
TAMANHO DAS CLASSES NA REDE ESTADUAL PAULISTA: A GESTÃO DA REDE PÚBLICA À MARGEM DAS DESIGUALDADES EDUCACIONAIS

CLASS SIZE IN THE SÃO PAULO STATE PUBLIC SCHOOLS (BRAZIL):
PUBLIC SYSTEM MANAGEMENT ON THE MARGIN OF EDUCATIONAL INEQUALITIES

TAMAÑO DE LA CLASE EN LAS ESCUELAS PÚBLICAS DEL ESTADO DE SÃO PAULO (BRASIL):
LA GESTIÓN DEL SISTEMA PÚBLICO AL MARGEN DE LAS DESIGUALDADES EDUCATIVAS

Rodrigo Travitzki¹
Fernando Luiz Cássio²

RESUMO

Este trabalho investiga o efeito do tamanho das classes dos Anos Finais do Ensino Fundamental da rede estadual paulista no desempenho dos estudantes na Prova Brasil, utilizando metodologias consagradas de estudos econômicos. O trabalho avalia o efeito do tamanho das classes em três estratos de nível socioeconômico (NSE) e os resultados sugerem que as políticas em vigor no estado tendem a reduzir o desempenho de dois terços dos alunos, além de aumentar a desigualdade entre os NSE. O trabalho também apresenta uma série histórica do tamanho das classes no Estado de São Paulo, identificando uma mudança de tendência a partir da Reorganização Escolar de 2015. Discute-se ainda, a partir dos resultados, a lógica da gestão regular da rede de ensino e o papel do acesso público aos dados e indicadores educacionais do estado na compreensão de políticas estruturais que operam à margem das desigualdades sociais.

PALAVRAS-CHAVE: Qualidade (Educação). Tamanho de classe. Políticas educacionais. Direito à educação. Economia da Educação. São Paulo (Estado).

ABSTRACT

This paper investigates the effect of class size of the final years of Elementary School of the State of São Paulo public system in the performance of students in a National Exam (Prova Brasil), using consecrated methodologies from economic studies. The study evaluates the effect of class size on three socioeconomic strata, and the results suggest that the São Paulo current policies tend to reduce the performances of two-thirds of students and increase inequality among strata. The paper also presents a historical series of class size in the State of São Paulo, identifying change in the trend from the School Reorganization of 2015. It also discusses the logic of ordinary management of the school system and the role of public access to educational data and indicators of the state in understanding structural policies that operate on the margins of social inequalities.

¹ Doutor em Educação – Universidade de São Paulo (USP) - Pós-Doutorado - Universidade da Beira Interior, UBI, Portugal. Email: travitzki@usp.br

² Doutor em Química – Universidade de São Paulo (USP) - Professor Adjunto - Centro de Ciências Naturais e Humanas Universidade Federal do ABC – UFABC - Santo André, SP – Brasil. Email: fernando.cassio@ufabc.edu.br

KEYWORDS: Quality (Education). Class Size. Education Policy. Right to Education. Education Economics. São Paulo (Brazilian State).

RESUMEN

En este trabajo se investiga el efecto del tamaño de la clase al final de Educación Primaria de las escuelas públicas de São Paulo en el rendimiento de los estudiantes en la Prova Brasil, utilizando metodologías de los estudios económicos. El estudio evalúa el efecto del tamaño de la clase en tres estratos socioeconómicos (NSE) y los resultados sugieren que las políticas en vigor en el estado tienden a reducir el rendimiento de dos tercios de los estudiantes, además de incrementar la desigualdad entre los NSE.

PALABRAS CLAVE: Calidad (Educación). Tamaño de Clase. Políticas Educativas. Derecho a la Educación. Economía de la Educación. São Paulo (Estado de Brasil).

1 INTRODUÇÃO

No rescaldo das discussões sobre a proposta de Reorganização da rede estadual paulista de 2015, algumas questões ainda estão por debater. Os argumentos utilizados pela Secretaria da Educação do Estado de São Paulo (SSE-SP) para defender a medida foram pontualmente refutados por diversos grupos de pesquisadores do estado (PÓ *et al.*, 2015; CÁSSIO *et al.*, 2016). Há muito a entender, contudo, sobre a política de gestão da rede que culminou, mas não se esgota, na Reorganização Escolar em São Paulo.

Em artigo recente, Cássio *et al.* (2016) analisaram a distribuição do tamanho das classes e o seu fluxo de abertura/fechamento na rede paulista entre 2015 e 2016, concluindo que embora a Reorganização tenha sido judicialmente suspensa, a lógica que a presidiu não é diferente da que vigora na gestão ordinária da rede estadual. Portanto, a confirmação da existência de uma “reorganização silenciosa”, perseguida por movimentos sociais, sindicatos e setores da imprensa em 2016, tem relevância menos científica do que jornalística. O problema mais importante, quase sempre obscurecido pelo ruído da recepção pública às determinações da SEE-SP, é entender as políticas que, em médio e longo prazo, regulam a dinâmica da rede estadual. Se uma reorganização drástica e amplamente divulgada gera reações indesejadas na comunidade, a gestão da rede pode, com o acesso a 25 anos de dados educacionais consistentes no estado, perseguir o caminho pretendido em poucos anos.

Neste trabalho exploraremos um aspecto específico desse caudal de informações: o tamanho das classes na rede estadual paulista. Este é um dos temas que circundou o debate sobre a potencial “reorganização silenciosa” de 2016: estaria o Governo do Estado de São Paulo aumentando o número de alunos por classe nas escolas estaduais? Teria sido isso causado pelo fechamento de classes ou pela transferência compulsória de estudantes entre unidades escolares? Não pretendemos responder essas perguntas aqui³, mas oferecer uma leitura da política de tamanho de classes na rede paulista e de seu estreito diálogo com os resultados das pesquisas sobre “determinantes do desempenho escolar” provenientes do campo econômico.

Utilizando as mesmas metodologias de trabalhos de economistas sobre o assunto, investigamos a influência da variável “tamanho da classe” na nota de matemática (Prova Brasil 2013) de estudantes do Ensino Fundamental da rede estadual de São Paulo, observando-a em três estratos de nível socioeconômico (NSE). A partir dos resultados, apresentamos uma série histórica e discutimos algumas tendências para a política de tamanho das classes no estado nos anos precedentes à proposta de Reorganização da rede paulista.

Os estudos econômicos sobre o tamanho das classes

A última síntese de indicadores educacionais da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) aponta um relativo consenso na literatura internacional sobre os determinantes do desempenho escolar: “em geral, são fracas as evidências do efeito das diferenças de tamanho das classes na performance dos estudantes” (OECD, 2016, p. 394). Malgrado o breve comentário, os debates acerca desse assunto permanecem acirrados na literatura especializada. Em primeiro lugar, pois se pode discordar politicamente da primazia dos determinantes do “desempenho escolar” sobre os determinantes da **qualidade na educação**, e esta discordância (qualidade subsumida ao desempenho em testes) anuncia diferenças quase inconciliáveis entre as abordagens de

³ Apesar de Cássio *et al.* (2016) terem mostrado que as médias do número de alunos por classe, frequentemente utilizadas pela Secretaria da Educação do Estado de São Paulo como propaganda positiva da gestão, forneçam uma ideia enganosa sobre a distribuição do tamanho das classes no conjunto das escolas, e de fato o número de classes com a lotação máxima aumentou na rede estadual paulista na transição 2015-2016.

economistas e de educadores para o problema do tamanho das classes. Em segundo lugar, há uma série de divergências metodológicas dentro do próprio campo dos estudos econométricos referentes ao tema. Como, de fato, o maior volume de trabalhos relacionados à influência do tamanho das classes nos indicadores educacionais vem dos estudos econômicos, convém posicionar alguns desses debates antes de nos aproximarmos da política de tamanho de classes da rede paulista.

Se existe algum tipo de consenso a esse respeito na literatura, é a dificuldade de estudar o efeito da variável explicativa “tamanho de classe”. É fácil entender o porquê: quando se estuda fenômenos sociais complexos, isolar o efeito de uma única variável pode representar uma tarefa árdua. No caso do tamanho das classes, a variável-resposta utilizada na maior parte dos trabalhos é a nota obtida pelos estudantes em avaliações externas, geralmente em matemática. A dificuldade começa quando se tenta estabelecer relações causais entre o tamanho das classes e o desempenho nessas avaliações. Nossa intuição pedagógica diz que, quanto menor for uma classe, maior será a atenção individual recebida pelos estudantes e, provavelmente, melhor será a qualidade das aprendizagens nessas classes. Se assumirmos, por fim, que essas aprendizagens seriam mensuráveis por avaliações externas, esperaríamos índices melhores para classes menores.

Ocorre que a maioria dos estudos pertinentes ao tamanho de classes restringe-se a ser observacional, pois é difícil realizar experimentos sociais randomizados com um controle adequado das variáveis.⁴ O melhor exemplo de experimento desse tipo foi o projeto STAR (Student/Teacher Achievement Ratio), iniciado no Tennessee, em 1985, e que produziu fortes evidências de que classes menores no início da escolarização levam a melhores performances das crianças em testes cognitivos, sendo tais efeitos persistentes ao longo da

⁴ Baseados na coleta e análise de dados educacionais, **estudos observacionais** são mais baratos do que estudos experimentais, e geralmente mais aceitáveis do que estes por razões éticas (GOLDSTEIN, 2013). As abordagens observacionais, no entanto, impõem uma grande dificuldade no controle das variáveis, como quando se estuda a influência do tamanho da classe nas proficiências mensuradas por avaliações externas: a performance de estudantes em uma prova, por exemplo, é influenciada por diversas variáveis não-observadas que enviesam os resultados. O problema central consiste na comparabilidade entre grupos que receberam o “tratamento” (alunos estudando em classes pequenas, pessoas que foram beneficiadas por certa política pública, etc.) e grupos-controle (não afetados pela variável investigada). Uma das soluções encontradas para lidar com esses problemas reside nas **abordagens quasi-experimentais**, que, considerando variáveis de pré-tratamento (o *background* familiar, dados demográficos, etc.), propõem uma seleção dos grupos que serão comparados, mimetizando, na medida do possível, algumas condições de controle experimental.

vida escolar (MOSTELLER, 1995). Os debates sobre o tamanho de classes nos Estados Unidos são sobejamente politizados e polarizados, o mais famoso deles entre os economistas da educação Eric A. Hanushek e Alan B. Krueger (MISHEL; ROTHSTEIN, 2002).⁵ Apesar de suas diferenças políticas, “eles parecem concordar que uma classe de tamanho menor *pode* ter influência em *algumas* circunstâncias” (RICE, 2002, p. 91, grifos originais). Com efeito, outros estudos sugerem que efeitos positivos de classes menores são mais pronunciados para estudantes de famílias de baixa renda, indicando a possibilidade de políticas localizadas de redução no tamanho das classes serem mais eficientes do que uma política universal no mesmo sentido (idem, p. 92).

Estudos sobre o tamanho de classes foram realizados em diversos países (URQUIOLA; VERHOOGEN, 2009, p. 179), e foge ao escopo deste trabalho apresentar uma revisão completa a respeito. Algumas dessas pesquisas, no entanto, por abordarem o viés dos estratos socioeconômicos nas relações tamanho de classe/desempenho educacional, devem ser destacadas. A primeira delas é a de Angrist e Lavy (1999), realizada em Israel, cuja conclusão é de que a correlação positiva entre o tamanho de classes e o desempenho escolar (observada em pesquisas nos Estados Unidos, por exemplo) é um artefato da associação entre as classes menores e a proporção de alunos em condições desfavoráveis. Ao proporem um controle mais adequado das variáveis (considerando o *background* familiar dos estudantes, por exemplo), os autores encontraram uma associação negativa entre as classes maiores e o desempenho dos alunos, reforçando que este resultado, válido para um país como Israel, não necessariamente se aplica a países como Estados Unidos. Da mesma forma, deixam claro que os “fracos efeitos” do tamanho das classes no desempenho escolar dos estudantes – também observados nos países em desenvolvimento e utilizados por economistas da linha de Hanushek para defender que os insumos escolares não têm efeitos causais na aprendizagem dos estudantes – poderiam ser fruto da baixa qualidade do controle das variáveis nos estudos observacionais realizados.

A partir de dados do sistema educacional francês, Piketty e Valdenaire (2006)

⁵O primeiro questiona a apropriação dos resultados do experimento STAR pelos formuladores de políticas públicas de educação, defendendo a tese de que políticas de redução de tamanho de classe são muito caras e pouco efetivas. O segundo sugere que estas mesmas políticas seriam custo-efetivas a longo prazo. A diferença essencial entre as posições dos dois economistas baseia-se na importância dada à variável tamanho de classe em relação a outros determinantes do desempenho escolar (qualificação e fixação do corpo docente, por exemplo).

mencionam que os métodos “ingênuos” (com controle de variáveis precário) subestimam o verdadeiro impacto causal da redução do tamanho das classes, e que o viés estatístico dos estudos que concluem pelos “efeitos limitados (leia-se, nulos) de classes reduzidas sobre a performance dos estudantes têm provavelmente contribuído para alimentar certo ceticismo no debate público a respeito de políticas especiais para escolas em contextos desfavoráveis” (idem, p. 9). Também neste estudo, os resultados assinalam um impacto diferenciado do tamanho das classes para estudantes socialmente desfavorecidos. Talvez pela notoriedade recente dos trabalhos de Thomas Piketty, este é o único estudo mencionado pela OCDE (2016, p. 394) para relativizar a posição de uma parcela da literatura – amparada pela Organização, mas contraintuitiva para os educadores – de que o tamanho das classes influencia pouco o desempenho escolar.

Em um trabalho metodologicamente inovador, Urquiola e Verhoogen (2009) estudaram o sistema educacional do Chile, incluindo as escolas privadas, mostrando que a procura por escolas com número maior ou menor de alunos por classe, em um sistema em que a metade das escolas é privada, é influenciada por fatores como renda e escolaridade familiar. Famílias de NSE altos tendem a procurar escolas privadas com classes menores (produtos de diferenciação dessas escolas no mercado). Do mesmo modo, certas escolas públicas são sistematicamente procuradas por famílias de NSE mais baixos, e essas também possuem classes menores, por serem menos procuradas. Enquanto que as escolas subsidiadas (em sistema de *voucher*) que atendem as classes médias, sobretudo as operadas por grandes grupos empresariais da educação, trabalham com classes bem maiores. Esse resultado demonstra a dificuldade de estudar o efeito do tamanho das classes no desempenho escolar sem considerar adequadamente os diversos fatores que influenciam a dinâmica do tamanho das classes.⁶

No Brasil, alguns estudos de Naércio de Menezes-Filho sobre os determinantes do desempenho escolar têm sido bem recebidos por diversos grupos e instituições, geralmente vinculados ao capital privado. Em linhas gerais, esses estudos dialogam com as ideias de

⁶No ano 2000, o MEC publicou estudo sobre a influência do tamanho da classe no desempenho escolar a partir dos dados do Saeb 1997, concluindo que o tamanho das classes não fazia diferença nas proficiências observadas (WEISELFISZ, 2000) e fazendo coro com os estudos de Hanushek nos Estados Unidos. Ao confrontarmos o texto deste estudo com a discussão metodológica disponível na literatura mais recente, concluímos ser provável que este resultado brasileiro seja enviesado por um controle inadequado das variáveis e dos grupos de comparação.

Hanushek e concluem que o tamanho das classes influencia menos o desempenho escolar dos estudantes do que, por exemplo, o tempo de permanência na escola (MENEZES-FILHO, 2011).⁷ Não por acaso, os mesmos grupos que acolhem essas conclusões – que, como vimos, estão longe de ser majoritárias na literatura sobre economia da educação – são também os principais apoiadores das políticas de escola em tempo integral, tais como a polêmica Medida Provisória n. 746/2016, que institui a reforma do Ensino Médio no Brasil.

Metodologias para o estudo do tamanho das classes e a proposta deste trabalho

Como a literatura tem mostrado, a dificuldade de controle das variáveis que influenciam o tamanho das classes e a impossibilidade de realizar experimentos randomizados exigem a adoção de metodologias que permitam trabalhar com dados públicos e indicadores educacionais (dados observacionais) de forma controlada, selecionando aquilo que será comparado e tentando isolar as variáveis para estudar seus efeitos e minimizar os viesamentos.

Várias metodologias *quasi*-experimentais têm sido empregadas para investigar os efeitos de tamanho de classe sobre o desempenho escolar, destacando-se a técnica de variáveis instrumentais (ANGRIST; LAVY, 1999; PIKETTY; VALDENAIRE, 2006; CAMARGO, 2012)⁸ e o método de pareamento com escore de propensão (*propensity score matching*) que foi generalizado para variáveis de tratamento contínuas por Hirano e Imbens (2004), método empregado por Oliveira (2010) e neste trabalho. O uso desse método permite, segundo Oliveira (OLIVEIRA, 2010, p. 68),

determinar, entre as crianças que estão em turmas de 35 alunos, as que mais se assemelham àquelas que estão em turmas de 25 alunos em termos das características observáveis que, por hipótese, influem tanto nos resultados escolares quanto no tamanho da classe.

⁷ Menezes-Filho orientou algumas teses e dissertações sobre o tema, incluindo os trabalhos de Franco (2008), que reconhece que a elevada estratificação do sistema educacional brasileiro traz dificuldades para se obter *rankings* confiáveis; e Oliveira (2010), que realizou um estudo de custo-efetividade das políticas de redução do tamanho da classe.

Este último trabalho, como o de Camargo (2012), realizado no Rio Grande do Sul, aborda a dificuldade de lidar com os *peer effects* (efeitos dos pares) nesse tipo de estudo (quando se aumenta o número de alunos em uma classe, pode haver um efeito positivo dos pares na proficiência de todo o grupo).

⁸Esse método visa isolar o efeito da variável desejada utilizando uma nova variável (dita instrumental) que seja altamente correlacionada com aquilo que se deseja estudar, mas que não seja correlacionada com características não observadas que afetam os resultados dos indivíduos (CAMARGO, 2012, p. 27).

A autora acrescenta que “se o pareamento é adequado as variáveis não observadas relevantes estão igualmente distribuídas nos grupos de tratamento e de controle e seu efeito é anulado” (idem, p. 69). Estudos de avaliação de projetos da Fundação Itaú Social têm não apenas utilizado o método, como também oferecido jornadas de formação metodológica para gestores de políticas públicas e produzido literatura própria a respeito (MENEZES-FILHO, 2016).

Buscando forjar um grupo-controle semelhante ao grupo de tratamento em termos de características que sejam observáveis, o pareamento com escore de propensão é severamente criticado por King e Nielsen (2016), que argumentam que a técnica se aproxima de um *design* experimental de padrões inferiores ao desejável, desprezando informações e gerando novos vieses estatísticos (isto constituiria um paradoxo, em se tratando de um método que aspiraria a eliminá-los). Em suma, os autores recomendam aos usuários dessa metodologia conhecerem bem as limitações do método e informarem nos seus textos o que foi realizado.

O objetivo deste trabalho, além de utilizar o método para analisar os dados da rede paulista, é problematizar o seu uso na avaliação de políticas públicas de educação, a exemplo da política de tamanho de classes. Como mostraremos, ao se adicionar uma única variável em uma análise estratificada, o método produz resultados bastante distintos.

2 DADOS E MÉTODOS

Neste estudo replicamos o método do pareamento com escore de propensão generalizado (GPS, *generalized propensity score*) (HIRANO; IMBENS, 2004) utilizado por Oliveira (2010), que estimou o efeito do tamanho das classes sobre a proficiência em matemática (Saeb 2005) de estudantes dos Anos Iniciais do Ensino Fundamental de escolas urbanas (estaduais e municipais) do Brasil. Não apenas o método, bem como a base de dados e as variáveis aqui utilizadas são praticamente idênticas as do trabalho de Oliveira.

Cabe, no entanto, destacar algumas diferenças entre os dois trabalhos. Aqui, a análise centrou-se nos alunos dos Anos Finais do Ensino Fundamental de escolas estaduais e urbanas do Estado de São Paulo. Ressalte-se que, embora restrito a um único estado da federação, nosso estudo utiliza uma amostra maior que a de Oliveira (2010), pois em 2005

não havia Prova Brasil – enquanto o Saeb é amostral, a Prova Brasil é censitária para as escolas públicas de Ensino Fundamental. Outra diferença importante entre os dois estudos é que neste se propõe o desdobramento dos resultados em três níveis socioeconômicos (NSE) diferentes.

As análises foram realizadas por meio do *software* R, de livre acesso e código aberto. Além de ser amplamente reconhecido na academia, o *software* pode ser obtido sem custos, ampliando o caráter democrático e a reprodutibilidade dos estudos. Foram utilizados especialmente os pacotes R *causaldrf* (GALAGATE; SCHAFFER, 2015), *weights* (PASEK *et al.*, 2016), *plyr* (WICKHAM, 2011) e *Hmisc* (HARRELL JR.; DUPONT, 2016).

Dados utilizados

Analisamos os microdados da Prova Brasil 2013 dos alunos dos Anos Finais do Ensino Fundamental de escolas estaduais urbanas do Estado de São Paulo. A amostra Total contém 337.524 registros, representando 418.589 alunos distribuídos em 12.582 classes e 3.473 escolas. Na análise realizada, a variável-resposta é a nota de matemática do aluno, enquanto a variável de tratamento é o tamanho da classe. As 47 variáveis “candidatas” a pré-tratamento são provenientes do estudo de Oliveira (2010), recebendo inclusive os mesmos nomes – motivo pelo qual não as descrevemos aqui. Algumas das variáveis empregadas naquele trabalho não puderam ser replicadas em virtude de mudanças na base de dados. Apenas duas variáveis foram alteradas em relação ao estudo de Oliveira: em vez de utilizar a média e o desviopadrão da nota de matemática, utilizamos a média e o desviopadrão da escolaridade familiar do aluno (variáveis *turma.escolarid.media* e *turma.escolarid.sd*). O objetivo destas variáveis é ajudar a controlar o efeito dos pares na estimação do efeito do tamanho da classe. Segue a lista das variáveis de pré-tratamento:

- 1. aluno e contexto familiar:** *jareprovado*, *aluatrasado*, *alunaobranco*, *aluhomem*, *trabalhanao*, *preescola*, *compcasa*, *moracomamae*, *escmae_0*, *escmae_4i*, *escmae_11*, *escmae_48*, *escmae_sup*, *escmae_ns*, *mora_2*, *mora_5*, *mora_5m*;
- 2. diretor:** *dtfunc_4*, *dtfunc_15*, *dtfunc_15m*, *tdir_2*, *tdir_4*, *tdir_5*, *dcargah_20*, *dcargah_30*, *dcargah_40*, *dcargah_40m*;
- 3. professor:** *escdoc_em*, *escdoc_supped*, *escdoc_supmat*, *escdoc_supout*, *templec9*, *templec10m*;

4. classe e escola: selecao_sim, selecao_out, selecao_nao, ft_homo, ft_hetero, ft_nao, segurancasim, conservbom, violencianao, compas, intas, capital, turma.escolarid.media e turma.escolarid.sd.

TABELA 1 – Descrição de algumas variáveis utilizadas

	Média	EP	Variância	N _{alunos}	N _{obs}
turma.tamanho	34,132	0,008	23,883	418.589	337.524
nota.MT	245,642	0,073	2220,389	418.589	337.524
turma.escolarid.media	7,776	0,002	2,011	418.564	337.516
escdoc_supmat (formação docente)	0,621	0,001	0,235	407.676	328.783
compas (computador na escola)	0,593	0,001	0,241	409.031	329.820
compcasa (computador em casa)	0,828	0,001	0,143	412.946	333.094
aluatrasado (defasagem idade/série)	0,345	0,001	0,226	414.250	334.133
jareprovado (reprovação)	0,198	0,001	0,159	408.723	329.769

Fonte: Elaborado a partir dos microdados da Prova Brasil 2013, MEC/Inep.

Alunos de 9º ano do Ensino Fundamental de escolas estaduais urbanas do Estado de São Paulo. EP = Erro Padrão da média; N_{alunos} = número de alunos representados na amostra; N_{obs} = número de observações.⁹

Para a construção da série histórica, foram utilizados dados do Censo Escolar (no nível das classes) e dados (no nível das escolas) obtidos do Cadastro de Escolas da SEE-SP via Lei n. 12.527/2011 (Lei de Acesso à Informação) dos meses de maio de 2014 a 2016, devido ao fato de maio ser o mês de referência para a consolidação dos dados do Censo Escolar no estado.

⁹Cabe notar que, embora os dados da Prova Brasil sejam censitários, nem todos os estudantes acabam participando da avaliação, de forma que é necessário o uso de um peso individual para ponderar as diversas técnicas estatísticas. Neste trabalho, foi utilizada a variável ‘peso_aluno_MT’ para tal fim.

Seleção das variáveis de pré-tratamento

A seleção de variáveis de pré-tratamento consiste em um ponto crítico nos métodos baseados em escore de propensão, visto haver risco de alterar os resultados dependendo dos controles incluídos no modelo (KING; NIELSEN, 2016). Neste enfoque, realizamos um procedimento objetivo de seleção de variáveis. Em primeiro lugar, utilizamos variáveis já estabelecidas em estudo anterior com uma base de dados semelhante (OLIVEIRA, 2010). Em segundo lugar, a partir deste rol de variáveis, foram selecionadas as variáveis associadas à resposta e ao tratamento para cada uma das quatro amostras estudadas (Total, NSE baixo, NSE médio e NSE alto).¹⁰

Aplicação do método GPS

Para cada uma das quatro amostras, ajustamos um modelo linear associando o tamanho da classe às variáveis de pré-tratamento. A partir deste modelo, estimamos o GPS para cada observação (aluno). Testamos a capacidade do GPS de balancear as variáveis de pré-tratamento. Na prática, isso significa reduzir (idealmente anular) a diferença entre as médias de cada variável quando se compara alunos em classes de tamanhos diferentes. Foram obtidos resultados satisfatórios neste sentido. Por fim, ajustamos o modelo da esperança condicional da resposta, ou seja, o modelo que prediz a resposta média esperada para cada tamanho de classe, levando em conta o viés de seleção “detectado” com o GPS. Tal procedimento gera a **função resposta à dose**¹¹, que pode ser observada mais adiante nas Figuras 1 e 2.¹²

3 O EFEITO DO TAMANHO DAS CLASSES EM SÃO PAULO

A primeira parte desta seção trata dos resultados relativos ao conjunto total dos alunos analisados. Em seguida, replicamos os resultados separados em três estratos de NSE. A Figura 1 mostra qual seria a nota de matemática (Y) para cada aluno dependendo do tamanho de sua classe (T). Observa-se que a relação entre Y e T não é linear, sugerindo um

¹⁰Mais detalhes podem ser encontrados no Apêndice.

¹¹A curva que relaciona o “tratamento” (estar em uma classe com determinado tamanho) à variável-resposta (ter uma certa nota em matemática). Quanto melhor a qualidade do pareamento, mais segura será a relação causal sugerida por esta curva.

¹²Mais detalhes sobre a aplicação do método podem ser encontrados no Apêndice.

“ponto ótimo” estimado em 34 alunos (linha pontilhada no gráfico), que seria o tamanho ideal de classe para os Anos Finais do Ensino Fundamental em São Paulo, tendo como referência a nota de matemática da Prova Brasil 2013. Este número é maior do que o recomendado pelo Conselho Nacional de Educação (30 alunos por classe)¹³, porém um pouco menor do que o permitido no estado (35 alunos).¹⁴ É um resultado, portanto, que suporta a atual política estadual de tamanho de classes.¹⁵

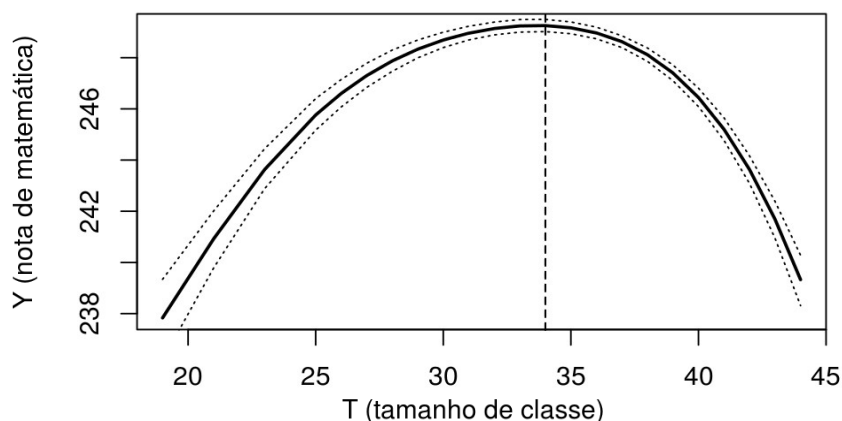


FIGURA 1 – Efeito do tamanho da classe (T) sobre a nota de matemática (Y) para a amostra Total, 9º ano do EF, rede estadual paulista (função resposta à dose). As linhas pontilhadas representam o intervalo de confiança (95%) das estimativas

Fonte: Elaborado a partir dos microdados da Prova Brasil 2013 – MEC/Inep; método GPS (HIRANO; IMBENS, 2004).

Em certa medida, a Figura 1 confirma o resultado obtido por Oliveira (2010, p. 102) para os Anos Iniciais do Ensino Fundamental. A pesquisadora havia estimado um tamanho ótimo (para aquela etapa) de 30 alunos por classe, menor do que estimamos para os Anos

¹³Segundo o Parecer CNE/CEB n. 08/2010.

¹⁴Segundo a Resolução SE n. 02/2016.

¹⁵Mais adiante mostraremos que a função resposta à dose para a amostra Total não fornece uma descrição adequada do fenômeno, pois deixa de levar em conta as diferenças socioeconômicas (Figura 2). Com efeito, os resultados deste estudo, quando tomados em conjunto, não suportam a atual política de tamanho de classes no Estado de São Paulo.

Finais – o que é compatível com a legislação estadual¹⁶ e a literatura: tipicamente, os anos iniciais da escolaridade pedem classes menores (PIKETTY; VALDENNAIRE, 2006).

A importância do nível socioeconômico para a estimação do efeito

Resta verificar em que medida estes resultados, obtidos com a amostra Total de alunos, mantêm-se inalterados ao analisarmos os diferentes contextos socioeconômicos. Para isso, desdobramos a amostra em três estratos de NSE de tamanho semelhante utilizando a variável “escolaridade familiar média por turma” como indicador (*proxy*) de NSE. A descrição de algumas variáveis para os três NSE se encontra na Tabela 2. A Figura 2 apresenta o mesmo tipo de resultado da Figura 1, no entanto, considera as diferenças de NSE no estado.

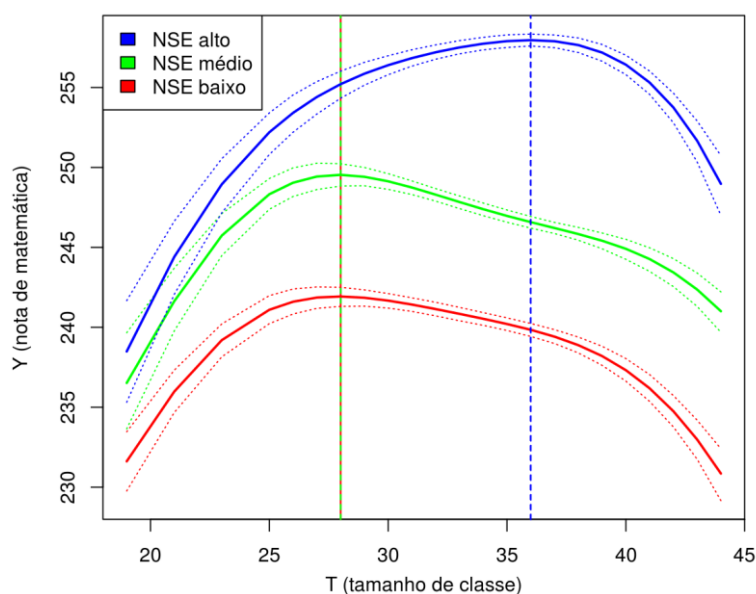


FIGURA 2 – Efeito do tamanho da classe (T) sobre a nota de matemática (Y) em três NSE, 9º ano do EF, rede estadual paulista (função resposta à dose). As linhas pontilhadas representam o intervalo de confiança (95%) das estimativas

Fonte: Elaborado a partir dos microdados da Prova Brasil 2013, MEC/Inep; método GPS (HIRANO;IMBENS, 2004).

¹⁶ A Resolução SE n. 02/2016 define em 30 o número máximo de alunos por classe nos Anos Iniciais do Ensino Fundamental na rede estadual paulista, enquanto o Parecer CNE/CEB n. 08/2010 sugere que 24 alunos por classe seria o tamanho ideal das classes para esta etapa.

Comparando este resultado com o da Figura 1, a amostra Total parece seguir um comportamento mais próximo ao do estrato superior de NSE (curva azul), diferenciando-se dos demais. Percebe-se tal fato não apenas pelo formato das curvas, como também pelos diferentes pontos ótimos estimados: a diferença entre o ponto ótimo dos dados completos (34) é de 2 alunos em relação ao NSE superior (36) e de 6 alunos em relação aos outros dois terços da população (28). Some-se ainda o fato de que, para todos os estratos, a moda da distribuição é de 35 alunos, ou seja, o tamanho de classe mais encontrado foi invariavelmente de 35 alunos por classe, corroborando o resultado de Cássio *et al.* (2016). Ao se observar a Figura 2, nota-se que nas classes de 35 alunos há uma grande diferença na nota de matemática esperada, quando se compara o NSE alto com os outros dois estratos.

No aspecto metodológico, estes resultados, como tantos outros já mencionados (PIKETTY; VALDENAIRE, 2006; TRAVITZKI *et al.*, 2014), salientam a importância de levar em conta o NSE ao investigar fenômenos educacionais. No que toca às políticas públicas voltadas direta ou indiretamente ao tamanho de classes na rede estadual paulista, nossos resultados sugerem que:

- a) há uma relação significativa entre o tamanho de classe e o desempenho escolar. Esta relação não é linear e sugere um ponto ótimo para o tamanho das classes, ou seja, em relação ao desempenho escolar, é desejável que as classes não sejam muito grandes nem muito pequenas;
- b) o ponto ótimo varia em diferentes contextos socioeconômicos. No terço superior de NSE, o tamanho ideal seria de 36 alunos nos Anos Finais do Ensino Fundamental da rede paulista. Nos dois terços inferiores, contudo, o tamanho ideal seria de 28 alunos;
- c) de forma geral, os resultados obtidos com a amostra Total de alunos se assemelham mais aos resultados obtidos para o terço superior de nível socioeconômico do que para os dois terços inferiores.

Nos diversos testes realizados (com diferentes variáveis e modelos) foram observadas variações nos pontos ótimos (com um ou dois alunos por classe de deslocamento para a esquerda ou para a direita nas curvas). Isso reflete, por um lado, as limitações inerentes às metodologias *quasi*-experimentais, provenientes, em última análise, de limitações nos dados disponíveis. Por outro lado, pelo menos dois resultados se

confirmaram de forma sistemática: (1) a forma das curvas sempre sugere um ponto ótimo; (2) o terço superior de NSE sempre apresenta ponto ótimo maior que os dois terços inferiores, que possuem valores iguais ou próximos entre si. Alguns testes realizados para as redes municipais sugeriram um ponto ótimo menor, contudo são necessários mais estudos neste sentido.

Em suma, os resultados sugerem que as políticas de tamanho de classe da rede paulista, ainda que fossem executadas com perfeição, teriam o efeito esperado em apenas um terço dos estudantes. Nessa premissa, uma política que seja efetivamente voltada para a melhoria do desempenho escolar de todos os estudantes precisa superar a visão da população como um bloco homogêneo e adaptar-se aos diferentes contextos socioeconômicos.

4 A GESTÃO DO TAMANHO DAS CLASSES NA REDE PAULISTA

Para ajudar a entender como a gestão do tamanho das classes na rede estadual paulista atinge estudantes de diferentes contextos socioeconômicos, desdobramos a descrição das variáveis da Tabela 1 nos três estratos de NSE (Tabela 2). A Figura 3 ilustra a comparação entre os níveis com três dessas variáveis. Embora todas as variáveis sejam significativamente diferentes ($p < 0,01$) entre os estratos – até porque a amostra é grande –, a magnitude da variação foi menor no tamanho de classes do que nas outras variáveis. Em outras palavras, os alunos dos três NSE são bastante diferentes entre si, mas estudam em classes de tamanho médio semelhante. Cumpre notar, por outro lado, uma diferença quando se observa não a média, mas a dispersão do tamanho das classes. Na Tabela 2 verifica-se que a variância do tamanho das classes é maior nos estratos inferiores de NSE. Tendo em conta o formato das curvas nas Figuras 1 e 2, uma variância maior tende a prejudicar o desempenho dos alunos.

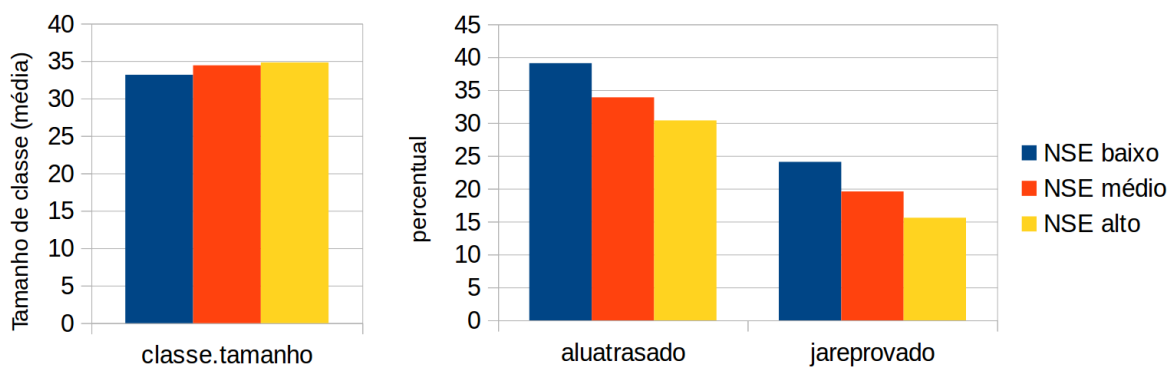


FIGURA 3 – Comparação de algumas variáveis nos três estratos de NSE

Fonte: Elaborado a partir dos microdados da Prova Brasil 2013, MEC/Inep.

TABELA 2 – Descrição de algumas variáveis por estrato socioeconômico

NSE baixo					
	Média	EP	Variância	N_{alunos}	N_{obs}
turma.tamanho	33,174	0,014	29,074	142.526	111.442
nota.MT	237,291	0,12	2056,713	142.526	111.442
turma.escolarid.media	6,238	0,002	0,573	142.526	111.442
escdoc_supmat	0,608	0,001	0,238	138.169	108.058
compas	0,601	0,001	0,24	139.570	109.135
compcasa	0,743	0,001	0,191	140.386	109.813
aluatrasado	0,391	0,001	0,238	140.887	110.196
jareprovado	0,241	0,001	0,183	138.863	108.648
NSE médio					
	Média	EP	Variância	N_{alunos}	N_{obs}
turma.tamanho	34,434	0,013	22,373	138.249	111.315
nota.MT	244,583	0,124	2134,348	138.249	111.315
turma.escolarid.media	7,798	0,001	0,124	138.249	111.315
escdoc_supmat	0,618	0,001	0,236	134.797	108.494
compas	0,582	0,001	0,243	135.248	108.910
compcasa	0,835	0,001	0,138	136.429	109.877
aluatrasado	0,339	0,001	0,224	136.820	110.192
jareprovado	0,196	0,001	0,158	134.894	108.663
NSE alto					
	Média	EP	Variância	N_{alunos}	N_{obs}
turma.tamanho	34,823	0,012	18,464	137.775	114.755
nota.MT	255,345	0,129	2308,514	137.775	114.755

turma.escolarid.media	9,346	0,002	0,476	137.775	114.755
escdoc_supmat	0,639	0,001	0,23	134.671	112.219
compas	0,596	0,001	0,241	134.174	111.763
compcasa	0,909	0,001	0,083	136.117	113.400
aluatrasado	0,304	0,001	0,212	136.530	113.741
jareprovado	0,156	0,001	0,132	134.952	112.454

Fonte: Elaborado a partir dos microdados da Prova Brasil 2013, MEC/Inep.

Alunos de 9º ano do Ensino Fundamental de escolas estaduais urbanas do Estado de São Paulo. EP = Erro Padrão da média; N_{alunos} = número de alunos representados na amostra; N_{obs} = número de observações.

Dadas as diferenças entre o percentual de alunos já reprovados ou com defasagem idade/série nos três NSE – em contraste com a semelhança no tamanho de classe – e considerando as curvas da Figura 2, verifica-se a adoção de uma política que é reprodutora ou mesmo ampliadora das desigualdades educacionais na rede estadual paulista. Aliás, o tamanho de classe mais comum em todos os estratos, 35 alunos, está entre os que mais aumentam as diferenças de desempenho entre o terço superior e os dois terços inferiores na Figura 2. Mais do que ignorar as diferenças entre os NSE, do ponto de vista do tamanho das classes a política parece estar orientada pelo perfil do estrato superior.

As conclusões até aqui estão fundamentadas nos mesmos pressupostos, dados e metodologias dos estudos econômicos que tomam o desempenho dos alunos em avaliações externas como única medida da qualidade do ensino, e cujos resultados têm aproximado agentes públicos e privados no planejamento das políticas para a gestão das redes públicas. Como se vê, as políticas de tamanho de classe na rede paulista parecem dialogar muito mais com os economistas da educação que privilegiam a minimização do custo das políticas em vez da ampliação de seu alcance aos diferentes níveis socioeconômicos.

O tamanho das classes na rede paulista e o problema do acesso à informação

Os resultados anteriores se referem a 2013, os dados mais recentes disponíveis da Prova Brasil. Utilizando as informações do Censo Escolar (MEC/Inep) e da própria SEE-SP (Cadastro de Escolas), traçamos uma série histórica do tamanho das classes no Estado de

São Paulo, importante para a compreensão das recentes políticas educacionais paulistas e, em particular, de um dos aspectos centrais da proposta de Reorganização de 2015. A Figura 4 apresenta as medianas do tamanho das classes na rede estadual paulista entre 2007 e 2016 para as duas etapas do Ensino Fundamental e para o Ensino Médio.¹⁷ Observa-se que diminuíram sistematicamente de 2007 a 2015 e aumentaram pela primeira vez (exceto EF-AI) justamente na transição 2015-2016.

Em relação ao movimento de diminuição, muitas são as razões que poderiam explicar a evolução do tamanho das classes na direção de classes menores, entre as quais a retração do volume de matrículas na rede estadual, fatores demográficos, políticas específicas de diminuição do tamanho das classes, etc. Os argumentos demográficos, utilizados pela SEE-SP para justificar a proposta da Reorganização, já foram refutados na primeira parte do estudo de Cássio *et al.* (2016). Já o volume de matrículas no Estado de São Paulo foi objeto de um levantamento recente, mostrando que as matrículas na rede estadual vêm diminuindo gradualmente desde 2007 no Ensino Fundamental, concomitantemente ao crescimento das redes municipal e privada no estado (TODOS PELA EDUCAÇÃO, 2016). Por outro lado, uma retração no número absoluto de matrículas ao longo dos anos diz muito pouco sobre a distribuição – espacial e socioeconômica – dos estudantes em uma rede tão complexa como a do Estado de São Paulo.

¹⁷As medianas fornecem, para os objetivos aqui propostos, informações mais acuradas sobre a distribuição do tamanho das classes do que as médias, que quase sempre apresentam valores menores em virtude das (raras) classes muito pequenas, descrevendo de forma enviesada a realidade da rede.

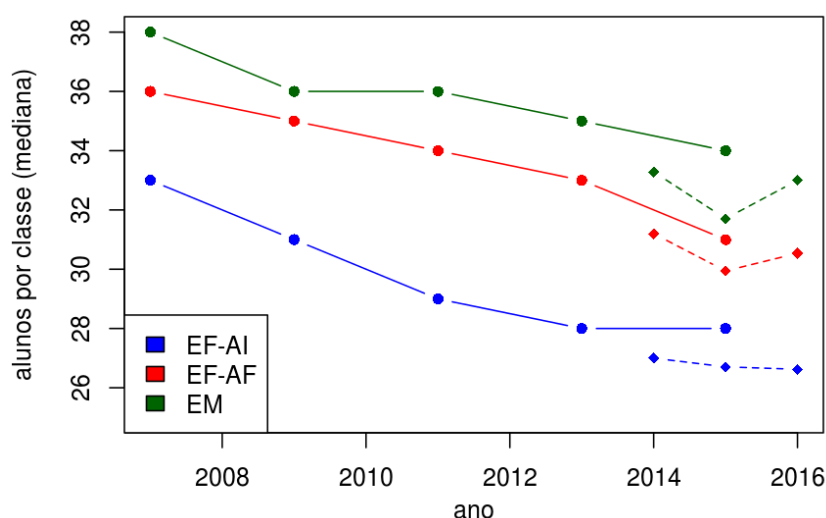


FIGURA 4 – Mediana do tamanho das classes por etapa (rede estadual paulista, 2007-2016). Linhas contínuas: dados desagregados (por classe) provenientes do Inep. Linhas pontilhadas: dados agregados (por escola) provenientes da SEE-SP

Fonte: Elaborado a partir dos microdados do Censo Escolar, MEC/Inep, 2007-2015; e do Cadastro de Escolas da SEE-SP, 2014-2016.

Ainda que essas variações pudessem ser utilizadas para defender uma Reorganização Escolar que readequasse o tamanho das classes como resposta aos movimentos de depleção de matrículas, observa-se, na Figura 4 (linhas cheias), um decréscimo do número de alunos por classe para um mínimo, em 2015, muito próximo ao preconizado pelo Parecer CNE/CEB n. 08/2010 para o tamanho ideal das classes no Brasil. Frise-se ainda que nos referimos ao tamanho ideal das classes sem considerar os diferentes contextos socioeconômicos, que, como vimos, importam uma considerável complexidade ao problema.

Entre 2014 e 2016, nas linhas pontilhadas da Figura 4, nota-se um aumento da mediana do tamanho das classes na rede estadual, o primeiro movimento de aumento desde 2007. Isto indicia – tal como já havíamos observado com a evolução da moda do número de alunos por classe (CÁSSIO *et al.*, 2016) – uma gestão que visa formar classes maiores e que, diante dos resultados apresentados na seção anterior, opera à margem das desigualdades sociais, ignorando os efeitos do tamanho de classe nos diferentes estratos socioeconômicos. Nesse prisma, o presente estudo pode ajudar a SEE-SP a reformular sua política de tamanho de classes, agora com vistas à promoção do desempenho escolar de “todos” os estudantes.

Embora tanto as linhas contínuas e pontilhadas da Figura 4 apresentem, em teoria, a mesma informação, é preciso pontuar que as diferenças entre estas se relacionam a diferenças de agregação dos dois conjuntos de dados. Os dados do Censo Escolar (linhas contínuas) estão desagregados no nível das classes, enquanto os dados da SEE-SP estão agregados no nível das escolas: para estes últimos, só é possível saber o número médio de alunos por classe em cada escola, dividindo o número de matrículas pelo número de classes. Não é possível, portanto, avaliar as classes propriamente ditas, apenas médias em cada escola. Em decorrência do nível de agregação dos dados, as medianas de tamanho de classe com dados da SEE-SP são sistematicamente menores se comparadas às do Inep, mais fiéis à realidade.

Além disso, a SEE-SP divulga o tamanho das classes da rede através de médias, não informando medianas ou modas, procedimento análogo ao de muitas redes públicas no Brasil. No entanto, a média não descreve adequadamente distribuições assimétricas como o tamanho de classe (observado nos dados do Censo Escolar) – neste caso, as médias são sempre menores do que as medianas, por sua vez menores do que a moda. Neste sentido, se a política de gestão do tamanho das classes da rede paulista trabalha à margem das desigualdades educacionais, também a política de disponibilização de dados opera de forma a ocultá-las sob a homogeneidade das médias e dos números absolutos.

5 CONCLUSÃO

A análise da amostra de 337.524 alunos sugere que há uma relação não-linear entre o tamanho de classe e o desempenho escolar, em que se observa um “ponto ótimo”. Considerando a população total, o ponto ótimo para os Anos Finais do Ensino Fundamental é de 34 alunos: exatamente a média do tamanho de classe. Sob essa ótica, as políticas educacionais paulistas se mostram consistentemente fundamentadas em evidências – ao menos em média.

Contudo, levando em conta as diferenças de NSE presentes no estado, a análise revelou que o terço superior de NSE apresenta um ponto ótimo de 36 alunos, enquanto nos dois terços inferiores seria 28. Adicionalmente, em classes de 34 alunos observou-se considerável diferença de desempenho entre os três estratos de NSE. Tais resultados sugerem que as políticas vigentes no estado, de caráter homogeneizante, têm o condão de

reduzir o desempenho escolar de todos os estratos (especialmente dos dois terços inferiores) e de acentuar as desigualdades educacionais existentes.

A análise da série histórica do tamanho das classes mostra, por seu turno, que a Reorganização Escolar de 2015 rompeu a longa tendência de redução das classes nos Anos Finais do Ensino Fundamental e no Ensino Médio da rede estadual paulista. Tendo em vista os resultados anteriores, este novo movimento no sentido de aumentar o tamanho das classes deve prejudicar o desempenho escolar de dois terços dos estudantes nos Anos Finais do Ensino Fundamental – e justamente daqueles que mais necessitam dos serviços públicos de educação.

REFERÊNCIAS

ANGRIST, Joshua D.; LAVY, Victor. Using maimonides' rule to estimate the effect of class size on scholastic achievement. **The Quarterly Journal of Economics**, v. 114, n. 2, p. 533-575, 1999.

CAMARGO, Juliana. **O efeito do tamanho da turma sobre o desempenho escolar: uma avaliação do impacto da "enturmação" no Ensino Fundamental do Rio Grande do Sul**. Dissertação (Mestrado em Economia). Porto Alegre: Universidade Federal do Rio Grande do Sul, 2012.

CÁSSIO, Fernando L.; CROCHIK, Leonardo; DI PIERRO, Maria Clara; STOCO, Sergio. Demanda social, planejamento e direito à educação: Uma análise da rede estadual de ensino paulista na transição 2015-2016. **Educação & Sociedade**, Campinas, v. 37, n. 137, p. 1089-1119, 2016.
FRANCO, Ana Maria de Paiva. **Os determinantes da qualidade da educação no Brasil**. Tese (Doutorado em Economia). São Paulo: Universidade de São Paulo, 2008.

GALAGATE, Douglas; SCHAFFER, Joseph L. **Estimating average dose response functions using the R Package *causaldrf***. 30 nov. 2015. Disponível em: <https://cran.r-project.org/web/packages/causaldrf/vignettes/Using_causaldrf.pdf>. Acesso em: 14 nov. 2016.

GOLDSTEIN, Harvey. Evaluating educational changes: a statistical perspective. **Ensaio: Avaliação e Políticas Públicas em Educação**. Rio de Janeiro, v. 21, n. 78, p. 101-114, 2013.

HARRELL JR., Frank E.; DUPONT, Charles. **Hmisc: Harrell miscellaneous**. Disponível em: <<https://cran.r-project.org/package=Hmisc>>. Acesso em: 27 nov. 2016.

HIRANO, Keisuke; IMBENS, Guido W. The propensity score with continuous treatments. In: GELMAN, Andrew; MENG, Xiao-Li. (Ed.). **Applied Bayesian modeling and causal inference from incomplete-data perspectives: an essential journey with Donald Rubin's statistical family**. Chichester, UK: John Wiley & Sons, 2004. p. 73-84.

KING, Gary; NIELSEN, Richard. **Why propensity scores should not be used for matching**. 2016. Disponível em: <j.mp/PScore>. Acesso em: 13 nov. 2016.

MENEZES-FILHO, Naércio. Os determinantes do desempenho escolar no Brasil. In: DUARTE, Pedro Garcia; SILBER, Simão D.; GUILHOTO, Joaquim J. M. (Org.). **O Brasil do século XXI**. São Paulo: Saraiva, 2011. v. 1. p. 231-256.

MENEZES-FILHO, Naércio (Org.). **Avaliação econômica de projetos sociais**. São Paulo: Fundação Itaú Social, 2016.

MISHEL, Lawrence; ROTHSTEIN, Richard. (Ed.). **The class size debate**. Washington, D.C.: Economic Policy Institute, 2002.

MOSTELLER, Frederick. The Tennessee study of class size in the early school grades. **Critical Issues For Children and Youths**, v. 5, n. 2, p. 113-127, 1995.

OECD. **Education at a Glance 2016: OECD indicators**. Paris: OECD Publishing, 2016.

OLIVEIRA, Jaqueline Maria de. **Custo-efetividade de políticas de redução do tamanho da classe e ampliação da jornada escolar: uma aplicação de estimadores de *matching***. Rio de Janeiro: BNDES, 2010.

PASEK, Josh. *et al.* **Weights: Weighting and Weighted Statistics**. Disponível em: <<https://cran.r-project.org/package=weights>>. Acesso em: 27 nov. 2016.

PIKETTY, Thomas; VALDENNAIRE, Mathieu. **L'impact de la taille des classes sur la réussite scolaire dans les écoles, collèges et lycées français: Estimations à partir du panel primaire 1997 et du panel secondaire 1995**. Paris: Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche / Direction de l'évaluation et de la prospective, 2006. (Les dossiers, n. 173)

PÓ, Marcos Vinicius; YAMADA, Erika; XIMENES, Salomão; LOTTA, Gabriela; ALMEIDA, Wilson Mesquita de. **Análise da política pública de Reorganização Escolar proposta pelo governo do Estado de São Paulo**. São Bernardo do Campo: Universidade Federal do ABC, 2015. Disponível em: <<https://blogdosalomaoximenes.files.wordpress.com/2015/12/anc3a1lise-da-reorganizac3a7c3a3o-escolar-sp.pdf>>. Acesso em: 15 nov. 2016.

RICE, Jennifer King. Making the evidence matter: Implications of the class size research debate for policy makers. In: MISHEL, Lawrence; ROTHSTEIN, Richard. (Ed.). **The Class Size Debate**. Washington, D.C.: Economic Policy Institute, 2002. p. 89-94.

TODOS PELA EDUCAÇÃO. **Análise da evolução do número de matrículas e turmas de Ensino Fundamental e Médio da rede estadual paulista**. São Paulo: Todos Pela Educação, jun. 2016.

TRAVITZKI, Rodrigo; CALERO; Jorge; BOTO; Carlota. What does the National High School Exam (ENEM) tell Brazilian society? **CEPAL Review**, v. 113, p. 157-174, 2014.

URQUIOLA, Miguel; VERHOOGEN, Eric. Class-size caps, sorting, and the regression-discontinuity design. **The American Economic Review**, v. 99, n. 1, p. 179-215, 2009.

WEISELFISZ, Jacobo. **Tamanho da turma: faz diferença?** Brasília, DF: Fundescola / MEC, 2000. (Série estudos, n. 12)

WICKHAM, Hadley. The split-apply-combine strategy for data analysis. **Journal of Statistical Software**, v. 40, n. 1, p. 1–29, 2011.

APÊNDICE

Detalhamento da aplicação do método GPS

O procedimento descrito a seguir foi realizado separadamente em cada uma das quatro amostras estudadas (Total, NSE baixo, NSE médio e NSE alto). Dentre as 47 variáveis disponíveis, selecionamos, com regressão linear, aquelas significativamente associadas tanto ao tratamento ($p < 0,01$) quanto à resposta ($p < 0,001$).

Definidas as variáveis de pré-tratamento, estimamos o GPS para cada aluno. Na amostra Total, o GPS apresentou média 0,059 e desvio padrão 0,019, com mínimo de zero, máximo de 0,077 e os seguintes quartis: 0,049; 0,067; e 0,075.

Estimamos, com isso, a esperança da resposta condicionada ao tratamento. Foi utilizado um modelo cúbico, que apresentou melhor ajuste do que o modelo linear. A Tabela 3 apresenta os resultados do modelo para as quatro amostras. Nota-se que os três coeficientes relativos ao GPS se mostraram significativos ($p < 0,001$, na maioria dos casos),

evidenciando a presença de viés de seleção associado ao tamanho das classes. Os intervalos de confiança foram estimados por *bootstrap* com 1.000 replicações.

TABELA 3 – Modelos da esperança da resposta condicionada ao tratamento

	Total		NSE baixo		NSE médio		NSE alto	
	coef.	p-valor	coef.	p-valor	coef.	p-valor	coef.	p-valor
Intercepto	504,748	0,000	134,869	0,000	110,370	0,113	363,194	0,000
T	61,382	0,000	27,399	0,000	25,907	0,000	47,809	0,000
log(gps)	37,823	0,000	37,470	0,000	24,433	0,000	27,668	0,000
log(gps)²	1,286	0,000	4,465	0,000	2,599	0,000	1,014	0,003
T²	1,729	0,000	0,744	0,000	0,670	0,000	1,274	0,000
log(gps)³	0,062	0,000	0,262	0,000	0,137	0,000	0,029	0,047
T³	0,017	0,000	0,007	0,000	0,006	0,000	0,012	0,000
T:log(gps)	1,037	0,000	0,466	0,000	0,260	0,002	0,663	0,000
R²	0,007		0,005		0,003		0,004	
p-valor	0,000		0,000		0,000		0,000	

Fonte: Elaborado a partir dos microdados da Prova Brasil 2013, MEC/Inep.

Agradecimentos

Os autores agradecem aos membros da Rede Escola Pública e Universidade, sem cuja generosa interlocução este trabalho não teria sido possível.ⁱ

ⁱRevisão gramatical do texto por: Maira Fátima de Oliveira Nobre