

EFICIÊNCIA NAS ESCOLAS PÚBLICAS BRASILEIRAS: UMA ANÁLISE DE FRONTEIRA PARCIAL QUANTÍLICA EM DOIS ESTÁGIOS

Maria Inês Rosa da Rocha Caetani
Mestranda do PPGE/PUCRS

Victor Rodrigues de Oliveira
Mestrando do PPGDE/UFPR

RESUMO

O objetivo deste estudo foi mensurar a eficiência das escolas públicas brasileiras e investigar os seus determinantes. A mensuração da eficiência é fundamental para a formulação de políticas públicas, uma vez que é necessário compreender quais os insumos mais importantes para melhorar a qualidade da educação. Para tanto, propôs-se uma nova abordagem semiparamétrica para a mensuração da eficiência da gestão escolar, que contorna o problema da presença de *outliers* na amostra. Por sua vez, os determinantes da eficiência escolar são avaliados por meio do método de misturas finitas. Este estudo é o primeiro a empregar uma metodologia que permite a observação de escolas públicas supereficientes para o caso brasileiro. Para alcançar o objetivo proposto foram utilizados os microdados do Censo Escolar 2011 e do SAEB 2011. Os resultados indicaram, de forma geral, que quanto maior o percentual de alunos que frequentaram a creche e/ou a pré-escola, de escolas com alto grau de defasagem idade-série, de escolas que adotaram o sistema de progressão escolar e de escolas localizadas na área urbana menor será a medida de eficiência. Além disso, notou-se que a realocação das “dotações” escolares é primordial para aumentar o nível de eficiência escolar, corroborando a noção de que os resultados das escolas são decorrentes de um bom balanceamento entre insumos e produtos.

Palavras-chave: eficiência escolar; misturas finitas; abordagem quantílica.

ABSTRACT

The aim of this study was to measure the efficiency of Brazilian public schools and investigate its determinants. The measurement of efficiency is crucial for the formulation of public policy, since it is necessary to understand what are the most important inputs to improve the quality of education. Therefore, we proposed a new semiparametric approach for measuring the efficiency of school management, which circumvents the problem of the presence of outliers in the sample. In turn, the determinants of school efficiency are evaluated by the method of finite mixtures. This study is the first to employ a methodology that allows the observation of super-efficient public schools for the Brazilian case. To achieve the proposed objective were used microdata from the 2011 School Census and SAEB 2011. The results showed, in general, the higher the percentage of students who attended the daycare and/or preschool, schools with high gap age-grade, schools that have adopted the system of school progress and school located in urban lower the efficiency measure. Furthermore, it was noted that the relocation of the "appropriations" school is paramount to increase the level of school efficiency, supporting the notion that school results are due to a good balance between inputs and outputs.

Key-words: school efficiency; finite mixture; quantile approach.

JEL Classification: C14, I21, I28.

Área 12 - Economia Social e Demografia Econômica

1 Introdução

A educação é a possibilidade de mudança e de emancipação social e é através das escolas, que exercem como requisito indispensável formar indivíduos preparados para empregos formais e regulamentados, que se pode atingir a equidade social. Muitas reformas como a adoção de políticas redistributivas e as transformações na organização e na gestão da educação pública ocorreram no sistema de ensino brasileiro nas últimas décadas com o objetivo de melhorar a excelência do ensino e ampliar o acesso da população à escolaridade. Entretanto, as ações não foram o suficiente para que a sociedade civil possa desfrutar de uma educação universal e de qualidade devido à ineficiência na aplicação dos recursos.

Muitos estudos têm sido desenvolvidos avaliando e questionando a compensação dos esforços que têm sido despendidos pelas autoridades públicas para alcançar a universalização e a eficiência do ensino público brasileiro. Os resultados encontrados, em geral, sugerem a inadequada gestão pública dos recursos, colocando inclusive em dúvida a necessidade do mecanismo de transferência, FPM – Fundo de Participação dos Municípios e FDE – Fundo de Participação dos Estados, que de acordo com a Constituição de 1988, Art. 161, tem o objetivo de promover o equilíbrio socioeconômico entre municípios e estados brasileiros.

De acordo com Gasparini & Miranda (2011), em uma avaliação do FPM ao longo do território nacional há no país um potencial tributário que não é explorado. Conforme os autores, a prestação dos serviços de cada município poderia ser inteiramente financiada com receitas próprias. Somente os municípios brasileiros com mais de 50 mil habitantes têm a necessidade de recorrer à complementariedade FPM o que representa 7,5% dos integrantes que recebem as verbas. Além disso, não há uma boa administração dos gastos públicos que é demonstrado pela verba efetivamente gasta e a eficiência do impacto produzido. As estimativas registraram um desperdício de R\$ 16 bilhões nos municípios do interior do Brasil em 2000.

A crença de que o setor público gasta muito e não consegue resolver os problemas da sociedade brasileira, trouxe a discussão à luz da necessidade do alcance da maior eficiência no gasto público e suscitou a adoção de técnicas para avaliar o direcionamento dos recursos para o setor educacional. Um modelo bastante utilizado é análise envoltória de dados (DEA — *data envelopment analysis*) que permite verificar se os recursos orçamentários de cada município estão sendo utilizados de forma que reflitam nas variáveis sociais fundamentalmente na educação, na saúde e na cultura. Os trabalhos que utilizam essa metodologia, em geral, sugerem que a eficiência educacional não está relacionada à quantidade de recursos, mas como esses recursos são aplicados. No entanto, de acordo com Faria, Jannuzzi & Silva (2008), uma localidade pode aplicar bem os seus recursos e mesmo assim não conseguir atingir bons indicadores sociais porque os gastos não são o suficiente para atingir os melhores resultados.

Por outro lado, o fato de uma localidade ter baixo desenvolvimento econômico não significa que não tenha recursos para investir na prestação de seus serviços. O estado da Paraíba, por exemplo, de acordo com Almeida & Gasparini (2011, p. 622), “[que] possui um baixo nível de desenvolvimento econômico, grandes disparidades internas e uma população com escolaridade média baixa” apresentou, até o ano de 2000, melhora nos seus índices educacionais, com maior desempenho dos alunos, mas que frente aos recursos aplicados poderia ter obtido um resultado mais satisfatório que aquele alcançado.

Outras evidências de ineficiência são encontradas no trabalho de Junior, Irff & Benegas (2011) que fez uma avaliação dos gastos públicos municipais no estado do Ceará, utilizando a mesma metodologia DEA. O estudo estimou o nível de eficiência dos serviços prestados à população do estado, concluindo que os indicadores sociais de 55% dos municípios sofrem impactos positivos frente aos recursos aplicados. No entanto, o estudo considerou que o estado é ineficiente nos gastos públicos.

Já Zoghbi *et al.* (2009) demonstraram através da comparação de indicadores, construídos para o setor público de ensino concernente à educação nos níveis fundamental e médio, entre os estados brasileiros que Ceará é um estado que apresenta o menor índice de gastos, mas que está dentro da fronteira de produção da educação fundamental. De acordo com esses autores nem sempre o melhor desempenho denota maior eficiência relativa dos estados brasileiros. O desempenho depende também da melhora de outros indicadores socioeconômicos como saúde, infraestrutura, distribuição de renda.

Não é de admirar, portanto que o Brasil ainda seja considerado um país de baixo nível educacional. A despeito do maior número de indivíduos frequentando a escola, a qualidade do ensino deixa a desejar. De acordo com Gramani & Duarte (2011) e apontado por Veloso (2009, p. 10), de 1995 a 2007 os indicadores quantitativos apresentaram evolução considerável com um aumento do percentual de crianças entre 7 e 14 anos que frequentavam a escola de 93% para 98%, na faixa de 15 a 17 anos subiram de 64% para 80% e a taxa de conclusão do ensino médio elevou-se de 17% para 44%. Por outro lado, os indicadores qualitativos não apresentaram efeito tão satisfatório. De acordo com os autores, o resultado da prova do Sistema de Avaliação da Educação Básica (SAEB) exibiu desempenho dos alunos abaixo do esperado nos anos de 1995 e 2001. Nesse estudo também é assinalado que o Brasil investiu em educação o mesmo percentual que a Coreia do Sul no mesmo período, mais que a Espanha e mais que o Japão, evidenciando a ineficiência dos recursos aplicados.

De acordo com o exposto, acredita-se que para alcançar maior eficiência dos gastos em educação é necessário estimular a complementação orçamentária com recursos da própria comunidade e desenvolver um modelo de gestão que seja baseado na combinação de um planejamento, de controle central da formulação e da implementação das políticas e da definição de custos mínimos para todos os municípios. Por meio da formulação de políticas educacionais, envolvendo estratégias eficientes de gestão e financiamento talvez seja possível alcançar a expansão qualitativa da educação e uma maior equidade social. Neste sentido desenvolveu-se este estudo que objetiva mensurar a eficiência das escolas públicas brasileiras e compreender quais os seus determinantes, de forma a contribuir para a formulação de políticas. Para alcançar este objetivo será utilizado o método de fronteira parcial quantílica desenvolvido por Aragon, Daouia & Thomas-Agnan (2005), que permite a presença de escolas “supereficientes”. O método de medida está sendo devidamente demonstrado e detalhado na seção 2.

Além desta introdução este artigo está dividido em quatro partes: na próxima seção demonstram-se os procedimentos metodológicos, na seção três apresenta-se a base de dados e o tratamento da amostra, na seção quatro, o resultado do modelo e na seção 5 serão apresentadas as conclusões finais.

2 Eficiência Produtiva nas Escolas

Apesar do consenso com relação à ideia de que os ganhos sociais advindos da melhoria da educação superam os privados, ainda se conhece pouco sobre as formas de efetivamente melhorar a qualidade do ensino. Neste sentido, a abordagem teórica importante é a função de produção ou fronteira de eficiência na educação formulada pela primeira vez por Coleman *et al.* (1966). O objetivo do relatório Coleman era encontrar evidências de que o baixo desempenho dos alunos negros era resultado da insuficiência de recursos em suas escolas. A partir disso, houve uma expansão do número de estudos que procuram compreender como as diversas características dos alunos, das escolas e das famílias se combinam para determinar o desempenho dos alunos.

No âmbito dos insumos educacionais merecem destaque os trabalhos de Hanushek (1986) e Hanushek & Luque (2003) que utilizam exames de proficiência em uma amostra de escolas norte-americanas e em um painel de países, respectivamente, comparando os resultados dos testes padronizados com insumos dos professores, dos alunos ou das turmas. Os resultados indicaram que as variáveis citadas não apresentaram efeito sobre o desempenho educacional, nem nos países desenvolvidos nem nos em desenvolvimento.

A mensuração da função de produção da educação é fundamental para a formulação de políticas públicas. Como ressaltado por Sacerdote (2001), “*it is clearly difficult to think about improving student outcomes in primary and secondary schools until we know which inputs matter*”. Entretanto, para mensurar a eficiência são necessárias algumas alterações na teoria. Para tanto, considere uma função de produção da educação como especificado abaixo:

$$y_i = f(x_i, z_i, d_i) \quad (1)$$

onde y_i é o rendimento dos alunos na escola i , x_i são os insumos educacionais disponibilizados pela escola, z_i é o *background* familiar e d_i representam as “dotações”. Os estudos citados consideram a função de produção como um ajuste ideal para os valores médios da distribuição da variável dependente e, portanto, a estimação correta desta função deve passar pela média condicional $E(Y|X)$. Dessa forma, o produto efetivamente observado, y_i , pode estar acima ou abaixo da função estimada. Nesta abordagem, considera-se que, na média, as observações são eficientes, isto é, situam-se sobre a fronteira de produção. É possível incorporar a eficiência na fronteira de produção através de um índice θ_i que multiplica a equação (1) como segue:

$$y_i = \theta_i f(x_i, z_i, d_i) \quad (2)$$

A partir da equação (2) é necessário encontrar uma forma de obter-se a função de produção sem considerar que os pontos encontrados são eficientes na média. De forma paramétrica, o usual é considerar a distribuição dos erros como uma normal truncada de erros não positivos, método conhecido como fronteiras estocásticas. Assim, o erro é decomposto em duas partes: a dos erros normais (v_i) e a dos erros não negativos (u_i) (DELGADO & MACHADO, 2007). A fronteira de produção pode ser estimada a partir da equação que segue:

$$\ln(y_i) = x_i \beta + v_i - u_i \quad (3)$$

onde $\ln(y_i)$ representa o logaritmo do produto, x_i representa os insumos, v_i representa os erros aleatórios de acordo com uma distribuição normal e u_i representa os erros não positivos a partir de uma normal truncada ou de uma exponencial. A medida de eficiência, θ_i , é obtida por meio de $\exp(-u_i)$ ¹.

A equação (2) também pode ser estimada de forma não paramétrica, que não utiliza erros aleatórios. A estimação é feita por meio das técnicas denominadas *Free Disposal Hull* (FDH) proposta por Deprins, Simar & Tulkens (1984) e *Data Envelopment Analysis* (DEA) proposta por Charnes, Cooper & Rhodes (1978). Enquanto a DEA assume uma tecnologia convexa e emprega programação linear, o FDH é baseado no princípio de dominância fraca e afasta-se do pressuposto de convexidade inerente à DEA. Ou seja, o FDH “envelopa” os dados por meio da hipótese de *non-convex staircase-hull*. Estas abordagens permitem a hierarquização das unidades de análise por meio da alocação dos insumos e, conseqüentemente, do produto obtido. Uma ampla literatura sobre eficiência se desenvolveu, destacando-se os estudos de Wilson (2005) e de Afonso & Aubyn (2006). Estes trabalhos incorporam os desenvolvimentos recentes do método DEA, de forma a obterem-se índices mais robustos.

Os prós e os contras de abordagens paramétricas e não paramétricas têm sido intensamente debatidos. A primeira tem sido criticada por confiar em suposições restritivas relativas à forma funcional e à distribuição dos erros aleatórios e por permitir somente um único produto. As análises não paramétricas, por seu turno, têm sido criticadas por serem determinísticas e, principalmente, por serem extremamente vulneráveis a *outliers* e a erros de mensuração. Os estudos sobre eficiência que adotam um enfoque não paramétrico têm migrado para o chamado *partial frontier approaches*, a saber, a análise de eficiência

¹ Os principais estudos sobre fronteiras estocásticas são os de Aigner & Chu (1968), de Aigner, Lovell & Schmidt (1977) e de Greene (1993).

order-m (CAZALS; FLORENS & SIMAR, 2002) e *order- α* (ARAGON; DAOUIA & THOMAS-AGNAN, 2005). Estas análises “parciais” generalizam o FDH ao permitirem que observações “supereficientes” sejam alocadas além da fronteira de possibilidade de produção. Assim, as medidas de eficiência são menos vulneráveis a *outliers*.

Este estudo adotará o método proposto por Aragon, Daouia & Thomas-Agnan (2005). Para tanto, é necessário compreender o conceito de análise de fronteira parcial, como proposto por Tauchmann (2012). Considere uma amostra de N unidades de tomada de decisão (*decision making units* – DMU). Para cada DMU, $i = 1, \dots, N$, um conjunto de insumos produtivos x_{i1}, \dots, x_{iK} e um conjunto de produtos y_{i1}, \dots, y_{iL} são observados. O objetivo de uma medida de eficiência é calcular um escore θ_i para cada DMU. Normalmente, duas variantes são calculadas: i) a eficiência orientada para os insumos θ_i^{inp} , isto é, o fator pelo qual o consumo de insumos da DMU i pode ser reduzido proporcionalmente sem alterar o produto resultante; e ii) a eficiência orientada para o produto θ_i^{out} , isto é, o fator pelo qual a geração do produto pode ser aumentada sem alteração do consumo dos insumos. Enquanto a primeira mensura a distância radial relativa na direção dos insumos, a segunda mensura a distância radial relativa na direção dos produtos. Para os modelos *full frontier* todas as DMUs são “envelopadas” pela fronteira de possibilidade de produção, isto é, $\theta_i^{inp} \in (0, 1]$ e $\theta_i^{out} \in [1, \infty)$. Dessa forma, as DMUs são caracterizadas por escores de eficiência que assumem o valor um, enquanto desvios para baixo (orientação para os insumos) e para cima (orientação para os produtos) da unidade indicam ineficiência. Por outro lado, as abordagens de fronteira parcial permitem que se supere (orientação para os insumos) o valor de um. Neste estudo os modelos estimados são todos orientados para os insumos, uma vez que se pretende obter resultados voltados para diminuir o consumo dos insumos, mas manter o nível de eficiência dos produtos obtidos.

Como o enfoque da fronteira parcial generaliza a abordagem FDH, discutir-se-á brevemente este último método. No método FDH, a eficiência orientada para os insumos é estimada ao comparar-se cada DMU, $i = 1, \dots, N$, com todas as outras DMUs, $j = 1, \dots, N$, que produzem, pelo menos, tanto quanto ela. O conjunto de pares de DMUs na amostra que satisfaz a condição $y_{lj} \geq y_{li}$ para qualquer l é denotado por B_i . Entre os pares de DMUs aquele que exibe o consumo mínimo de insumos serve como referência para i e $\hat{\theta}_i^{FDH}$ é calculado como o uso relativo dos insumos:

$$\hat{\theta}_i^{FDH} = \min_{j \in B_i} \left\{ \max_{k=1, \dots, K} \left\{ \frac{x_{kj}}{x_{ki}} \right\} \right\} \quad (4)$$

As DMUs que apresentam um consumo mínimo de insumos entre todos os seus pares serve como sua própria referência. Neste caso $\hat{\theta}_i^{FDH}$ assume o valor um. Todavia, até mesmo uma única DMU nos dados que apresente pequeno consumo de insumos pode tornar as demais ineficientes. Assim, o FDH é altamente sensível a *outliers* e a erros de medição.

2.1 O Método de Fronteira Parcial Quantílica

O método *order- α* , $\varphi(x)$, proposto por Aragon, Daouia & Thomas-Agnan (2005) generaliza o FDH. Para tanto, considere um espaço de probabilidade (Ω, A, \mathcal{P}) sobre o qual o vetor de insumos X e o vetor de produtos Y estão definidos. Neste enfoque definem-se o conjunto ψ no qual estão contidos o vetor de insumos, o vetor de produtos e a distribuição conjunta dos dois vetores e um subconjunto de ψ : $\psi^* = \{(x, y) \in \psi | F_X(x) > 0\}$. A partir desta definição o valor da função de produção coincide com o último ponto de um quantil qualquer, isto é, dado que $X \leq x$ tem-se que:

$$q_1(x) = \inf\{y \geq 0 | F(y|x) = 1\} \quad (5)$$

A equação (5) sugere introduzir o conceito de função de produção de ordem contínua $\alpha \in [0,1]$, como uma função quantílica de ordem α da função que determina Y dado que X não excede um dado nível de insumos. Esta função assume, para um dado nível de insumos x , o seguinte valor:

$$q_\alpha(x) = F^{-1}(\alpha|x) = \inf\{y \geq 0 | F(y|x) \geq \alpha\} \quad (6)$$

Esta função quantílica condicional é o limiar (*threshold*) do produto educacional que é excedido por $100(1 - \alpha)\%$ das escolas que utilizam menos insumos do que o nível x . A função $F^{-1}(\cdot |x)$ é a inversa generalizada de $F(\cdot |x)$. Se a função de distribuição $F(\cdot |x)$ é estritamente crescente a sua inversa coincide com a inversa generalizada. Assumindo que para todo x tal que $F_x(x) > 0$, a função de distribuição condicional $F(\cdot |x)$ é estritamente crescente no intervalo $[0, \varphi(x)]$. Então,

$$\forall (x, y) \in \psi^* \text{ tem-se que } y = q_\alpha(x) \text{ com } \alpha = F(y|x) \quad (7)$$

Da propriedade (7) observa-se que qualquer plano de produção (x, y) no subconjunto ψ^* pertence a alguma curva quantílica de ordem α . Assim, a função quantílica q_α mensura a eficiência do plano de produção (x, y) ao compará-lo com todos que utilizam o mesmo nível de insumos x , bem como com os que usam níveis menores que x . Para estimar a função quantílica condicional q_α é útil utilizar a função quantílica empírica condicional obtida ao inverter-se a função de distribuição empírica condicional $\hat{F}(\cdot |x)$:

$$\hat{q}_{\alpha,n} = \hat{F}^{-1}(\alpha|x) = \inf\{y | \hat{F}(y|x) \geq \alpha\} \quad (8)$$

Para obter-se o estimador em (8) assumamos que N_x seja o número de observações X_i menor ou igual a x , isto é, $N_x = \sum_{i=1}^n 1(X_i \leq x)$ e para $j = 1 \dots, N_x$ denotemos por $Y_{(j)}$ uma estatística de ordem j das observações Y_i tal que $X_i \leq x: Y_{(i_1)} \leq Y_{(i_2)} \leq \dots \leq Y_{(i_{N_x})}$. Assim, para x tal que $N_x \neq 0$,

$$\hat{F}(y|x) = \frac{\sum_{i|X_i \leq x} 1(Y_i \leq y)}{N_x} = \frac{\sum_{j=1}^{N_x} 1(Y_{(j)} \leq y)}{N_x} \quad (9)$$

Portanto,

$$\hat{F}(y|x) = \begin{cases} 0 & \text{se } y < Y_{(i_1)} \\ \frac{k}{N_x} & \text{se } Y_{(i_k)} \leq y \leq Y_{(i_{k+1})}, 1 \leq k \leq N_x - 1 \\ 1 & \text{se } y \geq Y_{(i_{N_x})} \end{cases} \quad (10)$$

Para $\alpha > 0$

$$\hat{\varphi}_n(x) = Y_{(i_{N_x})} = \max_{i|X_i \leq x} Y_i, \quad \hat{q}_{\alpha,n}(x) = \begin{cases} Y_{(i_{\lfloor \alpha N_x \rfloor})} & \text{se } \alpha N_x \in \mathbb{N} \\ Y_{(i_{\lfloor \alpha N_x \rfloor})} & \text{caso contrário} \end{cases} \quad (11)$$

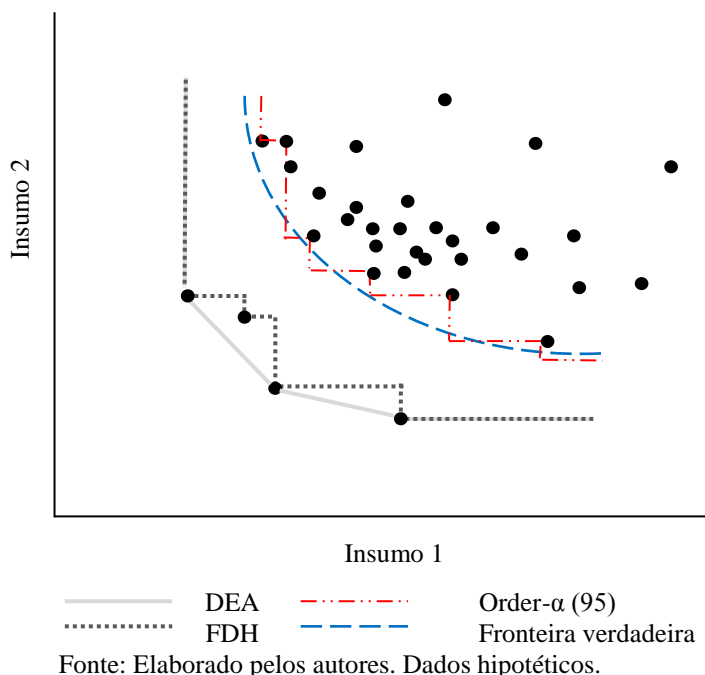
onde $\lfloor \alpha N_x \rfloor$ denota a parte inteira de αN_x . Reescrevendo (11) nos termos de (4) obtém-se o escore de eficiência:

$$\hat{\theta}_{\alpha i}^{OA} = P(100 - \alpha)_{j \in B_i} \left\{ \max_{k=1, \dots, K} \left\{ \frac{x_{kj}}{x_{ki}} \right\} \right\} \quad (12)$$

Para garantir a precisão das estimativas, os intervalos de confiança da medida de eficiência, equação (12), serão obtidos por *bootstrap* com 1000 replicações.

A Figura 1 fornece uma ilustração gráfica dos métodos não paramétricos discutidos acima. Para o DEA e o FDH, as DMUs irregulares abrangem as fronteiras estimadas, tornando todo o resto das DMUs altamente ineficientes. De fato, as observações regulares não afetam as fronteiras estimadas pelo DEA e pelo FDH. Por outro lado, o método *order- α* permite que as DMUs se localizem fora da fronteira de possibilidade de produção estimada.

Figura 1: **Fronteira de possibilidade de produção (isoquantas) estimadas não parametricamente**



Para estimar o índice de eficiência como proposto em (12) é necessário escolher o α ótimo e, assim, reduzir o viés da medida de eficiência. Há três métodos (dois locais e um global) para detectar descontinuidades na amostra e determinar o valor ótimo de α :

1. α para o qual a série diferenciada duas vezes assume um valor mínimo.
2. Valores de α para os quais os valores negativos persistem após suavização repetida da série que foi duas vezes diferenciada.
3. α que minimiza o critério BIC ao dividir a série em duas funções e ajustá-las por meio de um termo de segunda ordem.

Estes critérios de escolha podem ser aproximados pela função² que segue:

$$\gamma^*(T^{\alpha,x}, F) = \sup_{(x_0, y_0) \in \mathbb{R}_+^2} |IF((x_0, y_0); T^{\alpha,x}, F)| = \frac{\max(\alpha, 1 - \alpha)}{f(q_\alpha(x)|x)F_X(x)} \quad (13)$$

onde o funcional T é avaliado por meio de uma função de distribuição empírica distribuída igualmente com respeito a distância de Prohorov e IF é uma função de influência definida a partir da primeira derivada de Gâteaux. Este funcional permite (i) descrever o efeito de uma influência infinitesimal de mudanças nas coordenadas dos vetores ponderada pela função densidade de probabilidade e (ii) avaliar a influência relativa de observações individuais sobre o valor da estimativa.

² Para mais detalhes sobre o método de detecção de *outliers* ver Daouia & Ruiz-Gazen (2006).

2.2 Modelo de Misturas Finitas

O primeiro trabalho a observar a existência das misturas finitas foi o de Feller (1943). Estes modelos representam a adição de densidades de S distintas populações para analisar a heterogeneidade não observada. Observa-se que algumas variáveis aleatórias, Y , podem ser heterogêneas entre grupos, entretanto são homogêneas intra-grupos. Em decorrência da heterogeneidade as variáveis têm uma distribuição de probabilidade que difere entre os grupos, embora se pressuponha que todas derivem de uma mesma distribuição paramétrica, denotada por $p(y|\theta)$, sendo que o parâmetro θ difere entre os grupos. Assim, para uma população composta por K subpopulações uma variável S denomina os grupos que assumem os valores no conjunto $\{1, \dots, K\}$.

De forma semelhante a Y , as componentes S também podem ser extraídas de uma amostra aleatória cuja probabilidade é ξ_S . Se S é conhecido Y seguirá uma distribuição $p(y|\theta_S)$ e a distribuição conjunta $p(y, S)$ é dada por:

$$p(y, S) = p(y|S)p(S) = p(y|\theta_S)\xi_S \quad (14)$$

Conforme Cameron e Trivedi (2005), um modelo de misturas finitas com S componentes tem a seguinte representação:

$$f(y_i|\theta_p, \xi_p) = \sum_{p=1}^S \xi_p f_p(y_i|\theta_p), \quad 0 < \xi_p < 1, \quad \sum_{p=1}^S \xi_p = 1 \quad (15)$$

onde $f_p(y_i|\theta_p)$, $p = 1, 2, \dots, S$ é a s -ésima densidade e ξ_p a probabilidade da s -ésima densidade. O modelo de misturas finitas é semiparamétrico uma vez que não é necessário fazer hipóteses sobre as distribuições de cada subpopulação. Contudo, diferentemente de Y , as componentes em Y é capturada pelos diferentes parâmetros associados a cada componente. Para determinar o número de componentes foram utilizados os critérios de informação de Akaike (AIC) e o Bayesiano (BIC) e o teste de razão de verossimilhança modificado (*likelihood ratio test* – LRT).

Chen & Kalbfleisch (2005) sugerem a aplicação do LRT para determinar o número de componentes em um modelo de misturas finitas. Para um modelo com duas componentes suponha que $\phi(x; \mu, \sigma)$ é a distribuição normal com média μ e desvio-padrão σ . Neste caso tem-se:

$$f_Y(y; \xi, \mu_1, \mu_2, \sigma) = \xi\phi(y; \mu_1, \sigma) + (1 - \xi)\phi(y; \mu_2, \sigma) \quad (16)$$

O problema a ser testado é H_0 : f_Y é normalmente distribuída contra H_1 : f_Y é distribuída como em (16). Neste caso, a função de verossimilhança é dada por:

$$l_n(\xi, \mu_1, \mu_2, \sigma) = \sum_{i=1}^n \log(\xi\phi(y_i; \mu_1, \sigma) + (1 - \xi)\phi(y_i; \mu_2, \sigma)) + 2 \log(4\xi(1 - \xi)) \quad (17)$$

Para um modelo com s componentes a estatística LRT é definida como segue:

$$l_n(\mu_1, \dots, \mu_m, \sigma) = \sum_{i=1}^n \log(\xi_1\phi(y_1; \mu_1, \sigma) + \dots + \xi_r\phi(y_r; \mu_r, \sigma)) + C \log(\prod_{i=1}^r \xi_i) \quad (18)$$

Os autores mostram que o teste LRT é assintoticamente distribuído como:

$$q\chi_0^2 + \frac{1}{2}\chi_1^2 + (1 - q)\chi_2^2 \quad (19)$$

onde a proporção q depende da distribuição utilizada.

3 Fonte de Dados e Tratamento da Amostra

Neste estudo serão empregados os dados do Sistema Nacional de Avaliação da Educação Básica (SAEB) e do Censo Escolar, ambas conduzidas pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) vinculado ao Ministério da Educação do Brasil. Essas duas fontes de dados tem por objetivo compor as variáveis de *produtos*, *insumos*, *dotação* e *variáveis socioeconômicas* dos alunos matriculados em determinada escola. A unidade básica de análise deste estudo é a escola. Desse modo, as informações sobre alunos se referem à média das características dos mesmos no estabelecimento de ensino.

Na primeira etapa do SAEB são selecionadas escolas com seu conjunto de turmas e alunos da 4ª série e da 8ª série do Ensino Fundamental e da 3ª série do Ensino Médio. Na segunda etapa foram selecionadas turmas em cada uma das séries, dentro das escolas selecionadas. Uma vez selecionada uma turma para participar da avaliação, todos os alunos da turma faziam parte automaticamente da amostra e cada aluno presente no dia da avaliação foi submetido às provas das disciplinas de língua portuguesa e matemática.

O Censo Escolar é um levantamento de dados estatístico-educacionais de âmbito nacional realizado todos os anos e coordenado pelo INEP e com a participação de todas as escolas públicas e privadas do país. Abrange o ensino regular (educação Infantil e ensinos fundamental e médio), a educação especial e a educação de jovens e adultos (EJA). O Censo Escolar coleta dados sobre estabelecimentos, matrículas, funções docentes, movimento e rendimento escolar. Essas informações são utilizadas para traçar um panorama nacional da educação básica e servem de referência para a formulação de políticas públicas e execução de programas na área da educação, incluindo os de transferência de recursos públicos como merenda e transporte escolar, distribuição de livros e uniformes, implantação de bibliotecas, entre outros.

Utilizar-se-ão os dados da 4ª série e da 8ª série do ensino público e os resultados das provas de Matemática e de Língua Portuguesa para evitar oscilações bruscas decorrentes de turmas pequenas e comportamentos específicos de alguma série. A análise considerou somente as escolas públicas, pois o tratamento da amostra reduziu consideravelmente as escolas de ensino privado. Além disso, não se consideraram as observações que não continham informações em todas as variáveis utilizadas.

As estatísticas descritivas³ das variáveis utilizadas para o cálculo da medida de eficiência por série estão apresentadas na Tabela 1. Para o primeiro estágio são três variáveis de produto: proficiência em Matemática, proficiência em Língua Portuguesa e matrículas por escola. No caso dos insumos são cinco variáveis: número total de salas por escola, indicador de infraestrutura, uma medida para capturar a presença concomitante de um conjunto de recursos de multimídia, o número total de professores com graduação nas disciplinas avaliadas por escola e o número total de professores que fizeram cursos de aperfeiçoamento por escola.

Com relação às notas observa-se que elas são bastante simétricas e revelam um comportamento normal. A variável *matrículas* corresponde ao total de alunos matriculados por escola e é utilizada como indicador de produto. Esta variável merece destaque devido à dispersão significativa, distribuição assimétrica à direita, com o mínimo de doze alunos em uma pequena escola para a 4ª série até um máximo de 8.153

³ Os dados das provas do SAEB são desagregados por alunos. Como o objetivo é apresentar informações por escolas, os resultados são reportados pela média. Portanto, a distribuição na tabela apresentada é a distribuição de médias por escola e não dos alunos.

alunos na maior escola em termos de matrícula no ensino fundamental. As médias de matrículas se elevam ao longo das séries e o desvio-padrão também aumenta consideravelmente.

Tabela 1: Estatísticas Descritivas do 1º Estágio

Variáveis	4ª série				8ª série			
	μ	σ	Min	Max	μ	σ	Min	Max
Produtos								
Português	190,9	24,9	95,0	318,8	229,3	23,5	140,6	339,9
Matemática	208,8	27,6	130,6	357,1	237,9	24,6	154,1	374,3
Matrículas	503,2	353,2	12	3.903	760,0	492,4	24	8.153
Insumos								
Salas	9,95	6,33	1	103	12,21	7,76	1	115
Infraestrutura	6,08	1,29	0	9	6,54	1,36	0	9
Multimídia	5,33	1,39	0	7	5,73	1,18	0	7
Graduação	2,99	2,44	0	27	4,96	3,24	0	35
Curso	2,61	2,14	0	23	3,95	2,87	0	29

Fonte: SAEB 2011 e Censo Escolar 2011. Elaborado pelos autores.

No caso dos insumos a variável *infraestrutura* provém da soma de nove variáveis *dummies* obtidas junto aos microdados do Censo Escolar, a saber: se o sistema de água da escola provém de rede pública (*água*); se a escola é ligada à rede pública de energia elétrica (*energia*); se a escola está ligada à rede pública de esgoto (*esgoto*); se há coleta periódica de lixo pela prefeitura (*lixo*); se o sanitário está dentro da escola (*sanitário*); se a escola possui biblioteca (*biblioteca*); se a escola possui sala de leitura (*sala de leitura*); se a escola possui cozinha (*cozinha*); e se a escola possui laboratório de ciências (*laboratório*). A variável *multimídia* é construída pelo método de *principal components analysis* (PCA) e combina sete variáveis *dummies* que indicam a presença de: equipamento de televisão (*TV*); equipamento de videocassete (*videocassete*); equipamento de DVD (*DVD*); parabólica (*parabólica*); copiadora (*copiadora*); impressora (*impressora*); e computadores (*computadores*).

A variável *graduação* mensura o número de professores que possuem licenciatura em matemática e português por escola e por série. A variável *salas*, por sua vez, mensura o número total de salas existentes na escola. Por fim, a variável *curso* indica o número de professores que fizeram cursos de aperfeiçoamento e de atualização por escola.

A Tabela 2 reúne informações com relação às escolas públicas. As variáveis *financeiro*, *rotatividade*, as que compõem o indicador de *infraestrutura*, as que compõem a variável *multimídia*, *alimentação*, *progressão* e *urbano* representam o conjunto de dotações.

A variável *cor* indica a porcentagem de alunos que se autodeclararam brancos e amarelos. Observa-se que para as escolas de ensino fundamental a minoria é composta por alunos brancos, 34% na 4ª série e 38% na 8ª série. Nota-se também que a maioria das mães possui no mínimo a 8ª série do ensino fundamental completa.

A variável *creche e pré-escola* foram inseridas para avaliar a hipótese proposta por Curi e Menezes-Filho (2009). As evidências do estudo destes autores indicou que os alunos que frequentaram a pré-escola têm uma maior probabilidade de concluir os quatro ciclos escolares e a creche tem relação positiva e significativa apenas com a conclusão do ensino médio e do ensino universitário. Além disso, a pré-escola está associada com um aumento de um ano e meio de escolaridade e de 16% na renda, independentemente da sua relação com a educação. Finalmente, os alunos que frequentaram a pré-escola têm um desempenho escolar melhor, medido por testes de proficiência, na 4ª e 8ª séries do ensino fundamental.

Tabela 2: Estatísticas Descritivas do 2º Estágio

Variáveis	4ª série				8ª série			
	μ	σ	Min	Max	μ	σ	Min	Max
Branco	0,35	0,18	0	1	0,38	0,20	0	1
Idade	14,90	0,55	13,33	19,00	14,88	0,50	12,32	18,63
Computador	0,55	0,30	0	1,66	0,65	0,32	0	2
Jornal	0,62	0,18	0	1	0,67	0,15	0	1
Livros	0,90	0,10	0	1	0,87	0,09	0	1
Trabalho fora	0,19	0,13	0	1	0,17	0,09	0	1
Trabalha em casa	0,85	0,13	0	1	0,81	0,12	0	1
Creche	0,38	0,20	0	1	0,32	0,18	0	1
Pré-escola	0,40	0,20	0	1	0,47	0,17	0	1
Não reprovado	0,70	0,21	0	1	0,67	0,17	0	1
Faz lição	0,98	0,05	0	1	0,94	0,06	0	1
Capital socioeconômico	2,10	0,45	0	4,15	2,16	0,44	0	4,20
Educação mãe	0,53	0,21	0	1	0,52	0,21	0	1
Financeiro	0,52	0,50	0	1	0,48	0,50	0	1
Rotatividade	0,74	0,44	0	1	0,57	0,50	0	1
Água	0,86	0,35	0	1	0,87	0,33	0	1
Energia	0,99	0,08	0	1	1,00	0,07	0	1
Esgoto	0,54	0,50	0	1	0,55	0,50	0	1
Lixo	0,92	0,27	0	1	0,92	0,27	0	1
Sanitário	0,97	0,17	0	1	0,98	0,14	0	1
Laboratório	0,07	0,25	0	1	0,30	0,46	0	1
Cozinha	0,98	0,13	0	1	0,98	0,13	0	1
Biblioteca	0,47	0,50	0	1	0,63	0,48	0	1
Sala de leitura	0,28	0,45	0	1	0,31	0,46	0	1
TV	0,97	0,17	0	1	0,99	0,12	0	1
Videocassete	0,55	0,50	0	1	0,65	0,48	0	1
DVD	0,95	0,21	0	1	0,97	0,17	0	1
Parabólica	0,41	0,49	0	1	0,57	0,50	0	1
Copiadora	0,59	0,49	0	1	0,62	0,49	0	1
Impressora	0,92	0,28	0	1	0,96	0,19	0	1
Computadores	0,95	0,23	0	1	0,98	0,14	0	1
Alimentação	0,99	0,11	0	1	0,98	0,13	0	1
Progressão	0,41	0,49	0	1	0,35	0,48	0	1

Fonte: SAEB 2011 e Censo Escolar 2011. Elaborado pelos autores.

A variável *capital socioeconômico* foi obtida por PCA com a extração de um componente principal a partir da combinação das seguintes variáveis: o número de aparelhos de televisões na casa do aluno; o número de aparelhos de rádio na casa do aluno; o número de aparelhos de DVD na casa o aluno; o número de geladeiras na casa do aluno; e o número de carros da família do aluno.

4 Resultados do Modelo

O modelo empregado neste estudo pode ser dito duplamente robusto, uma vez que adota um método não paramétrico para obter o escore de eficiência para cada escola e, em seguida, o modelo semiparamétrico de misturas finitas para investigar os determinantes da eficiência escolar. A primeira etapa para estabelecer a eficiência das escolas é determinar o valor ótimo de α que minimizará o viés das estimativas e também permitirá a obtenção correta da proporção de *outliers* presente na amostra. O critério adotado escolhido foi aquele que minimiza a proporção de *outliers*, isto é, o valor de α deste estudo indicou a

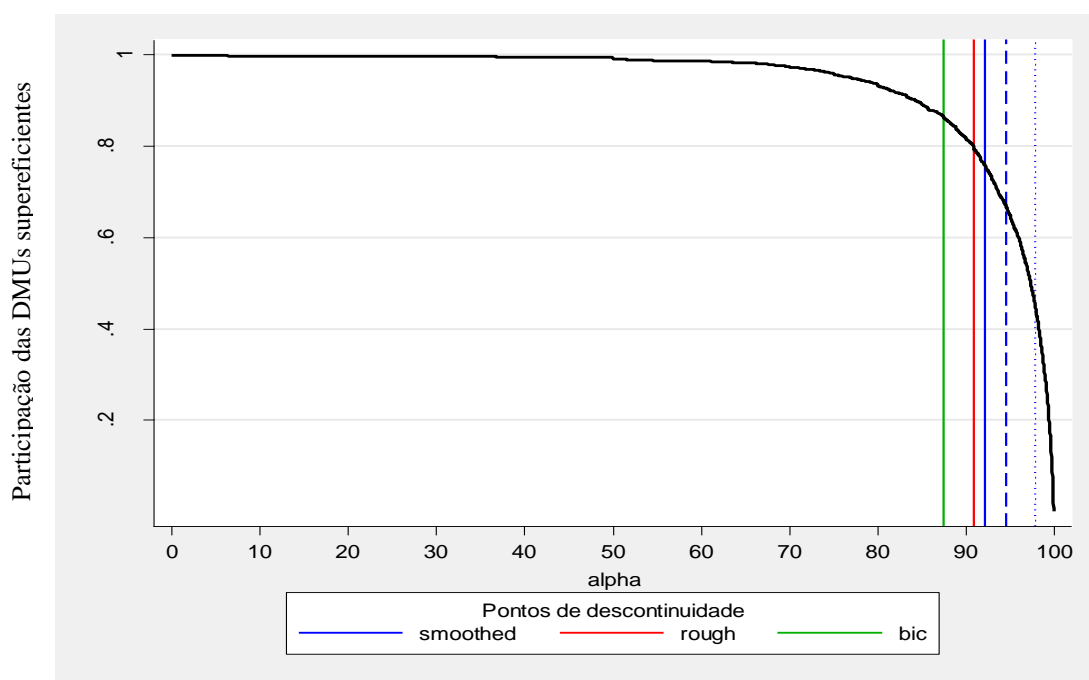
presença de 45,31% de DMUs classificadas como supereficientes. Estes resultados podem ser observados na Tabela 3 e na Figura 2.

Tabela 3: Teste para determinar o α ótimo

Teste	Valor de α	Proporção de <i>outliers</i> (%)
<i>Prime suggestion by criterion smooth</i>	0,9212	75,83
<i>Second-rate suggestion by criterion smooth</i>	0,9449	67,02
<i>Third-rate suggestion by criterion smooth</i>	0,9782	45,31
<i>Prime suggestion by criterion rough</i>	0,9090	79,93
<i>Prime suggestion by criterion BIC</i>	0,8571	86,15

Fonte: Elaborado pelos autores.

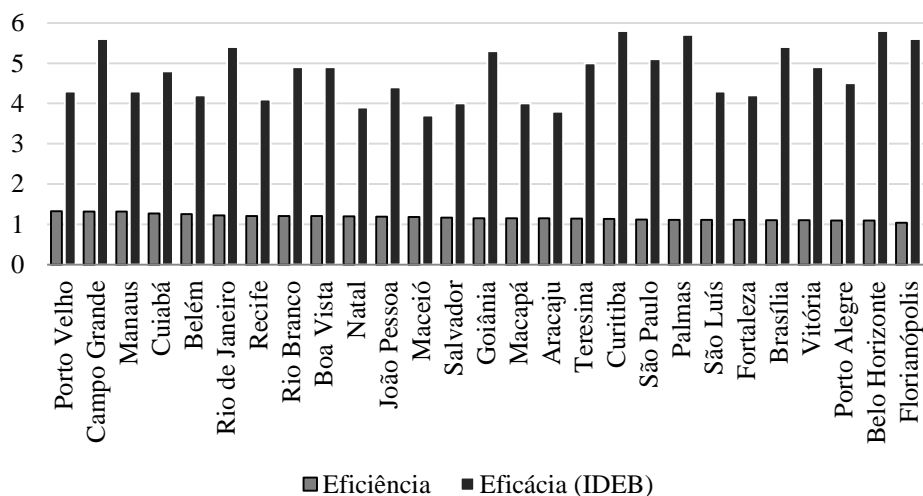
Figura 2: Detecção de outliers e participação das DMUs supereficientes



Fonte: SAEB 2011 e Censo Escolar 2011. Elaborado pelos autores.

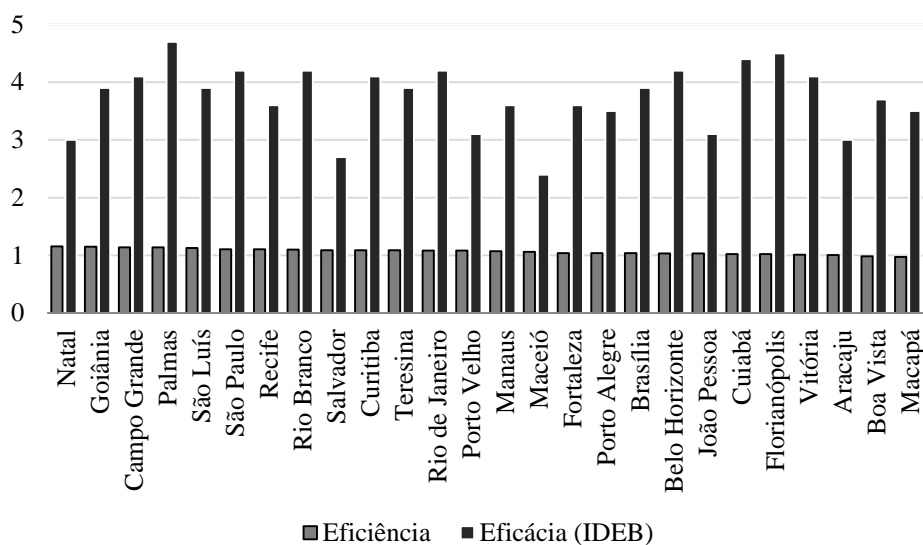
Ao analisar-se a eficiência educacional nas escolas públicas *vis-à-vis* a eficácia das mesmas (IDEB), nota-se uma correlação muito baixa entre estas variáveis para a 4ª série (1,42%) e uma associação linear fraca para a 8ª série (11,50%). Os Gráficos 1 e 2 apresentam um resumo dos indicadores de eficiência e de eficácia por cidade. O critério para a classificação das cidades é a média das eficiências por cidade, que pode esconder muitas nuances, mas permite uma caracterização geral das cidades. A eficiência foi corrigida por *bootstrap* de maneira que se obtêm o intervalo de confiança para o índice e uma distribuição mais confiável após a aplicação da técnica. Cabe ressaltar que a medida de eficiência adotada neste estudo, como proposta por Aragon, Daouia e Thomas-Agnan (2005), permite a presença de *outliers*. Assim, esta medida pode assumir um valor superior à unidade, diferentemente dos métodos DEA e FDH. Deve-se observar que as medidas de eficiência são resultado de um “bom balanceamento entre insumos e produtos” (DELGADO & MACHADO, 2007, p. 447).

Gráfico 1: Ranking segundo a eficiência e o IDEB (Índice de Desenvolvimento da Educação Básica) para a 4ª série



Fonte: SAEB 2011 e Censo Escolar 2011; INEP/MEC. Elaborado pelos autores.

Gráfico 2: Ranking segundo a eficiência e o IDEB (Índice de Desenvolvimento da Educação Básica) para a 8ª série



Fonte: SAEB 2011 e Censo Escolar 2011; INEP/MEC. Elaborado pelos autores.

Como se observa nos Gráficos 1 e 2 a capital que conseguiu adotar estratégias que combina eficiência e eficácia é Rio Branco para a 4ª série e Palmas para a 8ª série do ensino fundamental. Por outro lado, observa-se que Florianópolis é a capital com maior eficácia para a 4ª série, porém está entre as menos eficientes. De forma geral, observa-se que as capitais mais eficientes não são sempre as mais eficazes.

Após a determinação dos índices de eficiência considerando-se três produtos e cinco insumos – 1º estágio – é realizada uma análise de regressão destes índices para cada escola, onde se incluem como covariadas as condições físicas e de localização das escolas e as variáveis de dotação – 2º estágio. O objetivo desta última etapa é encontrar as variáveis que podem afetar as medidas de eficiência escolar e que não estão sob o controle direto das escolas. Como o índice calculado está definido no intervalo $(0, \infty)$ não há censura, como no caso do DEA e do FDH, e, portanto, não se utilizará um modelo *tobit*, diferentemente

do que foi realizado por Delgado & Machado (2007). Para alcançar o objetivo proposto empregou-se o modelo de misturas finitas, que permite o controle da heterogeneidade não observada.

Na Tabela 4 se apresentam as estimativas do teste modificado de razão de verossimilhança para a comparação de dois, três e quatro componentes do modelo de misturas finitas. O critério de seleção será com base no p -valor e na estatística qui-quadrado. Os resultados indicaram a presença de multimodalidade no índice de eficiência da gestão escolar do ensino público. Os critérios AIC e BIC conduziram para um modelo com três componentes, corroborado pelo p -valor. Dessa forma, a influência das covariadas sobre a medida de eficiência não é resultado de uma única distribuição.

Tabela 4: Critério AIC e BIC e Teste Modificado de Razão de Verossimilhança

Série	Componentes	AIC	BIC	Modelos	Teste LR (p -valor)
4ª série	2 (modelo 1)	-9.864,77	-8.863,99	-	-
	3 (modelo 2)	-12.547,98	-11.042,73	Modelo 2 <i>versus</i> Modelo 1	0,000
	4 (modelo 3)	-13.410,80	-11.401,09	Modelo 3 <i>versus</i> Modelo 2	0,008
8ª série	2 (modelo 1)	-14.453,73	-13.493,84	-	-
	3 (modelo 2)	-15.497,43	-14.065,46	Modelo 2 <i>versus</i> Modelo 1	0,000
	4 (modelo 3)	-16.219,36	-14.480,54	Modelo 3 <i>versus</i> Modelo 2	0,210

Fonte: Elaborado pelos autores.

Foram realizadas análises para a 4ª série e a 8ª série do ensino fundamental. Como foram obtidos muitos resultados serão dispostos somente os principais⁴. A Tabela 5 apresenta os resultados do modelo de misturas finitas para a 4ª série. O teste para verificar a existência de heterogeneidade não observada na amostra indicou a sua presença e, portanto, este resultado reforça a ideia de que se deve investigar a eficiência das escolas considerando-se mais de uma distribuição.

A primeira componente agrupa as escolas públicas com os menores índices de eficiência, o que corresponde a 46,12% da amostra analisada neste estudo. Observa-se que a variável *idade* está relacionada de forma negativa com a eficiência da escola. Este resultado está capturando, de forma indireta, a defasagem idade-série, que permite avaliar a qualidade da educação⁵. Deve-se destacar também a importância da variável *jornal* que representa uma *proxy* para o capital educacional do aluno. As variáveis inseridas para capturar a hipótese de Curi e Menezes-Filho (2009) indicaram que os alunos que frequentaram a creche e/ou a pré-escola estão relacionados de forma negativa à eficiência das escolas. Também se observou que a variável capital socioeconômico, para capturar as condições financeiras dos alunos, foi significativa e apresentou o sinal esperado somente para esta componente.

A segunda componente inclui as escolas que foram classificadas como supereficientes pelo método de fronteira parcial quantílica. Estas escolas representam aproximadamente 18% da amostra estudada. A variável *jornal* apresentou um sinal positivo e significativo, porém a magnitude desta variável é superior ao observado para a primeira componente. Além disso, percebe-se que a variável *financeiro*, que foi inserida para capturar a influência da escassez de recursos na escola, apresentou um sinal negativo. Cabe destacar que as escolas que adotaram um sistema de progressão escolar (ou ciclos) apresentam, em média, uma eficiência menor do que aquelas que não o adotaram.

⁴ Todos os resultados podem ser obtidos junto aos autores.

⁵ A literatura tem destacado que a qualidade da educação pode ser avaliada sobre vários pontos de vista, utilizando testes padronizados como o SAEB, a Prova Brasil e o PISA. Indicadores de qualidade, além dos resultados dos testes, incluem a implantação de programas de redução de evasão escolar e de reprovação, a falta de alunos e de professores e a defasagem idade-série.

Tabela 5: Resultados do modelo de misturas finitas para a 4ª série

Variáveis	Componente 1	Componente 2	Componente 3
Branco	-0,0119 (0,00757)	-0,0267 (0,0512)	-0,0183 (0,0151)
Idade	-0,0225*** (0,00310)	-0,00142 (0,0211)	-0,0169*** (0,00593)
Computador	0,00476 (0,00695)	0,0197 (0,0449)	0,00996 (0,0129)
Jornal	0,0202*** (0,00602)	0,0809** (0,0397)	0,00409 (0,0114)
Livros	-0,0151 (0,0114)	-0,0103 (0,0702)	0,0544** (0,0213)
Trabalho fora	0,00514 (0,00825)	0,0604 (0,0559)	0,00637 (0,0158)
Trabalha em casa	0,0123 (0,00919)	0,00135 (0,0564)	-0,000106 (0,0164)
Creche	-0,0145* (0,00749)	-0,134** (0,0520)	-0,0349** (0,0151)
Pré-escola	-0,0157** (0,00739)	-0,137*** (0,0497)	-0,0282* (0,0148)
Não reprovado	-0,0110 (0,00717)	0,0747 (0,0476)	0,0190 (0,0133)
Faz lição	-0,0189 (0,0214)	0,140 (0,128)	0,0476 (0,0511)
Capital socioeconômico	0,0101** (0,00435)	0,0267 (0,0289)	0,000586 (0,00794)
Educação mãe	0,00812 (0,00723)	0,0262 (0,0471)	0,0284** (0,0136)
Financeiro	-0,000320 (0,00193)	-0,0327** (0,0149)	-0,0120*** (0,00393)
Rotatividade	-0,00392* (0,00215)	0,0281* (0,0165)	0,0152*** (0,00437)
Alimentação	-0,0329** (0,0132)	-0,0793 (0,0812)	-0,0173 (0,0282)
Progressão	-0,00107 (0,00259)	-0,0511*** (0,0196)	-0,00382 (0,00523)
Urbano	-0,00817** (0,00386)	1,33e-05 (0,0268)	-0,0155** (0,00775)
Proporção da amostra	46,12%	17,94%	35,94%
Teste de heterogeneidade não observada		4.270,38 [0,0000]	

Fonte: SAEB 2011 e Censo Escolar 2011. Elaborado pelos autores.

A terceira componente, responsável por pouco menos de 36% da amostra, agrupa as escolas com um escore intermediário de eficiência, ou seja, são as instituições de ensino com um nível de eficiência superior às da primeira componente, porém inferior às da segunda componente. Neste caso, a variável *jornal* deixa de ser significativa e a variável *livros*, que indica o percentual de alunos por escola que leu um livro ou mais durante o ano, apresentou um sinal positivo e significativo a 5%. A variável *educação da mãe* corrobora a literatura que aponta que as condições familiares importam para o rendimento das crianças. Assim, estas escolas absorvem filhos de pais com maior escolaridade, ou, pais mais instruídos optam, com maior frequência, por matricular seus filhos nestas escolas. A variável *rotatividade* apresentou um sinal não esperado, indicando que estas escolas estão associadas com uma maior rotatividade de professores.

Para a 8ª série do ensino fundamental, conforme a Tabela 6, o teste de heterogeneidade não rejeita a hipótese de diferença entre as variáveis para cada componente. A primeira componente contém as escolas com os menores escores de eficiência e representa pouco menos de 78% da amostra. As variáveis *computador* e *livros* apresentaram o sinal esperado e, portanto, indicam que maiores níveis de capital

educacional estão associados a maiores índices de eficiência. Também se observa que a variável que indica a proporção de alunos que trabalha fora de casa só foi significativa para esta componente e apresentou o sinal esperado. Assim, as escolas mais eficientes são aquelas que têm menor proporção de alunos que exercem tarefas fora de casa. Merecem destaque as variáveis *não reprovado* e *faz lição*, principalmente a magnitude desta última. Este resultado significa que o aumento de 1% na proporção de alunos que fazem as lições de casa designadas pelo professor aumenta o escore de eficiência, em média, em 0,12 pontos. Por outro lado, as escolas que adotaram programa de progressão escolar aumentam sua medida de ineficiência. Isto significa que os programas de progressão continuada reduzem a taxa de reprovação, mas, ao mesmo tempo, afetam negativamente o desempenho dos alunos nas disciplinas, uma vez que alunos com ainda pouca habilidade em determinado conteúdo são introduzidos em novos conteúdos cujo aprendizado depende do conteúdo anterior (DELGADO & MACHADO, 2007, p. 455).

Tabela 6: **Resultados do modelo de misturas finitas para a 8ª série**

Variáveis	Componente 1	Componente 2	Componente 3
Branco	-0,00473 (0,00730)	-0,00868 (0,0840)	0,0339 (0,0361)
Idade	-0,0248*** (0,00341)	-0,0138 (0,0317)	-0,0232 (0,0190)
Computador	0,0185*** (0,00653)	-0,0613 (0,0729)	0,0147 (0,0289)
Jornal	0,00611 (0,00723)	0,0853 (0,0784)	0,0220 (0,0334)
Livros	0,0575*** (0,0124)	-0,204 (0,144)	-0,0277 (0,0552)
Trabalho fora	-0,0382*** (0,0108)	0,133 (0,114)	-0,0407 (0,0464)
Trabalha em casa	0,00419 (0,00877)	0,0512 (0,0910)	0,0102 (0,0398)
Creche	-0,0388*** (0,00878)	-0,278*** (0,0900)	-0,00940 (0,0415)
Pré-escola	-0,0437*** (0,00804)	-0,242*** (0,0796)	-0,0103 (0,0390)
Não reprovado	0,0536*** (0,00869)	-0,0579 (0,0838)	0,0582* (0,0342)
Faz lição	0,119*** (0,0159)	-0,149 (0,185)	-0,0914 (0,0718)
Capital socioeconômico	-0,0173*** (0,00473)	0,00543 (0,0531)	-0,0132 (0,0190)
Educação mãe	0,000926 (0,00736)	-0,0752 (0,0802)	-0,0271 (0,0359)
Financeiro	-0,000764 (0,00174)	-0,0175 (0,0222)	0,0135 (0,00844)
Rotatividade	-0,00686*** (0,00177)	0,00136 (0,0230)	-0,00105 (0,00789)
Alimentação	-0,0217*** (0,00825)	0,212* (0,115)	-0,0200 (0,0996)
Progressão	-0,0141*** (0,00294)	-0,0182 (0,0403)	-0,0150 (0,0146)
Urbano	-0,00309 (0,00366)	-0,0727* (0,0400)	-0,0315** (0,0147)
Proporção da amostra	77,78%	13,03%	9,19%
Teste de heterogeneidade não observada		8.985,00 [0,0000]	

Fonte: SAEB 2011 e Censo Escolar 2011. Elaborado pelos autores.

Este resultado pode favorecer a ideia de que a formação de turmas de acordo com as estratégias conhecidas como *boutique*⁶ pode aumentar o desempenho dos alunos e, portanto, a qualidade do ensino, que se refletirá sobre a eficiência média das escolas.

As escolas com maior escore de eficiência estão concentradas na segunda componente e representam aproximadamente 13% das escolas públicas. As escolas deste grupo que possuem merenda escolar são, em média, mais eficientes. Os resultados das variáveis *creche* e *pré-escola* não corroboram a hipótese de Curi e Menezes-Filho (2009). Este resultado indica que quanto maior a proporção de alunos que iniciou a trajetória escolar na creche e/ou na pré-escola menor é a eficiência escolar, fator explicado também pela desigualdade de oportunidade no acesso a essa fase da vida escolar, uma vez que menos de 50% dos alunos frequentaram a pré-escola, por exemplo. Para a terceira componente, responsável por menos de 10% da amostra, as escolas com maior percentual de alunos que não foram reprovados apresentam um melhor desempenho em termos de eficiência. Observa-se que as escolas urbanas são, em média, menos eficientes que as escolas rurais. As escolas rurais possuem, de forma geral, menos recursos que as localizadas nas cidades, todavia, elas não são prejudicadas de forma significativa por causa desta escassez de recursos. Este resultado permite inferir que as escolas urbanas não conseguem atingir um balanceamento adequado entre insumos e produtos, diferentemente das demais, que por conta da baixa disponibilidade dos mesmos, utilizam de forma mais eficiente os recursos a sua disposição.

Por fim, com relação às “dotações” das escolas observaram-se resultados interessantes. Enquanto a presença de sala de leitura, de biblioteca e de computadores, por exemplo, aumenta a eficiência das escolas, a subutilização de aparelhos de televisão (e seus complementos, isto é, aparelho de videocassete e antena parabólica) e de laboratórios de ciências não apresenta o mesmo efeito. Como se adotou uma abordagem que parte da hipótese de livre disponibilidade de insumos compreende-se que a readequação da relação entre insumos e produtos é necessária. Para algumas escolas há a presença de recursos de ensino em quantidade superior a sua capacidade de utilização, enquanto observa-se o efeito inverso em outras escolas.

5 Considerações finais

O objetivo deste estudo foi mensurar a eficiência das escolas públicas brasileiras e investigar os seus determinantes. A mensuração da eficiência é fundamental para a formulação de políticas públicas, uma vez que é necessário compreender quais os insumos mais importantes para melhorar a qualidade da educação. Para tanto, propôs-se uma nova abordagem semiparamétrica para a mensuração da eficiência da gestão escolar, que contorna o problema da presença de *outliers* na amostra. Este método permite a construção de uma fronteira de possibilidade de produção (ou fronteira de eficiência na educação) com a presença de escolas supereficientes na determinação do balanceamento entre insumos e produtos escolares. A escolha da proporção destas escolas ocorre de forma endógena. Por sua vez, os determinantes da eficiência escolar são avaliados por meio do método de misturas finitas, que permite a presença de uma amostra composta por mais de uma distribuição de probabilidade. Para alcançar o objetivo proposto foram utilizados os microdados do Censo Escolar 2011 e do SAEB 2011.

Os resultados do teste de heterogeneidade não observada indicaram que o método de misturas finitas é adequado para estudar os determinantes da eficiência escolar. As escolas consideradas ineficientes, com relação à quarta série, são àquelas que apresentam o maior percentual de alunos com idade acima de 10 anos, que apresentam escassez de recursos financeiros e as localizadas na área urbana. Por outro lado, a leitura de jornais e de livros, melhores condições financeiras e maiores níveis educacionais da mãe estão associados de forma positiva ao indicador de eficiência das escolas. Para a oitava série observou-se que os alunos que trabalham fora de casa, os alunos mais velhos, os alunos que iniciaram a trajetória escolar na creche e/ou na creche/pré-escola, as escolas com maiores índices de rotatividade de professores, as

⁶ As estratégias do tipo *boutique* são aquelas em que os alunos estudam em um ambiente homogêneo, de forma que os professores têm maior facilidade para a transmissão do conteúdo (HIDALGO-HIDALGO, 2011).

instituições que adotaram o sistema de ciclos e as escolas localizadas na área urbana apresentam menores índices de eficiência. Além disso, notou-se que as escolas com maiores percentuais de alunos que possuem computadores, que não foram reprovados, que fazem a lição e que leem um maior número de livros apresentaram maiores índices de eficiência.

Por fim, cabe ressaltar que este estudo é o primeiro a empregar uma metodologia que permite a observação de escolas públicas supereficientes para o caso brasileiro. Esta abordagem foi combinada com o método de misturas finitas, que permite avaliar os condicionantes da eficiência escolar por meio de diferentes grupos. Além disso, notou-se que a realocação das “dotações” escolares é primordial para aumentar o nível de eficiência escolar, corroborando a noção de que os resultados das escolas são decorrentes de um bom balanceamento entre insumos e produtos, como defendido por Delgado & Machado (2007).

Referências Bibliográficas

- [1] AFONSO, A. & AUBYN, M. ST. (2006). Cross-country efficiency of secondary education provision: a semiparametric analysis with non-discretionary inputs. *Economic Modelling*, 23(3): 476-491.
- [2] AIGNER, D. J. & CHU, S. F. 1968. On estimating the industry production function. *American Economic Review*, 58(4): 826-839.
- [3] AIGNER, D. J., LOVELL, C. A. K. & SCHMIDT, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6(1): 21-37.
- [4] ALMEIDA, A. T. C. & GASPARINI, C. E. (2011). *Gastos Públicos Municipais e Educação Fundamental na Paraíba: Uma Avaliação usando DEA*. Documentos Técnicos, 42(3), BNB.
- [5] ARAGON, Y., DAOUIA, A. & THOMAS-AGNAN, C. (2005). Nonparametric frontier estimation: a conditional quantilebased approach. *Econometric Theory*, 21(2): 358-389.
- [6] CAMERON, A. & TRIVEDI, P. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.
- [7] CAZALS, C., FLORENS, J. & SIMAR, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, 106(1): 1-25.
- [8] CHARNES A., COOPER, W. W. & RHODES, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6): 429-444.
- [9] CHEN, J. & KALBFLEISCH, J. D. (2005). Modified likelihood ratio test in finite mixture models with a structural parameter. *Journal of Statistical Planning and Inference*, 129(1-2): 93-107.
- [10] COLEMAN, J. S., CAMPBELL, E. Q., HOBSON, C. J., MCPARTLAND, J., MOOD, A. M., WEINFELD, F. D. & YORK, R. L. (1966). *Equality of educational opportunity*. Washington, D.C.: U.S. Government Printing Office.
- [11] CURI, A. Z. & MENEZES-FILHO, N. A. (2009). A Relação entre Educação Pré-Primária, Salários, Escolaridade e Proficiência Escolar no Brasil. *Estudos Econômicos*, 39(4): 811-850.
- [12] DARAIO, C. & SIMAR, L. (2007). *Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications*. New York: Springer.
- [13] DAOUIA, A. & RUIZ-GAZEN, A. (2006). Robust Nonparametric Frontier Estimators: Qualitative Robustness and Influence Function. *Statistica Sinica*, 16: 1233-1253.
- [14] DEPRINS, D., SIMAR, L. & TULKENS, H. (1984). Measuring labor efficiency in post offices. In *The Performance of Public Enterprises: Concepts and Measurement*, ed. MARCHAND, M., PESTIEAU P. & TULKENS, H. Amsterdam: Elsevier.
- [15] DELGADO, V. M. S. & MACHADO, A. F. (2007). Eficiência das escolas públicas estaduais de Minas Gerais. *Pesquisa e Planejamento Econômico*, 37(3): 427-464.
- [16] FARIA, F. P., JANNUZZI, P. M. & SILVA, S. J. (2008). Eficiência dos gastos municipais em saúde e educação: uma investigação através da análise envoltória no estado do Rio de Janeiro. *Revista de Administração Pública*, 42(1): 155-177.

- [17] FELLER, W. (1943). *On a general class of contagious distributions*. *Annals of Mathematical Statistics*, v. 14, p. 389-400.
- [18] GASPARINI, C. E. & MIRANDA, R. B. (2011). Transferências, equidade e eficiência municipal no Brasil. *Planejamento e políticas públicas*, 36: 311-342.
- [19] GRAMANI, M. C. N. & DUARTE, A. L. C. M. (2011). O impacto do desempenho das instituições de educação básica na qualidade do ensino superior. *Ensaio: Avaliação de Políticas Públicas*, 19(72): 679-702.
- [20] HANUSHEK, E. A. (1986). The economics of schooling: production and efficiency in public schools. *Journal of Economic Literature*, 24(3): 1.141-1.177.
- [21] HANUSHEK, E. A. & LUQUE, J. (2003). Efficiency and equity in schools around the world. *Economics of Education Review*, 22(5): 481-502.
- [22] HIDALGO-HIDALGO, M. (2011). On the optimal allocation of students when peer effects are at work: tracking vs. Mixing. *SERIEs*, 2(1):31-52
- [23] JUNIOR, S. P. M., IRFFI, G. & BENEGAS, M. (2011). Análise Da Eficiência Técnica Dos Gastos Com Educação, Saúde E Assistência Social Dos Municípios Cearenses. *Planejamento e Políticas Públicas*, 36: 87-113.
- [24] SACERDOTE, B. (2001). Peer effects with random assignment: results for Dartmouth roommates. *The Quarterly Journal of Economics*, 116(2): 681-704.
- [25] TAUCHMANN, H. (2012). Partial frontier efficiency analysis for Stata. *The Stata Journal*, 12(3): 461-478.
- [26] VELOSO, F. (2009). 15 anos de avanços na educação no Brasil: onde estamos? In: VELOSO, F. et al. (Org.). *Educação básica no Brasil: construindo o país do futuro*. Rio de Janeiro: Campus, 2009. p. 3-24.
- [27] WILSON, P. W. (2005). *Efficiency in education production among PISA countries with emphasis on transitioning economies*. Texas: University of Texas, 40 p.
- [28] ZOGHBI *et al.* (2009). Mensurando o desempenho e a eficiência dos gastos estaduais em educação fundamental e média. *Estudos Econômicos*, 39(4): 785-809.