

## **CONDITIONAL FDH EFFICIENCY, INCOME DISPERSION AND MARKET IMPERFECTIONS: THE CASE OF THE BRAZILIAN AGRICULTURAL CENSUS OF 2006**

**Geraldo da Silva e Souza  
Eliane Gonçalves Gomes**

Brazilian Agricultural Research Corporation (Embrapa)  
PqEB, Av. W3 Norte Final, 70770-901, Brasília, DF, Brazil  
geraldosouza@embrapa.br  
eliane.gomes@embrapa.br

### **ABSTRACT**

In this article we assess the effect of market imperfections and income inequality on rural production efficiency. The analysis is carried out using the notion of stochastic conditional efficiency computed in terms of free disposal hull (FDH) efficiency measurements. Free disposal hull and conditional FDH are output oriented with variable returns to scale. They are evaluated for rural production at county level, considering as output the rank of rural gross income and as inputs the ranks of land expenses, labor expenses, and expenses on other technological factors. The conditional frontier is dependent on income dispersion and market imperfections resulting from credit, technical assistance, and environmental, social and demographic indices. The econometric approach is based on fractional regression models and general method of moments (GMM). Overall, market imperfection variables act to reduce performance, and income dispersion is positively associated with technical efficiency.

**KEYWORDS: conditional FDH, GMM, fractional regression, rural income dispersion**

**Main area: DEA – Data Envelopment Analysis**

### **RESUMO**

In this article we assess the effect of market imperfections and income inequality on rural production efficiency. The analysis is carried out using the notion of stochastic conditional efficiency computed in terms of free disposal hull (FDH) efficiency measurements. Free disposal hull and conditional FDH are output oriented with variable returns to scale. They are evaluated for rural production at county level, considering as output the rank of rural gross income and as inputs the ranks of land expenses, labor expenses, and expenses on other technological factors. The conditional frontier is dependent on income dispersion and market imperfections resulting from credit, technical assistance, and environmental, social and demographic indices. The econometric approach is based on fractional regression models and general method of moments. Overall, market imperfection variables act to reduce performance, and income dispersion is positively associated with technical efficiency.

**PALAVRAS-CHAVE: FDH condicional, GMM, regressão fracionária, dispersão da renda rural**

**Main area: DEA – Análise Envoltória de Dados**

## 1. Introduction

Recent studies of Brazilian agriculture based on the Agricultural Census of 2006 suggest a positive association between the (income) Gini index and production efficiency. Souza et al. (2015), using the Gini index as the dependent variable, found that the variable returns to scale (VRS) data envelopment analysis (DEA) score (Banker et al., 1984) is highly significant in a nonlinear fractional regression model relating income dispersion to DEA and other covariate determinants of market imperfections. The model was fitted to each of the five Brazilian regions (south, southeast, north, northeast and center-west). The DEA variable was dominant in all regions and the market imperfection variables vary in intensity by region. As Alves and Souza (2015) point out, due to market imperfections, the small farmers sell their products at lower values and buy inputs at higher prices. As a result they are not able to adopt technologies at relatively higher prices. Conditions leading to market imperfections and affecting agricultural production are schooling, access to credit, access to water and electricity, and development level in general. Thus, it is important to identify and quantify the effects of market imperfections in order to guide public policies and to reduce income concentration in the rural areas.

In another regression study relating ranks of income concentration as a linear function of ranks on labor, land and technological input expenses, Alves et al. (2013) found that, for Brazil, all variables were statistically significant and that technological inputs dominated the relationship. They concluded that technology was responsible for the observed income concentration. Indeed, they report that 11% of farms were responsible for 87% of the value of the rural production, according to the 2006 agricultural census. In this article they emphasize the need for further studies quantifying the effect of technology in the presence of market imperfections. Other studies deserve mentioning. Ney and Hoffman (2008) analyzed the contribution of agricultural and non-farm activities to the inequality of rural income distribution in Brazil, observing two evidences: the participation of each sector in the household earnings of different income strata delimited by percentiles, and the decomposition of Gini coefficients. Their results show that agriculture and rural non-farm activities contribute, respectively, to reduce and to increase rural income inequalities in Brazil. Ney and Hoffmann (2009) assessed the effects of rural income determinants, in particular of human capital and physical capital. Ferreira and Souza (2007) analyzed the participation and the contribution of the household income 'retirements and pensions' for the inequality of the distribution of the per capita household income in Brazil and rural Brazil, in the period from 1981 to 2003. Neder and Silva (2004) estimated poverty indexes and income distribution in rural areas based on the National Survey for Household Sampling for the period 1995–2001. They reported a drop in the rural income concentration in some Brazilian states. These facts were not confirmed by the Brazilian 2006 agricultural census, if one restricts attention to rural income.

We contribute to this literature, evaluating production in the proper way, that is, conditioning efficiency on the covariates of importance (market imperfection determinants). Output is the rural income and the inputs are labor, land and capital expenses. The flavor of the analysis is stochastic and involves the conditional efficiency probability models proposed by Daraio and Simar (2007). The measure of efficiency considered is FDH. Our present approach differs from Souza et al. (2015) and Souza and Gomes (2014). In the former, income dispersion is the dependent variable; in the latter, income distribution is not considered in their model. The statistical analysis we carry out is similar to Souza and Gomes (2014) and mimics the approach of Souza and Gomes (2015).

Our discussion proceeds as follows. Section 2 introduces the variables used in the production process. Section 3 and Section 4 deal with methodology. Section 5 shows and discusses the statistical results. Finally, Section 6 summarizes and concludes the article.

## 2. Production Variables and Covariates

The main source for this work is the Brazilian Agricultural Census of 2006 (IBGE, 2012a). Other sources used are listed below on a case-by-case basis. For the inputs and the output we worked with monetary values. The choice of monetary values as opposed to

quantities arose from the fact that using values allows for the aggregation of all agricultural outputs and inputs in the production process.

Farm data were pooled to form averages for each county. A total of 4,965 counties provided valid data for our analysis. This figure represents 89.3% of the total number of Brazilian counties. 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 straightforward and do not require further explanation. They are measured on a farm level, as provided by the census, and aggregated by county.

The contextual variables considered are the Gini index, the proportion of farmers who received technical assistance, total financing per farm, and performance county indexes in the social, environmental, and demographic dimensions. These performance 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 et al. (2013). The idea was also used by the National Confederation of Agriculture (Confederação Nacional da Agricultura, 2013) 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).

Table 1: Description of the production variables

Variable	Components	Unit	Notes
Y (output)	Value of production of cattle, swine, goats, equines, buffaloes, donkeys, mules, sheep, other birds, rabbits, apiculture, sericulture, raniculture, aquaculture, horticulture, flowers, forestry, agro industry, permanent crops, temporary crops, extractive activities	Reais	-
Land	4 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 6 percent a year. Depreciation periods: Machines – 15 years, Planted forests – 20 years, Permanent cultures – 15 years, Improvements – 50 years, and Animals – 5 years

#### Gini index

As a measure of income dispersion we use the county Gini index. This is defined as follows. If  $x_i$  is the rural income of farm  $i$ , the index is  $g/2\bar{x}$ , where  $g = (1/n^2) \sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|$  and  $\bar{x}$  is the mean of the  $x_i$ . The Gini index varies in the interval [0,1) with values close to 1 indicating more intense income concentrations.

### Social dimension

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

The data used in the social dimension were extracted from the Brazilian Demographic Census 2010 (IBGE, 2012b), Brazilian Agricultural Census 2006 (IBGE, 2012a) and from the databases of the National Institute of Research and Educational Studies – INEP, referring to education in 2009 (INEP, 2012) and the Ministry of Health, 2011 data (Ministério da Saúde, 2011).

### Demographic dimension

The variables comprising the demographic dimension capture aspects of the population dynamics, which relates to rural development. These are the 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 (IBGE, 2012b).

### Environmental dimension

The variables comprising the environmental dimension are the proportion of farmers practicing the technique of vegetation fires, which use agrochemicals, practicing crops rotation, minimum tillage, 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 (IBGE, 2012a).

All variables within each dimension are measured in such way so as to correlate positively with the given dimension. They were rank transformed and normalized by the maximum value in the dimension, 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. FDH Unconditional and Conditional Measures of Technical Efficiency**

We begin with the notion of FDH efficiency and related measures. Daraio and Simar (2007) and Bădin et al. (2012) discuss a measure of conditional efficiency closely related to FDH and probability theory, which we use to assess the influence of covariates in the production process. Convexity of the technology set is not required. Consider production observations  $(x_j, y_j)$ ,  $j = 1, \dots, n$ , of  $n$  producing units (DMUs). The input vector  $x_j$  is a vector in  $R^3$  with non-negative components with at least one strictly positive. The output vector  $y_j$  is a non-negative point in  $R$ . The technical efficiency FDH of DMU  $\tau$  is taken relative to the frontier of free disposability (free disposal hull) of the set (1).

$$\Psi = \left\{ (x, y) \in R_+^{3+1}, y \leq \sum_{j=1}^n \gamma_j y_j, x \geq \sum_{j=1}^n \gamma_j x_j, \sum_{j=1}^n \gamma_j = 1, \gamma_j \in \{0, 1\}, j = 1 \dots n \right\} \quad (1)$$

The input-oriented FDH is given by (2) and the output oriented is given by (3).

$$\hat{\theta}(x_\tau, y_\tau) = \text{Min} \left\{ \theta; y_\tau \leq \sum_{j=1}^n \gamma_j y_j, \theta x_\tau \leq \sum_{j=1}^n \gamma_j x_j, \sum_{j=1}^n \gamma_j = 1, \gamma_j \in \{0, 1\} \right\} \quad (2)$$

$$\hat{\lambda}(x_\tau, y_\tau) = \text{Max} \left\{ \lambda; \lambda y_\tau \leq \sum_{j=1}^n \gamma_j y_j, x_\tau \geq \sum_{j=1}^n \gamma_j x_j, \sum_{j=1}^n \gamma_j = 1, \gamma_j \in \{0,1\} \right\} \quad (3)$$

One can show that

$$\hat{\theta}(x_\tau, y_\tau) = \text{Min}_{j=1\dots n} \left\{ \text{Max}_{i=1,2,3} \left\{ \frac{x_j^i}{x_\tau^i} \right\} \right\}, \hat{\lambda}(x_\tau, y_\tau) = \text{Max}_{j=1\dots n} \left\{ \frac{y_j}{y_\tau} \right\} \quad (4)$$

A very interesting interpretation of the FDH arises when the production process is described by a probability measure, defined on the product space  $R_+^{3+1}$  by random variables  $(X, Y)$ . For efficiency purposes, one is interested in the probability of dominance given by (5).

$$H_{XY}(y, x) = P(Y \geq y, X \leq x). \quad (5)$$

Notice that

1.  $H_{XY}(y, x)$  gives the probability that a unit operating at input-output level  $(x, y)$  is dominated, i.e. that another unit produces at least as much output while using no more of any input than the unit operating at  $(x, y)$ ;
2.  $H_{XY}(y, x)$  is monotone, non-decreasing in  $x$  and monotone non-increasing in  $y$ .

The support of  $H_{XY}(x, y)$  is the technology set  $\Omega$  given by (6), where  $\Omega = \{(x, y) : x \text{ can produce } y\}$ .

$$H_{XY}(y, x) = 0 \quad \forall (x, y) \notin \Omega, \quad (6)$$

We then have (7) or (8).

$$H_{XY}(x, y) = P(X \leq x | Y \geq y)P(Y \geq y) = F_{X|Y}(x | y)S_Y(y) \quad (7)$$

$$H_{XY}(x, y) = P(Y \geq y | X \leq x)P(X \leq x) = S_{Y|X}(y | x)F_X(x). \quad (8)$$

New concepts of efficiency measures can be defined for the input-oriented and output-oriented cases, assuming  $S_Y(y) > 0$  and  $F_X(x) > 0$ . For input orientation we have (9) and for output orientation (10).

$$\theta(x, y) = \inf \{ \theta | F_{X|Y}(\theta x | y) > 0 \} = \inf \{ \theta | H_{XY}(\theta x, y) > 0 \} \quad (9)$$

$$\lambda(x, y) = \sup \{ \lambda | S_{Y|X}(\lambda y | x) > 0 \} = \sup \{ \lambda | H_{XY}(\lambda y, x) > 0 \}. \quad (10)$$

Since the support of the joint distribution of  $(X, Y)$  is the technology set, boundaries of  $\Omega$  can be defined in terms of conditional distributions.

In our empirical work we use the output orientation. The Farrell-Debreu stochastic output efficiency measure for given levels of input  $(x)$  and output  $(y)$  is given in (10) and is nonparametrically estimated by (11).

$$\hat{\lambda}_n(x, y) = \sup \left\{ \lambda \mid \hat{S}_{Y|X,n}(\lambda y | x) > 0 \right\} \quad (11)$$

$$\text{where } \hat{S}_{Y|X,n}(y | x) = \frac{\sum_{i=1}^n I(x_i \leq x, y_i \geq y)}{\sum_{i=1}^n I(x_i \leq x)}.$$

It can be shown that (12) coincides with the FDH estimator.

$$\hat{\lambda}_n(x, y) = \text{Max}_{i: x_i \leq x} \left\{ \frac{y_i}{y} \right\} \quad (12)$$

The estimated FDH production set is very sensitive to outliers, and consequently so are the estimated efficiency scores. Here we avoid the outlier problem by transforming the data into ranks before the analysis.

To assess the significance of a continuous contextual variable  $Z$  of dimension  $m$  on the output-oriented efficiency measurement, we condition on  $Z = z$ , to obtain (13),

$$\lambda(x, y | z) = \sup \left\{ \lambda \mid S_Y(\lambda y | x, z) > 0 \right\}, \quad (13)$$

where  $S_Y(y | x, z) = \text{Prob}(Y \geq y | X \leq x, Z = z)$ .

The nonparametric kernel estimate proposed by Daraio and Simar (2007) is defined by (14), where  $K(\cdot)$  is the kernel and  $h_n$  is the bandwidth. The bandwidth selection can be carried out using likelihood cross validation as described in Silverman (1986), but other methods can also be used.

$$\hat{S}_{Y,n}(y | x, z) = \frac{\sum_{i=1}^n I(x_i \leq x, y_i \geq y) K_{h_n}(Z_i - z)}{\sum_{i=1}^n I(x_i \leq x) K_{h_n}(Z_i - z)} \quad (14)$$

For the multivariate case one could use a bounded multivariate kernel or a product of univariate bounded kernels. An easy choice in the latter case would be the product of marginal triangular (Epanechnikov) kernels with support  $[-1, 1]^k$ .

Plugging in the estimator  $\hat{S}_{Y,n}(y | x, z)$  in equation (14), we get the conditional FDH efficiency measure for the output-oriented case as in (15).

$$\hat{\lambda}_n(x, y | z) = \sup \left\{ \lambda \mid \hat{S}_Y(\lambda y | x, z) > 0 \right\}. \quad (15)$$

Therefore, for a multivariate bounded kernel we could use the estimator presented in (16), where  $\tau_n$  is the bandwidth and  $\hat{\sigma}$  is the square root of the average variance of the covariates.

$$\hat{\lambda}_n(x_j, y_j | z_j) = \max_{\{i: x_i \leq x_j, \|z_i - z_j\| \leq \tau_n \hat{\sigma}\}} \left\{ \frac{y_i}{y_j} \right\} \quad (16)$$

With a product of univariate bounded kernels, following Bădin et al. (2012), one could also use the formulation in (17), where  $(\tau_{1n}, \dots, \tau_{mn})$  is the vector of marginal univariate bandwidths of the covariates.

$$\hat{\lambda}(x_j, y_j | z_j) = \max_{\{i: x_i \leq x_j, |z_i^1 - z_j^1| \leq \tau_{1n}, \dots, |z_i^m - z_j^m| \leq \tau_{mn}\}} \left\{ \frac{y_i}{y_j} \right\} \quad (17)$$

Statistical inference will be dependent on how we choose kernels and bandwidths. In our approach we obtained a better fit using a multivariate bounded kernel and a single bandwidth, following Silverman (1986) for the bandwidth choice. The multivariate Epanechnikov kernel is given by (18), where  $u$  is a point in a  $d$ -dimensional space and  $c_d$  is the volume of the unit sphere. In our application  $d=6$  and  $c_6 = (1/6)\pi^3$ .

$$K(u) = \begin{cases} (1/(2c_d))(d+2)(1-u'u) & \text{if } u'u < 1 \\ 0 & \text{otherwise} \end{cases} \quad (18)$$

The optimal window width for the smoothing of normally distributed data with unit variance is given by (19).

$$h_{\text{opt}} = A(K)n^{-1/(d+4)} \quad (19)$$

$$A(K) = \left\{ 8c_d^{-1}(d+4)(2\sqrt{\pi})^d \right\}^{1/(d+4)}$$

If the data is not transformed for unit variances, the indicated procedure is to choose a single-scale parameter  $\hat{\sigma}$  and to use the value  $\hat{\sigma}h_{\text{opt}}$  for the window width. A possible choice is

$$\hat{\sigma}^2 = \frac{1}{d} \sum_{i=1}^d \hat{\sigma}_i^2 \text{ as above, where } (\hat{\sigma}_1^2, \dots, \hat{\sigma}_d^2) \text{ is the vector of variances of the covariates.}$$

#### 4. Statistical Inference and Covariate Assessment

For the assessment of the influence of the covariates on efficiency measurements, Daraio and Simar (2007) suggest a nonparametric statistical analysis using the ratio

$$q_n(x_j, y_j | z_j) = \frac{\hat{\lambda}_n(x_j, y_j | z_j)}{\hat{\lambda}_n(x_j, y_j)} \text{ as the response variable. The underlying suggested statistical}$$

model (Daraio and Simar, 2007) is  $E(q(x_j, y_j | z_j)) = G(\mu_j)$ ,  $\mu_j = z_j'\theta$ , assuming the vector of covariates as exogenous.

A natural parametric version of this model is to consider the flexible fractional regression specification of Ramalho et al. (2010). One may also assume that observations are correlated and that some of the covariates are endogenous. Typical choices for  $G(\cdot)$  are  $G(\mu) = \Phi(\mu)$ ,  $G(\mu) = e^\mu / (1 + e^\mu)$ , and  $G(\mu) = 1 - e^{-e^\mu}$ . Here  $\Phi(\cdot)$  denotes the distribution function of the standard normal. Let  $h$  be a vector of strictly exogenous variables available for use as instruments.

$$\text{If } u_{jn}(\theta, z_j) = \frac{\hat{\lambda}_n(x_j, y_j | z_j)}{\hat{\lambda}_n(x_j, y_j)} - G(z_j'\theta) \text{ and assuming the moment condition}$$

$E\{h \otimes u_{jn}(\theta, z_j)\} = 0$  one may estimate the parameter  $\theta$  by the generalized method of moments – GMM (Gallant, 1987; Greene, 2011). Standard errors for the estimators can be computed controlling for heteroskedasticity and correlation among DMUs. See Stata (2013). An endogeneity and goodness-of-fit test is performed testing for overidentifying restrictions.

## 5. Statistical Findings

Table 2 shows GMM results for the fitting of the model defined by  $G(\mu) = \Phi(\mu)$ . The other parametric alternatives for responses in (0,1) give similar results. Marginal effects are computed as  $\frac{dG(\mu)}{d\mu} \theta_j | \mu = z' \theta$ ,  $\frac{dG(\mu)}{d\mu} = \phi(\mu)$ , where  $\phi(\cdot)$  is the standard normal density function. The Hansen specifications test statistics with 2 degrees of freedom is 3.921 (p-value=0.141).

The interpretation of the sign of a covariate ( $Z$ ) coefficient is taken from Daraio and Simar (2007, p.115): “an increasing regression corresponds to a favorable environmental factor and a decreasing regression indicates an unfavorable factor.” We see from Table 2 that the financing, social and demographic factors and the Gini index are favorable. These variables show the correct signs. The environmental variable is not significant. The influence of the Gini index points to a strong association with technical efficiency, confirming the results of previous studies reporting that technology is responsible for rural income concentration in the country. We see the negative sign of technical assistance as an indication that rural extension in the country does not lead to higher efficiency levels, very probably due to market imperfections that inhibit the use of proper technologies to increase rural productivity.

Table 2: GMM estimation when expected response is defined by the standard normal distribution function

Variable	Coefficient	Robust standard error	z	P> z	95% Confidence interval	
Constant	-1.7826	0.5587	-3.19	0.0010	-2.8776	-0.6877
Financing	2.1058	0.4424	4.76	0.0000	1.2388	2.9728
Environmental	-0.0227	0.2473	-0.09	0.9270	-0.5073	0.4619
Social	1.8832	0.5104	3.69	0.0000	0.8829	2.8835
Demographic	0.8030	0.2996	2.68	0.0070	0.2159	1.3902
Tech. Assist.	-3.2529	0.8673	-3.75	0.0000	-4.9528	-1.5529
Gini index	2.4775	0.8132	3.05	0.0020	0.8837	4.0713

Note: Exogenous covariates are environment, social and demographic indexes. Endogenous covariates are financing, technical assistance, social index and income dispersion. Instruments are regional dummies, and the environmental index, power supply, demographic index, and demographic<sup>2</sup>. Standard errors are adjusted for correlation and heteroskedasticity.

The intensity of the marginal effects on the response may be appreciated by the median values of the factor  $\phi(\mu)$ . Table 3 shows 5-number summaries for each region and the corresponding box plots are shown in Figure 1. The region that will benefit the most from public policies reducing market imperfections is the northeast, closely followed by the north region.

Table 3: 5-number summary statistics for the regional marginal effects

Region	Minimum	25% Percentile	Median	75% Percentile	Maximum	Number of observations
North	0.0028	0.1643	0.2765	0.3613	0.3989	405
Northeast	0.0031	0.2107	0.2908	0.3580	0.3989	1666
Southeast	0.0027	0.1156	0.2100	0.3114	0.3989	1531
South	0.0015	0.1628	0.2656	0.3501	0.3989	1143
Center-west	0.0024	0.0598	0.1218	0.1988	0.3984	220

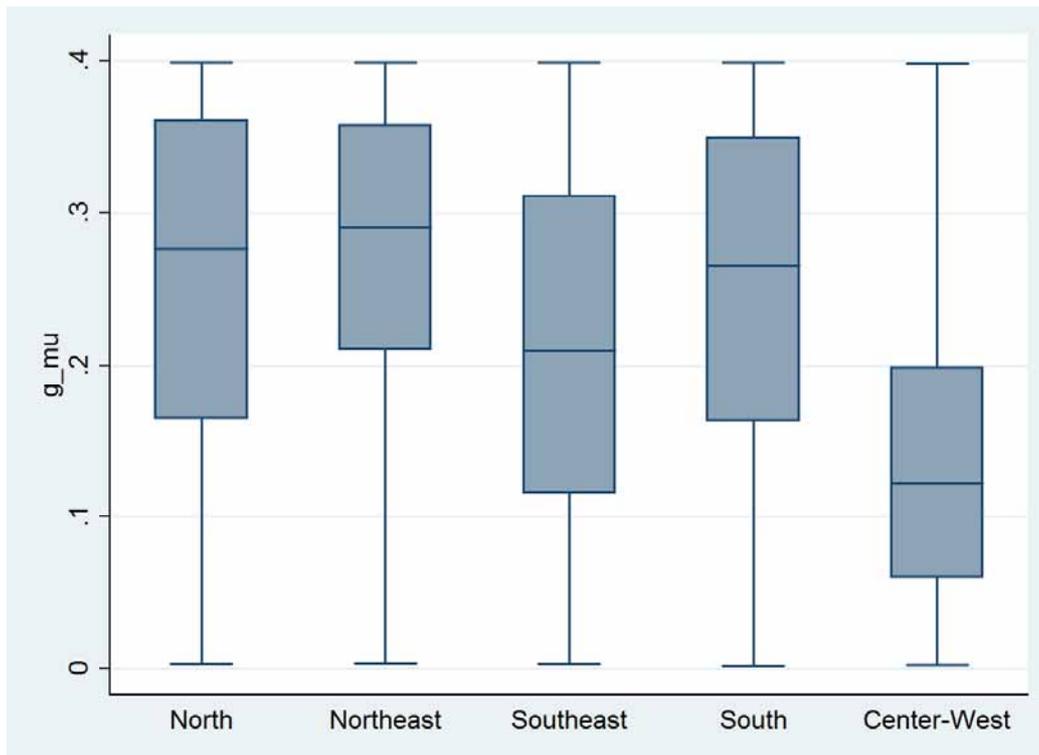


Figure 1: Regional marginal effects box plots.

## 6. Summary and Conclusions

Using the Brazilian 2006 agricultural census data and other official sources we examined the influence of market imperfection variables and income dispersion on the efficiency of rural production in Brazil at a county level. As a measure of technical efficiency we took the FDH performance score. The analysis is based on the conditional frontier assuming a stochastic data-generating process for the frontier. The statistical analysis is carried out using fractional regression and GMM, and postulates a nonlinear relationship between the ratio of the conditional to unconditional FDH frontiers and the set of conditioning variables.

We conclude that market imperfections are unfavorable for production, creating difficulties for small farms in adopting technological intensive inputs, and that technology is strongly associated with efficiency. This observation responds to the fact that efficient counties show a high degree of income concentration. The regions in Brazil likely to be most affected by issues of income concentration and market imperfections are the north and northeast, and thus public policies should mainly be guided to these regions.

## References

- Alves, E., Souza, G.S. (2015), Pequenos estabelecimento em termos de área também enriquecem? Pedras e troços, *Revista de Política Agrícola* (forthcoming).
- Alves, E., Souza, G.S., Rocha, D.P. (2013), Desigualdade nos campos sob a ótica do censo agropecuário 2006, *Revista de Política Agrícola*, 22, 67–75.
- Bădin, L., Daraio, C., Simar, L. (2012), How to Measure the Impact of Environmental Factors in a Nonparametric Production, *European Journal of Operations Research*, 223:818–833.
- Banker, R.D., Charnes, A., Cooper, W.W. (1984), Some models for estimating technical scale inefficiencies in Data Envelopment Analysis, *Management Science*, 30(9), 1078–1092.
- Confederação Nacional da Agricultura. *Índice de Desenvolvimento Rural CNA*, 2013 (documento não publicado).
- Daraio, C., Simar, L. *Advanced Robust and Nonparametric Methods in Efficiency Analysis*, Springer, New York, 2007.
- Embrapa. *Balanco Social da Pesquisa Agropecuária Brasileira*, Brasília, Secretaria de Comunicação/Secretaria de Gestão Estratégica, 2001.

- Ferreira, C.R., Souza, S.C.I.** (2007), As aposentadorias e pensões e a concentração dos rendimentos domiciliares per capita no Brasil e na sua área rural: 1981 a 2003, *Revista de Economia e Sociologia Rural*, 45(4), 985–1011.
- Gallant, A.R.** *Nonlinear Statistical Models*, Wiley, New York, 1987.
- Greene, W.H.** *Econometric Analysis*, 7th ed., Prentice Hall, New Jersey, 2011.
- IBGE.** *Censo Agropecuário 2006.* 2012a. Disponível em: <http://www.ibge.gov.br/home/estatistica/economia/agropecuaria/censoagro/>. Acesso em: 24 jan. 2012.
- IBGE.** *Censo Demográfico 2010.* 2012b. Disponível em: <http://censo2010.ibge.gov.br/>. Acesso em: 24 jan. 2012.
- INEP.** *Nota Técnica do Índice de Desenvolvimento da Educação Básica.* 2012. Disponível em: <http://ideb.inep.gov.br/resultado/>. Acesso em: 24 jan. 2012.
- Ministério da Saúde.** *IDSUS - Índice de Desempenho do SUS.* Ano 1, 2011. Disponível em: <http://portal.saude.gov.br/>. Acesso em: 02 mar. 2012.
- 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., Souza, G.S.** (2004), Uma metodologia alternativa para mensuração de pressão sobre o mercado de câmbio, *Estudos Econômicos*, 34:73–99.
- Neder, H.D., Silva, J.L.M.** (2004), Pobreza e distribuição de renda em áreas rurais: uma abordagem de inferência, *Revista de Economia e Sociologia Rural*, 42(3), 469–486.
- Ney, M.G., Hoffmann, R.** (2008), A contribuição das atividades agrícolas e não-agrícolas para a desigualdade de renda no Brasil rural, *Economia Aplicada*, 12(3), 365–393.
- Ney, M.G., Hoffmann, R.** (2009), Educação, concentração fundiária e desigualdade de rendimentos no meio rural brasileiro, *Revista de Economia e Sociologia Rural*, 47(1), 147–181.
- Ramalho, E.A., Ramalho, J.J.S., Henriques, P.D.** (2010), Fractional regression models for second stage DEA efficiency analyses, *Journal of Productivity Analysis*, 34:239–255.
- 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.
- Silverman, B.W.** *Density Estimation*, Chapman and Hall, New York, 1986.
- Souza, G.S., Alves, E., Gomes, E.G.** (2015). Determinantes da dispersão da renda no meio rural (documento não publicado).
- Souza, G.S., Gomes, E.G.** (2014), Analyzing the influence of contextual variables in the Brazilian agricultural production by means of conditional FDH and fractional regression models, *Program and Abstracts – XVII Congresso Latino-Iberoamericano de Investigación Operativa*, p. 40.
- Souza, G.S., Gomes, E.G.** (2015), Management of agricultural research centers in Brazil: a DEA application using a dynamic GMM approach, *European Journal of Operational Research*, 240:81–824.
- 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*, 34:1–15.
- Stata.** *Base Reference Manual*, Stata Press, Texas, 2013.