

PONTIFÍCIA UNIVERSIDADE CATÓLICA  
DO RIO DE JANEIRO



Tomás Ruiz de Gamboa do Valle

**IDEB e Accountability: Efeitos da Publicação do Índice na  
Educação Pública Municipal Brasileira**

Monografia de Conclusão de Curso

Orientador: Arthur Bragança

Coorientador: Arthur Aguillar

Rio de Janeiro, Brasil

Novembro de 2021

# Sumário

<b>Sumário</b>	<b>3</b>	
<b>1</b>	<b>Introdução</b>	<b>4</b>
<b>2</b>	<b>Contexto Institucional</b>	<b>8</b>
<b>3</b>	<b>Dados</b>	<b>14</b>
<b>4</b>	<b>Metodologia</b>	<b>18</b>
<b>5</b>	<b>Resultados</b>	<b>23</b>
<b>6</b>	<b>Conclusão</b>	<b>29</b>
<b>7</b>	<b>Referências</b>	<b>30</b>

# 1 Introdução

Atualmente, é inquestionável, na literatura, a importância da qualidade da educação, não só pelos seus impactos diretos, mas muito também pelas suas diversas externalidades positivas em áreas como saúde, criminalidade e “boa cidadania” (Lochner, 2011). Além disso, se trata de um determinante central para redistribuição de rendimentos (Peracchi, 2006) e crescimento econômico (Hanushek e Woessmann, 2008). Diante dessa realidade, é essencial entender os desafios enfrentados pelo Brasil na área da educação e, de forma igualmente importante, quais políticas são eficazes e custo-efetivas para superá-los.

Nesse sentido, em sua obra “Os determinantes do desempenho escolar do Brasil” (2007), Naércio Menezes-Filho busca elencar os mais relevantes fatores da performance acadêmica no país. Em resumo, o principal conjunto de determinantes identificado foi o de características pessoais e familiares, como cor da pele, escolaridade dos pais e renda. Em segundo lugar, temos a já conhecida diferença entre a rede pública e privada, bem como a menos discutida disparidade entre os colégios do primeiro grupo. Este último fator também foi documentado por Felício e Fernandes (2005) e denominado “Efeito Escola”. Para Menezes-Filho, a gestão tem papel central neste efeito; ele defende em sua publicação que as soluções para a educação brasileira já estão presentes e implementadas em algumas instituições públicas ao redor do país, restando identificá-las e replicá-las, o quanto possível. Somam-se a este fato as evidências empíricas apresentadas em Camargo et al. (2014) de que a publicação das notas do Exame Nacional do Ensino Médio (ENEM) gera, no setor privado (em que há incentivos de mercado), uma melhoria de performance devido a um choque de *accountability* que mobiliza a gestão. Nesse sentido, parece plausível que algo semelhante possa ocorrer no sistema público, mas através de outro sistema de incentivos.

Nesse contexto, o Índice de Desenvolvimento da Educação Básica (IDEB) foi criado em 2007 pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) para medir a qualidade de aprendizado no país e estabelecer metas para a melhoria do ensino. Deixando de lado, por ora, a simplicidade da meta estabelecida de atingir 6.0 até 2022 mesmo com a atual base de dados disponível, o indicador passou a ter papel importante no sistema educacional brasileiro. Além da sua função original

de monitoramento da qualidade educacional, pelo fato de o IDEB ser público e objetivo, pode gerar um choque de *accountability* que faria com que a sua existência incentivasse melhorias nas escolas da rede pública.

Nesse sentido, este trabalho propõe a investigação de um dos diversos canais possíveis, através do qual um choque de *accountability* ocasionado pela publicação do IDEB pode afetar de fato a qualidade da educação brasileira. Em uma perspectiva mais abrangente, existe uma gama razoável de possíveis mecanismos, seja via investimentos ou via organização e gestão (Akhtari, Moreira e Trucco, 2017; Ferraz, Finan e Moreira, 2012). A hipótese proposta a ser testada é de que a publicação do IDEB de um município (ou de uma escola) representaria um choque de *accountability* que, por sua vez, levaria a prefeitura (ou a gestão) a alterar a alocação de recursos monetários e/ou humanos disponíveis de modo a gerar um melhor resultado no indicador. Levando em consideração o Brasil e o Mundo, já é possível coletar um forte conjunto de evidências empíricas de que sistemas de *accountability* podem mover de fato os agentes e, conseqüentemente, os indicadores. Um exemplo brasileiro interessantíssimo está associado à figura do político oportunista, já que o eleitorado brasileiro tende a ver muito positivamente investimentos em educação (Sakurai, 2009), ainda que seu usufruto possa ser objeto de conflitos intergeracionais nas famílias (Bursztyn e Coffman, 2012). Nesse mérito, evidências empíricas indicam que maiores investimentos em educação estão associados a maiores chances de reeleição do incumbente (Assunção e Estevan, 2019) e, em especial, que as prefeituras que foram capazes de aumentar um ponto no IDEB de seus municípios entre 2005 e 2007 aumentaram sua probabilidade de reeleição em algo entre 5 e 10 pontos percentuais em 2008 (Firpo, Pieri e Souza, 2016).

Desde a proposição e implementação do IDEB em 2007, por parte de Reynaldo Fernandes, o próprio voltou a escrever sobre o indicador em um contexto de *accountability* por algumas vezes até os dias de hoje, das quais destaco as mais intimamente ligadas com o tema deste trabalho: Fernandes (2007); Fernandes & Grimaud (2009); Fernandes (2016) e Fernandes & Grimaud (2020). Estes trabalhos não só fornecem excelentes contextualizações e referências teóricas como também, ao longo do tempo, revelam os avanços e descobertas feitos pelo criador e/ou colegas. Ao fim desta saga, Fernandes já publicou uma série de excelentes argumentos a favor da publicação do IDEB e do sistema de *accountability*

que o mesmo protagoniza, debateu as principais críticas e até as deu razão, quando necessário. Em seus trabalhos mais recentes, tenta avaliar o efeito deste sistema que ajudou a instalar na melhoria das escolas brasileiras ao longo dos anos, de forma convincente mas majoritariamente qualitativa. Nesse sentido, este trabalho pretende contribuir com esta literatura fornecendo o que, até onde vai o conhecimento do autor, seria a primeira estimativa quantitativa do efeito de *accountability* de se publicar o IDEB ao nível da escola.

Para isso, foram utilizados uma série de dados disponibilizados pelo INEP. Dados acerca das Turmas, das Escolas e dos Docentes por elas responsáveis foram obtidos do Censo Escolar, enquanto as notas do IDEB e do SAEB foram obtidas da base do IDEB em si. Depois de limpar os dados, aproveito a descontinuidade causada pela regra de participação do SAEB, que requer um mínimo de 20 alunos na última série do Ensino Fundamental I para que haja aplicação da prova. Como essa nota é necessária para gerar o IDEB, é possível usar esse *cutoff* arbitrário no tamanho da série para estimar um Fuzzy RDD que avalia o efeito da publicação do IDEB (tratamento de *accountability*) na Taxa de Aprovação desta mesma última série no período de 2009 a 2015 - período escolhido para esta análise justamente pela estabilidade da regra de *cutoff*.

Desse modo encontro resultados significativos, com 99% de confiança, que revelam um impacto de 2,09 pontos percentuais (LATE) na taxa de aprovação média daquelas escolas que, por pouco, publicam o IDEB quando comparadas àquelas que, por pouco, não publicam. Além disso, temos que a primeira publicação do IDEB (em um passado recente) oferece um efeito ainda maior na Taxa de Aprovação, de 4,37 pontos percentuais. Este segundo efeito pode indicar que há uma maior pressão nos primeiros anos de divulgação, que resulta em maiores resultados, bem como pode representar evidência de *gaming* e *score inflation*. Serão necessárias maiores investigações futuras para determinar qual é o caso. De todo o modo, os efeitos encontrados sugerem que é positivo para as escolas publicarem o IDEB, especialmente em um país marcado por alta reprovação que, por sua vez, é uma das principais causas da evasão.

O restante deste documento está dividido em 6 seções. A segunda consiste em uma contextualização institucional reunindo as informações necessárias acerca do sistema educacional brasileiro e, principalmente, do IDEB para o melhor entendimento do restante do trabalho. O documento conta, também, com uma seção de Metodologia, subdividida

entre “Arcabouço Teórico”, que fornece toda a teoria necessária para a realização do RDD e “Estratégia Empírica” que de fato explicita como a estimação foi feita neste caso em particular. Além disso, apresento: capítulo de Dados, com maiores detalhes acerca das bases utilizadas neste trabalho; Resultados, que reúne os principais resultados obtidos nas estimações; Conclusão, onde discutirei os resultados obtidos, suas implicações e suas restrições; e Referências onde apresento as fontes consultadas na feitura deste trabalho e de suas versões futuras.

## 2 Contexto Institucional

### I. Indicadores Educacionais no Brasil

Desde o final dos anos 2000, a educação básica brasileira foi amplamente considerada universal – ou quase – em termos de acesso. No entanto, trata-se ainda de uma afirmação que dispõe de controvérsias, como as críticas embasadas pelo relatório acerca da pobreza na infância e na adolescência publicado pelo braço brasileiro do Fundo das Nações Unidas para a Infância (UNICEF, sigla em inglês) em 2019. Foi documentado pelo órgão internacional que 6,5% dos brasileiros entre 4 e 17 anos de idade não tem qualquer acesso à educação, considerada “privação extrema”. Além desta parcela que da juventude que tem seu direito constitucional à educação negligenciado, outros 13,8% desta população foi considerada em “privação intermediária” pois vão à escola, mas seguem analfabetos ou não estudam em uma série apropriada para sua idade. Dentro deste contexto, parece razoável que, embora acesso à educação ainda seja uma questão neste país, a qualidade educacional provida aos 93,5% das crianças e adolescentes brasileiros seja a prioridade para o Ministério da Educação e Cultura (MEC) ao longo dos últimos anos.

Para uma compreensão adequada da proposta e implementação do IDEB, faz-se necessário um breve histórico das métricas de qualidade educacional no Brasil. O primeiro marco importante neste campo foi a implementação do Sistema de Avaliação da Educação Básica (SAEB), idealizado nos anos 1980 e realizado pela primeira vez no início dos anos 1990. Consistia em submeter amostras nacionais, regionais e estaduais de estudantes do último ano do Ensino Fundamental I (4<sup>a</sup> série/5<sup>o</sup> ano) e do Ensino Fundamental II (8<sup>a</sup> série/9<sup>o</sup> ano) a testes padronizados de proficiência em Matemática e Português, com algumas poucas exceções ao longo dos anos (Fernandes & Gremaud, 2009). No ano de 1995, a avaliação passou a incorporar a Teoria de Resposta ao Item (TRI), metodologia que calcula o valor de cada item da prova de forma mais independente de modo a garantir comparabilidade das notas ao longo dos anos. Nos anos seguintes, os resultados do SAEB ajudaram a trazer luz para importantes fraquezas do sistema educacional brasileiro (juntamente com PISA e outros exames internacionais) além de ter

iniciado a cultura de avaliação da qualidade educacional através de testes padronizados (Fernandes & Gremaud, 2009).

Uma década após a implementação do TRI, em 2005, foi criada a Prova Brasil para expandir o SAEB e aumentar a capilaridade das informações providas (Portaria Ministerial nº 931, de 21 de março de 2005). Nesse momento, o MEC acreditava que os resultados do SAEB, ainda que muito informativos, falhavam em capturar e identificar apropriadamente a diversidade e heterogeneidade observada dentro de cada estado e município. A Prova Brasil, então, foi desenvolvida para complementar as amostras do SAEB com a testagem em caráter censitário de sua população-alvo: turmas da 4ª série/5º ano e da 8ª série/9º ano de escolas públicas e urbanas com mais de 30 alunos matriculados. Essa mudança foi de suma importância pois permitiu que cada rede pública, inclusive a municipal, e que diversas escolas pudessem descobrir – e comparar – seus indicadores de qualidade educacional. Além disso, transformou um indicador que até então servia apenas para diagnósticos gerais em algo capaz de implementar um sistema de *accountability* e um conjunto de incentivos associados. No ano de 2007, o SAEB e a Prova Brasil foram unificados, utilizando então os mesmos questionários e metodologias de correção. Dessarte, a diferença que se manteve entre eles foi o fato de que o primeiro avalia os níveis nacional, regional e estadual através de amostragem enquanto a segunda cobre o nível municipal e da escola via uma abordagem censitária. Ademais, é importante salientar que a aderência a ambas as avaliações é voluntária por parte das escolas e entes governamentais. Ainda assim, o MEC conseguiu à época que todos os estados e 99% das secretarias municipais de educação se comprometessem a participar.

A última peça deste sistema em desenvolvimento foi o Exame Nacional do Ensino Médio (ENEM). Criado inicialmente em 1998 como uma avaliação voluntária para um diagnóstico geral do Ensino Médio – que não era abrangido pelo SAEB/Prova Brasil – o ENEM rapidamente se tornou o principal exame de admissão para o Ensino Superior em todo o país, registrando 4 milhões de participantes oriundos de cerca de 95% das escolas do país ainda em 2008 (Fernandes & Gremaud, 2009). Assim, pela primeira vez, era possível observar, medir e comparar o desempenho e a qualidade das três etapas da educação básica através de testes padronizados. Anos mais tarde, em 2017 a Prova Brasil passaria a incorporar o Ensino Médio em seu censo e oficialmente abandonaria este nome



para seguir unificado sob a sigla SAEB, deixando o ENEM essencialmente como prova de admissão universitária.

## II. O Indicador de Desenvolvimento da Educação Básica (IDEB)

A essa altura, o Brasil já dispunha em alguma medida daquilo que Hanushek e Raymond (2004; 2005) descreveram como sistema de “*accountability* fraca”, dado que os indicadores educacionais supracitados eram amplamente divulgados e propriamente identificados por escola ou rede pública de ensino. Tamanha disponibilidade de dados, especialmente no contexto da educação pública brasileira onde a escolha ou troca de instituição de ensino é extremamente limitada, permitia e encorajava que pressões políticas e sociais fossem aplicadas aos diretores, prefeitos e secretários de educação. Essa foi a primeira implementação nacional de uma estrutura de *accountability* na educação brasileira, embora experiências regionais já tivessem explorado estes mecanismos previamente (Andrade, 2008; Brooke, 2006).

Em 2007, Reynaldo Fernandes, então Presidente do INEP, braço de pesquisa do MEC, propôs e implementou o Indicador de Desenvolvimento da Educação Básica (IDEB). A ideia era complementar os resultados obtidos via SAEB/Prova Brasil de forma a gerar uma medida mais completa e eficaz de qualidade escolar. Fernandes justificou a necessidade de uma medida complementar, dentre outras coisas, com o argumento de que os testes padronizados sozinhos poderiam prover um incentivo perverso para que diretores e secretários de educação implementassem medidas que aumentam as taxas de retenção com o objetivo de aumentar as notas de uma dada série/ano ou diferentes estratégias para inflar as suas notas.

É importante salientar que a crítica de Fernandes ao uso dos testes padronizados não significava que ele descreditava o método, especialmente dadas as robustas evidências já disponíveis na época que mostravam o poder de um indicador deste tipo. Naquele momento, havia quem argumentasse a favor de um “indicador de insumos e processos” para a educação. Contudo, no início daquela mesma década, Eric Hanushek e Margaret Raymond (2002) haviam publicado um artigo acerca de sistemas de *accountability* educacionais e concluído que este tipo de indicador, embora popular, tendia a ser menos preciso para avaliações de performance estudantil e podia ser mais facilmente inflado sem ganho real de educação

(Hanushek & Raymond, 2002; Hanushek, 2002). Além disso, ainda que os testes fossem admitidamente limitados, haviam fortes evidências empíricas de que este tipo de notas padronizadas se correlaciona com importantes variáveis, como rendimentos na vida adulta (Bishop, 1989 and 1992; O’Neil 1990; Grogger and Eide 1993, Blackburn and Neumark, 1993 and 1995; Murnane, Willet and Levy, 1995; Neal and Johnson, 1996; Murnane et al., 2000; Murnane et al, 2001; Zax and Rees, 1998) e anos acumulados de educação (Currie and Thomas, 1999), que por sua vez está correlacionado com diversos indicadores importantes de saúde e criminalidade, dentre outros.

Nesse contexto, a ideia proposta por Fernandes foi de combinar os testes padronizados com uma medida de “fluxo de progressão” – em particular, a taxa média de aprovação da etapa de ensino em questão. Dessarte, o IDEB foi criado como produto da média padronizada das notas de Matemática e Português no SAEB/Prova Brasil do último ano de cada etapa de ensino (4<sup>a</sup> série/5<sup>o</sup> ano; 8<sup>a</sup> série/9<sup>o</sup> ano; 3<sup>o</sup> ano/3<sup>a</sup> série – EM) para cada unidade avaliada (escola, município, estado, região, país),  $N_{i,e}$  e a taxa média de aprovação da etapa de ensino inteira, para a mesma unidade,  $P_{i,e}$ . Assim, temos:

$$IDEB_{i,e} = N_{i,e} \times P_{i,e}$$

Em que  $i$  representa a unidade estudada dentre as opções supracitadas e  $e$  representa a etapa de ensino avaliada.

Uma das características mais interessantes deste índice criado em 2007 é o fato de que  $P_{i,e} = 1/T_{i,e}$ , em que  $T_{i,e}$  representa o tempo médio necessário para completar um ano escolar na unidade de avaliação  $i$  no etapa de ensino  $e$ , permitindo uma interpretação mais intuitiva do índice. Outra funcionalidade admirável está associada ao fato de que a fórmula acima é na verdade uma função Cobb-Douglas sob a especificação particular em que  $\alpha = \beta = 1$ , de forma que  $N_{i,e}$  e  $P_{i,e}$  possuem uma taxa de substituição de -1. Há duas principais razões para tal particularidade: i. Para que o indicador construído seja facilmente interpretado e compreendido, explicitando o *trade-off* entre as variáveis. Dessa forma, permite que tomadores de decisão compreendam a compensação necessária para manter o  $IDEB_{i,e}$  em um valor fixo caso  $N_{i,e}$  ou  $P_{i,e}$  sejam alterado; e ii. Para criar um incentivo para que diretores e secretários de educação minimizem as taxas de retenção

sem desconsiderar o aprendizado de fato realizado pelos estudantes.

Pouco tempo depois da implementação em 2007, o IDEB se tornou o mais relevante indicador de qualidade escolar no Brasil e se consolidou como ferramenta crucial para pesquisadores e *policymakers* no campo da educação – posição que mantém até os dias de hoje. Adicionalmente, o MEC criou um sistema de metas para cada unidade (escolas, municípios, estados, regiões e Brasil), por etapa de ensino, do sistema brasileiro. O principal objetivo deste programa é equalizar a qualidade escolar brasileira com aquela da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), quando deste cálculo, em 2022, no bicentenário da Independência do país. Para tal, foi feito um considerável esforço acadêmico para replicar os componentes do IDEB no grupo de países da OCDE de forma que foi obtido o valor 6,0, estabelecido então como média nacional (INEP, 2007). Ademais, com a intenção de reduzir a desigualdade entre escolas e municípios em termos de IDEB, cada unidade avaliada recebeu um conjunto diferente de metas bianuais (frequência de publicação do índice) de forma que fosse exigido das escolas com piores performances no indicador uma melhora mais rápida (com a promessa de auxílio técnico e financeiro da União) de tal forma que o esforço combinado das dezenas de milhares de escolas e dos mais de 5.500 municípios levariam o país ao grande marco de 6,0.

### III. Condições para a Publicação do IDEB

Os seguintes capítulos deste trabalho se dedicam a explorar o possível efeito de *accountability* gerado pela publicação do IDEB de uma dada unidade educacional. Para isso, é importante elucidar as regras e condições envolvidas na divulgação – ou não – do indicador e é a isto que esta seção tem como objetivo. Primeiramente, é necessário destacar que a análise empírica foi conduzida ao nível da escola, para todo o Brasil mas apenas para o Ensino Fundamental I (anos iniciais). Assim, esta seção somente se propõe a discorrer acerca das condições relevantes dentro deste recorte.

Embora elegantemente simples em sua fórmula, o poderoso indicador ainda não foi universalizado, de forma que diversas escolas e alguns municípios não possuem IDEB publicado sequer uma vez até os dias de hoje. Isso ocorre pois, diferentemente de  $P_{i,e}$  que é provido pelo Censo Escolar que abrange todas as escolas e unidades federativas,  $N_{i,e}$

é construído com as notas do SAEB/Prova Brasil, que não possuem o mesmo caráter universal. Assim, em poucas palavras, a publicação do IDEB depende da disponibilidade de resultados destes testes padronizados – salvo raras exceções de não conformidade com o Censo Escolar.

Como mencionado anteriormente, em 2005 o SAEB ganha um segundo braço, a Prova Brasil, que conduzia testagens de caráter censitário apenas em escolas públicas de Ensino Fundamental (I e II) do sistema público **urbano** regular que tivessem 30 ou mais alunos matriculados no último ano de cada uma das etapas de ensino avaliadas. No biênio seguinte as regras se mantiveram e em 2009 houve uma importante mudança (Portaria Inep nº 87, de 7 de maio de 2009): Prova Brasil agora avaliava de forma censitária escolas públicas de Ensino Fundamental (I e II) dos sistemas públicos **urbano e rural** regulares que somassem 20 ou mais alunos na 4ª série/5º ano (EFI) ou na 8ª série/9º ano (EFII).

Essa regulamentação se manteve com poucas mudanças até 2015, com a grande exceção do período sendo oriunda da Portaria Inep nº 410, de 03 de novembro de 2011, que permitia que os municípios que fizeram a transição do Ensino Fundamental de 8 para 9 anos requisitassem a não divulgação de seus resultados. Além disso, vale mencionar que neste período de 2009 a 2015 não havia então publicação do IDEB para escolas exclusivamente de Educação Profissional, EJA ou Educação Especial. De 2017 em diante, a Prova Brasil deixa de existir e vira oficialmente SAEB, incluindo então escolas regulares urbanas e rurais com pelo menos 10 alunos matriculados na última série de cada etapa de ensino, inclusive ensino médio.

A totalidade deste estudo foi conduzido com dados referentes ao intervalo 2009-2015, no qual há maior estabilidade de “regras de entrada”. Além disso, como explicado no capítulo seguinte, estas condições pautaram decisões importantes de amostragem e da construção das variáveis de tratamento para a formação de um grupo de controle com contrafactual crível.

## 3 Dados

A análise empírica foi realizada com uma base de dados em painel ao nível da escola, compreendendo todo o Brasil de 2009 a 2015. Esse painel foi construído utilizando diferentes dados educacionais originalmente disponibilizados publicamente pelo INEP. O período foi selecionado por ser o maior intervalo sem mudanças significativas na regra de entrada para o SAEB e, por conseguinte, do IDEB. Ainda assim, as bases de dados publicadas apresentam consideráveis diferenças e algumas questões de compatibilidade ao longo do período. Nesse contexto, foi de grande importância o trabalho da ONG *open source* brasileira Base dos Dados, que disponibiliza gratuitamente versões compatibilizadas e integradas de diversas bases de dados nacionais e internacionais. Todos os dados utilizados neste trabalho foram obtidos através da plataforma da organização, BD+.

### I. IDEB

A variável de interesse foi obtida com a base ao nível da escola consolidada pela Base dos Dados. Como trata-se de uma junção de bases, cada observação representa um etapa de ensino (Ensino Fundamental – anos iniciais; Ensino Fundamental – Anos Finais; e Ensino Médio) de uma escola brasileira que teve o IDEB publicado ao menos uma vez de 2005 a 2019. Neste trabalho, utilizo apenas os dados referentes ao Ensino Fundamental - anos iniciais, de forma que o arquivo disponibilizado apresenta: identificador único e universal da escola, rede de ensino responsável, município, variáveis de aprovação (último ano e média da etapa de ensino), notas do SAEB e sua média padronizada, IDEB e sua projeção. Vale salientar que a publicação do IDEB é bienal em anos ímpares.

Devido a diferenças no modelo de gestão que afetam os incentivos, foram mantidas apenas as escolas da rede pública municipal que, dentro do período de análise, representam cerca de 65% das escolas do país e por volta de 82% das escolas públicas. Além disso, foram selecionados apenas os anos de 2009 a 2015, como supracitado. Demais filtragens foram aplicadas posteriormente com auxílio de variáveis do Censo Escolar, como exposto a seguir.

Além da limpeza, também foram construídas nesta base de dados algumas alternativas de variável de tratamento, como: *dummies* para “entrada” (publicou IDEB no ano t mas não publicou em t-2) e “saída” (não publicou IDEB no ano t mas publicou em t-2); anos de tratamento (quantidade de anos publicando IDEB); e medidas de sinalização como publicar IDEB acima, em torno ou abaixo da média municipal.

## II. Censo Escolar

Os dados do Censo Escolar têm extrema importância neste trabalho pois, diferentemente do SAEB e do IDEB, tornam públicas informações de todas as escolas, sem regra de entrada. Assim a base oferece a maior parte das características disponíveis que são observáveis tanto nas escolas do grupo de controle quanto do grupo de tratamento. Os dados utilizados foram colhidos de três diferentes bases do Censo Escolar da Base dos Dados: i. ao nível da escola trazendo informações de infraestrutura, equipamentos, administração e serviços educacionais oferecidos; ii. ao nível da turma trazendo dados de atividades realizadas, número de alunos e necessidades especiais dos discentes; e iii. ao nível do docente, informando a qualificação e escolarização, além características pessoais como idade, sexo e etnicidade.

Pelas características acima, as variáveis do Censo Escolar foram utilizadas para filtrar a amostra. Foram mantidas apenas as observações referentes ao último ano do Ensino Fundamental I (4<sup>a</sup> série/5<sup>o</sup> ano) da rede municipal regular em atividade no ano do Censo e com regulamentação em dia. Foram retiradas da amostra também escolas com atendimento exclusivo para jovens com necessidades especiais (turmas especiais via de regra não realizam SAEB), bem como escolas em situações diferenciadas como educação indígena, devido às diferenças de operacionais e de contexto e ao fato de que não costumam realizar SAEB também. A ideia aqui foi criar os grupos de controle e de tratamento com características estruturais similares garantindo que o primeiro tenha as características necessárias para ter o IDEB publicado exceto a quantidade de alunos. Além disso, foram mantidos apenas os dados de 2009 a 2015, como supracitado.

### III. Indicadores Educacionais

Por fim, foi utilizada a base de dados de Indicadores Educacionais da Base dos Dados, que reúne uma série de indicadores oferecidos pelo INEP em sua plataforma em uma única base ao nível da escola de 2007 a 2020 (muitas das variáveis não possuem observações ao longo de todo o intervalo). Daqui foram retiradas variáveis como média de alunos por turma, média de horas-aula, taxa de distorção idade-série, taxa de abandono, adequação de formação do docente, dentre outros. Estas informações foram então adicionadas a base de dados já existente respeitando a amostragem já feita.

### IV. Base de Dados Final

Assim, a base final reúne todas as variáveis e informações expostas até aqui e consiste em um painel ao nível da escola abrangendo o Brasil de 2009 a 2015 e contendo 75.649 escolas municipais regulares. Em todos os anos observados, cerca de um terço das escolas da amostra possuem IDEB publicado e os outros dois terços não. A Tabela 1 a seguir apresenta as estatísticas descritivas dividindo entre tratados (Publicou IDEB = 1) e não tratados (Publicou IDEB = 0).

Tabela 1 – Estatísticas Descritivas por Status de Tratamento - 2009, 2011, 2013 e 2015

<b>Publicou IDEB == 1</b>	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Número de Alunos	98.949	62,9	41,5	1,0	32,0	81,0	625,0
Entrou no IDEB (D)	108.095	0,2	0,4	0	0	0	1
Taxa de Aprovação	108.095	89,7	8,7	16,7	85,2	96,3	100,0
Indicador de Rendimento	108.095	0,9	0,1	0,1	0,8	1,0	1,0
Número de Docentes	92.451	8,9	9,7	1,0	2,0	12,0	172,0
Share de Docentes Não-Brancos	92.451	0,6	0,4	0,0	0,2	1,0	1,0
Quantidade PCs Adms	23.911	3,2	3,4	1,0	1,0	4,0	78,0
Internet (D)	52.499	0,8	0,4	0,0	1,0	1,0	1,0
Share de Docentes sem Pós-Grad	86.369	0,6	0,4	0,0	0,3	1,0	1,0
Sala da Diretoria (D)	53.220	0,9	0,3	0,0	1,0	1,0	1,0
Sala dos Professores (D)	53.220	0,8	0,4	0,0	1,0	1,0	1,0
Quadra de Esportes (D)	53.220	0,5	0,5	0,0	0,0	1,0	1,0
Share de Docentes sem EF	92.451	0,001	0,02	0,0	0,0	0,0	1,0
EF de 8 anos (D)	98.949	0,2	0,4	0,0	0,0	0,0	1,0

<b>Publicou IDEB == 0</b>	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Número de Alunos	34.391	13,5	10,4	1,0	9,0	16,0	509,0
Entrou no IDEB (D)	46.976	0,0	0,0	0,0	0,0	0,0	0,0
Taxa de Aprovação	5.257	89,2	11,4	0,0	84,3	97,6	100,0
Indicador de Rendimento	5.249	0,9	0,1	0,2	0,8	1,0	1,0
Número de Docentes	30.591	2,7	3,1	1,0	1,0	4,0	212,0
Share de Docentes Não-Brancos	30.591	0,7	0,4	0,0	0,2	1,0	1,0
Quantidade PCs Adms	12.466	2,2	2,6	1,0	1,0	2,0	48,0
Internet (D)	71.204	0,2	0,4	0,0	0,0	0,0	1,0
Share de Docentes sem Pós-Grad	25.419	0,6	0,4	0,0	0,0	1,0	1,0
Sala da Diretoria (D)	85.567	0,3	0,5	0,0	0,0	1,0	1,0
Sala dos Professores (D)	85.567	0,2	0,4	0,0	0,0	0,0	1,0
Quadra de Esportes (D)	85.567	0,1	0,3	0,0	0,0	0,0	1,0
Share de Docentes sem EF	30.591	0,000	0,02	0,0	0,0	0,0	1,0
EF de 8 anos (D)	34.391	0,2	0,4	0,0	0,0	0,0	1,0



# 4 Metodologia

## I. Arcabouço Teórico

A metodologia empregada neste estudo para identificação dos efeitos é a conhecida como RDD (*Regression Discontinuity Design*) e é atribuída a Donald Campbell, psicólogo da educação, que o utilizou em diversos estudos desde 1960 quando estimou o efeito de receber prêmios de mérito (com base em nota) em variáveis educacionais futuras. No entanto, a popularidade do método começou a crescer no início dos anos 2000 e provavelmente está em seu pico histórico. A principal vantagem desta metodologia se deve ao fato de que apresenta uma forma amplamente crível de eliminar vieses de seleção e consequentemente permite estimar efeitos de tratamento para uma dada subpopulação da amostra. Esta seção apresenta um breve apanhado acerca do arcabouço teórico utilizado no RDD.

Para conduzir um estudo deste tipo, é necessário encontrar um contexto em que haja uma variável contínua ou aproximadamente contínua (como idade) associada a atribuição do tratamento na qual haja um ponto de *cutoff* em que a probabilidade de tratamento “salte” – chamamos essa variável de *running variable* e denotamos por  $R$ . Esse *cutoff* ( $c_0$ ) pode ser oriundo de regras sociais, burocracias ou essencialmente quaisquer outras fontes, desde que seja conhecida, minimamente precisa e livre de manipulações. Esta última exigência precisa ser empiricamente testada. Ademais, a depender da precisão mencionada teremos diferentes tipos de RDD. Caso ultrapassar o *cutoff* leve a probabilidade de tratamento de 0 para 1 exatamente temos um RDD do tipo “Sharp”, onde o tratamento é uma função determinística da *running variable*  $R$  de tal forma que  $T_{i,t} = 1$ , se  $R_{i,t} \geq c_0$  e  $T_{i,t} = 0$ , se  $R_{i,t} < c_0$ .

Por outro lado, caso haja um salto, mas não haja essa relação determinística, é empregado um RDD do tipo “Fuzzy” onde a descontinuidade é observada na probabilidade condicional de tratamento associada a  $R$ , na forma:

$$\lim_{R_{i,t} \rightarrow c_0} \Pr(T_{i,t} = 1 \mid R_{i,t} = c_0) \neq \lim_{c_0 \leftarrow R_{i,t}} \Pr(T_{i,t} = 1 \mid R_{i,t} = c_0)$$

Nesse caso, a variável  $R$  não determina perfeitamente a variável de tratamento  $T$ , de modo que é necessário empregar uma estratégia similar a uma Variável Instrumental (VI) para eliminar a endogeneidade nas estimações, sendo este o diferencial do Fuzzy RDD. Nesse sentido, faz-se necessário que as variáveis respeitem, também as condições de VI, o que normalmente não é problemático dado que se costuma instrumentar o tratamento  $T$  com a variável  $Z_{i,t} = 1$  se  $R_{i,t} \geq c_0$  ( $Z_{i,t} = 0$  em caso contrário).

Algumas precauções são importantes para a utilização destas metodologias. A primeira e principal é a credibilidade de que, existe um conjunto de observações ao redor do *cutoff* que é tão próxima em características observáveis e não-observáveis, pode-se argumentar, que o *cutoff* arbitrário pode ser interpretado como uma atribuição aleatória de tratamento. Isto é, existe um grupo de não tratados que chamaremos de “quase-tratados” que estavam logo abaixo do *cutoff* mas não entraram e um outro grupo, “quase-não-tratados” que recebe o tratamento, mas se encontra bastante próximo do *cutoff*. Através do RDD se calculará então a diferença de  $Y$  (variável de resultado) nesta subamostra altamente local e encontrará o LATE (Efeito de Tratamento Médio Local, sigla em inglês). Esta dinâmica requer que as amostras sejam bastante grandes para trabalhar com este método, pois só entraram no cálculo as observações próximas do *cutoff*.

Se houver credibilidade na narrativa acima, resta essencialmente verificar dois outros pontos para que argumentar que há identificação correta do efeito causal. O primeiro teste que deve ser realizado para defender a credibilidade do experimento ficou conhecido como teste de McCrary, que visa verificar algumas das possíveis vulnerabilidades do experimento, em especial a manipulação do tratamento. Para que o racional exposto acima valha é necessário que as unidades observadas não estejam manipulando seu próprio status de tratamento e o teste de McCrary verifica se a densidade das observações é contínua ao redor do *cutoff* – caso contrário, há evidência de manipulação.

Por fim, como a hipótese de identificação é baseada em argumentos e dificilmente pode ser testada, é comum que se busque mostrar sua credibilidade. Assim, a segunda forma que isso é feito é através de um “teste de balanceamento”, que neste caso consiste

na análise do comportamento das demais variáveis utilizadas ao redor do mesmo *cutoff*. As demais variáveis utilizadas como controles não podem apresentar saltos ao redor do *cutoff*, dado que devem ser características pré-tratamento ou não correlacionadas com o tratamento e não deveriam ser afetadas dessa forma.

## II. Estratégia Empírica

Esta seção é dedicada a explicar e descrever os procedimentos e métodos realizados para estimação dos efeitos. Vale salientar novamente que esta análise foi feita apenas com escolas regulares em situação de funcionamento ativa, adequadamente regulamentadas e que oferecem turmas de 1º ao 5º ano (ou 1ª a 4ª série) do Ensino Fundamental (etapa de ensino de sigla EFI). Como existem diversas diferenças em termos de organização, gestão e incentivos entre EFI, EFII e Ensino Médio, escolho trabalhar apenas com a primeira etapa, principalmente pois: i. devido a fase da vida em que se encontram os alunos, há menos diferenças de *background* entre eles do que em qualquer outra fase, bem como menos possíveis influências externas que poderiam afetar seu desempenho; e iii. outros estudos apontam que é a etapa de ensino que responde mais expressivamente a este tipo de política e conseqüentemente onde se espera medir efeitos com mais clareza (Fernandes & Gremaud, 2020).

Caso a publicação – ou não – do IDEB de uma dada escola  $i$  em um dado ano  $t$  fosse totalmente aleatória, bastaria estimar um simples modelo como  $Y_{i,t} = \alpha + \beta T_{i,t} + \epsilon_{i,t}$  por Mínimos Quadrados Ordinários e teríamos, de forma simples, o efeito desejado. Como o que acontece de fato é que o número de alunos matriculados no último ano do Ensino Fundamental I (anos iniciais) tem papel determinante nesta atribuição de tratamento, um modelo como o acima ofereceria uma estimativa viesada do efeito de tratamento. De posse dessa atribuição um pouco mais complexa, há necessidade, também, de uma metodologia mais sofisticada do que a acima.

No caso estudado, o número de alunos cursando a 4ª série/5º ano é o que torna – ou não – uma escola elegível para o SAEB o que significa que a mesma participará do exame, salvo raras exceções. Oficialmente, 20 ou mais alunos nesta série/ano são suficientes para tal. Uma vez que a escola participa do SAEB e obtém uma nota média padronizada  $N_{i,t}$ , basta multiplicar pelo Indicador de Rendimento (média das taxas de aprovação de

todos os anos da etapa de ensino),  $P_{i,t}$  e temos o *IDEB* da escola  $i$  no ano  $t$  que será então publicado. Apesar de relativamente raros dentro da amostra, existem cenários em que escolas com menos de 20 alunos publicam IDEB e cenários em que escolas com mais de 20 não publicam (possíveis razões brevemente discutidas no Contexto Institucional). Assim a atribuição do tratamento não é função determinística do número de alunos (nossa *running variable*) e se faz necessário um Fuzzy RDD, como exposto na seção anterior.

Assim, temos que primeiro definir a relação entre a *running variable* e o status de tratamento em si, como abaixo:

$$T_{i,t} = \delta 1\{R_{i,t} \geq 20\} + f(R_{i,t}) + X_{i,t}\Gamma + \theta_t + \mu_{i,t} \quad (4.1)$$

De tal forma que  $R_{i,t}$  é o número de alunos da escola  $i$  no ano  $t$ ,  $1\{R_{i,t} \geq 20\}$  é um função indicadora que assume o valor 1 para escolas com 20 ou mais alunos,  $f(R_{i,t})$  é uma função suave do número de alunos que pode variar em ambos os lados da descontinuidade,  $\theta_t$  representa efeitos fixos de ano,  $\mu_{i,t}$  é o termo de erro e  $X$  é um vetor de controles que incluem: total de docentes, *share* de docentes não-brancos, quantidade de computadores administrativos, internet (*dummy*), *share* de docentes sem pós-graduação, sala para diretoria (*dummy*), sala dos professores (*dummy*), quadra de esportes (*dummy*), *share* de docentes com Ensino Fundamental Incompleto e *dummy* que indica se o EF é de 8 anos (1) ou 9 anos (0).

Importante ressaltar que foram realizadas também versões desta regressão em que  $T_{i,t}$  é a variável de Entrada no IDEB, *dummy* que retorna 1 caso a escola tenha publicado IDEB no ano  $t$  sem ter publicado no ano  $t - 2$  (IDEB é bienal), e não simplesmente Publicou IDEB (D).

Assim, temos como segundo estágio:

$$Y_{i,t} = \alpha + \beta T_{i,t} + f(R_{i,t}) + X_{i,t}\lambda + \rho_t + \epsilon_{i,t} \quad (4.2)$$

Em que instrumentamos o tratamento  $T_{i,t}$  da escola  $i$  no ano  $t$  com a equação (5.1) acima;  $\alpha$  é o intercepto,  $\beta$  é o coeficiente associado ao tratamento,  $Y_{i,t}$  é a variável de

resultado que, nas especificações principais, foi a Taxa de Aprovação (referente ao último ano do EFI) e o restante dos termos é análogo aos de (5.1) expostos acima.

Bem como em qualquer regressão descontínua, a hipótese de identificação é que conforme nos aproximamos da descontinuidade pela esquerda ou pela direita, temos um subconjunto de cada lado com escolas essencialmente idênticas antes do tratamento. Assim, espera-se que na ausência do mesmo, estas escolas todas teriam resultados muito similares e por isso as não tratadas dentro deste subconjunto podem ser utilizadas como contrafactuais para as tratadas.

Importante ressaltar que em nenhuma das especificações e amostragens o teste de McCrary indicou qualquer problema de manipulação. Além disso, nenhuma das variáveis aqui utilizadas como controle possui saltos ao redor do *cutoff*.

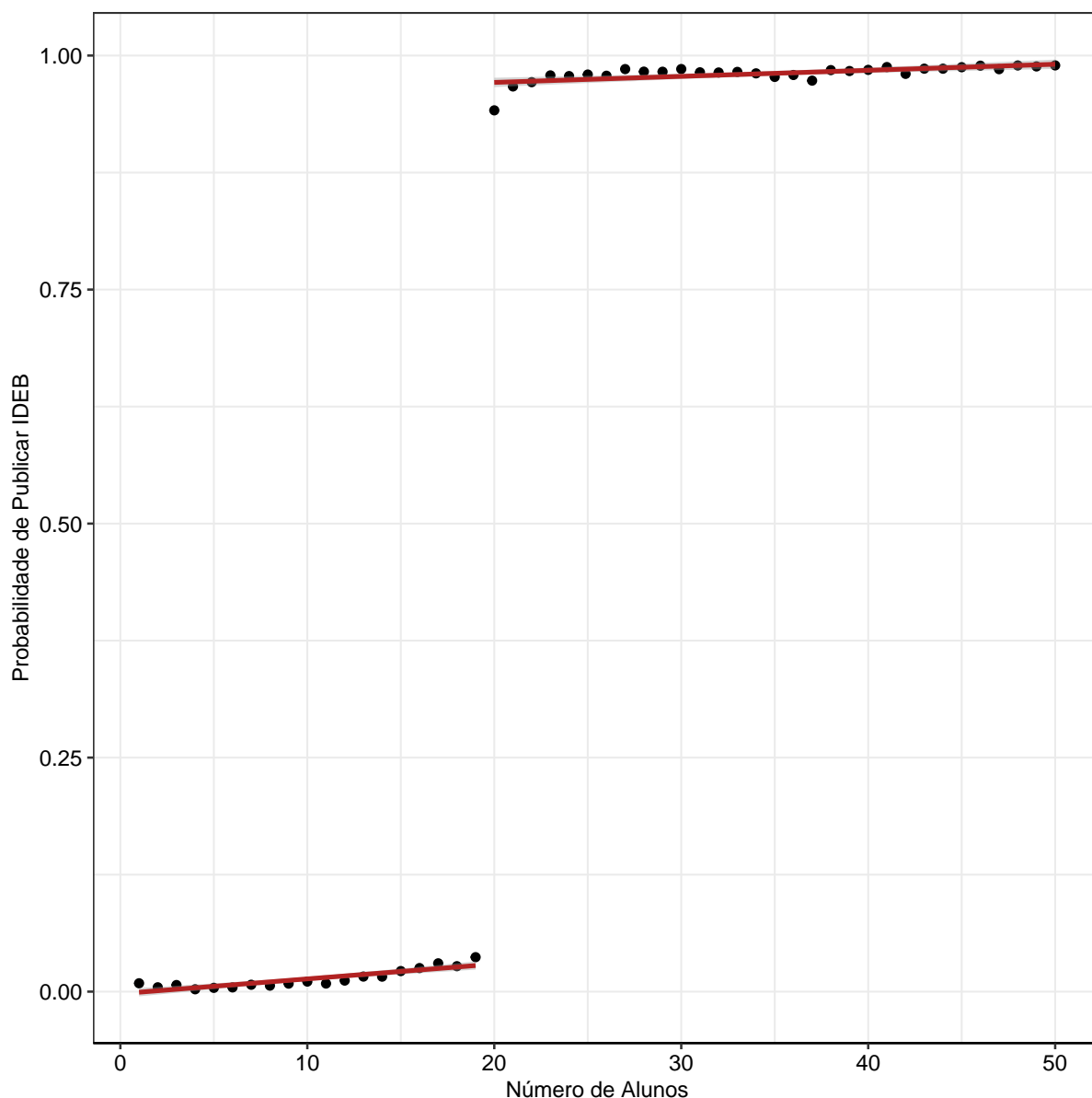
# 5 Resultados

## I. Número de Alunos e a Probabilidade de Publicar IDEB

Primeiro, é importante entender a relação entre a *running variable* e o status factual de tratamento (Publicou IDEB = 1 ou Publicou IDEB = 0). Talvez a forma mais interessante de fazê-lo seja através de um gráfico de probabilidade condicional, que associa cada valor da variável *running* (número de alunos) a uma probabilidade de publicar IDEB.

Olhando para a Figura 1, é possível observar um nítido salto de pouco mais de 85 pontos percentuais na probabilidade de ter o IDEB publicado quando vamos de 19 para 20 alunos. Além disso, vemos que não há outliers e que há uma leve inclinação positiva nas tendências, tanto antes quanto depois do *cutoff*, fortalecendo a ideia de que, em toda a amostra, quanto mais alunos, mais chance de publicar o IDEB. De posse deste gráfico, podemos ver que a escolha das variáveis e do contexto para este RDD estão bastante adequados. Usaremos o desenho Fuzzy para lidar com as pequenas variações que geram essa inclinação positiva.

Figura 1 – Probabilidade Condicional de Publicar IDEB por Número de Alunos



A Tabela 2 traz os resultados de Primeiro Estágio sob 6 especificações distintas. Embora haja uma significância estatística muito boa em todos os coeficientes encontrados, fica nítido que as especificações que mostram uma relação mais forte entre a variável de tratamento e número de alunos são aquelas em que o tratamento é definido pela variável Publicou IDEB (D). Isso se deve ao fato de que a variável de Entrada conta como tratados apenas aquelas escolas que publicaram IDEB no ano  $t$  mas não haviam publicado no ano  $t - 2$ . Ainda que tenha relação forte, a correspondência entre tratamento designado e realizado é naturalmente menor.

Além disso, vale destacar que a especificação (1), na qual são utilizados apenas efeitos fixos de ano como controle rendeu o coeficiente mais próximo de 1, bem como o maior  $R^2$ , 0,71.

Tabela 2 – Resultados - Primeiro Estágio

	<i>Variável Dependente</i>					
	Publicou IDEB (D)			Entrou no IDEB (D)		
	(1)	(2)	(3)	(4)	(5)	(6)
Número de Alunos $\geq 20$	0,821*** (0,008)	0,700*** (0,021)	0,701*** (0,022)	0,463*** (0,022)	0,360*** (0,035)	0,335*** (0,039)
Band Width	6	6	6	6	6	6
Controles Escola	não	sim	sim	não	sim	sim
Controles Docentes	não	não	sim	não	não	sim
Efeitos Fixos (Ano)	sim	sim	sim	sim	sim	sim
Observations	15.326	3.749	2.911	15.326	3.749	2.911
$R^2$	0,71	0,57	0,58	0,12	0,18	0,14

*Note:*

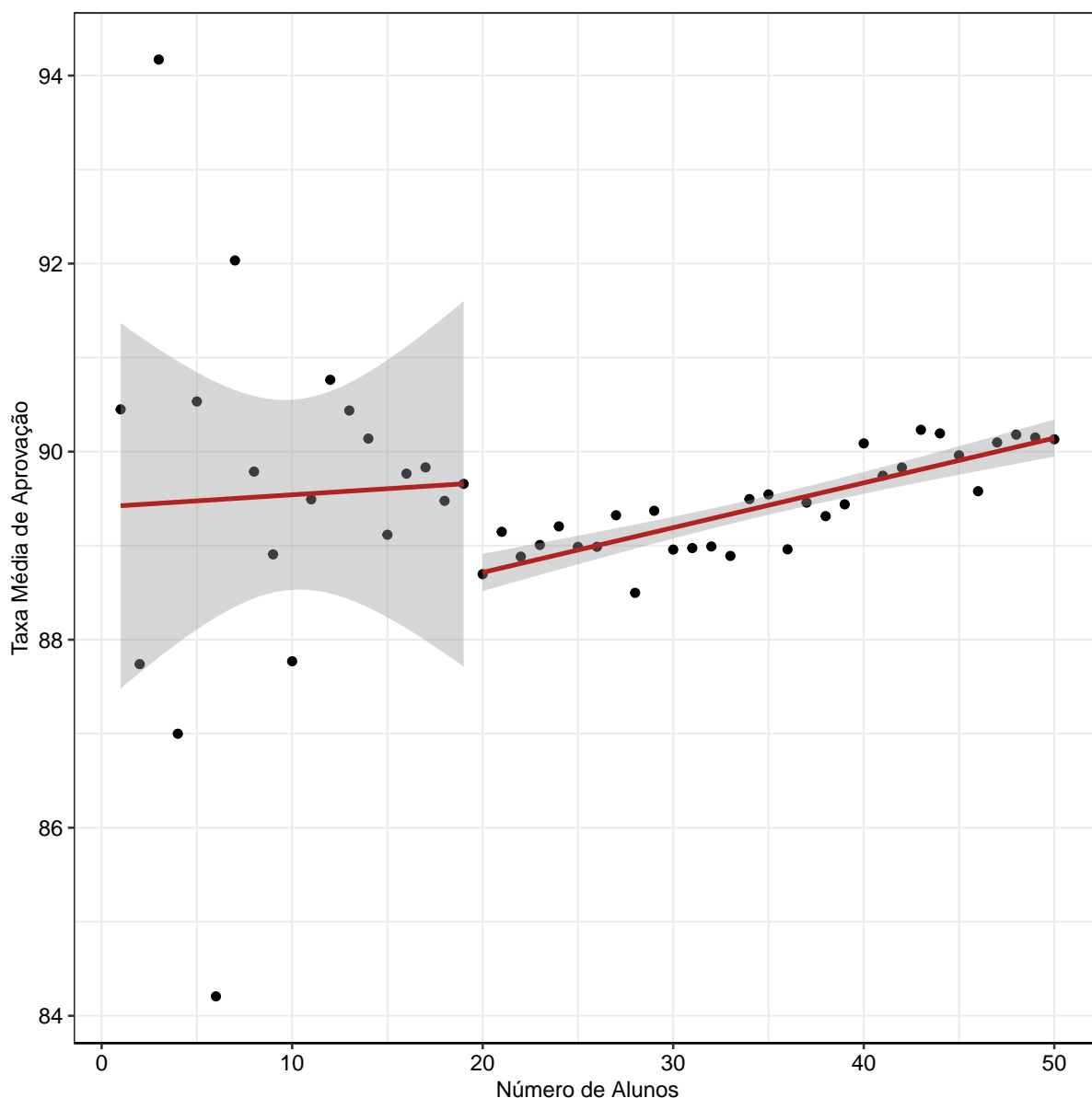
\* $p < 0,1$ ; \*\* $p < 0,05$ ; \*\*\* $p < 0,01$

## II. Publicação do IDEB e Aprovação Escolar

Agora, vamos analisar cuidadosamente o que podemos inferir acerca da Taxa de Aprovação (sempre medida no último ano do EFI) e a fora que se relaciona com o a publicação do IDEB. A Figura 2 mostra uma nítida mudança de padrão a partir de 20 alunos na série/ano. Entre 0 e 19 alunos, observamos pontos espalhados com distribuição aparentemente consideravelmente aleatória – sobe e desce muito rápido e sem padrão. Dos 20 alunos em diante, no entanto, as Taxas Médias de Aprovação por Número de Alunos formam uma tendência clara de crescimento e o conjunto de pontos tem aparência coesa e uniforme. Embora a taxa de aprovação seja sempre coletada pelo Censo Escolar, independentemente do SAEB e do IDEB, dentre as escolas menores e que não publicam o IDEB, realmente não parece haver padrão claro que explique o comportamento. Ainda assim, a tendência calculada por uma regressão linear simples é positiva (com grandes intervalos de confiança) e a média é alta em relação às escolas com 20 a 30 alunos, ainda que isto não seja muito informativo por si só.



Figura 2 – Taxa Média de Aprovação por Número de Alunos



A Tabela 3 nos trás os resultados de 3 especificações diferentes de controles para 2 especificações diferentes de tratamento. Quando utilizamos apenas os efeitos fixos de ano, em(1), temos um efeito negativo e estatisticamente significativo com 99% de confiança, condizente com o que observamos no gráfico, composto pelas mesmas duas variáveis. Quando incluimos controles de Escola, como regime de Ensino Fundamental (8 vs 9 anos) e variáveis de infraestrutura como presença de quadra de esportes e disponibilidade de internet, vemos o efeito desaparecer. Assim, no modelo (2), temos efeito não significativo estatisticamente e de baixa magnitude.

No entanto, esse panorama muda quando juntamos a este conjunto de controles as variáveis dos docentes referentes a qualificação, como *share* de docentes sem pós graduação e *share* de docentes com EF incompleto, bem como referentes a características pessoais, por exemplo *share* de docentes não-brancos. Agora, levando em conta as condições físicas escolares e as informações disponíveis acerca dos docentes responsáveis pelo ensino, vemos que na verdade o efeito de publicar o IDEB é positivo na Taxa de Aprovação, bastante mais intuitivo.

Agora, nos debrucemos sobre as especificações (4)-(6). Como debatido anteriormente, a escolha da entrada no IDEB como tratamento enfraquece o primeiro estágio, contudo não o suficiente para invalidá-lo ou para perder a confiança de 99%. Tendo dito isto, os resultados seguem um padrão muito similar às especificações (1), (2) e (3) em termos de sinal e significância, em algum nível reforçando esta narrativa. Talvez o mais interessante aqui sejam as magnitudes de (4) e (6) pois nos levam a crer que as escolas que “entram” no IDEB após ao menos um período sem participar estão experienciando efeitos mais fortes em qualquer que seja a direção.

Por fim, foquemos na especificação e número (6): temos um coeficiente positivo, significativo com 99% de confiança e que apresenta um valor pouco mais de 2 vezes maior que (3). Isto indica que o subgrupo do tratamento Publicar IDEB que não havia publicado no biênio anterior parece receber um impacto maior do que o grupo como um todo e conseqüentemente que o seu conjunto complementar (Publicou IDEB em  $t$  e tinha publicado em  $t - 2$ ). As possíveis implicações destes – e outros – resultados serão melhor debatidas no capítulo seguinte.

Tabela 3 – Resultados - Fuzzy RDD

	<i>Variável Dependente</i>					
	Taxa de Aprovação					
	(1)	(2)	(3)	(4)	(5)	(6)
Publicou IDEB (D)	-2,05*** (0,56)	0,32 (0,97)	2,09*** (1,03)			
Entrou no IDEB (D)				-3,64*** (0,99)	0,631 (1,90)	4,37*** (2,21)
Band Width	6	6	6	6	6	6
Controles Escola	não	sim	sim	não	sim	sim
Controles Docentes	não	não	sim	não	não	sim
Efeitos Fixos (Ano)	sim	sim	sim	sim	sim	sim
Observations	15.326	3.749	2.911	15.326	3.749	2.911
R <sup>2</sup>	0,06	0,13	0,18	0,05	0,13	0,13

*Note:*

\*p&lt;0,1; \*\*p&lt;0,05; \*\*\*p&lt;0,01

## 6 Conclusão

Os principais resultados obtidos indicam que, embora um primeiro olhar possa revelar uma relação negativa entre aprovação e publicação do IDEB, a relação se torna positiva quando os controles necessários são inseridos. Por se tratar da taxa de aprovação do último ano do EFI, este resultado do modelo (3) associa nossa métrica de *accountability* com maior taxa de aprovação em 2,09 pontos percentuais no ano de conclusão desta etapa de ensino em um país marcado pela repetência e consequente evasão em etapas posteriores.

Além disso, temos que o efeito é ainda maior quando se publica o IDEB sem tê-lo feito nos últimos 2 anos (4,37 pontos percentuais), talvez indicando que o choque inicial de saída do não observável para o público é mais forte do que o efeito estático de permanecer neste grupo. Outra visão interessante deste efeito que poderia explicar a maior magnitude na “entrada” é o fato de que o Indicador de Rendimento (que tem essa taxa como componente) é *input* para o IDEB de forma que escolas que passam a publicá-lo podem querer inflar suas taxas de aprovação para inflar seu primeiro IDEB. Essa no entanto, seria uma tática de *gaming* e *score inflation* que mereceria testes de robustez associados. Assim um *follow-up* deste estudo teria muito o que explorar, como esta especulação e a possível consequência destes efeitos na taxa de evasão dos anos posteriores, que são extremamente interessantes e merecem maiores investigações. Até lá, no entanto, não podemos tirar grandes conclusões. Outro vulnerabilidade que precisa ser levantada é o fato do efeito ser extremamente local, oferecendo ameaça a validade externa do experimento.

Ainda assim, é importante destacar que, até onde vai o conhecimento do autor, este é o primeiro estudo que conseguiu calcular efeitos significantes de *accountability* do IDEB no Brasil ao nível da Escola. Claro que sempre há melhorias e expansões a serem feitas, muitas das quais já idealizadas. Fora isso, é preciso salientar que resultados como este e oriundos de eventuais *follow-ups* deste estudo tem grande relevância no desenho de política pública, mostrando o potencial que um sistema de *accountability* tem para mover os agentes públicos e privados na direção correta quando bem construído.

## 7 Referências

ANDRADE, Eduardo de Carvalho. **"School accountability" no Brasil: experiências e dificuldades**. Brazilian Journal of Political Economy, v. 28, n. 3, p. 443-453, 2008.

AKHTARI, Mitra; MOREIRA, Diana; TRUCCO, Laura. **Political turnover, bureaucratic turnover, and the quality of public services**. In: Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association. National Tax Association, 2017. p. 1-84.

ASSUNÇÃO, Matheus; ESTEVAN, Fernanda. **Do Voters Reward Politicians for Education Expenditures?**. 2019.

BISHOP, John H. **Is the test score decline responsible for the productivity growth decline?**. The American Economic Review, p. 178-197, 1989.

BLACKBURN, McKinley L.; NEUMARK, David. **Omitted-ability bias and the increase in the return to schooling**. Journal of labor economics, v. 11, n. 3, p. 521-544, 1993.

BLACKBURN, McKinley L.; NEUMARK, David. **Are OLS Estimates of the Return to Schooling Biased Downward? Another Look**. The Review of Economics and Statistics, p. 217-230, 1995.

BORUSYAK, Kirill; JARAVEL, Xavier. **Revisiting event study designs**. Available at SSRN 2826228, 2017.

BROOKE, Nigel. **O futuro das políticas de responsabilização educacional no Brasil**. Cadernos de Pesquisa, v. 36, n. 128, p. 377-401, 2006.

BURSZTYN, Leonardo; COFFMAN, Lucas C. **The schooling decision: Family preferences, intergenerational conflict, and moral hazard in the Brazilian favelas**. Journal of Political Economy, v. 120, n. 3, p. 359-397, 2012.

CAMARGO, B., CAMELO, R., FIRPO, S., PONCZEK, V. P. **Information, market incentives, and student performance.** 2014.

CURRIE, Janet; THOMAS, Duncan. **Early test scores, socioeconomic status and future outcomes.** National bureau of economic research, 1999.

DE CHAISEMARTIN, Clement; D'HAULTFOEUILLE, Xavier. **Two-way fixed effects estimators with heterogeneous treatment effects.** American Economic Review, v. 110, n. 9, p. 2964-96, 2020.

FELÍCIO, Fabiana; FERNANDES, Reynaldo. **O efeito da qualidade da escola sobre o desempenho escolar: uma avaliação do ensino fundamental no estado de São Paulo.** Anais do XXXIII Encontro Nacional de Economia, 2005.

FERNANDES, Reynaldo. **Índice de desenvolvimento da educação básica (IDEB).** Textos para discussão, n. 26, p. 29-29, 2007.

FERRAZ, Claudio; FINAN, Frederico. **Exposing corrupt politicians: the effects of Brazil's publicly released audits on electoral outcomes.** The Quarterly journal of economics, v. 123, n. 2, p. 703-745, 2008.

FERRAZ, Claudio; FINAN, Frederico. **Electoral accountability and corruption: Evidence from the audits of local governments.** American Economic Review, v. 101, n. 4, p. 1274-1311, 2011.

FERRAZ, Claudio; FINAN, Frederico; MOREIRA, Diana B. **Corrupting learning: Evidence from missing federal education funds in Brazil.** Journal of Public Economics, v. 96, n. 9-10, p. 712-726, 2012.

FIRPO, Sergio; PIERI, Renan; SOUZA, André Portela. **Electoral impacts of uncovering public school quality: Evidence from Brazilian municipalities.** Economia, v. 18, n. 1, p. 1-17, 2017.

GOODMAN-BACON, Andrew. **Difference-in-differences with variation in treatment timing.** National Bureau of Economic Research, 2018.

HANUSHEK, Eric A. **Publicly provided education.** Handbook of public eco-

nomics, v. 4, p. 2045-2141, 2002.

HANUSHEK, Eric A.; RAYMOND, Margaret E. **Does school accountability lead to improved student performance?**. Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management, v. 24, n. 2, p. 297-327, 2005.

HANUSHEK, Eric A.; RAYMOND, Margaret E. **The effect of school accountability systems on the level and distribution of student achievement**. Journal of the European Economic Association, v. 2, n. 2-3, p. 406-415, 2004.

HANUSHEK, Eric A.; RAYMOND, Margaret E. **Improving educational quality: How best to evaluate our schools?**. In: Education in the 21st century: Meeting the challenges of a changing world. 2002. p. 195-247.

HANUSHEK, Eric A.; WOESSMANN, Ludger. **The role of cognitive skills in economic development**. Journal of economic literature, v. 46, n. 3, p. 607-68, 2008.

HÖLMSTROM, Bengt. **Moral hazard and observability**. The Bell journal of economics, p. 74-91, 1979.

INEP. **Nota metodológica sobre a compatibilização de desempenhos do PISA com a escala do SAEB**. 2007.

GROGGER, Jeff; EIDE, Eric. **Changes in college skills and the rise in the college wage premium**. Journal of Human Resources, p. 280-310, 1995.

JOHNSON, William R.; NEAL, Derek A. **The Role of Pre-Market Factors in Black-White Wage Differences**. National Bureau of Economic Research, 1995.

LOCHNER, Lance. **Non-production benefits of education: Crime, health, and good citizenship**. National Bureau of Economic Research, 2011.

MENEZES-FILHO, Naércio Aquino. **Os determinantes do desempenho escolar do Brasil**. IFB, 2007.

MURNANE, Richard J.; WILLETT, John B.; LEVY, Frank. **The growing importance of cognitive skills in wage determination**. National Bureau of Economic

Research, 1995.

MURNANE, Richard J.; WILLET, John. B.; DUHALDEBORDE, Yves; TYLER, John H. **How important are the cognitive skills of teenagers in predicting subsequent earnings?**. Journal of Policy Analysis and Management, v. 19, n. 4, p. 547-568, 2000.

MURNANE, Richard J.; WILLET, John. B.; BRAATZ, M. Jay; DUHALDEBORDE, Yves. **Do different dimensions of male high school students' skills predict labor market success a decade later? Evidence from the NLSY**. Economics of Education Review, v. 20, n. 4, p. 311-320, 2001.

O'NEILL, June. **The role of human capital in earnings differences between black and white men**. Journal of economic Perspectives, v. 4, n. 4, p. 25-45, 1990.

PERACCHI, Franco. **Educational wage premia and the distribution of earnings: An international perspective**. Handbook of the Economics of Education, v. 1, p. 189-254, 2006.

SAKURAI, Sergio Naruhiko. **Ciclos políticos nas funções orçamentárias dos municípios brasileiros: uma análise para o período 1990-2005 via dados em painel**. Estudos Econômicos (São Paulo), v. 39, n. 1, p. 39-58, 2009.

Carneiro, Pedro, Flavio Cunha, and James J. Heckman. **“Interpreting the evidence of family influence on child development.”** The economics of early childhood development: lessons for economic Policy (2003).

Cunha, Flavio, Heckman, James J., Lochner, Lance, Masterov, Dimitriy V. (2006). **“Interpreting the evidence on life cycle skill formation”**. Handbook of the Economics of Education, 1, 697-812.

Curi, Andréa Zaitune, and Naércio Aquino Menezes-Filho. **“A relação entre educação pré-primária, salários, escolaridade e proficiência escolar no Brasil.”** Estudos Econômicos (São Paulo) 39.4 (2009): 811-850.

Currie, Janet. **“Early childhood education programs.”** Journal of Economic perspectives 15.2 (2001): 213-238.



Currie, Janet, and Enrico Moretti. "Mother's education and the intergenerational transmission of human capital: Evidence from college openings." *The Quarterly journal of economics* 118.4 (2003): 1495-1532.

Garces, Eliana, Duncan Thomas, and Janet Currie. "Longer-term effects of Head Start." *American economic review* 92.4 (2002): 999-1012.

Gertler, Paul, and Lia C. Fernald. "**The medium term impact of Oportunidades on child development in rural areas.**" Unpublished manuscript, University of California at Berkeley (2004).

Hanushek, Eric A., and Ludger Woessmann. "**The role of cognitive skills in economic development.**" *Journal of economic literature* 46.3 (2008): 607-68.

Lochner, Lance. "**Non-production benefits of education: Crime, health, and good citizenship**". No. w16722. National Bureau of Economic Research, 2011.

Menezes-Filho, Naércio Aquino. "**Os determinantes do desempenho escolar do Brasil**". IFB, 2007.

Peracchi, Franco. "**Educational wage premia and the distribution of earnings: An international perspective.**" *Handbook of the Economics of Education* 1 (2006): 189-254.

PERACCHI, Franco. **Educational wage premia and the distribution of earnings: An international perspective.** *Handbook of the Economics of Education*, v. 1, p. 189-254, 2006.

MENEZES-FILHO, Naércio Aquino. **Os determinantes do desempenho escolar do Brasil.** IFB, 2007.

LOCHNER, Lance. **Non-production benefits of education: Crime, health, and good citizenship.** National Bureau of Economic Research, 2011.

HANUSHEK, Eric A.; WOESSMANN, Ludger. **The role of cognitive skills in economic development.** *Journal of economic literature*, v. 46, n. 3, p. 607-68, 2008

FERNANDES, Reynaldo. **A universalização da avaliação e a criação do**

**Ideb: pressupostos e perspectivas.** Em Aberto, v. 29, n. 96, 2016.

FERNANDES, Reynaldo; GREMAUD, Amaury Patrick. **Avaliação educacional em larga escala e accountability: uma breve análise da experiência brasileira.** Pesquisa e Debate em Educação, v. 10, n. 1, p. 1103-1137, 2020.

FERNANDES, Reynaldo; GREMAUD, Amaury Patrick. **Qualidade da educação: avaliação, indicadores e metas.** Educação básica no Brasil: construindo o país do futuro. Rio de Janeiro: Elsevier, v. 1, p. 213-238, 2009.

LUCAS, Adrienne M.; MBITI, Isaac M. **Effects of school quality on student achievement: Discontinuity evidence from kenya.** American Economic Journal: Applied Economics, v. 6, n. 3, p. 234-63, 2014.

BAUER, Adriana. **Estudos sobre Sistemas de Avaliação Educacional no Brasil: um retrato em preto e branco.** Revista@ mbienteeducação, v. 5, n. 1, p. 7-31, 2017.

BAUER, Adriana; CASSETTARI, Nathalia; OLIVEIRA, Romualdo Portela de. **Políticas docentes e qualidade da educação: uma revisão da literatura e indicações de política.** Ensaio: Avaliação e Políticas Públicas em Educação, v. 25, p. 943-970, 2017.