

Schettini, Bernardo P.

Working Paper

Eficiência técnica dos municípios Brasileiros na educação pública: Escores robustos e fatores determinantes

Texto para Discussão, No. 2001

Provided in Cooperation with:

Institute of Applied Economic Research (ipea), Brasília

Suggested Citation: Schettini, Bernardo P. (2014) : Eficiência técnica dos municípios Brasileiros na educação pública: Escores robustos e fatores determinantes, Texto para Discussão, No. 2001, Instituto de Pesquisa Econômica Aplicada (IPEA), Brasília

This Version is available at:

<https://hdl.handle.net/10419/121713>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

2001

TEXTO PARA DISCUSSÃO

EFICIÊNCIA TÉCNICA DOS MUNICÍPIOS BRASILEIROS NA EDUCAÇÃO PÚBLICA: ESCORES ROBUSTOS E FATORES DETERMINANTES

Bernardo P. Schettini



2001

TEXTO PARA DISCUSSÃO

Brasília, agosto de 2014

EFICIÊNCIA TÉCNICA DOS MUNICÍPIOS BRASILEIROS NA EDUCAÇÃO PÚBLICA: ESCORES ROBUSTOS E FATORES DETERMINANTES

Bernardo P. Schettini¹

1. Técnico de Planejamento e Pesquisa da Diretoria de Estudos e Políticas Macroeconômicas (Dimac) do Ipea.
E-mail: bernardo.schettini@ipea.gov.br.

Governo Federal

**Secretaria de Assuntos Estratégicos da
Presidência da República**
Ministro Marcelo Côrtes Neri

ipea Instituto de Pesquisa
Econômica Aplicada

Fundação pública vinculada à Secretaria de Assuntos Estratégicos da Presidência da República, o Ipea fornece suporte técnico e institucional às ações governamentais – possibilitando a formulação de inúmeras políticas públicas e programas de desenvolvimento brasileiro – e disponibiliza, para a sociedade, pesquisas e estudos realizados por seus técnicos.

Presidente

Sergei Suarez Dillon Soares

Diretor de Desenvolvimento Institucional

Luiz Cezar Loureiro de Azeredo

Diretor de Estudos e Políticas do Estado, das Instituições e da Democracia

Daniel Ricardo de Castro Cerqueira

Diretor de Estudos e Políticas Macroeconômicas

Cláudio Hamilton Matos dos Santos

Diretor de Estudos e Políticas Regionais, Urbanas e Ambientais

Rogério Bouer Miranda

Diretora de Estudos e Políticas Setoriais de Inovação, Regulação e Infraestrutura

Fernanda De Negri

Diretor de Estudos e Políticas Sociais

Herton Ellery Araújo

Diretor de Estudos e Relações Econômicas e Políticas Internacionais

Renato Coelho Baumann das Neves

Chefe de Gabinete

Bernardo Abreu de Medeiros

Assessor-chefe de Imprensa e Comunicação

João Cláudio Garcia Rodrigues Lima

Ouvidoria: <http://www.ipea.gov.br/ouvidoria>

URL: <http://www.ipea.gov.br>

Texto para Discussão

Publicação cujo objetivo é divulgar resultados de estudos direta ou indiretamente desenvolvidos pelo Ipea, os quais, por sua relevância, levam informações para profissionais especializados e estabelecem um espaço para sugestões.

© Instituto de Pesquisa Econômica Aplicada – **ipea** 2014

Texto para discussão / Instituto de Pesquisa Econômica Aplicada.- Brasília : Rio de Janeiro : Ipea , 1990-

ISSN 1415-4765

1. Brasil. 2. Aspectos Econômicos. 3. Aspectos Sociais.
I. Instituto de Pesquisa Econômica Aplicada.

CDD 330.908

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade do(s) autor(es), não exprimindo, necessariamente, o ponto de vista do Instituto de Pesquisa Econômica Aplicada ou da Secretaria de Assuntos Estratégicos da Presidência da República.

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

JEL: C14, C24, H52, I22.

SUMÁRIO

SINOPSE

ABSTRACT

| | |
|-----------------------------|----|
| 1 INTRODUÇÃO | 7 |
| 2 ESTRATÉGIA EMPÍRICA | 10 |
| 3 BASE DE DADOS | 18 |
| 4 RESULTADOS..... | 28 |
| 5 CONCLUSÃO | 41 |
| REFERÊNCIAS | 42 |

SINOPSE

Este artigo estima a eficiência técnica relativa dos municípios brasileiros na educação pública por meio da técnica DEA, aliada a regressões tobit e com o auxílio de métodos de reamostragem. O método *jackstrap* possibilitou chegar a escores de eficiência robustos, enquanto o *bootstrap* viabilizou estimar o viés destes índices. Com efeito, estimativas kernel das densidades evidenciaram alterações notáveis na distribuição da eficiência. A análise econométrica contou com o algoritmo que corrige a correlação serial dos escores. Conforme esperado, os fatores ambientais alteram de maneira relevante a dispersão da eficiência no Brasil. Em um passo final, foram implementados testes de comparação de grupos baseados em informações referentes a certas práticas de gestão na área da educação pública.

Palavras-chave: eficiência técnica; ensino fundamental público; DEA em dois estágios; fatores ambientais; práticas de gestão.

ABSTRACT

This paper estimates the relative technical efficiency of Brazilian municipalities in the public education using the DEA model along with tobit regressions with the aid of resampling methods. The jackstrap method enabled the estimation of robust efficiency scores, while the bootstrap allowed the estimation of the bias in those indexes. In effect, kernel density estimates portray remarkable changes in the distribution of efficiency. The econometric analysis relied upon the algorithm which corrects the serial correlation in the scores. It is noteworthy how accounting for environmental factors alters the dispersion of efficiency in Brazil. In a final step, we implement group comparisons tests based on information relative to certain management practices in the area of public education.

Keywords: technical efficiency; primary public education; two-stage DEA; environmental factors; management practices.

1 INTRODUÇÃO

O Plano de Desenvolvimento da Educação – PDE (Brasil, 2007) enxergava a necessidade de elevar até 7% do PIB a despesa pública em educação. Esta proporção aparecia como meta no Projeto de Lei do Plano Nacional da Educação (PNE) para o período de 2011 a 2020.¹ Durante a tramitação no Congresso Nacional, esta parcela foi elevada para 10% do produto interno bruto (PIB).

Segundo dados do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), o “investimento direto em educação” era de 4,5% do PIB em 2007 (Maciel, 2012).² De acordo com informações extraídas do Sistema de Informações sobre Orçamentos Públicos em Educação (Siope), o gasto total nas subfunções típicas era de 4,2% do PIB no mesmo ano. Estes valores cresceram para 5,3 e 4,82% do PIB em 2011. A experiência mostra que a expansão desordenada do gasto público leva ao desperdício, isto é, à ineficiência. A expansão pretendida requer, portanto, planejamento e práticas de gestão adequadas.

Boa parte das despesas na área da educação é executada pelas prefeituras e secretarias de educação municipais. De acordo com dados do Siope, esta parcela foi de 46% em média entre 2007 e 2011. Deste total, mais de 80% é destinado ao ensino fundamental, pois os municípios têm atuação prioritária nesta área. Estudos sobre eficiência sugerem haver substancial desperdício de recursos nesta esfera. Por exemplo, Rocha *et al.* (2013) estimam que os resultados medidos pelo índice de desenvolvimento da educação básica (IDEB) poderiam ser alcançados com um gasto por aluno cerca de 40% menor em média. Há uma série de estudos sobre eficiência técnica na educação básica no Brasil, conforme síntese feita na sequência.

Zoghbi *et al.* (2011) mensuraram a eficiência dos municípios de São Paulo no ensino fundamental por meio da técnica não paramétrica *free disposal hull* (FDH).

1. O Projeto de Lei do Plano Nacional da Educação (PNE) encontra-se disponível em: <<http://goo.gl/4vDlq2>>.

2. Maciel (2012) descreve a metodologia utilizada pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) no cálculo de indicadores financeiros da área da educação pública. Segundo a terminologia adotada, o “investimento total em educação” é a soma de todos os recursos aplicados pelo setor público (União, estados e municípios), incluindo a estimativa da aposentadoria futura do pessoal ativo no setor. O “investimento direto” exclui, além da estimativa da aposentadoria futura, o pagamento de bolsas de estudo, financiamento estudantil e transferências para entidades privadas.

Eles mostraram a distribuição dos escores de eficiência entre as regiões administrativas do estado, bem como analisaram a correlação com o PIB *per capita*, entre outras variáveis. Rosano-Peña, Albuquerque e Marcio (2012) estudaram os municípios de Goiás com o auxílio da técnica de análise envoltória de dados (em inglês, *data envelopment analysis* – DEA), cabendo observar que a mensuração levou em consideração a existência de insumos não discricionários (sem recorrer a métodos econométricos). Delgado e Machado (2007), por sua vez, analisaram a eficiência das escolas públicas estaduais de Minas Gerais no ensino fundamental e médio, utilizando o método DEA em dois estágios, que usa instrumental econométrico para estimar a influência de fatores socioeconômicos. Trigo (2010) analisou a eficiência na educação fundamental e média das escolas públicas e privadas de todo o Brasil. Para tal, usou fronteiras estocásticas, pois a influência de pontos extremos é menor nesta técnica paramétrica.³

Gasparini e Ramos (2003) fizeram estudo pioneiro com a técnica DEA para estudar a eficiência dos estados e do Distrito Federal no ensino médio. A análise de regressão feita no segundo estágio incluiu fatores socioeconômicos que influenciam o desempenho dos sistemas escolares. Zoghbi *et al.* (2009) utilizaram a técnica FDH para comparar o desempenho e os escores de eficiência dos estados e do Distrito Federal na área de educação fundamental e média. Alves Júnior (2010) adotou um procedimento em três estágios que combina DEA e fronteira estocástica para estudar a eficiência dos estados e do Distrito Federal de 1995 a 2008 nesses mesmos níveis. Benegas (2012), por sua vez, inovou ao utilizar o modelo *network* DEA considerando também tanto a educação fundamental como o ensino médio. Tal modelo concebe a produção educacional em duas etapas. A primeira garante a abrangência da oferta (o gasto público gera estabelecimentos de ensino e docentes). Estes produtos intermediários são insumos na segunda etapa do processo (efetividade do ensino), que resulta nas notas nos testes padronizados e nas taxas de conclusão nos ciclos escolares.

3. Vale notar que estudos que envolvem municípios enfrentam o desafio de lidar com a existência de diversas observações influentes. Isto pode ocorrer devido a erro de mensuração nos dados ou porque certos municípios apresentam desempenho excepcional e deslocam a fronteira de produção, possivelmente por se defrontarem com uma tecnologia distinta. O fato é que as técnicas não paramétricas como *free disposal hull* (FDH) e de análise envoltória de dados (DEA) são bastante sensíveis à presença de tais pontos extremos. Limitação semelhante devem enfrentar estudos que enfocam as escolas. Muitas vezes a saída é fazer estimativas para municípios ou escolas de determinada região.

Em relação ao ensino superior, Marinho, Resende e Façanha (1997) utilizaram a técnica DEA para estimar a eficiência das universidades federais brasileiras. Posteriormente, Façanha e Marinho (1999) relacionaram a eficiência estimada com o modelo de financiamento destas instituições de ensino à luz das teorias dos contratos e incentivos. Façanha e Marinho (2001) fizeram estimativas DEA incluindo instituições de ensino superior do governo e privadas.

Neste estudo, foi analisada a eficiência relativa dos municípios de todo o Brasil na área da educação por meio de dados referentes ao primeiro ciclo do ensino fundamental (1º ao 5º ano). Para tal, utilizou-se a técnica DEA aliada a regressões tobit para levar em consideração os determinantes socioeconômicos das ineficiências. Por fim, analisou-se a influência de alguns fatores ligados à administração no nível das secretarias de educação e das próprias escolas.

Seguindo Stošić e Sampaio de Sousa (2003) e Sampaio de Sousa e Stošić (2005), foi utilizada a técnica denominada *jackstrap*, que possibilitou identificar uma série de pontos extremos que deslocam a fronteira da possibilidade de produção. Estimativas kernel da densidade evidenciaram o impacto de tal procedimento sobre toda a distribuição da eficiência. Da mesma forma, seguindo a recomendação de Simar e Wilson (1998), utilizou-se a técnica *bootstrap* para estimar o viés e corrigir os escores. Novamente, a comparação das distribuições revelou alteração na densidade. Em relação à análise econométrica, implementou-se o algoritmo eficiente de Simar e Wilson (2007) para encontrar escores livres da influência de fatores ambientais. No passo final, foram implementados testes F e Kolmogorov-Smirnov para comparar a distribuição dos escores de certos grupos criados a partir de informações acerca da gestão na área da educação pública.

Afora essa introdução, o artigo compreende quatro seções. A seção 2 mostra a estratégia empírica. Na seção 3, são descritos os dados e reportadas certas estatísticas descritivas. Na seção 4 foram apresentados os resultados encontrados. A seção 5 faz as considerações finais.

2 ESTRATÉGIA EMPÍRICA

Esta seção apresenta o conjunto de métodos utilizados nas análises dos dados municipais referentes à educação fundamental. A estratégia empírica seguida consiste na aplicação da técnica DEA aliada à análise econométrica. Afora isso, os passos descritos na sequência expõem as técnicas de reamostragem empregadas para lidar com observações influentes, corrigir o viés das estimativas, assim como viabilizar a inferência estatística.

2.1 Método semiparamétrico

As análises de eficiência têm como principal referencial, na área de economia, a teoria da produção. No caso da educação fundamental, as principais unidades tomadoras de decisão (DMUs) são as prefeituras municipais, uma vez que grande parte das decisões acerca da alocação de recursos orçamentários foge do escopo de atuação da administração das escolas.⁴

O modelo empírico pressupõe que os N municípios se defrontam com uma tecnologia comum (Ψ). No que se segue, denotam-se por $x_n = (x_{n1}, \dots, x_{np}) \in \mathbb{R}_+^p$ e $y_n = (y_{n1}, \dots, y_{nq}) \in \mathbb{R}_+^q$ os vetores de insumos e produtos da DMU n ($n = 1, \dots, N$). No conjunto de dados $\mathcal{C} = \{(x_n, y_n); n = 1, \dots, N\}$, o conjunto de possibilidades de produção (CPP) é formado pelos planos (x, y) factíveis:

$$\mathcal{P} = \{(x, y) \in \mathbb{R}_+^{p+q} | x \text{ pode produzir } y\}. \quad (1)$$

Essa tecnologia não é observável, mas a técnica DEA possibilita chegar a uma estimativa $\hat{\mathcal{P}}$ a partir dos dados. Para tal, adota o expediente de programação linear matemática aliado ao princípio da extrapolação mínima.

Segundo tal princípio, a estimativa $\hat{\mathcal{P}}$ é o menor subconjunto no espaço $\mathbb{R}_+^p \times \mathbb{R}_+^q$ que contém os dados (x_n, y_n) , para todo n , e satisfaz certos pressupostos que podem variar de acordo com a aplicação.

4. As análises empreendidas se apoiam no conceito de eficiência técnica. Isto diz respeito à utilização ótima dos insumos em face da tecnologia de produção, mas desconsidera os vetores de preços necessários para analisar a eficiência alocativa, uma vez que estes não se encontram disponíveis no âmbito municipal.

Entre essas condições de regularidade, as mais comuns são: *i*) o livre descarte de insumos e produtos; *ii*) a convexidade do conjunto de possibilidades de produção; e *iii*) algum tipo de retorno de escala.

A primeira condição significa ser possível produzir menos com mais, isto é, $(x, y) \in \mathcal{P}$, $x' \geq x$ e $y' \leq y \Rightarrow (x', y') \in \mathcal{P}$. De acordo com a segunda, combinações convexas de planos de produção são factíveis, ou seja, $(x, y) \in \mathcal{P}$, $(x', y') \in \mathcal{P}$, $\alpha \in [0, 1] \Rightarrow \alpha \cdot (x, y) + (1 - \alpha) \cdot (x', y') \in \mathcal{P}$. A terceira condição de regularidade atribui Υ - retornos de escala à tecnologia de produção, em que as hipóteses possíveis são $\Upsilon = crs$ (retornos constantes de escala), $\Upsilon = drs$ (não crescentes), $\Upsilon = irs$ (não decrescentes) ou $\Upsilon = vrs$ (variáveis).

A tecnologia que será encontrada é a seguinte (Bogetoft e Otto, 2011):

$$\hat{\mathcal{P}}(\Upsilon) = \left\{ (x, y) \in \mathbb{R}_+^{p+q} \mid \exists n \in \Lambda^N(\Upsilon) : x \geq \sum_{n=1}^N \lambda_n x_n, y \leq \sum_{n=1}^N \lambda_n y_n \right\}, \quad (2)$$

Em que, por exemplo:⁵

$$\Lambda^N(crs) = \left\{ \lambda \in \mathbb{R}_+^N \mid \sum_{n=1}^N \lambda_n \text{ livre} \right\},$$

$$\Lambda^N(vrs) = \left\{ \lambda \in \mathbb{R}_+^N \mid \sum_{n=1}^N \lambda_n = 1 \right\}.$$

É usual trabalhar com medidas radiais de eficiência, que enfocam mudanças proporcionais nos insumos ou produtos. Na orientação pelos produtos, o escore de eficiência de Farrel da DMU o é dado pelo escalar:

$$\delta_o = \delta(x_o, y_o \mid \mathcal{P}) = \max\{\delta \in \mathbb{R}_+ \mid (x_o, \delta y_o) \in \mathcal{P}\}. \quad (3)$$

5. O conjunto de possibilidades de produção (CPP) mais amplo é o de retornos constantes, pois permite mudanças arbitrárias de escala ($\sum_{n=1}^N \lambda_n$ livre). A tecnologia de retornos não crescentes permite mudanças radiais somente na direção da origem ($\sum_{n=1}^N \lambda_n \leq 1$). O contrário é válido para a de retornos não decrescentes ($\sum_{n=1}^N \lambda_n \geq 1$). A tecnologia de retornos variáveis envelope melhor os dados, sendo definida pela intersecção destes dois últimos CPPs ($\sum_{n=1}^N \lambda_n = 1$).

Deve estar claro que, para essa DMU, a estimativa do escore ($\hat{\delta}_o$) irá refletir a distância euclidiana (paralela aos eixos dos produtos e perpendicular aos eixos dos insumos) do ponto (x_o, y_o) até a fronteira estimada da tecnologia de produção subjacente aos dados.

O programa DEA é dado por:

$$\begin{aligned} \hat{\delta}_o &= \max_{\delta, \lambda_1, \dots, \lambda_N} \delta & (4) \\ \text{s.t.} \quad x_{oi} &\geq \sum_{n=1}^N \lambda_n x_{ni}, & (i = 1, \dots, p) \\ \delta y_{oj} &\geq \sum_{n=1}^N \lambda_n y_{nj}, & (j = 1, \dots, q) \\ \lambda &\in \Lambda^N(\gamma). \end{aligned}$$

Portanto, o problema de otimização descrito envolve, para cada DMU o , encontrar $N + 1$ variáveis a partir de uma função objetivo linear e $p + q + 1$ restrições lineares.

Regressões podem ser utilizadas para lançar luz acerca dos determinantes exógenos dos escores, bem como para corrigir tais estimativas e obter *rankings* ajustados pela influência desses fatores ambientais (por exemplo, as condições socioeconômicas locais, que estão fora do controle dos gestores locais no curto prazo).

Esses fatores podem ser resumidos pelo vetor z_n de variáveis exógenas, que não comparecem na função de produção, mas influenciam o processo produtivo de maneira indireta. Tipicamente, essas análises econométricas são feitas por meio de uma equação *log*-linear:

$$\ln \delta_n = z_n \beta + \varepsilon_n, \quad (n = 1, \dots, N) \quad (5)$$

em que β denota o vetor de parâmetros que precisa ser estimado e ε_n é um termo de erro *iid* não correlacionado com z_n .

Os escores δ_n são realizações de uma variável censurada (pois $\delta_n = 1$ e $\ln \delta_n = 0$ para $z_n \beta + \varepsilon_n \leq 0$). Isto motiva o uso do modelo tobit. Porém, como a real ineficiência não é observável, a variável-resposta é formada pelos índices DEA no modelo empírico, isto é:

$$\ln \hat{\delta}_n = z_n \beta + \xi_n \geq 0, \quad (n = 1, \dots, N) \quad (6)$$

em que ξ_n é uma variável *nid* truncada à esquerda no ponto $-Z_n \beta$.

Essa regressão pode ser vista como a decomposição da ineficiência em uma parcela que é exógena ($z_n \beta$) e outra parte que reflete o desempenho da administração local (ξ_n). Desta forma, as estimativas de máxima verossimilhança para β possibilitam chegar na “ineficiência pura” (não correlacionada com fatores socioeconômicos), cabendo normalizar esses escores residuais:

$$\tilde{\delta}_n = \exp(\xi_n - \xi_{\min}), \quad (n = 1, \dots, N) \quad (7)$$

em que ξ_{\min} denota o mínimo de ξ_n (DMU mais eficiente).

Este modelo empírico pode conduzir a análises equivocadas, pois os escores são sensíveis à presença de *outliers*. Além disso, os estimadores apresentados não estão livres de viés. Apesar de consistentes, a convergência é muito lenta no modelo não paramétrico. Há também questões ligadas à eficiência dos estimadores que precisam ser resolvidas. O restante da seção descreve a maneira escolhida para lidar com essas dificuldades práticas.

2.2 Outliers

A principal vantagem da técnica da DEA é a flexibilidade na adequação aos dados. Por seu turno, a principal fraqueza reside no fato de não levar em consideração perturbações aleatórias nos dados, decorrentes, por exemplo, de erros de mensuração nas variáveis. Na ótica DEA, qualquer desvio em relação à fronteira de produção significa ineficiência.

Esses fatores fazem com que a mensuração pretendida seja sensível à presença de *outliers*, uma vez que a fronteira é constituída por combinações lineares dos pontos extremos da amostra. Stošić e Sampaio de Sousa (2003) propuseram uma maneira de encontrar escores robustos a observações influentes que é atrativa quando a quantidade de DMUs é elevada, pois não depende da inspeção visual dos dados.⁶

6. Sampaio de Sousa e Stošić (2005) exemplificaram a utilização dessa técnica num exercício empírico com dados municipais brasileiros. Santos, Cribari-Neto e Sampaio de Sousa (2007) e Sampaio de Sousa, Cribari-Neto e Stošić (2005) fizeram aplicações semelhantes para mensurar a eficiência nas principais áreas de atuação das prefeituras. Delgado e Machado (2007) utilizaram-na nas análises das escolas estaduais de Minas Gerais.

Essa técnica é denominada de *jackstrap*, pois combina um esquema *jackknife* com a reamostragem estocástica *bootstrap*, o que reduz o custo computacional. Estes expedientes são utilizados para calcular a influência de cada observação sobre a fronteira de produção.

A alavancagem da DMU o se refere ao impacto da sua retirada da amostra sobre a mensuração da eficiência das demais DMUs, podendo ser medida pelo desvio-padrão dos escores antes e depois:

$$\ell_o = \sqrt{\frac{\sum_{n=1; n \neq o}^N (\hat{\delta}_n^{o^-} - \hat{\delta}_n)^2}{N-1}}, \quad (8)$$

em que $\{\hat{\delta}_n | n = 1, \dots, N\}$ se refere aos escores calculados para cada DMU, utilizando todos os dados, enquanto $\{\hat{\delta}_n^{o^-} | n = 1, \dots, N; n \neq o\}$ denota os escores calculados sem a observação da DMU o .⁷

Os passos relevantes, conforme Stošić e Sampaio de Sousa (2003) estão expostos a seguir.

1. Selecionar aleatoriamente um subconjunto de L DMUs e calcular as alavancagens $\tilde{\ell}_l$; em que l assume L valores distintos sorteados do conjunto $\{1, \dots, N\}$, pois não há reposição.
2. Repetir o passo anterior B vezes para calcular as alavancagens $\tilde{\ell}_{lb}$ ($b = 1, \dots, B$), notando que, para B grande o suficiente, cada DMU l será selecionada $n_l \approx BL/N$ vezes.
3. Calcular a média para cada DMU l , sendo

$$\tilde{\ell}_l = \frac{\sum_{b=1}^{n_l} \tilde{\ell}_{lb}}{n_l}, \quad (9)$$

e a média global das alavancagens,

7. Para N grande, é praticamente proibitivo calcular alavancagens para todas as unidades tomadoras de decisão (DMUs), uma vez que isto exigiria resolver $N \times (N - 1)$ programas lineares.

$$\tilde{\ell} = \frac{\sum_{l=1}^N \tilde{\ell}_l}{N}. \quad (10)$$

É possível atribuir menor chance de selecionar DMUs muito influentes no cálculo de distribuições empíricas via reamostragem, adotando-se uma função de probabilidade que penaliza a alavancagem elevada. Esses valores possibilitam também identificar e eliminar observações muito influentes para refazer os cálculos via DEA, sendo necessário escolher um valor de limiar.

Além de combinar essas duas ideias, a função (11) de probabilidade (*Heaviside step function*) apresentou os resultados mais satisfatórios nas simulações reportadas.

$$P(\tilde{\ell}_o) = \begin{cases} 1, & \tilde{\ell}_o < \tilde{\ell} \log N \\ 0, & \tilde{\ell}_o \geq \tilde{\ell} \log N, \end{cases} \quad (11)$$

em que $\tilde{\ell}^* = \tilde{\ell} \log N$ é o valor do limiar (que leva em consideração o tamanho da amostra de dados).

2.3 Viés e imprecisão

A análise econométrica descrita anteriormente – modelo de regressão (6) – pode ser comprometida, pois o estimador de β tende a ser enviesado, além de impreciso, devido aos problemas descritos no que se segue.

É importante notar que, embora a estimativa DEA dos escores seja consistente, a taxa de convergência é muito baixa e depende da dimensionalidade do plano de produção (maldição da dimensionalidade):

$$\hat{\delta}_o = \delta_o + \mathcal{O}_p \left(N^{-\frac{2}{p+q+1}} \right). \quad (12)$$

As análises descritas anteriormente são válidas assintoticamente. No entanto, três problemas precisam ser considerados para melhorar as propriedades do estimador em amostras finitas.

Em primeiro lugar, os resíduos ξ_n são serialmente correlacionados. Isto decorre do fato de que os índices DEA dependem das observações de (x, y) de maneira complexa por intermédio do CPP. Esta estrutura de correlação, que é desconhecida, deve contaminar a distribuição dos resíduos.

Segundo, as variáveis que integram o plano de produção (x_n, y_n) são correlacionadas com os fatores ambientais no vetor z_n – caso contrário, não haveria motivação para as análises econométricas. Isto, por sua vez, faz com que ξ_n seja correlacionado com as variáveis de z_n .

Em terceiro lugar, os próprios escores são enviesados para baixo, pois subestimam a real ineficiência, uma vez que o CPP estimado via DEA constitui uma aproximação interior da tecnologia. Este viés tende a ser maior em regiões esparsas do CPP, pois há menos informação em relação à localização da fronteira. Isto constitui outro canal pelo qual o resíduo ξ_n pode estar correlacionado com (x_n, y_n) e, por conseguinte, com z_n .

Esses problemas podem ser contornados com o auxílio de técnicas de reamostragem. Contudo, chama-se atenção para o fato de que o *bootstrap* usual não se aplica no caso da técnica DEA (Simar e Wilson, 1998).⁸

Simar e Wilson (2007) propuseram um algoritmo (que aplica *bootstrap* duas vezes nos dados) para corrigir o viés e aumentar a precisão, conforme descrito a seguir.

1. Utilizando os dados originais $\mathcal{P} = \{(x_n, y_n); n = 1, \dots, N\}$, calcular os escores $\hat{\delta}_n = \hat{\delta}(x_n, y_n | \hat{\mathcal{P}})$ ($n = 1, \dots, N$) com base no programa (4).
2. Utilizar o método de máxima verossimilhança para obter estimativa $\hat{\beta}$ de β e $\hat{\sigma}_\varepsilon$ de σ_ε na regressão truncada de $\ln \hat{\delta}_n$ em z_n , conforme (6), utilizando as $M < N$ observações que obtiveram $\hat{\delta}_n > 1$.
3. Repetir os passos (a) a (d) L_j vezes para obter N conjuntos de estimativas *bootstrap*

$$\mathcal{A}_n = \{\hat{\delta}_{nb}^*\}_{b=1}^{L_1} :$$

8. Não faz sentido reamostrar B vezes diretamente $\hat{\delta}_n$ para calcular os momentos necessários e realizar testes de hipóteses, pois tal procedimento considera que toda diferença entre os escores é aleatória. Assim, também não é adequado reamostrar o conjunto \mathcal{C} e, para cada réplica \mathcal{C}^b , calcular a tecnologia $\hat{\mathcal{P}}^b$ e os escores $\hat{\delta}_n^b$ ($n = 1, \dots, N$), uma vez que isto irá produzir $\hat{\delta}_n^b \leq 1$ sempre que $(x_n, y_n) \notin \hat{\mathcal{P}}^b$.

- a) para $n = 1, \dots, N$, sortear ε_i da distribuição $N(0, \hat{\sigma}_\varepsilon^2)$ truncada em $(-z_n \hat{\beta})$;
 - b) computar $\ln \delta_n^* = z_n \hat{\beta} + \varepsilon_n$ ($n = 1, \dots, N$);
 - c) definir, $x_n^* = x_n$, $y_n^* = y_n \hat{\delta}_n / \delta_n^*$ ($n = 1, \dots, N$); e
 - d) computar $\hat{\delta}_n^* = \delta(x_n, y_n | \hat{\mathcal{P}}^*)$ ($n = 1, \dots, N$), em que $\hat{\mathcal{P}}^*$ é obtido de (2) substituindo-se (x, y) por (x^*, y^*) .
4. Utilizando as estimativas em \mathcal{A}_n e o estimador original $\hat{\delta}_n$, calcular o estimador com correção de viés $\hat{\hat{\delta}}_n = 2\hat{\delta}_n - \frac{1}{B} \sum_{b=1}^B \delta_{n,b}^*$ ($n = 1, \dots, N$).⁹
 5. Utilizar o método da máxima verossimilhança para estimar a regressão truncada de $\ln \hat{\hat{\delta}}_n$ em z_n e obter as estimativas $(\hat{\hat{\beta}}, \hat{\hat{\sigma}})$.
 6. Repetir os passos (a) a (c) L_2 vezes para obter um conjunto de estimativas *bootstrap* $\mathcal{B}_n = \{\hat{\beta}^*, \hat{\sigma}_\varepsilon^*\}_{b=1}^{L_2}$:
 - a) para $n = 1, \dots, N$, sortear ε_i da distribuição $N(0, \hat{\hat{\sigma}})$ truncada em $(-z_n \hat{\hat{\beta}})$;
 - b) computar $\ln \delta_n^{**} = z_n \hat{\hat{\beta}} + \varepsilon_n$ ($n = 1, \dots, N$); e
 - c) utilizar o método da máxima verossimilhança para estimar a regressão truncada de $\ln \delta_n^{**}$ em z_n e obter as estimativas $(\hat{\hat{\beta}}^*, \hat{\hat{\sigma}}^*)$.
 7. Utilizar os valores de *bootstrap* em \mathcal{B} e as estimativas originais $\hat{\beta}, \hat{\sigma}$ para construir intervalos de confiança para cada elemento de β e para σ_ε .

Vale ressaltar que, como o passo 1 gera a estimativa $\hat{\delta}_n$ consistente de δ_n , segue que o método da máxima verossimilhança utilizado no passo 2 produz a estimativa $\hat{\beta}$ do vetor β , que também é consistente. Contudo, a usual taxa de convergência \sqrt{N} não é verificada. Em face disso, os passos 3 e 4 aplicam o *bootstrap* nestes dados para chegar à estimativa $\hat{\hat{\delta}}_n$ com correção de viés. Finalmente, os passos 5 e 6 executam *bootstrap* paramétrico no modelo de regressão não linear, enquanto o passo 7 utiliza os percentis da distribuição empírica para construir intervalos de confiança para os parâmetros relevantes.

9. Vale lembrar que $\text{viés}_{\mathcal{P},n} = E_{\mathcal{P}}(\hat{\delta}_n) - \delta_n$. O estimador *bootstrap* calcula $\widehat{\text{viés}}_n = \bar{\delta}_n^* - \hat{\delta}_n$, em que $\bar{\delta}_n^* = \frac{1}{B} \sum_{b=1}^B \delta_{n,b}^*$ como aproximação para $\text{viés}_{\hat{\mathcal{P}},n} = E_{\hat{\mathcal{P}}}(\hat{\delta}_n^*) - \hat{\delta}_n$ (Simar e Wilson, 1998). Segue que a eficiência com correção de viés é dada por $\hat{\hat{\delta}}_n = \hat{\delta}_n - \widehat{\text{viés}}_n = 2\hat{\delta}_n - \bar{\delta}_n^*$.

3 BASE DE DADOS

Esta seção mostra os dados utilizados nas análises empreendidas. Inicialmente, é feita uma breve descrição das variáveis. Na sequência, reportam-se os resultados da análise de componentes principais. A última parte mostra algumas estatísticas descritivas.

3.1 Descrição das variáveis

As principais fontes consultadas para a análise de eficiência foram o INEP e o Fundo Nacional de Desenvolvimento da Educação (FNDE). As análises de regressão, assim como a comparação de distribuições por intermédio de indicadores de gestão, contaram com dados de outros órgãos, como a Secretaria do Tesouro (STN), o Instituto Brasileiro de Geografia e Estatística (IBGE) e o Programa das Nações Unidas para o Desenvolvimento (PNUD).

As bases de dados do INEP utilizadas foram o censo da educação básica (Censo Escolar) e a Avaliação Nacional do Rendimento Escolar (Prova Brasil), que consiste na parte quase censitária do Sistema de Avaliação da Educação Básica (SAEB), abrangendo as escolas públicas urbanas e rurais do ensino fundamental que tenham pelo menos vinte alunos matriculados nas séries avaliadas. No caso do FNDE, as análises contaram com dados do Siope.

A definição realista do rol de insumos e produtos da educação é relevante para as análises de eficiência.¹⁰ A tabela 1 descreve as variáveis empregadas no estudo. Vale ressaltar que elas se referem às médias extraídas das escolas cuja dependência administrativa é municipal. Estes dados enfocam os anos iniciais do ensino fundamental (do 1º ao 5º ano).

Em relação aos produtos, a opção seguida consistiu em trabalhar com três variáveis que captam dimensões distintas da qualidade na área da educação: proficiência em matemática (*nota_mt*), inverso da taxa de distorção idade-série (*inv_tx_distorcao*) e o inverso da taxa de abandono (*inv_tx_abandono*).

10. Se o CPP abranger insumos e produtos que importam pouco, isto deve levar à classificação incorreta de DMUs como eficientes em decorrência da ausência de pares para a comparação ou porque algumas unidades se especializaram em produtos irrelevantes. Vale notar também que a técnica sofre da maldição da dimensionalidade – ver seção 2, equação (12).

TABELA 1
Insumos e produtos (2011)

| Variável | Definição | Descrição | Fonte |
|-------------------------------------|---|---|---------------------------------|
| <i>prof_es</i> | Número de professores com ensino superior por matrícula | É feita a contagem do número de professores com ensino superior e do número de matrículas separadamente para as séries iniciais (1ª ao 5ª ano) do EF nas escolas de dependência administrativa municipal. | Censo Escolar/INEP |
| <i>horas_aula</i> | Número médio de horas-aula diárias | O INEP disponibiliza o número médio de horas-aula de forma separada para os anos iniciais do EF nas escolas de dependência administrativa municipal. | Censo Escolar/INEP |
| <i>num_salas</i> | Número de salas existentes por matrícula | É feita a contagem do número total de salas existentes e o número total de alunos no EF nas escolas de dependência administrativa municipal. | Censo Escolar/INEP |
| <i>pca_k</i> | Infraestrutura: instalações auxiliares disponíveis nas escolas | PCA na proporção de escolas que contam com biblioteca, sala de leitura, laboratório de informática, laboratório de ciências, quadra de esportes, sala de professores e sala da diretoria. | Censo Escolar/INEP |
| <i>gasto_aluno</i> | Gasto em MDE pelo município por matrícula no ensino fundamental | Considera-se, no orçamento municipal, a função 12 (Educação), subfunção 361 (Ensino Fundamental). Enquadram-se nos gastos com MDE aqueles previstos na LDB, art. 70. O denominador é dado pelo número de matrículas. A variável usada foi a média do indicador entre 2008 e 2011. | SIOPE/FNDE e Censo Escolar/INEP |
| <i>nota_mt</i> | Proficiência média em matemática dos alunos do ensino fundamental | Nota média em matemática que obtiveram os alunos do EF no 5ª ano do EF nas escolas de dependência administrativa municipal. O INEP pondera os alunos, as turmas e as escolas para fazer os cálculos de média. | Prova Brasil/INEP |
| <i>inv_tx_distorcao¹</i> | Inverso da taxa de distorção idade-série no ensino fundamental | Essa taxa indica, entre os alunos matriculados, a proporção de matrículas fora da idade correta. O INEP disponibiliza essa taxa de distorção das escolas municipais separadamente para as séries iniciais do EF nas escolas municipais. | Censo Escolar/INEP |
| <i>inv_tx_abandono²</i> | Inverso da taxa de abandono na etapa do ensino fundamental | Essa taxa indica a porcentagem de alunos que deixaram de frequentar a escola após a data de referência do Censo Escolar. O INEP faz os cálculos separadamente para as séries iniciais do EF nas escolas municipais. | Censo Escolar/INEP |

Elaboração do autor.

Obs.: EF = ensino fundamental, MDE = manutenção e desenvolvimento do ensino, LDB = Lei nº 9.394, de dezembro de 1996, conhecida como Lei de Diretrizes e Bases da Educação.

Notas: ¹ Vamos considerar o Censo do ano t e a série k do ensino fundamental, cuja idade adequada é de i anos. A taxa calculada pelo INEP é dada pelo quociente entre o número de alunos que, no ano t , completam $i + 2$ anos ou mais (nascimento antes de $t - |i + 1|$), e a matrícula total na série k .

² Taxa de abandono = $[ABA / (APR + REP + ABA)]$; onde: ABA = número de alunos que deixaram de frequentar; APR = número de alunos que aprovaram; e REP = número de alunos que reprovaram. Nos casos em que ABA = 0, imputamos o valor 1.000 para essa variável.

No que se refere aos insumos, além do gasto médio por aluno na etapa de ensino estudada (*gasto_aluno*), selecionaram-se quatro variáveis: professores com ensino superior por matrícula (*prof_es*), horas-aula diárias (*horas_aula*), salas de aula por matrícula (*num_salas*) e infraestrutura (*pca_k*).

A variável de infraestrutura foi elaborada por meio da análise de componentes principais (em inglês, *principal component analysis* – PCA), aplicada na proporção de escolas que contam com biblioteca, sala de leitura, laboratório de informática, laboratório de ciências, sala de professores e sala da diretoria.¹¹ Os resultados referentes a tal análise serão reportados adiante.

11. As condições das instalações escolares (energia e água pública, esgoto sanitário, banheiros etc.) não comparecem entre os insumos, pois fica a impressão de que a variabilidade relevante desaparece quando se tira a média municipal. Equipamentos audiovisuais de apoio ao ensino (televisão, videocassete, DVD e retroprojeto) foram desconsiderados, uma vez que é bastante questionável a influência exercida por estes itens. No caso de computador, internet e banda larga, observa-se proporção elevada de não respostas.

TABELA 2
Variáveis exógenas (2010-2011)

| Variável | Descrição e fonte dos dados |
|------------------------|---|
| <i>urbana</i> | % de escolas localizadas em área urbana (Censo Escolar/INEP) |
| <i>lmat</i> | Número de matrículas em escolas públicas municipais no ensino fundamental (em log) (Censo Escolar/INEP) |
| <i>renda_pc</i> | Renda média municipal per capita (em log) (PNUD) |
| <i>tx_des</i> | Taxa de desocupação da população com 10 anos ou mais de idade (% da população economicamente ativa que estava desocupada na semana anterior à data do Censo, mas que havia procurado trabalho no mês anterior) (PNUD) |
| <i>pob_crianca</i> | % de crianças pobres (indivíduos com até 14 anos de idade que têm renda domiciliar per capita igual ou inferior a R\$ 140,00 mensais) (PNUD) |
| <i>gini</i> | Coefficiente de gini (desigualdade na distribuição de indivíduos segundo a renda domiciliar per capita) (PNUD) |
| <i>trab_aluno</i> | % de alunos que trabalha fora de casa (Prova Brasil/INEP, questionário socioeconômico do aluno) |
| <i>prof_outro_trab</i> | % de professores que tem outra atividade fora da área de educação para complementar a renda (Prova Brasil/INEP, questionário do professor) |
| <i>tx_agua_esgoto</i> | % de pessoas em domicílios com abastecimento de água e esgotamento sanitário inadequados (PNUD) |
| <i>exp_vida</i> | Longevidade: esperança de vida ao nascer (em anos) (PNUD) |
| <i>mort_infantil</i> | Mortalidade infantil (número de crianças que não deverão sobreviver ao primeiro ano de vida em cada 1.000 crianças nascidas vivas) (PNUD) |
| <i>tx_analf</i> | Taxa de analfabetismo (não sabe ler nem escrever um bilhete simples) entre pessoas de 15 anos ou mais (PNUD) |
| <i>reuniao_pais</i> | % de pais ou responsáveis que frequentam as reuniões escolares "sempre" ou "quase sempre" (Prova Brasil/INEP, questionário socioeconômico do aluno) |

Elaboração do autor.

Em relação às demais variáveis que contribuíram para as análises empíricas, acredita-se que seja útil trabalhar com a distinção entre as covariadas exógenas e certos indicadores de gestão, conforme a seguir.

A tabela 2 mostra as variáveis exógenas que comparecem na regressão, sendo a proporção de escolas urbanas, extraída do Censo Escolar, incluída para levar em consideração as condições distintas enfrentadas no campo e na cidade, assim como o número de matrículas.

Porém, a maior parte das variáveis são indicadores socioeconômicos obtidos no *Atlas do Desenvolvimento Humano 2013* elaborado pelo PNUD e por instituições parceiras com base nos dados do Censo Demográfico de 2010. Afora estes fatores, consideraram-se características socioeconômicas das pessoas envolvidas na produção educacional extraídas dos questionários respondidos pelos alunos e professores que participaram da Prova Brasil.

TABELA 3
Indicadores de gestão (2011)

| Variável | Descrição | Indicador ($I = 1$) |
|-----------------------|--|--|
| <i>rat_transf</i> | Transferências correntes recebidas em relação ao total de receitas tributárias do município (média entre 2008 e 2011) (fonte FINBRA/STN) | $I(rat_transf) = 1$ se valor da variável é maior que a mediana |
| <i>plano_munic</i> | Plano municipal de educação (fonte MUNIC/IBGE) | $I(plano_munic) = 1$ se elaborou plano de educação |
| <i>consorcio</i> | Consórcio público ou administrativo intermunicipal, com o estado ou a União (fonte MUNIC/IBGE) | $I(consorcio) = 1$ se há consórcio firmado |
| <i>cons_pub</i> | Consórcio público intermunicipal, com o estado ou a União (fonte MUNIC/IBGE) | $I(cons_pub) = 1$ se há consórcio público firmado |
| <i>cons_adm</i> | Consórcio administrativo intermunicipal, com o estado ou a União (fonte MUNIC/IBGE) | $I(cons_adm) = 1$ se há consórcio administrativo firmado |
| <i>conv_priv</i> | Convênio de parceria com o setor privado (fonte MUNIC/IBGE) | $I(conv_priv) = 1$ se há convênio |
| <i>apoio_priv</i> | Apoio do setor privado ou de comunidades (fonte MUNIC/IBGE) | $I(apoio_priv) = 1$ se conta com esse tipo de apoio |
| <i>reduc_abandono</i> | Programa de redução das taxas de abandono ou de reprovação (fonte Prova Brasil/INEP, questionário do diretor) | $I(reduc_abandono) = 1$ se % é maior que o limite do primeiro quartil |
| <i>reforco</i> | Desenvolve regularmente programa de apoio ou reforço de aprendizagem (fonte Prova Brasil/INEP, questionário do diretor) | $I(abandono) = 1$ se % é maior que a mediana |
| <i>dir_indicacao</i> | Diretor assumiu o cargo através de indicação (política, técnica ou outra) (fonte Prova Brasil/INEP, questionário do diretor) | $I(dir_indicacao) = 1$ se % é maior que o limite do primeiro quartil |
| <i>prof_vinculo</i> | Mais de 50% dos professores têm vínculo estável (fonte Prova Brasil/INEP, questionário do diretor) | $I(prof_vinculo) = 1$ se % é maior que o limite do primeiro quartil |

Elaboração do autor.

A tabela 3 descreve as variáveis utilizadas para comparar a distribuição dos escores de eficiência entre municípios que adotam diferentes práticas de gestão. Inicialmente examina-se se os resultados guardam relação com a composição das receitas das prefeituras. Para tal, considerou-se a razão entre as receitas de transferências correntes e a arrecadação tributária, calculada a partir das informações da base Finanças do Brasil (FINBRA), disponibilizada pela STN.

Essa análise contou principalmente com indicadores da base Pesquisa de Informações Básicas Municipais (MUNIC) do IBGE, referentes à área da educação, como a existência de consórcio com outras esferas do governo e convênio com o setor privado. Vale ressaltar que estas variáveis se referem à gestão na esfera das secretarias de educação. Além destes indicadores, consideraram-se variáveis extraídas do questionário da Prova Brasil respondido pelo diretor, bem como a existência de programas de redução das taxas de abandono e reforço escolar, para examinar a influência de variáveis relacionadas à administração dentro da escola.

3.2 PCA

A aplicação da técnica de estatística multivariada denominada PCA surgiu da necessidade de se incorporar ao modelo empírico a variabilidade percebida na infraestrutura média das escolas entre os municípios sem aumentar demais a dimensionalidade do vetor de insumos.

Faz-se uma breve exposição, seguindo Mingoti (2005), para auxiliar na interpretação dos resultados reportados na sequência.¹²

O vetor x_k de dimensão $p_k \times 1$ contém as variáveis que indicam a proporção de escolas que contam com as instalações auxiliares em um dado município. O vetor de médias é indicado por μ_k , enquanto $\Sigma_{p_k \times p_k}$ denota a matriz de covariância. Os autovalores de Σ são indicados por $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{p_k}$, sendo seus respectivos autovetores normalizados dados por e_i ($i = 1, 2, \dots, p_k$).

Seja $\tilde{x}_k = O'x_k$, em que $O = [e_1 \ e_2 \ \dots \ e_{p_k}]$ é a matriz ortogonal ($p_k \times p_k$) formada pelos autovetores de Σ (matriz de *loadings*). A matriz de covariância de \tilde{x}_k é dada por $\Omega = O'\Sigma O$, em que Ω é diagonal, sendo formada pelos autovalores ordenados de Σ .¹³ A j -ésima componente principal de Σ é dada por $\tilde{x}_{jk} = e_j'x_{jk}$, sendo $\text{var}[\tilde{x}_{jk}] = \lambda_j$ e $\text{cov}[\tilde{x}_{jk}, \tilde{x}_{lk}] = 0$ sempre que $j \neq l$. Desta forma, a matriz de dados $X_{p_k} (p_k \times N)$ deve ser multiplicada por O' para se chegar aos valores numéricos das componentes principais, conhecidos como escores.¹⁴

As estatísticas descritivas das variáveis de infraestrutura consideradas são mostradas na tabela 4 ($N = 5.563$; $p = 7$). A PCA permitiu reduzir esta dimensionalidade, pois consiste em um resumo parcimonioso da variabilidade encontrada nos dados.

12. Para uma descrição da análise de componentes principais (PCA), ver, por exemplo, Mardia, Kent e Bibby (1979).

13. Pelo teorema da decomposição espectral, pode-se dizer que Ω é similar a Σ , pois ambas têm o mesmo determinante e o mesmo traço.

14. É importante perceber que cada autovetor promove uma rotação distinta nos dados, sendo cada uma ortogonal às demais. A primeira destas rotações é a combinação linear padronizada (a soma dos quadrados dos pesos é igual à unidade) das variáveis, que obtém a maior variância possível. A segunda combinação novamente maximiza a variância, mas é restrita pela exigência de ser ortogonal ao primeiro autovetor. E assim por diante.

TABELA 4
Estatística descritiva (infraestrutura)

| Variável | N = 5.563 | | | | |
|------------------------|-----------|------|-------|----------|--------------|
| | min | max | μ | σ | σ/μ |
| <i>biblioteca</i> | 0,00 | 1,00 | 0,21 | 0,24 | 1,15 |
| <i>sala_leitura</i> | 0,00 | 1,00 | 0,11 | 0,17 | 1,63 |
| <i>lab_informatica</i> | 0,00 | 1,00 | 0,28 | 0,24 | 0,88 |
| <i>lab_ciencias</i> | 0,00 | 1,00 | 0,02 | 0,07 | 3,84 |
| <i>quadra_esportes</i> | 0,00 | 1,00 | 0,17 | 0,22 | 1,28 |
| <i>sala_professor</i> | 0,00 | 1,00 | 0,39 | 0,29 | 0,76 |
| <i>sala_diretoria</i> | 0,00 | 1,00 | 0,52 | 0,31 | 0,60 |

Fonte: Censo Escolar 2011.
Elaboração do autor.

Na prática, trabalhou-se apenas com a primeira componente principal da matriz de covariância reportada na tabela 5. Como todas as variáveis estão na mesma escala, não há necessidade de se trabalhar com a matriz de correlação. Por isso, a padronização adotada consistiu em centralizar as variáveis em torno de seus respectivos valores médios.

TABELA 5
Matriz de covariância (infraestrutura)

| | <i>biblioteca</i> | <i>sala_leitura</i> | <i>lab_informatica</i> | <i>lab_ciencias</i> | <i>quadra_esportes</i> | <i>sala_professor</i> | <i>sala_diretoria</i> |
|------------------------|-------------------|---------------------|------------------------|---------------------|------------------------|-----------------------|-----------------------|
| <i>biblioteca</i> | 0,0563 | 0,0028 | 0,0290 | 0,0050 | 0,0247 | 0,0397 | 0,0380 |
| <i>sala_leitura</i> | 0,0028 | 0,0304 | 0,0138 | 0,0022 | 0,0125 | 0,0204 | 0,0206 |
| <i>lab_informatica</i> | 0,0290 | 0,0138 | 0,0598 | 0,0050 | 0,0282 | 0,0425 | 0,0448 |
| <i>lab_ciencias</i> | 0,0050 | 0,0022 | 0,0050 | 0,0055 | 0,0047 | 0,0054 | 0,0053 |
| <i>quadra_esportes</i> | 0,0247 | 0,0125 | 0,0282 | 0,0047 | 0,0470 | 0,0399 | 0,0369 |
| <i>sala_professor</i> | 0,0397 | 0,0204 | 0,0425 | 0,0054 | 0,0399 | 0,0864 | 0,0654 |
| <i>sala_diretoria</i> | 0,0380 | 0,0206 | 0,0448 | 0,0053 | 0,0369 | 0,0654 | 0,0976 |

Elaboração do autor.

A tabela 6 traz os resultados encontrados pela PCA. A componente principal 1 (PC1), que está relacionada ao maior autovalor, resume a maior parte da variância total (62,9%) e tem uma interpretação simples, uma vez que todas as variáveis apresentam o mesmo sinal, podendo ser vista como o vetor de pesos para o índice de infraestrutura médio. A dispersão do índice PC1 é mostrada no gráfico 1.¹⁵

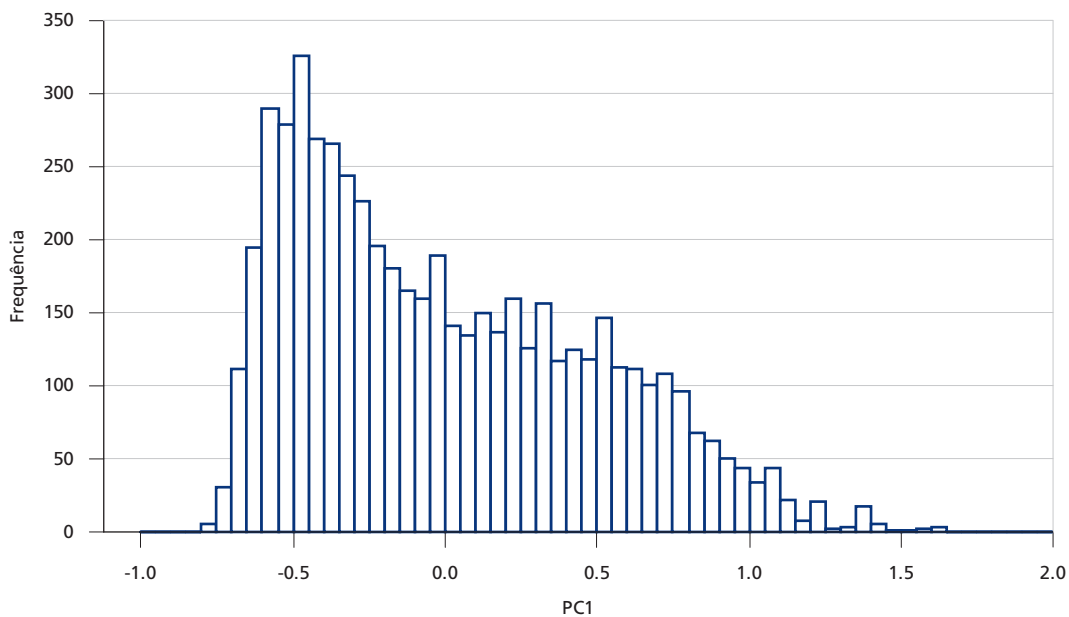
15. A variável *pca_k* é formada pelos escores de PC1 normalizados (normalização linear no intervalo [0, 1]). Esta variável satisfaz a restrição de positividade do programa. Vale notar que os municípios que obtiveram valores iguais a zero foram excluídos da amostra de dados devido à ausência de dados para outras variáveis (tabela 7).

TABELA 6
Componentes principais (infraestrutura)

| | PC1 | PC2 | PC3 | PC4 | PC5 | PC6 | PC7 |
|------------------------|------|-------|-------|-------|-------|-------|-------|
| <i>biblioteca</i> | 0,34 | 0,76 | -0,19 | -0,04 | -0,04 | 0,51 | 0,07 |
| <i>sala_leitura</i> | 0,16 | -0,51 | 0,32 | -0,11 | -0,11 | 0,76 | 0,05 |
| <i>lab_informatica</i> | 0,38 | 0,11 | 0,56 | 0,68 | -0,20 | -0,17 | 0,03 |
| <i>lab_ciencias</i> | 0,05 | 0,04 | 0,05 | 0,00 | 0,04 | 0,07 | -0,99 |
| <i>quadra_esportes</i> | 0,33 | 0,06 | 0,40 | -0,34 | 0,77 | -0,13 | 0,06 |
| <i>sala_professor</i> | 0,54 | -0,07 | 0,03 | -0,56 | -0,54 | -0,33 | -0,02 |
| <i>sala_diretoria</i> | 0,56 | -0,36 | -0,62 | 0,32 | 0,25 | -0,01 | -0,01 |

Elaboração do autor.

GRÁFICO 1
Histograma dos escores de PC1



Elaboração do autor.

As demais componentes principais (PC2 a PC7) apresentaram sinais contrários para algumas variáveis. De acordo com a interpretação usual, PC2 (10,2% da variância total) compara a proporção de escolas que tem biblioteca com os itens sala de leitura e sala da diretoria, enquanto PC7 (1,2% da variância) reflete essencialmente a proporção de escolas que conta com laboratório de ciências no município. Para interpretar as demais componentes, basta aplicar raciocínio semelhante.

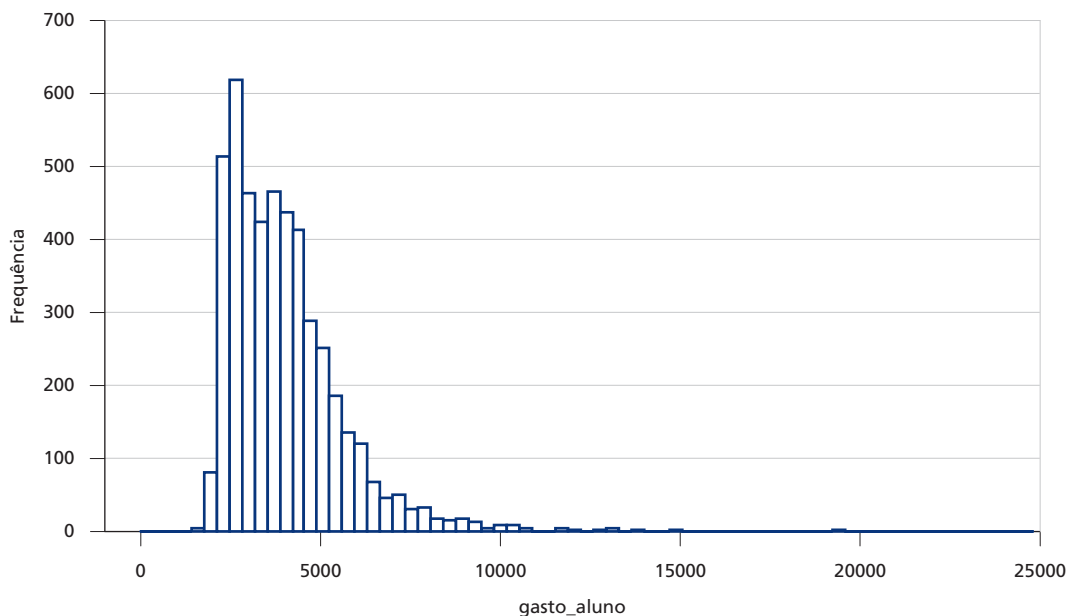
3.3 Estatística descritiva

Os escores PC1 normalizados apresentam baixa correlação com o gasto por aluno ($r \approx 0,35$). A associação do gasto por aluno é maior com o número de salas por matrícula ($r \approx 0,5$), que também mensura a infraestrutura das escolas. Embora seja baixa a correlação com o número médio de horas-aula ($r \approx 0,32$), o mesmo não se pode afirmar quanto ao número de professores com ensino superior por matrícula ($r \approx 0,52$). Estas ordens de magnitude encontradas motivaram a inclusão do gasto por aluno como um insumo adicional na função de produção. O gráfico 2 evidencia que a distribuição desta variável é bastante assimétrica.¹⁶

GRÁFICO 2

Histograma do gasto por aluno

(Em R\$ de 2011)

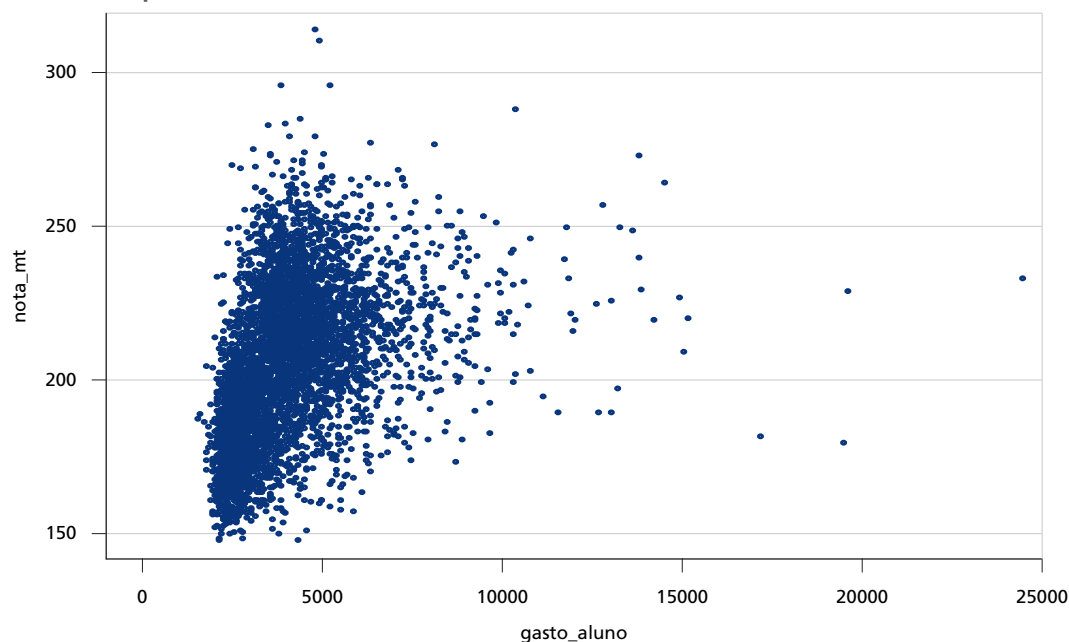


Fonte: Sistema de Informações sobre Orçamentos Públicos em Educação (Siope).
Elaboração do autor.

16. Há sinais de erros de mensuração nas despesas em manutenção e desenvolvimento do ensino (MDE) por matrícula, como valores muito elevados em certas localidades. É impossível visualizar a situação de alguns municípios, como Magda (SP), São João de Itacema (SP), Jateí (MS), Pinhal da Serra (RS), Abdon Batista (SC) e Mirante do Paranapanema (SP), que tiveram valores acima de R\$ 15 mil por ano. Vários municípios de Alagoas, Ceará, Maranhão e Pará reportaram valores médios abaixo de R\$ 2 mil (preços de 2011).

Cumpra-se notar que a relação entre as despesas em manutenção e desenvolvimento do ensino (MDE) por matrícula e a nota média em matemática na Prova Brasil é relativamente fraca ($r \approx 0,43$). O mesmo pode-se dizer em relação aos demais produtos. A associação com o inverso da taxa de abandono é semelhante ($r \approx 0,45$), porém o inverso da taxa de distorção apresentou correlação mais baixa ($r \approx 0,35$).

GRÁFICO 3
Gasto por aluno *versus* nota de matemática na Prova Brasil



Fonte: Siope e Prova Brasil 2011.
Elaboração do autor.

A fim de ilustrar esse fato, o gráfico 3 mostra a relação do gasto por aluno e o desempenho médio no teste padronizado por município. Fica claro que tal associação é mais relevante para níveis baixos de gasto em MDE. Isto reforça a ideia de que municípios à direita da distribuição são possivelmente ineficientes, pois gastam mal os recursos públicos.

TABELA 7
Estatística descritiva (insumos e produtos)

| Variável | N = 4.757 | | | | |
|-------------------------|-----------|-----------|----------|----------|--------------|
| | min | max | μ | σ | σ/μ |
| <i>prof_es</i> | 0,00 | 0,14 | 0,04 | 0,02 | 0,51 |
| <i>horas_aula</i> | 3,50 | 9,70 | 4,31 | 0,50 | 0,12 |
| <i>num_salas</i> | 0,00 | 0,18 | 0,01 | 0,01 | 1,23 |
| <i>pca_k</i> | 0,01 | 1,00 | 0,32 | 0,21 | 0,64 |
| <i>gasto_aluno</i> | 1.511,44 | 24.456,35 | 4.007,08 | 1.673,08 | 0,42 |
| <i>nota_mt</i> | 147,94 | 314,46 | 204,29 | 25,61 | 0,13 |
| <i>inv_tx_distorcao</i> | 1,46 | 333,33 | 9,35 | 13,93 | 1,49 |
| <i>inv_tx_abandono</i> | 4,29 | 1.000,00 | 370,90 | 403,15 | 1,09 |

Fonte: Censo Escolar, Siope e Prova Brasil.
Elaboração do autor.

A tabela 7 faz a descrição dos insumos e produtos da educação pública utilizados no programa DEA. A tabela 8 apresenta as estatísticas relevantes das variáveis exógenas utilizadas no modelo econométrico. Por fim, a tabela 9 faz análise semelhante para as variáveis de gestão.

TABELA 8
Estatística descritiva (variáveis exógenas)¹

| Variável | N = 4.757 | | | | |
|------------------------|-----------|------|-------|----------|--------------|
| | min | max | μ | σ | σ/μ |
| <i>urbana</i> | 0,00 | 1,00 | 0,47 | 0,33 | 0,70 |
| <i>lmat</i> | -2,63 | 6,29 | 0,36 | 1,17 | 3,28 |
| <i>lrenda_pc</i> | -0,04 | 3,02 | 1,46 | 0,50 | 0,34 |
| <i>tx_des</i> | 0,00 | 0,42 | 0,07 | 0,04 | 0,53 |
| <i>pob_crianca</i> | 0,00 | 0,85 | 0,34 | 0,22 | 0,65 |
| <i>gini</i> | 0,28 | 0,80 | 0,50 | 0,07 | 0,13 |
| <i>trab_aluno</i> | 0,00 | 0,62 | 0,17 | 0,08 | 0,47 |
| <i>prof_outro_trab</i> | 0,00 | 1,00 | 0,08 | 0,16 | 1,90 |
| <i>tx_agua_esgoto</i> | 0,00 | 0,85 | 0,10 | 0,13 | 1,36 |
| <i>exp_vida</i> | 0,65 | 0,79 | 0,73 | 0,03 | 0,04 |
| <i>mort_infantil</i> | 0,09 | 0,47 | 0,20 | 0,07 | 0,37 |
| <i>tx_analf</i> | 0,01 | 0,44 | 0,16 | 0,10 | 0,60 |
| <i>reuniao_pais</i> | 0,05 | 1,00 | 0,59 | 0,11 | 0,19 |

Elaboração do autor.

Nota: ¹ Variáveis que denotam % foram expressas como proporção. O número de matrículas foi transformado para milhares antes de tirar o logaritmo. Da mesma forma, a renda por habitante foi convertida para centenas de R\$. A expectativa de vida e a mortalidade infantil também foram divididas por 100.

Cabe observar que o total de 4.757 observações é menor que o total de municípios existentes no Brasil (5.565 em 2011). Isto se deve principalmente ao fato de que, mesmo após tirar a média das variáveis relevantes dentro de cada município (ignorando a falta de dados para algumas escolas), detectaram-se ausências de dados para algumas variáveis.

TABELA 9
Estatística descritiva (indicadores de gestão)

| Variável | N = 4.757 | | | | | | |
|-----------------------|-----------|------------|---------|-------|------------|--------|-----------|
| | min | 1º quartil | mediana | μ | 3º quartil | max | N (l = 1) |
| <i>rat_transf</i> | 0,56 | 11,94 | 22,68 | 27,78 | 36,82 | 759,88 | 2.378 |
| <i>plano_munic</i> | 0,00 | 0,00 | 1,00 | 0,62 | 1,00 | 1,00 | 2.927 |
| <i>consorcio</i> | 0,00 | 0,00 | 0,00 | 0,34 | 1,00 | 1,00 | 1.625 |
| <i>cons_pub</i> | 0,00 | 0,00 | 0,00 | 0,13 | 0,00 | 1,00 | 620 |
| <i>cons_adm</i> | 0,00 | 0,00 | 0,00 | 0,24 | 0,00 | 1,00 | 1.149 |
| <i>conv_priv</i> | 0,00 | 0,00 | 0,00 | 0,13 | 0,00 | 1,00 | 639 |
| <i>apoio_priv</i> | 0,00 | 0,00 | 0,00 | 0,08 | 0,00 | 1,00 | 401 |
| <i>reduc_abandono</i> | 0,00 | 0,38 | 0,79 | 0,66 | 1,00 | 1,00 | 2.378 |
| <i>reforco</i> | 0,00 | 0,67 | 1,00 | 0,80 | 1,00 | 1,00 | 3.556 |
| <i>dir_indicacao</i> | 0,00 | 0,57 | 1,00 | 0,75 | 1,00 | 1,00 | 3.571 |
| <i>prof_vinculo</i> | 0,00 | 0,71 | 1,00 | 0,81 | 1,00 | 1,00 | 3.540 |

Elaboração do autor.

4 RESULTADOS

Na sequência reportam-se os resultados encontrados por meio da técnica DEA em dois estágios. Inicialmente, mostra-se o impacto da reamostragem nos escores de eficiência. Em seguida, apresentam-se os resultados das regressões destes índices nos fatores ambientais, bem como a dispersão da ineficiência residual no território brasileiro. Por fim, avalia-se a existência de diferenças entre grupos formados por características institucionais relacionadas a práticas de gestão na área da educação pública.¹⁷

17. A implementação da técnica *jackstrap* descrita em Stošić e Sampaio de Sousa (2003) e Sampaio de Sousa e Stošić (2005) contou com o aplicativo disponibilizado por Borko Stošić em sua página pessoal na internet: <<http://goo.gl/wyZKFQ>>. A aplicação do método DEA semiparamétrico se valeu dos códigos em R escritos por Paul Wilson. O pacote FEAR é disponibilizado em sua página pessoal: <<http://goo.gl/34vngt>>. Ver também Wilson (2008).

4.1 Reamostragem

A base de dados compreende 4.757 municípios. Estas DMUs foram submetidas ao teste de sensibilidade dos escores com o auxílio da técnica *jackstrap*. Para tal, foram construídas $B = 2.000$ réplicas de amostras aleatórias (bolhas) de tamanho $L = 500$ para calcular a alavancagem média de cada DMU, assim como o valor de limiar, conforme os passos descritos na seção 2.2. Apesar do modelo empírico se basear em retornos variáveis de escala, excluíram-se os municípios apontados como *outliers* nos três cenários mais comuns para a tecnologia de produção.¹⁸

A técnica *jackstrap* identificou 371 DMUs que podem ser consideradas influentes (7,8% do total). Destas, 292 foram identificadas sob a tecnologia de retornos constantes, enquanto sob retornos variáveis e não crescentes de escala foram encontrados, respectivamente, 294 e 272 pontos influentes. Como esperado, estas DMUs contavam, tipicamente, com valores baixos para determinados insumos ou elevados para certos produtos. Como exerciam influência excessiva sobre a fronteira DEA, de acordo com a regra adotada, foram excluídas da base de dados, que passou a contar com 4.386 municípios.

A tabela 10 mostra o total de municípios por estado considerados influentes sob retornos constantes (*crs*), variáveis (*vrs*) e não crescentes (*nrs*). Apresenta também a alavancagem das DMUs excluídas (*leverage*), quantas vezes estas foram selecionadas em uma bolha (*total hits*), assim como o número de vezes que tiveram alavancagem diferente de zero (*leverage hits*). Estes números dizem respeito a médias dos municípios excluídos dentro de cada estado.

18. Sob a tecnologia de retornos constantes de escala, a alavancagem das DMUs que se situam na fronteira tende a ser maior. Isto faz com que a alavancagem média e o valor de limiar da alavancagem também sejam mais elevados. Como consequência, certas DMUs consideradas influentes sob a tecnologia de retornos variáveis e não crescentes não são identificadas quando implementada a técnica *jackstrap* somente para retornos constantes de escala.

TABELA 10
Resultados da técnica *jackstrap*¹

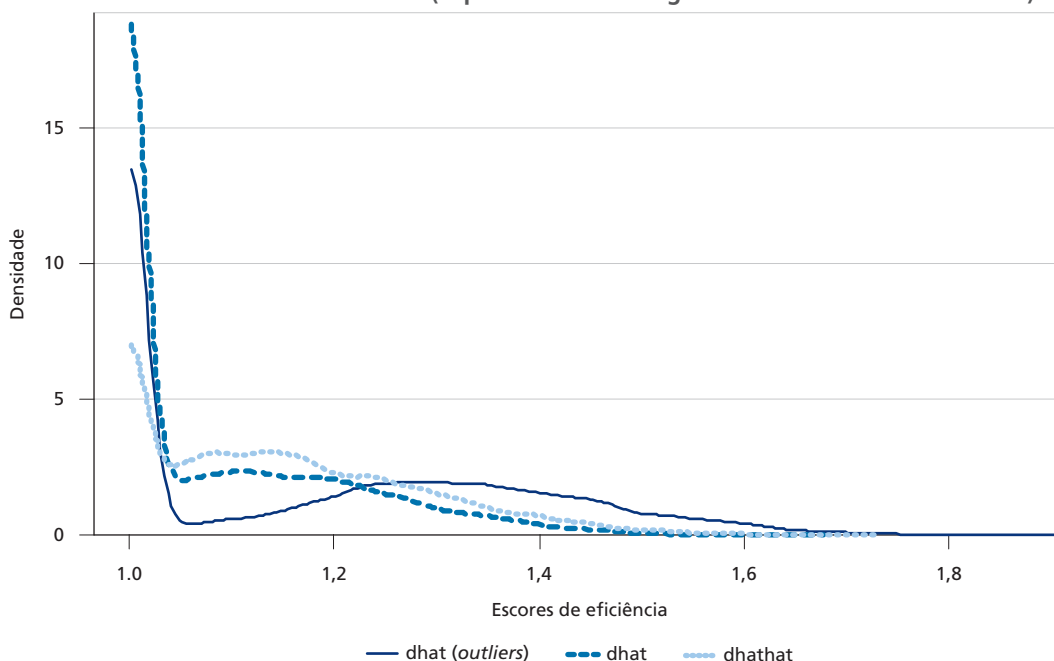
| Região | UF | Nº de municípios | crs (limiar = 0.975×10^{-3}) | | | vrs (limiar = 0.930×10^{-3}) | | | nrs (limiar = 0.756×10^{-3}) | | |
|--------|----|------------------|--|------------|---------------|--|------------|---------------|--|------------|---------------|
| | | | Leverage $\times 10^3$ | Total hits | Leverage hits | Leverage $\times 10^3$ | Total hits | Leverage hits | Leverage $\times 10^3$ | Total hits | Leverage hits |
| CO | GO | 2 | 1,99 | 226,5 | 179,0 | 1,11 | 226,5 | 172,5 | 1,14 | 226,5 | 173,5 |
| | MT | 2 | 1,44 | 207,0 | 175,5 | 0,43 | 207,0 | 123,0 | 0,21 | 207,0 | 94,0 |
| | MS | 1 | 1,91 | 189,0 | 150,0 | 1,56 | 189,0 | 151,0 | 1,93 | 189,0 | 157,0 |
| NE | AL | 4 | 0,55 | 215,0 | 110,3 | 1,62 | 215,0 | 143,3 | 0,41 | 215,0 | 102,5 |
| | BA | 61 | 2,85 | 211,9 | 137,6 | 3,34 | 211,9 | 165,7 | 2,33 | 211,9 | 137,3 |
| | CE | 29 | 6,32 | 211,0 | 169,3 | 5,24 | 211,0 | 164,3 | 4,80 | 211,0 | 153,2 |
| | MA | 11 | 2,95 | 202,5 | 130,0 | 3,03 | 202,5 | 157,5 | 1,67 | 202,5 | 133,4 |
| | PB | 2 | 3,73 | 195,5 | 177,0 | 4,03 | 195,5 | 177,0 | 3,12 | 195,5 | 176,0 |
| | PE | 13 | 2,09 | 212,0 | 128,6 | 2,48 | 212,0 | 150,6 | 2,29 | 212,0 | 142,4 |
| | PI | 8 | 3,20 | 210,8 | 122,9 | 3,34 | 210,8 | 155,9 | 2,65 | 210,8 | 110,5 |
| N | AC | 2 | 0,93 | 228,0 | 133,5 | 1,37 | 228,0 | 160,5 | 1,34 | 228,0 | 159,0 |
| | AM | 8 | 5,32 | 214,1 | 170,5 | 5,35 | 214,1 | 170,9 | 4,91 | 214,1 | 167,5 |
| | PA | 31 | 3,63 | 206,9 | 154,8 | 4,11 | 206,9 | 165,3 | 2,86 | 206,9 | 153,8 |
| | RO | 1 | 1,39 | 193,0 | 150,0 | 0,51 | 193,0 | 101,0 | 0,46 | 193,0 | 98,0 |
| | TO | 4 | 1,97 | 201,0 | 167,8 | 0,78 | 201,0 | 120,3 | 0,60 | 201,0 | 105,0 |
| SE | ES | 2 | 2,10 | 204,0 | 167,5 | 0,75 | 204,0 | 140,5 | 0,45 | 204,0 | 148,0 |
| | MG | 64 | 2,45 | 209,9 | 150,5 | 2,66 | 209,9 | 160,8 | 2,83 | 209,9 | 161,4 |
| | RJ | 17 | 1,37 | 213,6 | 148,5 | 1,34 | 213,6 | 157,4 | 1,36 | 213,6 | 156,9 |
| | SP | 46 | 2,84 | 213,1 | 144,7 | 2,02 | 213,1 | 158,2 | 1,86 | 213,1 | 159,2 |
| S | PR | 37 | 3,19 | 209,9 | 166,4 | 2,29 | 209,9 | 157,3 | 1,86 | 209,9 | 146,5 |
| | RS | 7 | 1,48 | 203,7 | 162,1 | 1,11 | 203,7 | 152,4 | 0,77 | 203,7 | 139,4 |
| | SC | 19 | 2,74 | 208,1 | 166,8 | 1,47 | 208,1 | 143,1 | 1,02 | 208,1 | 130,9 |

Elaboração do autor.
Nota: ¹ $L = 500$; $B = 2.000$.

Em relação ao impacto da exclusão desses municípios, vale chamar atenção para o fato de que a ineficiência média passou de 23,6% para 11,7%, isto é, caiu praticamente pela metade. No extremo da distribuição, nota-se que a ineficiência máxima diminuiu de 90,9% para 67,8%. Já o número de DMUs eficientes elevou-se de 1.371 para 1.508 (28,8% e 34,4% do total de prefeituras consideradas em cada um destes exercícios). Sem dúvida, a estimativa deste estudo para a expansão radial factível dos produtos da educação é conservadora. Este nível de desperdício não é alto se comparado a resultados de outros estudos sobre a mesma temática.

GRÁFICO 4

Estimativas kernel das densidades (impacto da reamostragem sobre os escores de eficiência)¹



Elaboração do autor.

Nota: ¹Kernel gaussiano aliado ao método *reflection*.

Como exposto na seção 2.3, a técnica DEA resulta em escores viesados em amostras finitas, subestimando a real ineficiência. Visando corrigir esta distorção, empregou-se o método *bootstrap*, que possibilitou estimar os escores livres de viés ($\hat{\delta}$) via reamostragem. Com isso, a ineficiência média encontrada passou para 16,9%, enquanto seu valor máximo chegou a 73,1%. O gráfico 4 evidencia o impacto das técnicas de reamostragem usadas sobre toda a distribuição dos escores.

A tabela 11 mostra um *ranking* simplificado dos dez municípios mais eficientes (painel A) e das dez localidades mais ineficientes (painel B), assim como os valores dos insumos e produtos destas DMUs. Entre os dez mais eficientes, sete estão localizados no estado de São Paulo. Já entre os dez menos eficientes, cinco são do Rio Grande do Norte e outros três também são da região Nordeste.

Cabe observar, no entanto, que esse *ranking* não faz qualquer distinção em relação às realidades socioeconômicas enfrentadas pelas populações desses municípios, ao passo que há ampla evidência na literatura sobre educação e eficiência acerca da importância de tais fatores.

TABELA 11
Ranking parcial (bootstrapped escores)

| UF | Município | $\hat{\delta}$ | Insumos | | | | | Produtos | | |
|-------------------------------|---------------------|----------------|---------|------------|-----------|-------------|-------|----------|------------------|-----------------|
| | | | prof_es | horas_aula | num_salas | gasto_aluno | pca_k | nota_mt | inv_tx_distorcao | inv_tx_abandono |
| Painel A: 10 mais eficientes | | | | | | | | | | |
| SP | Emilianópolis | 1,00 | 0,06 | 5,00 | 0,05 | 4.801,27 | 0,65 | 200,58 | 9,62 | 1.000,00 |
| GO | Lagoa Santa | 1,00 | 0,07 | 4,80 | 0,04 | 10.263,69 | 0,58 | 215,01 | 6,25 | 1.000,00 |
| SP | Luís Antônio | 1,00 | 0,06 | 5,00 | 0,01 | 4.484,82 | 0,73 | 204,06 | 6,10 | 1.000,00 |
| GO | Aporé | 1,00 | 0,06 | 8,10 | 0,03 | 8.838,58 | 0,64 | 227,59 | 6,49 | 1.000,00 |
| PR | Lupionópolis | 1,00 | 0,08 | 4,70 | 0,02 | 6.460,28 | 0,61 | 219,86 | 6,49 | 1.000,00 |
| SP | Mineiros do Tietê | 1,00 | 0,07 | 7,20 | 0,02 | 5.417,19 | 0,75 | 222,81 | 13,16 | 1.000,00 |
| SP | Nova Luzitânia | 1,00 | 0,05 | 5,00 | 0,04 | 5.584,47 | 0,77 | 215,24 | 13,89 | 1.000,00 |
| SP | Platina | 1,00 | 0,06 | 5,20 | 0,03 | 6.205,28 | 0,61 | 222,52 | 13,33 | 1.000,00 |
| SP | Taruá | 1,00 | 0,06 | 6,70 | 0,01 | 5.934,65 | 0,74 | 227,87 | 6,94 | 1.000,00 |
| SP | Ipeúna | 1,00 | 0,05 | 5,80 | 0,03 | 5.379,76 | 0,59 | 218,79 | 12,20 | 1.000,00 |
| Painel B: 10 menos eficientes | | | | | | | | | | |
| RN | Lagoa d'Anta | 1,61 | 0,04 | 4,50 | 0,01 | 3.029,53 | 0,24 | 158,71 | 2,57 | 17,54 |
| TO | Ananás | 1,62 | 0,03 | 4,30 | 0,01 | 5.469,67 | 0,34 | 160,85 | 9,35 | 250,00 |
| PB | Pedro Régis | 1,64 | 0,05 | 4,30 | 0,01 | 3.034,84 | 0,25 | 157,06 | 2,64 | 17,54 |
| RN | São Bento do Trairi | 1,64 | 0,04 | 4,50 | 0,01 | 3.541,99 | 0,15 | 161,96 | 3,70 | 50,00 |
| SE | General Maynard | 1,68 | 0,04 | 4,10 | 0,01 | 5.225,63 | 0,39 | 158,98 | 2,54 | 40,00 |
| RN | Várzea | 1,69 | 0,05 | 4,00 | 0,02 | 3.802,69 | 0,23 | 150,25 | 3,23 | 18,52 |
| RN | Lagoa de Velhos | 1,72 | 0,04 | 4,50 | 0,01 | 3.619,74 | 0,23 | 154,64 | 2,68 | 31,25 |
| PB | Poço Dantas | 1,73 | 0,04 | 4,30 | 0,01 | 3.606,93 | 0,31 | 151,46 | 2,51 | 30,30 |
| RN | Messias Targino | 1,73 | 0,04 | 4,50 | 0,02 | 3.904,70 | 0,17 | 157,50 | 3,68 | 21,28 |
| MT | Santa Cruz do Xingu | 1,73 | 0,06 | 4,00 | 0,03 | 5.487,76 | 0,42 | 157,70 | 4,00 | 16,39 |

Elaboração do autor.

4.2 Fatores ambientais

A influência dos fatores ambientais nos índices de eficiência foi estimada com o auxílio de métodos econométricos. Tal análise se justifica pelo fato de as variáveis consideradas estarem, no curto prazo, fora do controle dos gestores municipais. A tabela 11 apresenta os resultados da regressão tobit. Cumpre notar que estas estimativas não levam em consideração nem o viés dos escores, nem a correlação serial entre eles, que compromete a eficiência dos estimadores.

De acordo com esses resultados, a proporção de escolas urbanas, a renda *per capita* municipal, a proporção de professores que complementa a renda em atividade não ligada à área da educação, além de todos os indicadores das condições de saneamento e saúde da população local (acesso à água e saneamento inadequados, expectativa de vida e mortalidade infantil), não influenciam o retorno dos recursos públicos aplicado nas escolas.

TABELA 12
Regressão tobit nos fatores ambientais

| Variável dependente: $\ln \delta$ | Tobit 0 | | | |
|-----------------------------------|---------|------------------------|-----------------------|--------|
| | beta | std error ¹ | 95% C.I. ¹ | |
| | | | inf | sup |
| <i>Intercepto</i> | -0,170 | 0,266 | -0,692 | 0,352 |
| <i>urbana</i> | -0,017 | 0,010 | -0,037 | 0,002 |
| <i>lmat</i> | -0,012 | 0,002 | -0,016 | -0,007 |
| <i>renda_pc</i> | -0,030 | 0,020 | -0,069 | 0,008 |
| <i>tx_des</i> | 0,554 | 0,063 | 0,430 | 0,677 |
| <i>pob_crianca</i> | -0,103 | 0,051 | -0,203 | -0,002 |
| <i>gini</i> | 0,355 | 0,059 | 0,240 | 0,471 |
| <i>trab_aluno</i> | 0,281 | 0,031 | 0,220 | 0,342 |
| <i>prof_outro_trab</i> | 0,020 | 0,013 | -0,006 | 0,045 |
| <i>tx_agua_esgoto</i> | -0,034 | 0,024 | -0,081 | 0,013 |
| <i>exp_vida</i> | 0,150 | 0,337 | -0,510 | 0,810 |
| <i>mort_infantil</i> | 0,143 | 0,122 | -0,096 | 0,382 |
| <i>tx_analf</i> | 0,329 | 0,051 | 0,230 | 0,428 |
| <i>reuniao_pais</i> | -0,205 | 0,020 | -0,245 | -0,166 |

N = 4,386
LogLik = 564.34, R² = 0.0492, σ = 0.1040

Elaboração do autor.

Nota: ¹Erros-padrão assintóticos.

A tabela 13 apresenta os resultados seguindo o algoritmo de Simar e Wilson (2007). Na regressão 1, a taxa de pobreza entre as crianças não é diferente de zero, enquanto a expectativa de vida e a mortalidade infantil passaram a ser significantes. É importante notar que, enquanto na estimativa inicial os fatores ambientais explicavam menos de 5% da variância dos escores de eficiência, agora estas mesmas variáveis exógenas explicam quase 30% da variância total. A regressão 2 refaz as estimativas apenas com as variáveis significantes a 5%.

TABELA 13
Regressões tobit nos fatores ambientais (*bootstrap duplo*)

| Variável dependente: $\ln \delta$ | Tobit 1 (variáveis ambientais) | | | | Tobit 2 (variáveis significantes) | | | |
|-----------------------------------|--|------------------------|-----------------------|--------|---|------------------------|-----------------------|--------|
| | beta | std error ¹ | 95% C.I. ¹ | | beta | std error ¹ | 95% C.I. ¹ | |
| | | | inf | sup | | | inf | sup |
| <i>Intercepto</i> | -0,295 | 0,021 | -0,593 | -0,026 | -0,275 | 0,020 | -0,590 | -0,029 |
| <i>urbana</i> | -0,012 | 0,000 | -0,018 | 0,003 | - | - | - | - |
| <i>lmat</i> | -0,014 | 0,000 | -0,013 | -0,009 | -0,015 | 0,000 | -0,014 | -0,009 |
| <i>lrenda_pc</i> | -0,014 | 0,000 | -0,028 | 0,017 | - | - | - | - |
| <i>tx_des</i> | 0,420 | 0,002 | 0,281 | 0,430 | 0,407 | 0,001 | 0,272 | 0,414 |
| <i>pob_crianca</i> | -0,043 | 0,001 | -0,080 | 0,039 | - | - | - | - |
| <i>gini</i> | 0,205 | 0,001 | 0,079 | 0,211 | 0,192 | 0,000 | 0,099 | 0,187 |
| <i>trab_aluno</i> | 0,225 | 0,000 | 0,148 | 0,220 | 0,232 | 0,000 | 0,154 | 0,225 |
| <i>prof_outro_trab</i> | 0,014 | 0,000 | -0,004 | 0,027 | - | - | - | - |
| <i>tx_agua_esgoto</i> | 0,013 | 0,000 | -0,016 | 0,043 | - | - | - | - |
| <i>exp_vida</i> | 0,418 | 0,033 | 0,106 | 0,824 | 0,352 | 0,031 | 0,105 | 0,795 |
| <i>mort_infantil</i> | 0,293 | 0,005 | 0,162 | 0,431 | 0,282 | 0,005 | 0,163 | 0,429 |
| <i>tx_analf</i> | 0,200 | 0,001 | 0,107 | 0,227 | 0,214 | 0,001 | 0,129 | 0,219 |
| <i>reuniao_pais</i> | -0,132 | 0,000 | -0,124 | -0,081 | -0,137 | 0,000 | -0,128 | -0,086 |
| | N = 4,386 | | | | N = 4,386 | | | |
| | LogLik = 4323.00, R ² = 0.2826, σ = 0.0903 | | | | LogLik = 4324.34, R ² = 0.2867, σ = 0.09028 | | | |

Elaboração do autor.

Nota: ¹Erros-padrão e intervalos de confiança *bootstrap*.

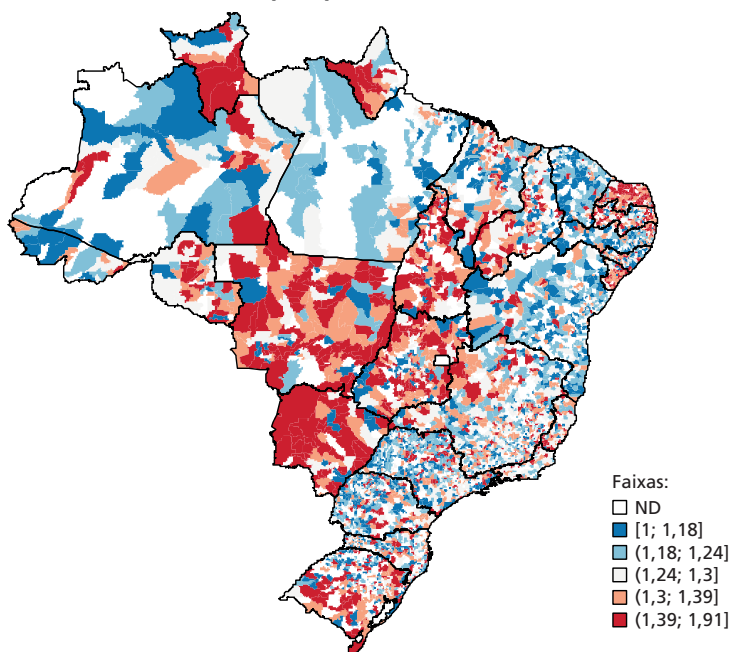
Mesmo trabalhando com retornos variáveis de escala, observou-se que, quanto maior o número de matrículas na rede municipal, menor é a ineficiência. Isto sugere a existência de ganhos de escala na educação pública no Brasil.

Em relação às demais variáveis, nota-se que a maior parte delas apresentou o sinal esperado segundo a teoria. A exceção fica a cargo da expectativa de vida, pois níveis mais elevados de esperança de vida ao nascer estão associados a maiores desperdícios, o que surpreende. Contudo, a mortalidade infantil obteve o sinal correto, corroborando com a ideia de que a saúde na infância está associada a uma educação pública mais eficiente. Embora o acesso inadequado da população a água e esgoto não tenha sido significativo, vale notar que a mortalidade infantil também está correlacionada com a oferta de saneamento básico.

A condição socioeconômica da população está positivamente relacionada à eficiência da rede de ensino. Apesar de a renda *per capita* não ter sido significativa, esta dimensão importante comparece no modelo empírico por intermédio da taxa de desemprego e da desigualdade de renda, além da proporção de crianças que trabalham. A taxa de analfabetismo também atua nesta arena. Certamente, a influência de tais fatores retrata a importância dos insumos familiares na vida escolar. Por fim, destaca-se que melhores resultados são atingidos quando a frequência dos pais a reuniões escolares é alta.

A dispersão da ineficiência residual (δ) – normalização empreendida conforme equação (7) da seção 2.1 – no território brasileiro, de acordo com os quintis da distribuição, é mostrada no mapa 1, que possibilita visualizar a distribuição regional da ineficiência propriamente dita, isto é, livre da influência de fatores socioeconômicos, entre outras variáveis ambientais. A escala de cores adotada enfoca os quintis extremos da distribuição.

MAPA 1
Escores de ineficiência residuais por quintis no Brasil



Elaboração do autor.
Obs.: ND = não disponível.

A região Centro-Oeste do país concentra boa parte dos desperdícios, pois em Goiás, Mato Grosso e Mato Grosso do Sul são raros os casos de municípios com escores localizados no 1º ou 2º quintil da distribuição. Contudo, é possível visualizar *clusters* do 4º e 5º quintis nas demais regiões. Entretanto, existem estados com redes municipais de ensino razoavelmente eficientes em quase todas as regiões (a exceção fica a cargo do Centro-Oeste).

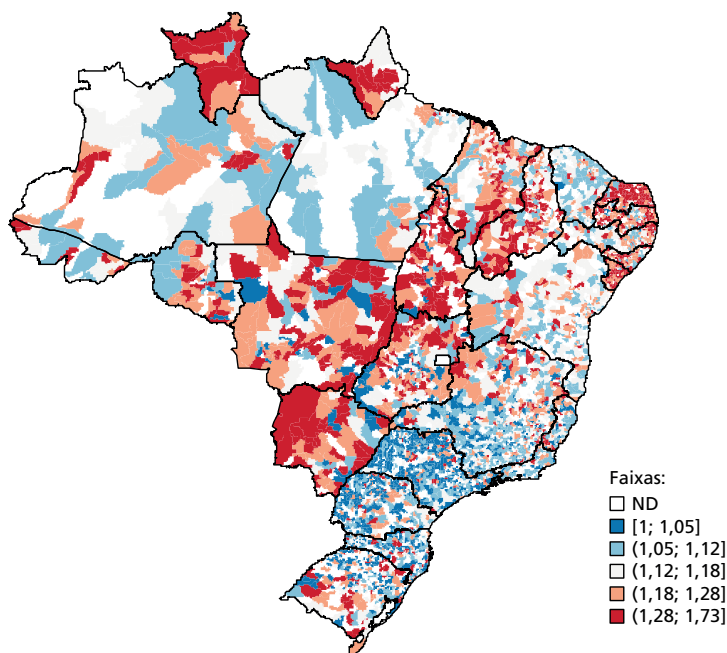
No Norte do país, os municípios do Amapá, Rondônia, Roraima e Tocantins são claramente ineficientes. Já no Amazonas, Acre e Pará – apesar da alta proporção de estimativas não disponíveis (NDs) por falta de dados ou por caracterizarem *outliers*, sobretudo no Pará –, há prefeituras que aparentam aproveitar bem os recursos disponíveis. Na região Nordeste, os destaques positivos ficam a cargo do Ceará e da Bahia (ambos têm alta proporção de NDs), além de Pernambuco, ao passo que no Rio Grande do Norte e em Sergipe, além de certas localidades da Paraíba e do Piauí, o nível de desperdício aparenta ser bem elevado.

No Sudeste, as redes de ensino municipais apresentam bom nível de eficiência em São Paulo. Minas Gerais apresentou alta dispersão dos municípios no *ranking* criado segundo os quintis. Já as redes municipais do Espírito Santo e do Rio de Janeiro tiveram, em grande medida, desempenho mediano (3º quintil) – cabendo notar que no Rio de Janeiro a proporção de NDs é alta. Por fim, no Sul do país, destacou-se a situação do Rio Grande do Sul, devido à predominância de municípios no 4º e 5º quintis. Nesta região, há proporção alta de municípios com redes de ensino eficientes no Paraná.

O mapa 2 mostra a dispersão dos escores ($\hat{\delta}$) calculados pelo programa DEA. Este mapa deve ser contrastado com o mapa 1, que apresenta os índices de ineficiência residual ($\tilde{\delta}$). Fica claro que os fatores ambientais fazem toda a diferença na mensuração da eficiência relativa. O Nordeste deixou de ser predominantemente ineficiente, enquanto certas localidades do Sul e Sudeste tiveram visível queda no *ranking* apurado pelos quintis. O próprio Norte, em certa medida, se beneficiou da correção empreendida nos escores. A situação do Centro-Oeste, que não era boa, ficou ainda pior em termos relativos.

MAPA 2

Escores de ineficiência DEA por quintis no Brasil¹



Elaboração do autor.
Obs.: ND = não disponível.

4.3 Variáveis de gestão

Os fatores ambientais considerados nas análises prévias estão fora do controle dos gestores municipais. No entanto, há práticas de gestão da prefeitura e das próprias escolas que podem impactar as taxas de aproveitamento dos recursos disponíveis. Com o intuito de verificar se há correlação entre os indicadores de gestão e os escores de eficiência relativa, empregaram-se duas estratégias empíricas diferentes, conforme mostrado na sequência.

A maneira mais natural de testar as hipóteses de interesse é colocando as novas variáveis no modelo econométrico e examinando as significâncias destes fatores. A tabela 14 mostra as estimativas obtidas com a implementação do algoritmo descrito na seção 2.2. De acordo com estas estimativas, a razão entre transferências e receitas tributárias e a existência de consórcio público, consórcio administrativo e convênio privado, além da proporção de diretores que chegaram ao cargo por indicação e a proporção de professores que complementam a renda com outra atividade fora da área de educação, são indiferentes no que se refere à eficiência das redes municipais de ensino.

TABELA 14
Regressões tobit nos fatores ambientais e indicadores de gestão (*bootstrap duplo*)

| Variável dependente: $\ln \delta$ | Tobit 3 (variáveis ambientais e indicadores de gestão) | | | | Tobit 4 (variáveis significantes) | | | |
|-----------------------------------|--|------------------------|-----------------------|--------|--|------------------------|-----------------------|--------|
| | beta | std error ¹ | 95% C.I. ¹ | | beta | std error ¹ | 95% C.I. ¹ | |
| | | | inf | sup | | | inf | sup |
| <i>Intercepto</i> | -0,191 | 0,020 | -0,508 | 0,060 | -0,157 | 0,020 | -0,489 | 0,070 |
| <i>urbana</i> | -0,009 | 0,000 | -0,016 | 0,006 | - | - | - | - |
| <i>lmat</i> | -0,015 | 0,000 | -0,014 | -0,009 | -0,015 | 0,000 | -0,014 | -0,009 |
| <i>lrenda_pc</i> | -0,019 | 0,000 | -0,033 | 0,012 | - | - | - | - |
| <i>tx_des</i> | 0,414 | 0,002 | 0,278 | 0,425 | 0,403 | 0,001 | 0,272 | 0,413 |
| <i>pob_crianca</i> | -0,053 | 0,001 | -0,090 | 0,027 | - | - | - | - |
| <i>gini</i> | 0,207 | 0,001 | 0,083 | 0,214 | 0,181 | 0,000 | 0,092 | 0,178 |
| <i>trab_aluno</i> | 0,217 | 0,000 | 0,141 | 0,213 | 0,222 | 0,000 | 0,147 | 0,218 |
| <i>prof_outro_trab</i> | 0,010 | 0,000 | -0,006 | 0,024 | - | - | - | - |
| <i>tx_agua_esgoto</i> | 0,007 | 0,000 | -0,020 | 0,038 | - | - | - | - |
| <i>exp_vida</i> | 0,365 | 0,033 | 0,072 | 0,790 | 0,275 | 0,031 | 0,037 | 0,726 |
| <i>mort_infantil</i> | 0,248 | 0,005 | 0,124 | 0,394 | 0,215 | 0,005 | 0,105 | 0,369 |
| <i>tx_analf</i> | 0,181 | 0,001 | 0,091 | 0,211 | 0,183 | 0,001 | 0,102 | 0,193 |
| <i>reuniao_pais</i> | -0,123 | 0,000 | -0,116 | -0,074 | -0,128 | 0,000 | -0,120 | -0,078 |
| <i>rat_transf</i> | 0,000 | 0,000 | 0,000 | 0,000 | - | - | - | - |
| <i>plano_munic</i> | -0,010 | 0,000 | -0,013 | -0,004 | -0,010 | 0,000 | -0,013 | -0,004 |
| <i>cons_pub</i> | -0,008 | 0,000 | -0,013 | 0,001 | - | - | - | - |
| <i>cons_adm</i> | -0,004 | 0,000 | -0,009 | 0,002 | - | - | - | - |
| <i>conv_priv</i> | 0,000 | 0,000 | -0,007 | 0,007 | - | - | - | - |
| <i>apoio_priv</i> | 0,014 | 0,000 | 0,003 | 0,020 | 0,013 | 0,000 | 0,002 | 0,019 |
| <i>reduc_abandono</i> | -0,017 | 0,000 | -0,021 | -0,007 | -0,017 | 0,000 | -0,021 | -0,007 |
| <i>reforco</i> | -0,031 | 0,000 | -0,037 | -0,019 | -0,031 | 0,000 | -0,037 | -0,019 |
| <i>dir_indicacao</i> | -0,004 | 0,000 | -0,010 | 0,003 | - | - | - | - |
| <i>prof_vinculo</i> | 0,000 | 0,000 | -0,008 | 0,009 | - | - | - | - |
| | <i>N</i> = 4,386 | | | | <i>N</i> = 4,386 | | | |
| | <i>LogLik</i> = 4375.06, <i>R</i> ² = 0.3007, σ = 0.0892 | | | | <i>LogLik</i> = 4378.12, <i>R</i> ² = 0.3050, σ = 0.0892 | | | |

Elaboração do autor.

Nota: ¹Erros-padrão e intervalos de confiança *bootstrap*.

O apoio privado foi significativo, porém o seu sinal é positivo. É importante deixar claro que essa regressão não está imune a problemas de seletividade amostral, impossibilitando inferências causais. Dito de outra forma, o resultado encontrado pode provir do fato de que este tipo de auxílio acontece de forma mais frequente em locais onde o aproveitamento dos insumos é insatisfatório.

As principais práticas de gestão correlacionadas com a eficiência das redes de ensino são os programas de reforço escolar e aqueles voltados para a redução das taxas de abandono. A existência de plano municipal de educação provavelmente sinaliza planejamento nesta área, pois o parâmetro estimado é significativo e apresentou o sinal esperado.

Para checar a robustez de tais resultados, a outra estratégia empírica seguida consistiu em comparar as distribuições dos índices de eficiência para grupos formados a partir das mesmas informações sobre gestão. Para tal, aplicaram-se os testes F (paramétrico) e Kolmogorov-Smirnov (não paramétrico), sendo ambos unilaterais. Os resultados encontrados são apresentados na tabela 15. Em relação às estimativas econométricas, as indicações são bastante semelhantes. A exceção fica a cargo da existência de consórcio público, que, de acordo com os dois testes empíricos considerados, é significativo pelo menos a 5%. Cabe notar, por fim, que a existência de convênio é significativa a 10%, de acordo com o teste Kolmogorov-Smirnov, mas o teste F não rejeita a nula a níveis convencionais.

TABELA 15
Testes de hipóteses (comparação entre grupos)¹

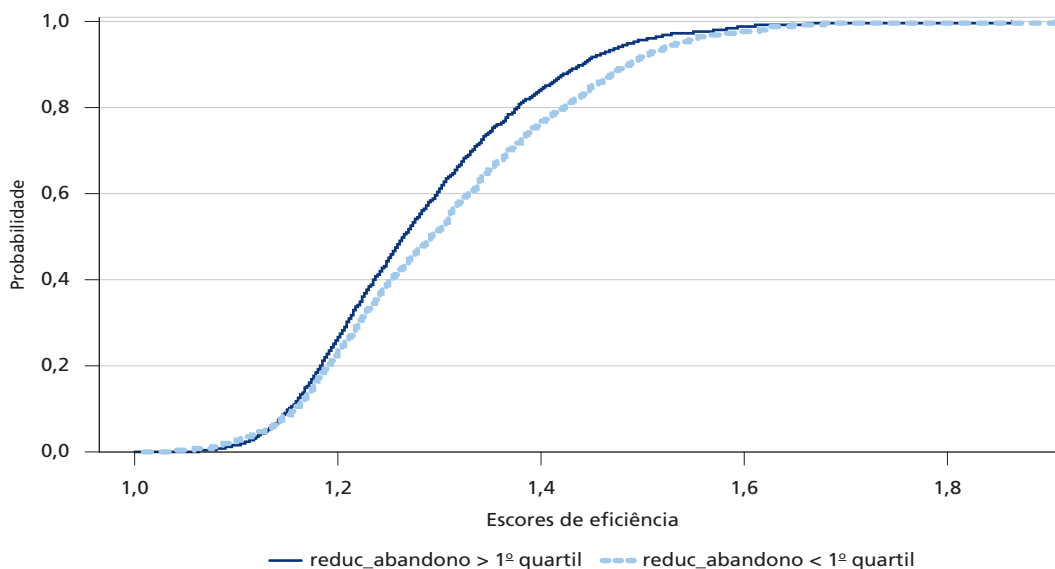
| | Teste F | | | Kolmogorov-Smirnov | | |
|-----------------------|---------|-------------------------|---------|--------------------|----------------------------|---------|
| | F_N | F_N^1 (alpha=0.05) | p-valor | D_{KS} | D_{KS}^1 (alpha=0.05) | p-valor |
| <i>plano_munic</i> | 1,13 | 1,07 | 0,00 | 0,07 | 0,03 | 0,00 |
| <i>cons_pub</i> | 1,08 | 1,07 | 0,00 | 0,06 | 0,06 | 0,04 |
| <i>cons_adm</i> | 1,04 | 1,09 | 0,20 | 0,03 | 0,04 | 0,21 |
| <i>conv_priv</i> | 1,04 | 1,11 | 0,26 | 0,05 | 0,06 | 0,09 |
| <i>apoio_priv</i> | 0,95 | 1,14 | 0,75 | 0,02 | 0,07 | 0,85 |
| <i>reduc_abandono</i> | 1,19 | 1,08 | 0,00 | 0,10 | 0,02 | 0,00 |
| <i>reforco</i> | 1,18 | 1,08 | 0,00 | 0,10 | 0,02 | 0,00 |
| <i>dir_indicacao</i> | 1,00 | 1,09 | 0,54 | 0,02 | 0,02 | 0,51 |
| <i>prof_vinculo</i> | 1,02 | 1,08 | 0,34 | 0,03 | 0,02 | 0,17 |

Elaboração do autor.

Nota: ¹No caso das variáveis binárias da base MUNIC, a implementação dos testes é direta. No caso das variáveis contínuas do FINBRA (razão entre transferências e receitas tributárias) e da Prova Brasil (proporções de escolas que adotam programas de reforço escolas e voltados para redução das taxas de abandono, proporção de diretores que chegaram ao cargo por meio de indicação e proporção de professores que complementam a renda), as variáveis foram dicotomizadas de acordo com a mediana ou o 1º quartil. Ver tabela 3.

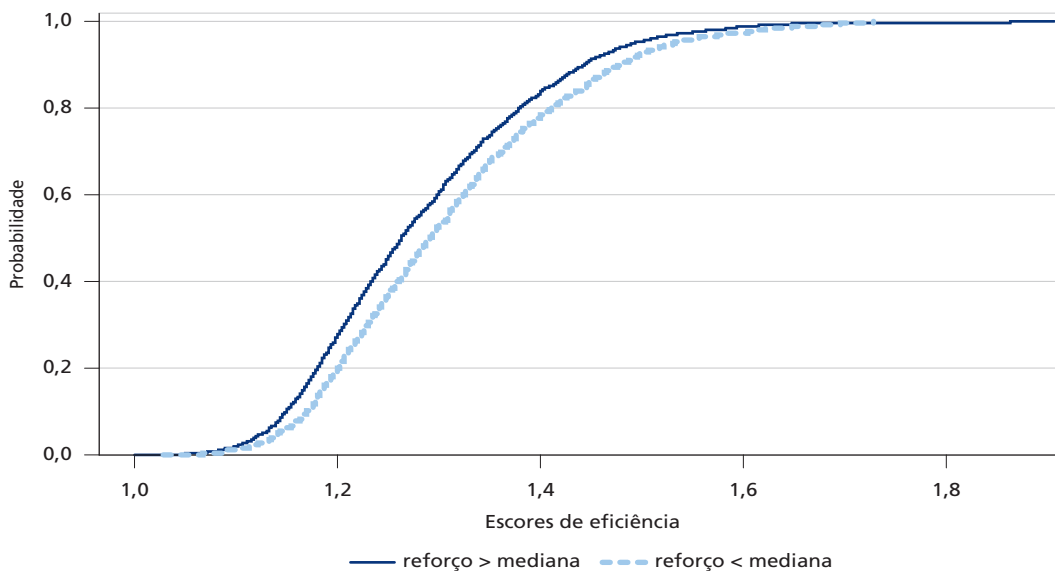
Os gráficos 5 e 6, considerando-se as principais práticas de gestão identificadas, retratam as diferenças existentes nas distribuições acumuladas empíricas dos escores de eficiência. Fica claro que a eficiência das redes de ensino é mais elevada em municípios onde é mais alta a proporção de escolas que administram programas de redução das taxas de abandono, sendo a mesma conclusão válida para programas de reforço escolar. Portanto, do ponto de vista da política pública, um interessante caminho a se seguir seria disseminar experiências exitosas com estes programas.

GRÁFICO 5
Proporção de escolas com programas de redução das taxas de abandono e distribuição acumulada dos escores



Elaboração do autor.

GRÁFICO 6
Proporção de escolas com programas de reforço escolar e distribuição acumulada dos escores



Elaboração do autor.

5 CONCLUSÃO

O PNE prevê substancial elevação do gasto público na área da educação até 2020. Segundo dados do INEP e do Siope, esta despesa era da ordem de 5% do PIB em 2011. A meta é atingir 10% do PIB até 2020, isto é, praticamente dobrar o volume de recursos como razão do produto. O ensino fundamental abarca cerca de 60% dos recursos da área da educação, mas tem experimentado queda nas matrículas nos últimos anos (fenômeno demográfico). As prefeituras e secretarias municipais de educação são as principais responsáveis pela aplicação destes recursos, sendo bastante duvidosa a capacidade de planejamento das prefeituras.

Nesse cenário, a mensuração da eficiência realizada por este estudo procurou evidenciar a existência de potenciais melhoras na aplicação dos recursos públicos nessa área. Para tal, consideraram-se o valor gasto por aluno e os insumos de capital e trabalho disponíveis nas localidades examinadas. Em relação aos produtos da educação, trabalhou-se com as notas em um teste padronizado, o inverso da taxa de distorção e da taxa de abandono. Além de considerar este rol amplo de insumos e produtos, a fim de não superestimar as ineficiências, excluiu-se do banco de dados uma série de municípios que, devido a erro de mensuração nas variáveis ou simplesmente em decorrência de desempenho excepcional, exercem muita influência sobre a fronteira de produção. De outro lado, extraiu-se o viés de subestimação dos escores com o auxílio de técnicas de reamostragem.

As análises desses resultados foram empreendidas com o auxílio de técnicas econométricas que indicaram que fatores ambientais que estão fora do alcance do gestor municipal no curto prazo estão correlacionados com os escores de eficiência. Condições ruins de saúde (mortalidade infantil elevada), além da pouca instrução da população local (taxa de analfabetismo), influenciam de maneira negativa a eficiência. Da mesma forma, o desemprego e a desigualdade de renda estão negativamente relacionados à eficiência. Porém, melhores resultados são alcançados quando a população infantil se dedica apenas ao estudo (criança fora do trabalho) e os pais têm participação na vida escolar de seus filhos. Em relação à gestão na área da educação, os resultados encontrados indicaram que a existência de plano municipal de educação está relacionada a melhores resultados, possivelmente por sinalizar algum planejamento neste campo. Além disso, programas de redução das taxas de abandono e de reforço escolar possibilitam melhores resultados.

REFERÊNCIAS

- ALVES JUNIOR, J. **Avaliação de eficiência na aplicação de recursos educacionais em presença de condicionantes externos e efeitos aleatórios**: uma abordagem em três estágios. 2010. Dissertação (Mestrado) – Universidade de Brasília, Brasília, 2010.
- BENEGAS, M. O uso do modelo NetWork DEA para avaliação da eficiência técnica do gasto público em ensino básico no Brasil. **Economia**, v. 13, p. 569-601, 2012.
- BOGETOFT, P.; OTTO, L. **Benchmarking with DEA, SFA, and R**. New York: Springer, 2011.
- BRASIL. Ministério da Educação. **O Plano de Desenvolvimento da Educação**: razões, princípios e programas. Brasília: MEC, 2007. 43 p.
- DELGADO, V.; MACHADO, A. Eficiência das escolas públicas estaduais de Minas Gerais. **Pesquisa e planejamento econômico**, v. 37, n. 3, p. 427-464, 2007.
- FAÇANHA, L.; MARINHO, A. Instituições de ensino superior governamentais e particulares: avaliação comparativa da eficiência. **Revista de administração pública**, v. 35, n. 6, p. 83-105, 2001.
- _____. Instituições federais de ensino superior: modelos de financiamento o incentivo à eficiência. **Revista brasileira de economia**, v. 53, n. 3, p. 357-386, 1999.
- GASPARINI, C.; RAMOS, F. Efetividade e eficiência no ensino médio brasileiro. **Economia aplicada**, v. 7, n. 2, p. 389-411, 2003.
- MACIEL, W. **Metodologia para o cálculo do indicador investimento público em educação em relação ao PIB de 2000 a 2010**. Brasília: INEP, 2012. (Texto para Discussão, n. 34).
- MARDIA, K.; KENT, J.; BIBBY, J. **Multivariate analysis**. London: Academic Press, 1979.
- MARINHO, A.; RESENDE, M.; FAÇANHA, L. Brazilian federal universities: relative efficiency evaluation and data envelopment analysis. **Revista brasileira de economia**, v. 51, n. 4, p. 489-508, 1997.
- MINGOTI, S. **Análise de dados a partir de métodos de estatística multivariada**: uma abordagem aplicada. Belo Horizonte: Editora da UFMG, 2005.
- ROCHA, F. *et al.* **É possível atingir as metas para a educação sem aumentar os gastos?** Uma análise para os municípios brasileiros. Brasília: STN, 2013. (Texto para Discussão, n. 15).
- ROSANO-PEÑA, C.; ALBUQUERQUE, P.; MARCIO, C. A eficiência dos gastos públicos em educação: evidências georreferenciadas nos municípios goianos. **Economia aplicada**, v. 16, n. 3, p. 421-443, 2012.

SAMPAIO DE SOUSA, M.; CRIBARI-NETO, F.; STOŠIĆ, B. Explaining DEA technical efficiency scores in an outlier corrected environment: the case of public services in Brazilian municipalities. **Brazilian review of econometrics**, v. 25, n. 2, p. 287-313, 2005.

SAMPAIO DE SOUSA, M.; STOŠIĆ, B. Technical Efficiency of the Brazilian municipalities: correcting nonparametric frontier measurements for outliers. **Journal of productivity analysis**, v. 24, n. 2, p. 157-181, 2005.

SANTOS, F.; CRIBARI-NETO, F.; SAMPAIO DE SOUSA, M. Uma avaliação da eficiência do gasto público no Brasil. **Revista brasileira de estatística**, v. 68, n. 228, p. 7-55, 2007.

SIMAR, L.; WILSON, P. Estimation and inference in two-stage, semi-parametric models of production processes. **Journal of econometrics**, v. 136, n. 1, p. 31-64, 2007.

_____. Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models. **Management science**, v. 44, n. 1, p. 49-61, 1998.

STOŠIĆ, B.; SAMPAIO DE SOUSA, M. **Jackstrapping DEA scores for robust efficiency measurement**. In: ENCONTRO BRASILEIRO DE ECONOMETRIA, 25., Porto Seguro, Bahia: SBE, 2003.

TRIGO, P. P. **Avaliação da eficiência técnica no ensino básico brasileiro**. 2010. Dissertação (Mestrado) – Universidade de São Paulo, Ribeirão Preto, 2010.

WILSON, P. FEAR: A package for frontier efficiency analysis with R. **Socio-economic planning sciences**, v. 42, n. 4, p. 247-254, 2008.

ZOGHBI, A.; MATTOS, E.; ROCHA, F.; ARVATE, P. Uma análise da eficiência nos gastos em educação fundamental para os municípios paulistas. **Planejamento e políticas públicas**, n. 36, p. 9-61, 2011.

_____. Mensurando e avaliando a eficiência dos gastos estaduais em educação fundamental e média. **Estudos econômicos**, v. 39, p. 1-20, 2009.

EDITORIAL

Coordenação

Cláudio Passos de Oliveira

Supervisão

Everson da Silva Moura

Reginaldo da Silva Domingos

Revisão

Clícia Silveira Rodrigues

Idalina Barbara de Castro

Laetícia Jensen Eble

Leonardo Moreira de Souza

Marcelo Araujo de Sales Aguiar

Marco Aurélio Dias Pires

Olavo Mesquita de Carvalho

Regina Marta de Aguiar

Barbara Pimentel (estagiária)

Jessyka Mendes de Carvalho Vasquez (estagiária)

Karen Aparecida Rosa (estagiária)

Tauânara Monteiro Ribeiro da Silva (estagiária)

Editoração

Bernar José Vieira

Cristiano Ferreira de Araújo

Daniella Silva Nogueira

Danilo Leite de Macedo Tavares

Diego André Souza Santos

Jeovah Herculano Szervinsk Junior

Leonardo Hideki Higa

Capa

Luis Cláudio Cardoso da Silva

Projeto Gráfico

Renato Rodrigues Buenos

The manuscripts in languages other than Portuguese published herein have not been proofread.

Livraria do Ipea

SBS – Quadra 1 - Bloco J - Ed. BNDES, Térreo.

70076-900 – Brasília – DF

Fone: (61) 3315-5336

Correio eletrônico: livraria@ipea.gov.br

Composto em adobe garamond pro 12/16 (texto)
Frutiger 67 bold condensed (títulos, gráficos e tabelas)
Impresso em offset 90g/m² (miolo)
Cartão supremo 250g/m² (capa)
Brasília-DF

Missão do Ipea

Aprimorar as políticas públicas essenciais ao desenvolvimento brasileiro por meio da produção e disseminação de conhecimentos e da assessoria ao Estado nas suas decisões estratégicas.



ipea Instituto de Pesquisa
Econômica Aplicada

Secretaria de
Assuntos Estratégicos