve{y}_{2B}$$

where rearranging terms

$$\ln \breve{y}_{2B} = \frac{-0.6243 + 1.5089 \ln \tilde{y}_{1i}^{\max}}{1.5089}$$

As we did with output 1, we also measure output slack for output 2 in absolute values as follows:

$$\mathrm{OS}_{2B} = \breve{y}_{2B} - \tilde{y}_{2B}^{\max}$$

Table 4 summarizes the changes in inefficiency values for DMUs breaking monotonicity. In this educational example, extra distances are moderately low but for the highest values the imposition of monotonicity on outputs shifts some DMUs to a more realistic radial efficiency. With respect to output slacks, we observe important potential gains for Spanish students, especially in mathematics. Average output slacks in mathematics and reading account for 0.64 and 0.52 SDs of test scores of all students respectively. Nevertheless, for some pupils their projection to the production frontier without output slacks supposes to increase mathematics and reading scores up to 2.84 and 1.30 SDs, respectively.

Finally, we focus our attention on the relationship between school ownership and efficiency. Table 5 reports student level average efficiency before and after imposing monotonicity on the output distance function as well as mean output slacks. We run different statistical tests to look for significant performance differences regarding school ownership. First, we test through a one-way Analysis of Variance (ANOVA) if there are significant differences in technical efficiency by school ownership before and after the imposition of monotonicity. Second, we employ a t-test for paired samples to contrast whether or not mean technical efficiency remains constant before and after the imposition of monotonicity. Finally, in order to test mean output slacks differences by school ownership, and taking into account that the output slack variables are not normally distributed, we employ a Kruskal–Wallis test.

What can we learn from this comparison? On the one hand, once school inputs, student background

Table 4. Descriptive statistics for estimated new efficiencies in DMUs where monotonicity is imposed

Distance slack	N	Mean	SD	Minimum	Maximum
Extra distance $\ln \hat{D}_{\mathrm{O}i}^{\mathrm{extra}}(x, y)$	204	0.018	0.041	1.92E−07	0.288
Output slacks ($\mathrm{OS}_{\mathrm{maths}}$)	194	53.19	51.81	0.95	235.72
Output slacks ($\mathrm{OS}_{\mathrm{reading}}$)	10	39.08	31.46	8.46	96.49

Table 5. Efficiency, output slacks and school ownership

School type	N	Mean efficieny[a] before imposing monotonicity	Mean efficieny[b] after imposing monotonicity	Mean mathematic slack[d] after imposing monotonicity	Mean reading slack[e] after imposing monotonicity
Private, government independent	237	0.8921	0.8920	2.2713	0.1968
Private, government dependent	829	0.8873	0.8866	3.7841	0.2093
Government	1383	0.8857	0.8846	4.8032	0.1234
All	2449	0.8869[c]	0.8860	4.2132	0.1596

Notes: [a]Mean differences among school types are not statistically significant, at 95% level, with F-test = 1.088 (p-value = 0.337). Variances are distributed homogenously, at 95% level, with p-value associated to the Levene's test = 0.375.
[b]Mean differences among school types are not statistically significant, at 95% level, with F-test = 1.485 (p-value = 0.227). Variances are distributed homogenously, at 95% level, with p-value associated to the Levene's test = 0.188.
[c]There is a significant difference (paired t-test = 4.265) between the efficiency scores before and after the imposition of monotonicity.
[d]The p-value associated with the Kruskal–Wallis test equals 0.082 revealing statistically significant mean differences for the mathematic slack by school ownership only at 90% level.
[e]The p-value associated with the Kruskal–Wallis test equals 0.429 revealing no statistically significant mean differences for the reading slack by school ownership.

and peer group, are taken into account, there are no mean technical efficiency differences by school ownership neither before nor after applying our approach to impose monotonicity (this result is consistent with Perelman and Santin (2008)). On the other hand, according to the mathematic slack, we observe that students attending public schools have a slight, but significant at 90% level, greater potential of improvement in mathematics to reach the production frontier according to their inputs endowment. In addition to this, differences in reading slack are not significant by school ownership. Finally, it is interesting to remark that the paired *t*-test points out a statistically significant difference between the technical efficiency scores estimated before and after the imposition of monotonicity.

From our point of view, without all these proposed corrections for those DMUs breaking off the monotonicity on outputs assumption, the first estimated radial technical efficiency results obtained with the initial stochastic frontier analysis could be misleading in most empirical applications.

VI. Concluding Remarks

The violation of the monotonicity on outputs assumption is not admissible from the point of view of production theory. In order to avoid this economic inconsistency in empirical parametric frontier estimation studies, in this article we provide a methodology, based on the computation of the estimated output distance function derivatives, to easily impose monotonicity on outputs. The final target of this approach is to enhance efficiency estimations with parametric distance functions as well as to calculate the measurement of parametric *output slacks* proposed in the article.

The example in education reveals that around a nonnegligible 8.33% of DMUs break monotonicity on outputs especially in mathematics (pupils with high results in reading with respect to a relative low performance in mathematics), representing 7.92% of total. Although the extra-distance measurements obtained in this application are of a modest importance, we find that nonnegligible output slacks in both outputs. This result points out to high potential educational gains in Spain especially in the mathematics learning process. Furthermore, we do not find technical efficiency mean differences by school ownership. However, we observe that students attending public schools have a slight, but significant at 90% level, greater mathematics slack pointing out to a major potential of improvement in mathematics.

We do think that these new proposed efficiency corrections may concern practitioners in future empirical applications to obtain unbiased interpretable efficiency results.

Acknowledgements

The authors are grateful to Knox Lovell, Kris Kerstens and Luis Orea for helpful discussions and encouragement. A preliminary version of this article was published in the Working Papers Collection of Fundación de las Cajas de Ahorros – FUNCAS – (Spanish Saving Banks Foundation) as the Working Paper 459.

References

Afonso, A. and St. Aubyn, M. (2006) Cross-country efficiency of secondary education provision: a semi-parametric analysis with non-discretionary inputs, *Economic Modelling*, **23**, 476–91.

Aigner, D. J., Lovell, C. A. K. and Schmidt, P. (1977) Formulation and estimation of stochastic production function models, *Journal of Econometrics*, **6**, 21–37.

Banker, R. D., Janakiraman, S. and Natarajan, R. (2004) Analysis of trends in technical and allocative efficiency: an application to Texas public school districts, *European Journal of Operational Research*, **154**, 477–91.

Borts, G. H. and Mishan, E. J. (1962) Exploring the 'uneconomic region' of the production function, *The Review of Economic Studies*, **29**, 300–12.

Coelli, T. and Perelman, S. (1999) A comparison of parametric and non-parametric distance functions: with application to European railways, *European Journal of Operational Research*, **117**, 326–39.

Coelli, T. and Perelman, S. (2000) Technical efficiency of European railways: a distance function approach, *Applied Economics*, **32**, 1967–76.

Coelli, T., Prasada Rao, D. S., O'Donnell, C. J. and Battese, G. E. (2005) *An Introduction to Efficiency and Productivity Analysis*, Springer, New York.

Cordero-Ferrera, J. M., Pedraja-Chaparro, F. and Salinas-Jiménez, F. J. (2008) Measuring efficiency in education: an analysis of different approaches for incorporating non-discretionary inputs, *Applied Economics*, **40**, 1323–39.

Färe, R., Grosskopf, S. and Lovell, C. A. K. (1984) *The Measurement of Efficiency in Production*, Kluwer-Nijhoff Publishing, Boston.

Färe, R., Grosskopf, S., Lovell, C. A. K. and Yaisawarng, S. (1993) Derivation of shadow prices for undesirable outputs: a distance function approach, *Review of Economics and Statistics*, **75**, 374–80.

Färe, R. and Primont, D. (1995) *Multi-Output Production and Duality: Theory and Applications*, Kluwer Academic Publishers, Boston.

Ferrier, G. D., Rosko, M. D. and Valdmanis, V. G. (2006) Analysis of uncompensated hospital care using a DEA model of output congestion, *Health Care Management Science*, **9**, 181–8.

Flegg, A. T., Allen, D. O., Field, K. and Thurlow, T. W. (2004) Measuring the efficiency of British universities: a multi-period data envelopment analysis, *Education Economics*, **12**, 231–49.

Hanushek, E. A., Rivkin, S. G. and Taylor, L. L. (1996) Aggregation and the estimated effects of school resources, *The Review of Economics and Statistics*, **78**, 611–27.

Lovell, C. A. K., Richardson, S., Travers, P. and Wood, L. L. (1994) Resources and functionings: a new view of inequality in Australia, in *Models and Measurement of Welfare and Inequality* (Ed.) W. Eichhorn, Springer Verlag, Berlin, pp. 787–807.

O'Donnell, C. J. and Coelli, T. J. (2005) A Bayesian approach to imposing curvature on distance functions, *Journal of Econometrics*, **126**, 493–523.

OECD (1999) *Classifying Educational Programmes: Manual for ISCED-97 Implementation in OECD Countries*, Organisation for Economic Co-operation and Development, Paris.

OECD (2001) *Manual for the PISA 2000 Database*, Organisation for Economic Co-operation and Development, Paris.

OECD (2002) *PISA 2000 Technical Report*, Organisation for Economic Co-operation and Development, Paris.

Perelman, S. and Santin, D. (2008) Measuring educational efficiency at student level with parametric stochastic distance functions: an application to Spanish PISA results, *Education Economics*. DOI: 10.1080/09645290802470475.

Rosenzweig, M. and Wolpin, K. I. (1994) Are there increasing returns to the intergenerational production of human capital? Maternal schooling and child intellectual achievement, *Journal of Human Resources*, **29**, 670–93.

Santín, D. (2006) La medición de la eficiencia de las escuelas: una revisión crítica, *Hacienda Pública Española/Revista de Economía Pública*, **177**, 57–83.

Shephard, R. W. (1970) *Theory of Cost and Production Functions*, Princeton University Press, Princeton.