Available online at http://ijdea.srbiau.ac.ir

Int. J. Data Envelopment Analysis (ISSN 2345-458X)

Vol.5, No.2, Year 2017 Article ID IJDEA-00422, 12 pages
Research Article

International Journal of Data Envelopment Analysis

Science and Research Branch (IAU)

Combining Time DEA Scores Using a Dynamic Panel Data Model

Geraldo da Silva e Souza[a], Eliane Gonçalves Gomes[b*]

[(a,b)] *Embrapa, Secretaria de Inteligência e Relações Estratégicas – SIRE*
Parque Estação Biológica, Av. W3 Norte final, 70770-901, Brasília, DF, Brazil

Received September, 22, 2016, Accepted March, 25, 2017

Abstract
We define a combined DEA score to evaluate efficiency in agricultural research. The production model we propose considers efficiency measurements under variable returns to scale for each year in the period 2012–2017. We postulate a first-order autoregressive process in the presence of covariates, to explain efficiency. Powers of the autocorrelation coefficient estimated assuming a dynamic panel specification, are used as weights to determine a combined efficiency score. A higher weight is given to recent efficiency measurements. We use a fractional regression model to investigate the statistical significance of covariates on the combined score further.

Keywords: Time series, DEA, Fractional regression, Agricultural research.

*.Corresponding author Email: eliane.gomes@embrapa.br

1. Introduction

Since 1996, the Brazilian Agriculture Research Corporation (Embrapa) has been monitoring the production performance of its research centers by using a Data envelopment analysis (DEA) model [1–5]. Recently, the research centers' evaluation system has been reviewed, and the efficiency component has gained renewed importance in the whole performance evaluation process. New goals contemplate performance for a time interlude and must accommodate different efficiency components computed within each year.

Current DEA literature includes a plethora of models dealing with the performance measure in a time-series context. The combination with cross-sections is also possible. Malmquist DEA [6] is an instance of DEA analysis for panel data. Lynde and Richmond [7] consider a model for the study of time-series data on inputs and outputs, allowing the inclusion of shifting technology into the DEA framework. Dynamic DEA models and dynamic network DEA models are discussed in Tone [8]. It is not common, however, to model the DEA responses as evolving, satisfying a statistical time-series model where the dependent variable follows a stochastic process. We intend to explore this feature of time-series DEA.

We present a method to combine a series of DEA measurements computed in each point in time into a single score, reflecting average efficiency in the period. The method assigns a weight to the efficiency in each year. The weight sequence decreases with time, attributing more importance to recent years by assuming that the efficiency responses, for the panel of decision-making units (DMUs), follow a first-order autoregressive process. The autocorrelation coefficient is estimated by a generalized method of moments (GMM) method [9, 10], assuming a dynamic panel. To the best of our knowledge, the application is new in the DEA context.

Finally, we investigate the effect of contextual classification variables on the efficiency score, with the objective of relating best productions practices with control variables.

2. Production Data

Embrapa considers as outputs of its production system, 50 real-valued production indicators of output. Each of its 41 research centers provides a 50-dimensional response vector, corresponding to a three-dimensional input vector defined by expenses in labor, capital and other inputs. The output is a single univariate response. These data (output and inputs) are the production dataset, which comprises a 1×41 output matrix Y, and a 3×41 input matrix X.

Reducing the dimension of the output vector to a single quantity demands transformations, allowing aggregation and a proper weighting system. One approach to realizing this goal is via multivariate analysis and rank transformation or eliminating scale, transforming each indicator in a dimensionless index. Here, we achieve this transformation by measuring each research result, whether input or output, in per capita mean units. Division by a common constant is not likely to affect the validity of DEA production assumptions, as reported in Olesen et al. [11]. Instead of using a common set of weights, we allow each research center to have its particular weight structure. That is, we leave the process of weights determination to each research center, under the supervision of Embrapa's managers. The weighting system determination follows a hierarchical structure. The 50 transformed indices are split into eight groups, and weights are assigned to each variable within each group and for each group. The final individual weights are multiplicative.

Let $0 \le a_{ji}^{o}, \sum_{i=1}^{k_j} a_{ji}^{o} = 1$ be the weights assigned by unit o within group j for indicator i. There are eight groups, 41 units (research centers), and k_j indices for group j. Let $0 \le w_j^{o}, \sum_{j=1}^{8} w_j^{o} = 1$ be the set of group's weights for unit o. For period t, let y_{ji}^{ot} denote the value of indicator i for category j for unit o. Let $\bar{y}_{ji}^{t}$ be the mean for period t of attribute i in group j. The output score for unit o for period t is given by:

$$y^{ot} = \sum_{j=1}^{8} w_j^{o} \left(\sum_{i=1}^{k_j} a_{ji}^{o} \left(\frac{y_{ji}^{ot}}{\bar{y}_{ji}^{t}} \right) \right) \qquad (1)$$

Expenses on labor, capital and other inputs are normalized by the period means and consider as indices $x_{\upsilon}^{ot} = \frac{d_{\upsilon}^{ot}}{\bar{d}_{\upsilon}^{t}}$, where d_{υ}^{ot} denotes expenses on item υ (labor, capital and other input expenses) by unit o in period t, and $\bar{d}_{\upsilon}^{t}$ signifies the average item υ expenses in period t. Table 1 shows the input and output data matrices for the year 2012. Type is a categorization of the research centers, based on their research focus: research on agricultural products (Product), on agricultural specific themes (Thematic), and on issues related to environmental and ecological aspects (Ecological), respectively.

Table 1 Production data for 2012

Unit	Type	$X1$	$X2$	$X3$	Y
DMU_01	Thematic	1.2596	1.6022	1.8564	1.2262
DMU_02	Product	0.9640	0.6145	0.6243	1.8892
DMU_03	Thematic	0.9248	0.8927	1.1162	0.6032
DMU_04	Thematic	1.1331	1.1152	0.3523	0.6292
DMU_05	Product	0.7788	0.7047	0.8194	0.4341
DMU_06	Product	0.9784	1.0158	0.2970	0.2680
DMU_07	Thematic	1.1117	1.1892	1.1169	0.6684
DMU_08	Product	0.9212	0.9100	0.7916	0.3073
DMU_09	Thematic	1.1944	1.2588	1.6828	3.5340
DMU_10	Product	1.1585	1.0735	1.1758	0.8485
DMU_11	Product	0.8747	0.9734	1.3100	0.5487
DMU_12	Product	0.8772	0.5409	1.0461	1.2287
DMU_13	Product	0.7949	0.9082	1.1966	0.8758
DMU_14	Thematic	1.1740	1.1902	1.0041	1.4603
DMU_15	Thematic	1.1367	0.8285	1.5117	1.0967
DMU_16	Product	1.0151	1.2114	0.6319	1.3719
DMU_17	Product	0.9043	0.7265	0.9144	0.6894
DMU_18	Thematic	1.1939	0.7222	0.8624	0.8810
DMU_19	Product	0.9414	1.0763	1.5327	1.3961
DMU_20	Product	0.8337	0.9570	1.3607	1.0567
DMU_21	Product	0.8756	0.7547	0.9975	0.6188
DMU_22	Thematic	1.3235	0.9921	1.0863	6.9858

DMU_23	Product	0.9538	1.2033	1.4893	1.1431
DMU_24	Ecological	0.9543	0.7536	0.7585	0.3006
DMU_25	Ecological	0.8879	0.7614	1.0881	0.2147
DMU_26	Ecological	1.0321	1.2781	0.0000	0.0642
DMU_27	Ecological	0.9148	1.0708	1.1193	1.1552
DMU_28	Ecological	1.1476	1.5037	1.1289	0.3723
DMU_29	Ecological	1.1508	1.1897	0.6865	0.4400
DMU_30	Ecological	0.9241	0.7893	0.5842	0.3452
DMU_31	Ecological	1.1111	1.2159	0.8454	0.4286
DMU_32	Ecological	0.8099	0.8341	0.4112	0.3961
DMU_33	Ecological	0.9490	2.4303	0.3894	0.6524
DMU_34	Ecological	0.8915	0.9107	0.8409	0.9681
DMU_35	Ecological	1.1379	0.8255	0.8870	2.3713
DMU_36	Ecological	1.0619	0.8323	0.7585	1.1690
DMU_37	Ecological	0.7845	0.9097	0.7904	0.4018
DMU_38	Ecological	1.0048	0.7678	0.5118	1.4188
DMU_39	Product	0.9722	0.8572	1.2233	1.3563
DMU_40	Product	0.8448	0.7719	2.2280	0.9666
DMU_41	Thematic	1.0975	0.8366	1.9721	0.9905

3. Methodology

The response variable in our analysis is the classical input-oriented DEA measure of technical efficiency, computed under the assumption of variable returns to scale (DEA-VRS) [12]. If $Y^t = (y^{1t}, y^{2t}, \ldots, y^{41t})$ is the output matrix, and $X^t = \begin{pmatrix} d_1^{1t}, d_1^{2t}, \ldots, d_1^{41t} \\ d_2^{1t}, d_2^{2t}, \ldots, d_2^{41t} \\ d_3^{1t}, d_3^{2t}, \ldots, d_3^{41t} \end{pmatrix}$ is the input matrix, for period *t*, the DEA-VRS technical efficiency $\hat{\theta}^{ot}$ for unit *o* is the solution of the following linear programming problem:

$$
\begin{aligned}
& Min_{\lambda,\zeta}\ \zeta \\
& Y^t \lambda \geq y^{ot} \\
& X^t \lambda \leq \zeta x^{ot}, (x^{ot})' = (d_1^{ot}, d_2^{ot}, d_3^{ot}) \quad (2) \\
& \lambda' e = 1,\ \lambda' = (\lambda_1, \ldots, \lambda_{41}), \\
& e' = (1, \ldots, 1),\ \lambda_i \geq 0
\end{aligned}
$$

The DEA estimates can be shown to be weakly consistent within years [13]. Under a deterministic frontier assumption in the context of univariate outputs, the DEA estimate is strongly consistent and is a nonparametric maximum likelihood estimate [14. 15]. Assuming independent production decisions under the same production function, these considerations justify the use of DEA responses in regression analysis when covariates are not endogenous or separable [16].

Through time, we assume that the DEA measurements follow the dynamic panel data":

$$
\hat{\theta}^{ot} = \rho \hat{\theta}^{o(t-1)} + z'_{ot}\beta + u_o + \varepsilon_{ot}, \quad o = 1, \ldots, 41,\ \ t = 2012, \ldots, 2017 \qquad (3)
$$

Here $0 < \rho < 1$ is the autoregressive parameter, z_{ot} is a *q*-vector of strictly exogenous variables, β is the corresponding parameter vector, u_o are the random panel level effects (research centers), and ε_{ot} are iid (independent and identically distributed) errors over the whole sample with constant variance. Both u_o and ε_{ot} are assumed to be independent for each *o* over all *t*. Therefore, the lagged dependent variables are correlated with the unobserved panel level effects, making standard estimation inconsistent [17]. With many panels and few periods, we follow the GMM approach suggested by Arellano and Bover [9] and Blundell and Bond [10]. The model accommodates less restrictive assumptions, regarding the covariates as endogeneity. A key assumption regarding the residual evolution through time is the non-existence of second-order autocorrelation in the differenced series, which can be tested following Arellano and Bond [18]. Exploiting the autoregressive structure, we propose the final efficiency estimate as the following, where ρ^h represents the correlation between efficiencies distant *h* periods apart.:

$$eff_o = \left(\frac{\hat{\theta}^{o(2017)} + \rho\hat{\theta}^{o(2016)} + \rho^2\hat{\theta}^{o(2015)} + \rho^3\hat{\theta}^{o(2014)} + \rho^4\hat{\theta}^{o(2013)} + \rho^5\hat{\theta}^{o(2012)}}{1+\rho+\rho^2+\rho^3+\rho^4+\rho^5}\right), \quad o = 1,\ldots,41 \tag{4}$$

Higher-order processes can be considered and tested in the framework of dynamic panels. The correlation structure will be less trivial.

In order to assess the significance of factor variables on the response eff_o, we use a standard fractional regression model [19]. Let an observed response $\hat{\theta} = eff_o$ with values in (0,1] be dependent on a vector of covariates *w*. It is assumed that $E(\hat{\theta} \mid w) = G(w\delta)$, where *G*(.) is typically a probability distribution function. The model is well-defined, even when $\hat{\theta} = eff_o$ puts positive probability mass at one. The unknown parameter δ is then estimated by quasi-maximum likelihood (QML), maximizing $\sum_{i=1}^{n}\left(\begin{array}{l}\hat{\theta}_i \log\left(G\left(w_i\delta\right)\right) + \\ (1-\hat{\theta}_i)\log\left(1-G\left(w_i\delta\right)\right)\end{array}\right)$ [19].

Under the correct specification of the mean function $\sqrt{n}\left(\hat{\delta}-\delta\right) \xrightarrow{d} N(0,V)$. V is estimated as below in (5). The QML estimator is efficient within the class of estimators containing all linear, exponential family-based QML and weighted nonlinear least squares estimators:

$$\begin{aligned}
&\hat{V} = \left(\hat{A}\right)^{-1}\hat{B}\hat{A}, \\
&\hat{A} = 1/n\sum_{i=1}^{n}\left(\hat{g}_i^2 / \hat{G}_i(1-\hat{G}_i)\right)w_i'w_i \\
&\hat{B} = 1/n\sum_{i=1}^{n}\left(\hat{u}_i^2\hat{g}_i^2 / (\hat{G}_i(1-\hat{G}_i))^2\right)w_i'w_i \\
&\hat{G}_i = G(w_i'\hat{\delta}), \quad \hat{g}_i = G'(w_i'\hat{\delta}), \quad \hat{u}_i = \hat{\theta}_i - \hat{G}_i
\end{aligned} \tag{5}$$

These formulas appear in Ramalho et al. [20]. The calculations may be performed with the use of Stata 15 [21], where the method is implemented.

4. Statistical Results

Table 2 shows DEA-VRS efficiency measurements for each year and the combined estimate (column 'Combined efficiency'). The panel efficiency graphs are shown in Figure 1. One can see that units 2, 5, 13 and 32 are efficient through the period. Units 10, 18 and 34 show an increasing trend. Units 4, 21, 26 and 40 show a decreasing trend. For other units we do not identify a clear trend. The apparent volatility through time calls for an overall measure to capture average performance.

The Arellano–Bover/Blundell–Bond estimation is computed using Stata v.14 software [17]. Additional covariates included in the model specification are a time dummy variable for 2016, two dummy variables for type (base is

Thematic) and two dummy variables indicating size (base is Large). Research centers were classified into three groups of size, using cluster analysis (Ward's method) applied to the evolution of total input expenses.

Table 2: Efficiency scores, age and size (Size: 3 = large, 2 = medium, 1 = small; Age: 0 > 10 years, 1 ≤ 10 years)

Unit	Size	Age	Technical efficiency						Combined efficiency
			2012	2013	2014	2015	2016	2017	
DMU_1	3	0	0.6552	0.7492	0.6606	0.6429	0.6295	0.6202	0.6414
DMU_2	1	0	1.0000	1.0000	1.0000	0.9834	0.9695	1.0000	0.9899
DMU_3	1	0	0.8515	0.8654	0.8532	0.8647	0.8287	0.8240	0.8392
DMU_4	1	1	0.8563	1.0000	0.8360	0.6975	0.6909	0.6826	0.7345
DMU_5	3	0	1.0000	1.0000	1.0000	0.9957	1.0000	1.0000	0.9993
DMU_6	1	1	0.9563	0.7648	0.6852	0.8587	0.7546	0.7851	0.7855
DMU_7	2	0	0.7199	0.7077	0.7084	0.8160	0.7141	0.7544	0.7446
DMU_8	1	0	0.8571	0.8780	0.8889	0.8624	0.8428	0.8433	0.8543
DMU_9	1	0	0.8581	0.8894	0.8423	0.8913	0.6651	0.6953	0.7571
DMU_10	2	0	0.7027	0.7095	0.7769	0.9732	0.9144	0.9552	0.8994
DMU_11	2	0	0.8951	0.9510	0.9838	0.9898	0.8902	0.9815	0.9545
DMU_12	3	0	1.0000	1.0000	0.9754	0.9236	0.8597	1.0000	0.9508
DMU_13	2	0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
DMU_14	1	0	0.7442	0.7220	0.6779	0.7135	0.6507	0.6111	0.6588
DMU_15	2	0	0.7497	0.7499	0.7239	0.7179	0.7876	0.7726	0.7594
DMU_16	2	0	0.8696	0.8106	0.8373	0.8641	0.7800	0.8083	0.8166
DMU_17	3	0	0.9107	0.8902	0.9219	0.9737	0.8877	0.8955	0.9095
DMU_18	1	0	0.8290	0.8053	0.8308	0.8246	0.9431	0.9670	0.9053
DMU_19	2	0	0.8923	0.8729	0.8823	0.8996	0.7813	0.8277	0.8402
DMU_20	3	0	0.9723	0.9538	0.9857	0.9298	0.9257	0.9467	0.9448
DMU_21	2	0	0.9116	0.9289	0.9353	0.9167	0.9085	0.9534	0.9308
DMU_22	1	0	1.0000	1.0000	1.0000	1.0000	0.9067	0.9172	0.9468
DMU_23	2	0	0.8577	0.8364	0.8435	0.8648	0.8234	0.8356	0.8393
DMU_24	2	0	0.9040	0.9288	0.8697	0.9637	0.8562	0.8835	0.8925
DMU_25	3	0	0.8936	0.9031	0.9498	0.9071	0.8897	0.9318	0.9156
DMU_26	1	1	1.0000	1.0000	1.0000	0.8256	0.7114	0.7934	0.8255
DMU_27	3	0	0.9065	0.8715	0.8699	0.8785	0.8344	0.8444	0.8551
DMU_28	2	0	0.6819	0.7058	0.6980	0.6945	0.6865	0.7195	0.7023
DMU_29	1	0	0.7019	0.7767	0.7052	0.7340	0.6901	0.6699	0.6982
DMU_30	1	0	0.9327	0.9062	0.8661	0.8579	0.8462	0.8462	0.8586
DMU_31	1	0	0.7167	0.8187	0.9510	0.8658	0.9122	0.8619	0.8744
DMU_32	2	0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
DMU_33	1	1	0.9273	0.6522	0.7361	0.8490	0.7612	0.7810	0.7801
DMU_34	1	0	0.9290	0.9000	0.9063	0.8753	1.0000	1.0000	0.9592

DMU_35	1	0	0.8565	0.9069	0.9134	1.0000	0.7691	0.8532	0.8670
DMU_36	2	0	0.8287	0.8046	0.7778	0.8016	0.7977	0.9846	0.8668
DMU_37	3	0	0.9959	0.9869	0.9264	0.9090	0.9245	1.0000	0.9578
DMU_38	3	0	0.9737	1.0000	1.0000	0.9992	0.9259	1.0000	0.9805
DMU_39	1	0	0.8790	0.9583	1.0000	0.9970	0.9155	0.8468	0.9139
DMU_40	1	0	0.9713	0.9147	0.9268	0.9389	0.8243	0.8532	0.8780
DMU_41	2	0	0.7650	0.9288	1.0000	1.0000	0.7149	0.7609	0.8264

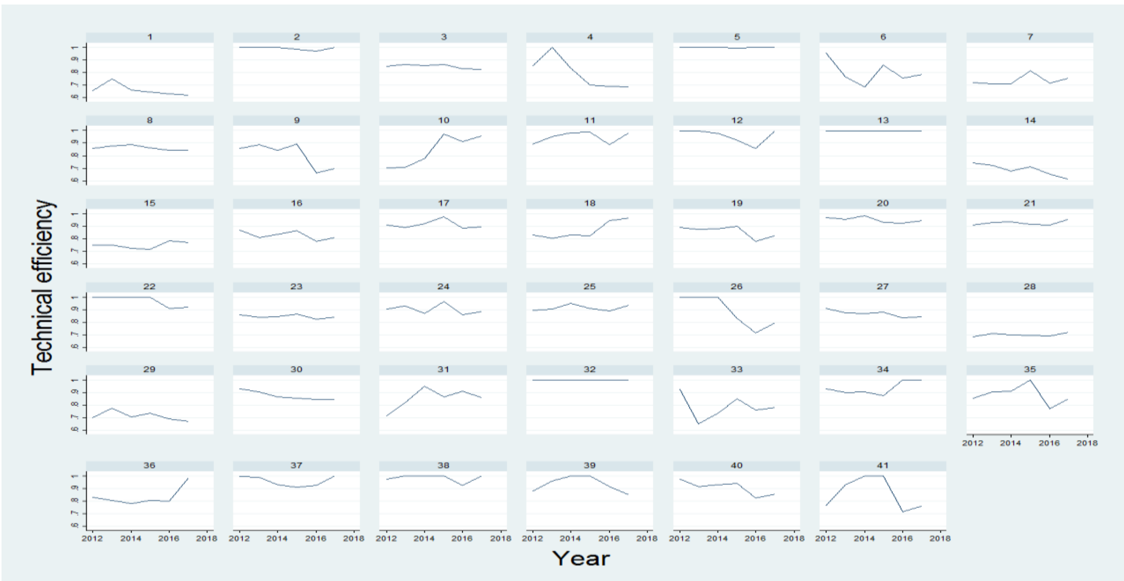


Fig. 1

We now analyze the significance of type, size and time effects on the DEA responses. Only a single time dummy was included (2016), to account for a reduction in the overall efficiency level observed in 2016. This effect can be detected by computing the yearly averages (Figure 1). We see that type and size are nonsignificant effects. The results are summarized in Table 3. Exclusion of type and size leads to the final estimates shown in Table 4. The panel data parameters of Table 4 were used to compute the combined efficiency scores shown in Table 1. We see that the condition for stationarity holds since the autoregressive parameter satisfies $0 < \hat{\rho} < 1$. The Arellano–Bond autocorrelation test has a *p*-value of 90.1%, and there is no evidence of a second-order autocorrelation, which would invalidate the model specification.

Table 3: Preliminary dynamic panel estimation

	Coefficient	Standard Error	z	P>|z|	[95% Confidence	Interval]
L1	0.4806	0.1530	3.14	0.002	0.1808	0.7804
Type_Ecological	-1.0494	8.5045	-0.12	0.902	-17.7180	15.6192
Type_Product	0.7175	4.9223	0.15	0.884	-8.9301	10.3650
Size_Small	-1.9984	11.4951	-0.17	0.862	-24.5284	20.5316
Size_Medium	-1.7458	11.3044	-0.15	0.877	-23.9021	20.4105
Time_2016	-0.0504	0.0128	-3.94	0.000	-0.0754	-0.0253
Constant	2.0365	10.0591	0.20	0.840	-17.6790	21.7520

Table 4: Final dynamic panel estimation

	Coefficient	Standard Error	z	P>|z|	[95% Confidence	Interval]
L1	0.6643	0.0883	7.53	0.000	0.4913	0.8373
Time_2016	-0.0558	0.0131	-4.25	0.000	-0.0815	-0.0301
Constant	0.3003	0.0759	3.96	0.000	0.1516	0.4490

Arellano–Bond test for zero autocorrelation in first-differenced errors

Order	z	Prob > z
1	-3.2497	0.0012
2	0.12404	0.9013

H0: no autocorrelation

A joint analysis of type, size and age (a dummy variable indicating whether the research center has been in operation for less than 10 years, age = 1) is then performed by applying fractional regression, assuming the probit or the logistic response to explain the combined efficiency score. The model below assumes the quasi-likelihood function, where $\Phi(.)$ is the standard normal or the logistic distribution function; $size_1, size_2$ are dummies for small and medium research centers, and age is the indicator of whether a research center is aged more than 10 years. A further classification of type was considered in the analysis. We do not detect the importance of this effect in the panel regression. The corresponding dummy variables for ecological and product are $type_1, type_2$, respectively. In the following equation, the betas (β) are parameters to be estimated:

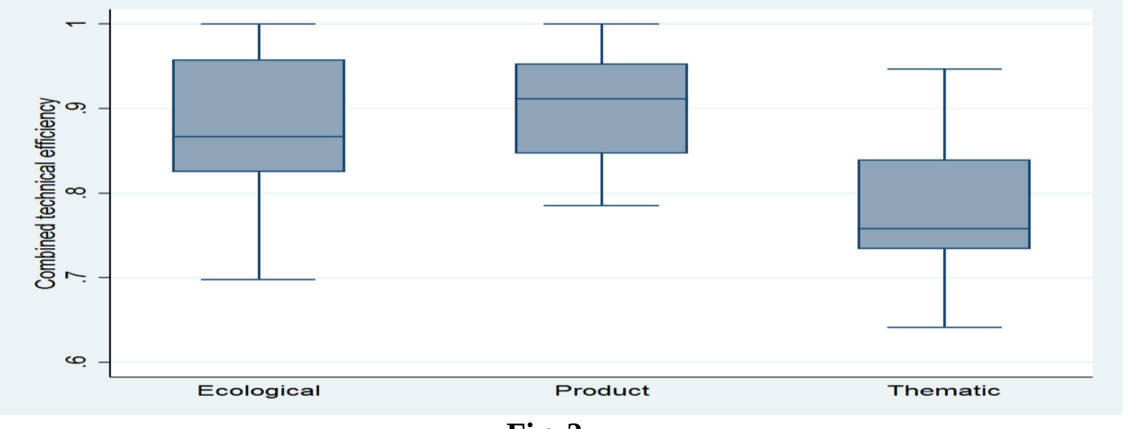
$$\ln L = \sum_{j=1}^{41} escore_j \ln\left(\Phi\begin{pmatrix}\beta_0+\beta_1 size_1+\beta_2 size_2\\ +\beta_3 age+\beta_4 type_1+\beta_5 type_2\end{pmatrix}\right)+ (1-escore_j)\ln\begin{pmatrix}1-\\ \Phi\begin{pmatrix}\beta_0+\beta_1 size_1+\beta_2 size_2+\\ \beta_3 age+\beta_4 type_1+\beta_5 type_2\end{pmatrix}\end{pmatrix} \qquad (6)$$

Table 5 shows the results of the analysis, assuming the logistic distribution. Results with the normal distribution function are similar. We computed bootstrap standard errors (1,975 replications, seed = 1211) instead of the QML estimates. Confidence intervals are bias-corrected. We see that the joint analysis conveys the same impression as the marginal chi-square analyses.

Table 5: Fractional logit regression for combined efficiency score

	Coefficient	Bias	Standard Error	[Bias-corrected 95% Confidence Interval]	
Age	-0.5922	-0.0191	0.2863	-1.2026	-0.0667
Size_Small	-0.2519	-0.0754	0.4474	-1.2010	0.5041
Size_Medium	-0.3752	-0.0749	0.4246	-1.2748	0.3363
Type_Ecological	0.6050	-0.0161	0.3232	-0.0410	1.2016
Type_Product	0.9640	0.0030	0.3011	0.3405	1.5174
Constant	1.6077	0.0992	0.5059	0.9684	2.8162

We see significant type and age effects, but not a size effect. Figure 2 and 3 are box-plots describing the observations on the combined efficiency considering, separately, type and age effects. They are not related to the fractional regression model, but are in close agreement. Comparing the medians (center of the boxes), one can observe the dominance of the Product type and of the older research centers.



Fig, 2

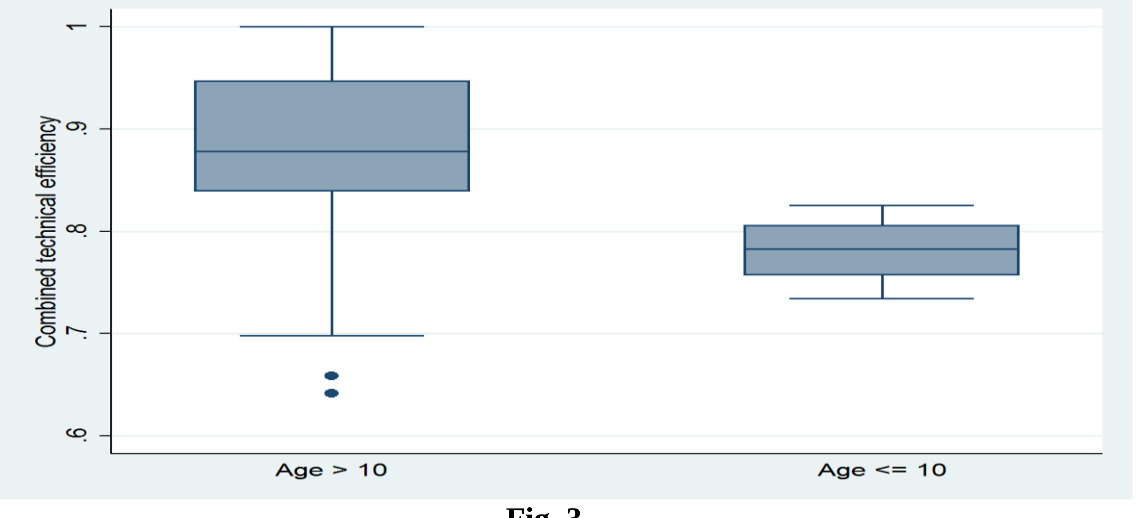


Fig. 3

5. Concluding Remarks

We were successful in modeling Embrapa's production system by applying a deterministic frontier DEA model. The model is justified, given the nature of the response where it is less likely to observe idiosyncratic than deterministic errors. Indeed, a stochastic frontier using the whole sample and a similar specification does not seem to converge.

A better approach was achieved by modeling the DEA time measurements for each research center as a dynamic first-order autoregressive panel, including covariates effects. This idea has appeal since it assumes a common autoregressive coefficient. With only a few time observations for each research center, it is not sensible to estimate separate coefficients. The common estimated autoregressive coefficient is used to define a sequence of weights that decrease over time, reflecting the decreasing importance of lagged efficiency scores. The Arellano–Bond test validates the dynamic model.

The final combined scores show a strong association with age and type, but not with the size of a research center. The size of the research center is important since production variables were normalized according to the number of employees, to make units more comparable, reducing unwanted scale effects from potentially biasing the results. Fractional regression consubstantiates this approach.

We also notice from the fractional regression that previous experience with the production evaluation process has a positive effect on the combined efficiency score. It was not possible to include this effect on the dynamic panel, due to collinearities in differences. Research centers classified as Product are more efficient than the others (Ecological and Thematic). This classification of the research centers has been an object of discussion in the company. Based on the fractional regression, we see that in terms of efficiency levels, the classification is not unreasonable.

References

[1] Souza GS, Alves E, Avila AFD. Technical efficiency of production in agricultural research. Scientometrics 1999;46(1):141-60. DOI: 10.1007/BF02766299

[2] Souza GS, Gomes EG, Magalhães MC, Avila AFD. Economic efficiency of Embrapa's research centers and the influence of contextual variables. Pesqui Oper. 2007;27(1):15-26. DOI: 10.1590/S0101-74382007000100002

[3] Avila AFD, Gomes EG, Souza GS, Filho, RCP, de Souza MO. Avaliação de desempenho de unidades de pesquisa agropecuária: métricas e resultados da experiência da Embrapa. Brasília: Embrapa - Secretaria de Gestão Estratégica; 2013. Documentos 16. ISSN: 1679-4680.

[4] Souza GS, Gomes EG. Scale of operation, allocative inefficiencies and separability of inputs and outputs in agricultural research. Pesqui Oper. 2013; 33(3):399-415. DOI: 10.1590/S0101-74382013005000007

[5] Souza GS, Gomes EG. Management of agricultural research centers in Brazil: a DEA application using a dynamic GMM approach. Eur J Oper Res. 2015;240(3):819-24. DOI: 10.1016/j.ejor.2014.07.027

[6] Färe R, Grosskopf S, Norris M. Productivity growth, technical progress, and efficiency change in industrialized countries: Reply. Am Econ Rev. 1997;87(5):1040-43. Available from: http://www.academicroom.com/article/productivity-growth-technical-progress-and-efficiency-change-industrialized-countries

[7] Lynde C, Richmond J. Productivity and efficiency in the UK: a time series application of DEA. Econ Model. 1999;16(1):105-22. DOI: 10.1016/S0264-9993(98)00034-0

[8] Tone, K. Advances in DEA theory and applications: with extensions to forecasting models. Oxford: John Wiley & Sons; 2017.

[9] Arellano M, Bover O. Another look at the instrumental variable estimation of error-components models. J Econom. 1995;68(1):29-51. DOI: 10.1016/0304-4076(94)01642-D

[10] Blundell R, Bond S. Initial conditions and moment restrictions in dynamic panel data models. J Econom. 1998;87(1):115-43. DOI: 10.1016/S0304-4076(98)00009-8

[11] Olesen OB, Petersen NC, Podinovski VV. Efficiency analysis with ratio measures. Eur J Oper Res. 2015;245(2):446-62. DOI: 10.1016/j.ejor.2015.03.013

[12] Banker RD, Charnes A, Cooper WW. Some models for estimating technical scale inefficiencies in data envelopment analysis. Manag Sci. 1984. 30(9):1078-92. DOI: 10.1287/mnsc.30.9.1078

[13] Kneip A, Park BU, Simar, L. A note on the convergence of nonparametric DEA estimators for production efficiency scores. Econom Theory 1998;14(6):783-93. DOI: 10.1017/S0266466698146042

[14] Banker RD. Maximum likelihood, consistency and data envelopment analysis: a statistical foundation. Manag Sci. 1993;39(10):1265-73. DOI: 10.1287/mnsc.39.10.1265

[15] Souza GS. Statistical properties of data envelopment analysis estimators of

production functions. Braz Rev Econom. 2001; 21(2):291-322. DOI: 10.12660/bre.v21n22001.2753

[16] Simar L, Wilson PW. Two-stage DEA: *caveat emptor*. J Product Anal. 2011;36(2):205-18. DOI: 10.1007/s11123-011-0230-6

[17] Stata. Stata base reference manual, release 14. College Station, Texas: Stata Press; 2015.

[18] Arellano M, Bond S. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Rev Econ Stud. 1991;58(2):277-97. DOI: 10.2307/2297968

[19] Papke LE, Wooldridge JM. Econometric methods for fractional response variables with an application to 401(k) plan participation rates. J Appl Econom. 1996;11(6): 619-32. DOI: 10.1002/(SICI)1099-1255(199611)11:6%3C619::AID-JAE418%3E3.0.CO;2-1

[20] Ramalho, E.A., Ramalho, J.J.S., Henriques, P.D. Fractional regression models for second stage DEA efficiency analyses. J Product Anal. 2010;34:239–255. DOI: I 10.1007/s11123-010-0184-0

[21] Stata. Base Reference Manual. Texas: Stata Press, 2017. Available from: https://www.stata.com/bookstore/base-reference-manual/