

## Capítulo 6

# Determinantes da concentração da renda rural no Cerrado

Geraldo da Silva e Souza

Eliane Gonçalves Gomes

Eliseu Roberto de Andrade Alves

## Introdução

Ao analisar os dados do Censo Agropecuário 2006, Alves et al. (2013) verificaram a existência de forte concentração na renda rural e relataram que 11% dos estabelecimentos (cerca de 500 mil estabelecimentos) respondem por 87% do valor total da produção rural. Embora tal nível de concentração seja também observado em outros países em termos percentuais (11% nos Estados Unidos e de 14% na Europa), é de interesse a identificação de fatores associados à concentração da renda rural no Brasil, medida pelo índice de Gini. Parece óbvio que a inclusão do grupo de produtores excluídos do processo de produção mais rentável aumentaria substancialmente a produção agropecuária brasileira, favorecendo a competitividade dessa produção nos mercados interno e externo.

Segundo Alves e Souza (2015), há evidências de que a concentração de renda resulta do uso da tecnologia. Outros fatores de importância que causam concentração estão relacionados a imperfeições de mercado que criam barreiras ao acesso e ao uso de tecnologia. Alves e Souza (2015) também enfatizaram que as imperfeições de mercado forçam os pequenos produtores a vender sua produção a preços menores do que deveriam em face de seus custos de produção. Tipicamente, o pequeno produtor é obrigado a comprar insumos sob condições desfavoráveis.

As imperfeições de mercado resultam das condições de desenvolvimento no entorno da propriedade rural. São resultantes de assimetrias existentes no acesso ao crédito para a produção, infraestrutura logística e social, disponibilidade e acesso à informação, extensão rural e assistência técnica, entre outros. É uma condição geral que afeta toda agricultura brasileira, em particular no Cerrado.

Outros estudos sobre fatores determinantes da concentração da renda rural no Brasil são, por exemplo, Neder e Silva (2004), Ney e Hoffmann (2008, 2009), Helfand et al. (2009) e Souza et al. (2016). Nossa contribuição a essa literatura está na caracterização estatística da dependência entre concentração de renda e fatores contextuais – tecnologia e condições de entorno dos estabelecimentos – para o Cerrado.

Esses resultados servem ao propósito de orientar ações de políticas públicas e assistência técnica para o bioma. A variável dependente na análise é o índice de Gini calculado em cada município, tomando-se como base a distribuição da renda bruta dos estabelecimentos rurais. Como *proxy* para tecnologia, utiliza-se um escore de desempenho municipal, definido por meio de um modelo DEA (do inglês *Data Envelopment Analysis*) com retornos variáveis à escala. Na modelagem DEA, a renda bruta é a variável produto e os insumos são gastos com terra, trabalho e capital. Medições de DEA, portanto, refletem eficiência econômica. Como fatores representativos de imperfeições de mercados, consideram-se indicadores de meio ambiente, educação, saúde, infraestrutura e demografia. A variável DEA é potencialmente endógena.

A importância de estudar o bioma Cerrado separadamente está em sua representatividade para a agricultura brasileira. De acordo com o Censo Agropecuário 2006, o Cerrado abrange 15% dos estabelecimentos rurais, concentra 32% da renda bruta rural do País e 37% da terra agricultável.

## Variáveis de produção, variáveis contextuais e dados

A fonte de dados principal utilizada é o Censo Agropecuário 2006. Os dados de produção para a modelagem DEA foram agregados por município e transformados em ranks. Têm-se dados válidos para 1.098 municípios, que representa 99,6% do total de municípios no bioma Cerrado e 19,8% do número total de municípios do Brasil. No modelo DEA, a unidade tomadora de decisão (DMU) é o município. A variável resposta do modelo de produção é o rank da renda bruta municipal total e as variáveis de insumos são os ranks dos totais municipais das despesas dos estabelecimentos com terra, trabalho e capital, conforme descrição da Tabela 1 (Souza et al., 2016).

As variáveis contextuais consideradas são a medida de eficiência econômica DEA e os indicadores descritos a seguir, que seguem a mesma definição descrita em Souza et al. (2016) e em nota técnica não publicada da Confederação Nacional da Agricultura (CNA) (Confederação Nacional da Agricultura, 2013).

1. Índice de desempenho do Sistema Único de Saúde (SUS), como indicador de saúde. Este indicador afere o desempenho do SUS, levando em conta as dimensões de acesso aos serviços, de efetividade da atenção básica, ambulatorial, hospitalar e emergencial no município. Essa variável reflete tanto o acesso a bens públicos (saúde) quanto a melhor expectativa de vida da população do município.
2. Existência de energia elétrica (indicador de infraestrutura). Proporção de domicílios rurais dotados de energia elétrica no município.
3. Índice de desenvolvimento da educação básica (indicador de educação). É um indicador que avalia o desempenho dos estudantes em exames padronizados, combinado com informações sobre o rendimento escolar. No caso dessa variável, foram consideradas as notas médias obtidas por alunos da rede pública,

**Tabela 1.** Descrição das variáveis de produção.

Variável	Componente	Unidade	Observação
Produto	Renda bruta do estabelecimento	Real	Valor total da produção agropecuária do estabelecimento
Terra	4% do valor da terra	Real	Valor de aluguel da terra
Trabalho	Salários ou outras formas de compensação pagos à família e aos trabalhadores contratados	Real	-
Capital (insumos tecnológicos)	Maquinário, melhorias no estabelecimento, aluguéis de equipamentos, valor de culturas permanentes, valor dos animais, valor das matas plantadas, valor de sementes, valor de sal e forragens, valor de medicamentos, adubação e pesticidas, despesas com combustíveis, eletricidade, armazenagem, matérias-primas e outros gastos	Real	Todos os valores de estoques foram depreciados à taxa de 6% ao ano num período determinado que varia com o tipo do estoque: <ul style="list-style-type: none"> <li>• máquinas – 15 anos</li> <li>• matas plantadas – 20 anos</li> <li>• culturas permanentes – 15 anos</li> <li>• benfeitorias – 50 anos</li> <li>• animais – 5 anos</li> </ul>

Fonte: Souza et al. (2016).

tanto rural como urbana, do município em 2009. Esse índice reflete a qualidade educacional da população.

4. Assistência técnica (indicador de assistência técnica). Proporção de estabelecimentos que receberam assistência técnica.
5. Indicador demográfico (indicador demográfico). É um construto definido por meio de uma média ponderada de ranks de vários atributos. As variáveis componentes da dimensão demográfica captam os aspectos relacionados com a dinâmica da população que tende a acompanhar o desenvolvimento rural. São elas: razão entre a população rural e urbana do município; índice de migração; razão de dependência (i.e., participação relativa da população considerada inativa – 0 a 14 anos e 60 anos e mais de idade – na população potencialmente ativa – 15 a 59 anos de idade); taxa de envelhecimento rural (razão entre a população acima de 60 anos e a população total do município); e média de moradores por domicílio rural.
6. Indicador ambiental (indicador ambiental). É um construto definido por meio de uma média ponderada de ranks dos atributos: complementar da participação da área de pastagem plantada degradada e área degradada na área de lavouras, pecuária e terras degradadas; complementar da proporção de estabelecimentos que responderam as queimadas como prática agrícola utilizada; proporção de estabelecimentos que responderam o plantio em curva de nível como prática agrícola utilizada; proporção de estabelecimentos que responderam a rotação de culturas como prática agrícola utilizada; proporção de estabelecimentos com

sistema de preparo do solo por cultivo mínimo; proporção de estabelecimentos com sistema de preparo do solo por plantio direto; complementar da proporção de estabelecimentos que utilizam agroquímicos para controle de pragas e doenças; proporção de áreas de florestas e sistemas agroflorestais; e complementar da proporção de domicílios em que o lixo foi jogado em terreno baldio ou logradouro ou jogado em rio, lago ou mar.

Os indicadores ambiental e demográfico foram determinados via média ponderada de ranks normalizados pelo máximo, com pesos definidos por correlações múltiplas. Isto é, a  $i$ -ésima componente tem peso relativo  $R_i^2 / \sum_j R_j^2$ , sendo  $R_i^2$  o coeficiente de correlação múltipla obtido da regressão da  $i$ -ésima variável nas demais variáveis.

## Abordagem metodológica

Descrevem-se nesta seção os aspectos metodológicos desenvolvidos neste capítulo. Inicia-se a discussão com uma breve introdução ao conceito do índice de Gini, sua estimação e propriedades estatísticas. Na sequência, apresentam-se o modelo de produção e o conceito subjacente de eficiência de produção bem como os modelos econométricos utilizados na identificação dos fatores determinantes da concentração e da avaliação da endogeneidade da tecnologia.

### Índice de concentração de Gini

As medidas de concentração propõem-se a avaliar a variabilidade relativa de um atributo. Exemplos típicos nesse contexto são dados pelo coeficiente de variação e pelo índice de concentração de Gini. Os conceitos aqui apresentados seguem Kendall e Stuart (1977). Seja  $Y$  uma variável aleatória não negativa do tipo discreto ou absolutamente contínuo, com função massa de probabilidades (densidade)  $w(\cdot)$  concentrada em algum subconjunto de  $R^+$ . No caso absolutamente contínuo, define-se o desvio absoluto médio por

$$\Delta = \iint_{R^+ \times R^+} |u-v| w(u)w(v) du dv$$

No caso discreto, quando  $Y$  tem massa probabilística concentrada no subconjunto  $\{y_1, \dots, y_N\}$ , duas definições são possíveis:

$$\Delta = \frac{1}{N^2} \sum_{i=1}^N \sum_{k=1}^N |y_i - y_k| w(y_i)w(y_k)$$

ou

$$\Delta = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{k=1}^N |y_i - y_k| w(y_i)w(y_k)$$

A distinção entre os conceitos é irrelevante se  $N$  é grande.

O índice de concentração de Gini populacional ( $G$ ) é a medida de variação relativa que se obtém normalizando  $\Delta$  pelo valor esperado  $E(Y) = \mu$ , i.e.,  $G = \Delta/2\mu$ . Tem-se, assim,  $0 \leq G \leq 1$ . A quantidade  $G$  está associada ao conceito de área de concentração definida pela curva de Lorenz, conforme se segue. Sejam

$$F(x) = \int_0^x w(t)dt, \phi(x) = \frac{1}{\mu} \int_0^x tw(t)dt$$

É evidente que ambas as funções assumem valores no intervalo  $[0,1]$ . Seja  $\psi(\cdot)$  a função inversa de  $F(\cdot)$ .

Define-se

$$h(f) = \frac{1}{\mu} \int_0^{\psi(f)} tw(t)dt$$

para  $f$  em  $(0,1)$ . Esta é a função de Lorenz. Tem-se:

$$\frac{dh}{df} = \frac{\Psi(f)}{\mu} > 0$$

$$\frac{d^2h}{df^2} = \frac{1}{\mu w(\psi(f))} > 0$$

Segue que a função de Lorenz é crescente e convexa. O gráfico de  $h(\cdot)$  denomina-se curva de Lorenz.

A área delimitada pela reta  $h = f$  e a curva de Lorenz denominam-se área de concentração, representada aqui por  $S$ . Pode-se demonstrar (Kendall; Stuart 1977) que  $S = \Delta/4\mu$ . Portanto,  $S = G/2$ . Como  $0 \leq S \leq 1/2$ , segue que  $0 \leq G \leq 1$ .

Para uma amostra aleatória  $y_1, \dots, y_n$  de  $Y$ , os estimadores amostrais de  $G$  podem ser definidos como a seguir:

$$2\bar{y}\hat{G} = \begin{cases} g = \hat{\Delta} = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n |y_i - y_j| \\ \text{ou} \\ \hat{\Delta} = \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j=1}^n |y_i - y_j| = \frac{2}{n(n-1)} \sum_{i < j} |y_i - y_j| \end{cases}$$

Em ambos os casos,  $\hat{G}$  é um estimador consistente, mas viesado de  $G$ . Considerando o primeiro estimador, tem-se que

$$E(\hat{\Delta}) = \frac{2\sigma}{\sqrt{\pi}} \frac{n(n-1)}{n^2}$$

para a distribuição normal  $N(\mu, \delta^2)$  Para a distribuição exponencial com parâmetro  $\lambda > 0$ ,

$$E(\hat{\Delta}) = (1/\lambda) \frac{n(n-1)}{n^2}$$

Para a distribuição uniforme no intervalo  $[0, b]$ ,

$$E(\hat{\Delta}) = (b/3) \frac{n(n-1)}{n^2}$$

Aqui se segue Souza (1977) e faz-se uso da fórmula com

$$g = (1/n^2) \sum_{i=1}^n \sum_{j=1}^n |y_i - y_j|$$

## Modelo de produção e tecnologia

Considere um processo de produção composto de 1.098 DMUs (municípios). Cada DMU faz uso do vetor de insumos  $(x_1, x_2, x_3)$  – despesas com terra, trabalho e outros insumos tecnológicos (capital) – para produzir o nível de *output*  $y$  (renda rural bruta). As variáveis de produção são transformadas em ranks. Seja  $Y = (y_1, y_2, \dots, y_{1.098})$  o vetor produto e  $X$  a matriz  $3 \times 1.098$  de utilização de insumos. A  $r$ -ésima coluna de  $X$  define o vetor de insumos utilizado pela DMU  $r$  para produzir  $y_r$ .

A eficiência econômica aqui considerada  $\phi^*(x_o, y_o)$  – medida DEA de desempenho para o município  $o$  com vetor de produção  $(x_o, y_o)$ , orientada a produto com retornos variáveis à escala (Banker et al., 1984) – é dada por:

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

sujeito a

$$a) Y\lambda \geq \phi y_o$$

$$b) X\lambda \leq x_o$$

$$c) \lambda > 0, \lambda 1 = 1, \phi \text{ livre}$$

Observa-se que a transformação em ranks é usual em estatística não paramétrica (Conover, 1999) e reduz a influência de observações atípicas no modelo DEA, além de manter a monotonicidade do escore de desempenho medido pela eficiência DEA. Neste estudo, considerou-se que a medida de desempenho como definida anteriormente cristaliza o efeito tecnológico na variável resposta (índice de concentração). Para mais detalhes sobre uso do DEA, ver Cooper et al. (2011).

## Análise estatística

Os modelos estatísticos que serão utilizados exploram o fato de que o índice de Gini assume intervalos de valores em (0,1) para distribuições com massa probabilística não concentrada em uma constante e que a medida DEA é potencialmente endógena. Faz-se uso conjunto de regressão fracionária (Ramalho et al., 2010) com o método de momentos generalizado – GMM (Gallant, 1987; Davidson; MacKinnon, 1993; Greene, 2011).

Seja  $\hat{\theta}_\tau$  o índice de Gini para o município,  $\tau$  e  $w_\tau$  o vetor de variáveis contextuais, incluindo tecnologia. Supõe-se que  $\hat{\theta}_\tau = H(w_\tau \delta) + \varepsilon_\tau$ , em que  $H(\cdot)$  é uma função de distribuição de probabilidades. Tal dependência permite o estudo do comportamento do índice de concentração como função monótona do construto linear  $w_\tau \delta$ . O vetor paramétrico  $\delta$  é estimado por GMM. Essa abordagem engloba a potencial endogeneidade de qualquer variável, em particular da *proxy* de tecnologia. Como vetor de variáveis instrumentais  $z_\tau$  faz-se uso de indicadores de desenvolvimento municipal. A condição de momentos é dada pela expressão  $E\{z_\tau \otimes [\theta_\tau - H(w_\tau \delta)]\} = 0$ , onde  $\otimes$  denota produto direto. Escolhas típicas para  $H(\cdot)$  obtêm-se das distribuições logística, normal padrão e inversa da distribuição do valor extremo, para as quais  $H(u) = e^u/(1 + e^u)$ ,  $H(u) = \Phi(u)$ , e  $H(u) = 1 - e^{-e^u}$  respectivamente. A função  $\Phi(u)$  é a função de distribuição normal padrão.

Os efeitos marginais das covariáveis são calculados com a utilização das fórmulas seguintes.

$$\partial H(\mu_\tau) / \partial x = [H'(\mu_\tau)] \times (\partial \mu_\tau / \partial x)$$

$$H'(\mu_\tau) = \exp[\mu_\tau - \exp(\mu_\tau)]$$

$$\mu_\tau = w_\tau \delta$$

O índice de Gini e a *proxy* para tecnologia captam as mesmas imperfeições de mercado e, portanto, espera-se uma associação significativa e positiva entre essas variáveis. A hipótese básica é que valores elevados da eficiência são indicativos de acesso à informação e à tecnologia, e esses componentes estariam disponíveis apenas para os produtores com alto nível de renda.

## Resultados estatísticos

Na Tabela 2 mostram-se as correlações de Spearman entre todos os pares de variáveis consideradas na análise. Todas as correlações são não negativas, o que sugere uma relação monotônica entre o índice de Gini e o construto linear  $w_r \delta$ .

**Tabela 2.** Correlações de Spearman.

	Gini	Tec	Ambi	Demo	Infra	Assitec	Saúde	Educa
Gini	1,0000	0,4596	0,1649	0,1588	0,0521	0,1139	0,0149	0,0959
		<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>0,1273</i>	<i>0,0008</i>	<i>0,6625</i>	<i>0,0049</i>
Tec	0,4596	1,0000	0,2379	0,5191	0,3585	0,4366	0,0052	0,3592
	<i>&lt;0,0001</i>		<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>0,8799</i>	<i>&lt;0,0001</i>
Ambi	0,1649	0,2379	1,0000	0,4451	0,3410	0,4178	0,2658	0,4243
	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>		<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>
Demo	0,1588	0,5191	0,4451	1,0000	0,5130	0,6967	0,1806	0,5769
	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>		<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>
Infra	0,0521	0,3585	0,3410	0,5130	1,0000	0,5904	0,3752	0,6678
	<i>0,1273</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>		<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>
Assitec	0,1139	0,4366	0,4178	0,6967	0,5904	1,0000	0,2946	0,6224
	<i>0,0008</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>		<i>&lt;0,0001</i>	<i>&lt;0,0001</i>
Saúde	0,0149	0,0052	0,2658	0,1806	0,3752	0,2946	1,0000	0,4579
	<i>0,6625</i>	<i>0,8799</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>		<i>&lt;0,0001</i>
Educa	0,0959	0,3592	0,4243	0,5769	0,6678	0,6224	0,4579	1,0000
	<i>0,0049</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	<i>&lt;0,0001</i>	

Gini = índice de concentração; Tec = tecnologia (DEA score); Ambi = indicador ambiental; Demo = indicador demográfico; Infra = infraestrutura; Assitec = assistência técnica; Saúde = indicador de saúde; Educa = indicador de educação; p-valores em itálico.

Apresentam-se a seguir análises de variância com investigação dos efeitos regionais estimados para cada variável contextual. A região Sul foi eliminada por conter apenas um município no bioma Cerrado. Nas Tabelas 3 e 4 mostra-se a análise de variância da variável tecnológica – score DEA. A análise indica diferenças significativas entre regiões. A maior média é para a região Centro-Oeste, seguida de Sudeste, Norte e Nordeste. Todos os pares de diferenças entre regiões são significativos.

Nas Tabelas 5 e 6 verifica-se a análise de variância do indicador ambiental. A análise indica diferenças significativas entre regiões. A maior média é para a região Centro-Oeste, seguida de Sudeste, Norte e Nordeste. Todos os pares de diferenças entre regiões são significativos.



**Tabela 3.** Análise de variância – escore DEA (tecnologia).

Fonte	Graus de liberdade	Soma de quadrados	Quadrado médio	F	Pr > F
Regiões	3	14,0034	4,6678	72,89	<0,0001
Erro	855	54,7556	0,0640	-	-
Total corrigido	858	68,7590	-	-	-

**Tabela 4.** Médias de mínimos quadrados por região – escore DEA (tecnologia).

Região	Médias de mínimos quadrados	Desvio-padrão
Centro-Oeste	0,7498	0,0242
Norte	0,3210	0,0234
Nordeste	0,4454	0,0181
Sudeste	0,6060	0,0121

**Tabela 5.** Análise de variância – indicador ambiental.

Fonte	Graus de liberdade	Soma de quadrados	Quadrado médio	F	Pr > F
Regiões	3	2,3280	0,7760	142,38	<0,0001
Erro	855	4,6599	0,0055	-	-
Total corrigido	858	6,9879	-	-	-

**Tabela 6.** Médias de mínimos quadrados por região – indicador ambiental.

Região	Médias de mínimos quadrados	Desvio-padrão
Centro-Oeste	0,5578	0,0071
Norte	0,4859	0,0068
Nordeste	0,4158	0,0053
Sudeste	0,5361	0,0035

Nas Tabelas 7 e 8 apresenta-se o resultado da análise de variância para o indicador demográfico. O padrão é similar aos observados anteriormente.

**Tabela 7.** Análise de variância – indicador demográfico.

Fonte	Graus de liberdade	Soma de quadrados	Quadrado médio	F	Pr > F
Regiões	3	8,1971	2,73244448	203,27	<0,0001
Erro	855	11,4930	0,0134	-	-
Total Corrigido	858	19,6900	-	-	-

**Tabela 8.** Médias de mínimos quadrados por região – indicador demográfico.

Região	Médias de mínimos quadrados	Desvio-padrão
Centro-Oeste	0,6768	0,0111
Norte	0,4916	0,0107
Nordeste	0,3761	0,0083
Sudeste	0,5789	0,0056

A análise de variância para o indicador educacional pode ser depreendida das Tabelas 9 e 10. Nesta componente a maior média é observada na região Sudeste, seguidas de Centro-Oeste, Norte e Nordeste. Todas as diferenças são significativas.

**Tabela 9.** Análise de variância – indicador de educação.

Fonte	Graus de liberdade	Soma de quadrados	Quadrado médio	F	Pr > F
Regiões	3	40,5067	13,5022	481,98	<0,0001
Erro	855	23,9520	0,0280	-	-
Total corrigido	858	64,4587	-	-	-

**Tabela 10.** Médias de mínimos quadrados por região – indicador de educação.

Região	Médias de mínimos quadrados	Desvio-padrão
Centro-Oeste	0,5238	0,0160
Norte	0,4071	0,0155
Nordeste	0,22836	0,01196
Sudeste	0,75184	0,00801

O indicador de saúde também difere entre regiões (Tabelas 11 e 12). A maior média está na região Sudeste, seguida das regiões Norte, Centro-Oeste e Nordeste. As diferenças das médias dos pares regionais são significativas. Esses resultados não concordam com o padrão do País apresentado em Souza et al. (2016), em que a relação é Sul > Sudeste > Nordeste = Centro-Oeste > Norte.

**Tabela 11.** Análise de variância – indicador de saúde.

Fonte	Graus de liberdade	Soma de quadrados	Quadrado médio	F	Pr > F
Regiões	3	22,7947	7,5983	148,34	<0,0001
Erro	855	43,795328	0,0512	-	-
Total corrigido	858	66,5900	-	-	-

**Tabela 12.** Médias de mínimos quadrados por região – indicador de saúde.

Região	Médias de mínimos quadrados	Desvio-padrão
Centro-oeste	0,3247	0,0217
Norte	0,4714	0,0209
Nordeste	0,2624	0,0162
Sudeste	0,6384	0,0108

O indicador de infraestrutura é altamente correlacionado com as imperfeições de mercado, particularmente educação e assistência técnica. Esse indicador também difere entre regiões, conforme Tabelas 13 e 14. A maior média é da região Sudeste, seguida de Centro-Oeste, Nordeste e Norte.

**Tabela 13.** Análise de variância – indicador de infraestrutura.

Fonte	Graus de liberdade	Soma de quadrados	Quadrado médio	F	Pr > F
Regiões	3	40,1329	13,3777	256,7	<0,0001
Erro	855	44,5569	0,0521	-	-
Total corrigido	858	84,6898	-	-	-

**Tabela 14.** Médias de mínimos quadrados por região – indicador de infraestrutura.

Região	Médias de mínimos quadrados	Desvio-padrão
Centro-Oeste	0,3596	0,0219
Norte	0,1161	0,0211
Nordeste	0,1714	0,0163
Sudeste	0,6163	0,0109

A assistência técnica também difere entre regiões (Tabelas 15 e 16) e a ordem das médias é Sudeste = Centro-Oeste > Norte > Nordeste.

**Tabela 15.** Análise de variância – indicador de assistência técnica.

Fonte	Graus de liberdade	Soma de quadrados	Quadrado médio	F	Pr > F
Regiões	3	34,0292	11,3431	263,33	<0,0001
Erro	855	36,8297	0,04308	-	-
Total corrigido	858	70,8588	-	-	-

**Tabela 16.** Médias de mínimos quadrados por região – indicador de assistência técnica.

Região	Médias de mínimos quadrados	Desvio-padrão
Centro-Oeste	0,6431	0,0199
Norte	0,4403	0,0192
Nordeste	0,1730	0,0148
Sudeste	0,65578	0,00993

Finalmente, o indicador de concentração de Gini sugere níveis elevados de concentração da renda em todas as regiões. Médias são superiores a 80% (Tabelas 17 e 18).

**Tabela 17.** Análise de variância – índice de Gini.

Fonte	Graus de liberdade	Soma de quadrados	Quadrado médio	F	Pr > F
Regiões	3	0,1775	0,0592	7,71	<0,0001
Erro	855	6,5616	0,0077	-	-
Total corrigido	858	6,7391	-	-	-

**Tabela 18.** Médias de mínimos quadrados por região – índice de Gini.

Região	Médias de mínimos quadrados	Desvio-padrão
Centro-Oeste	0,8391	0,0084
Norte	0,8091	0,0081
Nordeste	0,8048	0,0063
Sudeste	0,8353	0,0042

Na Tabela 19 apresentam-se os resultados da estimação via GMM. Foram utilizadas 12 variáveis instrumentais, quais sejam: proporção de estabelecimentos com destino adequado de lixo, taxa de envelhecimento da população rural, existência de energia elétrica, índice de desenvolvimento da educação básica, suprimento de água, índice de desempenho do sistema de saúde, indicador ambiental, indicador demográfico e quatro variáveis indicadoras regionais (Norte, Nordeste, Centro-Oeste e Sul). A função objetivo do método de estimação é mínima para a distribuição inversa do valor extremo. O teste J de Hansen de adequabilidade do ajuste (Greene, 2011) com três graus de liberdade (0,596) não é significativo ( $p$ -valor = 0,897). A correlação entre valores observados e preditos é 53,2%. O indicador de saúde não foi significativo e, portanto, foi retirado do modelo, mas não do conjunto de variáveis instrumentais.

As covariáveis atuam no sentido oposto às imperfeições de mercado. Valores baixos desses indicadores implicam eficiência técnica próxima a um para estabelecimentos

**Tabela 19.** Resultados da estimação via método de momentos generalizado (GMM).

Variável	Coefficiente	Desvio-padrão	z	P> z	Intervalo de confiança a 95%	
Tecnologia	0,7773	0,2654	2,93	0,003	0,2571	1,2976
Meio ambiente	0,4615	0,1312	3,52	0,000	0,2045	0,7186
Demografia	-0,4082	0,2365	-1,73	0,084	-0,8718	0,0553
Educação	-0,1532	0,0730	-2,10	0,036	-0,2962	-0,0101
Infraestrutura	-0,0915	0,0441	-2,08	0,038	-0,1779	-0,0051
Norte	0,0224	0,0506	0,44	0,658	-0,0767	0,1216
Nordeste	-0,1179	0,0441	-2,67	0,008	-0,2044	-0,0315
Centro-Oeste	-0,1316	0,0439	-3,00	0,003	-0,2176	-0,0455
Sul	0,2796	0,0361	7,75	0,000	0,2089	0,3503
Constante	0,2907	0,0811	3,58	0,000	0,1317	0,4497

com alto nível de renda e concentração. Nesse contexto, esperam-se sinais negativos para as variáveis de imperfeição de mercado e positivo para tecnologia. Nota-se que a variável ambiental, como aqui medida, está associada à tecnologia e deveria ter sinal positivo. De fato, observa-se esse padrão. Nota-se um forte efeito regional. O efeito demográfico é marginalmente significativo e os demais são estatisticamente significantes.

A intensidade de qualquer efeito marginal, para o município  $\tau$ , é determinada pelo fator  $\exp[\mu_\tau - \exp(\mu_\tau)]$ . Observa-se que as regiões mais beneficiadas com redução das imperfeições de mercado seriam Norte e Nordeste. Na Tabela 20 especifica o quantitativo esperado, calculado na intensidade média regional de cada variável contextual. Por exemplo, o efeito do aumento de uma quantidade  $dtec$  no nível eficiência técnica levaria ao incremento de 0,2401  $dtec$  no índice de concentração de renda esperado para a região Norte. Por sua vez, o aumento de uma quantidade  $deduc$  no nível de educação reduziria a concentração da renda em 0,0473  $deduc$  na mesma região.

**Tabela 20.** Efeitos marginais no índice de Gini, *ceteris paribus*.

Região	Tecnologia	Meio ambiente	Demografia	Educação	Infraestrutura
Norte	0,2401	0,1428	-0,1261	-0,0473	-0,0284
Nordeste	0,2393	0,1423	-0,1257	-0,0471	-0,0283
Sudeste	0,2246	0,1335	-0,1179	-0,0442	-0,0266
Sul	0,1422	0,0845	-0,0747	-0,0280	-0,0168
Centro-Oeste	0,2238	0,1331	-0,1175	-0,0441	-0,0265

## Considerações finais

Estudou-se a concentração da renda rural bruta no Cerrado, considerando a distribuição municipal do índice de Gini. A concentração é elevada e em média supera 80%. Para os municípios do Cerrado, as maiores concentrações ocorrem nas regiões Centro-Oeste e Sudeste. A eficiência técnica é maior no Centro-Oeste. Imperfeições de mercado diferem regionalmente como indicado nos modelos de análise de variância. As condições mais desfavoráveis concentram-se nas regiões Norte e Nordeste, que apresentam valores mínimos em todos os atributos estudados, com exceção do indicador de saúde.

Os efeitos marginais são mais intensos nas regiões Norte e Nordeste, como evidenciado pela estimação via método de momentos generalizado. Como consequência, essas regiões seriam as mais beneficiadas por políticas públicas de inclusão produtiva via redução de imperfeições de mercado nos municípios correspondentes.

O escore DEA – *proxy* escolhida para uso de tecnologia – é dominante no modelo econométrico explicativo da concentração da renda rural. O escore tecnológico está também fortemente associado às imperfeições de mercado, que são também responsáveis pela concentração de renda. Assim, o uso de políticas públicas para eliminar desigualdades nas áreas rurais do Cerrado passa pela eliminação das imperfeições de mercado. Se essa eliminação não ocorrer, a extensão rural não será capaz de atingir de forma adequada os produtores rurais, particularmente aqueles do bioma Cerrado.

O modelo estatístico, com base em regressão fracionária e GMM, passa nos testes de especificação e confirma a relação positiva e estatisticamente significativa entre tecnologia e concentração de renda rural. Como esperado, as melhorias das condições de entorno dos estabelecimentos podem reduzir a concentração de renda em todas as regiões.

## Referências

- ALVES, E.; SOUZA, G. da S. Pequenos estabelecimentos em termos de área também enriquecem? Pedras e Tropeços. **Revista de Política Agrícola**, v. 24, n. 3, p. 7-21, 2015.
- ALVES, E.; SOUZA, G. S.; ROCHA, D. P. Desigualdade nos campos sob a ótica do censo agropecuário 2006. **Revista de Política Agrícola**, v. 22, p. 67-75, 2013.
- BANKER, R. D.; CHARNES, A.; COOPER, W. W. Some models for estimating technical scale inefficiencies in data envelopment analysis. **Management Science**, v. 30, n. 9, p. 1078-1092, 1984. DOI: 10.1287/mnsc.30.9.1078.
- CONFEDERAÇÃO NACIONAL DA AGRICULTURA. **Índice de Desenvolvimento Rural CNA**. [S.l.: s.n.], 2013.
- CONOVER, M. J. **Practical nonparametric statistics**. 3rd ed. New York: Wiley, 1999.
- COOPER, W. W.; SEIFORD, L. M.; ZHU, J. **Handbook on data envelopment analysis**. 2nd ed. New York: Springer, 2011.

- DAVIDSON, R.; MACKINNON, J. G. **Estimation and inference in econometrics**. New York: Oxford University Press, 1993.
- GALLANT, A. R. **Nonlinear statistical models**. New York: Wiley, 1987.
- GREENE, W. H. **Econometric analysis**. 7th ed. New Jersey: PrenticeHall, 2011.
- HELFAND, S. M.; ROCHA, R.; VINHAIS, H. E. F. Pobreza e desigualdade de renda no Brasil rural: uma análise da queda recente. **Pesquisa e Planejamento Econômico**, v. 39, n. 1, p. 59-80, 2009.
- KENDALL, M.; STUART, A. **The advanced theory of statistics**. London: Griffin, 1977. 484 p.
- NEDER, H. D.; SILVA, J. L. M. Pobreza e distribuição de renda em áreas rurais: uma abordagem de inferência. **Revista de Economia e Sociologia Rural**, v. 42, n. 3, p. 469-486, 2004. DOI: 10.1590/S0103-20032004000300004.
- NEY, M. G.; HOFFMANN, R. A contribuição das atividades agrícolas e não-agrícolas para a desigualdade de renda no Brasil rural. **Economia Aplicada**, v. 12, n. 3, p. 365-393, 2008. DOI: 10.1590/S1413-80502008000300002.
- NEY, M. G.; HOFFMANN, R. Educação, concentração fundiária e desigualdade de rendimentos no meio rural brasileiro. **Revista de Economia e Sociologia Rural**, v. 47, n. 1, p. 147-181, jan./mar. 2009. DOI: 10.1590/S0103-20032009000100006.
- RAMALHO, E. A.; RAMALHO, J. J. S.; HENRIQUES, P. D. Fractional regression models for second stage DEA efficiency analyses. **Journal of Productivity Analysis**, v. 34, n. 3, p. 239-255, 2010.
- SOUZA, G. S.; GOMES, E. G.; ALVES, E. R. A. Determinantes da dispersão da renda no meio rural brasileiro. **Blucher Proceedings**, v. 2, n. 1, p. 173-184, 2016.
- SOUZA, J. **Estatística econômica e social**. Rio de Janeiro: Campus, 1977. 229 p.