

Título do capítulo	CAPÍTULO 2 – UMA VISÃO DA PRODUÇÃO DA AGRICULTURA BRASILEIRA COM BASE EM DADOS RECENTES DO CENSO AGROPECUÁRIO
Autores(as)	Geraldo da Silva e Souza Eliane Gonçalves Gomes Eliseu Roberto de Andrade Alves
DOI	http://dx.doi.org/10.38116/978-65-5635-011-0/cap2

Título do livro	UMA JORNADA PELOS CONTRASTES DO BRASIL: CEM ANOS DO CENSO AGROPECUÁRIO
Organizadores(as)	José Eustáquio Ribeiro Vieira Filho José Garcia Gasques
Volume	-
Série	-
Cidade	Brasília
Editora	Instituto de Pesquisa Econômica Aplicada (Ipea)
Ano	2020
Edição	-
ISBN	978-65-5635-011-0
DOI	http://dx.doi.org/10.38116/978-65-5635-011-0

© Instituto de Pesquisa Econômica Aplicada – ipea 2020

As publicações do Ipea estão disponíveis para *download* gratuito nos formatos PDF (todas) e EPUB (livros e periódicos). Acesse: <http://www.ipea.gov.br/portal/publicacoes>

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto de Pesquisa Econômica Aplicada ou do Ministério da Economia.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

UMA VISÃO DA PRODUÇÃO DA AGRICULTURA BRASILEIRA COM BASE EM DADOS RECENTES DO CENSO AGROPECUÁRIO

Geraldo da Silva e Souza¹
Eliane Gonçalves Gomes²
Eliseu Roberto de Andrade Alves³

1 INTRODUÇÃO

Há fortes indícios de concentração de renda na agricultura brasileira. Os dados do Censo Agropecuário 2006 indicaram que a renda rural bruta estava concentrada em alguns estabelecimentos rurais com renda anual declarada superior a 200 salários mínimos. Essa classe compreendia 27.306 estabelecimentos rurais, representando 0,62% do total de estabelecimentos no Brasil, e era responsável por 51% do valor total da produção na época. Para o Censo Agropecuário 2017, esses números não mudaram muito. A mesma classe de renda compreende cerca de 24 mil estabelecimentos, representando 0,60% do total. O índice de concentração de Gini no nível do estabelecimento saltou de 0,85, em 2006, para 0,90, em 2017. As medianas dos índices municipais de Gini passaram de 0,84, 0,91, 0,90, 0,81 e 0,91 para as regiões Norte, Nordeste, Sudeste, Sul e Centro-Oeste em 2006 para 0,80, 0,86, 0,84, 0,78 e 0,87 em 2017, respectivamente. Vemos algumas melhorias em nível municipal, principalmente na região Sul. No entanto, no geral, a concentração ainda é alta. A evidência é que a mão de obra qualificada e os insumos tecnológicos continuam a ser os impulsionadores da inserção produtiva e da redução da pobreza nos campos.

Para reduzir diferenças e aumentar a produtividade, é necessário o controle das imperfeições de mercado. Conforme enfatizam Souza e Gomes (2019), as imperfeições de mercado são a principal causa que inibe o acesso dos agricultores à tecnologia e, portanto, à inclusão produtiva. São o resultado de assimetrias no acesso ao crédito para produção, infraestrutura, disponibilidade de informações, extensão rural, assistência técnica, entre outros (Alves e Souza, 2015). Com melhores condições de produção, os grandes produtores são capazes de negociar melhores preços de insumos e produtos. Souza, Gomes e Alves (2017) destacam que os pequenos agricultores tipicamente vendem seus produtos a valores relativamente mais baixos e compram insumos a preços relativamente mais altos, tal fato dificulta em muito o acesso à tecnologia. Essa negociação desfavorável pode levar a preços mais altos para a adoção de melhores tecnologias e para o uso de mão de obra qualificada.

Neste capítulo, identificam-se, no nível do município, fatores que afetam diretamente a produção. Os dados utilizados são do Censo Agropecuário 2017. A análise é semelhante à de Souza e Gomes (2019) para o Censo Agropecuário 2006. O processo de modelagem postula uma representação de Cobb-Douglas para a função de produção, em uma abordagem típica de fronteira estocástica. Os fatores de imperfeição do mercado afetam a fronteira e o componente de ineficiência. No geral, o estudo tem o objetivo de indicar orientações para políticas públicas direcionadas ao controle das variáveis associadas às imperfeições de mercado.

2 DADOS

A fonte de dados utilizada neste estudo é o Censo Agropecuário 2017. Os dados dos estabelecimentos rurais foram agregados por município. Temos dados válidos para 5.236 municípios, o que representa 94% do total de municípios no Brasil (5.570). Esses municípios compreendem 4.916.083 estabelecimentos rurais, 97% do total de estabelecimentos investigados no censo (5.073.324).

1. Professor do Departamento de Estatística da Universidade de Brasília (UnB). *E-mail*: <geraldosouza@unb.br>.

2. Pesquisadora da Empresa Brasileira de Pesquisa Agropecuária (Embrapa). *E-mail*: <eliane.gomes@embrapa.br>.

3. Pesquisador da Embrapa. *E-mail*: <eliseu.alves@embrapa.br>.

A produção é a declarada como receita bruta da agricultura e os insumos são terra, mão de obra e insumos tecnológicos. Todas as variáveis foram transformadas em *rank* e, na sequência, medidas em *logs*. *Trabalho* é medido como o *log* do *rank* municipal da soma de despesa com salário e despesa com contratação de serviços. As despesas com os insumos terra e tecnologia não estão disponíveis diretamente nos dados do censo e foram construídas *proxies* com escores apropriados. Para a terra, como *proxy* considerou-se o *log* do *rank* municipal da soma da área utilizada na agricultura nos estabelecimentos com a terra arrendada. Para a tecnologia (insumos tecnológicos), fez-se uso de um modelo fatorial aplicado a muitas variáveis. Um único escore tecnológico foi obtido por meio da média ponderada das classificações (*ranks*) dessas variáveis. Os pesos aplicados foram as comunalidades relativas derivadas do modelo fatorial. Quatro fatores foram retidos com o uso da regra do *valor próprio mínimo = 1*. Os insumos tecnológicos incluem: despesas com sementes de mudas, sal, ração, medicamentos, adubos, corretivos, agrotóxicos; despesas com combustíveis, energia elétrica, outras despesas, transporte da produção; capacidade de armazéns, infláveis, graneliros, silos; número de tratores, semeadeiras ou plantadeiras, colheitadeiras, adubadeiras ou distribuidoras de calcário; número de caminhões, utilitários, automóveis, motos, aviões e aeronaves de uso agrícola; quantitativo de animais (bovinos, bubalinos, equinos, asininos, muares, suínos, caprinos, ovinos, galinhas, galos, frangas, frangos e pintos, codornas, patos, gansos, marrecos, perdizes e faisões, coelhos, perus, avestruz); e patrimônio (caixas de abelha, total de pés de lavouras permanentes, total de pés de silvicultura).

As variáveis contextuais escolhidas foram definidas por um índice municipal de uso relativo de práticas agrícolas ecológicas, crédito financeiro, participação em cooperativas, educação e assistência técnica. Essas variáveis são todas *proxies* para imperfeições de mercado.

O índice de práticas agrícolas amigáveis ao meio ambiente é também uma média ponderada de ordenações resultante de um modelo fatorial, com pesos definidos por comunalidades relativas. Os totais observados incluem, entre outras, categorias de manejo florestal, outras práticas ambientais (plantio em contornos, rotação de culturas, descanso do solo, conservação de encostas) e preparação do solo. O índice é uma média ponderada dessas variáveis. Nessa dimensão, um único fator foi retido pela regra do *valor próprio mínimo = 1*.

O índice de crédito financeiro é definido por meio do *rank* obtido na classificação da proporção de agricultores que recebem crédito financeiro. O índice de cooperativismo é o *rank* obtido na classificação da proporção de agricultores membros de cooperativas agrícolas. O índice de educação ou alfabetização é a classificação (*rank*) da proporção de agricultores alfabetizados. A assistência técnica é medida pela classificação (*rank*) da proporção de agricultores que receberam assistência técnica.

Deve-se notar que, no nível dos estabelecimentos rurais, todas as variáveis contextuais são variáveis indicadoras.

3 METODOLOGIA

As referências básicas para a discussão nesta seção são Khumbhakar e Lovell (2000), Coelli, Rao e O'Donnell (2005), Greene (2011) e Stata (2019). Os modelos de fronteira de produção estocástica apareceram pela primeira vez na literatura no artigo de Aigner, Lovell e Schmidt (1977).

Na análise econométrica de um modelo de fronteira de produção estocástica, inicialmente se postula a existência de uma função de produção real não negativa $f(x, z, \theta)$, dependente do vetor de insumos x de dimensão k , do vetor de variáveis contextuais z de dimensão g , e de um vetor de parâmetros θ de dimensão d . Sem a presença de erros idiossincráticos e de ineficiência, a produção máxima y_j que pode ser obtida pelo município j com a utilização do nível de insumos x_j , dadas as covariáveis z_j , vem dado por $y_j = f(x_j, z_j, \theta)$.

A inclusão de ineficiência no processo produtivo pressupõe a existência de uma componente estocástica $\eta_j \in (0,1)$, de modo que a produção observada obtém-se da equação $y_j = f(x_j, z_j, \theta)\eta_j$. Se η_j é próximo a 1, isto quer dizer que o município j tem sua produção próxima do nível ótimo definido pela função de produção $f(x, z, \theta)$. Quando $\eta_j < 1$, o município j não produz o máximo possível definido pela tecnologia disponível incorporada na função de produção $f(x, z, \theta)$.

Tipicamente, as observações de produção também contêm variações aleatórias de pequena magnitude que deslocam o produto. É comum postular a presença de choques estocásticos na relação de produção por meio da inclusão, na modelagem, de variáveis aleatórias reais v_j , de modo que $y_j = f(x_j, z_j, \theta)\eta_j \exp(v_j)$.

A especificação anteriormente citada é equivalente ao modelo estatístico $\ln y_j = \ln f(x_j, z_j, \theta) + v_j - u_j$, em que u_j é uma variável aleatória positiva que representa a componente de ineficiência, isto é, $u_j = -\ln(\eta_j)$.

Uma função de produção de uso comum na *teoria da produção* é dada pela especificação Cobb-Douglas $f(x, z, \theta) = C \prod_{v=1}^k x_v^{\beta_v} \exp(z'w)$. Tem-se $\theta = (\beta, w)$, em que $\beta_v > 0$ é a elasticidade do fator de produção x_v e w_l é o parâmetro associado ao efeito da l -ésima componente de z . Portanto, na forma logarítmica, $\ln(y_j) = \ln(C) + \sum_{v=1}^k \beta_v \ln(x_{v,j}) + \sum_{l=1}^g w_l z_{l,j} + v_j - u_j$.

Especificações distintas para as componentes de erro conduzem a modelos de fronteira estocástica alternativos. Supõe-se que v_j seja independente de u_j . Os erros idiossincráticos representam uma amostra aleatória da distribuição normal, com média zero e variância σ^2 . Os u_j representam realizações independentes da distribuição exponencial com variância σ_u^2 , da distribuição meia-normal obtida pelo truncamento positivo da distribuição normal com média zero e variância σ_u^2 , ou ainda da distribuição normal truncada resultante do truncamento positivo da distribuição normal com média μ e variância δ^2 . Valores esperados dessas distribuições são σ_u^2 no caso exponencial, $\sqrt{2/\pi}\sigma_u$ para a meia-normal e $\mu + \phi\lambda$, com $\lambda = \phi(\mu/\delta)/\Phi(\mu/\delta)$ para a normal truncada, em que $\phi(\cdot)$ e $\Phi(\cdot)$ representam a função densidade de probabilidade e a função de distribuição da distribuição normal-padrão, respectivamente.

O vetor paramétrico θ é estimado pelo método da verossimilhança máxima. O processo inferencial é assintótico. As funções de verossimilhança para cada um dos casos que serão considerados são dadas por:

$$\text{modelo normal-exponencial: } L(\theta^*) = \sum_{j=1}^n \left\{ -\ln \sigma_u + \frac{\sigma^2}{2\sigma_u^2} + \ln \left(\frac{-\varepsilon_j - \frac{\sigma^2}{\sigma_u}}{\sigma} \right) + \frac{\varepsilon_j}{\sigma} \right\}, \quad (1)$$

$$\text{modelo normal-meia-normal: } L(\theta^*) = \sum_{j=1}^n \left\{ \frac{1}{2} \ln \left(\frac{2}{\pi} \right) - \ln(\sigma_s) + \ln \left(-\frac{\rho\varepsilon_j}{\sigma_s} \right) - \frac{\varepsilon_j^2}{2\sigma_s^2} \right\}, \quad (2)$$

modelo normal-normal truncada:

$$L(\theta^*) = \sum_{j=1}^n \left\{ \frac{1}{2} \ln(2\pi) - \ln(\sigma_s) - \ln \Phi \left(\frac{\mu}{\sigma_s \sqrt{\gamma}} \right) + \ln \Phi \left(\frac{(1-\gamma)\mu - \gamma\varepsilon_j}{\sigma_s \sqrt{\gamma(1-\gamma)}} \right) \right\} - \sum_{j=1}^n \left\{ \frac{1}{2} \left(\frac{\varepsilon_j + \mu}{\sigma_s} \right)^2 \right\} - \sum_{j=1}^n \left\{ \frac{1}{2} \left(\frac{\varepsilon_j + \mu}{\sigma_s} \right) \right\}. \quad (3)$$

Nas expressões anteriores, $\varepsilon_j = v_j - u_j$ é a diferença $\ln y_j - \ln f(x_j, z_j, \theta)$, $\sigma_s^2 = \sigma^2 + \sigma_u^2$, $\rho = \sigma_u/\sigma$ e $\gamma = \sigma_u^2/\sigma_s^2$. O parâmetro θ^* inclui θ e qualquer parametrização adicional presente na componente de ineficiência.

Efeitos associados às variáveis contextuais nas componentes de ineficiência são modelados por meio da utilização dos parâmetros presentes na especificação da componente de ineficiência. Para as distribuições exponencial e meia-normal, postula-se que $\sigma_u^2 = \exp(m'b)$, em que m é o vetor de covariáveis e b o vetor paramétrico correspondente. Para a distribuição normal truncada, postula-se que $\mu = m'b$. Em qualquer caso, a média esperada da distribuição é uma função monotônica do construto linear $m'b$. É possível também modelar dependência na componente de variância idiossincrática σ^2 .

A medida de eficiência técnica é definida por (Stata 2019):

$$te_j = \left(\frac{1 - \Phi(\sigma_* - \mu_{*j}/\sigma_*)}{1 - \Phi(-\mu_{*j}/\sigma_*)} \right) \exp\left(-\mu_{*j} + \frac{1}{2}\sigma_*^2\right), \quad (4)$$

em que $\mu_{*j} = -\varepsilon_j \frac{\sigma_u^2}{\sigma_s^2}$, $\sigma_* = \frac{\sigma_u \sigma}{\sigma_s}$ para o modelo normal-meia-normal, $\mu_{*j} = -\varepsilon_j - \frac{\sigma_u^2}{\sigma_s^2}$, $\sigma_* = \sigma$ para o caso normal-exponencial, e $\mu_{*j} = \frac{-\varepsilon_j \sigma_u^2 + \mu \sigma^2}{\sigma_s^2}$, $\sigma_* = \frac{\sigma_u \sigma}{\sigma_s}$ para o caso normal-normal truncada.

4 RESULTADOS ESTATÍSTICOS

4.1 Estatísticas descritivas

Como as variáveis são medidas como *log* dos *ranks*, valores médios ou medianos para o país não fazem sentido. A estimação da função de produção com o uso de *ranks*, contudo, é possível para o país e a abordagem tem um cunho de natureza não paramétrica (Conover, 1999). As comparações entre regiões também são viáveis e robustas (Conover, 1999). Nesse contexto, as tabelas 1 e 2 mostram os resumos de cinco números (mínimo, primeiro quartil, mediana, terceiro quartil e máximo) regionais obtidos para as variáveis de produção (renda bruta, terra, trabalho e tecnologia) e para as variáveis contextuais (assistência técnica, alfabetização, cooperativismo, crédito e meio ambiente), respectivamente. Note-se que a ordem de grandeza em que os atributos são medidos é a mesma devido ao uso dos *ranks* (*logs*) na detecção de diferenças.

Na tabela 1, percebe-se que a renda é mais intensa na região Centro-Oeste, seguida pelas regiões Sul, Norte, Sudeste e Nordeste. A variação é menor no Centro-Oeste e no Sul, e a maior variação ocorre no Nordeste. Os valores medianos das regiões Norte e Sudeste não diferem muito. A tecnologia segue a mesma ordem. O Centro-Oeste domina claramente. A região Norte tem a maior variabilidade. O uso da terra é mais intenso no Centro-Oeste, seguido pelas regiões Norte, Sudeste, Nordeste e Sul. O Centro-Oeste e o Norte dominam e mostram variações menores. O trabalho é mais intenso na região Centro-Oeste, seguida pelas regiões Sudeste, Sul, Norte e Nordeste. O Centro-Oeste também tem a menor variabilidade. A evidência resultante das observações das estatísticas descritivas dos fatores de produção é uma clara dominância das regiões Centro-Oeste e Sul. O gráfico 1 ilustra o comportamento das distribuições empíricas regionais das quatro variáveis de produção.

TABELA 1
Resumo de cinco números para as variáveis de produção – Grandes Regiões

		Renda	Terra	Trabalho	Tecnologia
Centro-Oeste Número de municípios = 458	Mínimo	6,347	5,352	6,361	6,872
	1º quartil	8,173	8,319	8,141	8,004
	Mediana	8,461	8,513	8,420	8,243
	3º quartil	8,578	8,589	8,539	8,408
	Máximo	8,623	8,623	8,623	8,574
Nordeste Número de municípios = 1.573	Mínimo	3,258	2,639	1,386	5,961
	1º quartil	6,586	7,106	6,640	7,583
	Mediana	7,295	7,805	7,437	7,829
	3º quartil	7,828	8,241	8,003	8,052
	Máximo	8,622	8,621	8,621	8,532
Norte Número de municípios = 423	Mínimo	3,466	2,773	3,296	5,847
	1º quartil	7,550	8,139	7,240	7,563
	Mediana	8,083	8,449	7,954	8,034
	3º quartil	8,361	8,547	8,298	8,282
	Máximo	8,612	8,622	8,621	8,580

(Continua)

(Continuação)

		Renda	Terra	Trabalho	Tecnologia
Sudeste Número de municípios = 1.598	Mínimo	1,609	0,693	3,689	5,819
	1ª quartil	7,519	7,277	7,748	7,594
	Mediana	8,035	7,881	8,182	7,864
	3ª quartil	8,348	8,229	8,446	8,100
	Máximo	8,620	8,617	8,622	8,550
Sul Número de municípios = 1.184	Mínimo	4,394	3,555	4,127	6,009
	1ª quartil	7,962	7,121	7,502	7,862
	Mediana	8,254	7,799	7,999	8,058
	3ª quartil	8,438	8,209	8,309	8,213
	Máximo	8,616	8,619	8,619	8,544

Fonte: Microdados do Censo Agropecuário 2017.

Elaboração dos autores.

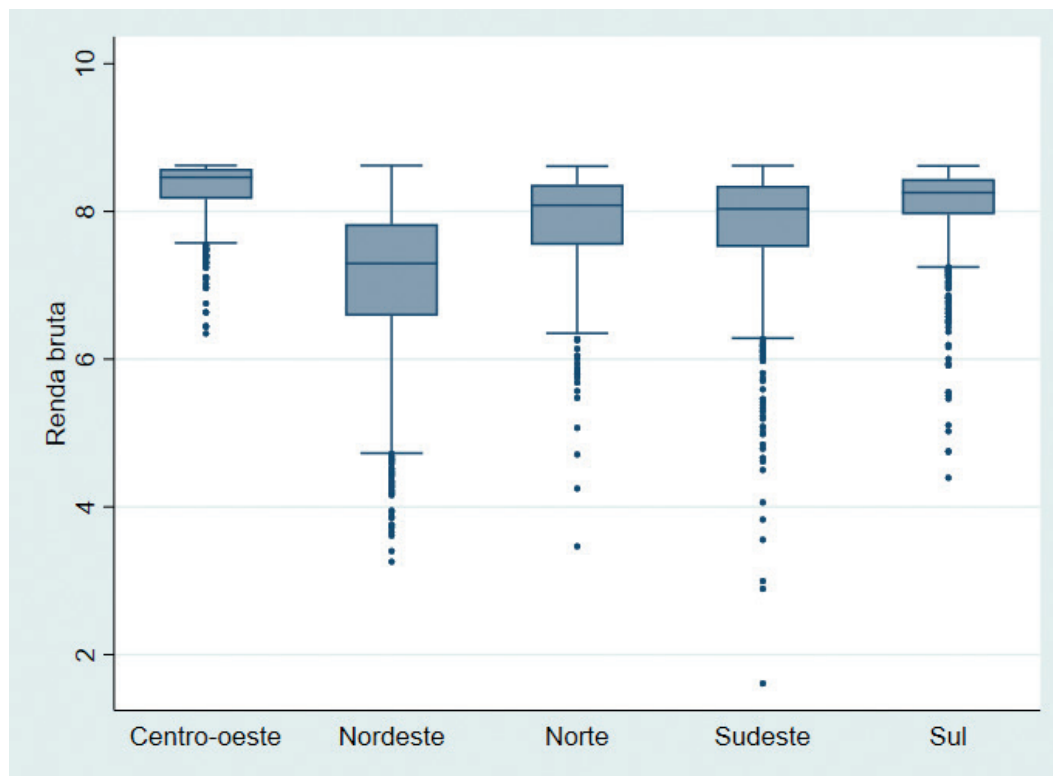
Obs.: 1. A compilação dos microdados foi feita na sala de acesso restrito do Instituto Brasileiro de Geografia e Estatística (IBGE). A base municipal compilada foi disponibilizada em 20 de fevereiro de 2020 pelo instituto aos autores (Processo nº 0001727.00000215/2018-13).

2. Renda: *log do rank* da receita bruta. Terra: *log do rank* da soma da área utilizada na agricultura com a terra arrendada. Trabalho: *log do rank* da soma da despesa com salário com a despesa com contratação de serviços. Tecnologia: *log do rank* do construto definido conforme seção 2.

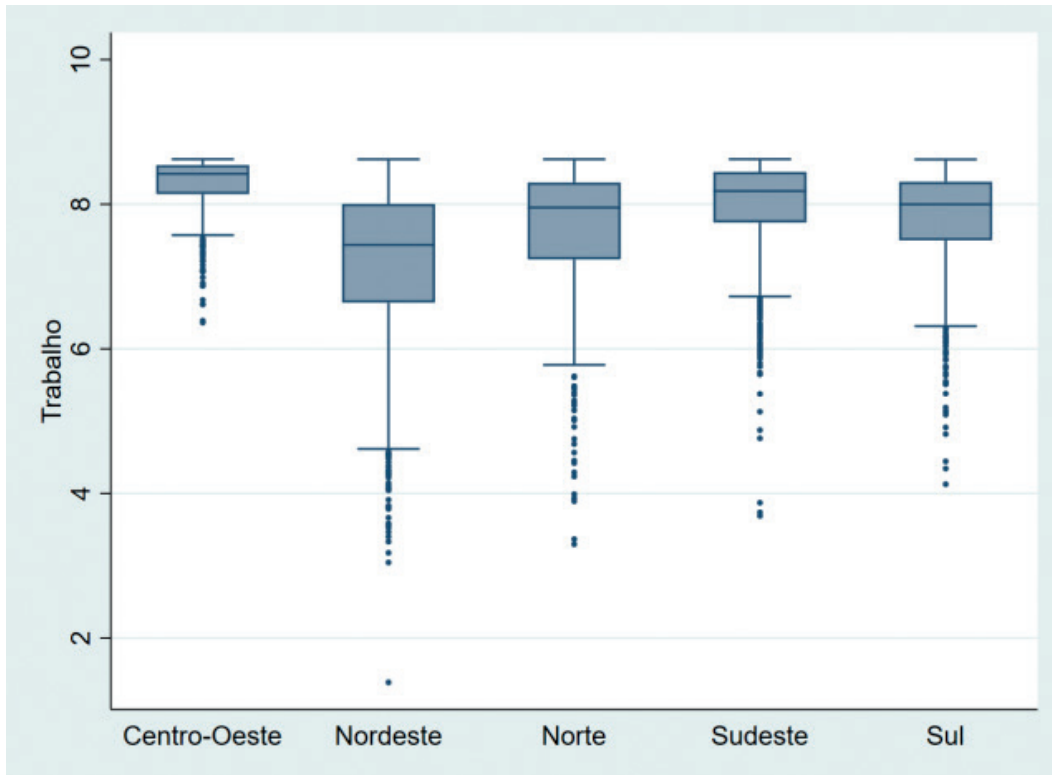
GRÁFICO 1

Diagramas de caixa (*boxplots*) das variáveis de produção – Grandes Regiões

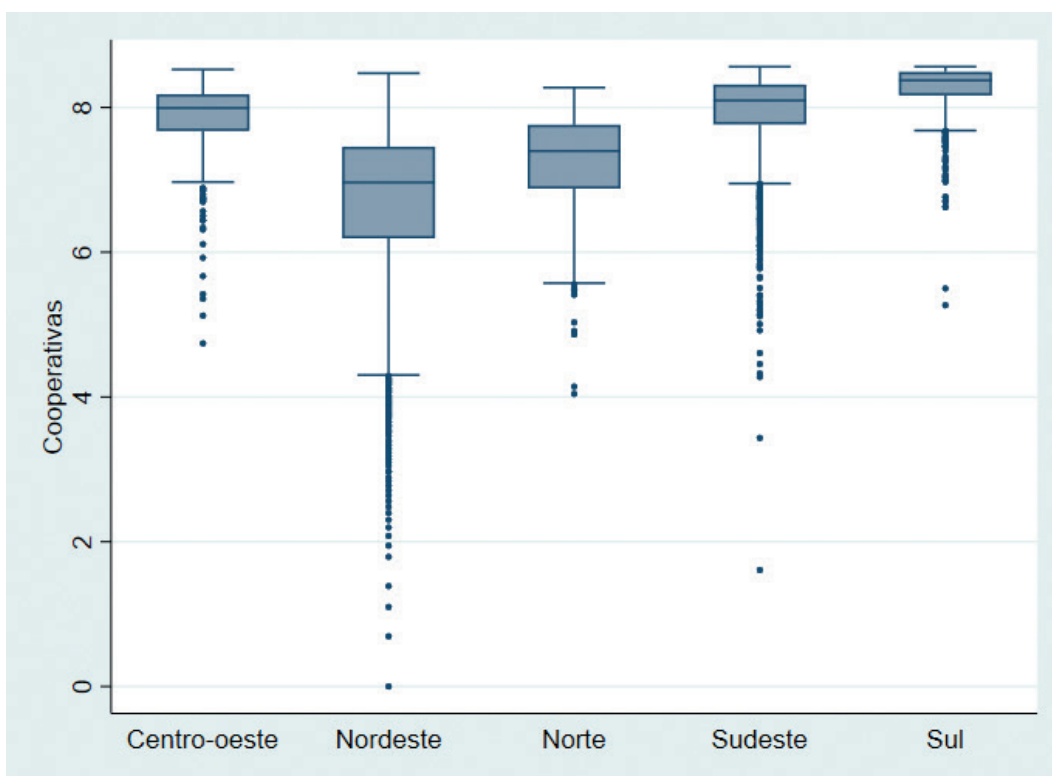
1A – Renda bruta



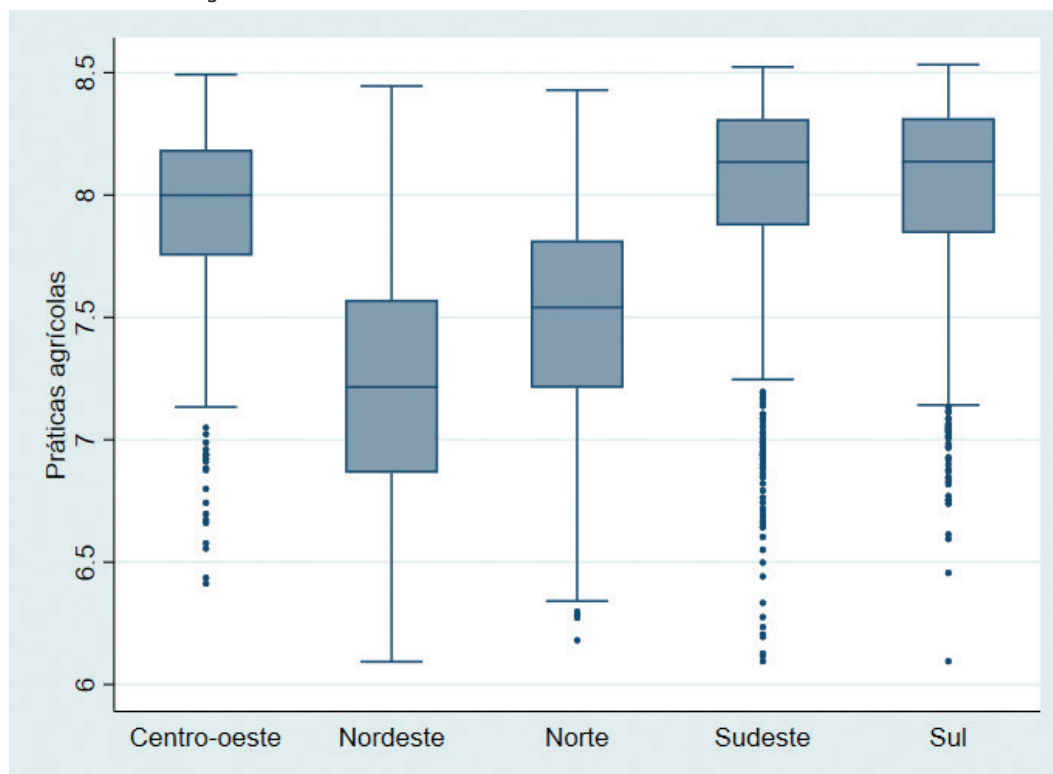
1B – Trabalho



1C – Cooperativas



1D – Práticas agrícolas



Fonte: Microdados do Censo Agropecuário 2017.

Elaboração dos autores.

Obs.: 1. A compilação dos microdados foi feita na sala de acesso restrito do IBGE. A base municipal compilada foi disponibilizada em 20 de fevereiro de 2020 pelo instituto aos autores (Processo nº 0001727.00000215/2018-13).

2. Figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Na tabela 2, observa-se o domínio da região Sul em quase todas as covariáveis, seguida pela região Sudeste. A exceção é na variável crédito, na qual o Sul é dominado pelo Centro-Oeste. A região Nordeste, tipicamente, mostra as pontuações mais baixas, exceto em crédito, para o qual o Norte tem as pontuações mais baixas. Do ponto de vista da variabilidade, o Sul mostra os menores valores para crédito, participação em cooperativas e assistência técnica. A região Norte tem a maior variabilidade no crédito. As variabilidades nas variáveis cooperativismo e alfabetização são maiores no Nordeste. As medidas de variabilidade para assistência técnica e meio ambiente são equivalentes para as regiões Norte e Nordeste. O gráfico 2 apresenta as variáveis contextuais.

TABELA 2
Resumo de cinco números para as variáveis contextuais – Grandes Regiões

	Cooperativas	Crédito	Alfabetização	Assistência técnica	Práticas agrícolas	
Centro-Oeste Número de municípios = 458	Mínimo	4,741	3,664	5,517	5,273	6,412
	1º quartil	7,677	7,468	7,805	7,460	7,753
	Mediana	7,992	7,955	8,010	7,822	7,999
	3º quartil	8,181	8,255	8,197	8,142	8,185
	Máximo	8,525	8,563	8,562	8,561	8,492
Nordeste Número de municípios = 1.573	Mínimo	0,000	0,693	0,000	0,693	6,093
	1º quartil	6,192	6,982	6,038	6,188	6,866
	Mediana	6,964	7,629	6,736	6,983	7,215
	3º quartil	7,457	8,040	7,175	7,518	7,572
	Máximo	8,474	8,539	8,159	8,508	8,445

(Continua)

(Continuação)

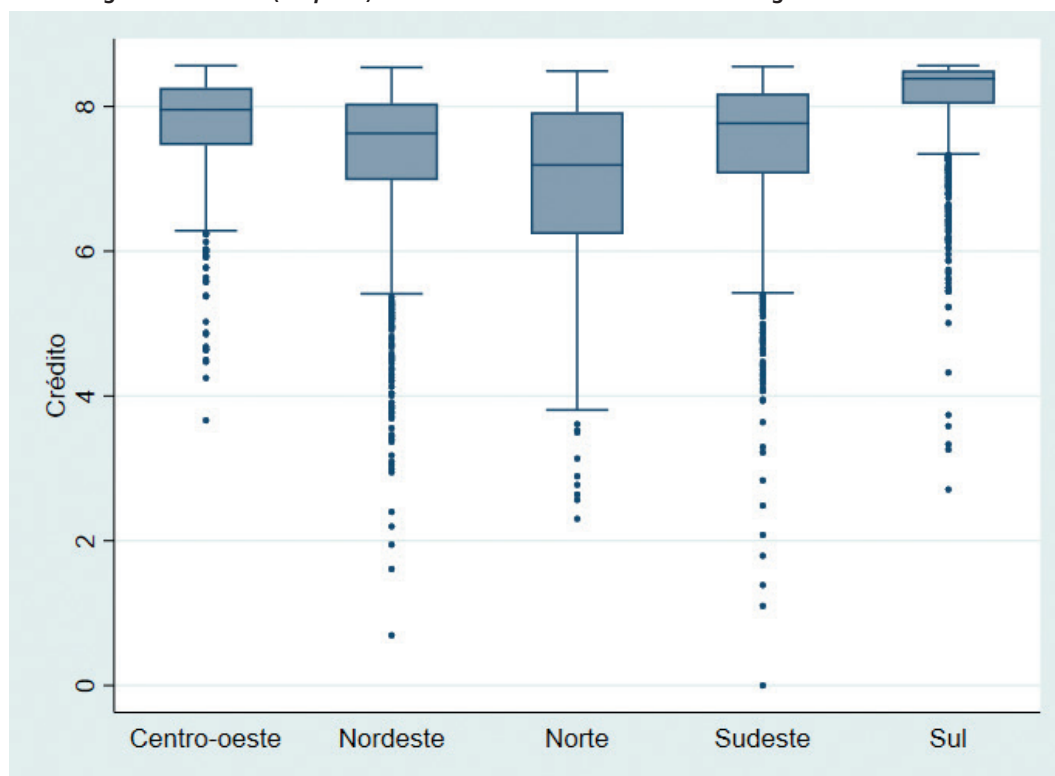
		Cooperativas	Crédito	Alfabetização	Assistência técnica	Práticas agrícolas
Norte Número de municípios = 423	Mínimo	4,043	2,303	2,079	0,000	6,181
	1º quartil	6,880	6,237	7,351	6,738	7,215
	Mediana	7,398	7,191	7,559	7,348	7,541
	3º quartil	7,757	7,918	7,733	7,733	7,814
	Máximo	8,275	8,489	8,496	8,460	8,428
Sudeste Número de municípios = 1.598	Mínimo	1,609	0,000	4,691	2,565	6,095
	1º quartil	7,769	7,072	7,846	7,770	7,875
	Mediana	8,096	7,768	8,144	8,099	8,135
	3º quartil	8,315	8,180	8,350	8,330	8,310
	Máximo	8,563	8,552	8,562	8,563	8,523
Sul Número de municípios = 1.184	Mínimo	5,268	2,708	7,485	5,700	6,095
	1º quartil	8,167	8,036	8,105	8,123	7,845
	Mediana	8,375	8,384	8,323	8,345	8,137
	3º quartil	8,490	8,498	8,472	8,480	8,314
	Máximo	8,563	8,563	8,562	8,563	8,532

Fonte: Microdados do Censo Agropecuário 2017.

Elaboração dos autores.

Obs.: A compilação dos microdados foi feita na sala de acesso restrito do IBGE. A base municipal compilada foi disponibilizada em 20 de fevereiro de 2020 pelo instituto aos autores (Processo nº 0001727.00000215/2018-13).

GRÁFICO 2

Diagramas de caixa (*boxplots*) das variáveis contextuais – Grandes Regiões

Fonte: Microdados do Censo Agropecuário 2017.

Elaboração dos autores.

Obs.: 1. A compilação dos microdados foi feita na sala de acesso restrito do IBGE. A base municipal compilada foi disponibilizada em 20 de fevereiro de 2020 pelo instituto aos autores (Processo nº 0001727.00000215/2018-13).

2. Figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

4.2 Estimação da fronteira estocástica

Considere-se a especificação Cobb-Douglas⁴ seguinte da função de produção:

$$\begin{aligned} \ln(y_j) = & \ln(C) + \beta_1 \ln(x_{1j}) + \beta_2 \ln(x_{2j}) + \beta_3 \ln(x_{3j}) + \\ & w_1 \ln(g_{1j}) + w_2 \ln(g_{2j}) + w_3 \ln(g_{3j}) + w_4 \ln(g_{4j}) \\ & + w_5 d_{1j} + w_6 d_{2j} + w_7 d_{3j} + w_8 d_{4j} \end{aligned} \quad (5)$$

em que os x 's representam os fatores de produção terra, trabalho e tecnologia, os g 's representam as variáveis contextuais cooperativas, crédito, alfabetização e assistência técnica e os d 's são variáveis indicadoras das regiões geográficas Norte, Nordeste, Sudeste e Sul. O Centro-Oeste não é considerado para eliminar singularidade da matriz de *dummies*.

No caso exponencial e meia-normal, a variância da componente de ineficiência (σ_u^2) é uma função de todas as covariáveis g 's, d 's, e práticas agrícolas. A variância idiossincrática (σ^2) é uma função dos d 's. No caso da distribuição normal truncada para a componente de ineficiência, supõe-se um modelo com média dependente de todas as variáveis contextuais. Os valores obtidos para a função *log-verossimilhança* calculada nos estimadores ótimos foram -1.442,53, -1.619,16 e -1.675,04 para as especificações exponencial, meia-normal e normal truncada, respectivamente. A melhor representação é dada, portanto, pela combinação normal-exponencial. A tabela 3 mostra os resultados do processo de estimação.

A variável *crédito* tem sinal negativo nas especificações da fronteira e da ineficiência. A indicação aqui é que o crédito excessivo pode diminuir a renda inicialmente, mas o aumento da renda resultante do aumento maior da eficiência pode equilibrar o efeito total, implicando um efeito líquido positivo final na renda. Todas as outras covariáveis mostram os sinais esperados e afetam significativamente a resposta. A assistência técnica não é significativa na média da ineficiência, mas é altamente significativa e positiva na fronteira de produção.

R^2 para o ajuste da fronteira é de 0,803 sem a inclusão da componente estocástica. Levando-se em conta o ajuste da componente de erro exponencial, obtém-se $R^2=0,792$. Nenhuma das três especificações estocásticas da fronteira pode ser rejeitada em favor do modelo de fronteira pura. Nas três especificações, as variâncias respectivas diferem significativamente de zero. A fronteira mostra retornos estritamente decrescentes de escala ($0,868 \pm 0,017$). A intensidade do retorno representa um aumento discreto em relação ao Censo Agropecuário 2006 ($0,777 \pm 0,014$), conforme reportado por Souza e Gomes (2019) e, neste contexto, os Censos Agropecuários 2006 e 2017 são compatíveis e enfatizam a correção da abordagem aqui adotada para a definição das *proxies* das variáveis de produção. A elasticidade dominante pertence à tecnologia, seguida pelo trabalho e pela terra. As elasticidades relativas, como em 2006, são dominadas pela tecnologia ($53,3\% \pm 1,9\%$), seguida pela mão de obra ($41,0\% \pm 1,5\%$) e pela terra ($5,7\% \pm 1,0\%$). Em relação aos dados de 2006 (Souza e Gomes, 2019), a tecnologia permaneceu aproximadamente constante ($58,7\% \pm 2,2\%$), a mão de obra aumentou ($29,7\% \pm 1,6\%$) e a terra diminuiu acentuadamente ($11,6\% \pm 1,8\%$). Isso mostra a agricultura brasileira menos intensiva em terras, mantendo constante o já elevado nível de tecnologia e migrando para uma força de trabalho mais qualificada.

4. Não testamos a adequabilidade da especificação Cobb-Douglas contra formas mais gerais. A rigor, isso deveria ser feito contra a única forma realmente flexível que é a forma flexível de Fourier. Isso envolveria o uso de restrições paramétricas artificiais e não permitiria comparações com outros estudos como em Souza e Gomes (2019). Optamos por fixar a forma mais simples da superfície resposta e testar a adequabilidade do conjunto superfície resposta-distribuição da eficiência.

TABELA 3
Resultados da análise de fronteira estocástica – modelo normal-exponencial

	Coefficiente	Desvio-padrão	z	P> z	Intervalo de confiança a 95%	
Função de produção						
Trabalho	0,3556	0,0120	29,52	0,000	0,3320	0,3792
Terra	0,0496	0,0086	5,75	0,000	0,0327	0,0665
Tecnologia	0,4630	0,0230	20,17	0,000	0,4180	0,5080
Cooperativas	0,0312	0,0080	3,89	0,000	0,0155	0,0470
Crédito	-0,0395	0,0068	-5,81	0,000	-0,0528	-0,0262
Alfabetização	0,0431	0,0106	4,08	0,000	0,0224	0,0639
Assistência técnica	0,0516	0,0085	6,09	0,000	0,0350	0,0682
Norte	0,0257	0,0322	0,80	0,425	-0,0374	0,0889
Nordeste	-0,1412	0,0257	-5,49	0,000	-0,1916	-0,0907
Sul	0,0967	0,0134	7,24	0,000	0,0705	0,1228
Sudeste	0,0027	0,0130	0,21	0,835	-0,0228	0,0283
Constante	0,6064	0,1730	3,50	0,000	0,2672	0,9455
$\ln \sigma^2$						
Norte	2,5728	0,1706	15,08	0,000	2,2384	2,9072
Nordeste	2,4606	0,1580	15,57	0,000	2,1509	2,7703
Sul	1,3411	0,1452	9,24	0,000	1,0565	1,6256
Sudeste	1,5475	0,1379	11,23	0,000	1,2773	1,8176
Constante	-4,9330	0,1287	-38,33	0,000	-5,1853	-4,6808
$\ln \sigma_u^2$						
Cooperativas	-0,1856	0,0598	-3,10	0,002	-0,3029	-0,0684
Crédito	-0,4890	0,0451	-10,84	0,000	-0,5774	-0,4006
Alfabetização	-0,2983	0,0655	-4,55	0,000	-0,4267	-0,1699
Assistência técnica	0,0519	0,0542	0,96	0,338	-0,0543	0,1580
Práticas agrícolas	-1,0153	0,0903	-11,24	0,000	-1,1923	-0,8384
Norte	-0,4669	0,4169	-1,12	0,263	-1,2841	0,3502
Nordeste	0,9556	0,2145	4,46	0,000	0,5352	1,3760
Sul	0,7271	0,1773	4,10	0,000	0,3796	1,0745
Sudeste	1,7010	0,1602	10,62	0,000	1,3871	2,0149
Constante	11,1292	0,8944	12,44	0,000	9,3761	12,8822

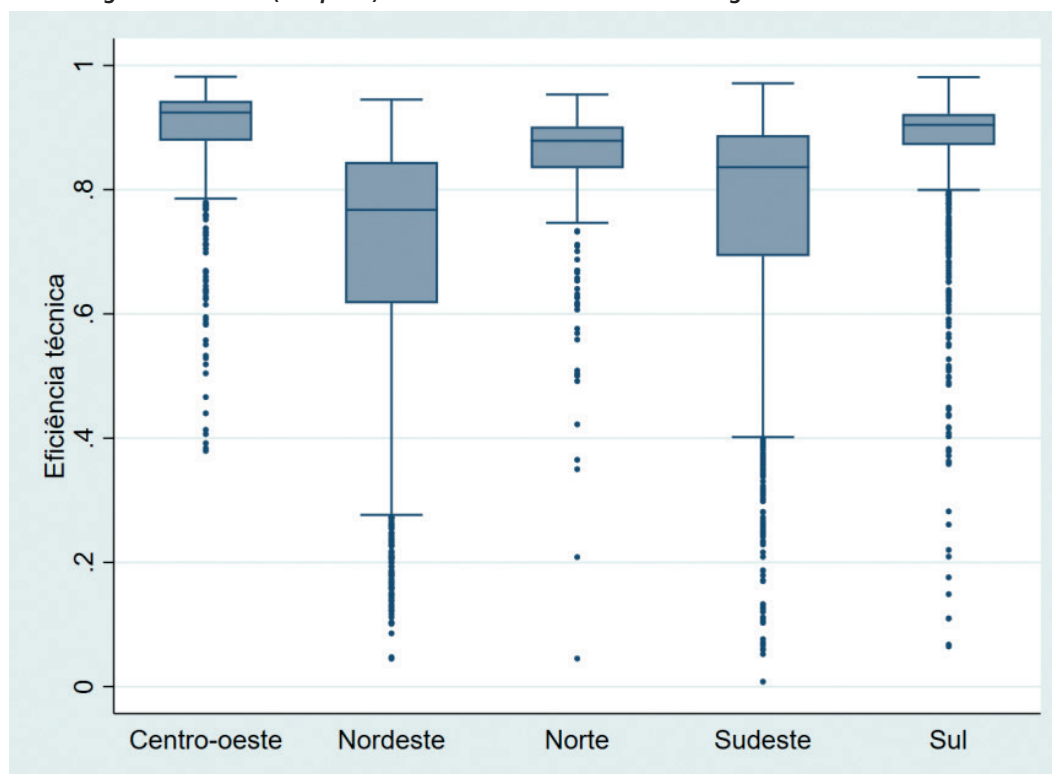
Fonte: Microdados do Censo Agropecuário 2017.
Elaboração dos autores.

Obs.: A compilação dos microdados foi feita na sala de acesso restrito do IBGE. A base municipal compilada foi disponibilizada em 20 de fevereiro de 2020 pelo instituto aos autores (Processo nº 0001727.00000215/2018-13).

O gráfico 3 mostra a distribuição da eficiência técnica. A região Centro-Oeste domina (mediana = 0,923), seguida pelas regiões Sul (0,901), Norte (0,882), Sudeste (0,834) e Nordeste (0,763). As maiores variações ocorrem nas regiões Sudeste e Nordeste. Esses resultados não são semelhantes aos de 2006. O Sul foi mais eficiente que o Centro-Oeste em 2006 (Souza e Gomes, 2019). A região menos eficiente pelas análises dos dados dos Censos Agropecuários 2006 e 2017 foi o Nordeste. Também se nota um aumento relativo da eficiência no Norte e uma diminuição relativa no Sudeste.

Os níveis tecnológicos medidos pelas variáveis indicadoras regionais mostram o menor valor para o Nordeste. O melhor valor é para o Sul. As outras regiões mostram valores semelhantes.

GRÁFICO 3
Diagramas de caixa (*box plots*) da eficiência técnica – Grandes Regiões



Fonte: Microdados do Censo Agropecuário 2017.

Elaboração dos autores.

Obs.: 1. A compilação dos microdados foi feita na sala de acesso restrito do IBGE. A base municipal compilada foi disponibilizada em 20 de fevereiro de 2020 pelo instituto aos autores (Processo nº 0001727.00000215/2018-13).

2. Figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

5 OBSERVAÇÕES FINAIS

Ajustou-se um modelo de fronteira estocástica aos dados municipais observados no Censo Agropecuário 2017. Além de avaliar as elasticidades dos insumos, este estudo teve como objetivo investigar os efeitos das variáveis associadas a imperfeições de mercado na produção. As imperfeições de mercado provêm de diferentes realidades na produção, as quais são vivenciadas por pequenos e grandes agricultores. Postulou-se que o acesso ao crédito, a associação em cooperativas, o nível de alfabetização, o acesso à assistência técnica e as práticas agrícolas ecológicas são os principais fatores necessários para reduzir as imperfeições de mercado e promover a melhoria da renda.

A agricultura brasileira, em nível municipal, continuou com alta concentração de renda em 2017, medida pelo índice de Gini. As elasticidades relativas são dominadas pela tecnologia, seguida pela mão de obra, indicando que, como em 2006, os fatores para a inclusão produtiva e a redução da pobreza são a tecnologia e uma força de trabalho qualificada. Parte da elasticidade relativa do uso da terra foi transferida para o trabalho, indicando uma redução adicional na importância da terra em 2017 em relação a 2006. Para a tomada de decisões em políticas públicas, é importante aumentar a alfabetização, o acesso ao crédito e a associação em cooperativas, incentivar boas práticas ambientais agrícolas e fortalecer a assistência técnica. Esta última será viável com imperfeições de mercado reduzidas via a extensão rural.

No geral, os dados do Censo Agropecuário 2017 indicam que o Centro-Oeste é a região com melhor desempenho e o Nordeste, com o pior. Boas práticas agrícolas (ambientais) mostram a maior variabilidade entre as regiões. Essas incluem as variáveis plantio em curvas de nível, rotação de culturas, descanso do solo, conservação de encostas e taludes, manejo florestal e preparação do solo. Vê-se aqui espaço para um aumento significativo de eficiência e, portanto, indiretamente, de renda.

REFERÊNCIAS

- AIGNER, D.; LOVELL, C. A. K.; SCHIMIDT, P. Formulation and estimation of stochastic frontier production function models. **Journal of Econometrics**, v. 6, p. 21-37, 1977.
- ALVES, E.; SOUZA, G. S. Pequenos estabelecimentos em termos de área também enriquecem? Pedras e tropeços. **Revista de Política Agrícola**, v. 24, p. 7-21, 2015.
- COELLI, T. J.; RAO, D. S. P.; O'DONNELL, C. J. **An introduction to efficiency and productivity analysis**. 2nd ed. New York: Springer, 2005. 367 p.
- CONOVER, W. J. **Practical nonparametric statistics**. 3rd ed. New York: Wiley, 1999. 584 p.
- GREENE, W. H. **Econometric analysis**. 7th ed. New Jersey: Prentice Hall, 2011.
- KHUMBHAKAR, S. C.; LOVELL, C. A. K. **Stochastic frontier analysis**. Cambridge: Cambridge University Press, 2000. 344 p.
- SOUZA, G. S.; GOMES, E. G. A stochastic production frontier analysis of the Brazilian agriculture in the presence of an endogenous covariate. *In*: PARLIER, G.; LIBERATORE, F.; DEMANGE, M. (Orgs.). **Operations research and enterprise systems**. 1st ed. Cham: Springer, n. 966, p. 3-14, 2019.
- SOUZA, G. S.; GOMES, E. G.; ALVES, E. R. A. Conditional FDH efficiency to assess performance factors for Brazilian agriculture. **Pesquisa Operacional**, v. 37, p. 93-106, 2017.
- STATA. **Stata base reference manual release 16**. College Station: Stata Press, 2019. 2911 p.