Heterogeneity in Agricultural Factor Productivity Across and Within Farm Size Groups in Brazil

Pedro Gabriel Meiners* — Universidade Federal do Paraná, Departamento de Economia

Marcelo de Oliveira Torres† — Universidade de Brasília, Departamento de Economia

Abstract

A key issue in evaluating agricultural factor productivity in Brazil is to account for the existing wide variation in farm sizes with respect to planted area. Econometric studies have dealt with it by using flexible production function models from which the resulting factor productivity measures can be fitted at different data values ranging from small to large farms. In some studies farms of different sizes are pooled together and parameter estimates are completely invariant to size. While others use dummies or perform separate regressions by land size group, then allowing for the parameter estimates to differ accordingly to size. In this study, we use a quantile regression approach to show that in fact there exists significant variation across and within groups in factor productivity not only because farms of different sizes use different input proportions but because there are inter- and intra-group farm-specific effects. Our results suggest therefore that the resulting average measures of factor productivity of previous studies may then be poorly describing factor productivity in the Brazilian agriculture likely leading to confounding conclusions.

***Keywords*—** Agriculture, Production, Quantile Regression, Brazil, Heterogeneity

Resumo

Um importante aspecto para a análise quantitativa da produtividade dos fatores empregados na agricultura brasileira é a extrema variação no tamanho das propriedades. Estudos econométricos mais recentes tratam dessa questão estimando funções de produção flexíveis, nas quais as medidas de produtividade dos fatores construídas com os parâmetros estimados são avaliadas em diferentes valores médios variando de pequenos a grandes produtores. Em alguns estudos, os parâmetros estimados são invariantes quanto ao tamanho da propriedade. Enquanto outros, usam *dummies* ou realizam regressões separadas por grupos de tamanho da área plantada (geralmente predeterminados pela disponibilidade de dados do IBGE), assim permitindo que os parâmetros estimados difiram de acordo com os grupos. Além de confirmar a existência de significativa variação na produtividade dos fatores entre os grupos, este estudo demonstra, usando regressões quantílicas, que os parâmetros estimados diferem por estrato de tamanho também dentro de cada grupo. Nossos resultados portanto sugerem que há significativa heterogeneidade tecnológica dentro dos grupos e que estudos que assumem homogeneidade intra-grupo podem estar sub- ou superestimando a real contribuição de cada fator para a agricultura brasileira.

***Palavras-chave*—** Agricultura, Produção, Regressão Quantílica, Brasil, Heterogeneidade
JEL code: Q12, C21, D24
Classificação ANPEC: Área 11- Economia Agrícola e Meio Ambiente

*pg.meiners@gmail.com
†Corresponding author, motorres@hotmail.com.br

1 Introduction

Agricultural productivity plays a key role in developing economies, as it is connected to food security, income and poverty eradication. In Brazil, it is no different, but the importance of agriculture sector goes beyond. Brazil is one of the largest exporters of agricultural products in the world and the sector as whole accounts for more than 30% of all Brazilian exports (Comex Stat). The sector's performance is therefore an essential element in the sustenance of a positive balance of payments every year and has helped the country to more quickly overcome economic downturns (Brazilian Central Bank).

It should not be surprising that researchers have spent a great deal of time studying the Brazilian agriculture performance. In fact, the productivity of Brazilian farms has been the focus of several studies lately. Some examples of parametric approaches are Moreira et al. (2007), Rada et al. (2018) and Helfand et al. (2017) that focus on the relationship between farm size and the total factor productivity (TFP) growth; Bragagnolo et al. (2010) and Rada and Valdes (2012) that investigate technical efficiency patterns across farmers at state level using stochastic frontier methods; and Helfand et al. (2011) that examines the causality between agricultural performance and poverty. An example of non-parametric approach is Gasques et al. (2010, 2014) which calculates TFP growth at state level using a Torniqvist index, or Avila et al. (2015) which measures the impacts of the Brazilian Company of Agricultural Research (EMBRAPA) along its fifteen years of existence using a multidimensional approach. A more comprehensive review on agricultural studies focused on Brazil can be found in Machado et al. (2020).

While this existing literature has made important contributions in disentangling the key production factors for the Brazilian agricultural performance, gaps remain, especially those related to the modeling of the wide variation in farm sizes with respect to planted area that exists in Brazil. In particular, the econometric studies have dealt with them by using flexible production function models from which the resulting factor productivity measures can be fitted at different data values ranging from small to large farms. In some studies farms of different sizes are pooled together and parameter estimates that enter into the productivity measures are completely invariant to size. It is then implicitly assumed that there are no farm-specific effects related to size and factor productivity measures differ across farms of different sizes only because small, medium and large farms may use different input proportions. Other studies include dummies by size or perform separate regressions by land size group (LASG), allowing for parameter estimates as well as input proportions to differ accordingly to farm size. It is then implicitly assumed that there may be across- but not within-group farm-specific effects.

By building upon Rada et al.'s modeling approach and using data from the 2006 agricultural census provided by the Brazilian Institute of Geography and Statistics (IBGE), we develop an empirical production model based on a fully flexible functional form and apply quantile regression techniques to look for statistically significance in the parameter estimates across nine percentiles of the data within LASG (0-5; 5-20; 20-100, 100-500; and >500ha). Using a bootstrap tool we calculate the output-input elasticities and investigate the differences across and within any LASG.

Results show the presence of significant heterogeneity in factor productivity across and within all LASGs. They indicate that modeling approaches that pool farms of different sizes together will yield biased estimators for factor productivity elasticities. Also, as factor productivity vary widely within LASGs, these biases do not disappear in studies that accommodate for size heterogeneity by including dummies of by performing separate regressions according to size.

The remainder of this article is organized as follows. Section 2 provides explanations on the empirical models. Section 3 discusses the data and variable construction, also display descriptive statistics and testing. Section 4 presents results, and Section 5 concludes.

2 Empirical Model

A comprehensive historical review of the empirical methods used in the investigation of agricultural productivity and efficiency up to the first decade of the 2000's is provided by Darku et al. (2013). In it its is shown that the main studies date back to the 1950's but only after the 1980's that the modern era of applied studies on agricultural productivity measurement picked up. In this period studies such as Bagi (1982), Kawagoe and Hayami (1985), Bravo-Ureta (1986), Aly et al. (1987), Tauer and Belbase (1987) and Kumbhakar et al. (1989), shifted the focus to more complex models based on Stochastic Frontier Analysis (SFA) and Data Enveloping

Analysis-DEA (DEA). Results tended to point out that larger farms are more efficient than smaller ones as they adopt new technology faster due to better access to credit, information and markets. A similar trend is verified when agriculture in developed countries is compared to developing ones.

During the 1990's SFA and DEA became standard approaches to the study of technical and allocative efficiency across regions and farms of different sizes (*e.g.*: Bravo-Ureta and Rieger (1990); Dawson and Woodford (1991); Kumbhakar et al. (1991); Kumbhakar and Heshmati (1995); Hallam and Machado (1996); Thiele and Brodersen (1999)). The results showed that farmers were becoming more efficient, but not necessarily because their areas under crop were getting larger. Factors such as education and age were getting more important, with a positive and a negative impact on productivity respectively. During the years 2000 studies started to focus on comparing different types of agricultural production (organic vs. conventional farms), on determining the effect of governmental policies on efficiency and on the comparison of different econometric methodologies (*e.g.*: Lansink et al. (2002); Paul et al. (2004); Tipi and Rehber (2006); Bravo-Ureta et al. (2007); Serra et al. (2008); Hassine-Belghith (2009)).

In this last decade, methodological advances have allowed researchers to measure the effects of agricultural productivity growth on poverty alleviation and on economic growth (*e.g.*:Schneider and Gugerty (2011); Gollin (2010); Cao and Birchenall (2013)) and to quantify the gaps on productivity performance between countries with a focus on the factors that may narrow them such as knowledge and education (*e.g.*:Gollin et al. (2011); Block (2014); Bustos et al. (2016); Davis et al. (2012)). Last but not least, database construction efforts turned longer historical series of data on climate variables readily available which allowed for the measurement of the impacts of climate change on agricultural productivity as in Gornall et al. (2010).

As already briefly mentioned in Section 1, our methodology fits in this most recent set of the international literature body. More specifically, as in Rada et al., we also assume that the agricultural output-input relationships follow a translog functional form, but differently we do not impose *a priori* constant returns to scale. Assuming Hicks neutral technology the basic empirical production function may be then represented by the following stochastic model:

$$\begin{aligned} \ln Y_i = \alpha_0 &+ \sum_{k=1}^{4} \beta_k \ln X_{ki} + \frac{1}{2}\sum_{k=1}^{4}\sum_{h=1}^{4} \gamma_{kh} \ln X_{ki} \ln X_{hi} + \sum_{w=1}^{2} \hat{\omega}_w \ln W_{wi} \\ &+ \frac{1}{2}\sum_{w=1}^{2}\sum_{z=1}^{2} \hat{\zeta}_{wz} \ln W_{wi} \ln W_{zi} + \sum_{w=1}^{2}\sum_{k=1}^{4} \hat{\mu}_{wk} \ln W_{wi} \ln X_{ki} + u_i \end{aligned} \tag{1}$$

Where X_{ki} is a vector for the independent variables (capital stock, land, labor and purchased inputs) for any i municipality and W_{wi} a vector for the climate control variables (average precipitation and average temperature in municipality i in the base year). The Greek letters α, β, γ, ω, ζ and μ are parameters to be estimated and u_i is a i.i.d. error. At first we estimate the equation (1) by ordinary least squares (OLS) for all municipalities and independently of farm sizes. We then group all farms in each LASG (0–5 ha, 5–20 ha, 20–100 ha, 100–500 ha, and 500- ha) and run OLS regressions for each LASG. In this case the model becomes

$$\begin{aligned} \ln Y_{ci} = \alpha_0 &+ \sum_{k=1}^{4} \beta_{ck} \ln X_{cki} + \frac{1}{2}\sum_{k=1}^{4}\sum_{h=1}^{4} \gamma_{ckh} \ln X_{cki} \ln X_{chi} + \sum_{w=1}^{2} \omega_w \ln W_{wi} \\ &+ \frac{1}{2}\sum_{w=1}^{2}\sum_{z=1}^{2} \zeta_{wz} \ln W_{wi} \ln W_{zi} + \sum_{w=1}^{2}\sum_{k=1}^{4} \mu_{cwk} \ln W_{wi} \ln X_{cki} + u_{ci} \end{aligned} \tag{2}$$

In eq.(2), the vector X_{cki} contains the same number of independent variables as in eq 1, but now it is also indexed to each LASG c. Y_{ci} refers to the sum of the value of agricultural production across all farmers in municipality i that fall into LASG c. For example, for the 0-5 ha LASG, we perform an OLS regression of Y_{0-5i} on the amount of labor ($x_{labor0-5i}$), capital stock ($x_{capitalstock0-5i}$), land ($x_{land0-5i}$) and purchased inputs ($x_{PurImp0-5i}$), controlled by the climate conditions as in eq. (1). Observe that the Greek letters β, γ, and μ are also indexed by LASG c.

Given the estimates obtained from running the models represented by eqs. (1) and eq.(2) we calculate elasticities of scale and of the output value with respect to land, capital, labor and purchased inputs by eq. (3).

Using those estimates, the elasticity of output Y with respect to input k is defined as eq. (4).

$$\hat{\psi}_{X_{ki}} = \frac{\partial \ln Y_i}{\partial \ln X_{ki}} = \hat{\beta}_k + \hat{\gamma}_{kk} \ln X_{ki} + \sum_{h=1}^{3} \hat{\gamma}_{jh} \ln X_{hi} + \sum_{w=1}^{2} \hat{\mu}_{wk} \ln W_{wi} \tag{3}$$

$$E[\hat{\psi}_{X_k}] = \frac{\sum_{i=1}^{n} \hat{\psi}_{X_{ki}}}{n} \tag{4}$$

Output elasticities for each input and LASG c is calculate using eq. (5) bellow and the mean obtained by eq. (6).

$$\hat{\psi}_{X_{cki}} = \frac{\partial \ln Y_{ci}}{\partial \ln X_{cki}} = \hat{\beta}_{ck} + \hat{\gamma}_{ckk} \ln X_{cki} + \sum_{h=1}^{3} \hat{\gamma}_{cjh} \ln X_{chi} + \sum_{w=1}^{2} \hat{\mu}_{cwk} \ln W_{wi} \tag{5}$$

$$E[\hat{\psi}_{X_{ck}}] = \frac{\sum_{i=1}^{n} \hat{\psi}_{X_{cki}}}{n} \tag{6}$$

Estimates of scale elasticities are also computed by using eqs. (3) and (5). With the former, that uses the aggregate data set (not separated by LASGs), the scale elasticity is defined as the mean of the sum of $\hat{\psi}_{X_{ki}}$ over all k inputs. With the latter the measure is is defined as the mean of the sum of $\hat{\psi}_{X_{cki}}$ over all k inputs for a given LASG c. Notice that the resulting elasticities are a combination of parameter estimates and data on input quantities and the climate variables. Therefore the elasticity values may vary due to differences in the values of the independent variables and in the estimated input marginal effects on output represented by the coefficient estimates. Inference, as suggested in Krinsky and Robb (1986, 1991), is performed through the use of a non-parametric bootstrap technique with 10,000 replications that allows us to calculate basic confidence intervals of 99%, also called Non-Studentized pivotal method (Carpenter and Bithell, 2000).

These OLS regressions, elasticity estimates and the inference methods allow us to investigate whether there is statistical heterogeneity between LASG's and between any LASG and the pooled model (represented by eq. (1)). Too be able however to address and verify the heterogeneity within each LASG we propose the use of the quantile regression (QR) technique (Koenker and Bassett (1978)).

The conditional quantile function (7) denotes a relationship between a quantile of the density distribution of the dependent variable lnY and the covariate vector $lnX^{\top}$. In here, $X^{\top}$ contains the inputs defined above as well as the climate variables and $\beta(\theta)$ contains their associated parameters in each quantile (θ).

$$Q_{\ln Y}(\theta|X) = \beta(\theta) \ln X^{\top} \tag{7}$$

The estimate for $\hat{\beta}_c(\theta)$ in a given LASG c can be obtained for any $\theta \in [0, 1]$ by the solution of the following minimization problem:

$$\min_{\beta_c \in \mathbb{R}} [\sum_{(i| \ln Y_{ci} > \ln X_{ci}^{\top} \beta_c)} \theta | \ln Y_{ci} - \ln X_{ci}^{\top} \beta_c| + \sum_{(i| \ln Y_{ci} < \ln X_i^{\top} \beta_c)} (1-\theta) | \ln Y_{ci} - \ln X_{ci}^{\top} \beta_c|] \tag{8}$$

If we apply this technique to our production function (1) the conditional quantile function can be represented by (9), where $F_u^{-1}(\theta)$ is the quantile of the error term distribution.

$$\begin{aligned} Q_{\ln Y}(\theta|X) &= \hat{\alpha}_0(\theta) + \sum_{k=1}^{4} \hat{\beta}_{ck}(\theta) \ln X_{cki} + \frac{1}{2} \sum_{k=1}^{4} \sum_{h=1}^{4} \hat{\gamma}_{ckh}(\theta) \ln X_{cki} \ln X_{chi} \\ &+ \sum_{w=1}^{2} \hat{\omega}_w(\theta) \ln W_{wi} + \frac{1}{2} \sum_{w=1}^{2} \sum_{z=1}^{2} \hat{\zeta}_{wz}(\theta) \ln W_{wi} \ln W_{zi} + \sum_{w=1}^{2} \sum_{k=1}^{4} \hat{\mu}_{cwk}(\theta) \ln W_{wi} \ln X_{cki} \\ &+ F_u^{-1}(\theta) \end{aligned} \tag{9}$$

Quantile input elasticities are then calculated for each quantile within a chosen LASG, using equation (10).The mean elasticities are computed as for eq. (11) and we employ the bootstrap process to attain converged mean values and confidence intervals, in each quantile, for all LASG's.

$$\hat{\psi}_{X_{cki}}(\theta) = \frac{\partial Q_{\ln Y}(\theta|X)}{\partial \ln X_{cki}} = \hat{\beta}_{ck}(\theta) + \hat{\gamma}_{ckk}(\theta) \ln X_{cki} + \sum_{h=1}^{5} \hat{\gamma}_{cjh}(\theta) \ln X_{chi} \tag{10}$$

$$E[\hat{\psi}_{X_{ck}}(\theta)] = \frac{\sum_{i=1}^{n} \hat{\psi}_{X_{cki}}(\theta)}{n} \tag{11}$$

By varying θ between (0, 1) in eq. (9) we obtain multiple elasticities for any quantile of output value conditional to the explanatory variables. Using this algorithm we calculate nine different values for each elasticity (from 10% to 90%), studied within each size-class proposed and use this results to analyse if there is heterogeneity present inside each LASG.

3 Data

3.1 Variable Construction

In order to represent the output and input variables we use municipal level data from the 2006 agricultural census provided by the Brazilian Institute of Geography and Statistics (IBGE). We consider all farmers in a given municipality of a specific LASG to be a "producer" that operates a "representative farm". All output and input numbers described below refer to the total output produced and input used along the whole period of reference of the 2006 IBGE agricultural census which is 01/31/2006 to 12/31/2006.

In eq. (1), the dependent variable $\ln Y_i$ is constructed as the natural logarithm of the total value of agricultural production, measured in the Brazilian currency, in municipality i. The independent variables are constructed for the four inputs k (capital stock, land, labor and purchased inputs) and the climate variables w. More specifically, $\ln X_{ki}$ is the natural logarithm of input k quantity and $\ln W_{wi}$ is the natural logarithm of climate variable w in municipality i. In equations (2) and (3) the dependent variable $\ln Y_{ci}$ also represents the natural logarithm of the total agricultural production monetary value, but aggregated only over farms with total land area belonging to the c LASG. $\ln X_{cki}$ is the natural logarithm of the input k quantity in the municipality i, also aggregated over farms with total land area belonging to LASG c. $\ln W_{wi}$ is defined as above. The proper construction of all these variables is discussed below.

For land we use total planted area (hectares) dedicated to perennial and annual crops in the period of reference. We consider the number of hectares in pasture to be implicitly represented by the number of cattle animals which is computed as part of the capital stock variable described in more detail below. The labor variable is the total number of hired and adult family workers.

Purchased inputs include fertilizers, pesticides, fuel, electricity among others and are only available as expenses in the 2006 IBGE Census. We are aware that the proper modeling of this variable in a production function context requires its transformation in a quantity index with the help of a price index based on the prices paid in the period of reference. Due to time constraints the construction of such an index was not feasible and purchased inputs are therefore computed as a monetary value rather than a quantity. For this reason we consider purchased inputs as a control variable in our production model.

To create a index for capital stock we used data for the quantity of animals, machinery and trees. As proposed by Hayami et al. (1971) animals are considered as a form of internal capital accumulation implying the existence of an infrastructure put in place to raising, feeding, breeding, slaughtering, sheltering and harvesting which then justifies considering the number of animals as part of capital stock. We start by assuming that a producer will slaughter an animal when the market price of its flash S_b is at least equal to the benefits of keeping it alive LV_b, i.e. the future gains of fattening or with any production by-products (milk, eggs etc) AY_b, minus the maintenance costs with species b, D_b. In equilibrium therefore we expect the following equivalence to hold:

$$SV_b \equiv AY_b - D_b \quad \rightarrow \quad SV_b \equiv LV_b \tag{12}$$

Assuming that the stock of animals in all farms was indeed at equilibrium level in 31/12/2006, we can then estimate the total value of living animals using the value of slaughtered animals, slaughtered numbers and number of animals in stock as the following.

$$LV_{cbi} = L_{cbi} * \frac{SV_{cbi}}{S_{cbi}} \quad or \quad LV_{cbi} = L_{cbi} * \frac{\sum_i SV_{cbi}}{\sum_i S_{cbi}} \tag{13}$$

Where LV_{cbi} is the total value of live animals of b species, inside i municipality, within each LASG c; SV_{cbi} is the value of slaughtered animals, L_{cbi} is the stock of living animals at the end of 2006, S_{cbi} is the number of slaughtered animal during 2006. The second equation is used when no animal of a particular specie was slaughtered in that year and municipality. We calculate this sub-index for three main species of farm animals: hog, chicken and cattle.

The index for machinery is constructed as in Moreira et al. (2007). The data in the census are in number of tractors of 100 hp or more, tractors with less than 100 hp, trucks, pick-up trucks, planters, harvesters and other agricultural machinery. To aggregate for all these types of machinery we use price data available at a monthly basis from the Institute of Agricultural Economics of São Paulo (IEA). These price data were only available for the year 2018 in a monthly basis and by using them to construct an index for 2006 we are in fact assuming that the price proportions were kept constant in time across municipalities and LASG's. All quantities of machinery were multiplied by its 2018 mean price and then divided by the price of a 'tractor of 100 hp or more' for being subsequently added together. As so, this variable yields the total value of machinery in units of 'tractors with 100hp or more'.

The existence of perennial crops implies a past investment incurred by the producer looking for capitalizing future gains. Using the method developed by Butzer et al. (2012), investments in a given orchard in the long term equilibrium equals the present value of the expected future income it generates. That is,

$$I = PV[E(Y_p - C_p)]. \tag{14}$$

Where Y_p is the total lifetime expected yield of the tree, and C_p the total lifetime costs p. As we only have data on production we assume, in accordance to Rada et. al. (Ibid.) that production costs account for 65% of gross revenues for Brazilian farms[1]. We further assume that all trees are in the middle of their productive lifetime ℓ, that interest rates are constant at 6%, and that the production is equal in all years. This approach enables us to construct sub-indices to 17 different perennial crops for the year 2006[2]. Using the production data and equation (14) we then arrive at the following equation:

$$PV_{cpi} = \sum_{t=1}^{\ell/2} \frac{0,35 * Y_{cpi}}{(1+0,06)^t}. \tag{15}$$

At this point, we have three different sub-indices for each type of capital stock: number of tractors with 100 hp or more, expected value of animal stock if slaughtered and the present value of perennial crops expected profits. To aggregate them we use standardized regression coefficients as weights (beta coefficients method) at the regional level. These weights are normalized to sum to unit before being applied to the data.

The regional estimated normalized weights for machinery, animals and trees, are respectively: Center-West (1.0367; 0.0051; -0.0418); Northeast (0.4893; 0.0747; 0.436); North (0.247; 0.237; 0.516); Southeast (0.518; 0.198; 0.284); South (0.518; 0.198; 0.284); and for Brazil as whole (0.552; 0.252; 0.196). Notice that the weights for the Center-West region are not well behaved as it is above 1 for machinery and negative for trees. They should all be positive and below 1. To overcome this problem we have used the weights for Brazil for the Center-West region.

Lastly the climate variables (average annual temperature measured in Celsius and total annual precipitation in millimeters) for each municipality in the year of 2007 were compiled from Rocha and Soares (2015).

3.2 Descriptive Statistics

A initial analysis of the census data provided in Table 1 shows the distribution of the input use across the multiple LASGs in the Brazilian agriculture. It becomes clear the strong heterogeneity in input use across farms of different sizes. The largest farms represent only 2% of all farms in Brazil but concentrate 39% of agricultural

[1] Butzer et al. uses the figure of 80%, we choose to comply with the Brazilian specific study value of 65%

[2] The productive life expectancy ℓ of each of the plant species follows approximations available in Butzer et al. appendix

area and produce 31% of all output, while these percentages for the smallest ones (0-5 ha) are 37%, 3% and 7%, respectively. As regards of labor force, 56% of all the employed workers belong to farms up to 20 ha Farms. The largest farms over 100 ha employ more than half of all purchased inputs and capital used in the Brazilian agriculture.

Table 1: Production and input use distribution by LASG

LASG (ha)	Number of Farms	Land Area (ha)	Labor Force (persons)	Purchased Inputs (R$1000)	Stock Capital (indexed units)	Total Production (R$1000)
0–5	1,735,615	1,917,812	4,659,860	2,788,405	2,672,774	10,501,654
	(37%)	(3%)	(31%)	(3%)	(7%)	(7%)
5–20	1,298,667	5,103,019	3,823,178	7,376,760	6,472,814	21,576,151
	(27%)	(10%)	(25%)	(10%)	(17%)	(15%)
20–100	1,196,948	10,804,917	3,985,668	12,480,511	10,030,734	34,180,602
	(25%)	(21%)	(26%)	(17%)	(27%)	(24%)
100–500	358,908	12,797,311	1,621,092	16,084,866	8,832,156	30,583,036
	(7%)	(25%)	(10%)	(22%)	(23%)	(21%)
>500	90,295	20,330,326	859,363	31,761,024	8,852,299	44,247,807
	(2%)	(39%)	(5%)	(45%)	(24%)	(31%)
Brazil	4,680,433	50,953,385	14,949,161	70,491,565	36,860,777	141,089,251

Variable specifications are in accordance to section's 3.1 designs. Land is total planted area in hectares. Numbers in parenthesis are shares of total Brazilian values. *Source: IBGE's 2006 Agricultural Census*.

Table 2 shows the input figures of Table 1 at a per hectare basis. As expected we can see that the small farm agriculture (0-5 ha) rely much more on labor relatively to all other LASG's with an average employment of 2.4 persons per hectare, while in the other LASG's this rate varies from 0.042 to 0.749. Surprisingly, however, Table 2 also shows that small farm agriculture (up to 20ha) also employ more capital and produce more output value at a per hectare basis. Expenditures on purchased inputs per hectare of land are relatively homogeneous across all LASG's varying from 1.2 to 1.6 units of purchased input per hectare.

Table 2: Input Use Relatively to Land Area by LASG

LASG (ha)	Labor Force (persons)	Purchased Inputs (R$1000)	Stock Capital (idexed units)	Total Production (R$1000)
0–5	2.429	1.453	1.393	5.475
5–20	0.749	1.445	1.265	4.228
20–100	0.368	1.155	0.928	3.163
100–500	0.126	1.256	0.690	2.389
>500	0.042	1.562	0.435	2.176

Variable specifications are in accordance to section's 3.1 designs. Land is total planted area in hectares. *Source: IBGE's 2006 Agricultural Census*.

At the municipality level, Table 3 reveals that on average municipalities allocated, in the reference year, 441ha, 1081ha, 2235ha, 2839ha and 8629ha to be respectively operated by small (0-5ha), medium-small (5-20ha), medium (20-100), medium-large (100-500ha) and large farms (>500ha). Also, as the farm size increases, the average number of employed labor in a representative municipality decreases while the average expenditure on purchased inputs, employed capital and output value increases. By comparing the number of standard deviations over the mean as an indicator of dispersion (SD/Mean), we see a relatively low dispersion within groups in the land use pattern, apart from the largest farm group with a indicator of 3.2 (27587/8629). The labor use pattern is more homogeneous within groups with the indicator ranging form 1.17 to 1.62. A much

Table 3: Descriptive Data by LASG

LASG (ha)	Land Area (ha)	Labor Force (persons)	Purchased Inputs (R$1000)	Stock Capital (indexed units)	Total Production (R$1000)
0–5					
Max	825	24878	28087	147484	235169
Min	0.03	10	0.12	0.012	0.27
Mean	440.97	1071.47	641.16	614.57	2414.72
SD	710.52	1777.32	1215.32	3830.64	6451.38
5–20					
Max	27240	13833	353071	351304	291113
Min	1	11	0.07	0.05	6
Mean	1081.14	809.99	1562.87	1371.35	4571.21
SD	1476.03	948.56	7132.23	7230.08	9922.2
20–100					
Max	68362	11891	67211	174688	319601
Min	3.5	6	0.37	0.10	9.25
Mean	2235.19	824.5	2581.81	2075.03	7070.87
SD	3003.98	954.64	4124.08	6184.18	12417.85
100–500					
Max	157312	7870	1081276	91961.55	876303.2
Min	1.2	5	0.23	0.007	6.10
Mean	2839.43	359.68	3568.86	1959.65	6785.67
SD	5248.15	475.91	18984.7	5084.45	18458.83
>500					
Max	478887	11141	1194391	134477	794353
Min	3.5	8	0.00	0.01	0.41
Mean	8629.17	364.75	13480.91	3757.34	18780.9
SD	27587.15	591.45	62589.98	8797.89	49524.76

Variable specifications are in accordance to section's 3.1 designs. Land is total planted area in hectares. Data is at municipality level. *Source: IBGE's 2006 Agricultural Census*.

higher dispersion rate is observed in the purchased inputs category, ranging from 1.59 in the 20-100ha group to 4.56, 4,54 and 5,32 in the 5-20ha, >500ha and 100-500 ha groups respectively. For capital stock, the indicator decreases sharply from 6.23 (3830/615) in the small farm group (0-5ha) and 5.27 in the 5-20ha group to 2.34 in the largest farm group (>500ha). As for output value, dispersion rates range from 1.75 to 2.72, which shows a relatively high dispersion rate intra-group in general.

3.3 Testing Nested Functional Forms, Slope and Monotonicity Properties

In order to verify the benefit of the translog functional form we employed an F-statistic to compare nested models. The results suggest rejection at a 99% of the more restrict Cobb-Douglas functional form in favor of the more flexible functional translog form. We also have assessed for differences in slope of coefficients between neighboring quantiles (*e.g.* $\theta = 0.10$ against $\theta = 0.20$) with the joint test for equality of slopes proposed by Koenker and Bassett (1982). The null hypothesis was rejected for all θs within every LASG.[3].

The monotocity property of the translog production function was checked by following Henningsen and Henning (2009)'s approach. Table 4 bellow shows the percentage points that do not fulfill the monotonicity

[3]Test results can be provided upon request.

condition for each prduction factor, which corresponds to the proportion of negative elasticities when employing a translog function form.

Table 4: Percentage of Monotonicity - Pooled and by LASG

Inputs	Pooled	0-5ha	5-20ha	20-100ha	100-500ha	>500ha
Land	1.7%	10.6%	5.3%	0.5%	1%	4.5%
Labor	16%	19.2%	33.8%	17.8%	20.2%	7.2%
Purchased	0%	0.3%	0%	0.2%	0%	0%
Capital	16.2%	6.2%	11.4%	10%	8.6%	5.6%

Source: Authors' estimations.

4 Results

We initially estimated our pooled model represented by eq. (1) and the models for each LASG c represented by eq. (2) by using standard OLS techniques and then by applying quantile regression techniques we estimate eq. (9). With the OLS and the quantile regression coefficient estimates for nine quantiles we then measure the elasticities of output with respect to land, capital labor and purchased inputs. They enable us to investigate for the heterogeneity in agricultural productivity between LASGs and between any LASG and the pooled estimate and within any given LASG.

Based on the OLS coefficient estimates the output-input elasticities measured with eqs. (3) and (5), are presented in Table 5 below. In general the aggregated measures show that, at the margin, purchased inputs have the largest contribution to output with an elasticity of 0.53, meaning that a 1% increase in purchased inputs implies an estimated increase of 0.53% in output. The second largest is capital followed by land and then labor. This order is also followed by the farms within the 5-20, 20-100 and 100-500ha intervals. For the smallest farms (0-5ha) labor comes in third and land in forth place, while for the largest (>500ha) land comes in second and capital in third place.

Table 5: Scale and Output-Input Elasticities - OLS Model

	Pooled	0-5ha	5-20ha	20-100ha	100-500ha	>500ha
$\hat{\psi}_A$	0.1789	0.0677	0.1821	0.1711	0.1883	0.2312
$\hat{\psi}_L$	0.0640	0.1718	0.0563	0.0480	0.0792	0.1723
$\hat{\psi}_I$	0.5294	0.5585	0.5324	0.5648	0.5082	0.4282
$\hat{\psi}_K$	0.1963	0.1860	0.2089	0.1972	0.1987	0.2231
$\sum \psi$	0.9687	0.9840	0.9797	0.9811	0.9744	1.0549

Source: Authors' estimations.

For the more dis-aggregated analysis based on OLS estimates by LASG, we can see that land productivity ranges from 0.07 to 0.23 when we move from the smallest to the largest farms with the intermediate ones (5-20, 20-100 and 100-500 ha) locked in a plateau around 0.18. For labor, the elasticity measures seem to follow an inverted-U pattern with the smallest and largest farms with the highest around 0.17 and the farms in the intermediate intervals with the lowest, around 0.06. Purchased inputs and capital importance to production are more homogeneous across farm sizes, with the latter increasing and the former decreasing in the largest farms (>500ha). As the comparison between the aggregated measures reached with the pooled model and the ones by LASG we can see that for land and labor aggregated estimates would tend to bias toward intermediate size farm productivity. Aggregated elasticity measures would also tend to respectively overestimate and underestimate the purchased inputs and capital contributions to productivity for the very large farms (>500ha).

As for returns to scale, measured with eqs. (4) and (6), results either reached with the pooled model or with the estimations for each LASG show that farms are operating pretty much under constant returns to scale with increasing returns only for the very largest (>500ha).

Next, we turn to the analysis of the heterogeneity within each LASG by using the QR method. As already highlighted above a benefit of using it is the ability to assess how the elasticity measures differ according to the different levels of production that each quantile represents, ranging from 0.10 to 0.90 within each LASG.

Figures from 1 through 5 below show the elasticity of scale and of the output with respect to capital, land, labor and purchased inputs. The shaded grey band along each curve represent a 99% confidence interval for the elasticity estimated values. The straight red line is the OLS estimates for each respective elasticity, where the two dashed lines represent 99% confidence intervals. All points and intervals obtained by bootstrapping with 10,000 sample replications. As a rule of thumb, if the gray area is not inside the dashed lines, the OLS model is not properly representing that section of the data, given the extreme heterogeneity in it.

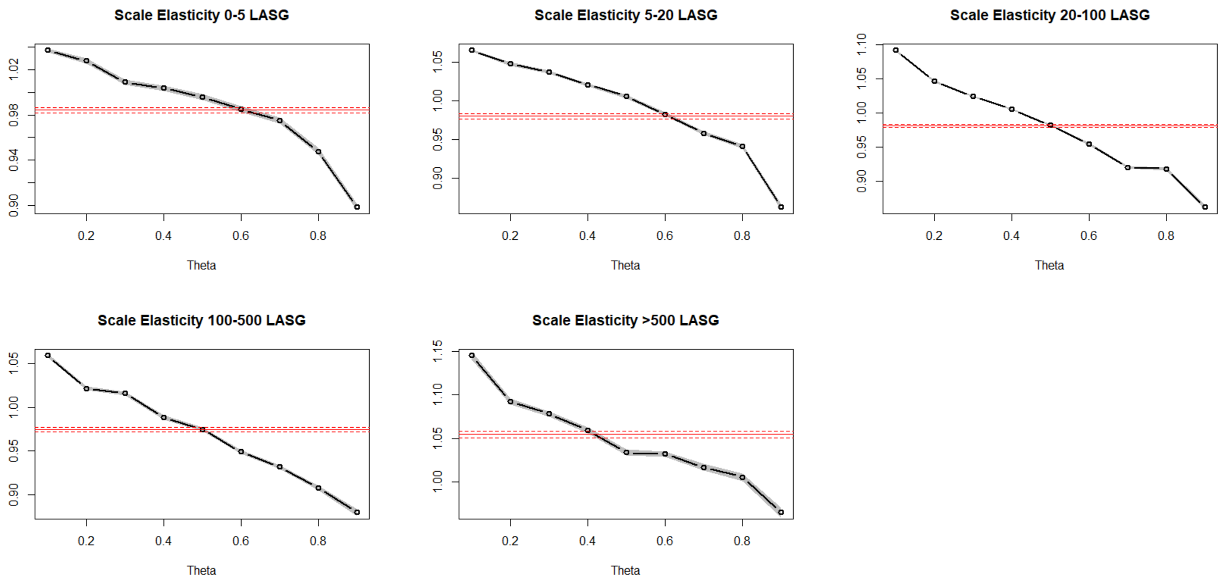

Figure 1: Scale Elasticities (y-axis) by Quantile (x-axis) for each LASG

In Figure 1 we can see that the economies of scale estimates are in general decreasing – in all LASGs – as the quantile increases meaning that farms with higher values of production tend to be more successful in exploiting gains of scale. More interesting however is the fact that all delineated curves cross the threshold value of 1. This means that in all LASGs there are farms operating under increasing, decreasing and constant returns to scale. For the LASGs 0-5, 20-100 and 100-500ha, the unitary elasticity is associated with the quantile 0.40, meaning that in these LASG's the botton 40% of farms in terms of value of production are experiencing increasing returns and the top 60% decreasing returns. For the LASG 5-20 and >500ha, the unitary elasticity is in the 0.55 and 0.85 quantiles, meaning that respectively the bottom 55% (top 45%) and 85% (top 15%) of the farms are operating under increasing (decreasing) returns to scale in these LASGs.

These results sharply contrast with the ones drawn from the OLS estimates in Table 5 that leads us to conclude in general for constant returns to scale no matter the farm size in terms of hectares. In fact, an analysis based on OLS would tend to underestimate the measures of economies of scale in the LASGs of 0-5 and 5-20ha for the bottom 60% of the farms and overestimate for the top 40%. In the 20-100 and 100-500 LASGs under- and overestimation happen in the bottom and top 50% respectively. For the uppermost largest farms (>500ha), OLS measure underestimate for the bottom 40% and overestimate for the top 60%.

Figure 2 shows the results for the elasticity of output with respect to capital by LASGs and quantiles. In general, the curves are negatively sloped suggesting that at the margin capital contributes more to production in the smaller farms. And again their comparison with the OLS estimates in Table 5 reveals that OLS method is very poor in taking into account heterogeneity. In general, OLS results would underestimate the contribution of capital to production at the margin for the farms in the bottom 45% and overestimate for the top 55%.

As for land, the graphs in Figure 3 do not show a homogeneous pattern across LASG's as in the Figures 1 and 2. In fact, the contribution of land to production at the margin follows a clear decreasing pattern for the smallest farms (0 and 5 ha), then follows a U pattern for the intermediate LASG's of 5-20 and 20-100 ha, with elasticities estimates reaching a minimum value at quantile 0.4, and then increasing for the largest farms,

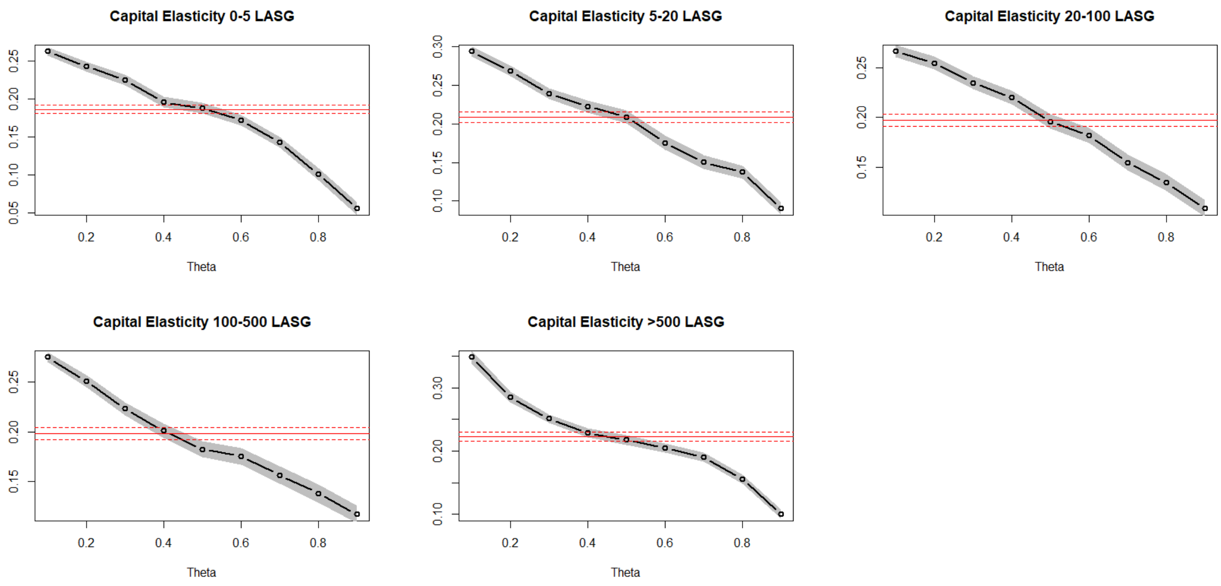

Figure 2: Capital Elasticities (y-axis) by Quantile (x-axis) for each LASG

with land area from 100-500 and >500 ha. Again, OLS estimates would be in general far off in describing the pattern of land contribution to output across farms within each LASG. For the smallest farms, OLS results would underestimate it for the bottom 20% and overestimate it for the top 80%. For the other quantiles, overestimation would occur in the bottom 60 up 70% of the farms.

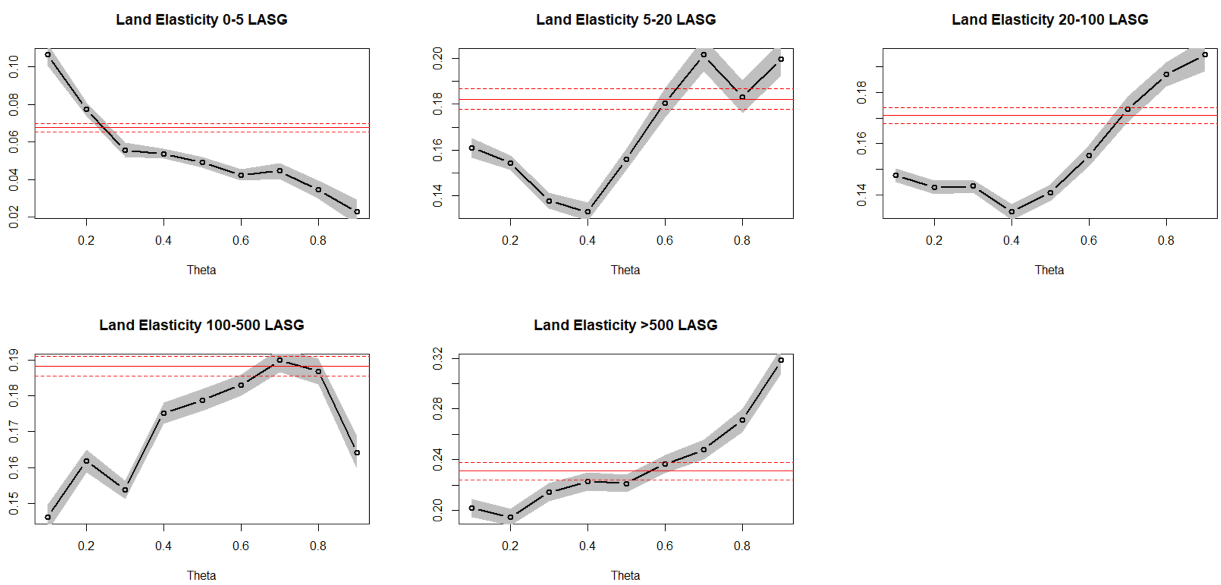

Figure 3: Land Elasticities (y-axis) by Quantile (x-axis) for each LASG

The next, Figure 4 shows the estimated output elasticities with respect to labor. For the smallest and largest LASGs, labor contribution to output at the margin increases with farm size. For intermediate farms, the elasticity measures follow either a decreasing pattern (20-100 ha) or a U pattern reaching the minimum value around quantiles 0.6 (5-20ha) and 0.7 (100-500 ha). Although the curves in Figure 4 also reveals that pooled estimation methods such as OLS would be a poor way to describe the productive pattern of an input across farms of different sizes, as seen in other Figures, the confidence intervals around the OLS labor elasticity estimates are much wider. In particular for the smallest farms (0-5 ha), this means a larger portion of quantiles for which the difference between a quantile estimates and the OLS estimate are not statistically significant.

Lastly, Figure 5 shows the productive patterns of purchased inputs across farms of different sizes within each

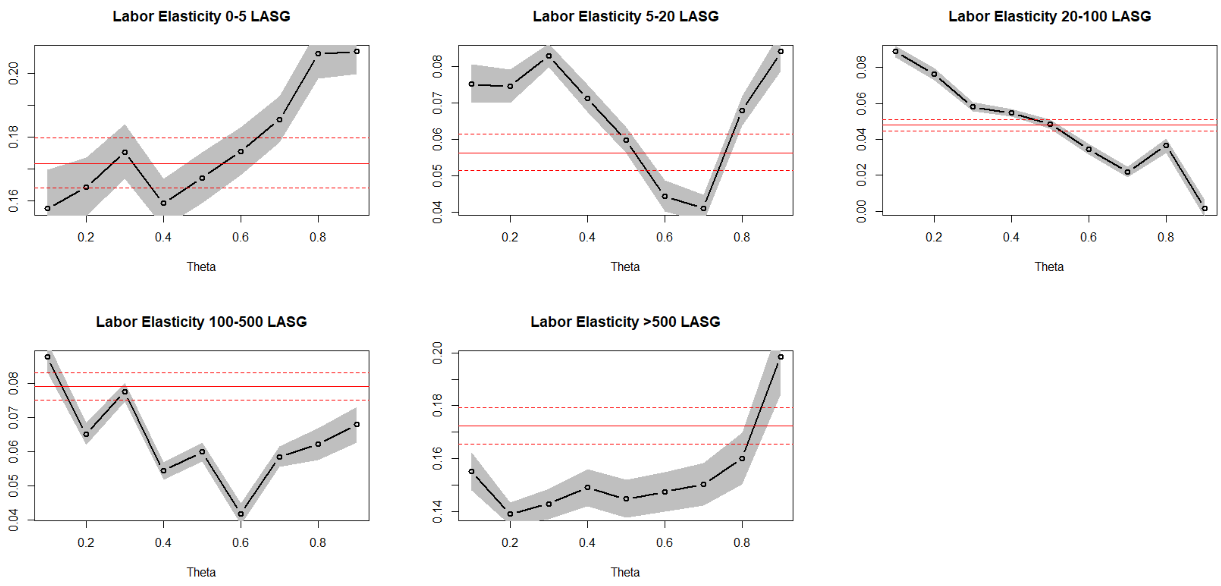

Figure 4: Labor Elasticities (y-axis) by Quantile (x-axis) for each LASG

LASG. At the margin, they are increasing for the farms with land area withing 0-5ha and follow a somewhat inverted-U shape for the farms in the other LASGs. With the peak being reached around the 0.4 and 0.5 quantiles for the farms in the 5-20 and 20-100ha intervals and in the 0.3 and 0.2 quantiles for the ones between 100-500ha and >500ha respectively. Overall we see that conclusions based on OLS would in general underestimate the productive pattern of purchased inputs to output compared to most of the quantile estimates in each LASG.

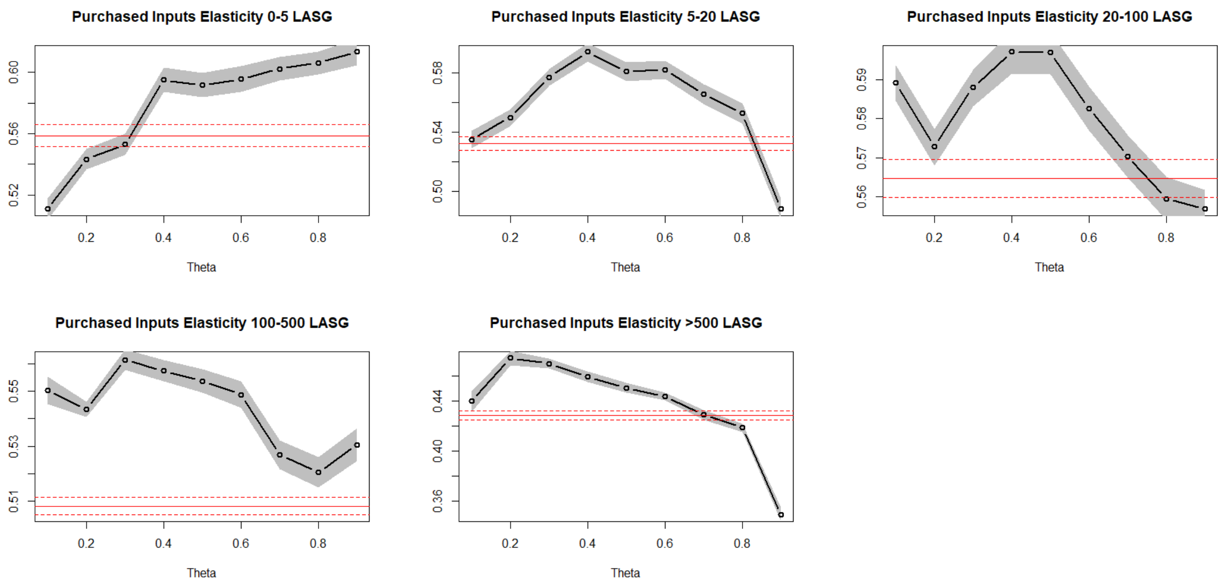

Figure 5: Purchased Inputs Elasticities (y-axis) by Quantile (x-axis) for each LASG

5 Conclusion

In this paper we have developed an empirical production model based on a fully flexible functional form and applied quantile regression and bootstrapping techniques to look at the heterogeneity in agricultural factor productivity in Brazil at the municipality level. We have estimated scale and output-input elasticities with respect to land, labor, capital and purchased inputs that are allowed to vary across and within groups of different farm sizes. Nested functional form specifications and the equality of production function slopes resulting from the quantile regression estimates have been performed along with the checking of monotonicity properties.

The results show the presence of significant heterogeneity in factor productivity across and within all farm size groups not only because farms of different sizes use different input proportions but rather because there are inter- and intra-group farm-specific effects, suggesting that 1) OLS modeling approaches that pool farms of different sizes together will yield biased estimators for factor productivity elasticities; and 2) these biases do not disappear in studies that accommodate for size heterogeneity by including dummies of by performing separate regressions according to size.

Further research should continue to focus on the heterogeneity across and within farms on two fronts. On one, by looking at the heterogeneity patterns in factor productivity when farms are allowed to be technically inefficient with the use of quantile frontier techniques as in (Chidmi et al., 2011; Kaditi and Nitsi, 2010). And on another, by performing a heterogeneity decomposition analysis that would allow for the disentangling of fixed from heterogeneous effects with the application of a quantile regression panel data approach as in Graham et al. (2015). This will require the updating of our 2006 database with data from the recently released 2017 IBGE Agricultural Census.

References

Aly, H., Belbase, K., Grabowski, R., and Kraft, S. (1987). The technical efficiency of illinois grain farms: An application of a ray-homothetic production function. *Southern Journal of Agricultural Economics*, 19. https://ageconsearch.umn.edu/record/29313/. Cited on page 2.

Avila, A. F. D., Rodrigues, G. S., Vedovoto, G. L., PENTEADO FILHO, R. d. C., and da Fonseca Junior, W. C. (2015). Embrapa experience on the impact assessment of agricultural r&d: 15 years using a multidimensional approach. In *Área de Informação da Sede-Artigo em anais de congresso (ALICE)*. In: IMPAR CONFERENCE, 2015, Paris. Impacts of agricultural research: towards ... https://www.alice.cnptia.embrapa.br/handle/doc/1036444. Cited on page 2.

Bagi, F. S. (1982). Relationship between farm size and technical efficiency in west tennessee agriculture. *Journal of Agricultural and Applied Economics*, 14(2):139–144. https://doi.org/10.1017/S0081305200024961. Cited on page 2.

Block, S. (2014). The decline and rise of agricultural productivity in sub-saharan africa since 1961. In *African Successes, Volume IV: Sustainable Growth*, pages 13–67. University of Chicago Press. https://www.nber.org/chapters/c13435. Cited on page 3.

Bragagnolo, C., Spolador, H. F., and Barros, G. (2010). Regional brazilian agriculture tfp analysis: a stochastic frontier analysis approach. *Revista Economia*, 11(4):217–242. http://www.anpec.org.br/revista/vol11/vol11n4p217_242.pdf. Cited on page 2.

Bravo-Ureta, B. E. (1986). Technical efficiency measures for dairy farms based on a probabilistic frontler function model. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 34(3):399–415. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1744-7976.1986.tb02220.x. Cited on page 2.

Bravo-Ureta, B. E. and Rieger, L. (1990). Alternative production frontier methodologies and dairy farm efficiency. *Journal of agricultural Economics*, 41(2):215–226. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1477-9552.1990.tb00637.x. Cited on page 3.

Bravo-Ureta, B. E., Solís, D., López, V. H. M., Maripani, J. F., Thiam, A., and Rivas, T. (2007). Technical efficiency in farming: a meta-regression analysis. *Journal of productivity Analysis*, 27(1):57–72. https://link.springer.com/article/10.1007/s11123-006-0025-3. Cited on page 3.

Bustos, P., Caprettini, B., and Ponticelli, J. (2016). Agricultural productivity and structural transformation: Evidence from brazil. *American Economic Review*, 106(6):1320–65. https://www.aeaweb.org/articles?id=10.1257/aer.20131061. Cited on page 3.

Butzer, R., Mundlak, Y., and Larson, D. (2012). *Measures of fixed capital in agriculture. In: Fuglie, K., Wang, S., and Ball, V. (Eds.), Productivity Growth in Agriculture: An International Perspective. CAB International, Cambridge, MA, USA. Chapter 15.* CABI. https://elibrary.worldbank.org/doi/abs/10.1596/1813-9450-5472. Cited on page 6.

Cao, K. H. and Birchenall, J. A. (2013). Agricultural productivity, structural change, and economic growth in post-reform china. *Journal of Development Economics*, 104:165–180. https://doi.org/10.1016/j.jdeveco.2013.06.001. Cited on page 3.

Carpenter, J. and Bithell, J. (2000). Bootstrap confidence intervals: when, which, what? a practical guide for medical statisticians. *Statistics in medicine*, 19(9):1141–1164. https://doi.org/10.1002/(SICI)1097-0258(20000515)19:9<1141::AID-SIM479>3.0.CO;2-F. Cited on page 4.

Chidmi, B., Solís, D., and Cabrera, V. E. (2011). Analyzing the sources of technical efficiency among heterogeneous dairy farms: A quantile regression approach. *Animal Production*, 13(2). http://animalproduction.net/index.php/JAP/article/view/315. Cited on page 13.

Darku, A. B., Malla, S., and Tran, K. C. (2013). Historical review of agricultural efficiency studies. *CAIRN Research Network*. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.646.4781&rep=rep1&type=pdf. Cited on page 2.

Davis, K., Nkonya, E., Kato, E., Mekonnen, D. A., Odendo, M., Miiro, R., and Nkuba, J. (2012). Impact of farmer field schools on agricultural productivity and poverty in east africa. *World Development*, 40(2):402–413. https://doi.org/10.1016/j.worlddev.2011.05.019. Cited on page 3.

Dawson, P. and Woodford, C. (1991). Generalized farm-specific technical efficiency in the england and wales dairy sector. *Oxford Agrarian Studies*, 19(1):53–60. https://doi.org/10.1080/13600819108424035. Cited on page 3.

Gasques, J. G., Bastos, E. T., Bacchi, M. R. P., and Valdes, C. (2010). Produtividade total dos fatores e transformações da agricultura brasileira: análise dos dados dos censos agropecuários. *A agricultura brasileira: desempenho, desafios e perspectivas*. https://ideas.repec.org/p/anp/en2010/184.html. Cited on page 2.

Gasques, J. G., Bastos, E. T., Valdes, C., and Bacchi, M. R. P. (2014). Produtividade da agricultura: resultados para o brasil e estados selecionados. *Revista de Política Agrícola*, 23(3):87–98. https://seer.sede.embrapa.br/index.php/RPA/article/view/943. Cited on page 2.

Gollin, D. (2010). Agricultural productivity and economic growth. *Handbook of agricultural economics*, 4:3825–3866. https://doi.org/10.1016/S1574-0072(09)04073-0. Cited on page 3.

Gollin, D., Lagakos, D., Waugh, M. E., et al. (2011). *The agricultural productivity gap in developing countries*. Leonard N. Stern School of Business, Department of Economics. https://www.aeaweb.org/conference/2013/retrieve.php?pdfid=238. Cited on page 3.

Gornall, J., Betts, R., Burke, E., Clark, R., Camp, J., Willett, K., and Wiltshire, A. (2010). Implications of climate change for agricultural productivity in the early twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1554):2973–2989. https://doi.org/10.1098/rstb.2010.0158. Cited on page 3.

Graham, B. S., Hahn, J., Poirier, A., and Powell, J. L. (2015). Quantile regression with panel data. Technical report, National Bureau of Economic Research. https://www.doi.org/10.3386/w21034. Cited on page 13.

Hallam, D. and Machado, F. (1996). Efficiency analysis with panel data: A study of portuguese dairy farms. *European review of agricultural economics*, 23(1):79–93. https://doi.org/10.1093/erae/23.1.79. Cited on page 3.

Hassine-Belghith, N. B. (2009). Exporting, technical efficiency and product quality: An empirical analysis of the agricultural sector in the mediterranean countries. *The Journal of Development Studies*, 45(5):769–788. https://doi.org/10.1080/00220380902753201. Cited on page 3.

Hayami, Y., Ruttan, V. W., et al. (1971). *Agricultural development: an international perspective.* Baltimore, Md/London: The Johns Hopkins Press. https://www.cabdirect.org/cabdirect/abstract/19721890134. Cited on page 5.

Helfand, S. M., Moreira, A. R. B., and Figueiredo, A. M. R. (2011). Explicando as diferenças de pobreza entre produtores agrícolas no brasil: simulações contrafactuais com o censo agropecuário 1995-96. *Revista de economia e sociologia rural*, 49(2):391–418. https://doi.org/10.1590/S0103-20032011000200006. Cited on page 2.

Helfand, S. M., Taylor, M. P., et al. (2017). The inverse relationship between farm size and productivity: Refocusing the debate. *Mode of access: http://economics. ucr. edu/pacdev/pacdev-papers/the_inverse_relationship. pdf.–Date of access*, 28. http://economics.ucr.edu/pacdev/pacdev-papers/the_inverse_relationship.pdf. Cited on page 2.

Henningsen, A. and Henning, C. H. (2009). Imposing regional monotonicity on translog stochastic production frontiers with a simple three-step procedure. *Journal of Productivity Analysis*, 32(3):217. https://doi.org/10.1007/s11123-009-0142-x. Cited on page 8.

Kaditi, E. A. and Nitsi, E. I. (2010). Applying regression quantiles to farm efficiency estimation. Technical report. https://core.ac.uk/download/pdf/6550611.pdf. Cited on page 13.

Kawagoe, T. and Hayami, Y. (1985). An intercountry comparison of agricultural production efficiency. *American Journal of Agricultural Economics*, 67(1):87–92. https://doi.org/10.2307/1240827. Cited on page 2.

Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, pages 33–50. https://www.doi.org/10.2307/1913643. Cited on page 4.

Koenker, R. and Bassett, G. (1982). Tests of linear hypotheses and l" 1 estimation. *Econometrica: Journal of the Econometric Society*, pages 1577–1583. https://www.doi.org/10.2307/1913398. Cited on page 8.

Krinsky, I. and Robb, A. L. (1986). On approximating the statistical properties of elasticities. *The Review of Economics and Statistics*, pages 715–719. https://www.doi.org/10.2307/1913398. Cited on page 4.

Krinsky, I. and Robb, A. L. (1991). Three methods for calculating the statistical properties of elasticities: a comparison. *Empirical Economics*, 16(2):199–209. https://doi.org/10.1007/BF01193491. Cited on page 4.

Kumbhakar, S. C., Biswas, B., and Bailey, D. (1989). A study of economic efficiency of utah dairy farmers: a system approach. *The review of Economics and Statistics*, pages 595–604. https://doi.org/10.2307/1928101. Cited on page 2.

Kumbhakar, S. C., Ghosh, S., and McGuckin, J. T. (1991). A generalized production frontier approach for estimating determinants of inefficiency in us dairy farms. *Journal of Business & Economic Statistics*, 9(3):279–286. https://doi.org/10.1080/07350015.1991.10509853. Cited on page 3.

Kumbhakar, S. C. and Heshmati, A. (1995). Efficiency measurement in swedish dairy farms: an application of rotating panel data, 1976–88. *American Journal of Agricultural Economics*, 77(3):660–674. https://doi.org/10.2307/1243233. Cited on page 3.

Lansink, A. O., Pietola, K., and Bäckman, S. (2002). Effciency and productivity of conventional and organic farms in finland 1994–1997. *European Review of Agricultural Economics*, 29(1):51–65. https://doi.org/10.1093/erae/29.1.51. Cited on page 3.

Machado, G. C., Bacha, C. J. C., and Johnston, F. L. (2020). Revisão sistemática dos trabalhos que calculam a ptf da agropecuária brasileira. *Revista de Política Agrícola*, 1(1):82. https://seer.sede.embrapa.br/index.php/RPA/article/view/1488. Cited on page 2.

Moreira, A. R., Helfand, S. M., and Figueiredo, A. M. (2007). Explicando as diferenças na produtividade agrícola no brasil. http://repositorio.ipea.gov.br/handle/11058/1966. Cited 2 times on pages 2 and 6.

Paul, C., Nehring, R., Banker, D., and Somwaru, A. (2004). Scale economies and efficiency in us agriculture: are traditional farms history? *Journal of Productivity Analysis*, 22(3):185–205. https://doi.org/10.1007/s11123-004-7573-1. Cited on page 3.

Rada, N., Helfand, S., and Magalhães, M. (2018). Agricultural productivity growth in brazil: Large and small farms excel. *Food Policy*. https://doi.org/10.1016/j.foodpol.2018.03.014. Cited 2 times on pages 2 and 3.

Rada, N. and Valdes, C. (2012). Policy, technology, and efficiency of brazilian agriculture. *USDA-ERS Economic Research Report*, (137). http://dx.doi.org/10.2139/ssrn.2112029. Cited on page 2.

Rocha, R. and Soares, R. R. (2015). Water scarcity and birth outcomes in the brazilian semiarid. *Journal of Development Economics*, 112:72–91. https://doi.org/10.1016/j.jdeveco.2014.10.003. Cited on page 6.

Schneider, K. and Gugerty, M. K. (2011). Agricultural productivity and poverty reduction: Linkages and pathways. *Libraries Test Journal*, 1(1):56–74. http://dx.doi.org/10.7152/12259. Cited on page 3.

Serra, T., Zilberman, D., and Gil, J. M. (2008). Farms' technical inefficiencies in the presence of government programs. *Australian Journal of Agricultural and Resource Economics*, 52(1):57–76. https://doi.org/10.1111/j.1467-8489.2008.00412.x. Cited on page 3.

Tauer, L. W. and Belbase, K. P. (1987). Technical efficiency of new york dairy farms. *Northeastern Journal of Agricultural and Resource Economics*, 16(1):10–16. https://doi.org/10.1017/S0899367X00000313. Cited on page 2.

Thiele, H. and Brodersen, C. M. (1999). Differences in farm efficiency in market and transition economies: empirical evidence from west and east germany. *European Review of Agricultural Economics*, 26(3):331–347. https://doi.org/10.1093/erae/26.3.331. Cited on page 3.

Tipi, T. and Rehber, E. (2006). Measuring technical efficiency and total factor productivity in agriculture: The case of the south marmara region of turkey. *New Zealand Journal of Agricultural Research*, 49(2):137–145. https://doi.org/10.1080/00288233.2006.9513703. Cited on page 3.