

## ARTIGOS

## A DESIGUALDADE SOCIOECONÔMICA AFETA MAIS MUNICÍPIOS MENOS FAVORECIDOS?

MARIA CRISTINA GRAMANI

**RESUMO**

*Este estudo busca analisar o panorama da educação básica em um país emergente que se caracteriza, por um lado, pelo elevado desenvolvimento econômico e, por outro, pela alta desigualdade socioeconômica. Um modelo de dois estágios é utilizado, sendo que o primeiro estágio usa variáveis diretamente relacionadas à educação para capturar a eficiência educacional de cada município e o segundo emprega a regressão Tobit a fim de estimar a influência das variáveis ambientais (não discricionárias) sobre a eficiência educacional encontrada no primeiro passo. Um agrupamento dos municípios em clusters foi implementado para assegurar uma comparação justa entre municípios homogêneos. Os resultados mostram discrepâncias significativas na influência de variáveis socioeconômicas no resultado educacional, dependendo da prosperidade de cada cluster.*

**DATA ENVELOPMENT ANALYSIS - DEA • EFICIÊNCIA EDUCACIONAL • EDUCAÇÃO BÁSICA • DESIGUALDADE SOCIAL**

## DOES SOCIOECONOMIC INEQUALITY AFFECT EDUCATION MORE IN LESS WEALTHY MUNICIPALITIES?

**ABSTRACT**

*This study attempts to capture the full picture of educational development in an emerging country that is characterized by both high economic development and high socioeconomic inequality. A two-step model is used in this study. The first step uses the variables that are directly related to education to capture the educational efficiency of each municipality; the second step uses a statistical Tobit model to estimate the influence of the non-discretionary variables on the educational efficiency found in the first step. A previous categorization by clusters is also implemented to ensure a fair comparison among homogeneous municipalities. The results show significant discrepancies in the influence of socioeconomic variables on educational outcome, which depends on the welfare of the cluster.*

**DATA ENVELOPMENT ANALYSIS - DEA • EDUCATIONAL EFFICIENCY • BASIC EDUCATION • SOCIAL INEQUALITIES**

## L'INÉGALITÉ SOCIO-ÉCONOMIQUE AFFECTE-T-ELLE PLUS L'ÉDUCATION DANS LES COMMUNES MOINS RICHES?

### RÉSUMÉ

*L'objectif de cet étude est d'analyser le panorama de l'enseignement obligatoire dans un pays émergent qui, d'un côté, se caractérise par son développement économique et, de l'autre, par sa grande inégalité socio-économique. Pour cela, nous avons choisi un modèle à deux niveaux, le premier utilise des variables directement liées à l'éducation, afin de juger de l'efficacité éducative de chaque commune; et le second, à l'aide de la méthode de régression Tobit, estime l'influence des variables environnementales (non discriminatoires) sur cette efficacité. Un regroupement des communes en clusters a été mis en place pour permettre une comparaison équitable entre communes homogènes. Les résultats montrent des écarts significatifs concernant l'influence des variables socio-économiques dans les résultats éducatifs, en fonction de la richesse du cluster.*

**DATA ENVELOPMENT ANALYSIS - DEA • EFFICACITÉ ÉDUCATIVE •  
ÉDUCATION BASIQUE • INÉGALITÉ SOCIALE**

## ¿LA DESIGUALDAD ECONÓMICA AFECTA MÁS A LOS MUNICIPIOS MENOS FAVORECIDOS?

### RESUMEN

*Este estudio busca analizar el panorama de la educación básica en un país emergente que se caracteriza, por un lado, por el elevado desarrollo económico y, por otro, por la alta desigualdad socioeconómica. Se utiliza un modelo de dos fases; la primera fase usa variables directamente relacionadas con la educación para capturar la eficiencia educacional de cada municipio, y la segunda emplea la regresión Tobit a fin de estimar la influencia de las variables ambientales (no discrecionales) sobre la eficiencia educacional encontrada en el primer paso. Se implementó un agrupamiento de los municipios en clusters para asegurar una justa comparación entre municipios homogéneos. Los resultados muestran significativas discrepancias en la influencia de variables socioeconómicas en el resultado educacional, en función de la prosperidad de cada cluster.*

**DATA ENVELOPMENT ANALYSIS - DEA • EFICIENCIA EDUCACIONAL •  
EDUCACIÓN BÁSICA • DESIGUALDAD SOCIAL**

**O** DESENVOLVIMENTO DE UM PAÍS É MEDIDO POR FATORES TAIS COMO DESEMPENHO econômico, qualidade de vida, sustentabilidade e igualdade. Países em desenvolvimento apresentam discrepâncias com relação a esses fatores; por exemplo, um país emergente pode ter um alto Produto Interno Bruto – PIB –, mas uma posição não satisfatória no Índice de Desenvolvimento Humano – IDH. Essa divergência é o caso, por exemplo, de Brasil, Índia e Rússia, que, em 2014, ocupavam a 7ª, 9ª e 10ª posições, respectivamente, em termos de PIB, mas estavam na 75ª, 130ª e 50ª no IDH.<sup>1</sup>

Mas a alta desigualdade pode também ser encontrada entre regiões do país. Por exemplo, o Brasil registra o terceiro maior índice de Gini na América Latina (quanto mais elevado esse índice, menos igualitária é a distribuição de renda), entretanto, possui cidades com baixo índice Gini de 0,28 (similar ao da Noruega, de 0,25) e outras onde esse valor atinge 0,8. Tais desigualdades significativas indicam que propostas de melhorias devem estar relacionadas a diferentes aspectos para cada uma das regiões de um país e, conseqüentemente, as políticas educacionais também devem ser diferentes. Rumberger e Palardy (2005), em estudo baseado em 14.199 estudantes de escolas americanas, mostram que políticas destinadas à melhoria da *performance* de uma região podem não melhorar ou até mesmo piorar o desempenho de outra região.

Portanto, torna-se importante a análise das discrepâncias entre regiões de países em desenvolvimento, como o Brasil, pois, a despeito do alto desenvolvimento econômico, as diferentes culturas e as grandes

<sup>1</sup> Disponível em: <<http://data.worldbank.org>>; <<http://www.br.undp.org>>. Acesso em: 03 maio 2017.

desigualdades entre regiões podem levar a um desempenho insatisfatório em áreas-chave como a educação.

Outro ponto importante é que o desempenho da educação é significativamente afetado por variáveis não discricionárias, sendo necessário usar fatores na análise que não estão diretamente relacionados com educação, como escolas, professores e gastos. Em países em desenvolvimento com alta desigualdade, aspectos ligados às condições básicas, tais como acesso a sistemas de esgoto e tratamento de água, e ao desenvolvimento social, como educação materna, analfabetismo, taxa de abandono, entre outros, devem ser considerados na análise da *performance* educacional. Essa investigação pode auxiliar na tomada de decisão do governo para determinar quais recursos realmente afetarão o desempenho da educação em cada região ou cidade de um país. Em um país em desenvolvimento, do tamanho de um continente e com alta desigualdade socioeconômica, como o Brasil, as potenciais melhorias devem diferir de região para região, tornando-se importante a questão: “a desigualdade socioeconômica afeta mais a educação em municípios menos ricos?”.

A contribuição deste artigo se divide em três aspectos principais. Primeiro, devido às discrepâncias existentes dentro do país, categorizamos os municípios brasileiros em *clusters* homogêneos para realizar uma comparação mais justa. Segundo, a fim de considerar a relação existente entre eficiência educacional e indicadores socioeconômicos, construímos um modelo de dois estágios, em que no primeiro são analisadas as diferentes necessidades dentro do país. Por fim, no segundo estágio, o modelo mostra o impacto das variáveis não discricionárias na eficiência educacional. No primeiro passo o modelo Data Envelopment Analysis – DEA – é usado para medir o uso dos recursos a fim de encontrar a melhor *performance* de cada município, e a regressão Tobit é usada no segundo passo a fim de estimar o impacto da alta desigualdade socioeconômica na eficiência educacional dos municípios.

A próxima seção traz uma breve revisão da literatura acerca da eficiência educacional, seguida por uma descrição do sistema educacional brasileiro, incluindo sua estrutura, financiamento e avaliação. Posteriormente são apresentados o modelo de dois estágios, o método de clusterização e a base de dados utilizada para o cálculo da eficiência educacional e a análise do impacto das condições socioeconômicas em países em desenvolvimento. Discutem-se os resultados empíricos do modelo de dois estágios e, finalmente, são tecidas as conclusões do estudo.

## EFICIÊNCIA EDUCACIONAL: REVISÃO DA LITERATURA

A literatura sobre eficiência educacional inclui vários estudos que utilizam variáveis diretamente relacionadas à educação, como infraestrutura,

professores e avaliações nacionais, bem como variáveis socioeconômicas, como ocupação e nível educacional dos pais, bolsas de estudo, diferentes contextos culturais e sociais, escolas rurais e urbanas, renda e pobreza (MANCEBON; BANDRES, 1999; AFONSO; AUBYN, 2006; PRIMONT; DOMAZLICKY, 2006; RASSOULI-CURRIER, 2007; HOMLUND; MCNALLY; VIARENGO, 2010). Considerando os determinantes da eficiência educacional em um ambiente com alta desigualdade, como nos países em desenvolvimento, Hauner (2008) observou diferenças significativas na eficiência educacional e na saúde entre as regiões da Rússia e sugeriu que, em média, os resultados atuais de saúde e educação poderiam ser produzidos com aproximadamente dois terços dos recursos atuais se as regiões menos eficientes competissem com as mais eficientes. Do mesmo modo, Rodríguez-Pose e Tselios (2011) revelaram uma forte correlação entre os níveis de escolaridade e desigualdade em 102 regiões da Europa Ocidental. Henriquez *et al.* (2012), analisando o sistema educacional altamente heterogêneo do Chile, verificaram a existência de escolas que atendem com sucesso alunos de baixa renda. Os estudos existentes no Brasil mostram que as regiões mais desenvolvidas economicamente (Sudeste e Sul) apresentam desempenho superior na educação. O anuário do movimento Todos Pela Educação (2012) corrobora essas diferenças de eficiências regionais do Brasil, com o Nordeste registrando as maiores desvantagens em termos de taxas de aprovação, taxas de abandono escolar e avaliação nacional.

Os países em desenvolvimento caracterizam-se por necessidades básicas urgentes relacionadas às condições de qualidade vida, como a infraestrutura e o acesso aos sistemas de esgotos e água tratada, em contraste com os países desenvolvidos, que devem melhorar em aspectos mais avançados. Murillo e Roman (2011, p. 29) ressaltam que

[...] a disponibilidade de infraestrutura e serviços básicos (água, eletricidade, esgoto), instalações didáticas (instalações desportivas, laboratórios, bibliotecas), bem como o número de livros na biblioteca e de computadores na escola, têm um efeito sobre a avaliação de alunos do ensino primário na América Latina, mas seu peso relativo varia significativamente de país para país.

As discrepâncias socioeconômicas entre as regiões de um país em desenvolvimento podem ser alarmantes. Por exemplo, segundo o Observatório do Plano Nacional de Educação,<sup>2</sup> entre os municípios do Nordeste brasileiro, o tratamento de água varia de 0,15% a 99% e o acesso a sistemas de esgoto varia de 0% a 100%. Assim, embora a contribuição da desigualdade entre regiões de um país seja considerada pequena para a desigualdade internacional (UNITED NATIONS, 2006), é importante compreender e defender políticas que abordem as disparidades

2

Disponível em: <<http://www.observatoriodopne.org.br/>>.  
Acesso em: 03 maio 2017.

regionais em nações emergentes, pois estas variáveis afetam significativamente a eficiência educacional e outros fatores.

Dois pontos importantes distinguem o presente trabalho de outros artigos. Primeiro, analisamos quase todos os municípios de um país em desenvolvimento com elevado PIB, mas posicionado muito abaixo do esperado no IDH. Em segundo lugar, em vez de comparar os municípios brasileiros por região, distribuímos os municípios em três *clusters* utilizando a metodologia *K-means* para assegurar uma comparação justa entre aqueles que são homogêneos. Assim como em Henriquez *et al.* (2012), no Brasil também é verificada a existência de escolas que atendem com sucesso alunos de baixa renda e, portanto, nesse estudo, não vamos considerar o Brasil por regiões, mas por *clusters* de municípios.

## O SISTEMA EDUCACIONAL BRASILEIRO

### A INFRAESTRUTURA

A educação básica no Brasil é dividida em dois níveis: o ensino fundamental e o ensino médio. O primeiro é dividido em fundamental I, que inclui os níveis 1-5 (alunos com 6 a 10 anos de idade) e fundamental II, que engloba os níveis 6-9 (alunos de 11 a 14 anos). O ensino médio contém os anos 1-3 (estudantes de 15 a 17 anos). A educação é obrigatória entre os 4 e os 17 anos, incluindo a educação pré-escolar, fundamental e média. O sistema educacional é obrigado a assegurar essas matrículas.

No Brasil, o ensino fundamental I é ministrado por professores de várias disciplinas que interagem com o mesmo grupo de alunos durante todo o ano letivo. No nível fundamental II, os professores são especialistas, sendo em geral um professor para cada matéria. Os professores da educação básica são obrigados a realizar cursos superiores, embora esta não seja a realidade no Brasil. Aproximadamente 30% dos professores do fundamental I não possuem ensino superior e cerca de 70% dos docentes do fundamental II não têm formação superior na área em que ensinam (Tabela 1).

**TABELA 1**  
**PORCENTAGEM DE ALUNOS E PROFESSORES DO ENSINO FUNDAMENTAL E INVESTIMENTO PÚBLICO NA EDUCAÇÃO - BRASIL - 2007-2011**

ALUNOS, PROFESSORES E INVESTIMENTO	2007	2008	2009	2010	2011
Percentual de alunos (6-14 anos) matriculados no ensino fundamental	92,6	93,8	98,3	-	93,7
Percentual de professores do ensino fundamental com formação superior	69,8	68,9	69,7	70,8	73,0
Percentual de professores do ensino fundamental I com formação superior	62,5	61,6	62,8	64,0	66,3
Percentual de professores do ensino fundamental II com formação superior	-	-	63,9	65,0	71,4
Percentual de professores do ensino fundamental II com formação superior na mesma área que atuam	-	-	14,2	15,3	28,7
Investimento público direto na educação básica por aluno (R\$)	3.073,5	3.524,2	3.813,9	4.353,0	4.916,3

Fonte: Observatório do PNE (<http://www.observatoriodopne.org.br/downloads>).

## FINANCIAMENTO

Segundo a Organização para Cooperação e Desenvolvimento Econômico – OCDE –, em seu relatório *Panorama da educação* de 2013, o investimento público brasileiro em educação cresceu significativamente na última década, passando de 3,5% do PIB em 2000 para 5,6% em 2010. No entanto, o nível de gastos continua abaixo da média nacional da OCDE, que foi de 6,3% em 2010 (OCDE, 2013). De acordo com o Instituto Nacional de Educação e Pesquisa Educacional Anísio Teixeira – Inep –, as escolas públicas são responsáveis pela educação da maioria da população, respondendo, no nível fundamental, por 87% dos estudantes brasileiros.<sup>3</sup>

O sistema de distribuição de financiamento da educação pública básica no Brasil tem sofrido mudanças. Segundo Ulyssea, Fernandes e Gremaud (2006), até meados dos anos 1990, o financiamento educacional estava completamente descentralizado e o volume de financiamento dependia exclusivamente da receita e do número de matrículas em cada unidade federal. Em 1996, foi criado o Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério – Fundef –, projetado para redistribuir os recursos da educação para os governos estaduais e municipais, de acordo com o número de alunos matriculados anualmente no ensino fundamental. Em 2007, o Fundef foi estendido a todos os níveis de educação básica (anteriormente limitado aos níveis fundamentais) e substituído pelo Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação – Fundeb. De acordo com as regras atuais, o montante a ser distribuído depende do custo por aluno de cada estado, que é calculado com base na receita estimada do fundo e no número de alunos

3

O Inep é uma agência federal do Ministério da Educação – MEC. Sua missão é promover estudos, pesquisas e análises do sistema educacional brasileiro para apoiar a formulação e implementação de políticas públicas de educação, utilizando parâmetros de qualidade e equidade, e produzir informações claras e confiáveis para gestores, pesquisadores, educadores e público em geral ([www.inep.gov.br](http://www.inep.gov.br)).

matriculados no ensino básico em escolas públicas (estaduais e municipais, segundo o último recenseamento escolar).<sup>4</sup>

Embora a expansão do Fundeb fosse importante para o país, “o desafio agora é disponibilizar esses recursos e permitir um fácil acesso a esses níveis de educação para todas as crianças e jovens do país” (FUNDO DAS NAÇÕES UNIDAS PARA A INFÂNCIA – UNICEF, 2009, p. 42).

## AVALIAÇÃO

A avaliação do desempenho da educação básica brasileira é recente. O Índice de Desenvolvimento da Educação Básica – Ideb – foi criado pelo Inep em 2007 para combinar, em um único indicador, dois conceitos relacionados à qualidade educacional: de fluxo e desempenho escolar em avaliações nacionais. O fluxo escolar é medido com base nos dados das aprovações obtidas no censo escolar e o desempenho nas avaliações é calculado por meio de dois testes nacionais: o Sistema Nacional de Avaliação da Educação Básica – Saeb – e a Prova Brasil. Uma instituição com melhor desempenho no teste e menores taxas de reprovação e abandono receberá uma classificação melhor. O Ideb fornece um resultado de 0 a 10 e é aplicado a estudantes de escolas públicas e privadas nos 5º e 9º anos da educação fundamental e no 3º ano do ensino médio. O cálculo para o índice é o seguinte:

$$IDEB_{ji} = N_{ji} * P_{ji} \quad (1)$$

$$0 \leq N_{ji} \leq 10 ; 0 \leq P_{ji} \leq 1 ; 0 \leq IDEB \leq 10 \quad (2),$$

onde:

$i$  = ano da avaliação nacional (Saeb e Prova Brasil);

$N_{ji}$  = proficiência média em português e matemática, padronizada com um intervalo de 0 a 10, para os alunos da escola  $j$  que foi obtida na avaliação nacional (Saeb e Prova Brasil) no ano  $i$ ;

$P_{ji}$  = indicador baseado na taxa de aprovação dos alunos da escola  $j$  relativa ao nível de educação, no ano  $i$ .

O uso de um indicador de *output* para medir o desempenho educacional (a avaliação nacional) é consenso na literatura (BRYK; THUM, 1989; RUMBERGER, 1995; NEAL, 1997; RUMBERGER; THOMAS, 2000; LEE; BURKAM, 2003).

## MÉTODO

### MODELO DEA EM DOIS ESTÁGIOS

A metodologia DEA em dois estágios é composta por duas etapas: a primeira avalia a eficiência educacional levando em conta variáveis diretamente relacionadas à educação; e a segunda consiste em utilizar

<sup>4</sup> Disponível em: <www.mec.gov.br>. Acesso em: 03 maio 2017.



a regressão dos fatores de desigualdade socioeconômica e determinar qual destes mais afeta a eficiência educacional.

Primeiro estágio: medida de eficiência por meio da técnica DEA

Para avaliar a eficiência educacional, usamos a técnica DEA, que se originou no trabalho de Charnes, Cooper e Rhodes (1978). A técnica DEA é um método não paramétrico que avalia a eficiência relativa de unidades de produção (*decision-making units* – DMUs), as quais podem ser representadas por divisões ou unidades administrativas, tais como municípios, desde que sejam homogêneas; isto é, esses grupos devem produzir os mesmos bens e serviços (*outputs*) usando os mesmos recursos (insumos), e eles variam apenas em sua intensidade/quantidade (DYSON et al., 2001). Para otimizar cada DMU e determinar a fronteira de eficiência, podem ser usados dois modelos clássicos de DEA: com retornos de escala constantes (*decision-making units* – CRS) ou com retornos de escala variáveis (*variable returns to scale* – VRS). O modelo CRS assume que um aumento em um *input* gera um crescimento proporcional no *output*. O modelo VRS permite ganhos na eficiência produtiva à medida que a escala de produção varia, assumindo que não há proporcionalidade entre as variações nos *inputs* e *outputs*.

Além disso, são possíveis dois tipos de otimização. De acordo com Coelli et al. (2005, p. 54), a eficiência técnica orientada para os *inputs* aborda a seguinte questão: “Em quanto as quantidades de *input* podem ser proporcionalmente reduzidas sem alterar as quantidades de *output* produzidas?”. A medida orientada a *outputs* aborda uma questão diferente: “Em quanto os *outputs* podem ser expandidos sem alterar as quantidades de *input* utilizadas?”.

Matematicamente, o modelo DEA orientado a *output*, VRS, calcula a eficiência de  $I$  DMUs usando  $S$  *inputs* e  $M$  *outputs*, para cada DMU a fim de determinar a maior produção de *outputs*, com os *inputs* mantidos constantes. A eficiência é definida por  $\phi$ .

$$\max_{\phi, \lambda} \phi,$$

$$\text{s.a. } -\phi q_i + Q\lambda \geq 0 \quad (3)$$

$$x_i - X\lambda \geq 0 \quad (4)$$

$$N1'\lambda = 1 \quad (5)$$

$$\lambda \geq 0 \quad (6)$$

onde:

$S$  representa a quantidade de *inputs*;

$M$  representa a quantidade de *outputs*;

$I$  representa a quantidade de DMUs;

$x_i$  é um vetor  $S \times 1$  de *inputs* da  $i$ -th DMU;

$q_i$  é um vetor  $M \times 1$  de *outputs* da  $i$ -th DMU;

$X$  é uma matriz  $S \times I$  de *inputs* e  $Q$  é uma matriz  $M \times I$  de *outputs*;  
 $\phi$  é um escalar  $1 \leq \phi < \infty$ ;  
 $\lambda$  é um vetor  $I \times 1$  de constantes;  
 $N_1$  é um vetor  $I \times 1$  de 1's.

Geralmente, na primeira etapa, a eficiência educacional é avaliada usando, como *inputs*, as variáveis relacionadas a alunos, gastos e professores e, como *outputs*, os resultados de avaliações nacionais (Quadro 1).

**QUADRO 1**  
**VARIÁVEIS USADAS EM MODELOS DE DOIS ESTÁGIOS, SEGUNDO OUTROS AUTORES**

AUTORES	PRIMEIRO ESTÁGIO		SEGUNDO ESTÁGIO
	DEA <i>INPUTS</i>	DEA <i>OUTPUTS</i>	VARIÁVEIS INDEPENDENTES
Ray (1991)	Professores em sala de aula por aluno Pessoal de apoio por aluno Pessoal administrativo por aluno	Nota de matemática Nota de argumentação Nota de escrita Nota de leitura	Rendimento <i>per capita</i> da educação dos pais Valor médio das unidades habitacionais ocupadas pelo proprietário Porcentagem de estudantes, na região, pertencentes a grupos étnicos minoritários Porcentagem de alunos de famílias que recebem ajuda e famílias com filhos dependentes Porcentagem de famílias com renda abaixo do nível de pobreza em 1979 Porcentagem de crianças na comunidade de famílias com apenas o pai, ou apenas a mãe
Kirjavainen e Loikkanen (1998)	Horas de ensino por semana Horas de não aula por semana Experiência de professores Educação de professores Nível de admissão ou nível educacional dos pais dos alunos	Número de alunos aprovados Número de concluintes Notas em disciplinas obrigatórias no exame de matrícula Notas em disciplinas adicionais no exame de matrícula	Tamanho da escola Tamanho da sala Número de escolas privadas Número de mulheres Heterogeneidade Nível de educação dos pais
Afonso e Aubyn (2006)	Razão professor-aluno Horas na escola	Nota no Pisa	PIB <i>per capita</i> Nível educacional dos pais
Bradley, Johnes e Millington (2001)	Proporção de alunos não elegíveis para refeições escolares gratuitas Proporção de professores qualificados	Proporção de 5 + GCSE (Certificado Geral de Ensino Secundário) notas Taxa de frequência	Independência da escola em relação ao controle das autoridades locais Grau de concorrência entre escolas não seletivas Composição de gênero da escola Densidade populacional Taxa de desemprego local Áreas com alta proporção de profissionais e gerentes Despesas com professores e livros Tamanho da escola
Agasisti (2013)	Relação aluno-professor Proporção de computadores conectados à web Indicadores, para cada escola, de média escolar.	Resultado do Pisa em matemática Resultado do Pisa em ciências	Local da escola Escola localizada em uma cidade pequena ou cidade grande Escolas acadêmicas, técnicas e privadas Tamanho da escola e tamanho da turma, em termos de número de alunos Porcentagem de meninas Pressão parental

Fonte: Elaboração da autora.

Assim, propomos um modelo em que as DMUs consideradas correspondem a 5.020 municípios brasileiros, três *inputs* e um *output*. Os *inputs* são a despesa por aluno, a percentagem de professores com formação superior (FORM\_SUP) e a percentagem de professores com formação superior nas áreas em que ensinam (FORM\_SUP\_ESP). O *output* é o resultado do Ideb, obtido ao final do fundamental II, para as escolas públicas de cada município, que capta não apenas a avaliação, mas também a taxa de aprovação dos alunos. O modelo de eficiência educacional utilizado nesse estudo é orientado para o *output*, pois tentamos maximizar o desempenho dos municípios, diante de seus recursos existentes.

Segundo estágio: regressão Tobit

É comum no segundo estágio usar uma técnica de regressão para avaliar os efeitos das variáveis ambientais, ou não discricionárias, nos resultados obtidos no primeiro estágio (RAY, 1991; KIRJAVAINEN; LOIKKANEN, 1998; BRADLEY; JOHNES; MILLINGTON, 2001; AFONSO; AUBYN, 2006; AGASISTI, 2013). De acordo com Liu *et al.* (2013, p. 899), “metodologicamente, os estudos de educação básica mostram preferência pelo método DEA contextual em dois estágios”. Simar e Wilson (2004) apresentam uma extensa lista de artigos usando a abordagem em dois estágios.

O Método dos Mínimos Quadrados não é apropriado para determinar o efeito dos aspectos ambientais sobre a eficiência educacional dos municípios, devido à natureza da variável dependente. No nosso caso, a variável dependente, o escore de eficiência, está entre 0 e 1 e a distribuição não é normal, mas censurada em 0 e 1; portanto, a regressão Tobit é mais apropriada para análise. Muitos estudos na literatura utilizam os métodos DEA e Tobit, como Kirjavainen e Loikkanen (1998), e aqui queremos meramente testar se algum fator socioeconômico tem poder explicativo para as diferenças de eficiência educacional.

Portanto, a variável dependente empregada no modelo Tobit neste trabalho é a de eficiência obtida pela DEA no primeiro estágio, enquanto as variáveis independentes são:

- acesso ao sistema de esgoto: representa condições básicas de qualidade de vida, principalmente para os países em desenvolvimento. De acordo com o Health Systems 20/20 (AGENCY FOR INTERNATIONAL DEVELOPMENT IN COOPERATION, 2012, p. 101), “Quase metade das pessoas no mundo em desenvolvimento têm uma ou mais das principais doenças ou infecções associadas com abastecimento de água inadequado e falta de saneamento”;
- educação materna: um grande número de estudos na literatura relata uma relação significativa entre educação materna e desigualdade socioeconômica (HANUSHEK; LUQUE, 2003; CHUDGAR; LUSCHEI, 2009; CHUDGAR; SHAFIK, 2010);

- taxa de abandono escolar: de acordo com Fernandes (2007, p. 7), “Um sistema ideal seria aquele em que todas as crianças e adolescentes tivessem acesso à escola, não perdessem tempo com repetição de ano, não saíssem da escola precocemente e no final, todos aprendessem”. No entanto, de acordo com um relatório da OCDE (2012), o Brasil ainda precisa superar as disparidades geográficas no número de alunos que abandonam e repetem de ano;
- índice de Gini: mede o grau de desigualdade na distribuição da renda familiar em um país (números menores são melhores); por exemplo, no Brasil, em 2012, esse índice variava de 0,28 a 0,80;<sup>5</sup>
- Renda *per capita*: trata-se de uma *proxy* para a riqueza do município de um país em desenvolvimento;
- taxa de analfabetismo: embora tenha melhorado no Brasil, a taxa de alfabetização ainda varia de 0% a 39% entre os municípios;
- defasagem escolar: indica a percentagem da população com idade entre 11 e 13 anos que frequentou os últimos anos do nível fundamental ou já completou este nível. Esta variável é uma *proxy* para a evolução do aluno nas escolas.

O modelo de regressão pode ser definido como:

$$y = \begin{cases} y^*; & 0 \leq y^* \leq 1 & (7) \\ 0; & y^* < 0 & (8) \\ 1; & 1 < y^* & (9) \end{cases}$$

$$y^* = \beta x_i + \mathcal{E}_i \quad (10)$$

onde:

$y$  é o escore de eficiência do modelo DEA;

$y^*$  é a variável latente (ou não observável);

$\beta$  são os parâmetros do modelo;

$x_i$  é o vetor das variáveis não discricionárias.

Dessa forma, o modelo de regressão pretende mostrar qual variável não discricionária tem um efeito significativo sobre a eficiência educacional dos municípios. No Brasil, é difícil identificar um único fator educacional ou socioeconômico que tem maior potencial de influência na melhoria da eficiência educacional, pois as dimensões continentais do país, as altas desigualdades e a disponibilidade de recursos básicos não são ideais.

## CLUSTERIZAÇÃO DOS DADOS: MÉTODO *K-MEANS*

Outros estudos usam, em geral, regiões para classificação; entretanto, como observado anteriormente, há grandes discrepâncias mesmo dentro de regiões do Brasil. Portanto, para garantir a homogeneidade das DMUs, agregamos os municípios por características comuns usando o algoritmo de agrupamento de *K-means*. Segundo Yang (2010, p. 239),

[...] os algoritmos de agrupamento são uma coleção de métodos para dividir um conjunto de  $n$  observações em *clusters* de modo que os membros do mesmo *cluster* têm características mais comuns do que membros de diferentes grupos.

A metodologia *K-means* é um algoritmo de agrupamento simples e frequentemente usado, proposto pela primeira vez em 1955 (JAIN; DUBES, 1988; DRINEAS; KANNAN; MAHONEY, 2006; JAIN, 2010). O algoritmo *K-means* é baseado em vários grupos (*clusters*) que são predefinidos ( $k$ ) de tal forma que o erro quadrático entre a média empírica de um *cluster* e os pontos no *cluster* é minimizado.

De acordo com Jain (2010), seja  $X = \{x_i\}$ ,  $i = 1, \dots, n$  o conjunto de  $n$  pontos  $d$ -dimensional a serem agrupados em um conjunto de  $K$  *clusters*,  $C = \{C_k, k = 1, \dots, K\}$ . Seja  $\mu_k$  a média do *clusters*  $C_k$ . O erro quadrático entre  $\mu_k$  e os pontos no *clusters*  $C_k$  é definido por:

$$J(C_k) = \sum_{x_i \in C_k} \|x_i - \mu_k\|_2^2. \quad (11)$$

O objetivo de *K-means* é minimizar a soma dos erros quadrados sobre todos os  $K$  *clusters*. Como os erros quadráticos sempre diminuem com o número de *clusters*  $K$  (com  $J(C) = 0$  quando  $K = n$ ), a soma pode ser minimizada apenas para um número fixo de *clusters*. Como a complexidade do problema de clusterização *K-means* é *NP-hard*, algoritmos heurísticos são geralmente utilizados.

Assim, antes de aplicar o modelo DEA, utilizamos o algoritmo *K-means* para agrupar os municípios em três *clusters*, considerando todas as variáveis educacionais e socioeconômicas. O uso de *clusters* pode diminuir os possíveis erros estatísticos que não são capturados pela DEA (por meio da existência de um *outlier* ou de medidas de erro).

## BASE DE DADOS E ESTATÍSTICA DESCRITIVA

Todos os dados utilizados neste estudo foram obtidos a partir de bases de dados públicas fornecidas por agências governamentais em seus *sites*, como mostra o Quadro 2.

**QUADRO 2**  
**DEFINIÇÃO E FONTE DAS VARIÁVEIS**

VARIÁVEL	DEFINIÇÃO	FONTE
Gasto por aluno	Despesas de educação no nível de ensino fundamental, por aluno em escola pública, para cada município	Fundo Nacional de Desenvolvimento da Educação - FNDE - Sistema Siope, 2010
FORM_SUP	Percentual de professores, no nível fundamental, com grau de ensino superior	< <a href="http://www.observatoriodopne.org.br/">http://www.observatoriodopne.org.br/</a> >, 2010
FORM_SUP_ESP	Percentual de professores, no nível fundamental II, com grau de ensino superior na área da disciplina que lecionam	< <a href="http://www.observatoriodopne.org.br/">http://www.observatoriodopne.org.br/</a> >, 2010
Ideb	Nota obtida no Índice de Desenvolvimento da Educação Básica, no último ano do fundamental II, nas escolas públicas	< <a href="http://ideb.inep.gov.br/">http://ideb.inep.gov.br/</a> >
Acesso ao sistema de esgoto	Percentual de escolas públicas, no nível fundamental, com acesso a sistemas de esgoto	< <a href="http://www.observatoriodopne.org.br/">http://www.observatoriodopne.org.br/</a> >, 2010
Educação materna	Percentual de mães, sem escolaridade no nível fundamental completa e com filhos menores de 15 anos (2010)	< <a href="http://www.pnud.org.br">www.pnud.org.br</a> >, 2010
Taxa de abandono	Taxa de abandono escolar para o nível fundamental de educação pública por município. A taxa de abandono escolar é definida como o percentual de alunos matriculados que, numa determinada série, não frequentam a escola durante o ano letivo.	Inep - Estatísticas Educacionais - Indicadores Educacionais, 2011
Índice de Gini	Índice Gini de renda familiar <i>per capita</i> de pessoas que vivem em uma determinada área geográfica, no ano em curso	IBGE - Instituto Brasileiro de Geografia e Estatística, 2010 < <a href="http://tabnet.datasus.gov.br/cgi/ibge/censo/cnv/ginibr.def">http://tabnet.datasus.gov.br/cgi/ibge/censo/cnv/ginibr.def</a> >
Renda	Renda <i>per capita</i> , 2010	IBGE - Instituto Brasileiro de Geografia e Estatística, 2010 < <a href="http://tabnet.datasus.gov.br/cgi/ibge/censo">http://tabnet.datasus.gov.br/cgi/ibge/censo</a> >
Defasagem escolar	Percentual da população de 11 a 13 anos de idade frequentando os anos finais do fundamental ou que já concluiu o fundamental	<i>Atlas de Desenvolvimento Humano</i> , 2010 < <a href="http://www.atlasbrasil.org.br">http://www.atlasbrasil.org.br</a> >
Taxa de analfabetismo	Taxa da população de 11 a 14 anos de idade que não consegue ler ou escrever um bilhete simples e o número total de pessoas nesta faixa etária, multiplicado por 100	<i>Atlas de Desenvolvimento Humano</i> , 2010 < <a href="http://www.atlasbrasil.org.br">http://www.atlasbrasil.org.br</a> >

Fonte: Elaboração da autora.

O presente estudo analisa a eficiência educacional dos municípios no nível fundamental. Esse nível foi escolhido porque, apesar das melhorias nas avaliações recentes, a taxa de crescimento parece ser lenta de acordo com os dados apresentados na Tabela 1.

A análise incluiu 5.020 municípios de um grupo de 5.565 que foram considerados. Foram excluídos 545 municípios devido à indisponibilidade de dados sobre os gastos por aluno e/ou informações sobre professores. Aplicando o algoritmo *K-means* do *software Stata* (considerando todas as variáveis do modelo proposto), os 5.020 municípios são divididos em três *clusters*: *cluster 1*, contendo 2.425 municípios; *cluster 2*, com 2.026 municípios; e *cluster 3*, englobando 569 municípios.

Todos os dados foram obtidos para o nível fundamental, com exceção daqueles sobre gasto por aluno e FORM\_SUP\_ESP. A variável despesa por aluno está disponível para o total da unidade federal; portanto, usamos a média ponderada do gasto por cada município de acordo com o seu respectivo número de matrículas, no nível fundamental. A variável FORM\_SUP\_ESP foi utilizada para o nível fundamental II porque os últimos três anos do ensino fundamental são aqueles em que a maioria dos alunos requer professores especializados.

A Tabela 2 traz a estatística descritiva para todas as variáveis que são separadas nos três *clusters*. Notavelmente, se considerarmos todos os municípios juntos, perderemos diferenças importantes, pois os *clusters* apresentam dissimilaridades significativas. Em média, o *cluster* 2 é o grupo com os menores valores de gasto por aluno, FORM\_SUP e FORM\_SUP\_ESP e as condições básicas de qualidade de vida são mais desprovidas; portanto, este *cluster* agrega, claramente, os municípios brasileiros mais desfavorecidos. Os *clusters* 1 e 3 são semelhantes em algumas variáveis, mas o *cluster* 3 tem uma vantagem em todas as variáveis, exceto no acesso a sistemas de esgoto.

Também é importante enfatizar que mesmo os municípios mais favorecidos necessitam de melhorias significativas no que diz respeito à satisfação das necessidades básicas. O percentual de escolas com acesso a sistemas de esgoto no *cluster* 3 varia de 0,1% a 100%, e a média desta variável é inferior a 50% para todos os *clusters*.

A Tabela 2 também apresenta a matriz de correlação entre todas as variáveis, incluindo tanto aquelas diretamente relacionadas à educação (utilizadas no primeiro estágio) quanto as ambientais (utilizadas no segundo estágio).

**TABELA 2**  
**ESTATÍSTICA DESCRITIVA**

VARIÁVEL	MÉDIA	DESV. PAD.	MÍN.	MÁX.	MATRIZ DE CORRELAÇÃO															
					1	2	3	4	5	6	7	8	9	10	11					
<b>Cluster 1</b>																				
1. FORM_SUP	78,6	15,6	2,20	100,0	1,00	0,42	0,44	0,14	0,32	0,43	-0,33	-0,29	0,45	-0,54	0,41					
2. FORM_SUP_ESP	25,6	18,4	0,10	100,0		1,00	0,35	0,05	0,52	0,23	-0,16	-0,22	0,26	-0,28	0,19					
3. Ideb	4,2	0,6	1,90	6,40			1,00	0,11	0,31	0,33	-0,51	-0,31	0,40	-0,46	0,31					
4. Gastos	4.635,6	443,5	3.909,52	5.601,4				1,00	0,02	0,16	-0,20	-0,03	0,17	-0,15	0,14					
5. Acesso a esgoto	41,5	37,8	0,10	100,0					1,00	0,23	-0,13	-0,23	0,33	-0,24	0,11					
6. Educação materna	83,2	8,2	36,74	98,6						1,00	-0,27	-0,27	0,44	-0,38	0,36					
7. Taxa de abandono	3,5	3,0	0,00	24,8							1,00	0,17	-0,40	0,38	-0,35					
8. Índice de Gini	0,5	0,0	0,29	0,76								1,00	-0,05	0,29	-0,28					
9. Renda	586,6	217,0	121,32	1.866,5									1,00	-0,46	0,35					
10. Taxa de analfabetismo	2,0	2,0	0,00	29,5										1,00	-0,50					
11. Defasagem escolar	87,9	6,0	50,87	100,0											1,00					
<b>Cluster 2</b>																				
1. FORM_SUP	55,1	25,5	0,10	100,0	1,00	0,54	0,46	0,43	0,39	0,37	-0,38	-0,36	0,54	-0,47	0,51					
2. FORM_SUP_ESP	11,3	12,9	0,10	90,5		1,00	0,42	0,36	0,51	0,28	-0,29	-0,34	0,50	-0,42	0,39					
3. Ideb	3,5	0,7	1,60	6,40			1,00	0,28	0,39	0,36	-0,65	-0,23	0,54	-0,52	0,37					
4. Gastos	3.184,4	453,1	1.868,85	3.909,2				1,00	0,34	0,34	-0,20	-0,34	0,47	-0,40	0,36					
5. Acesso a esgoto	18,5	26,6	0,10	100,0					1,00	0,32	-0,20	-0,35	0,56	-0,42	0,30					
6. Educação materna	74,6	10,2	22,41	97,2						1,00	-0,28	-0,34	0,49	-0,46	0,38					
7. Taxa de abandono	6,7	4,6	0,00	28,0							1,00	0,11	-0,39	0,43	-0,44					
8. Índice de Gini	0,52	0,06	0,32	0,80								1,00	-0,32	0,43	-0,39					
9. Renda	351,1	178,2	96,25	1.253,1									1,00	-0,63	0,45					
10. Taxa de analfabetismo	5,8	4,0	0,00	31,0										1,00	-0,63					
11. Defasagem escolar	80,5	9,0	34,62	98,31											1,00					
<b>Cluster 3</b>																				
1. FORM_SUP	84,1	12,4	33,3	100,0	1,00	0,31	0,19	-0,03	0,18	0,20	-0,07	-0,18	0,20	-0,31	0,26					
2. FORM_SUP_ESP	25,3	18,8	0,10	100,0		1,00	0,18	0,03	0,44	0,06	-0,09	-0,18	0,02	-0,15	0,06					
3. Ideb	4,3	0,6	2,60	6,20			1,00	-0,01	0,27	0,33	-0,49	-0,27	0,25	-0,27	0,26					
4. Gastos	6.572,7	1.183,2	5.604,3	18.773,4				1,00	0,03	0,11	0,00	-0,08	0,06	-0,05	0,11					
5. Acesso a esgoto	37,8	40,6	0,10	100,0					1,00	0,09	-0,21	-0,22	0,12	-0,09	0,03					
6. Educação materna	87,1	7,3	52,78	100,0						1,00	-0,22	-0,21	0,32	-0,29	0,27					
7. Taxa de abandono	2,3	2,7	0,00	23,4							1,00	0,14	-0,16	0,16	-0,27					
8. Índice de Gini	0,47	0,07	0,28	0,77								1,00	0,19	0,23	-0,29					
9. Renda	678,4	251,0	180,53	2.043,7									1,00	-0,29	0,18					
10. Taxa de analfabetismo	1,4	1,38	0,00	11,91										1,00	-0,38					
11. Defasagem escolar	90,3	6,50	40,22	100,0											1,00					

Fonte: Elaboração da autora.



## RESULTADOS E DISCUSSÃO

A Tabela 3 apresenta o percentual de municípios por região em cada *cluster* e a renda média *per capita*. Por exemplo, o *cluster 1* possui uma proporção significativa de municípios do Norte, Centro-Oeste, Sudeste e Sul, o *cluster 2* inclui mais municípios do Norte e do Nordeste e o *cluster 3* tem principalmente municípios do Centro-Oeste, Sudeste e Sul. Mas, claramente, o *cluster 2* agrega os municípios com menor renda de todas as regiões. É interessante notar que o *cluster 1* engloba não só a maioria dos municípios do Centro-Oeste, Sudeste e Sul, que são as regiões mais ricas do Brasil, mas também 49% dos municípios mais ricos do Norte (uma região desprivilegiada no Brasil). Isso mostra que a metodologia *K-means* agrega os municípios independentemente das regiões.

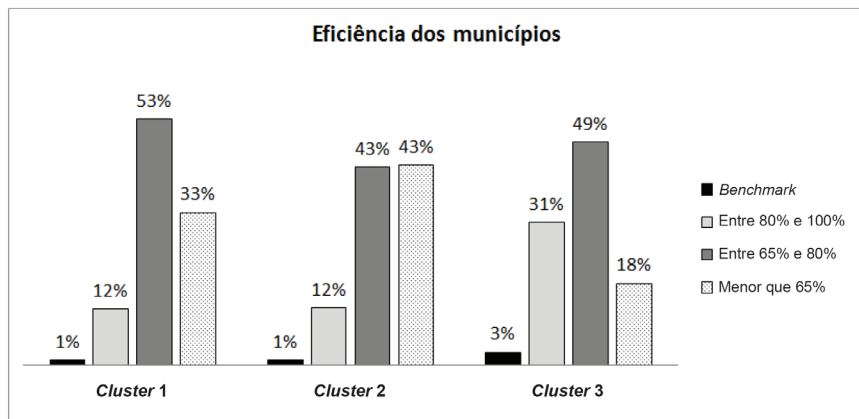
**TABELA 3**  
PORCENTAGEM DE MUNICÍPIOS POR REGIÃO E RENDA MÉDIA *PER CAPITA*,  
SEGUNDO *CLUSTERS*

<i>CLUSTERS</i>	NORTE	NORDESTE	CENTRO- -OESTE	SUDESTE	SUL
<b><i>Cluster 1</i></b>					
Municípios (%)	49,0	12,0	63,0	73,0	59,0
Renda <i>per capita</i> (R\$)	414,4	331,5	591,5	586,8	703,9
<b><i>Cluster 2</i></b>					
Municípios (%)	47,0	88,0	11,0	13,0	21,0
Renda <i>per capita</i> (R\$)	302,4	271,1	555,5	\$543,4	652,6
<b><i>Cluster 3</i></b>					
Municípios (%)	4,0	1,0	26,0	14,0	20,0
Renda <i>per capita</i> (R\$)	396,3	418,3	577,2	669,3	771,7

Fonte: Elaboração da autora.

O modelo DEA foi aplicado três vezes, uma para cada *cluster*, e a Figura 1 mostra os percentuais de municípios por escore de eficiência, para cada *cluster*. Por exemplo, no *cluster 1*, 1% dos municípios são *benchmarks* (com uma pontuação de eficiência de 100%), 12% têm uma pontuação de eficiência entre 80% e <100%, 53% estão entre 65% e <80% e 33% têm uma pontuação inferior a 65%. Claramente, o *cluster 2* agrega os municípios menos eficientes e inclui 43% daqueles com escore de eficiência inferior a 65%, enquanto o *cluster 3* tem o maior percentual de municípios eficientes, englobando 31% daqueles com escore de eficiência entre 80% e 100%.

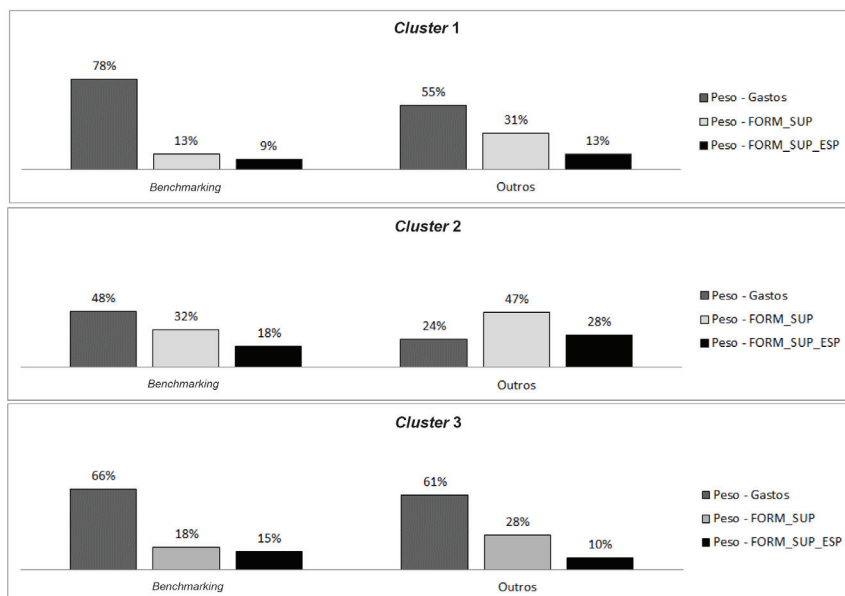
**FIGURA 1**  
**PERCENTUAL DE MUNICÍPIOS POR ESCORE DE EFICIÊNCIA**



Fonte: Elaboração da autora.

A Figura 2 apresenta os pesos que o modelo DEA gera para cada variável de *input*. Quanto maior seu peso, mais elevada é a contribuição da variável para a eficiência educacional. Esta variável é assim indicada como determinante da eficiência. Analisamos os pesos considerando dois grupos. O primeiro inclui os *benchmarks* (os municípios que são considerados *benchmarks* para 20 ou mais municípios) e o segundo agrega os demais municípios.

**FIGURA 2**  
**CONTRIBUIÇÕES DE INPUT (PESOS EM PORCENTAGEM) DO MODELO DEA**

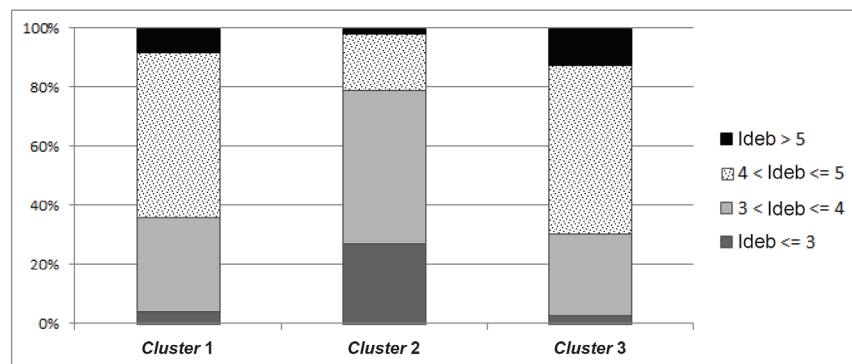


Fonte: Elaboração da autora.

Observa-se que, para os grupos de *benchmarking*, a variável relacionada aos gastos permanece com maior peso, principalmente para os *clusters* 1 e 3. No entanto, para o grupo “outros”, é clara a diferença entre os *clusters*. Assim como em Holmlund, McNally e Viarengo (2010), também encontramos evidências de heterogeneidade nos efeitos dos gastos em resultados educacionais, com efeitos mais fortes sobre o *cluster* economicamente mais desfavorecido. Para os *clusters* 1 e 3, a variável de gastos *per capita* apresenta o maior peso; ao contrário, para o *cluster* 2, que agrega os municípios mais desfavorecidos, a variável de gastos *per capita* recebe o menor peso do modelo DEA, ou seja, esta variável possui espaço significativo para melhoria neste *cluster*. Como o Ideb do *cluster* 2, em geral, é muito baixo (ver Figura 3), esses resultados sugerem que alguns municípios parecem gastar muito mais, mas obtêm o mesmo resultado baixo que os municípios que gastam menos. Claramente, estes resultados sugerem que: para os municípios menos ricos, outros fatores além do gasto *per capita* influenciam mais o Ideb; o gasto *per capita* não é o determinante de eficiência para o *cluster* menos favorecido; e para os municípios mais ricos, o gasto *per capita* é considerado um determinante eficiente da educação.

Vale ressaltar que a relação entre gasto e desempenho na educação é debatida desde o caso bem conhecido do trabalho de Hanushek (1986), que não encontrou uma forte relação entre gastos escolares e desempenho dos alunos. Anos mais tarde, Hedges, Laine e Greenwald (1994) reanalisaram os estudos de Hanushek e encontraram o oposto, ou seja, uma relação positiva entre os dólares gastos na educação e o desempenho dos alunos. Mas, como sugerido pelo Departamento de Educação do Reino Unido (UNITED KINGDOM, 2014), uma razão para os diferentes resultados na literatura pode ser devido às desigualdades consideradas, como no estudo de Holmlund, McNally e Viarengo (2010), em que o tamanho do efeito foi maior para crianças mais desfavorecidas.

**FIGURA 3**  
PERCENTUAL DOS MUNICÍPIOS, SEGUNDO RESULTADOS IDEB (IDEB VARIA DE 0 A 10)



Fonte: Elaboração da autora.

Além disso, a Figura 2 mostra que a variável FORM\_SUP\_ESP oferece os pesos mais baixos para os *clusters* 1 e 3. Esse dado sugere um possível potencial para melhorar os resultados dos municípios, aumentando o número de professores com graus de ensino superior nas áreas em que ensinam (mesmo para o *cluster* 2).

Finalmente, utilizando os escores de eficiência previamente calculados pelo modelo DEA, avaliamos agora a importância dos *inputs* não discricionários. A Tabela 4 apresenta os resultados da regressão Tobit para três especificações alternativas das equações (7)-(10), conforme mostra o Quadro 3, que apresenta os coeficientes estimados para cada um dos três modelos.

**QUADRO 3**  
VARIÁVEIS NÃO DISCRICIONÁRIAS UTILIZADAS NOS TRÊS MODELOS

VARIÁVEIS	MODELO 1	MODELO 2	MODELO 3
Acesso ao sistema de esgoto (%)	X	X	X
Educação materna (%)	X	X	X
Taxa de abandono (%)	X	X	X
Índice de Gini	X		X
Taxa de analfabetismo (%)		X	X
Defasagem escolar (%)		X	X

Fonte: Elaboração da autora.

**TABELA 4**  
RESULTADOS TOBIT

VARIÁVEIS	MODELO 1			MODELO 2			MODELO 3		
	CLUSTER 1	CLUSTER 2	CLUSTER 3	CLUSTER 1	CLUSTER 2	CLUSTER 3	CLUSTER 1	CLUSTER 2	CLUSTER 3
Acesso a esgoto (%)	.024*** (.0057)	.011 (.0105)	.013 (.0110)	.031*** (.0057)	.005 (.0101)	.021* (.1086)	.023*** (.0057)	.008 (.0103)	.013 (.0110)
Educação materna (%)	.059* (.0281)	.059* (.0264)	.165* (.0646)	.110*** (.0284)	.056* (.0258)	.201** (.0647)	.075** (.0286)	.0621* (.0261)	.160* (.0654)
Taxa de abandono (%)	-1.030*** (.0722)	-1.185*** (.0535)	-1.592*** (.1623)	-1.11*** (.0748)	-1.26*** (.0554)	-1.60*** (.1672)	-1.07*** (.0742)	-1.26*** (.0560)	-1.59*** (.1658)
Índice de Gini	-21.619*** (3.4274)	8.551 (4.181)	-23.05*** (6.470)				-24.14*** (3.5506)	6.325 (4.3285)	-22.00** (6.8609)
Renda	.0001 (.00011)	.005** (.00017)	.004* (.0018)	.0004 (.0011)	.003 (.0018)	.0018 (.00018)	.0023 (.0011)	.0035 (.0018)	.0041* (.0019)
Taxa de analfabetismo (%)				-.144 (.1275)	-.568*** (.0833)	-.4902 (.3505)	-.010 (.1278)	-.592*** (.0849)	-.289 (.3530)
Defasagem escolar (%)				-.084* (.0404)	-.313*** (.0330)	.042 (.0740)	-.128** (.0405)	-.305*** (.0335)	-.0135 (.0754)
Constante	76.61*** (3.072)	64.01*** (3.321)	72.15*** (6.802)	70.19*** (4.02)	98.72*** (3.464)	56.59*** (8.342)	87.67*** (4.743)	94.32*** (4.587)	74.07*** (9.903)
/sigma	9.87 (.143)	10.24 (.162)	10.10 (.306)	9.95 (.144)	10.01 (.158)	10.18 (.309)	9.85 (.142)	10.01 (.158)	10.09 (.306)
LR chi2(6)	424.14	635.41	161.51	389.22	728.14	151.98	435.01	730.27	162.18
Prob > chi2	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
N	2,425	2,025	569	2,425	2,025	569	2,425	2,025	569

Nota: Erro padrão entre parênteses. \* significante a 5%; \*\* significante a 1%; \*\*\* significante a 0,1%.

Fonte: Elaboração da autora.

Independentemente de *clusters* e modelos, a Tabela 4 mostra que os coeficientes de escolaridade materna e as taxas de abandono escolar são estatisticamente significativos, ou seja, um aumento na educação materna eleva a eficiência educacional e um crescimento na taxa de abandono escolar reduz a eficiência educacional. Estes resultados são esperados; muitos estudos podem ser encontrados que vinculam as taxas de abandono escolar a retrocessos educacionais e, em geral, os antecedentes socioeconômicos e a etnia são citados como fatores de abandono (RUMBERGER, 1995; GOLDSCHMIDT; WANG, 1999). No entanto, em geral, a renda não é estatisticamente significativa para todos os *clusters*.

Em contraste, a Tabela 4 mostra diferenças significativas nos efeitos das outras variáveis socioeconômicas nos três grupos. Para o grupo mais rico (*cluster* 3), o índice de Gini é estatisticamente significativo, mas as variáveis relacionadas aos fatores básicos da educação, como a taxa de analfabetismo e a defasagem escolar, não são estatisticamente significativas. Já para o *cluster* 2 (o mais pobre), a taxa de analfabetismo e a defasagem escolar são estatisticamente significativas, mas o índice de Gini não é. Ou seja, considerando os dois extremos – o *cluster* mais rico e o mais pobre –, o indicador de desenvolvimento social (Gini) é estatisticamente significativo para o rico e os fatores básicos de uma educação satisfatória (taxa de analfabetismo e defasagem escolar) são estatisticamente significativos para o mais pobre. Esses resultados sugerem que, para o *cluster* mais pobre, a principal preocupação são as políticas públicas que reduzam a taxa de analfabetismo e a defasagem escolar; aqui o mais importante é como o município está gastando o dinheiro, e não o quanto é gasto. Nos municípios mais pobres, a falta de uma política pública assertiva é muito mais sentida. Além disso, tal resultado corrobora aqueles obtidos no primeiro passo, que indicam que os municípios do *cluster* 2 necessitam de uma melhor gestão dos gastos *per capita*.

## CONCLUSÕES

Evidências fortes confirmam a importância em analisar a eficiência educacional categorizando os municípios em *clusters*, quando se considera um país com desenvolvimento econômico promissor, mas com alta desigualdade socioeconômica. Os resultados indicam claramente que, se considerarmos o Brasil como um todo, apenas obteremos uma classificação educacional geral, sendo que muitos indicadores-chave, como os fatores básicos da educação, provavelmente serão negligenciados.

Além da classificação de municípios brasileiros por *clusters*, nesse artigo propomos um modelo educacional em dois estágios. Primeiramente, o estudo avalia a eficiência educacional de 5.020 municípios brasileiros, evidenciando a heterogeneidade dos efeitos dos gastos em avaliações educacionais, semelhante a Holmlund, McNally e

Viarengo (2010), e que a qualidade dos professores merece mais atenção em todo o país. O segundo estágio utiliza a regressão Tobit para estimar a influência das variáveis não discricionárias na eficiência educacional. Os principais resultados sugerem que: o rendimento não tem influência significativa no resultado; a educação materna e as taxas de abandono têm grande influência sobre todos os grupos; e, para os *clusters* mais ricos, o índice de Gini tem grande influência na educação, enquanto as variáveis relacionadas com a evolução dos alunos na escola (taxa de analfabetismo e defasagem escolar) não têm influência significativa, comportamento oposto ao verificado para o *cluster* menos favorecido (mais pobre).

Esse último resultado pode ser utilizado para comparar países desenvolvidos com aqueles emergentes. Para os primeiros, o foco nas variáveis relacionadas aos fatores básicos de uma educação próspera, como a defasagem escolar e a taxa de analfabetismo, pode não fazer sentido porque as políticas públicas já existem e funcionam bem, sendo que um possível problema para esses países pode estar relacionado à desigualdade de renda (índice de Gini). No entanto, para os países em desenvolvimento e pobres, as variáveis referentes à forma de gerir os gastos e conceber políticas públicas adequadas merecem muito mais atenção. Esses resultados corroboram, por exemplo, aqueles encontrados por Commander, Davoodi e Lee (1997), que mostram que os países menos favorecidos obtêm resultados piores em suas políticas de prevenção da mortalidade infantil e em suas estatísticas de expectativa de vida. Esses resultados também corroboram os de Miranda e Mendes (2004), que indicam que essa mesma situação pode estar ocorrendo nos municípios brasileiros: aqueles com pior distribuição de renda são menos capazes de implementar políticas sociais efetivas.

## REFERÊNCIAS

- AFONSO, A.; AUBYN, M. St. Cross-country efficiency of secondary education provision: a semi-parametric analysis with non-discretionary inputs. *Economic Modelling*, v. 23, n. 3, p. 476-491, 2006.
- AGASISTI, T. The efficiency of Italian secondary schools and the potential role of competition: a data envelopment analysis using OECD-PISA 2006 data. *Education Economics*, v. 21, n. 5, p. 520-544, 2013. Disponível em: <[http://www.oecd.org/edu/Brazil\\_EAG2013%20Country%20Note.pdf](http://www.oecd.org/edu/Brazil_EAG2013%20Country%20Note.pdf)>. Acesso em: 03 maio 2017.
- AGENCY FOR INTERNATIONAL DEVELOPMENT IN COOPERATION. Health Systems 20/20. *The health system assessment approach: a how-to manual*, 2012. Disponível em: <<http://apps.who.int/medicinedocs/documents/s19838en/s19838en.pdf>>. Acesso em: 03 maio 2017.
- BRADLEY, S.; JOHNES, G.; MILLINGTON, J. The effect of competition on the efficiency of secondary schools in England. *European Journal of Operational Research*, Amsterdã, v. 135, n. 3, p. 545-568, 2001.
- BRYK, A. S.; THUM, Y. M. The effects of high school organization on dropping out: an exploratory investigation. *American Educational Research Journal*, Thousand Oaks, CA, v. 26, n. 3, p. 353-383, 1989.

- CHARNES, A.; COOPER, W. W.; RHODES, E. Measuring the efficiency of decision-making units. *European Journal of Operational Research*, Amsterdã, v. 2, n. 6, p. 429-442, 1978.
- CHUDGAR, A.; LUSCHEI, T. National income, income inequality and the importance of schools: a hierarchical cross-national comparison. *American Educational Research Journal*, Thousand Oaks, CA, v. 46, n. 3, p. 626-58, 2009.
- CHUDGAR, A.; SHAFIK, M. N. Family, community, and educational outcomes in South Asia. *Prospects*, Manchester, UK, v. 40, n. 4, p. 517-534, 2010.
- COELLI, T. J.; RAO, D. S. P.; O'DONNELL, C. J.; BATTESSE, G. E. *An introduction to efficiency and productivity analysis*. 2. ed. New York: Springer Science, 2005.
- COMMANDER, S.; DAVOODI, H. R.; LEE, U. J. *The causes of government and the consequences for growth and well-being*. Washington, DC: World Bank, 1997. (Policy, Research working paper, n. WPS 1785). Disponível em: <<http://documents.worldbank.org/curated/en/375251468766769196/The-causes-of-government-and-the-consequences-for-growth-and-well-being>>. Acesso em: 03 maio 2017.
- DRINEAS, P.; KANNAN, R.; MAHONEY, M. W. Fast Monte Carlo algorithms for matrices II: computing a low-rank approximation to a matrix. *SIAM Journal on Computing*, Philadelphia, PA, v. 36, n. 1, p. 158-183, 2006.
- DYSON, R. G.; ALLEN, R.; CAMANHO, A. S.; PODINOVSKI, V. V.; SARRICO, C. S.; SHALE, E. A. Pitfalls and protocols in DEA. *European Journal of Operational Research*, Amsterdã, v. 132, n. 2, p. 245-259, 2001.
- FERNANDES, R. *Índice de desenvolvimento da educação básica (Ideb): metas intermediárias para a sua trajetória no Brasil, estados, municípios e escolas*. Brasília, DF: MEC/Inep, 2007. Disponível em: <[http://download.inep.gov.br/educacao\\_basica/porta\\_ideb/o\\_que\\_sao\\_as\\_metas/Artigo\\_projecoes.pdf](http://download.inep.gov.br/educacao_basica/porta_ideb/o_que_sao_as_metas/Artigo_projecoes.pdf)>. Acesso em: 03 maio 2017.
- FUNDO DAS NAÇÕES UNIDAS PARA A INFÂNCIA – UNICEF. *Situação da infância e da adolescência brasileira em 2009*. O direito de aprender: potencializar avanços e reduzir desigualdades. Brasília, DF: Unicef, 2009. Disponível em: <[http://www.unicef.org/brazil/pt/siab\\_capitulos.pdf](http://www.unicef.org/brazil/pt/siab_capitulos.pdf)>. Acesso em: 03 maio 2017.
- GOLDSCHMIDT, P.; WANG, J. When can schools affect dropout behavior? A longitudinal multilevel analysis. *American Educational Research Journal*, Thousand Oaks, CA, v. 36, n. 4, p. 715-738, 1999.
- HANUSHEK, E. A. The economics of schooling: production and efficiency in public schools. *Journal of Economic Literature*, Washington, D.C., v. 24, n. 3, p. 1141-1177, 1986.
- HANUSHEK, E. A.; LUQUE, J. A. Efficiency and equity in schools around the world. *Economics of Education Review*, Denver, v. 22, n. 5, p. 481-502, 2003.
- HAUNER, D. Explaining differences in public sector efficiency: evidence from Russia's regions. *World Development*, Michigan, v. 36, n. 10, p. 1745-1765, 2008.
- HEDGES, L. V.; LAINE, R.; GREENWALD, R. Does money matter? A meta-analysis of studies of the effect of different school inputs on student outcomes. *Educational Researcher*, Thousand Oaks, CA, v. 23, n. 3, p. 5-14, 1994.
- HENRIQUEZ, F.; LARA, B.; MIZALA, A.; REPETTO, A. Effective schools do exist: low-income children's academic performance in Chile. *Applied Economics Letters*, Londres, v. 19, n. 5, p. 445-451, 2012.
- HOLMLUND, H.; MCNALLY, S.; VIARENGO, M. Does money matter for schools? *Economics of Education Review*, Denver, v. 29, n. 6, p. 1154-1164, 2010.
- JAIN, A. K. Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, v. 31, n. 8, p. 651-666, 2010.
- JAIN, A. K.; DUBES, R. *Algorithms for clustering data*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
- KIRJAVAINEN, T.; LOIKKANEN, H. A. Efficiency differences of Finnish senior secondary schools: an application of DEA and Tobit analysis. *Economics of Education Review*, Denver, v. 17, n. 4, p. 377-394, 1998.



- LEE, V. E.; BURKAM, D. T. Dropping out of high school: the role of school organization and structure. *American Educational Research Journal*, Thousand Oaks, CA, v. 40, n. 2, p. 353-393, 2003.
- LIU, J. S.; LU, L. Y. Y.; LU, W. M.; LIN, B. J. Y. A survey of DEA applications. *Omega*, v. 41, n. 5, p. 893-902, 2013.
- MANCIBON, M. J.; BANDRÉS, E. Efficiency evaluation in secondary schools: the key role of model specification and of ex post analysis of results. *Education Economics*, Londres, v. 7, n. 2, p. 131-52, 1999.
- MIRANDA, R. N.; MENDES, M. *Municípios em extrema pobreza: só dinheiro não resolve*. Brasília: Consultoria Legislativa do Senado Federal, 2004. (Textos para discussão, n. 15). Disponível em: <<https://www12.senado.leg.br/publicacoes/estudos-legislativos/tipos-de-estudos/textos-para-discussao/tl-15-municipios-em-extrema-pobreza-so-dinheiro-nao-resolve>>. Acesso em: 03 maio 2017.
- MURILLO, F. J.; ROMÁN, M. School infrastructure and resources do matter: analysis of the incidence of school resources on the performance of Latin American students. *School Effectiveness and School Improvement*, Londres, v. 22, n. 1, p. 29-50, 2011.
- NEAL, D. The effects of Catholic secondary schooling on educational achievement. *Journal of Labor Economics*, Chicago, v. 15, n. 1, p. 98-123, 1997.
- ORGANIZAÇÃO PARA COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO – OECD. Education at a glance, 2012. Disponível em: <[https://www.oecd.org/edu/EAG%202012\\_e-book\\_EN\\_200912.pdf](https://www.oecd.org/edu/EAG%202012_e-book_EN_200912.pdf)>. Acesso em: 03 maio 2017.
- ORGANIZAÇÃO PARA COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO – OECD. Education at a glance, 2013. Disponível em: <[http://www.oecd-ilibrary.org/education/education-at-a-glance-2013\\_eag-2013-en](http://www.oecd-ilibrary.org/education/education-at-a-glance-2013_eag-2013-en)>. Acesso em: 03 maio 2017.
- PRIMONT, D. F.; DOMAZLICKY, B. Student achievement and efficiency in Missouri schools and the No Child Left Behind Act. *Economics of Education Review*, Denver, v. 25, n. 1, p. 77-90, 2006.
- RASSOULI-CURRIER, S. Assessing the efficiency of Oklahoma public schools: a data envelopment analysis. *Southwestern Economic Review*, Texas, v. 34, n. 1, p. 131-144, 2007.
- RAY, S. C. Resource use efficiency in public schools: a study of Connecticut data. *Management Science*, Chicago, v. 37, n. 12, p. 1620-1628, 1991.
- RODRÍGUEZ-POSE, A.; TSELIOS, V. Mapping the European regional educational distribution. *European Urban and Regional Studies*, Pennsylvania, v. 18, n. 4, p. 358-374, 2011. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1014.1569&rep=rep1&type=pdf>>. Acesso em: 03 maio 2017.
- RUMBERGER, R. W. Dropping out of middle school: a multilevel analysis of students and schools. *American Educational Research Journal*, Thousand Oaks, CA, v. 32, n. 3, p. 583-625, 1995.
- RUMBERGER, R. W.; PALARDY, G. J. High school performance: test scores, dropout rates, and transfer rates as alternative indicators of high school performance. *American Educational Research Journal*, Thousand Oaks, CA, v. 42, n. 3, p. 3-42, 2005.
- RUMBERGER, R. W.; THOMAS, S. L. The distribution of dropout and turnover rates among urban and suburban high schools. *Sociology of Education*, v. 73, n. 1, p. 39-67, 2000.
- SHAH, A. *A practitioner's guide to intergovernmental fiscal transfers*. Washington, DC: World Bank, 2006. (World Bank Policy Research Working Paper, 4039).
- SIMAR, L.; WILSON, P. W. Performance of the bootstrap for DEA estimators and iterating the principle. In: COOPER, W.W.; SEIFORD, L. M.; ZHU, J. (Ed.). *Handbook on Data Envelopment Analysis*. Boston: Kluwer Academic, 2004. p. 265-298.
- TODOS PELA EDUCAÇÃO. *Anuário Brasileiro da Educação Básica*. São Paulo: Moderna, 2012. Disponível em: <<http://www.todospelaeducacao.org.br/biblioteca/1450/anuario-brasileiro-da-educacao-basica>>. Acesso em: 03 maio 2017.
- UNITED KINGDOM. Parliament. Department for Education. *What impact does school spending have on pupil attainment? A review of the recent literature*. Londres: Strategic Analysis and Research Division & Infrastructure Funding and Longitudinal Analysis Division, 2014. Disponível em:



<<http://www.parliament.uk/documents/commons-committees/Education/Impact-of-school-spending-on-pupil-attainment.pdf>>. Acesso em: 03 maio 2017.

ULYSSEA, G.; FERNANDES, R.; GREMAUD, A. P. O impacto do Fundef na alocação de recursos para a educação básica. *Pesquisa e Planejamento Econômico*, Brasília, DF, v. 36, n. 1, p. 109-136, 2006.

UNITED NATIONS. *Diverging growth and development*. New York: World Economic and Social Survey, 2006.

YANG, Z. Cross system bank branch evaluation using clustering and data envelopment analysis. In: HUANG, D. S. et al. (Ed.). *Advanced intelligent computing theories and applications*. Berlin, Heidelberg: Springer, 2010. p. 238-245. (Lecture Notes in Computer Science, v. 6215)

MARIA CRISTINA GRAMANI  
Insper, São Paulo, São Paulo, Brasil  
[mariacng@insper.edu.br](mailto:mariacng@insper.edu.br)