A NONPARAMETRIC APPROACH TO ASSESS THE DETERMINANTS OF PRODUCTION AND TECHNICAL EFFICIENCY IN THE BRAZILIAN AGRICULTURE

Geraldo da Silva e Souza
Eliane Gonçalves Gomes
Eliseu Roberto de Andrade Alves
Brazilian Agricultural Research Corporation (Embrapa)
PqEB, Av. W3 Norte final, 7070-901, Brasília, DF, Brazil
{geraldo.souza; eliane.gomes; eliseu.alves}@embrapa.br

ABSTRACT

We use nonparametric output oriented DEA-VRS model to estimate a production frontier for the Brazilian agriculture using county data from the 2006 Agricultural Census. The frontier allows the assessment of the elasticities of labor, land and technological inputs. Contextual variables related to environmental, social and demographic dimensions and indicators of technical assistance, credit and regional dummies are used to explain technical efficiencies by means of fractional regression and bootstrap, assuming the efficiencies to follow the inverse extreme-value distribution. Only environmental effects are non-significant.

KEYWORDS. Data envelopment analysis. Two-stage approach. Agriculture.

Main area. Data Envelopment Analysis

RESUMO

Neste artigo foi usado o modelo DEA-VRS com orientação a *output* para estimar a fronteira de produção para a agricultura brasileira, em base municipal, considerando-se os dados do Censo Agropecuário de 2006. A fronteira permite o estudo das elasticidades dos *inputs* terra, mão de obra e insumos tecnológicos. Variáveis contextuais relacionadas a indicadores de desenvolvimento social, demográfico e ambiental, além de acesso a crédito, assistência técnica e dummies regionais, foram usadas para explicar a eficiência técnica com o uso de modelos de regressão fracionada e *bootstrap*, assumindo-se que as eficiências seguem a distribuição inversa do valor extremo. Apenas os efeitos ambientais foram não significativos.

PALAVARAS CHAVE. Análise de envoltória de dados. Abordagem em dois estágios. Agricultura.

Área principal. Análise de Envoltória de Dados

1. Introduction

Recently much interest has been seen in the literature regarding the fit of production frontiers to agricultural data in the presence of technical effects. Souza GS et al. (2013) fit a stochastic frontier to primary census data, estimated input elasticities assuming a Cobb-Douglas production function and assessed the effects of research, technical assistance, regional dummies, and income classes on technical efficiencies modeled by means of a half-normal distribution. Avila et al. (2013), through total factor productivity (TFP) indexes, analyzed the evolution of the agricultural productivity using partial and total factor productivity measures by countries, sub-regions and Latin American Countries as a whole, emphasizing Brazil. They also used an econometric analysis to identify sources of TFP growth in agriculture. Other studies on TFP growth in Brazil are Gasques et al. (2012, 2013). At world level, it is worth mentioning Fuglie (2010) and Fuglie et al. (2012).

The main conclusion one may infer from the studies combining an estimation of a production frontier and contextual effects affecting agriculture technical efficiency is that technology dominates the production function and technical assistance is an important component to increase efficiency. The dominance of technology compared to land and labor is confirmed by Gasques et al. (2012) and Souza GS et al. (2013). Also, the current literature expounds the process through which the growth of agriculture leads to the development of other sectors and, in doing so, promotes gains in income and welfare. In order to have a strong agricultural sector, however, growth in productivity and strong agricultural research efforts are needed, both of which play key roles in poverty reduction in addition to propelling the agricultural sector and the economy as a whole (Christiaensen and Demery, 2007; Thirtle et al., 2003; Fan et al., 2008; Fan and Zhang, 2008).

Further to these developments, we propose the use of nonparametric to assess agriculture technical efficiency in Brazil, at a county level. Robust measures of efficiency are proposed based on ranks and used to estimate the frontier, and from it derive input elasticities. Agricultural partial indicators indexes for environmental, social and demographic variables are combined with regional effects, credit and technical assistance to explain Data Envelopment Analysis (DEA) measurements. The statistical analysis is carried out by means of fractional regression methods (Ramalho et al., 2010) and nonparametric bootstrap to handle the DEA induced correlation among counties. The approach is new in the context of Brazilian agriculture applications.

Our discussion proceeds as follows. In Section 2 we describe the production and the contextual variables used in the article. Sections 3 and 4 describe the models involved in the theoretical approach. Section 5 is on statistical results. Finally, Section 6 summarizes the main findings and concludes the article.

2. Production Variables, Covariates, and Data

The data components involved in this work drew from the Brazilian Agricultural Census 2006. There were three key types of variables necessary to conduct the nonparametric frontier analysis proposed here: inputs, outputs and selected explanatory variables for the inefficiency components of the production process.

For the inputs and outputs, data were collected drawing from value of expenditures in outputs and inputs. The choice of values as opposed to quantities arose from the fact that using the value of output allows for aggregation of all agricultural outputs.

Farm data were pooled to form averages for each county. A total of 5,474 counties provided valid data for our analysis. This figure represents 98.4% of the total county number. The decision-making unit (DMU) for our production analysis is the county.

Table 1 provides a complete list of inputs and outputs used to construct the production variables used in the analysis. The production variables used are straight-forward and do not require further explaining. They are measured on a farm level, as provided by the census, and aggregated by county.

Table 1: Variable descriptors.

Variable	Components	Unit	Notes
Y (output)	Value of production of cattle, swine, goats, equines, buffaloes, donkeys, asinine, mules, sheep, other birds, rabbits, apiculture, sericulture, raniculture, aquaculture, horticulture, flowers, forestry, agro industry, permanent crops, temporary crops, extractive activities.	Reais	-
Land	four percent of land expenses, the rent paid for the land	Reais	-
Labor	Salaries or other forms of compensation paid to family and hired laborers	Reais	-
Capital (technological inputs)	Machinery, improvements in the farm, equipment rental, value of permanent crops, value of animals, value of forests in the establishment, value of seeds, value of salt and fodder, value of medication, fertilizers, manure, pesticides, expenses with fuel, electricity, storage, services provided, raw materials, incubation of eggs and other expenses.	Reais	Value of permanent crops, forests, machinery, improvements on the farm, animals and equipment rental were depreciated at a rate of six percent over a number of years (varying according to the category).

The contextual variables considered are dummies for regional effects (South, Southeast, North, Northeast, and Center-West), proportion of farmers who received technical assistance, total financing per farm, and performance county indexes in the social, environmental, and demographic dimensions. These indexes require further comments. They have been considered in total or in part in Embrapa (2001), Monteiro and Rodrigues (2006), Rodrigues et al. (2010), and Souza MO et al. (2013). The idea was also used by the National Confederation of Agriculture (2012, personal communication) to develop an overall indicator of rural development. Our version of these quantities presented here are similar, but not coincident with these sources. The technique of index construction is based on the work of Moreira et al (2004).

Social dimension

The variables comprising the social dimension reflect the level of well-being, favored by factors as availability of water and electric energy in the rural residence. They reflect also indicators of the level of education, health, and poverty of the rural residences.

The data used in the social dimension were extracted from the Brazilian Demographic Census 2010, Brazilian Agricultural Census 2006 and from the databases of National Institute of Research and Educational Studies – INEP (referring to education in 2009) and the Ministry of Health (2011 data).

Demographic dimension

The variables comprising the demographic dimension capture aspects of the population dynamics, which relates to rural development. These are proportion of rural to urban population, average size of a rural family, aging rate, migration index, and the ratio of the inactive population

(0 to 14 years and 60 years or more of age) to the active population (15 to 59 years of age). The data source is the Brazilian Demographic Census 2010.

Environmental dimension

The variables comprising the environmental dimension are proportions of farmers practicing the technique of vegetation fires, that use agrochemicals, practicing crops rotation, practicing minimum tillage, practicing no-tillage, planting in contour lines, providing proper garbage disposal, proportion of forest and agro-forest areas, and proportion of degraded areas. The data source is the Brazilian Agricultural Census 2006.

All variables within each dimension are measured in such way to correlate positively with the given dimension. They were rank transformed and normalized by the maximum, which in this case is the number of counties. Each specific dimension index is a weighted average of the normalized variables comprising the dimension, with weights defined by the relative squared multiple correlation obtained in the regression of a variable with all others, i.e. if R_i^2 is the squared multiple correlation of the regression, considering the *i*th variable as the dependent variable in the dimension, its weight will be $R_i^2 / \sum_j R_j^2$.

3. DEA Models

Consider a production process composed of *n* DMUs. Each DMU uses varying quantities of *m* different inputs to produce varying quantities of *s* different outputs.

Denote by $Y = (y_1, y_2, ..., y_n)$ the production (output) vector of the *n* DMUs. The *r*th component of *Y* is the output of DMU *r*. Denote by $X = (x_1, x_2, ..., x_n)$ the $3 \times n$ input matrix. The *r*th column of *X* is the input vector of DMU *r*. The matrices $Y = (y_{ij})$ and $X = (x_{ij})$ must satisfy $p_{ij} \geq 0$, $\sum_i p_{ij} > 0$ and $\sum_j p_{ij} > 0$, where p is x or y. This condition is satisfied in our application. The DEA measure of technical efficiency of production, output oriented, under variable returns to scale for DMU $o \in \{1, 2, ..., n\}$, with production vector (x_o, y_o) is

$$\phi^*(x_o, y_o) = \max_{\phi,\lambda} \phi$$

subject to

$$\text{i) } Y\lambda \geq \phi y_o, \ \text{ ii) } X\lambda \leq x_o \text{ and iii) } \lambda \geq 0, \lambda 1 = 1, \ \phi \text{ free}$$

With a view to an output augmentation program, the question we ask is: what proportional rate ϕ can be uniformly applied to augment the output vector y_o, without increasing the input vector x_o? The solution ϕ^* is the largest ϕ with this property.

Production variables (output and three inputs) were rank transformed and normalized for the computation of technical efficiency. The rank transformation is a nonparametric approach that allows reasonable efficiency estimation without much influence of extremely large/low values of inputs and the output. Without it a DEA analysis will not provide useful results. The estimation with ranks here mimics their use in Nonparametric Analysis in the presence of non-normality, outliers, and heteroskedasticity (Connover, 1999).

4. Statistical Analysis of Factors Influencing Efficiency Scores

Care should be exercised in the statistical inference related to the assessment of the effects of contextual variables in DEA performance measurements. Firstly, DEA measures are correlated by the very nature of their computations. Secondly, the potential correlation of a covariate with the efficiency index error may invalidate the analysis in a manner similar to what happens with the use of ordinary least squares in the presence of endogenous independent variables. See Simar and Wilson (2007) for more details.

Here we are interested in assessing real differences in performance due to regional effects, financing, technical assistance, and social, demographic and environmental intensity scores. We do not expect endogeneity of the regional effects, and the social, environment and demographic indexes. Exogeneity should be verified for financing and technical assistance.

Ramalho et al. (2010) suggest the use of fractional regression to analyze DEA response models. We follow this approach here. In this context, let an observed DEA response $\hat{\theta}$ be dependent on a vector of covariates *w*. Ramalho et al. (2010) consider a one- and a two-part model, the models differing in the way the efficient units are treated. In the one-part model it is assumed that $E(\hat{\theta}\mid w)=G\left(w\delta\right)$, where *G*(.) is a probability distribution function. The model is well defined even when $\hat{\theta}$ puts positive probability mass at one. The unknown parameter δ is then estimated by quasi maximum likelihood (QML), maximizing $\sum_{i=1}^{n}\left(\hat{\theta}_i \log\left(G\left(w_i\delta\right)\right)+(1-\hat{\theta}_i)\log\left(1-G\left(w_i\delta\right)\right)\right)$. The two-part model uses the whole sample to estimate the model $\text{Prob}\left(\hat{\theta}_i=1\middle| w_i\right)=F\left(w_i'\beta\right)$, where β is an unknown parameter vector, and F is a known probability distribution function. For the second part it is assumed $E(\hat{\theta}_i\mid w_i)=G\left(w_i'\delta\right)$ for the responses in (0,1). Typical choices for F and *G* in both models are the logistic, probit and the inverse extreme-value distribution functions. These are given by $G\left(u\right)=e^u/(1+e^u)$, $G\left(u\right)=\Phi(u)$, and $G\left(u\right)=1-e^{-e^u}$, respectively. The function $\Phi(u)$ is the standard normal distribution function.

We favor the use of the one part model here since it assumes the same influence of covariates on the response and it is more parsimonious. In our application we see no reason to specify distinct models to efficient and inefficient DMUs.

For the one-part model, Papke and Wooldridge (1996) show that under the correct specification of the mean function $\sqrt{n}\left(\hat{\delta}-\delta\right)\xrightarrow{d}N(0,V)$. *V* is estimated using the following formulations. The QML estimator is efficient within the class of estimators containing all linear exponential family-based QML and weighted nonlinear least squares estimators (Ramalho et al, 2010).

$$\hat{V}=\left(\hat{\mathrm{A}}\right)^{-1}\hat{\mathrm{B}}\hat{\mathrm{A}}$$

$$\hat{\mathrm{A}}=\frac{1}{\mathrm{n}}\sum_{i=1}^{n}\frac{\hat{g}_i^2}{\hat{G}_i(1-\hat{G}_i)}w_i'w_i$$

$$\hat{\mathrm{B}}=\frac{1}{\mathrm{n}}\sum_{i=1}^{n}\frac{\hat{u}_i^2\hat{g}_i^2}{\left(\hat{G}_i(1-\hat{G}_i)\right)^2}w_i'w_i$$

$$\hat{G}_i=G(w_i'\hat{\delta}),\ \hat{g}_i=G'(w_i'\hat{\delta}),\ \hat{u}_i=\hat{\theta}_i-\hat{G}_i$$

The validity of the asymptotic distributional properties of the QML estimator is also dependent on a condition not pointed out by Ramalho et al. (2010). Observations are assumed to be uncorrelated. See Papke and Wooldridge (1996). For this reason, our choice to derive the distributional properties of $\hat{\delta}$ is the nonparametric bootstrap. In this context, we compute $\hat{\delta}$ by QML for repeated samples from the counties observations, with replacement. The process is repeated 5,000 times.

5. Statistical and Other Findings

We begin our discussion reporting some descriptive statistics related to the technical efficient measurements. The histogram of the county efficiency scores ('vrs') in Figure 1 suggests an underlying uniform distribution in the interval [0,1]. The mean efficiency is 0.516 and the median 0.517. The states with highest median scores are Mato Grosso (0.789), São Paulo (0.876), Distrito Federal (0.876), and Mato Grosso do Sul (0.889). The least efficient in the median are Piauí (0.104), Paraíba (0.172), Ceará (0.220), Rio Grande do Norte (0.228), and Bahia (0.229). Median value for the Northeast region is 0.219, for the North is 0.410, for the South 0.645, for the Southeast 0.666, and 0.747 for the Center-West. The North and Northeast are dominated clearly by the other regions. The regional differences are depicted by the box-plots in Figure 2, where one notices the heavy presence of outliers in the Northeast region.

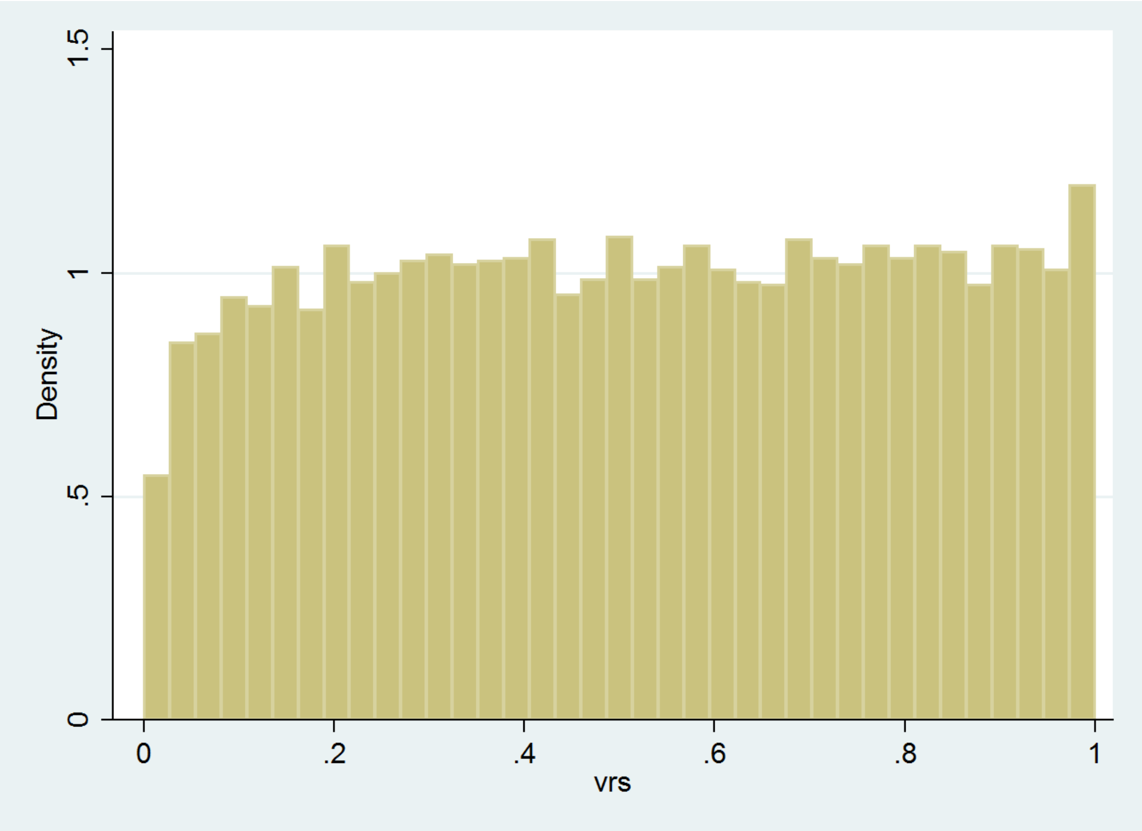

Figure 1: Histogram of efficiency scores ('vrs').

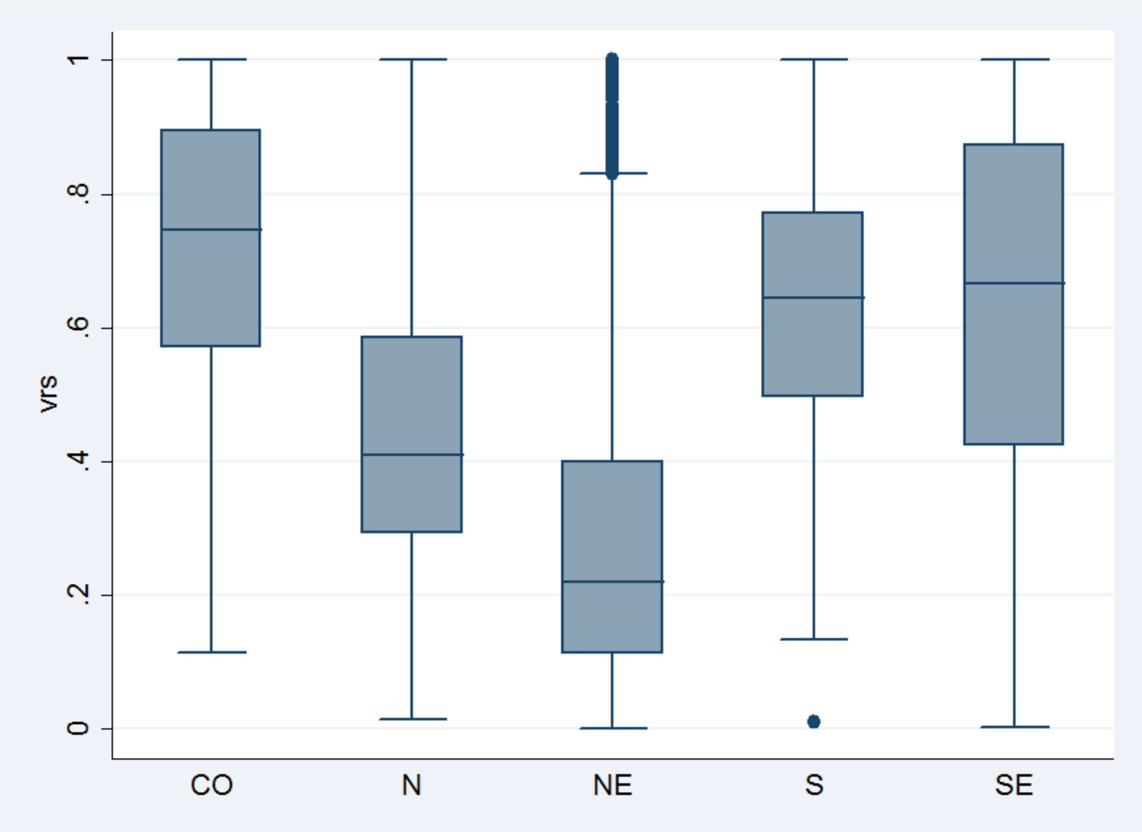

Figure 2. Efficiency ('vrs') box-plots by region ('CO' = Center-West; 'N' = North; 'NE' = Northeast; 'S' = South; 'SE' = Southeast).

The output augmentation program allows having an idea of the elasticity input effects. An approximation to these quantities can be obtained DEA-projecting on the frontier and estimating a Cobb-Douglas production function to the resulting data. Table 2 shows the results when the projection is carried out on the production values. R^2 is 0.539 and the elasticities are dominated by technological inputs, followed by labor and land. These results are consistent with the stochastic frontier approach of Souza GS et al. (2013), who found the technological input elasticity to be three times the labor elasticity and seven times the land elasticity. The corresponding figures here are two and nine, respectively.

Table 2. Input elasticities.

Variable	DF	Parameter estimate	Standard error	t Value	Pr > |t|
Intercept	1	7.37246	0.05116	144.10	< 0.0001
Labor	1	0.12987	0.00797	16.29	< 0.0001
Land	1	0.02655	0.00973	2.73	0.0064
Capital	1	0.24246	0.01218	19.91	< 0.0001

Our next step is the assessment of technical effects using the one part model of Ramalho et al. (2010) and Papke and Wooldridge (1996). We begin our discussion choosing among the three commonly used distribution functions – logistic, probit, and inverse extreme-value. General goodness of fit statistics are shown in Table 3. The inverse extreme-value distribution produces slightly better information criteria values. It is our choice here.

Table 3. Goodness of fit statistics for the choice of the distribution function to be used in the one-part model.

Model	AIC	BIC
Logistic	6480.2	6546.3
Probit	6477.5	6543.6
Inverse extreme-value	6458.1	6524.2

We proceed verifying exogeneity of financing and technical assistance. For this purpose we use the LM test of Heij et al. (2004). Residuals from the inverse extreme-value fractional model are regressed on the residuals and of the regressions of financing (r1) and technical assistance (r2) on the instrumental variables (regional effects and social, demographic and environmental indexes). This final regression, beyond these residuals, includes, additionally, all variables used in the initial model as independent variables. The null hypothesis of exogeneity is tested by the joint significance of the residuals r1 and r2. Considering the nonparametric bootstrap estimate (Stata, 2013) of the variance-covariance matrix of the parameters, the test statistic has a value of 2.05, with a p-value 0.359 derived from the chi-square distribution with two degrees of freedom, non significant.

Table 4 shows the estimation results from the fit of the inverse extreme-value fractional model with bootstrap standard errors.

Marginal effects are assessed through estimation of the quantities $\dfrac{\partial E(\hat{\theta}\,|\,w)}{\partial w_j} = \dfrac{dG(\mu)}{d\mu}\delta_j \,|\, \mu = w\delta$, $\dfrac{dG(\mu)}{d\mu} = \exp\{-\exp(\mu)\}\exp(\mu)$. It is seen that the effects depend on the parameters and on the values of the covariates. The maximum response is achieved when $\mu = 0$. The graphs in Figure 3 show marginal effects for financing and technical assistance. The maximum responses possible in each case are 0.467 and 0.200, respectively.

Table 4. Fractional regression estimation (SAS 9.3) using the inverse extreme-value distribution. Standard errors and confidence limits are based on 5,000 bootstrap replications.

Parameter	Estimate	Standard error	95% confidence interval	
			Lower limit	Upper limit
Intercept	-2.40170	0.0825	-2.5624	-2.2391
Center-west	0.02352	0.0240	-0.0231	0.0696
North	0.07023	0.0379	-0.0025	0.1443
Northeast	0.02194	0.0375	-0.0525	0.0959
South	-0.27030	0.0194	-0.3081	-0.2316
Financing	1.27050	0.0488	1.1749	1.3660
Technical assistance	0.54300	0.0488	0.4457	0.6378
Social index	0.66210	0.0661	0.5317	0.7920
Demographic index	1.47160	0.0790	1.3187	1.6299
Environment index	0.17360	0.1262	-0.0818	0.4230

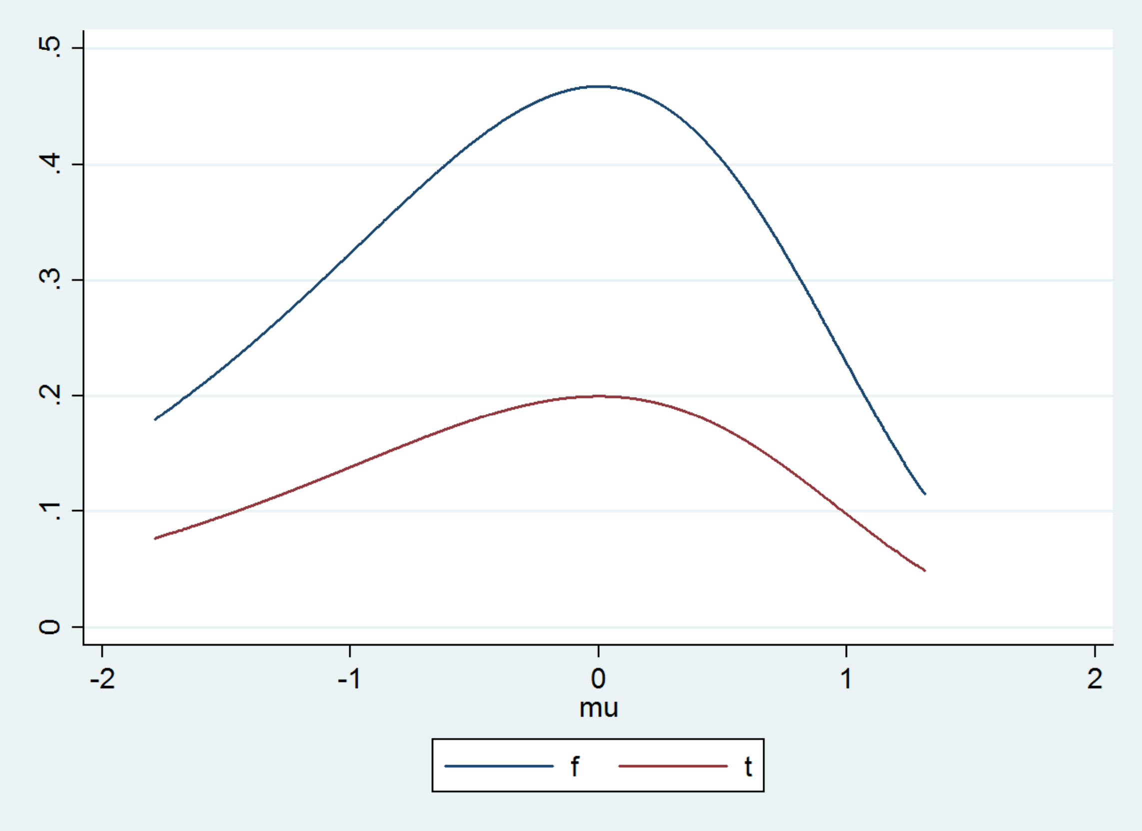

Figure 3: Partial effects of financing (‘f’) and technical assistance (‘t’).

On a regional basis we obtain the median responses shown in Table 5. The message here is that financing and extension programs, *ceteris paribus*, will have slower effects on the technical efficiency of production in the North and Northeast regions. Improvements on the social and demographic indexes of these regions may change this picture.

Table 5. Median responses of partial effects by region.

Region	Financing median effect	Technical assistance median effect
Center-West	0.435	0.186
North	0.403	0.172
Northeast	0.296	0.127
South	0.453	0.194
Southeast	0.415	0.177

6. Summary and Conclusions

We fitted a nonparametric production technology to county data derived from the Brazilian Agricultural Census of 2006 by means of data envelopment analysis. The output is the total value of the county agricultural production and the inputs are the expenditures on labor, land and technological inputs. Further to the analysis, we investigate the magnitude of a set of contextual variables on technical efficiency by means of fractional regression. These are regional dummies, financing, technical assistance and social, demographic and environmental indexes. Bootstrap methods were used to overcome the correlation among DMUs induced by the DEA scores calculations. Instrumental variable methods were used to account for potential endogeneity of some of the covariates.

We conclude that technological inputs dominate the production function, followed by labor and land. This is an indication that public policies leading to land access should be accompanied of proper rural extension if one is interested in improving production. Technical efficiency shows a strong dependency on all contextual variables with the exception of the environmental indexes. Financing and the demographic index dominate the relationship. On a regional basis financing and technical assistance show higher responses, *ceteris paribus*, for South, Center-West, and Southeast regions.

The urbanization, the shortage of land and labor and the export demand for farm products changed the organization of the agriculture in Brazil. This is now concentrated in specialized regions, where the production systems save land and labor by intensive use of technology. Rural establishments that are out of these regions, or poles, face very unfavorable conditions. It is imperative to improve social and demographic indexes in order to achieve maximum response with credit and rural extension actions.

References

Avila, A.F.D., Garagorry, F., e Cardoso, C.C. Produção e produtividade da agricultura brasileira: taxas de crescimento, comparações regionais e seus determinantes. In: In: Alves, E.R.A., , G.S., e Gomes, E.G.. (Org.), *Contribuição da Embrapa para o desenvolvimento da agricultura no Brasil*, 1ed., Embrapa Informação Tecnológica, Brasília, 87-124, 2013.

Christiaensen, L., and Demery, L. D*own to Earth: Agriculture and Poverty Reduction in Africa*, World Bank, Washington D.C., 2007.

Conover, M.J. *Practical nonparametric statistics*, 3rd ed., Wiley, New York, 1999.

Fan, S., and Zhang, X. (2008). Public Expenditure, Growth and Poverty Reduction in Rural Uganda, *African Development Review/Revue Africaine de Développement*, 20(3), 466-496.

Embrapa. (2001). *Balanço Social da Pesquisa Agropecuária Brasileira*. Brasília, DF: Embrapa, Secretaria de Comunicação, Secretaria de Gestão Estratégica.

Fan, S., Yu, B., and Jitsuchon, S. (2008). Does allocation of public spending matter in poverty reduction? Evidence from Thailand, *Asian Economic Journal*, 22(4), 411-430.

Fuglie, K. Total factor productivity in the global agricultural economy: evidence from FAO data." In: Pardey, P., Alston, J., and Babcock, B. (Eds.), *The Shifting Patterns of Agricultural Production and Productivity Worldwide*, The Midwest Agribusiness Trade Research and Information Center, Iowa, 63-95, 2010.

Fuglie, K., Wang, S.L., and Ball, V.E. *Productivity growth in agriculture: An international perspective.*, CAB International, Oxfordshire, 2012.

Gasques, J. G., Bastos, E.T., Valdes, C., e Bacchi, M.R.P. (2012). Produtividade da agricultura brasileira e os efeitos de algumas políticas, *Revista de Política Agrícola*, 3, 83-92.

Gasques, J. G., Bastos, E.T., Valdes, C., e Bacchi, M.R.P. Produtividade e crescimento: algumas comparações. In: Alves, E.R.A., Souza, G.S., e Gomes, E.G.. (Org.), *Contribuição da Embrapa para o desenvolvimento da agricultura no Brasil*, 1ed., Embrapa Informação Tecnológica, Brasília, 125-139, 2013.

Heij, C., De Boer, P., Franses, P.H., Kloek, T., and Van Dijk, H.K. *Econometric Methods with Applications in Business and Economics*. Oxford University Press, New York, 2004.

Monteiro, R.C., Rodrigues, G.S. (2006). A system of integrated indicators for socio-environmental assessment and eco-certification in agriculture – ambitec-agro, *Journal Technology Management & Innovation*, 1 (3), 47-59.

Moreira, T.B., Pinto, M.B., and Souza, G.S. (2004). Uma metodologia alternativa para mensuração de pressão sobre o mercado de câmbio, *Estudos Econômicos*, 34, 73-99.

Papke, L.E., and Wooldridge, J.M. (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates, *Journal of Applied Econometrics*, 11, 619-632.

Ramalho, E.A., Ramalho, J.J.S., and Henriques, P.D. (2010). Fractional regression models for second stage DEA efficiency analyses, *Journal of Productivity Analysis*, 34, 239-255.

Simar, L., and Wilson, P.W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes, *Journal of Econometrics*, 136(1), 31-64.

Rodrigues, G.S., Buschinelli, C.C.A., Avila, A.F.D. (2010). An environmental impact assessment system for agricultural research and development II: institutional learning experience at Embrapa, *Journal Technology Management & Innovation*, 5 (4), 38-56.

Souza, M.O., Marques, D.V., Souza, G.S. (2013), Uma metodologia alternativa para cálculo dos índices de impactos sociais e ambientais das tecnologias da Embrapa, *Cadernos do IME-Série Estatística*, Universidade do Estado do Rio de Janeiro – UERJ, 34, 1-15.

Souza, G.S., Alves, E.R.A., Gomes, E.G., Magalhães, E., e Rocha, D.P. Um modelo de produção para a agricultura brasileira e a importância da pesquisa da Embrapa. In: Alves, E.R.A., Souza, G.S., Gomes, E.G. (Org.), *Contribuição da Embrapa para o desenvolvimento da agricultura no Brasil*, 1ed., Embrapa Informação Tecnológica, Brasília, 49-86, 2013.

Stata. *Base Reference Manual*. Stata Press, Texas, 2013.

Thirtle, C., Lin, L., and Piesse, J. (2003). The impact of research-led agricultural productivity growth on poverty reduction in Africa, Asia and Latin America, *World Development*, 31(12), 1959-1975.