Análisis envolvente de datos de pronóstico de eficiencia en servicios de alimentos

Dynamic data envelopment analysis of food service efficiency forecast

Fernando Rojas^{1,2*}

Recibido 17 de julio de 2017, aceptado 22 de enero de 2018

Received: July 17, 2017 Accepted: January 22, 2018

RESUMEN

Proponemos una forma de análisis envolvente de datos que permite estimar la eficiencia de servicios de alimentos con salida estocástica relacionada con entradas predictivas en múltiples períodos mediante un modelo lineal generalizado auto-regresivo de media móvil (GLARMA). Esta forma de trabajo permite pronosticar el comportamiento futuro de la frontera de eficiencia de un *cluster* empresarial. Se da un ejemplo para los tomadores de decisiones cuando enfrentan este tipo de escenarios.

Palabras clave: DEA, eficiencia, GLARMA, pronóstico, servicios de alimentos.

ABSTRACT

We proposes a form of data envelopment analysis that allows estimating the efficiency of food services with stochastic outputs related to predictive inputs in multiple periods by generalized linear autoregressive moving average (GLARMA) model. This form of work allows forecasting the future behavior of the efficiency frontier of a cluster of companies. An example to the decision makers when they face scenarios like those described.

Keywords: DEA, efficiency, food service, forecast, GLARMA.

INTRODUCTION

Food services prepare food ration or food plate on a daily basis using raw materials, in a directed way to closed groups of users, see [5]. Productivity assessment has long been considered important for food service companies, see [3]. Data envelope analysis (DEA) corresponds to the most classic form of evaluation of productive efficiency in this and other areas. Interested readers can see [2] and [4].

The stochastic approach of DEA allows for both deterministic and random variables (RVs), producing

a single relative-to-best productivity index that relates all units under comparison, despite differing combinations of operating characteristics given that operating conditions are similar, see [6]. Because efficiency conceptually is dynamic and evolves over time, DEA models have been developed that capture the updating of inputs and outputs period after period, see [7]. However, very few and limited approaches manage to study the random and dynamic aspects that can impact the efficiency frontier in the foodservices industry. Then, we propose a Generalized linear autoregressive moving average (GLARMA) models that consider ARMA components but transforming

¹ Facultad de Farmacia. Escuela de Nutrición y Dietética. Universidad de Valparaíso. Valparaíso, Chile. E-mail: fernando.rojas@uv.cl

² Facultad de Farmacia. Centro de Micro-Bio Innovación. Universidad de Valparaíso. Valparaíso, Chile. E-mail: fernando.rojas@uv.cl

* Autor de correspondencia

the mean of the data with a link function in the line of generalized linear models (GLM).

We propose a form of DEA that allows estimating the efficiency of food service with stochastic outputs related to predictive inputs in multiple periods, with possibilities included to predict what is most likely to occur in terms of future efficiency. This study is organized as follows: Section 2 exposes the proposed methodology, whereas Section 3 illustrates it with a real-world case study; and Section 4 provides discussions, conclusions, limitations and future research.

METHODOLOGY

Static DEA to efficiency model

In deterministic DEA model a vector matrix characterizes the efficiency of a reference unit k in a cluster of N units, where each unit j or DMU_j has m inputs (X_{ij}) and s outputs (Y_{rj}):

$$\min \varphi; \text{ s.t. } X\lambda \leq \varphi X_k; Y\lambda \geq \varphi Y_k; \lambda'e = 1; \\ \lambda \geq 0; \varphi \geq 0, \quad (1)$$

where φ is a column vector with N elements, each of which is unity and the prime denotes transpose. Here the input (X_j) and output (Y_j) vectors ($j = 1, 2, \dots, N$) are all observed and in this deterministic parametric model assumes that there is no noise in the data like the classical DEA model and that the functional form parameters are deterministic given a priori ($Y = f(X, \beta)$), see [2], with an input oriented model. Here the reference unit k is compared with the other ($N - 1$) units in the cluster. Let $\lambda_{j^*} = (\lambda_{j^*})$ and φ^* be the optimal solutions of the above deterministic frontier model with all the slack variables zero. Then the reference unit k or DMU_k is technically efficient if $\varphi^* = 1$ and the first two sets of inequalities in equation (1) hold with equality. Thus the optimal value of φ^* provides a measure of technical efficiency (TE). If φ^* is positive but less than unity, then it is not technically efficient at the 100% level. To characterize the overall efficiency of the reference unit DMU_k one sets up the linear programming (LP) model as follows:

$$\min q'x; \text{ s.t. } X\lambda \leq x; Y\lambda \geq \varphi Y_k; \lambda'e = 1; \\ \lambda \geq 0; x \geq 0, \quad (2)$$

where q is an m -element vector of input prices as observed in the competitive market and x is an input vector to be optimally decided by DMU_k along with the weights λ_j . Here X_k and Y_k are the observed input and output vectors for the reference unit k , whereas x is the unknown decision vector to be optimally determined. Let λ^* and x^* be the optimal solution of LP model equation (2) with all slacks zero. Then the minimal input cost is given by $c_x^* = q'x^*$, whereas the observed cost of the reference unit is $c_x = q'X_k$

The dual problems corresponding to equations (2) and (1) appear as follows:

$$\max \alpha'Y_k + \xi; \text{ s.t. } \gamma \leq q; \gamma'X_j \geq \alpha Y_k + \xi; \\ J = 1, 2, \dots, N; \alpha, \gamma \geq 0; \xi \text{ free in sign, } (3)$$

and,

$$\max \alpha Y_k + \xi; \text{ s.t. } \gamma'X_j \geq \alpha Y_k + \xi; J = \\ 1, 2, \dots, N; \alpha, \gamma \geq 0; \xi \text{ free in sign. } (4)$$

Let asterisks denote optimal values and let DMU_k be efficient. Then it must follow from equation (3) that the production frontier for the k -th unit is as follows: $\alpha^*Y_k = \gamma^*X_k - \xi^*$ but since γ^* is constrained as $\gamma^* \leq q$, we must have $\alpha^*Y_k \leq q'X_k - \xi^*$. Thus so long as the actual inputs X_k are not equal to their optimal levels x , this efficiency gap measured by $(q'X_k - \xi^* - \alpha^*Y_k)$ may persist.

Dynamic efficiency with DEA model

Dynamic efficiency arises in most modern production processes, because they take several periods to adjust their inputs and outputs to the desired levels. When new technology is the source of productivity gain, the firms may take several periods to learn about the new technology and adopt it fully. Consider a dynamic extension of the overall efficiency model equation (4) in a simplified framework, see [6]. The overall efficiency model is then of the form:

$$\min J = \int_0^\infty \exp^{-\rho t} q'x(t) dt; \text{ s.t. } \sum_{j=1}^N X_j \lambda_j \\ \leq x(t); \sum_{j=1}^N Y_j \lambda_j \\ \geq Y_k(t); \lambda'e = 1; \\ \lambda \geq 0; x(t) \geq 0, \quad (5)$$

The decision variables are $\lambda = \lambda(t)$, $x(t)$ and the observed data are the input and output vectors $X = X_j(t)$, $Y = Y_j(t)$, where ρ is a positive rate of discount assumed to be known. Equations 3 and 4 that in 3 and 4, let asterisks indicate optimal values and the reference unit k be efficient over time. Then it must satisfy the following equality:

$$y^* = q(t).$$

Dynamic DEA model over time with modelling of outputs by GLARMA model

We introduce randomness in the dynamic DEA model by generalized linear autoregressive moving average regression model (GLARMA) that relationships between the outputs and inputs variables, when there have stochastic behavior over time. The assumption of independent and identically distributed (IID) RV for outputs $Y(t)$ depending over time t , for $t = 1, \dots, T$, can be violated in DEA. Unlike the classical DEA model where residuals are assumed for a regression model that relates output and inputs (v), as well as residuals to model the efficiency (γ), either additive ($Y = f(X; \beta) + v - \gamma$) or multiplicative ($Y = f(X; \beta) \exp(v) \exp(-\gamma)$), with unknown β parameters determined via maximum likelihood estimation, GLARMA models are derived from generalized linear model (GLM). Information on past time $\Omega(t-1)$ is summarized in the state variable $\eta(t)$. The systemic component of model GLARMA (p, q) is described by a link function of the mean where, $\eta(t) = g(\mu_Y(t))$ corresponds to the mean (μ) of the variable of interest to be predicted conditional to past information $\Omega(t-1)$ given by

$$\eta(t) = g(\mu_Y(t))$$

$$= X(t)^T \beta$$

$$+ \sum_{a=1}^p \varphi_a \{g(Y(t-a)) - X(t-a)^T \beta\} + \sum_{a=1}^q \theta_a \{g(Y(t-a)) - \eta(t-a)\}$$

where φ and θ corresponds to the ARMA components of a model of orders p and q , respectively. Note that $\beta(t)^T = (\beta_0, \beta_1, \dots, \beta_n)$ is a vector of coefficients associated with n observed covariates with dependence over time denoted by $X(t)^T = (X_{1,t}, X_{2,t}, \dots, X_{n,t})$, with $X_0 = 1$. This covariates can be inputs of dynamic DEA model as well as other determinants of output, such as dummy variables that indicate the presence or absence of some attribute or condition. The link function of the GLARMA model generally is the identity or logarithmic function and the corresponding model variance is assumed to be constant over time, see [1].

Then this type of model have versatility to fit an expected value predict ($E(Y(t)) = \widehat{Y}(t)$) to each time $t = 1, \dots, T$ and also forecast to h -forward step $t = T+1, \dots, T+h$. In particular if link function is identity GLARMA model to the expected value in each time t is expressed by:

$$\widehat{Y}(t) = X(t)^T \beta + \sum_{a=1}^p \varphi_a \{g(Y(t-a)) - X(t-a)^T \beta\} + \sum_{a=1}^q \theta_a \{g(Y(t-a)) - \eta(t-a)\} \quad (8)$$

and forecast in $T+1$ with GLARMA identity is expressed by:

$$\widehat{Y}(T+1) = X(T+1)^T \beta + \sum_{a=1}^p \varphi_a \{g(Y(T+1-a)) - X(T+1-a)^T \beta\} + \sum_{a=1}^q \theta_a \{g(Y(T+1-a)) - \eta(T+1-a)\} \quad (9)$$

Data Set We selected 10 food services that carry out business in closed care centers in the commune of Valparaíso, Chile, from 2001 and 2016 year. Based in the study of Reynolds and Thompson (2007) we selected the following variables to dynamic stochastic DEA models, see Table 1:

Table 1. Inputs and Output to dynamic stochastic DEA models.

Variable	Input/ Output	Measure as
Sales	Output	USD\$/year
Server wage	Input	USD\$/hour
Seats	Input	Number
Stand alone	Input	0 (stand alone) or 1 (adjacent to other building(s))

Computational framework

We have implemented the methodology in the R software (<http://www.r-project.org/>). Specifically, we have used the base package for descriptive statistics, glarma and forecast packages to perform statistical analysis on expected output and forecast of sales in food service value using time series data, respectively. We use Benchmarking package to solve the programming in the stochastic dynamic DEA method based on equation (5).

RESULTS

Figure 1 shows the autocorrelation function ACF plots for the time series of sales to each one of the companies of food service. The plots indicate that practically all of the components have time-dependence for their sales. A similar behaviour can be observed in the Partial ACF (PACF) plots, omitted due to lack of space.

Figure 2 show changes in dynamic stochastic DEA frontier of 9 food services in selected years, also of forecast the change in the frontier of efficiency in a future year.

Figure 1. ACF Plot of time series of sales of 9 food services.

Figure 2. Changes in dynamic stochastic DEA frontier of 9 food services in selected years.

Table 2 shows the set of indicated companies with the best fit to GLARMA models with parameters of the normal distribution chosen by the Akaike information criteria (AIC), displaying: parameter of standard deviation (SD) conditional over time to normal distribution of DPUT, autoregressive coefficient of order one (AR1), coefficient of explanatory variables showed in Table 1, intercept coefficients with it is standard error (SE) in parenthesis, and value of AIC.

DISCUSSIONS

The present proposal is useful for evaluating and forecasting changes in the frontiers of efficiency in the area of food service.

Through GLARMA modeling of time series that have a random behavior ordered in time of sales, seen as an output of the food service production process, it is possible to predict the future results of these companies employing process inputs as related predictors. The modeling offered allows updating the temporary characterization of the efficiency in this type of companies, which is a fundamental tool for the

Table 2. Estimated parameters of GLARMA model for indicated companies.

Company	SD	AR1(SE)	Seats(SE)	Stand alone(SE)	Intercept(SE)	AIC
1	0,72	0,41(0,34)	97(536)	540(397)	1,49(1,2)	42,34
2	0,51	-0,20(0,34)			0,86(0,45)	287,96
3	0,47	-0,18(0,53)		-0,39(0,38)	0,63(0,5)	29,78
4	0,37	0,0064(0,29)		0,36(0,29)	1,9(0,35)	21,55
5	0,64	0,28(0,25)			1,22(0,6)	35,63
6	0,42	0,41(0,21)			2,88(0,48)	23,84
7	0,49	0,11(0,49)	-0,07(0,06)		2,85(1,86)	29,59
8	0,45	0,36(0,19)	0,019(0,008)		-0,18(2,29)	27,41
9	0,53	0,75(0,02)	-0,004(0,006)		3,91(1,66)	32,13

strategic planning that the managers perform, as much of operations as in general level in these companies. The model is characterized by being very flexible regarding diverse probabilistic distributions that could be presenting the data of sale.

In this respect, classical regression models have important constraints on their assumptions, which are largely overcome with models such as GLARMA, which go in the GLM line, where variability is assumed directly on the response variable. The GLARMA statistical framework is very robust, providing important tools for diagnosing, selecting, relating, fitting and fore casting through this approach. In the present work, we have not gone into these aspects in order not to deviate from the objective to be addressed in order to construct a dynamic stochastic DEA model applicable to food services. Nevertheless the deepening in its thematic applicable in the areas of DEA and stochastic frontier analysis, can represent an excellent way of promoting research in the future.

The analysis of our results shows clearly that the efficiency frontier can change in the different periods (Figure 2), and that this is not maintained inactive in the different periods of time, nor in the future to forecast. Also, we can link the results of the different food service companies analyzed, know their temporal variability, and link their relationship predictors that explain either significantly or not significantly, using the GLARMA coefficients for each company under analysis (Table 2).

The methodology presented has the limitation of having been developed as a univariate model, that is, it only considers a response variable, unlike the classical and deterministic DEA that was conceived to consider a vector of multiple variables both as input and outputs. This is an extension field that must be explored to arrive at a multivariate stochastic model within the time series framework.

ACKNOWLEDGEMENTS

This research was partially funded by “Grants-CONICYT” of the National Commission for Scientific and Technological Research (CONICYT) government of Chile.

REFERENCES

- [1] M.A. Benjamin, R.A. Rigby D.M. and Stasinopoulos. “Generalized autoregressive moving average models”. Journal of the American Statistical Association. Vol. 98, pp. 214-223. 2003.
- [2] P. Bogetoft, L. Otto. “Benchmarking with Dea, Sfa and R”. Springer Science & Business Media. New York, USA. Vol. 157. 2010.
- [3] D. Reynolds. “Hospitality-productivity assessment using data-envelopment analysis”. The Cornell Hotel and Restaurant Administration Quarterly. Vol. 44, pp. 130–137. 2003.
- [4] D. Reynolds, G.M. Thompson and G. M. “Multiunit restaurant productivity assessment using three-phase data envelopment analysis”.

- [5] International Journal of Hospitality Management. Vol. 26, pp. 20–32. 2007.
- [6] F. Rojas and V. Leiva. “Inventory management in food companies with demand statistically dependent”. ARLA: Academia Revista Latinoamericana de Administración. Vol. 29, pp. 450–485. 2016.
- [7] J.K. Sengupta. “Efficiency analysis by stochastic data envelopment analysis”. Applied Economics Letters. Vol.7, pp. 379–383. 2000.
- [8] K. Tone and M. Tsutsui. “Dynamic dea: A slacks-based measure approach”. Omega. Vol. 38, pp. 145– 156. 2010.