

Desempenho em matemática no estado de São Paulo à luz de um indicador educacional do tipo Gini^I

ERIKA CAPELATO^{II}

JAMIL GOMES DE ABREU^{III}

ANDRÉ LUIZ CORRÊA^{IV}

<http://dx.doi.org/10.22347/2175-2753v15i47.3944>

Resumo

Neste trabalho analisamos o desempenho em matemática das regiões administrativas e da região metropolitana de São Paulo. Primeiramente, propomos um indicador educacional para medir o desempenho escolar. A seguir, obtemos o valor deste indicador usando a nota dos alunos do ensino médio da rede pública no Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (Saresp). Finalmente, usando regressão linear, investigamos a relação entre o desempenho em matemática e variáveis socioeconômicas. Os resultados mostraram que a região administrativa de São José do Rio Preto apresentou os melhores resultados do indicador educacional e a região metropolitana de São Paulo, os piores resultados. Além disso, observamos que o desempenho em matemática e as variáveis econômicas estão positivamente relacionadas.

Palavras-chave: Indicador Educacional; Matemática; Saresp.

Submetido em: 19/05/2022

Aprovado em: 29/05/2023

^I Este estudo é produto da pesquisa "Proposta de um indicador tipo Gini para medição do desempenho escolar", sob coordenação da primeira autora e financiada pela Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP.

^{II} Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP), São Paulo (SP), Brasil; <https://orcid.org/0000-0001-9039-6659>; e-mail: erika.capelato@unesp.br.

^{III} Universidade Federal de Lavras (UFLA), Lavras (MG), Brasil; <https://orcid.org/0000-0002-0353-461X>; e-mail: jamil.abreu@ufla.br.

^{IV} Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP), São Paulo (SP), Brasil; <https://orcid.org/0000-0001-5701-7549>; e-mail: andre.correa@unesp.br.

Performance in mathematics in the state of São Paulo in the light of a Gini educational indicator

Abstract

In this work we analyze the performance in mathematics of the administrative regions and the metropolitan region of São Paulo. First, we propose an educational indicator to measure school performance. Next, we obtain the value of this indicator using the grade of public high school students in the School Performance Assessment System of the State of São Paulo (Saresp). Finally, using linear regression, we investigated the relationship between math performance and socioeconomic variables. The results showed that the administrative region of São José do Rio Preto presented the best results of the educational indicator and the metropolitan region of São Paulo, the worst results. In addition, we observed that performance in mathematics and economic variables are positively related.

Keywords: Educational Indicator; Math; SARESP.

Rendimiento en matemáticas en el estado de São Paulo a la luz de un indicador educativo de Gini

Resumen

En este trabajo analizamos el desempeño en matemáticas de las regiones administrativas y de la región metropolitana de São Paulo. En primer lugar, proponemos un indicador educativo para medir el rendimiento escolar. A continuación, obtenemos el valor de este indicador a partir de la nota de los alumnos de enseñanza media pública en el Sistema de Evaluación del Desempeño Escolar del Estado de São Paulo (Saresp). Finalmente, usando regresión lineal, investigamos la relación entre el rendimiento en matemáticas y las variables socioeconómicas. Los resultados mostraron que la región administrativa de São José do Rio Preto presentó los mejores resultados del indicador educativo y la región metropolitana de São Paulo, los peores resultados. Además, observamos que el rendimiento en matemáticas y las variables económicas se relacionan positivamente.

Palabras clave: Indicador educativo; Matemáticas; Saresp.

Introdução

Uma vez que o nível educacional de um país diz muito sobre o seu desenvolvimento econômico e vice-versa, a avaliação da educação em larga escala e os indicadores têm assumido um papel marcante nas políticas educacionais brasileiras. Os indicadores educacionais são instrumentos que respondem à demanda da sociedade por transparência de resultados que mostrem a eficácia das escolas e, conseqüentemente, a qualidade educacional. Contudo, alguns indicadores educacionais, seja por suas características metodológicas, ou por serem analisados isolados de seu contexto, muitas vezes não mostram que existem diferenças no nível de desempenho entre grupos de alunos. A ausência destas diferenças é o que chamamos de equidade escolar.

Para o contexto educacional brasileiro encontramos na literatura diversos trabalhos que discutem o desempenho e a equidade escolar, ou ainda, que propõem indicadores capazes de mensurá-los (FRANCO; ORTIGÃO; ALBERNAZ; BONAMINO; AGUIAR; ALVES; SÁTYRO, 2007; HELENE; MARIANO 2020; MARTÍNEZ-GARRIDO; JAVIER MURILLO 2016; SOARES, 2005; SOARES; RODRIGUES; ERNICA, 2019; ZOGHBI; MATTOS; ROCHA; ARVATE, 2011).

O presente artigo se inscreve nesta literatura no sentido de propor um indicador do tipo Gini, que é capaz de captar o desempenho da totalidade de um grupo de alunos submetidos a uma avaliação e a equidade do desempenho dentro deste grupo. O aspecto distinto deste indicador e dos observados na literatura é que ele utiliza somente uma variável, a nota obtida em uma avaliação.

Além de propor o indicador educacional, este trabalho avança na tentativa de: (i) avaliar o desempenho escolar em matemática das regiões administrativas paulista e da região metropolitana de São Paulo a partir da nota obtida pelos alunos da rede pública de ensino da 3ª série do Ensino Médio no período 2008 a 2017 no Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (Saresp) e (ii) relacionar o desempenho em matemática, a partir do valor do indicador, com variáveis socioeconômicas destas regiões.

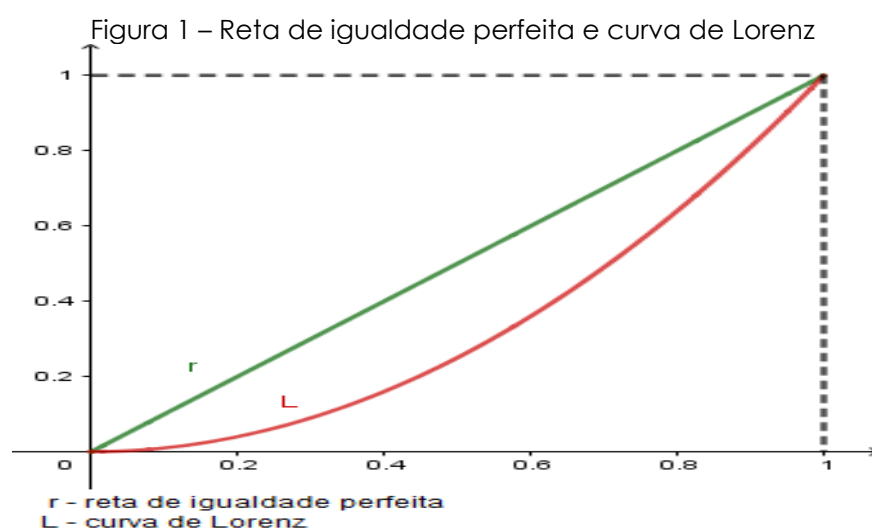
A justificativa para analisar o desempenho em matemática pode ser explicada com o mesmo argumento dos autores Franco, Ortigão, Albernaz, Bonamino, Aguiar, Alves e Sátyro (2007, p. 285), que relacionaram a escolha com “o caráter tipicamente escolar do conhecimento matemático, em comparação com o conhecimento da língua nativa, que, necessariamente é desenvolvido nos diversos

ambientes frequentados pelos jovens". A escolha por organizar e analisar os dados considerando as regiões administrativas paulista torna-se interessante, uma vez que os dados do Saresp não são divulgados desta maneira e, até onde conhecemos, a literatura não apresenta trabalhos que discutem resultados educacionais a partir desta classificação, possibilitando assim, o incentivo para propostas e reflexões de políticas públicas regionais.

Indicador educacional tipo Gini^v

Originalmente o índice de Gini foi utilizado para medir desigualdade na distribuição de renda e riqueza. Este índice, introduzido pelo estatístico italiano Corrado Gini (1912), é baseado na curva de Lorenz (1905).

A curva de Lorenz é o gráfico de uma função que associa a cada p no intervalo $[0,1]$ uma fração de um certo recurso (renda, riqueza, comida, terra, notas de uma avaliação, etc.) obtidos pelos $100p\%$ mais pobres de um grupo ou população. Ou seja, se a população está alinhada por ordem crescente das partes de seus bens em questão, então a cada p pertencente ao intervalo $[0,1]$ o número $L(p)$ é a fração da totalidade deste bem pertencente a fração p mais pobre da população. Seu gráfico situa-se abaixo da reta da função identidade que chamaremos de *reta de igualdade perfeita*. A curva de Lorenz seria, em teoria, a própria *reta de igualdade perfeita* $L(p) = p$ se todos recebessem a mesma fração de um certo recurso; veja Figura 1.



Fonte: Os autores (2022).

^v Os resultados desta seção foram apresentados e estão nos anais do XXXIX Congresso Nacional de Matemática Aplicada e Computacional - CNMAC, veja Capelato, Abreu (2020).

Se G denota o índice de Gini, ele será, por definição, a razão entre a área compreendida entre a curva de Lorenz e a reta de igualdade perfeita e a área abaixo desta reta, isto é,

$$G = 2 \int_0^1 (p - L(p)) dp \quad (1)$$

O valor de G é um número entre 0 e 1. Intuitivamente, quanto mais próxima a curva de Lorenz estiver da reta de igualdade perfeita mais igualitária é a distribuição do recurso; isso reflete numericamente numa menor área situada entre ambas as curvas, portanto num menor índice de Gini. Assim, G é tanto menor quanto mais igualitária é a distribuição de um recurso e tanto maior quanto menos igualitária for a distribuição deste mesmo recurso.

Através de uma conexão com a teoria de probabilidades, Farris (2010) introduz um certo parâmetro, denominado percentil do dólar médio, que aponta novas interpretações para o índice de Gini, permitindo aprofundar o entendimento do índice como medida de desigualdade a partir de uma outra perspectiva. O percentil do dólar médio descreve a posição ocupada pela pessoa que recebe o dólar médio de toda renda distribuída, ou seja, o dólar sorteado do montante de todo o recurso. Em seu artigo, o autor mostrou que a relação entre o índice de Gini e o percentil do dólar médio, indicado por p , é dado pela equação

$$p = \frac{G + 1}{2}.$$

Segundo dados do Banco Mundial, o Índice de Gini relativo à distribuição de renda no ano de 2015 é igual a 0,29 para a Suécia e 0,51 para o Brasil. Assim, o dólar médio na Suécia pertence à pessoa que ocupa a posição 65º e, no Brasil, à pessoa da posição 76º.

Inspirados no trabalho de Farris (2010), propomos um indicador para medir o desempenho escolar, ou seja, a habilidade que um grupo de pessoas demonstram numa avaliação cujo resultado é dado em forma de uma única nota.

Suponha que N indivíduos realizem um exame com quantidade total de M pontos. Sejam $0 = x_0 < x_1 < \dots < x_n = M$ uma partição do intervalo $[0, M]$ o qual corresponde a faixas de notas distribuídas a priori. Suponha que agrupemos os indivíduos em faixas, de modo que h_1 indivíduos tenham obtido notas no intervalo $[0, x_1)$, h_2 indivíduos no intervalo $[x_1, x_2)$, e, mais geralmente, h_i indivíduos no intervalo $[x_{i-1}, x_i)$. Além disso, suponha que a nota média no intervalo $[x_{i-1}, x_i)$ é β_i .

Desta forma, $N = \sum_{i=1}^n h_i$ é o número total de indivíduos que realizaram a avaliação e $\tau = M \cdot N$, o valor total de pontos que o exame poderia "distribuir". Seja $T = \sum_{i=1}^n \beta_i h_i$, o número total de pontos obtido pela totalidade dos indivíduos que fizeram o exame, o qual é uma fração do valor total τ . Além disso, com esta notação, temos o valor médio das notas obtidas na avaliação T/N o qual vamos denotar por \bar{X} .

Os números

$$p_j = \frac{1}{N} \sum_{i=1}^j h_i$$

São os n pontos (não necessariamente igualmente espaçados) ao longo do eixo horizontal entre 0 e 1, representando os percentis em ordem crescente de desempenho na avaliação. Por conveniência, definimos $p_0 = 0$.

O análogo da curva de Lorenz é a função L definida para estes valores particulares como: $L(p_0) = 0$ e

$$L(p_j) = \frac{1}{\tau} \sum_{i=1}^j \beta_i h_i \quad (1 \leq j \leq n). \quad (2)$$

Uma vez que esta representa a fração do total de pontos recebido pela fração p_i de indivíduos que receberam as menores notas no exame.

Com as notações apresentadas, podemos escrever a Tabela 1. A i -ésima linha contém informações do intervalo $[x_{i-1}, x_i)$ para $i = 1, \dots, n$ e p_j com $j = 1, \dots, n$.

Tabela 1 – Notas em exame com valores agregados

Intervalo de notas	Número de indivíduos no intervalo	Nota média no intervalo	p_j	$L(p_j)$
$[x_0, x_1)$	h_1	β_1	h_1/N	$(h_1 \cdot \beta_1)/\tau$
$[x_1, x_2)$	h_2	β_2	$(h_1 + h_2)/N$	$(h_1 \cdot \beta_1 + h_2 \cdot \beta_2)/\tau$
...
$[x_{i-1}, x_i)$	h_i	β_i	$(\sum_{k=1}^i h_k)/N$	$(\sum_{k=1}^i \beta_k h_k)/\tau$
...
$[x_{n-1}, x_n]$	h_n	β_n	1	T/τ

Fonte: Os autores (2022).

Diferentemente do caso da distribuição de renda, onde todos os recursos são distribuídos, aqui apenas parte de todos os recursos são distribuídos; na verdade, o recurso está disponível, mas deve ser “conquistado”, e sua distribuição completa não é obrigatória.

A maneira mais simples de preencher o resto da curva de Lorenz é por interpolação linear (FARRIS, 2010, p. 855). Claramente isto resulta em uma subestimação do índice de Gini, embora pequena se a agregação usar uma partição fina.

Sabemos que é um pouco complicado calcular o índice de Gini dado pela equação (1) por integração direta da equação (2). Em vez disso, vamos reinterpretar a equação (2) descobrindo uma conexão com a função de densidade de probabilidade.

A equação (2) pode ser reescrita como uma soma de Riemann da seguinte forma

$$L(p_j) = \frac{1}{\tau} \sum_{i=1}^j \beta_i h_i = \frac{1}{\tau} \sum_{i=1}^j \beta_i (p_i - p_{i-1}) N = \int_0^{p_j} \sigma(p) dp, \quad (3)$$

onde $\sigma(p) = \beta_j/M$ sempre que $p_{j-1} < p \leq p_j$.

Assim, é razoável usar a definição

$$L(p) = \int_0^p \sigma(q) dq, \quad (4)$$

Inicialmente definida no ponto p_j , para todo p , através de interpolação linear como mencionamos anteriormente. A natureza não decrescente de $\sigma(p)$ estabelece L como uma função não decrescente e convexa.

Note que

$$\int_0^1 \sigma(p) dp = L(1) = \frac{1}{\tau} \sum_{i=1}^n \beta_i h_i = \frac{T}{\tau}. \quad (5)$$

Tomando $s(p) = (\tau/T)\sigma(p)$ temos, para $p_{j-1} < p \leq p_j$

$$s(p) = \frac{\beta_i}{\bar{X}}. \quad (6)$$

Assim, a função s nos diz que parte da média dos pontos da avaliação pertence à porção de indivíduos no intervalo de percentil (p_{j-1}, p_j) .

Além disso, a função s satisfaz $s(p) \geq 0$ e $\int_0^1 s(p) dp = 1$, isto é, s é uma função densidade de probabilidade. O experimento que produz uma variável aleatória com esta função densidade é o seguinte: suponha que escolhamos um ponto aleatoriamente a partir do conjunto que contém todos os pontos atribuídos a todos os indivíduos submetidos ao exame. Registre o valor do percentil, p , da pessoa a que recebeu aquele ponto. Para este experimento, p é uma variável aleatória com densidade $s(p)$. Para ver isso, veja a equação (6). Por exemplo, a probabilidade de que um ponto escolhido aleatoriamente pertença a um indivíduo na faixa de percentil de p_{j-1} a p_j é exatamente a fração desses pontos em proporção ao todo, que é

$$\frac{\beta_i h_i}{T} = \frac{\tau}{T} \frac{\beta_i}{M} (p_i - p_{i-1}) = \int_{p_{i-1}}^{p_i} s(p) dp. \quad (7)$$

Seja \bar{p} o valor esperado da variável aleatória $p \in [0,1]$ com função densidade de probabilidade s ; isto é,

$$\bar{p} = \int_0^1 p \cdot s(p) dp. \quad (8)$$

Chamamos \bar{p} o percentil do ponto médio. Este parâmetro se comporta da seguinte forma: trata-se de um número entre 0,5 e 1; quanto mais bem distribuídos forem os pontos, mais este parâmetro estará próximo de 0,5 e quanto menos bem distribuídos forem os pontos, mais este parâmetro estará próximo de 1. Por exemplo, considere dois grupos de 100 alunos que se submeteram a um exame, cujos valores para \bar{p} sejam, respectivamente, iguais a 0,6 e 0,8; a definição de \bar{p} sugere que a média dos pontos (ou o ponto típico sorteado do montante de pontos) pertenceria ao indivíduo localizado na 60ª posição do primeiro grupo e ao indivíduo na 80ª posição do segundo grupo. Em particular, se ambos os grupos obtiveram desempenho aproximadamente equivalente em números absolutos (ou seja, valores de T aproximadamente iguais), fica evidente que o desempenho individual é mais homogêneo no primeiro grupo; evidentemente, isto é mais desejável do que o caso em que um pequeno número de indivíduos dentro do grupo apresenta um desempenho individual muito superior aos demais.

Considerando todas as notações e definições dadas, a seguir apresentamos os dois resultados principais que demonstram a metodologia do indicador educacional tipo Gini que estamos propondo:

Resultado 1. O índice de Gini G e o percentil do ponto médio \bar{p} estão relacionados por

$$\bar{p} = 1 - \frac{1 - G}{2L(1)}. \quad (9)$$

Demonstração. Substituindo a equação (4) na equação (1) temos que

$$G = 2 \int_0^1 (p - L(p)) dp = 2 \int_0^1 p dp - 2 \int_0^1 \int_0^p \sigma(q) dq dp = 1 - 2 \int_0^1 \int_q^1 \sigma(q) dp dq$$

$$\begin{aligned}
&= 1 - 2 \int_0^1 (1-q) \sigma(q) dq = 1 - 2 \int_0^1 \sigma(q) dq + 2 \int_0^1 q \sigma(q) dq \\
&= 1 - 2L(1) + 2 \int_0^1 q \frac{T}{\tau} s(q) dq = 1 - 2L(1) + 2L(1) \int_0^1 q s(q) dq
\end{aligned}$$

Como a última integral da igualdade acima é o valor esperado definido na equação (8) temos

$$G = 1 - 2L(1) + 2L(1)\bar{p} \quad (10)$$

isolando \bar{p} na equação (10) vemos que esta é equivalente à equação (9), como queríamos demonstrar.

Resultado 2. O índice de Gini G pode ser calculado a partir do conjunto de dados como

$$G = 1 - 2L(1) + \frac{1}{\tau} \sum_{i=1}^n \beta_i (p_{i-1} - p_i) h_i. \quad (11)$$

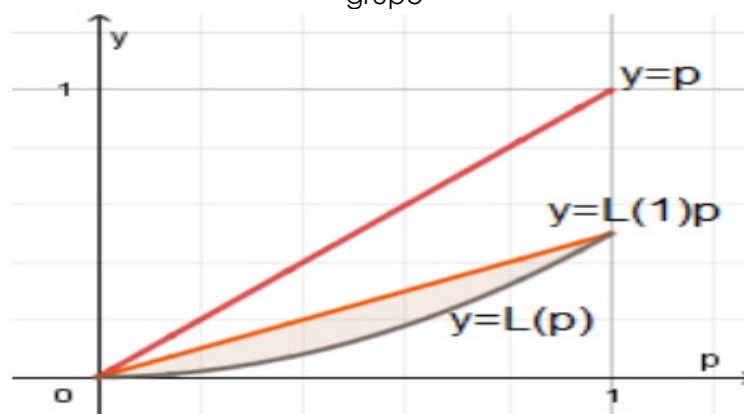
Demonstração. A demonstração da equação (11) se faz observando que \bar{p} , definido na equação (8), pode ser calculado diretamente quebrando a integral em subintervalos e usando a definição de s , ou seja,

$$\begin{aligned}
\bar{p} &= \int_0^1 p \cdot s(p) dp = \int_0^1 p \frac{\tau}{T} \sigma(p) dp = \frac{1}{L(1)} \int_0^1 p \sigma(p) dp = \frac{1}{L(1)} \sum_{i=1}^n \int_{p_{i-1}}^{p_i} p \sigma(p) dp \\
&= \frac{1}{L(1)} \sum_{i=1}^n \frac{\beta_i}{M} \int_{p_{i-1}}^{p_i} p dp = \frac{1}{L(1)} \sum_{i=1}^n \frac{\beta_i}{M} \frac{p_i^2 - p_{i-1}^2}{2} = \frac{1}{L(1)} \sum_{i=1}^n \frac{\beta_i}{M} \frac{p_i - p_{i-1}}{2} (p_i + p_{i-1}) \\
&= \frac{1}{2\tau L(1)} \sum_{i=1}^n \beta_i (p_i + p_{i-1}) h_i.
\end{aligned}$$

Substituindo este valor encontrado na equação (10) obtemos a equação (11).

Da equação (5) temos $L(1) = T/\tau$. Este valor trata-se de um número entre 0 e 1 que mede o desempenho (rendimento) do grupo. A menos que todas as pessoas obtenham nota máxima no exame, o valor $L(1)$ é menor que 1. Assim, o melhor que se pode esperar em termos de equidade é a reta $y = L(1)p$, que chamamos de reta de equidade intrínseca do grupo, pois isto significa que todos tiveram a mesma nota, veja Figura 2.

Figura 2 – Curva de Lorenz, reta de igualdade perfeita e reta de equidade intrínseca do grupo



Fonte: Os autores (2022).

Desta forma, podemos obter o que podemos chamar de índice de Gini intrínseco, o qual será dado, calculando a área entre as curvas $y = L(1)p$ e $y = L(p)$, por

$$g = 1 - \frac{1 - G}{L(1)}.$$

Observe que o valor de g varia entre 0 e 1 e sua interpretação coincide com a de G . Ainda que seja interessante olhar para o índice de Gini intrínseco, o fato de sua relação com G faz-nos considerar, para os cálculos do indicador educacional do tipo Gini, o seguinte:

- Rendimento do grupo de indivíduos que realiza a avaliação dado pela equação (5):

$$L(1) = T/\tau.$$

- Equidade na distribuição das notas dado pela equação (11):

$$G = 1 - 2L(1) + \frac{1}{\tau} \sum_{i=1}^n \beta_i (p_{i-1} - p_i) h_i.$$

A metodologia descrita nesta seção, mostra que, o melhor desempenho em uma avaliação ocorrerá sempre que tivermos, simultaneamente, o valor de $L(1)$ próximo de 1 (um) e o valor de G próximo de 0 (zero).

Indicador educacional tipo Gini como medida de desempenho em matemática nas regiões paulistas

Nesta seção, vamos usar a base de dados do Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (Saresp). O Saresp foi instituído pela Secretaria de Educação do Estado de São Paulo (SEE) através da Resolução da Secretaria de Educação Nº 27 de março de 1996^{vi}, tendo dois objetivos: (i) desenvolver um sistema de avaliação de desempenho dos alunos do ensino fundamental e médio das redes estadual e municipal que subsidie a Secretaria da Educação na tomada de decisão quanto à Política Educacional do Estado e (ii) verificar o desempenho dos alunos nos diferentes componentes curriculares (Português, Matemática, Ciências, História e Geografia), de modo a fornecer ao sistema de ensino, às equipes técnico-pedagógicas das Diretorias de Ensino e às Unidades Escolares informações que subsidiem a capacitação dos recursos humanos do magistério, a reorientação da proposta pedagógica e a articulação dos resultados da avaliação com o planejamento escolar.

A escala de proficiência do Saresp foi construída de forma empírica e comparativa a partir da adoção de uma distribuição de referência para as proficiências dos alunos. A referência foi a média do comportamento dos alunos de alguns países da Organização para Cooperação e Desenvolvimento Econômico (OCDE). Soares (2009) apresenta com detalhes como foi construída esta escala de proficiência.

^{vi} Disponível em http://siau.edunet.sp.gov.br/ItemLise/arquivos/27_1996.htm. Acesso em 12-05-2022.

Na Tabela 2 apresentamos os níveis de proficiência de matemática do Saresp que reúnem informações sobre o intervalo de pontuação, tanto para o Ensino Fundamental (EF), quanto para o Ensino Médio (EM).

Tabela 2 - Níveis de proficiência em Matemática - Saresp

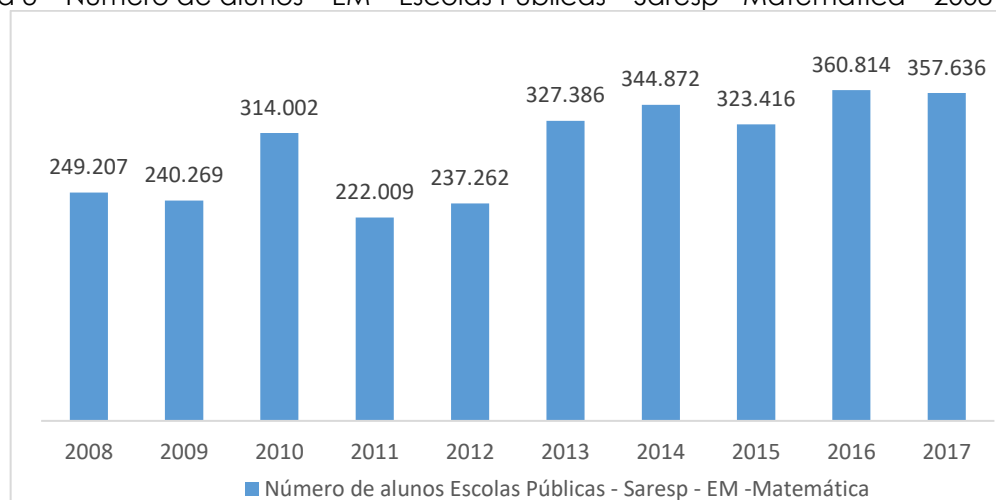
Níveis de proficiência	3º EF	5º EF	7º EF	9º EF	3ª EM
Abaixo do básico	<150	<175	<200	<225	<275
Básico	150 a <200	175 a <225	200 a <250	225 a <300	275 a <350
Adequado	200 a <250	225 a <250	250 a <300	300 a <350	350 a <400
Avançado	≥ 250	≥ 250	≥ 300	≥ 350	≥ 400

Fonte: SÃO PAULO (2019).

São Paulo (2019) aponta que o Saresp estabelece como padrão de desempenho esperado o nível adequado para cada um dos anos/séries e componentes curriculares avaliados. É importante mencionar que a proficiência dos alunos, tanto em Língua Portuguesa, quanto em Matemática, ancora-se nas mesmas escalas do Sistema de Avaliação Educacional Brasileiro (Saeb) portanto, os resultados obtidos pelos alunos paulistas nos dois exames tornam-se passíveis de comparação.

A SEE disponibiliza os resultados educacionais através dos micro dados de alunos do Saresp^{vii}. A Figura 3 mostra o número total de alunos da 3ª série do EM da rede pública na prova de Matemática do Saresp em cada ano do período.

Figura 3 – Número de alunos – EM – Escolas Públicas – Saresp - Matemática – 2008 a 2017



Fonte: Os autores (2022).

^{vii} Dados disponíveis em: <https://www.educacao.sp.gov.br/>. Acesso em Agosto de 2019.

De acordo com a Figura 3, podemos observar que o número de alunos foi menor no ano de 2011 e a maior participação do período ocorreu no ano de 2016.

Para que pudéssemos analisar o desempenho em matemática de forma regional, sistematizamos os dados do Saresp de acordo com as regiões paulistas estabelecidas pela Fundação Sistema Estadual de Análise de Dados (Seade). De acordo com o Seade, a divisão regional tem vistas à organização de atividade de planejamento, de prestação de serviços públicos e/ou gestão de políticas e programas públicos de responsabilidade compartilhada, quer entre Estado e municípios, quer entre municípios.

Atualmente, o Estado de São Paulo possui 645 municípios. Cada município está localizado ou em uma Região Administrativa (RA) ou na Região Metropolitana (RM) de São Paulo. As Regiões Administrativas são as seguintes: Araçatuba; Barretos; Bauru; Campinas; Central; Franca; Itapeva; Marília; Presidente Prudente; Registro; Ribeirão Preto; Santos; São José dos Campos; São José do Rio Preto e Sorocaba. Na Figura 4, construímos com o *software Qgis* um mapa do Estado de São Paulo, o qual foi elaborado a partir do agrupamento dos municípios que compõem cada RA e a RM de São Paulo.

Figura 3 – Divisão regional do Estado de São Paulo



Fonte: Os autores (2022).

A quantidade de alunos da 3ª série do EM na prova de Matemática do Saresp em cada região paulista está apresentada na Figura 4.

Figura 4 – Número de alunos por região paulista – Saresp – EM - Matemática

Região	Quantidade de alunos por ano no período 2008 - 2017									
	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Araçatuba	5720	5200	5476	4594	4370	5291	5230	4851	6055	5729
Barretos	2924	3346	2168	2645	2524	3262	3219	3349	3640	3188
Bauru	5933	6494	6746	5973	6218	7864	7872	6866	7948	7477
Campinas	36793	36776	48226	35651	35572	49893	51109	49707	53137	51146
Central	5045	4321	6350	2784	4808	6536	6164	5928	6838	6803
Franca	4233	5401	5113	4813	4171	5911	6128	6054	6693	6770
Itapeva	4866	4747	4216	3359	4131	5307	5589	5464	5838	5333
Marília	6435	5581	7983	6406	5453	8076	8043	6968	8234	8315
Presidente Prudente	7627	6283	6993	5908	6238	7107	7361	6906	7317	6855
Registro	2213	2349	2708	1651	2331	2736	2867	2990	3176	3070
Ribeirão Preto	6440	6158	10660	5945	6212	8710	8974	8904	9284	8083
Santos	8431	8477	11847	8414	8725	11728	12634	12470	12665	13229
Sorocaba	14308	14455	20519	14677	14042	19537	20628	18419	22030	23094
São José do Rio Preto	9410	9077	11898	8954	8999	11037	11522	10582	11709	11281
São José dos Campos	15163	12111	19718	12786	14721	18432	19782	19589	20173	20312
RM de São Paulo	113665	109493	143381	97449	108747	155959	165736	154369	176077	176951

Fonte: Os autores (2022).

Da Figura 4 é possível observar que a quantidade de alunos é diferente entre as regiões paulistas. Registro, exceto para o ano de 2010, é a menor região em números absolutos em cada ano do período, seguida de Barretos. Já a RM de São Paulo, considerando cada ano do período, é a que possui a maior participação em números absolutos seguida de Campinas.

Para calcularmos o valor do indicador educacional tipo Gini, precisamos conhecer o total de pontos da prova de Matemática do Saresp. Dada a especificidade da metodologia da Teoria de Resposta ao Item, na qual o Saresp é formulado, não é possível conhecer a nota máxima da avaliação, pois esta depende das características dos itens selecionados. Assim, adotamos para este valor a pontuação 454, a qual foi escolhida por ser a maior pontuação encontrada nos microdados no período analisado.

Conforme apresentamos na base metodológica do indicador, para fazermos o seu cálculo é necessário a organização da variável nota em uma tabela de distribuição de frequências, conforme a Tabela 1 apresentada neste texto. Para construir este tipo de tabela, seguimos o Critério de Sturges. Este critério pode ser encontrado em Toledo e Ovalle (1985).

Na Tabela 3 apresentamos os valores de $L(1)$ e os valores de G , para cada região paulista e para cada ano do período analisado.

Tabela 3 – Resultados dos indicadores

Valores de $L(1)$ e G no período 2008 - 2017											
Regiões		2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Araçatuba	G	0,44	0,45	0,45	0,43	0,44	0,43	0,44	0,41	0,42	0,42
	$L(1)$	0,62	0,61	0,60	0,64	0,62	0,61	0,62	0,65	0,64	0,64
Barretos	G	0,44	0,46	0,46	0,45	0,45	0,45	0,45	0,42	0,41	0,42
	$L(1)$	0,61	0,61	0,59	0,60	0,61	0,60	0,61	0,64	0,64	0,63
Bauru	G	0,44	0,45	0,45	0,45	0,44	0,45	0,45	0,42	0,43	0,43
	$L(1)$	0,62	0,60	0,61	0,61	0,62	0,61	0,61	0,64	0,63	0,63
Campinas	G	0,44	0,45	0,45	0,44	0,44	0,44	0,44	0,42	0,42	0,42
	$L(1)$	0,62	0,61	0,60	0,61	0,61	0,62	0,62	0,64	0,63	0,63
Central	G	0,44	0,44	0,44	0,44	0,44	0,44	0,43	0,41	0,42	0,42
	$L(1)$	0,62	0,62	0,62	0,61	0,62	0,62	0,63	0,65	0,64	0,64
Franca	G	0,44	0,45	0,45	0,44	0,45	0,45	0,45	0,41	0,43	0,42
	$L(1)$	0,62	0,61	0,60	0,61	0,61	0,61	0,61	0,64	0,63	0,63
Itapeva	G	0,44	0,45	0,52	0,44	0,44	0,44	0,43	0,40	0,42	0,41
	$L(1)$	0,62	0,61	0,51	0,61	0,62	0,62	0,62	0,65	0,64	0,64
Marília	G	0,43	0,46	0,46	0,45	0,45	0,46	0,45	0,42	0,43	0,43
	$L(1)$	0,63	0,60	0,60	0,61	0,60	0,60	0,61	0,64	0,63	0,63
Presidente Prudente	G	0,44	0,45	0,45	0,45	0,45	0,45	0,45	0,41	0,43	0,42
	$L(1)$	0,62	0,61	0,60	0,61	0,61	0,61	0,61	0,64	0,63	0,63
Registro	G	0,45	0,46	0,46	0,45	0,45	0,46	0,45	0,42	0,43	0,43
	$L(1)$	0,61	0,60	0,60	0,61	0,65	0,60	0,60	0,63	0,62	0,63
Ribeirão Preto	G	0,44	0,45	0,45	0,45	0,44	0,45	0,45	0,43	0,43	0,43
	$L(1)$	0,61	0,61	0,60	0,60	0,62	0,61	0,62	0,63	0,63	0,63
Santos	G	0,45	0,46	0,46	0,46	0,46	0,47	0,46	0,43	0,44	0,44
	$L(1)$	0,60	0,59	0,59	0,59	0,59	0,59	0,59	0,62	0,62	0,61
Sorocaba	G	0,44	0,46	0,46	0,45	0,45	0,46	0,45	0,42	0,42	0,41
	$L(1)$	0,61	0,60	0,60	0,60	0,61	0,60	0,61	0,64	0,64	0,64
São José do Rio Preto	G	0,44	0,44	0,45	0,44	0,44	0,44	0,43	0,40	0,41	0,41
	$L(1)$	0,62	0,62	0,61	0,62	0,62	0,63	0,63	0,66	0,65	0,65
São José dos Campos	G	0,44	0,45	0,45	0,45	0,45	0,45	0,45	0,42	0,42	0,39
	$L(1)$	0,61	0,60	0,60	0,61	0,61	0,61	0,61	0,64	0,63	0,63
São Paulo	G	0,46	0,47	0,47	0,47	0,47	0,46	0,47	0,44	0,45	0,45
	$L(1)$	0,60	0,59	0,58	0,58	0,59	0,59	0,59	0,61	0,60	0,61

Fonte: Os autores (2022).

Considerando que o melhor desempenho ocorre quando a região apresenta o maior valor para $L(1)$ e o menor valor para G , simultaneamente, observamos da Tabela 3 que as Regiões Administrativas de Araçatuba, Bauru, Campinas, Central, Franca, Itapeva, Marília, Presidente Prudente, Registro, Ribeirão Preto, Santos, São José do Rio Preto e a RM de São Paulo apresentaram o melhor desempenho em

Matemática no ano de 2015. Para a RA de Barretos, o melhor desempenho ocorreu no ano de 2016 e para a RA de Sorocaba em 2017.

Diferente das demais regiões paulistas, onde o maior valor de $L(1)$ ocorreu juntamente com o menor valor de G , a RA de São José dos Campos apresentou, no ano de 2015, os seguintes valores $L(1) = 0,64$ e $G=0,42$. Já para o ano de 2017, $L(1) = 0,63$ e $G=0,39$. No entanto, como os valores de $L(1)$ (desempenho absoluto do grupo de alunos) estão próximos, consideramos que o melhor desempenho em Matemática ocorreu no ano de 2017, uma vez que o indicador aponta que as notas foram distribuídas de forma mais equitativa.

Ainda de acordo com a Tabela 3 podemos observar que a RA de São José do Rio Preto esteve, em 90% do período analisado, com os melhores resultados do indicador educacional tipo Gini. Do lado oposto, a RM de São Paulo teve o pior desempenho em Matemática em 90% do período analisado.

Ainda que de forma tímida, a desigualdade no desempenho em Matemática entre as diferentes regiões paulistas pode ser verificada pelos resultados apresentados na Tabela 3. No ano de 2008 ocorreu o maior percentual de regiões com o mesmo desempenho em Matemática. Neste ano, 50% das regiões administrativas obtiveram os mesmos valores para $L(1)$ e G . No ano de 2014 o percentual de regiões com os mesmo desempenho foi de, aproximadamente, 44%. Nos anos de 2009, 2010, 2012 e 2015, este percentual foi de 37,5%. Para os demais anos do período, a semelhança entre o desempenho em Matemática nas regiões foi menor que 31%.

De modo geral, os valores da Tabela 3 mostram pouca evolução no valor do indicador educacional do tipo Gini, uma vez que o crescimento de $L(1)$ e a redução de G não ocorre, simultaneamente, de forma significativa no período. Este fato está em consonância com o cenário apresentado por outros indicadores.

Por exemplo, em Todos pela Educação (2020) os autores observaram, usando os dados do Saeb, que quando comparado com as outras etapas da Educação Básica, os baixos níveis de aprendizado dos alunos brasileiros são mais alarmantes no Ensino Médio e em Matemática. Observam ainda que em 2007 apenas 9,8% dos alunos do Ensino Médio possuíam nível adequado de aprendizagem em Matemática e, em 2017, este percentual reduziu para 9,1%.

Em avaliações internacionais, como é o caso do *Programme for International Student Assessment* (Pisa), a baixa aprendizagem dos estudantes do Ensino Médio em

Matemática também é observada. De acordo com INEP (2020), em 2009 a proficiência média em Matemática do Brasil no PISA foi de 386, em 2012, 2015 e 2018 foi de 389, 377 e 384, respectivamente, o que mostra que a aprendizagem em Matemática ficou estagnada neste período de dez anos. Considerando os países da América do Sul participantes da avaliação do Pisa em 2018, o Brasil ocupa, junto com a Argentina, o último lugar no *ranking* considerando a proficiência média em Matemática.

Uma pergunta natural a ser feita é a seguinte: como o desempenho em matemática se relaciona com variáveis socioeconômicas? Na próxima seção nos debruçamos na tentativa de responder a esta pergunta.

Relação entre o desempenho em matemática e variáveis socioeconômicas

Para descrever o perfil demográfico e econômico das regiões paulistas, vamos utilizar as seguintes variáveis: valores do indicador, densidade demográfica, produto interno bruto *per capita* (PIB *per capita*) e rendimento médio total dos empregos formais. Para a construção deste banco de dados para cada região paulista e, para o período de 2008 a 2017, utilizamos os dados disponíveis na Fundação Seade.

A análise da relação entre o desempenho em matemática (mensurado pelo indicador educacional tipo Gini, G e $L(1)$) e variáveis socioeconômicas, se deu através dos seguintes modelos de regressão linear utilizando-se o método dos Mínimos Quadrados Ordinários (MQO):

$$G = \beta_0 + \beta_1 \ln(\text{PIB_pc}) + \beta_2 \ln(\text{dens_demo}) + u \quad (\text{Modelo 1})$$

$$G = \beta_0 + \beta_1 \ln(\text{rend_med}) + \beta_2 \ln(\text{dens_demo}) + u \quad (\text{Modelo 2})$$

$$L(1) = \beta_0 + \beta_1 \ln(\text{PIB_pc}) + \beta_2 \ln(\text{dens_demo}) + u \quad (\text{Modelo 3})$$

$$L(1) = \beta_0 + \beta_1 \ln(\text{rend_med}) + \beta_2 \ln(\text{dens_demo}) + u \quad (\text{Modelo 4})$$

onde $\ln(\text{PIB_pc})$ denota o logaritmo do produto interno bruto *per capita*; $\ln(\text{rend_med})$ o logaritmo da renda média total dos empregos formais; $\ln(\text{dens_demo})$ o logaritmo da densidade demográfica; u é o termo de erro aleatório que contém os fatores não observados que afetam o indicador educacional e, β_i , com $i = 0, 1, 2$, os valores a serem estimados.

Modelos onde as três variáveis explicativas (PIB_{pc} , $rend_{med}$ e $dens_{demo}$) estivessem juntas não foram considerados pois teriam colinearidade elevada, uma vez que as variáveis PIB_{pc} e $rend_{med}$ possuem alta correlação. Ainda que o modelo MQO seja simples podemos observar que ele se mostrou eficaz para esta pesquisa e ainda é bastante utilizado na literatura (ALVES; SOARES, 2013; HELENE; MARIANO, 2020; SILVA; HASENBALG, 2000; MONTEIRO, 2015; PASSADOR; CALHADO, 2012; TRAVITZKI, 2017).

As estimativas dos modelos de regressão linear foram obtidas com o uso do software *Gretl*. Na Tabela 5 apresentamos os resultados dos modelos.

Tabela 5 – Resultado da regressão

MQO agrupado, usando 154 observações Incluídas 16 unidades de corte transversal Comprimento da série temporal: mínimo 4, máximo 10				
Modelo 1				
	Coefficiente	Erro padrão	Razão-t	p-valor
β_0	0,555402	0,0694041	8,002	0,0000
$\ln(dens_{demo})$	0,00663700	0,00143859	4,614	0,0000
$\ln(PIB_{pc})$	-0,0138874	0,00704428	-1,971	0,0505
R^2	0,138993		R^2 ajustado	0,127589
$F(2, 151)$	12,18801		p-valor(F)	0,000012
Modelo 2				
	Coefficiente	Erro padrão	Razão-t	p-valor
β_0	0,711142	0,0993743	7,156	0,0000
$\ln(dens_{demo})$	0,0101000	0,00210906	4,789	0,0000
$\ln(rend_{med})$	-0,0406009	0,0137905	-2,944	0,0038
R^2	0,164776		R^2 ajustado	0,153713
$F(2, 151)$	14,89491		p-valor(F)	1,25e-06
Modelo 3				
	Coefficiente	Erro padrão	Razão-t	p-valor
β_0	0,471873	0,0668902	7,054	0,0000
$\ln(dens_{demo})$	-0,00784280	0,00138648	-5,657	0,0000
$\ln(PIB_{pc})$	0,0172029	0,00678913	2,534	0,0123
R^2	0,192369		R^2 ajustado	0,181672
$F(2, 151)$	17,98332		p-valor(F)	9,8e-08
Modelo 4				
	Coefficiente	Erro padrão	Razão-t	p-valor
β_0	0,324990	0,0958926	3,389	0,0000
$\ln(dens_{demo})$	-0,0112744	0,00203517	-5,540	0,0000
$\ln(rend_{med})$	0,0438977	0,0133073	3,299	0,0012
R^2	0,138993		R^2 ajustado	0,127589
$F(2, 151)$	12,18801		p-valor(F)	0,000012

Fonte: Os autores (2022).

Comparando os valores dos Modelos 1 e 2 vemos que as variáveis explicativas do Modelo 2 são mais significativas. Por outro lado, ao compararmos os Modelos 3 e 4, vemos que as variáveis explicativas do Modelo 4 são mais significativas. Isto significa que as variáveis explicativas, densidade demográfica e rendimento médio total dos empregos formais, são as que melhores se ajustam ao modelo estimado pelo MQO, tanto quando a variável dependente é o indicador educacional tipo G , representado pelo Modelo 2, como quando a variável dependente é o desempenho do grupo de alunos $L(1)$, representado pelo Modelo 4.

Em resumo, a análise do Modelo 2 mostra que a equidade na distribuição das notas piora quando a densidade demográfica aumenta; porém, quando o rendimento médio total dos empregos formais aumenta a equidade na distribuição das notas melhora. Por exemplo, os valores mostram que um aumento de 10% no valor PIB per capita reduziria, aproximadamente 0,0014 do valor do indicador G . Um aumento de 10% na densidade demográfica aumentaria, aproximadamente 0,0007 do indicador G . O Modelo 4 mostra que, se a densidade demográfica aumentar o rendimento de todo o grupo diminui e, quando o rendimento médio total dos empregos formais aumentar, o desempenho do grupo também aumenta. Por exemplo, os valores mostram que um aumento de 10% na densidade demográfica diminui aproximadamente 0,001 do valor de $L(1)$ e, um aumento de 10% no valor do rendimento médio total de empregos formais aumenta aproximadamente 0,004 o valor de $L(1)$.

Os resultados obtidos, no que se refere ao sinal das variáveis explicativas, rendimento médio total dos empregos formais e PIB per capita, estão em consonância com teoria econômica que descreve a influência do capital humano sobre o crescimento econômico e o aumento da renda per capita. Esta teoria foi abordada em Becker (1964) e Mincer (1974), trabalhos pioneiros sobre o assunto.

Em relação ao sinal negativo da variável explicativa densidade demográfica, podemos dizer que se justifica a partir da conjectura que foi apresentada no trabalho Silva e Almeida (2012): as regiões mais densamente povoadas podem apresentar, uma maior demanda com educação e dificuldade com a gestão de recursos públicos destinados à educação.

De acordo com estimativas populacionais do ano de 2021 do Instituto Brasileiro de Geografia e Estatística (IBGE), a população do Estado de São Paulo é de

46.649.132^{viii} pessoas, sendo superior, tanto às populações dos outros estados brasileiros, quanto às populações de diversos países. Desta forma, isolando outros fatores, o Estado de São Paulo pode se enquadrar na conjectura de Estado com dificuldade na gestão de recursos públicos destinados à educação.

De fato, recentemente, foi publicado o artigo Alves, Araújo, Frio e Cordeiro (2020) que mostra a ineficiência do Estado de São Paulo na aplicação dos recursos financeiros para melhorar a educação. Usando dados do Índice de Desenvolvimento da Educação Básica (IDEB) os autores mostram que os resultados educacionais do IDEB para o Ensino Médio não acompanharam o crescimento dos gastos públicos aplicados nesta etapa, ou seja, apesar do estado de São Paulo sair do 6º para o 4º lugar na classificação do IDEB entre os anos de 2007 e 2017, o Estado ocupa a 25ª posição no ranking de eficiência do Ensino Médio, ficando à frente apenas do Rio Grande do Norte e da Bahia.

De forma geral, os resultados que obtivemos oferecem evidências de que o indicador educacional tipo Gini, quando relacionado com variáveis socioeconômicas, reforça a mesma assertiva obtida por outros trabalhos empíricos: é necessário melhorar as variáveis econômicas, na tentativa de ampliar o desempenho escolar (VIANA; LIMA, 2010; BARBOSA FILHO; PESSÔA, 2010; HELENE; MARIANO, 2020; FIRME, 2020).

Considerações finais

O indicador educacional proposto neste trabalho consegue mensurar a qualidade da educação no que diz respeito a dois pontos: (i) o desempenho absoluto do grupo de estudantes em relação às notas em uma avaliação e (ii) as desigualdades educacionais no grupo de estudantes, estas desigualdades são referentes a análise da equidade na "distribuição" das notas entre os estudantes em uma avaliação.

Através da análise dos resultados obtidos podemos reforçar a assertiva de que o Estado de São Paulo possui um desafio para melhorar a qualidade e a equidade do aprendizado de Matemática dos alunos do Ensino Médio das escolas públicas.

Assim, parece-nos de grande importância proporcionar discussões entre educadores, gestores e administradores do sistema educacional brasileiro ao redor

^{viii} Informações disponíveis em ibge.gov.br/cidades-e-estados/sp.html. Acesso em 02-02-2022.

de indicadores educacionais como este, ou seja, indicadores que mostram as desigualdades no processo educacional. Estas reflexões poderão ensejar políticas educacionais pautadas, por exemplo, nas ideias de Crahay (2013) ao responder como a escola pode ser mais justa e mais eficaz. Em suas reflexões o autor defende, entre outros aspectos, que o ensino seja organizado para que os objetivos e resultados finais sejam alcançados por todos os alunos. Além disso, o autor chama atenção para a ideologia pedagógica justa, cujo princípio é a igualdade de conhecimento entre os alunos, para isso aponta que os procedimentos educacionais devem ser pensados e voltados aos alunos com dificuldades mais significativas.

Além disso, é desejável que as regiões administrativas paulistas promovam discussões ao redor da relação entre educação e desenvolvimento econômico e social garantindo, assim, a formulação de políticas públicas de igualdade de oportunidades.

Uma vez que os modelos econométricos podem ser utilizados para avaliar políticas públicas e ajudar no processo de tomada de decisões, apontamos que uma limitação deste trabalho reside no uso de apenas três variáveis socioeconômicas como determinantes do desempenho em Matemática. A escolha destas variáveis e a ausência de outras no modelo se deu devido à falta de uniformização dos dados abertos para as regiões paulistas para o período analisado. O uso de outras variáveis como, o capital econômico, usualmente mensurado pela renda ou riqueza familiar, a desigualdade na distribuição de renda, a qualidade da infraestrutura das escolas, entre outras, poderia trazer formulações mais abrangentes para explicar o desempenho em Matemática e isto é objeto de investigações futuras.

Por fim, é preciso destacar o limite dos indicadores educacionais quando analisados isoladamente. As suas medidas, quando associadas a estudos qualitativos, tem potencial para quantificar a qualidade e eficiência do sistema educacional e contribuir para a sua melhoria, algo que ultrapassa o objetivo desta contribuição. A contribuição do artigo está na proposição de um indicador educacional do tipo Gini e na sua validação através dos dados sobre o desempenho em matemática dos estudantes do Estado de São Paulo. Os resultados e reflexões aqui apresentados, consideramos, que são um primeiro olhar para as características deste indicador e que motivam novas reflexões e pesquisas para avaliar se ele é frutífero para as discussões no cenário da educação básica no país.

Agradecimentos

A primeira autora agradece o apoio financeiro da Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP. Os resultados apresentados neste texto são produtos da pesquisa registrada no Processo Fapesp nº 2018/03327-4.

Referências

- ALVES, P. J. H.; ARAÚJO, J. M.; FRIO, G. S.; CORDEIRO, L. C. Uma análise dos resultados educacionais dos estados brasileiros. *Estudos em Avaliação Educacional*, São Paulo, v. 31, n. 78, p. 742-775, 2020. DOI: <https://doi.org/10.18222/eae.v31i78.7073>. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/7073>. Acesso em: 2 fev. 2022.
- ALVES, M. T. G.; SOARES, J. F. Contexto escolar e indicadores educacionais: condições desiguais para a efetivação de uma política de avaliação educacional. *Educação e Pesquisa*, São Paulo, v. 39, n. 1, p. 177-194, 2013. DOI: <https://doi.org/10.1590/S1517-97022013000100012>. Disponível em: <https://www.scielo.br/j/ep/a/PkVXrTbnCJDktQxLZNK7dDj/?lang=pt>. Acesso em: 2 fev. 2022.
- BARBOSA FILHO, F. de H.; PESSÔA, S. de A. Educação e crescimento: o que a evidência empírica e teórica mostra. *Revista Economia*, Brasília, DF, v. 11, n. 2, p. 265-303, 2010. Disponível em: https://www.anpec.org.br/revista/vol11/vol11n2p265_303.pdf. Acesso em: 2 fev. 2022.
- BECKER, G. S. *Human capital theory*. New York: Columbia University Press, 1964.
- CAPELATO, E; ABREU, J. G. Proposta de um indicador do tipo Gini para medição do desempenho escolar. *Proceeding Series of the Brazilian Society of Computational and Applied Mathematics*, v. 7, n. 1, 2020. DOI: <https://doi.org/10.5540/03.2020.007.01.0421>.
- CRAHAY, M. Como a escola pode ser mais justa e mais eficaz?. *Cadernos Cenpec: Pesquisa e Ação Educacional*, São Paulo, v. 3, n. 1, p. 9-40, 2013. DOI: <http://dx.doi.org/10.18676/cadernoscenpec.v3i1.202>. Disponível em: <https://cadernos.cenpec.org.br/cadernos/index.php/cadernos/article/view/202>. Acesso em: 2 fev. 2022.
- FIRME, V. de A. C. Evidências empíricas do impacto da alfabetização sobre o crescimento econômico. *Estudos em Avaliação Educacional*, São Paulo, v. 31, n. 78, p. 577-625, 2020. DOI: <https://doi.org/10.18222/eae.v31i78.7452>. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/7452>. Acesso em: 2 fev. 2022.
- FRANCO, C.; ORTIGÃO, I.; ALBERNAZ, A.; BONAMINO, A.; AGUIAR, G.; ALVES, F.; SÁTYRO, N. Qualidade e equidade em educação: reconsiderando o significado de "fatores intraescolares". *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 15, n. 55, p. 277-298, 2007. Disponível em: http://educa.fcc.org.br/scielo.php?pid=S0104-40362007000200007&script=sci_abstract&tlng=pt. Acesso em: 2 fev. 2022.
- FARRIS, F. A. The Gini index and measures of inequality. *The American Mathematical Monthly*, Menasha, v. 117, n. 10, p. 851-864, 2010. DOI: <https://doi.org/10.4169/000298910x523344>. Disponível em:

<https://www.tandfonline.com/doi/abs/10.4169/000298910X523344>. Acesso em: 2 fev. 2022.

GINI, C. Variabilità e mutabilità. In: PIZETTI, E.; SALVEMINI, T. (ed.). *Reprinted in Memorie di metodologica statistica*. Rome: Libreria Eredi Virgilio Veschi, 1912.

HELENE, O.; MARIANO, L. Educação e desigualdade na distribuição de rendas. *Educação & Sociedade*, Campinas, SP, v. 41, p. 1-12, 2020. DOI: <https://doi.org/10.1590/ES.223485>. Disponível em: <https://www.scielo.br/j/es/a/S4zzTrcT95THQFK8CvzmX4F/?lang=pt>. Acesso em: 2 fev. 2022.

INEP. Diretoria de Avaliação da Educação Básica. *Brasil no Pisa 2018*. Brasília, DF: INEP, 2020. Disponível em: https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_exames_da_educacao_basica/relatorio_brasil_no_pisa_2018.pdf. Acesso em: 2 fev. 2022.

LORENZ, M. O. Methods of measuring the concentration of wealth. *Publications American Statistical Association*, Boston, v. 9, n. 70, p. 209-219, 1905. DOI: <https://doi.org/10.1080/15225437.1905.10503443>. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/15225437.1905.10503443>. Acesso em: 2 fev. 2022.

MARTÍNEZ-GARRIDO, C.; JAVIER MURILLO, F. Investigación iberoamericana sobre enseñanza eficaz. *Revista Mexicana de Investigación Educativa*, [Ciudad de Mexico], v. 21, n. 69, p. 471-499, 2016. Disponível em: https://www.scielo.org.mx/scielo.php?pid=S1405-66662016000200471&script=sci_abstract&tlng=pt. Acesso em: 2 fev. 2022.

MINCER, J. *Schooling, experience, and earnings: human behavior & social institutions*. Cambridge: National Bureau of Economic Research, 1974.

MONTEIRO, J. Gasto público em educação e desempenho escolar. *Revista Brasileira de Economia*, Rio de Janeiro, v. 69, n. 4, p. 467-488, 2015. DOI: <https://doi.org/10.5935/0034-7140.20150022>. Disponível em: <https://bibliotecadigital.fgv.br/ojs/index.php/rbe/article/view/41538>. Acesso em: 2 fev. 2022.

PASSADOR, C. S.; CALHADO, G. C. Infraestrutura escolar, perfil socioeconômico dos alunos e qualidade da educação pública em Ribeirão Preto/SP. *Revista de Administração, Contabilidade e Economia da FUNDACE*, Ribeirão Preto, SP, v. 3, n. 2, p. 1-10, 2012. DOI: <https://doi.org/10.13059/racef.v3i2.44>. Disponível em: <https://www.fundace.org.br/revistaracef/index.php/racef/article/view/44>. Acesso em: 2 fev. 2022.

SÃO PAULO (Estado). *Sumário executivo do SARESP 2019: volume 1*. São Paulo: Secretaria de Educação do Estado de São Paulo, 2019. Disponível em: https://saresp.fde.sp.gov.br/Arquivos/SEED1903_sumario_2019_final_v2.pdf. Acesso em: 2 fev. 2022.

SILVA, J. L. M. da; ALMEIDA, J. C. L. Eficiência no gasto público com educação: uma análise dos municípios do Rio Grande do Norte. *Planejamento e Políticas Públicas*, Brasília, DF, n. 39, p. 219-242, 2012. Disponível em: <https://www.ipea.gov.br/ppp/index.php/PPP/article/view/320>. Acesso em: 2 fev. 2022.

SILVA, N. do V.; HASENBALG, C. Tendências da desigualdade educacional no Brasil. *Dados: Revista de Ciências Sociais*, Rio de Janeiro, v. 43, n. 3, p. 423-445, 2000. DOI: <https://doi.org/10.1590/S0011-52582000000300001>. Disponível em: <https://www.scielo.br/j/dados/a/pjz8g84tSdvrM6p3GbwdNtk/?lang=pt#>. Acesso em: 2 fev. 2022.

SOARES, J. F. Qualidade e equidade na educação básica brasileira: fatos e possibilidades. In: BROCK, C.; SCHWARTZMAN, S. (ed.). *Os desafios da educação no Brasil*. Rio de Janeiro: Nova Fronteira, 2005. p. 87-114.

SOARES, J. F. Índice de desenvolvimento da educação de São Paulo–IDESP. *São Paulo Perspec.*, São Paulo, v. 23, n. 1, p. 29-41, 2009. Disponível em: http://produtos.seade.gov.br/produtos/spp/v23n01/v23n01_03.pdf. Acesso em: 2 fev. 2022.

SOARES, J. F.; RODRIGUES, E. C.; ERNICA, M. *IDeA – Indicador de desigualdades e aprendizagens*: nota técnica. São Paulo: Fundação Tide Setubal, 2019.

TODOS PELA EDUCAÇÃO (Movimento). *Relatório anual de acompanhamento do Educação Já!:* balanço 2019 e perspectivas 2020. São Paulo: Todos pela Educação, 2020. Disponível em: https://www.todospelaeducacao.org.br/_uploads/_posts/417.pdf. Acesso em: 2 fev. 2022.

TOLEDO, G. L.; OVALLE, I. I. *Estatística básica*. 2. ed. São Paulo: Atlas, 1985.

TRAVITZKI, R. Qualidade com equidade escolar: obstáculos e desafios na educação brasileira. *REICE: Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, [S. l.], v. 15, n. 4, p. 27-49, 2017. DOI: <https://doi.org/10.15366/reice2017.15.4.002>. Disponível em: <https://revistas.uam.es/index.php/reice/article/view/8146>. Acesso em: 2 fev. 2023.

VIANA, G.; LIMA, J. F. de. Capital humano e crescimento econômico. *Interações*, Campo Grande, v. 11, n. 2, p. 137-148, 2010. DOI: <https://doi.org/10.1590/S1518-70122010000200003>. Disponível em: <https://www.scielo.br/j/inter/a/srrFK6rcbj7gwW6GMyVNHK/?lang=pt>. Acesso em: 2 fev. 2022.

ZOGHBI, A. C.; MATTOS, E.; ROCHA, F.; ARVATE, P. Uma análise da eficiência nos gastos em educação fundamental para os municípios paulistas. *Planejamento e Políticas Públicas*, Brasília, DF, n. 36, p. 9-61, 2011. Disponível em: <http://www.ipea.gov.br/ppp/index.php/PPP/article/view/221>. Acesso em: 2 fev. 2022.