

## AVALIAÇÃO DE EFICIÊNCIA DAS ESCOLAS PÚBLICAS DE ENSINO MÉDIO EM GOIÁS: UMA ANÁLISE DE DOIS ESTÁGIOS

GUILHERME RESENDE OLIVEIRA \*

ALEX FELIPE RODRIGUES LIMA †

SÉRGIO BORGES FONSECA JÚNIOR ‡

THIAGO MENDES ROSA §

### Resumo

O estudo verifica o nível de eficiência das escolas estaduais públicas goianas, por meio do desempenho na prova do Exame Nacional do Ensino Médio (ENEM) e o custo médio por aluno. Em um segundo momento, determina-se os condicionantes das eficiências. Para a avaliação de eficiência, utilizou-se o método não paramétrico de Análise Envoltória de Dados (DEA) e para o segundo estágio, a regressão beta inflacionada, que acomoda o intervalo dos resultados obtidos na primeira etapa. As evidências empíricas mostram que as escolas militares e mais especializadas tendem a ter desempenho superior às demais, enquanto as escolas com ensino em tempo integral, e com maior média de hora aula, tendem a ser mais ineficientes.

**Palavras-chave:** Eficiência; Escolas Estaduais; DEA; Regressão Beta Inflacionada.

### Abstract

The study verifies the level of efficiency of state public schools in Goiás, through the performance of national high school exam (ENEM) and the average cost per student. In a second stage, the conditions of efficiency are determined. For the evaluation of efficiency the Data Envelopment Analysis (DEA), a nonparametric method, is used and at second stage, the inflated beta regression is used, which accommodates the interval of results from the first step. The empirical evidence shows that military and more specialized schools tend to outperform others, while the schools with a full-time teaching and higher average class time, tend to be inefficient.

**Keywords:** Efficiency; State Schools; DEA; Beta Regression.

**JEL classification:** C52, I21.

**DOI:** <http://dx.doi.org/10.11606/1413-8050/ea118947>

\* Doutor em Economia pela UnB. Gerente de Inteligência Territorial da Agência Goiana de Assistência Técnica, Extensão Rural e Pesquisa Agropecuária (Emater). E-mail: resendego@yahoo.com.br

† Mestre em Estatística pela Universidade de Brasília (UnB). Pesquisador do Instituto Mauro Borges de Estatísticas e Estudos Socioeconômicos (IMB). E-mail: afelipe\_7@hotmail.com

‡ Mestre em Economia pela Universidade Federal de Uberlândia (UFU). Pesquisador em economia do IMB. E-mail: sergio-borges@msn.com

§ Mestre em Desenvolvimento Econômico pela Universidade Federal do Paraná (UFPR). Pesquisador da Companhia de Planejamento do Distrito Federal (Codeplan). E-mail: thiagomendes-rosa@outlook.com

## 1 Introdução

A afirmação de que a educação é a propulsora do desenvolvimento econômico é quase consenso nas últimas décadas. Embora haja profundas discussões sobre os vários determinantes do desenvolvimento econômico, acredita-se que a educação seja uma das peças chave nessa discussão. Nessa direção, Benegas et al. (2012) afirmam que o crescimento e desenvolvimento de longo prazo de uma nação passam inevitavelmente pela oferta abrangente e de qualidade da educação. Para Barros et al. (2011), a diferença de capital humano<sup>1</sup> é um dos fatores mais determinantes da desigualdade regional, pois uma parte do diferencial de renda tem relação com a diferença educacional.

De modo geral, a sociedade acredita que o setor público, ao executar a sua função alocativa, está gastando de maneira pouco parcimoniosa e ineficientemente. Com a educação, a situação não é diferente, são recorrentes as críticas quanto à qualidade do seu gasto público. De acordo com o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), nos anos 2000, os dispêndios com educação no Brasil passaram de 3,5% do PIB para um valor superior a 5%. Contudo, o aumento de recursos não foi acompanhado por uma melhoria qualitativa significativa da educação. É justamente na relação entre os gastos públicos e o desempenho educacional em Goiás que esse artigo está inserido.

De antemão um dado chama a atenção: o Censo Demográfico de 2010 (IBGE 2011) mostra que nos municípios do estado de Goiás, há elevada amplitude na taxa de alfabetização, que se situa entre 96% e 76%, com valor médio de 90%. No que diz respeito à qualificação, Oliveira (2016) mostra que um dos principais gargalos para o desenvolvimento do setor industrial goiano é a falta de mão de obra qualificada. Ainda, segundo esse autor, dependendo da atividade econômica e da região do estado, este é o principal entrave à atração de novos investimentos. Isso contribui ainda mais para a perpetuação da desigualdade regional. Nesse sentido, para reduzir a desigualdade regional de renda, é condição necessária, mas não suficiente, amenizar as divergências educacionais. Ademais, para resolver essa questão é necessário entender a base educacional da futura mão de obra qualificada.

O governo estadual é responsável por assegurar a educação no nível fundamental, usualmente provendo a etapa final, e oferecer, com prioridade, o ensino médio. De fato, estudos como Gonçalves & França (2013) e Zoghbi et al. (2009) revelam que nem sempre escolas com maior gasto por alunos obtêm melhor desempenho, seja na Prova Brasil, no Índice de Desenvolvimento da Educação Básica (IDEB), no Sistema de Avaliação do Ensino Básico (SAEB) ou no Exame Nacional do Ensino Médio (ENEM).

Salienta-se a diferença entre os termos eficácia e eficiência, pois o primeiro termo se refere diretamente ao desempenho e o segundo, ao desempenho relativamente ao uso de recursos. Tratando-se, especificamente neste artigo, da educação no nível médio, não necessariamente uma escola com boa nota no ENEM é eficiente, pois se ela receber recursos demasiadamente é possível que seus alunos tenham um elevado desempenho, mas não proporcionalmente à quantidade de recursos despendidos. Além disso, é possível que os alu-

---

<sup>1</sup>O capital humano pode ser entendido como um conjunto de capacidades, conhecimentos, competências, sejam estas cognitivas ou não, que favorecem o desempenho de atividades, de valor econômico ou não. Como a educação é um dos meios pelos quais o capital humano pode ser adquirido, ampliado e/ou desenvolvido, tem-se a relação entre capital humano e educação.

nos dessa escola possuam de antemão melhores indicadores socioeconômicos, educação dos pais, infraestrutura, entre outros fatores. Logo, é possível que a escola mais eficiente seja aquela que possua uma nota média com o uso de poucos recursos. Ou seja, a dimensão da eficiência a ser trabalhada neste artigo tem relação com a alocação ótima de recursos, portanto, faz uma conexão entre a magnitude de recursos utilizados e o desempenho na prova do ENEM.

Nesse sentido, eficiência é o uso racional dos meios dos quais se dispõe para alcançar um objetivo previamente determinado. Trata-se da capacidade de alcançar os objetivos e as metas programadas com o mínimo de recursos disponíveis e tempo, conseguindo, dessa forma, a sua otimização.

Há importantes distinções na forma de avaliar a eficiência. Os chamados métodos paramétricos supõem uma relação funcional pré-definida entre os recursos e o que foi produzido. Normalmente, usam médias para determinar o que poderia ter sido produzido. Por outro lado, os métodos não paramétricos não fazem apelo à função de produção e consideram que o máximo que poderia ter sido produzido é dado por meio da observação das unidades mais produtivas. Nesse artigo, considerou-se a técnica não paramétrica de Análise Envoltória de Dados (DEA). É importante salientar que a análise de eficiência cresceu no período recente, segundo Liu et al. (2013), mais de 5.000 artigos foram publicados nessa vertente. Além disso, a literatura nacional apresenta uma série de estudos sobre a análise de eficiência que utilizam essa mesma técnica.

São vários os estudos sobre a eficiência na educação (Faria et al. 2008, Zoghbi et al. 2009, Raposo et al. 2011, Benegas et al. 2012, Gonçalves & França 2013, Carvalho & Sousa 2014, Alves et al. 2014, Junior & Novi 2014, Salgado Junior & Novi 2015, Junior et al. 2015). Rosano-Peña et al. (2012) investigaram a eficiência dos gastos educacionais nos municípios goianos entre 2005 e 2009 por meio da análise envoltória de dados (DEA) com o uso do método de fronteira invertida, além da análise espacial. Eles afirmam que o nível de ineficiência global foi maior que 60%, indicando que apenas 20 municípios são eficientes. Três causas explicam esse resultado: ineficiência de escala, impacto do Entorno do DF (uma das microrregiões mais pobres do estado) e ineficiência de gestão.

Este trabalho apresenta algumas inovações, no que diz respeito aos estudos de eficiência, que são analisar as informações no maior nível de desagregação disponível e utilizar as informações de custo escolar, ambas inéditas na literatura. Os estudos que analisam a eficiência educacional, geralmente, estão desagregados em nível municipal, em razão da insuficiência de informações para análise mais granulares, como por aluno, turma ou escola. Logo, esta pesquisa contribui com a literatura ao avaliar a eficiência tendo como unidade de investigação as escolas públicas de Goiás que ofertam o ensino médio.

Justifica-se uma análise em nível microeconômico – no sentido de abordar as escolas como unidades individuais – na eficiência educacional, pois se sabe a priori que, embora existam diretrizes gerais para a educação, as ações individuais implementadas em determinadas escolas são responsáveis por desempenhos satisfatórios. Nesse sentido, uma escola eficiente pode servir de referência para outra que não é, de modo que boas práticas possam ser replicadas, especialmente nos casos com visível ineficiência. Carvalho & Sousa (2014) mostram que mesmo descontando os fatores ambientais e aleatórios (não controlados pelos gestores públicos), ainda persiste a necessidade de melhorias ligadas à gestão de cada uma das escolas. Uma vez determinada a

eficiência das escolas, será possível verificar quais são as variáveis socioeconômicas e institucionais (escola em tempo integral e escola militar exemplificam essa categoria) que explicam a eficiência por meio de uma regressão beta inflacionada, a qual é capaz de acomodar o intervalo de eficiência estimado por meio da DEA e verificar quais variáveis podem gerar máxima eficiência.

Nesse sentido, o segundo estágio investiga quais são os fatores que influenciam as diferenças nos escores de eficiência, devido aos contextos diferentes em que as escolas atuam. Por exemplo, duas escolas podem ter a mesma gestão de recursos, gastando exatamente a mesma quantia por aluno, todavia se uma delas estiver inserida em um contexto de realidade social mais privilegiada que a outra, a escola em melhor contexto social possivelmente apresentará melhores resultados, mesmo que o esforço gerencial de ambas as administrações seja o mesmo.

## 2 Dados e estratégia empírica

Os dados utilizados neste estudo foram obtidos a partir da compatibilização de duas bases de dados distintas. De um lado, uma das bases refere-se às características estruturais e institucionais das escolas públicas em Goiás que ofertam ensino médio. Esses dados foram obtidos junto à Secretaria de Estado de Educação, Cultura e Esporte do Estado de Goiás (SEDUCE-GO)<sup>2</sup>.

De outro lado, foram utilizados microdados referentes ao ENEM, obtidos junto ao INEP. Ambas as informações se referem ao ano de 2014, por serem os dados mais atualizados disponibilizados pela SEDUCE-GO. Foram selecionadas apenas escolas públicas de Goiás cujos alunos participaram do ENEM no ano de 2014, o que totalizou 356 escolas, localizadas no perímetro urbano e rural em 183 diferentes municípios<sup>3</sup>. Andrade (2011) ressalta que a grande dificuldade de um ranking baseado no produto é a escolha do indicador apropriado. E em relação ao Ensino Médio, existe somente um ranking das escolas, que é construído a partir dos resultados dos alunos do ENEM. Assim, existe certo consenso no uso dos resultados nos testes de proficiência para ranquear as escolas e, portanto, o problema da escolha do produto a ser analisado é mitigado.

Para analisar a eficiência dessas escolas, utilizou-se a metodologia de Análise Envoltória de Dados (DEA), que permite encontrar unidades de referência para a amostra verificada. Além disso, por meio do DEA, é possível avaliar múltiplos produtos como decorrência de vários insumos e sem a necessidade de estabelecer hipóteses a priori sobre a forma funcional, diferentemente das regressões econométricas (Boueri et al. 2015). O primeiro estágio é o mesmo utilizado por Junior & Novi (2014), que propõem uma metodologia para identificar os fatores que podem influenciar o desempenho de alunos.

Todavia, ao se utilizar esse método, é imprescindível ter cuidado com observações extremas e casos de outliers. Pois, se esses forem desconsiderados, corre-se o risco de elevar a fronteira de eficiência, tornando as demais observações aparentemente menos eficientes. Nesse sentido, foi realizada uma aná-

---

<sup>2</sup>Mediante acordo de sigilo quanto ao nome das escolas.

<sup>3</sup>O Estado de Goiás possui 246 municípios, portanto nem todos tiveram escola pública com aluno participante do ENEM no ano de 2014.

lise das estatísticas descritivas para remover possíveis *outliers*<sup>4</sup> Cribari-Neto & Pereira (2013).

O primeiro estudo sobre eficiência é de Farrell (1957), que desenvolveu um índice de eficiência técnica para a agricultura norte-americana. Seu método aprofundou o conceito de eficiência e permitiu encontrar uma unidade de referência (benchmarking). Nesse sentido, a ineficiência é entendida como a distância da produção de uma determinada unidade observacional, ou Decision Making Unit (DMU), em relação à unidade aceita como referência. Charnes et al. (1979) iniciaram e popularizam a técnica DEA, avaliando a eficiência de empresas assumindo retornos constantes de escala.

Outro trabalho importante foi o de Banker et al. (1984), o qual adotou retornos variáveis de escala. Portanto, na literatura, existem dois métodos clássicos para a Análise Envoltória de Dados, o retorno com escala constante (CCR) e o retorno com escala variável (VRS). Neste trabalho, foi escolhido o método VRS, pois se assume que um acréscimo no input poderá promover um acréscimo no output, não necessariamente proporcional, ou até mesmo um decréscimo (Ferrari & Cribari-Neto 2004). Vale ressaltar que não existe consenso na literatura se é melhor VRS ou CCR na área de Educação, portanto, ambos os métodos foram utilizados apesar de apenas um ter sido apresentado, devido à semelhança de resultados.

A análise foi feita com orientação baseada nos insumos, pois o objetivo do gestor escolar é administrar a escola, controlando os dispêndios, baseado no valor repassado para a Secretaria de Educação. Esta é chamada de eficiência técnica, que é a razão entre a quantidade de insumo utilizada para gerar determinado nível de produto, no caso, desempenho das avaliações e a quantidade mínima factível de insumos necessários (gasto por aluno) para produzir esse mesmo nível de produto.

É importante salientar que a análise de eficiência neste trabalho foi feita em dois estágios. Sendo que o primeiro estágio consiste na aplicação do modelo supracitado, DEA, enquanto o segundo consiste na verificação de quais são as variáveis independentes determinantes da eficiência das escolas, ou seja, são modelos de regressão cuja eficiência obtida no primeiro estágio é a variável dependente.

A Tabela 1 apresenta, na sua parte superior, quais são os produtos (*outputs*) utilizados na DEA, sendo o insumo (*input*) a relação gasto médio por aluno ao mês. Todas as escolas da amostra tiveram taxa de participação de pelo menos 50% no ENEM, de modo que a média da nota do ENEM é representativa do desempenho escolar. O gasto é formado pela agregação dos gastos com merenda e com pessoal – que contempla todos os tipos de ocupação em uma escola, como diretores, professores, auxiliares, administrativos, dentre outros. É importante salientar que não foram consideradas despesas com água, luz, telefone, pois estas foram inferiores em média a 2% do total de despesa de uma escola e não foram declaradas por todas as escolas. Além disso, os custos dos imóveis também foram desconsiderados, pois algumas unidades funcionam em prédio próprio e outras são alugadas, de modo que sua mensuração e, conseqüentemente, consideração no estudo foram impossibilitadas. Em situação análoga, por não serem representativos, não se considerou os repasses do

---

<sup>4</sup>A partir da análise, apenas uma escola, com gasto mensal por aluno inferior a R\$ 50,00 foi retirada da amostra.

Fundo Nacional do Desenvolvimento da Educação, a exemplo do Programa Dinheiro Direto na Escola.

Para estimar as equações do segundo estágio, utilizou-se a regressão beta inflacionado. Segundo Cribari-Neto & Pereira (2013) e Rosano-Peña et al. (2012), a adaptação dos modelos DEA para esse tipo de regressão é o mais adequado, pois o intervalo de eficiência estimado pelo DEA encontra-se entre 0 e 1. O modelo de regressão beta foi proposto por Ferrari & Cribari-Neto (2004) e baseia-se na suposição que a variável dependente possui distribuição beta, ou seja, a variável dependente está no intervalo (0,1). Nesse caso, dizemos que  $y$  tem distribuição Beta com média  $\mu$  e precisão  $\phi$  (denotado por  $y \sim \text{Beta}(\mu, \phi)$ ). Eles definiram uma parametrização diferente da usual para a obtenção de uma estrutura de regressão para a média da variável resposta, incluindo um parâmetro de precisão, que pode ser escrita da seguinte maneira:

$$f(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma(\phi(1-\mu))} y^{\mu\phi-1} (1-y)^{(1-\mu)\phi-1}, \quad 0 < y < 1, \quad (1)$$

em que  $0 < \mu < 1$ ,  $\phi > 1$  e  $\Gamma(\cdot)$  é uma função *gamma*. Nesse caso, o parâmetro de precisão é constante. Simas et al. (2010) estenderam esse modelo considerando uma estrutura de regressão para o parâmetro de precisão. Com isso, temos que o modelo de regressão beta com dispersão variável é definido sob a suposição que a média de  $y_t$  e o parâmetro de precisão satisfazem as seguintes relações:

$$g(\mu_t) = \sum_{i=1}^m x_{ti} \beta_i = \eta_t, \quad (2)$$

$$b(\phi_t) = \sum_{i=1}^q s_{ti} \lambda_i = \kappa_t, \quad (3)$$

em que  $\beta = (\beta_1, \dots, \beta_k)^T$  e  $\lambda = (\lambda_1, \dots, \lambda_k)^T$  são vetores de parâmetros de regressão desconhecidos,  $\eta_t$  e  $\kappa_t$  são preditores lineares e  $x_{t1}, \dots, x_{tk}$  e  $s_{t1}, \dots, s_{tk}$  são observações de covariáveis conhecidas.

Esse modelo faz parte da classe de modelos inflacionados em que a distribuição de probabilidade é uma mistura entre uma variável distribuída de forma contínua no intervalo (0,1) e uma distribuição degenerada concentrada em um dos limites desse intervalo (zero ou um), denotado por  $c$ . Se  $c=0$  temos o modelo de regressão beta inflacionado no ponto zero. Se  $c=1$  temos o modelo de regressão beta no ponto um.

Se  $y$  tem distribuição beta inflacionado, então sua função de densidade de probabilidade é dada por:

$$bi_c(y; \alpha, \mu, \phi) = \begin{cases} \alpha, & y = c, \\ (1-\alpha)f(y; \mu, \phi), & y \in (0, 1), \end{cases} \quad (4)$$

em que  $0 < \alpha < 1$  é o parâmetro de mistura da distribuição dado por  $\alpha = Pr(y = c)$ ,  $0 < \mu < 1$ ,  $\phi > 0$  e  $f(y; \mu, \phi)$  é a função de densidade  $\beta(\mu, \phi)$ .

Logo, o modelo de regressão beta inflacionado com dispersão variável é definido, supondo que a média condicional de  $y_t$ , a massa de probabilidade em  $c$  e o parâmetro de precisão satisfaz as seguintes relações:

**Tabela 1:** Variáveis utilizadas na análise de eficiência

Variáveis utilizadas no primeiro estágio	
Input	Gasto médio por Aluno
Outputs	Média da escola em Ciências da Natureza e suas Tecnologias no ENEM
	Média da escola em Ciências Humanas e suas Tecnologias no ENEM
	Média da escola em Matemática e suas Tecnologias no ENEM
	Média da escola em Linguagens, Códigos e suas Tecnologias no ENEM
	Média da escola em Redação no ENEM
Variáveis utilizadas no segundo estágio (modelos de regressão)	
Urbano ou rural	Refere-se à localização da escola no perímetro urbano ou rural
Indicador de Nível Socioeconômico (INSE)	É uma variável categórica disponibilizada pelo INEP que denota o nível socioeconômico das escolas a partir da aplicação de questionários aplicados juntos aos alunos (foi incluída dummies para as categorias, sendo como referência a categoria de nível baixo)
Tempo integral	<i>Dummy</i> que capta se a escola oferta ensino em tempo integral
Escola Militar	<i>Dummy</i> que capta se a escola é gerida pela polícia militar
Percentual de docentes com nível superior	É a participação relativa do número de docentes com ensino superior no total de docentes de uma escola

Fonte: SEDUCE-GO e INEP. Elaboração dos autores.

O uso do índice é sugerido por Alves et al. (2014) como variável explicativa ou de controle em estudos educacionais, ou ainda no planejamento de estudos comparativos para a definição dos grupos de escolas. Além disso, a ausência de um indicador como este induz a comparação entre escolas com alunados muito diferentes, como se vê nos rankings de escolas.

**Tabela 1:** Variáveis utilizadas na análise de eficiência (continuação)

Categorias para complexidade de gestão da escola	É uma variável construída pelo INEP que agrega características como o porte da escola, número de turnos de funcionamento, complexidade das etapas ofertadas pela escola e número de etapas/modalidades oferecidas (foram incluídas dummies, sendo a categoria de referência o Grupo 2)
Índice de regularidade do corpo docente	É uma variável construída pelo INEP que capta a assiduidade e regularidade da presença do professor na escola nos últimos 5 anos
Alunos por turma	Número médio de alunos por turma em uma determinada escola
Média de horas por aula diária	Número médio de horas por aula em uma determinada escola por dia
Laboratório de informática	<i>Dummy</i> que capta se a escola possui laboratório de informática
Laboratório de ciências	<i>Dummy</i> que capta se a escola possui laboratório de ciências
Biblioteca	<i>Dummy</i> que capta se a escola possui biblioteca

Fonte: SEDUCE-GO e INEP. Elaboração dos autores.

O uso do índice é sugerido por Alves et al. (2014) como variável explicativa ou de controle em estudos educacionais, ou ainda no planejamento de estudos comparativos para a definição dos grupos de escolas. Além disso, a ausência de um indicador como este induz a comparação entre escolas com alunados muito diferentes, como se vê nos rankings de escolas.



$$h(\alpha_t) = \sum_{i=1}^M z_{ti} \gamma_i = \zeta_t, \quad (5)$$

$$g(\mu_t) = \sum_{i=1}^m x_{ti} \beta_i = \eta_t, \quad (6)$$

$$b(\phi_t) = \sum_{i=1}^q s_{ti} \lambda_i = \kappa_t, \quad (7)$$

em que  $\gamma = (\gamma_1, \dots, \gamma_M)^T$ ,  $\beta = (\beta_1, \dots, \beta_m)^T$  e  $\lambda = (\lambda_1, \dots, \lambda_q)^T$  são vetores de parâmetros de regressão desconhecidos,  $x_{t1}, \dots, x_{tm}, z_{t1}, \dots, z_{tM}$ , e  $s_{t1}, \dots, s_{tq}$  são observações de covariáveis conhecidas. Nesse modelo de regressão, é possível utilizar diferentes tipos de funções de ligação, tais como *logit*, *probit*, *log-log* complementar, e *log-log* para  $\mu$  e  $\alpha$  e logarítmica ou raiz quadrada para  $\phi$ . Nesse caso,  $\mu_t$  é a média de  $y_t$  condicional em  $y_t \in (0, 1)$ .

As estimativas dos parâmetros nos dois modelos apresentados são obtidas por métodos numéricos de maximização da função de log-verossimilhança, usando algoritmo de otimização não linear (Algoritmo de Newton ou algoritmo quasi-Newton).

Os critérios de seleção adotados são baseados na função de Log-verossimilhança avaliada em seu ponto de máximo, acrescido de uma penalidade (referente ao número de parâmetros do modelo). Dessa forma, a adição ou remoção de covariáveis será baseada no critério de Akaike Generalizado (GAIC), em conjunto, com o procedimento *stepwise*. Esse método pode ser decomposto em três técnicas: *forward selection*, *backward elimination* e *stepwise regression*.

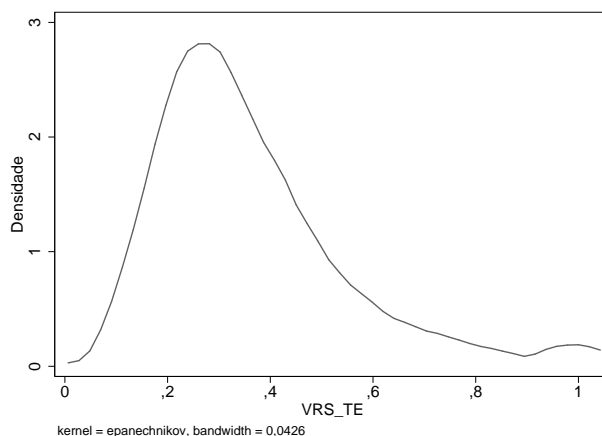
A primeira técnica consiste na adição sequencial de variáveis candidatas ao modelo. A segunda técnica se dá por meio de um modelo completo, contendo todas as variáveis explicativas candidatas, procedendo-se na eliminação (ou não) de parte dessas variáveis de maneira sequencial. A terceira técnica consiste em uma combinação das duas técnicas anteriores, assim, a cada adição de variáveis no modelo é realizado um procedimento para verificar se não há variáveis redundantes, mitigando assim a possibilidade da multicolinearidade entre os regressores. Dessa forma, o método *Stepwise* é considerado o melhor método para a seleção de variáveis explicativas (Facchin 2005).

### 3 Resultados

#### 3.1 Análise Envoltória de Dados (DEA)

A análise de eficiência revelou que as escolas públicas de Goiás possuem uma eficiência média de 0,36, valor insatisfatório, já que o indicador de eficiência varia entre 0 e 1. Além disso, apenas 2,3% das escolas analisadas possuem máxima eficiência. Em termos municipais, Taquaral de Goiás é a localidade mais eficiente, uma vez que a escola presente no município apresentou eficiência máxima. Em seguida, os únicos municípios que apresentam eficiência superior a 0,70 foram Paraúna, Cidade Ocidental e Águas Lindas de Goiás. No outro extremo, como municípios menos eficientes, tem-se São Miguel do Passa Quatro, Pilar de Goiás e Guaraitá, todos com valores inferiores a 0,12. Esses resultados sugerem que o Estado de Goiás, em geral, não consegue alocar recursos educacionais de maneira eficiente, sendo considerável a diferença entre

as unidades de eficiências medianas e de maiores eficiências. A Figura 1 apresenta a distribuição da eficiência estimada para as escolas de Goiás. Visualiza-se uma grande concentração de escolas com valores muito baixos de eficiência (cerca de 0,30) e um pequeno grupo de escolas que se destoam das demais, com eficiências superiores a 0,80. Isso mostra que existe um pequeno grupo de unidades que são altamente eficientes, contrastando com a grande maioria que se concentra em baixos níveis de eficiência.

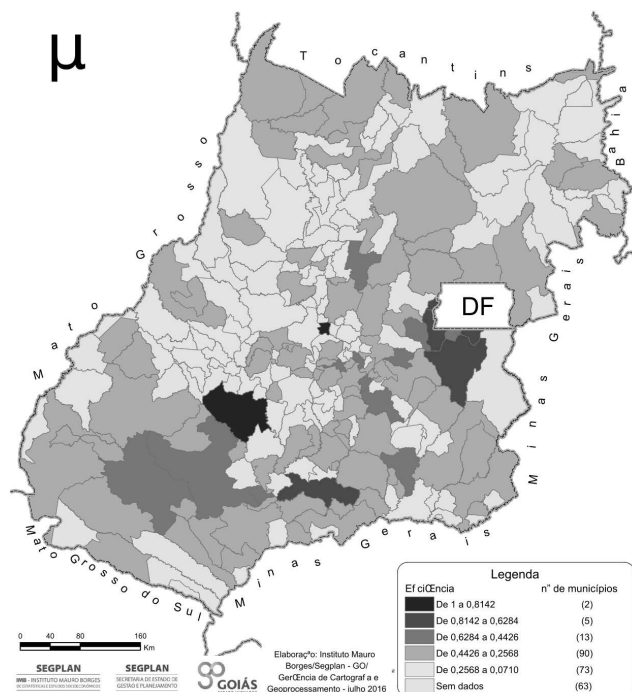


**Figura 1:** Densidade estimada da eficiência (VRS-TE) das escolas públicas de ensino médio de Goiás, 2014.

A Figura 2 apresenta a distribuição espacial da eficiência no Estado de Goiás. Percebe-se uma ausência de padrão espacial. Corrobora ainda esse aspecto, uma análise mais geral das 50 escolas mais e menos eficientes. Visualiza-se que, dentro do ranking das 50 escolas menos eficientes, há ausência de um padrão de dependência espacial entre os municípios, uma vez que apenas os municípios de Goiânia e Aragarças apresentaram mais de uma escola nesse ranking, com seis e duas escolas no total, respectivamente. Todavia, no outro extremo, há uma situação inversa, percebe-se que dentre as 50 escolas mais eficientes, 38 estão localizadas em apenas 11 dos 183 municípios analisados. Com destaque para Luziânia, Goiânia, Águas Lindas de Goiás e Anápolis, com respectivamente, oito, sete, cinco e quatro escolas dentre as 50 mais eficientes.

Uma vez que a DEA atribui toda a diferença existente entre a unidade eficiente e a unidade sob análise à ineficiência, investigar quais são os fatores que influenciam essas diferenças torna-se fundamental, dados os contextos diferentes em que as escolas atuam. Duas escolas podem apresentar a mesma gestão de recursos, gastando exatamente a mesma quantia por aluno, todavia uma delas pode estar inserida em um contexto de realidade social mais privilegiada que a outra. Nesse sentido, a escola em melhor contexto social possivelmente apresentará melhores resultados, mesmo que o esforço gerencial de ambas as administrações seja o mesmo. Com isso, faz-se necessário ponderar os resultados obtidos no DEA, já que nem toda diferença existente nos resultados obtidos por duas escolas pode ser atribuída à ineficiência (Junior et al. 2015, Boueri et al. 2015).

A Tabela 2 apresenta algumas estatísticas descritivas que destacam a diferença entre as escolas do primeiro e do último decil de acordo com a eficiência



**Figura 2:** Mapa da eficiência das escolas públicas de ensino médio de Goiás, 2014.

estimada, de modo a se lançar luz sobre a grande diferença observada na distribuição de eficiência no Estado.

A despeito dos grupos serem semelhantes na quantidade de escolas, a diferença entre a eficiência média é elevada, com uma razão de 5,6 vezes entre eles. É interessante notar que o grupo mais eficiente apresenta uma nota 20% superior em redação e cerca de 5% maior nas demais áreas do conhecimento, com um custo médio por aluno 76% mais baixo. Esse resultado preliminar destaca a enorme diferença entre os recursos aplicados pelo Governo de Goiás na educação e os resultados obtidos em termos de qualidade.

Com relação à estrutura física das escolas, percebe-se que as unidades mais eficientes apresentam maior presença de laboratórios de ciências (54% contra 33%), menor presença de biblioteca (89% contra 94%) e presença levemente superior de laboratórios de informática (89% contra 86%). Sendo assim, diferenças de estrutura física não parecem, em um primeiro momento, explicar importantes diferenças nos gastos. Um resultado interessante diz respeito às variáveis de fluxo escolar, nas quais as escolas mais eficientes apresentam piores resultados, com taxas de aprovação inferiores e de abandono superiores. Apesar dos resultados inferiores nesses indicadores, os resultados obtidos em termos de qualidade (notas do ENEM) são superiores, conforme destacado anteriormente.

A diferença de eficiência existente entre escolas com estruturas administrativas claramente distintas é outro ponto de destaque, uma vez que o grupo de escolas mais eficientes é composto por cerca de um quarto de escolas militares. Tais unidades militares<sup>5</sup> apresentaram eficiência média de 0,71, valor bem su-

<sup>5</sup>Neste estudo não foi considerada a contribuição comunitária voluntária que é realizada em

**Tabela 2:** Dados descritivos das escolas mais eficientes e menos eficientes

Variável	1º Decil	10º Decil
Nº de Escolas	36,0	35,0
Eficiência média (%)	14,0	79,0
Média Nota Ciências Naturais	463,0	479,0
Média Nota Ciências Humanas	515,0	536,0
Média Nota Matemática	436,0	460,0
Média Nota Redação	385,0	454,0
Média Nota Linguagem	478,0	503,0
Gasto por aluno por mês (R\$)	908,0	214,0
Laboratório de Ciências (%)	33,0	54,0
Biblioteca (%)	94,0	89,0
Urbano (%)	100,0	100,0
Laboratório de Informática(%)	86,0	89,0
Taxa de aprovação (%)	91,7	84,0
Taxa de reprovação (%)	5,3	11,5
Taxa de abandono (%)	3,0	4,5
Taxa de participação ENEM (%)	70,3	69,1
Docentes com Ensino Superior (%)	94,0	88,0
Colégio Militar (%)	0,0	25,7
Tempo Integral (%)	36,1	2,9
INSE	46,8	50,4
Alunos por turma	19,0	35,0
Média de horas aula	6,0	4,9

Fonte: SEDUCE-GO e INEP. Elaboração dos autores.

perior à média geral. Além disso, das oito escolas de máxima eficiência, quatro são militares. Ou seja, as escolas militares, além de muito eficazes<sup>6</sup>, são altamente eficientes, pois, para produzirem resultados satisfatórios, alocam de maneira mais eficiente seus fatores produtivos. Entretanto, faz-se necessário destacar que as escolas militares podem acarretar em vieses nos resultados obtidos, uma vez que nessas instituições existem processos seletivos para a entrada de alunos, o que muito provavelmente influenciam os resultados obtidos por elas em termos de nota. Os resultados dos indicadores socioeconômicos reforçam a possível diferença de composição entre os grupos, uma vez que as escolas mais eficientes trabalham com um público de realidade social mais favorecida (com um indicador de 50,4 contra 46,8).

No que diz respeito à localização rural ou urbana, um importante resultado é observado, em que consiste na ausência de escola rural dentre as de menor eficiência. Geralmente, atribui-se baixa eficácia às escolas rurais<sup>7</sup>, ou seja, que estas não possuem resultados satisfatórios em termos de desempenho em exames de proficiência educacional. Todavia, ao analisarmos esse desempenho condicionado aos seus recursos, vemos que é factível escolas rurais não terem os piores de níveis eficiência.

Finalmente, as variáveis mais associadas aos custos destacam melhores resultados para as escolas mais eficientes: elas apresentam apenas 3% de escolas

12 cotas mensais as quais são destinadas a prover despesas gerais, uma vez que esta não era oficialmente reconhecida como um custo pela SEDUCE-GO, portanto, não foi considerado na análise de eficiência.

<sup>6</sup>O desempenho (eficácia) das escolas militares foi superior ao das demais em todas as grandes áreas do ENEM, conforme Tabela A.1.

<sup>7</sup>Verificou-se que a eficácia das escolas rurais foi inferior a das urbanas em todas as grandes áreas do ENEM, conforme Tabela A.2.

com a oferta de ensino integral (36% para as menos eficientes), 35 alunos por turma (19 para as menos eficientes), um número inferior de horas aula (4,9 contra 6) e menor número de professores com ensino superior (88% contra 94%). Como o custo associado à manutenção de ensino integral é maior, existem ganhos de escala ao se alocar mais alunos por turma e professores com titulação mais baixa e que lecionam menos horas significam menores custos de mão de obra, tais resultados corroboram as estimativas de eficiência verificadas para o último decil. Após essa análise inicial da eficiência, o segundo estágio da análise conduz uma regressão beta inflacionado, de modo a se analisar mais especificamente quais são os fatores que podem explicar as diferentes eficiências observadas entre as escolas.

### 3.2 Modelos de Regressão

O segundo estágio busca identificar os determinantes da eficiência, isto é, quais variáveis ambientais, que fogem ao controle dos gestores públicos no curto prazo<sup>8</sup>, influenciam nos escores de eficiência. Nesse sentido, foram consideradas as variáveis apresentadas na Tabela 1, de modo que o escore de eficiência foi utilizado como variável dependente e as demais como independentes em diversas regressões testadas.

Como justificado na seção anterior, dentre os modelos estimados, o beta inflacionado se apresentou como o mais robusto em comparação às estimações de mínimos quadrados e tobit. Além de a variável dependente ser truncada (situar-se entre 0 e 1), os critérios de informação (BIC e AIC) e a análise residual, sugeriram o uso do modelo beta como o mais adequado (Hall 2000, Heller et al. 2006).

Nesse âmbito, salienta-se que a supracitada Figura 1, mostrou que a distribuição da densidade estimada dos *scores* de eficiência não possui distribuição normal, sendo levemente assimétrica à direita. Isso, em algum grau violaria os pressupostos de normalidade na distribuição da variável resposta do modelo *tobit* e dos mínimos quadrados ordinários. Além disso, ao se utilizar o modelo beta inflacionado é possível, por meio do modelo de máxima eficiência (massa de probabilidade igual a 1), estabelecer quais variáveis são responsáveis por gerar máxima eficiência. Assim, é possível identificar características diretamente associadas a um melhor desempenho, que, portanto, podem servir como *benchmarking*, ou seja, como boas práticas para as demais escolas (Hoff 2007, Kieschnick & McCullough 2003).

Assim, visualiza-se na Tabela 3, no primeiro modelo,  $\mu$ , que as variáveis relacionadas ao nível socioeconômico das escolas, denotadas pelas variáveis INSE, tendem a aumentar a eficiência média das escolas, com retornos crescentes à medida que se eleva o nível socioeconômico captado pelas diferentes variáveis (níveis) de INSE. Resultados similares são encontrados em vários estudos que versam sobre o desempenho, como Chechia & Andrade (2005), Fernandes & Gremaud (2009), Couri (2010), Raposo et al. (2011), entre outros.

A Tabela 3 mostra que as escolas com ensino em tempo integral, e com maior média de hora aula, tendem a ser mais ineficientes. Esse resultado sugere que a superutilização da estrutura físico-pedagógica para prover educa-

<sup>8</sup>As variáveis referentes à infraestrutura escolar, a exemplo da biblioteca ou laboratórios, podem ser controladas pelos gestores escolares no médio/longo prazo. Entretanto, os gestores não possuem a capacidade de adquirir esses fatores imediatamente, considerando, portanto, como dados.

ção em tempo integral, que se traduz na necessidade do número de docentes, ampliação da infraestrutura escolar, como salas de aula, quadras, laboratórios, não tem sido capaz de gerar incrementos substanciais na proficiência dos estudantes no ENEM *vis-à-vis* o aumento dos custos das escolas para implementarem tais ações.

Além disso, três variáveis que, à primeira vista, deveriam contribuir para uma maior eficiência das escolas são as relativas à complexidade das escolas, ao número de docentes com ensino superior e a sua regularidade. Todavia, seus coeficientes negativos revelam uma tendência oposta. Sendo assim, a mera disponibilização de docentes com nível educacional elevado em ambientes com boa complexidade, em termos de infraestrutura, sem projetos pedagógicos adequados, não tem sido capaz de promover resultados satisfatórios do ponto de vista da eficiência.

Ainda nessa discussão, o coeficiente negativo da variável da regularidade do docente, torna-se compreensível em um cenário em que os projetos educativos e pedagógicos não são bem traçados e implementados, fazendo com que a assiduidade do docente em sala de aula, por vezes, não gere uma elevação substancial da proficiência dos estudantes capaz de compensar o aumento dos custos gerados pelos mesmos.

O modelo de máxima eficiência,  $\alpha$ , fornece importante resultado, pois apresenta as características imprescindíveis que as escolas devem perseguir para obterem máxima eficiência. Por meio desse modelo, foi possível obter uma espécie de *benchmarking* das escolas de máxima eficiência.

Visualiza-se assim que a variável relacionada ao maior nível socioeconômico da escola é a que produz maiores incrementos na direção da máxima eficiência. É importante lembrar que essa variável contempla dimensões relacionadas ao nível socioeconômico da família, por meio de bens, renda, escolaridade dos pais, entre outras.

Assim, embora essa variável não seja de controle por parte das escolas, ela revela uma lacuna em que o poder público pode atuar, tentando criar políticas educacionais capazes de minimizar a perda de eficiência das escolas com baixo nível socioeconômico.

Para gerar eficiência máxima a variável complexidade da escola se mostrou indispensável. É importante compreender que essa variável contempla diversos indicadores como o número de matrículas, turnos em que a escola opera, quantas modalidades de ensino são ofertadas, dentre outras. Os níveis dois e três dentro de uma escala máxima, que vai até seis, denotam que as escolas de complexidade mais baixa e intermediária são as que devem ser perseguidas pelas escolas. Dessa forma, esse resultado sugere que as escolas devem se concentrar em um determinado segmento de ensino, não terem muitos alunos e não operarem em diversos turnos, para obterem maior eficiência. Portanto, os ganhos de escala com maior infraestrutura são mitigados pela perda de especialização na oferta escolar. Possivelmente, a mescla de alunos de diferentes faixas etárias e etapas de ensino, afetam o desempenho na prova do ENEM.

Ainda em relação ao modelo de máxima eficiência,  $\alpha$ , é característica das escolas de máxima eficiência turmas numerosas, resultado similar aos encontrados em estudos relacionados à eficácia (Camargo 2012, Leon & Menezes-Filho 2002, Fraiman 1997). Isso sugere que, em termos de eficiência, as turmas numerosas podem representar uma importante diminuição de custo das escolas sem afetar demasiadamente a eficácia da proficiência dos estudantes no ENEM.

**Tabela 3:** Estimativas dos parâmetros do modelo beta inflacionado

Modelo para média ( $\mu$ )			
Variáveis	Estimativa	Erro-padrão	<i>p</i> – valor
Intercepto	-0,959	0,226	0,000
INSE alto	1,226	0,673	0,069
INSE médio alto	0,342	0,055	0,000
INSE médio	0,306	0,040	0,000
INSE médio baixo	0,242	0,064	0,000
Tempo Integral	-0,487	0,073	0,000
Escola Militar	0,490	0,195	0,012
Docente com ensino superior	-0,005	0,001	0,000
Complexidade de gestão da escola 3	-0,297	0,083	0,000
Regularidade do docente	-0,209	0,048	0,000
Alunos por turma	0,059	0,003	0,000
Média hora aula	-0,088	0,019	0,000
Modelo de precisão ( $\phi$ )			
Variáveis	Estimativa	Erro-padrão	<i>p</i> – valor
Intercepto	20,944	0,848	0,000
INSE alto	-18,912	0,891	0,000
INSE médio alto	-17,593	0,197	0,000
INSE médio	-17,479	0,123	0,000
INSE médio baixo	-17,299	0,221	0,000
Tempo Integral	1,208	0,376	0,001
Escola Militar	-0,588	0,426	0,168
Complexidade de gestão da escola 2	-2,598	0,987	0,009
Complexidade de gestão da escola 3	-0,825	0,344	0,017
Complexidade de gestão da escola 4	-0,406	0,245	0,098
Alunos por turma	-0,075	0,006	0,000
Média hora aula	0,394	0,180	0,029
Laboratório de ciências	0,409	0,157	0,010
Modelo de máxima eficiência ( $\alpha$ , massa de probabilidade igual a 1)			
Variáveis	Estimativa	Erro-padrão	<i>p</i> – valor
Intercepto	-12,990	3,629	0,000
INSE alto	5,185	1,269	0,000
Complexidade de gestão da escola 2	4,269	1,670	0,011
Complexidade de gestão da escola 3	3,307	1,505	0,029
Alunos por turma	0,246	0,098	0,013
Critérios de Seleção dos modelos			
AIC: -618,148		BIC: -501,900	

Fonte: INEP e SEDUCE-GO. Elaboração dos autores.

Por fim, o modelo de precisão,  $\phi$ , está relacionado com a qualidade do ajuste dos regressores e sugere que as estimativas mais precisas na variável dependente (eficiência obtida no primeiro estágio) são obtidas principalmente nas escolas com maior número de alunos por turma.

#### 4 Resultados

Esse estudo constituiu um esforço inédito de avaliar a eficiência das escolas públicas de ensino médio no maior nível de desagregação disponível e com a variável custo-aluno. Os estudos anteriores, referentes à eficiência educacional (Raposo et al. 2011, Carvalho & Sousa 2014), utilizaram como insumos variáveis indiretamente relacionadas ao custo. Assim, as escolas foram tratadas como unidades individuais e a eficiência foi avaliada diretamente por meio dos recursos financeiros.

Em certa medida, algumas análises do estudo têm sua validade externa comprometida, pois os custos dos fatores, principalmente da mão de obra, podem variar de acordo com as UF's. Outra limitação da pesquisa é em relação ao viés de seleção sobre as escolas voltadas exclusivamente para o ensino médio. Uma vez que a prova não é obrigatória, os resultados de cada escola podem ser influenciados pelo nível de esforço e interesse dos seus alunos. Desse modo, os gestores daquelas escolas que possuem outras etapas de ensino, a exemplo do ensino fundamental, não focam as atividades escolares na realização da prova do ENEM. Contudo, como a comparação se deu apenas entre escolas que fazem a prova, esse viés é amenizado. Por fim, como Andrade (2011) ressalta, é possível que escolas que apresentam os melhores resultados no exame sejam aquelas que recebem os melhores alunos e não aquelas que mais podem colaborar para o desempenho deles.

De modo geral, os valores das eficiências encontrados foram baixos, a eficiência média das escolas analisadas foi de 0,36. A ausência de dependência espacial na eficiência das escolas corroborou em grande parte a justificativa deste estudo, de que o sucesso das escolas tidas como mais eficientes é reflexo de ações individuais dos gestores públicos e não de diretrizes gerais.

Os resultados relativos às variáveis de localização das escolas (urbana ou rural) e das escolas militares elucidaram a noção de eficiência que é obtida na Análise Envoltória de Dados (DEA). Percebe-se que os resultados em termos de desempenho (eficácia) em proficiência no ENEM, que tendem a ser piores nas escolas localizadas na zona rural e melhores para escolas militares, não podem ser confundidos com a noção de eficiência. Portanto, a noção de eficiência adotada neste estudo está associada à utilização de maneira ótima dos insumos restrito a uma determinada estrutura de gastos.

De modo geral, os modelos de regressão sugerem a necessidade de implementação de novos projetos educativo-pedagógicos. Foi mostrado que a presença de professores com maior educação formal e com assiduidade não tem sido capaz de promover efeitos satisfatórios no aumento da proficiência dos estudantes de Goiás no ENEM *vis-à-vis* ao seu custo. E que escolas em tempo integral e com carga horária elevada nos moldes atuais não se mostraram eficientes.

Por fim, outra contribuição deste estudo foi obtida a partir do modelo beta inflacionado, de máxima eficiência. Os modelos de regressão mostraram que as políticas públicas educacionais podem ser balizadas na direção de escolas



com ênfase exclusiva em poucos tipos de modalidade de ensino, com número de estudantes por escola não muito numerosos, isto é, escolas com ofertas mais especializadas. Destaca-se, ainda, a importância do nível socioeconômico dos alunos para a eficiência escolar, indicando que políticas públicas transversais, com foco nas famílias, têm grande potencial de melhoria na utilização dos recursos públicos educacionais.

Conclui-se que a análise de eficiência da educação em seu nível mais desagregado suscitou uma discussão pouco presente nas políticas educacionais, que é a transmissão de uma política idealizada e a sua materialização em termos de resultados. Por vezes, o gestor público, a seu bel-prazer ou no intuito de atender as reivindicações públicas, aumenta o número de docentes de uma escola, amplia a estrutura física da escola, por meio da criação de laboratórios, ampliação de modalidades educacionais, entre outros. Mas nem sempre tais ações se materializam no aumento da proficiência dos estudantes, o que pode não justificar determinadas implementações.

## 5 Agradecimentos

Os autores agradecem o apoio de suas respectivas instituições e da Secretaria de Estado de Educação, Cultura e Esporte do Estado de Goiás (SEDUCE-GO), pela cessão dos dados.

## Referências Bibliográficas

- Alves, M. T. G., Soares, J. F. & Xavier, F. P. (2014), 'Índice socioeconômico das escolas de educação básica brasileiras', *Ensaio: Avaliação e Políticas Públicas em Educação* 22(84), 671–703.
- Andrade, E. d. C. (2011), 'Rankings em educação: tipos, problemas, informações e mudanças', *Estudos Econômicos (São Paulo)* 41(2), 323–343.
- Banker, R. D., Charnes, A. & Cooper, W. W. (1984), 'Some models for estimating technical and scale inefficiencies in data envelopment analysis', *Management science* 30(9), 1078–1092.
- Barros, A. R. et al. (2011), Desigualdades regionais no brasil: natureza, causas, origens e soluções, Technical report.
- Benegas, M. et al. (2012), 'O uso do modelo network dea para avaliação da eficiência técnica do gasto público em ensino básico no brasil', *Economia* 13(3a).
- Boueri, R., Rocha, F. & Rodopoulos, F. (2015), 'Avaliação da qualidade do gasto público e mensuração da eficiência', *Rogério Boueri, Fabiana Rocha, Fabiana Rodopoulos (Organizadores)-Brasília: Secretaria do Tesouro Nacional*.
- Camargo, J. (2012), 'O efeito do tamanho da turma sobre o desempenho escolar: uma avaliação do impacto da "enturmação" no ensino fundamental do rio grande do sul'.
- Carvalho, L. D. B. d. & Sousa, M. d. C. S. d. (2014), 'Eficiência das escolas públicas urbanas das regiões nordeste e sudeste do brasil: uma abordagem em três estágios', *Estudos Econômicos (São Paulo)* 44(4), 649–684.

- Charnes, A., Cooper, W. W. & Rhodes, E. (1979), 'Measuring the efficiency of decision-making units', *European journal of operational research* 3(4), 339.
- Chechia, V. A. & Andrade, A. d. S. (2005), 'O desempenho escolar dos filhos na percepção de pais de alunos com sucesso e insucesso escolar', *Estudos de Psicologia* 10(3), 431–440.
- Couri, C. (2010), 'Nível socioeconômico e cor/raça em pesquisas sobre efeito-escola', *Estudos em Avaliação Educacional, São Paulo* 21(47), 449–472.
- Cribari-Neto, F. & Pereira, T. L. (2013), 'Avaliação da eficiência de administrações municipais no estado de São Paulo: Uma nova abordagem via modelos de regressão beta', *Rev. Bras. Biom* 31(2), 270–294.
- Facchin, S. (2005), 'Técnicas de análise multivariável aplicadas ao desenvolvimento de analisadores virtuais'.
- Faria, F. P., Jannuzzi, P. d. M., Silva, S. J. d. et al. (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*.
- Farrell, M. J. (1957), 'The measurement of productive efficiency', *Journal of the Royal Statistical Society. Series A (General)* 120(3), 253–290.
- Fernandes, R. & Gremaud, A. P. (2009), 'Qualidade da educação: avaliação, indicadores e metas', *Educação básica no Brasil: construindo o país do futuro. Rio de Janeiro: Elsevier* 1, 213–238.
- Ferrari, S. & Cribari-Neto, F. (2004), 'Beta regression for modelling rates and proportions', *Journal of Applied Statistics* 31(7), 799–815.
- Fraiman, L. P. (1997), 'A importância da participação dos pais na educação escolar dissertação (mestrado)'.
- Gonçalves, F. d. O. & França, M. T. A. (2013), 'Eficiência na provisão de educação pública municipal: uma análise em três estágios dos municípios brasileiros', *Estudos Econômicos (São Paulo)* 43(2), 271–299.
- Hall, D. B. (2000), 'Zero-inflated poisson and binomial regression with random effects: a case study', *Biometrics* 56(4), 1030–1039.
- Heller, G., Stasinopoulos, M., Rigby, B. et al. (2006), 'The zero-adjusted inverse gaussian distribution as a model for insurance claims', in 'Proceedings of the 21th International Workshop on Statistical Modelling', Vol. 226233.
- Hoff, A. (2007), 'Second stage DEA: Comparison of approaches for modelling the DEA score', *European Journal of Operational Research* 181(1), 425–435.
- IBGE (2011), *Sinopse do Censo Demográfico 2010*, Instituto Brasileiro de Geografia e Estatística - IBGE, Rio de Janeiro.  
**URL:** <http://www.ibge.gov.br/home/estatistica/populacao/censo2010/sinopse.pdf>
- Júnior, A. P. S. & Novi, J. C. (2014), 'Proposta metodológica: avaliação externa e desempenho dos alunos', *Revista Brasileira de Política e Administração da Educação-Periódico científico editado pela ANPAE* 30(3).

Junior, A. P. S., Novi, J. C., Ferreira, J., de Oliveira, M. M. B. & Miranda, P. S. (2015), 'Eficiência na gestão escolar: em busca das melhores práticas em escolas municipais brasileiras do ensino fundamental', *Revista Meta: Avaliação* 7(19), 85–122.

Kieschnick, R. & McCullough, B. D. (2003), 'Regression analysis of variates observed on (0, 1): percentages, proportions and fractions', *Statistical modelling* 3(3), 193–213.

Leon, F. L. L. d. & Menezes-Filho, N. A. (2002), 'Reprovação, avanço e evasão escolar no brasil'.

Liu, J. S., Lu, L. Y., Lu, W.-M. & Lin, B. J. (2013), 'A survey of dea applications', *Omega* 41(5), 893–902.

Oliveira, G. R. (2016), 'Três ensaios sobre avaliação de políticas públicas para o estado de goiás'.

Raposo, I. P., Soares, S., Maia, A. & Menezes, T. A. (2011), 'Public school efficiency using data envelopment analysis: an empirical application for the northeast of brazil', *Anais do 39 Encontro Nacional de Economia (ANPEC)* .

Rosano-Peña, C., Albuquerque, P. H. M. & Marcio, C. J. (2012), 'A eficiência dos gastos públicos em educação: evidências georreferenciadas nos municípios goianos', *Economia Aplicada* 16(3), 421–443.

Salgado Junior, A. P. & Novi, J. C. (2015), 'Práticas de gestão e destinação dos recursos financeiros em secretarias municipais da educação e escolas públicas municipais de ensino fundamental: um estudo multicaso', *Desenvolvimento em Questão* 14(33), 330–362.

Simas, A. B., Barreto-Souza, W. & Rocha, A. V. (2010), 'Improved estimators for a general class of beta regression models', *Computational Statistics & Data Analysis* 54(2), 348–366.

Zoghbi, A. C. P., Matos, E. H. C. d., Rocha, F. F. & Arvate, P. R. (2009), 'Mensurando o desempenho e a eficiência dos gastos estaduais em educação fundamental e média', *Estudos Econômicos (São Paulo)* 39(4), 785–809.

## Apêndice A

**Tabela A.1:** Média das notas do ENEM por área de conhecimento e tipo de escola

Tipo de escola	Ciências da Natureza e suas Tecnologias	Ciências Humanas e suas Tecnologias	Matemática e suas Tecnologias	Redação	Linguagens, Códigos e suas Tecnologias
Integral	479,67	536,02	455,74	446,38	500,21
Militar	492,90	547,31	476,02	489,73	515,48
Normal	462,51	517,23	439,27	407,86	484,03
Total	465,43	520,30	442,39	414,99	486,90

Fonte: Inep. Elaboração dos autores.

**Tabela A.2:** Média das notas do ENEM por área de conhecimento e localização da escola

Tipo de escola	Ciências da Natureza e suas Tecnologias	Ciências Humanas e suas Tecnologias	Matemática e suas Tecnologias	Redação	Linguagens, Códigos e suas Tecnologias
Rural	458,84	499,59	434,02	354,58	469,45
Urbana	465,54	520,65	442,53	416,02	487,20
Total	465,43	520,30	442,39	414,99	486,90

Fonte: Inep. Elaboração dos autores.