

**IMPACTO DA OLIMPÍADA BRASILEIRA DE
ESCOLAS PÚBLICAS (OBMEP) NO
DESEMPENHO EM MATEMÁTICA NA
PROVA BRASIL, ENEM E PISA.**

Camila M. Machado Soares

Elisabette Leo

Supervisão: José Francisco Soares

2014

Resumo

A Olimpíada Brasileira de Matemática para Escolas Públicas (OBMEP) estabeleceu-se nos últimos anos como importante política pública na área de Educação Básica. Diante disso, tornou-se importante avaliar seu impacto no aprendizado de todos alunos das escolas participantes. Este é o objetivo deste trabalho que para isso cria, primeiramente, uma medida de envolvimento de cada escola com a OBMEP. Essa medida é usada em modelo de regressão com a finalidade de medir o impacto do envolvimento das escolas participantes da OBMEP sobre a nota de Matemática dos alunos na Prova Brasil, no ENEM e no PISA. O modelo utilizado para o cálculo do impacto é especificado a partir de uma combinação conceitual do método de Diferenças em Diferenças com o de modelos hierárquicos. De maneira geral, observa-se que os alunos de escolas mais envolvidas com a OBMEP alcançam um desempenho em Matemática significativamente maior do que os alunos de outras escolas com alunado semelhante.

Introdução

A Olimpíada Brasileira de Matemática das Escolas Públicas (OBMEP) é uma iniciativa conjunta do Instituto de Matemática Pura e Aplicada (IMPA) e da Sociedade Brasileira de Matemática (SBM), financiada tanto pelo Ministério da Educação (MEC) e pelo Ministério de Ciência e Tecnologia (MCT). Desde sua primeira edição em 2005, a OBMEP consolidou-se como importante política pública na área de Educação Matemática, atingindo 18,8 milhões de inscritos e 99,89% dos municípios brasileiros, em 2013.

Os objetivosⁱ da Olimpíada foram definidos como: estimular e promover o aprendizado da Matemática nas escolas públicas; colaborar no aperfeiçoamento dos professores de matemática das escolas públicas, contribuindo assim para a sua valorização profissional; contribuir para a melhoria do ensino e da aprendizagem da Matemática nas escolas da rede pública; identificar jovens talentos e fornecer oportunidades para seu ingresso em cursos superiores nas áreas científicas e tecnológicas; promover a integração entre as escolas públicas, as universidades federais, os institutos de pesquisa e as sociedades científicas; contribuir para a inclusão social por meio da difusão de conhecimentos. Para atingir tais objetivos, a Coordenação Geral da OBMEP, e as escolas participantes, trabalham em conjunto não só na realização das provas de primeira e segunda fase, mas também na divulgação da iniciativa e no desenvolvimento de atividades preparatóriasⁱⁱ para as etapas da Olimpíada. As escolas participantes, orientadas pela Coordenação Geral, são responsáveis pela execução da primeira fase da Olimpíada, na qual todos os alunos do sexto ao nono ano do ensino fundamental e todos os alunos de todas as séries do ensino médio são estimulados a participar. Nessa fase não há limite no número de alunos inscritos. Já a segunda fase é executada somente pela Coordenação Geral, e dela participam os alunos classificados, por desempenho, na fase anterior. Os resultados da segunda fase definem os alunos premiados com menções honrosas ou medalhas de ouro, prata e bronze.

Como qualquer política pública, a OBMEP deve ter seus resultados registrados e seus impactos avaliados. Considerando a amplitude dos objetivos da Olimpíada, há muitos pontos cujos impactos poderiam ser analisados. O objetivo deste trabalho, porém, é avaliar um impacto específico: a influência da OBMEP nos resultados de desempenho

obtidos em testes padronizados do conjunto dos estudantes das escolas participantes. Mais especificamente, foi analisado o impacto do envolvimento das escolas com a OBMEP sobre os resultados obtidos em Matemática por alunos do nono ano na Prova Brasil dos anos de 2007, 2009 e 2011, no Exame Nacional do Ensino Médio (ENEM) dos anos de 2010, 2011 e 2012 e do Programme for International Student Assessment (PISA) em 2009.

Este trabalho inova em relação a trabalhos anteriores (SOARES, CANDIAN, 2011; BIONDI et al., 2009) em dois aspectos. Primeiramente, um indicador de envolvimento de cada escola com a OBMEP é construído com base na Teoria de Resposta ao Item. Num segundo momento, o impacto da OBMEP é estimado através de um modelo hierárquico que considera três grupos de escolas, definidos com base na trajetória de envolvimento das escolas com a OBMEP ao longo dos anos.

Além desta introdução, este trabalho é composto de uma seção dedicada à construção deste indicador de envolvimento, e à apresentação dos dados utilizados, uma seção de metodologia, em que o modelo aqui utilizado é explicitado. A quarta seção apresenta os principais resultados encontrados e por fim, a quinta seção conclui este relatório.

Dados

Medida de envolvimento da Escola com a OBMEP

Como houve uma quase universalização na participação das escolas públicas na OBMEP entre os anos de 2005 a 2013, é necessário medir o grau de envolvimento da escola com a Olimpíada, e não apenas registrar a participação. Nesse período o número de alunos inscritos na primeira fase saltou de 10,5 para 18,8 milhões e o percentual de municípios participantes de 93,5% para 99,89%. Diante disso, uma variável indicadora de participação tornou-se pouco informativa do real envolvimento da escola, além de apresentar pouca variação entre as escolas brasileiras já que quase todas estiveram envolvidas com a OBMEP em algum momento deste período.

Para construir uma medida de envolvimento da escola com a OBMEP, considerando os dados disponíveis, foram definidas quatro evidências do envolvimento: (a) número de inscritos e classificados para a segunda fase, (b) menções honrosas recebidas e (c) premiações (medalhas de ouro, prata e bronze) obtidos pela escola e (d) número de professores com envolvimento registrado. Cada uma dessas dimensões foi transformada em um indicador diferente. Posteriormente, estes quatro indicadores foram agregados em uma medida de envolvimento, utilizando-se para isso um modelo da Teoria de Resposta ao Item (TRI).

Para cada uma das escolas participantes na Olimpíada, e para cada um dos anos em análise, de 2005 a 2013, foram disponibilizadas pela coordenação Geral da OBMEP, informações sobre o número de alunos participantes na primeira fase, o número de alunos classificados para a segunda fase, a quantidade de menções honrosas e medalhas (ouro, prata e bronze) recebidas pela escola e por fim, o número de professores colaboradores.

Indicadores de envolvimento da escola

O indicador 1, denominado de “*Indicador de Participação*” foi construído agregando o número de alunos inscritos na primeira fase e o número de alunos classificados para a segunda fase da OBMEP. É importante destacar que a informação sobre o número de alunos participantes da primeira fase é reportada pela escola. A segunda fase, entretanto, ocorre em um ambiente externo à escola e é conduzida pela própria organização da OBMEP. Entretanto, o número de alunos participantes na segunda fase de cada escola obedece a uma regra exógena, definida pela organizaçãoⁱⁱⁱ. No entanto, a presença de alunos na segunda etapa também reflete a mobilização da escola que, por diferentes meios, incentiva e viabiliza a presença de seus alunos escolhidos para participar da segunda etapa. Por conta dessa característica em relação a ambas as informações, optou-se por criar um único indicador de participação da escola na OBMEP juntando as duas informações. Assim, escolas foram categorizadas como apresentando “*baixa participação*” em cada ano, caso não tivessem inscritos alunos na primeira fase ou não tivessem alunos classificados na segunda fase. As escolas que apresentassem alunos inscritos na primeira fase e alunos classificados para a segunda fase foram classificadas como “*alta participação*”.

Tabela 1 - Frequência do Indicador “Participação da Escola na OBMEP” para os anos 2005/13

ANO	Baixa Participação	Alta Participação
2005	50,9	49,1
2006	49,9	50,1
2007	40,0	60,0
2008	39,3	60,7
2009	33,4	66,6
2010	32,5	67,5
2011	32,5	67,5
2012	31,1	68,9
2013	28,2	71,8

Fonte: Elaboração das autoras

O indicador 2 representa o número de menções honrosas recebidas pelas escolas em cada uma de suas participações na OBMEP. Esse indicador foi dividido em quatro categorias: a primeira identifica as escolas que não tiveram nenhum aluno premiado; a segunda, as escolas em que apenas um aluno recebeu a menção honrosa; a terceira

indica que dois alunos receberam a menção honrosa; e finalmente a última categoria que indica que três ou mais alunos foram premiados com menção honrosa na escola. Esse indicador registra o fato de que recebimento de menções honrosas por alunos de uma determinada escola indica algum grau de envolvimento desta escola com a OBMEP, pois é o envolvimento da escola com a OBMEP que permite a alunos com conhecimento de matemática concretizar obter a premiação. Entretanto, a menção honrosa também é, em parte, fruto das características individuais do aluno que a recebeu e, portanto, este indicador indica também a presença de bons alunos na escola. Cabe ressaltar, porém, que o número de alunos premiados é muito pequeno em relação ao total de alunos que participam da Prova Brasil e, assim sendo, esse indicador capta mais o envolvimento da escola do que a seleção de alunos.

Tabela 2 - Frequência do Indicador “Recebimento de Menções Honrosas” - anos 2005/13

ANO	Sem Menção Honrosa	Com Uma Menção Honrosa	Com Duas Menções Honrosas	Com Três ou Mais Menções Honrosas
2005	79,5	9,7	4,5	6,3
2006	80,0	8,9	4,1	7,0
2007	80,1	9,7	4,2	6,0
2008	80,9	9,2	4,0	5,9
2009	79,1	10,6	4,3	5,9
2010	79,3	10,3	4,5	5,9
2011	79,2	10,6	4,4	5,8
2012	73,4	12,6	5,7	8,3
2013	74,2	12,7	5,5	7,6

Fonte: Elaboração das autoras

O indicador 3 é construído com as informações relativas às premiações com medalhas de ouro, prata e bronze. Foram definidas duas categorias: a primeira indica as escolas que não tiveram alunos premiados; a segunda indica as escolas com alunos que obtiveram uma ou mais medalhas, independentemente do tipo de medalha recebido. Como já destacado no parágrafo anterior em relação a menções honrosas, o recebimento de uma medalha indica não só o grau de envolvimento da escola com a Olimpíada, mas também a presença de bons alunos na escola. Mas, como também já exposto anteriormente, o número pouco significativo de alunos premiados faz com que o indicador 3, assim como o 2, seja mais um indicador de envolvimento do que de seleção de alunos.

Tabela 3 - Frequência do Indicador “Recebimento de Medalhas” para os anos de 2005/13

ANO	Sem alunos premiados	Uma ou mais premiações
2005	98,5	1,5
2006	98,6	1,4
2007	97,0	3,0
2008	97,3	2,7
2009	97,3	2,7
2010	97,0	3,0
2011	97,1	2,9
2012	95,6	4,4
2013	93,9	6,1

Fonte: Elaboração das autoras

Por fim, o indicador 4 indica o número de professores colaboradores cadastrados por cada escola^{iv}. Foram definidas três categorias: a primeira agrega as escolas que não possuem nenhum professor registrado junto à coordenação como participante na OBMEP ou escolas para as quais não havia informações sobre a participação de professores; a segunda, as escolas em que há um professor envolvido na OBMEP; e a terceira, as escolas em que há dois ou mais professores participantes. Cabe ressaltar que para os anos de 2005 e 2006, não há informações sobre a participação dos professores: todas as escolas, nestes dois anos, foram alocadas na primeira categoria. Este indicador capta claramente uma dimensão do envolvimento da escola com a OBMEP. Entretanto, o número de professores envolvidos é informado pela própria escola e este indicador possui algumas limitações que devem ser ressaltadas: não há uma clara definição de quem são e quais são as tarefas dos professores colaboradores e, além disso, um número significativo de escolas não registrou nenhum professor colaborador. Neste trabalho, assume-se que uma escola, com mais envolvimento com a OBMEP, registraria seus docentes junto à organização e que quanto maior o envolvimento da escola com a OBMEP, maior seria o número de professores envolvidos e registrados.

Tabela 4 - Frequência do Indicador “Professores Envolvidos na OBMEP” para os anos 2005/13

ANO	Sem professores participantes	Com um professor participante	Com Dois ou mais professores participantes
2005	100,0	0,0	0,0
2006	100,0	0,0	0,0
2007	80,1	12,2	7,7
2008	81,0	11,7	7,3
2009	67,1	19,0	13,9
2010	76,7	14,9	8,4
2011	64,6	9,9	25,5
2012	69,1	9,6	21,3
2013	68,7	8,8	22,5

Fonte: Elaboração das autoras

Para agregar os quatro indicadores em uma única medida de envolvimento da escola com a OBMEP, empregou-se um modelo da Teoria da Resposta ao Item (TRI). A TRI engloba uma série de modelos matemáticos cujo objetivo principal é a obtenção de medidas de construtos latentes baseadas na probabilidade de um indivíduo escolher uma determinada resposta a um item dicotômico e/ou ordinal^v. O modelo da TRI escolhido para a criação do índice de envolvimento da escola com a OBMEP foi desenvolvido por Samejima (1969). Trata-se de um modelo utilizado para estimação do nível de um construto latente através de itens cujas opções de respostas possuam uma ordenação. A expressão matemática é dada por:

$$\Pr = (X_{ij} = k | \theta) = \frac{1}{1 + \exp[-a_i(\theta_j - b_{i,k-1})]} - \frac{1}{1 + \exp[-a_i(\theta_j - b_{i,k})]}$$

X_{ij} é a resposta fornecida pelo j -ésimo indivíduo ao i -ésimo indicador ordinal; $k = 1, 2, \dots, m$ representa cada um dos níveis de X_i , sendo m a categoria associada ao maior valor de θ ; θ_j é a medida do indicador para o j -ésimo indivíduo; a_i é a inclinação e b_i, k é o ponto em que a probabilidade de resposta ao i -ésimo indicador ser maior ou igual a k é maior do que 50%.

Utilizando este modelo, construiu-se uma medida de envolvimento para cada escola em cada um dos anos de existência da Olimpíada, de 2005 a 2013.

Dados de Desempenho – Prova Brasil

Foram considerados, os resultados dos alunos do nono ano do Ensino Fundamental de escolas públicas participantes da Prova Brasil de Matemática de 2007, 2009 e 2011. Juntamente com a nota de Matemática destes alunos em cada um dos anos, foram também utilizadas informações retiradas do questionário socioeconômico, que acompanha a Prova Brasil e é preenchido pelos próprios alunos. Com base em tais informações, foram criadas variáveis indicadoras para o sexo^{vi} e a cor/raça^{vii} de cada aluno, além de uma indicadora de atraso escolar^{viii} e uma variável contínua referente ao nível socioeconômico dos alunos. Esta última variável foi criada a partir da metodologia utilizada por Alves e Soares (2009) e construída a partir da Teoria de Resposta ao Item (TRI) com informações sobre o nível de escolaridade dos pais, posse de bens de consumo duráveis e a presença de trabalhador doméstico na residência. Ainda utilizando informações do questionário socioeconômico da Prova Brasil, foram construídas variáveis que captam características das escolas: a proporção de alunas do sexo feminino, a proporção de alunos atrasados e o nível socioeconômico médio da escola. Também foi utilizada uma variável indicadora da dependência administrativa da escola, que assumia valor igual a 1 caso a escola fosse municipal e 0, caso estadual. Foram incluídos na análise dados de 5.681.424 alunos do nono ano do Ensino Fundamental que pertenciam a 35.000 escolas. Destas, 15643 eram escolas municipais e 19355, estaduais. O nível socioeconômico médio das escolas variava de 2,18 a 7,38, com média de 4,76^{ix}.

Dados de Desempenho - ENEM

Em relação ao ENEM, utilizaram-se neste estudo dados dos anos de 2010, 2011 e 2012 referentes aos resultados dos alunos de escolas públicas. Juntamente com a nota de Matemática destes alunos em cada um dos anos, foram também utilizadas informações retiradas do questionário socioeconômico, preenchido pelos próprios alunos. Com base em tais informações, foram criadas variáveis indicadoras para o sexo e a cor/raça de cada aluno, além de uma indicadora de atraso escolar e uma variável contínua referente ao nível socioeconômico dos alunos. Assim como nos dados da Prova Brasil a variável do nível socioeconômico foi criada a partir da metodologia utilizada por Alves e Soares (2009) e construída a partir da Teoria de Resposta ao Item (TRI) com informações sobre o nível de escolaridade dos pais, posse de bens de consumo duráveis e a presença de trabalhador doméstico na residência.

Ainda utilizando informações do questionário socioeconômico do ENEM, foram construídas variáveis agregadas no nível das escolas: a proporção de alunas do sexo feminino, a proporção de alunos atrasados e o nível socioeconômico médio da escola. Também foram utilizadas duas variáveis indicadoras da dependência administrativa da escola, uma para escolas estaduais e outra para escolas municipais.

Para os três anos analisados, foram coletadas informações de 3374468 alunos, pertencentes a 68604 escolas públicas. Destas, 1% eram escolas federais, 87% escolas estaduais e 12%, municipais.

Dados de Desempenho - PISA

Para a realização desta avaliação de impacto, utilizaram-se os dados de desempenho dos alunos brasileiros que participaram do PISA na edição de 2009. Neste ano, leitura foi a área enfatizada pelo programa. Entretanto, como em todas as edições, o teste incluía questões tanto de matemática como de ciência. Portanto, proficiências referentes às três áreas contempladas no PISA: leitura, matemática e ciências puderam ser calculadas. Neste estudo, analisaram-se especificamente os resultados em matemática de alunos de escolas públicas brasileiras, ou seja, de escolas federais, estaduais e municipais. Tomou-se como resultado em matemática a média simples dos cinco valores plausíveis para a proficiência em matemática divulgados pela Organização de Cooperação e de Desenvolvimento Econômico (OCDE)^x.

Juntamente com a medida de desempenho em Matemática, foram também utilizados indicadores contextuais, calculados pelo PISA com informações do questionário preenchido pelos alunos no momento de resposta aos itens do teste. Com base em tais informações, foram criadas variáveis indicadoras para o sexo^{xi} de cada aluno, atraso escolar^{xii} e uma variável contínua que capta o nível socioeconômico dos alunos. Esta última variável é denominada “Índice de status social, econômico e cultural” (PISA Index of Economic, Social and Cultural Status - ESCS)^{xiii}. O ESCS foi agregado para o nível das escolas, de modo a obter nível socioeconômico e cultural médio de cada escola. Para as duas primeiras variáveis, não há dados ausentes, o que agrega certa robustez às estimações.

Foram utilizadas informações do questionário socioeconômico do ENEM, para construir outras variáveis descritivas das escolas: a proporção de alunas do sexo feminino, a proporção de alunos com defasagem idade-série. Cabe destacar que optou-se por utilizar essas informações obtidas na base do ENEM uma vez que os alunos participantes do PISA deveriam estar cursando o primeiro semestre do segundo ano do Ensino Médio. Por fim, dados do Censo Escolar permitiram construir duas variáveis indicadoras da dependência administrativa da escola: municipal, estadual ou federal^{xiv}.

Para o ano de 2009, foram coletadas informações de 17796 alunos, pertencentes a 845 escolas públicas. Destas, 1,2% eram escolas federais, 85,6% escolas estaduais e 13,3%,

municipais. Assim como nos outros dois relatórios, as escolas estaduais formam representam a maioria da amostra.

Trajetórias de envolvimento

Como a medida de envolvimento com a OBMEP das escolas foi calculada para cada um dos anos entre 2005 e 2012, optou-se por construir trajetórias de envolvimento de cada escola. Para isso dicotomizou-se a medida de envolvimento. As escolas com medidas maiores do que a mediana da medida receberam o valor 1, representando alto envolvimento e as outras valor 0, indicativo de baixo envolvimento.

Após a dicotomização, obteve-se, para cada escola, e para cada ano, um indicador discreto de envolvimento. A partir desta informação, as trajetórias de envolvimento das escolas ao longo dos anos foram analisadas. Entende-se por trajetória de envolvimento o conjunto de indicadores de envolvimento atribuídos para cada escola em todos os anos em análise. Por exemplo, a trajetória de uma determinada escola entre 2005 e 2013 seria composta de um conjunto de nove 0s e 1s. Definiram-se, então, trajetórias boas, regulares e ruins, para cada um dos períodos analisados. A tabela 5 indica como essa classificação foi feita, com base no número de anos de envolvimento das escolas com a OBMEP. Por exemplo, quando o período de análise refere-se a 2005 a 2011, a trajetória de uma escola é classificada como boa caso esta escola apresente 6 ou 7 anos de alto envolvimento com a OBMEP.

Tabela 5 - Classificação das Trajetórias para os três períodos analisados

	2005-2007	2005-2009	2005-2011
Trajetória Boa	3 anos	4 ou 5 anos	6 ou 7 anos
Trajetória Regular	1 ou 2 anos	2 ou 3 anos	3, 4 ou 5 anos
Trajetória Ruim	Nenhum ano	0 ou 1 ano	0, 1 ou 2 anos

Fonte: Elaboração das autoras

Metodologia

Para medir o impacto da OBMEP nos resultados das três avaliações, utilizou-se um modelo hierárquico, que reflete algumas características do método de Diferenças em Diferenças. O método de Diferenças em Diferenças é muito utilizado em avaliações de impacto de projetos para os quais não existe uma seleção aleatória do grupo tratado. É, portanto, um método não experimental de análise de impacto. Este método supõe que os grupos de controle e tratamento irão diferir não só em termos do tratamento recebido, mas também que tais diferenças se apresentarão, principalmente, em termos de variáveis não observadas. O impacto é medido através de uma regressão linear com informações em dois períodos no tempo para os dois grupos (antes e depois do tratamento para tratados e não tratados). Com isso, controla-se por características não observáveis, evitando-se assim que a estimação seja viesada.

A escolha pelo método de Diferenças em Diferenças para analisar o impacto da OBMEP sobre o desempenho dos alunos em Matemática na Prova Brasil encontra dificuldades. Os dados obtidos em avaliações educacionais provêm de alunos que, na rotina pedagógica, pertencem a uma mesma turma ou escola. Nessa situação, há uma relação entre tais alunos pelo simples fato de frequentarem a mesma escola. Logo, a hipótese básica de um modelo de regressão, de independência entre as observações, é violada. Com isso, os erros padrão seriam estimados com um viés para baixo e poderiam ser encontrados resultados espúrios (HOX, 1995). Para lidar com casos como este, faz-se necessário utilizar técnicas como os modelos hierárquicos, também conhecidos como modelos de dois níveis, que permitem considerar a correlação entre as respostas de uma mesma escola.

Outra característica dos dados disponíveis impede a simples utilização deste método é que não há informações sobre desempenho e características de alunos/escolas para um período anterior à OBMEP. O ano de 2005 é o primeiro ano da Olimpíada e também da Prova Brasil. Além disso, pode-se afirmar que as escolas que participaram da OBMEP são intrinsecamente diferentes das que não participaram, como já destacado por Biondi *et al* (2009). Não há, assim, uma maneira direta de definir grupos de controle e tratamento comparáveis. Fica evidente, então, que não há como utilizar diretamente o método de Diferenças em Diferenças.

Diante destas características dos dados disponíveis, optou-se estimar um modelo hierárquico com algumas modificações, inspiradas pela ideia central ao método de Diferenças em Diferenças, comparou-se o desempenho em Matemática dos alunos em três trajetórias de envolvimento de suas respectivas escolas, depois de se controlar por características dos alunos, fora do controle da escola e por características das escolas fora também de seu controle.

Assim, o modelo hierárquico básico utilizado neste estudo pode ser descrito, no primeiro nível, individual, pela seguinte equação:

$$\text{Matemática}_{ij} = \pi_{0j} + \pi_{1j} * \text{FEMININO}_{ij} + \pi_{2j} * \text{AusenteSEXO}_{ij} + \pi_{3j} * \text{BRANCO}_{ij} + \pi_{4j} * \text{PRETO}_{ij} + \\ \pi_{5j} * \text{AMARELO}_{ij} + \pi_{6j} * \text{INDÍGENA}_{ij} + \pi_{7j} * \text{AusenteCOR}_{ij} + \pi_{8j} * \text{NSE}_{ij} + \pi_{9j} * \text{ATRASSO}_{ij} + \\ \pi_{10j} * \text{AusenteATRASSO}_{ij} + e_{ij}$$

em que i indexa os alunos, j a escola e k o ano de aplicação da Prova Brasil. Tanto o intercepto quanto as inclinações variam de acordo com a escola e com o tempo. Num segundo nível, o da escola, teríamos:

$$\pi_j = \beta_0 + \beta_1 * \text{TRAJETORIA_1}_j + \beta_2 * \text{TRAJETORIA_2}_j + \beta_3 * \text{MUNICIPAL}_j + \beta_4 * \text{PFEMININO}_j + \\ \beta_5 * \text{MNSE}_j + \beta_6 * \text{PATRASSO}_j + r_j$$

em que TRAJETORIA_1 e TRAJETORIA_2 são as variáveis indicadoras descritas acima, MUNICIPAL indica a dependência administrativa da escola, PFEMININO , a proporção de alunas, MNSE , o nível socioeconômico médio da escola e PATRASSO , a proporção de alunos atrasados.

Este modelo foi estimado, com pequenas modificações, para os três dados de desempenho utilizados.

Medida do Impacto da OBMEP

Prova Brasil.

Antes de apresentarmos os resultados da estimação do modelo acima explicitado, é interessante analisar como os resultados dos alunos de escolas em cada uma das três trajetórias diferem entre si, mesmo antes de controlar por diferenças observáveis entre as escolas e entre os alunos.

A tabela 6 apresenta o número de observações, a média, o valor mínimo e o máximo, além do desvio padrão, de cada uma das três categorias de trajetórias (boa, regular ou ruim). Esta simples análise exploratória já indica que alunos de escolas com boas trajetórias apresentam desempenho superior: a média alcançada por tais escolas é 26,10 pontos superior à encontrada em escolas com uma trajetória ruim de envolvimento com a OBMEP. O valor de 26,10 pontos equivale a cerca de um ano e meio de escolarização^{xv}.

Tabela 6 - Estatísticas Descritivas das Trajetórias - 2005/2011

	N	Mínimo	Máximo	Média	Desvio-Padrão
Trajecória Boa	585599	105,93	398,27	259,35	47,21
Trajecória Média	827197	105,93	398,27	243,69	45,63
Trajecória Ruim	540690	106,61	398,27	233,25	44,49

Fonte: Elaboração das autoras

Na tabela 7, são apresentados os resultados obtidos com a estimação do modelo explicitado na seção anterior. Cada linha desta tabela apresenta os coeficientes estimados, seguidos dos desvios-padrão estimados entre colchetes. Observa-se que, para todos os períodos analisados, a importância, significativa e positiva, do envolvimento da escola na nota média de Matemática na Prova Brasil. O impacto desse envolvimento, além disso, aumenta com o passar dos anos: em 2011, o indicador da trajetória de envolvimento, que agrega informações desde 2005, atinge seu valor mais alto.

Tabela 7 - Coeficientes das variáveis incluídas na especificação do modelo base. Variável dependente: proficiência em matemática na Prova Brasil nos anos de 2007, 2009 e 2011

Variáveis	2007	2009	2011
Constante	158,62	161,75	162,23
	[1,39]	[1,37]	[1,39]
Feminino	-9,34	-10,74	-7,64
	[0,06]	[0,06]	[0,06]
Ausente Sexo	-16,57	-13,82	-21,41
	[0,27]	[0,31]	[0,27]
Branco	4,37	4,12	4,14
	[0,07]	[0,07]	[0,07]
Preto	-4,89	-5,02	-5,82
	[0,10]	[0,10]	[0,10]
Amarelo	2,64	2,69	2,93
	[0,16]	[0,16]	[0,17]
Indígena	-1,25	-1,97	-1,70
	[0,17]	[0,18]	[0,21]
Ausente Cor	-3,95	-5,40	-7,98
	[0,24]	[0,23]	[0,12]
Atraso	-19,52	-18,83	-16,50
	[0,07]	[0,07]	[0,06]
Ausente Atraso	-12,00	2,18	-18,84
	[0,41]	[0,13]	[0,64]
NSE	1,99	2,07	2,21
	[0,03]	[0,03]	[0,03]
Proporção Alunas	,43	,26	,30
	[0,02]	[0,02]	[0,01]
Proporção Atraso	,06	,09	,01
	[0,01]	[0,01]	[0,01]
MNSE	10,37	10,98	10,71
	[0,18]	[0,17]	[0,18]
Municipal	2,66	2,81	5,28
	[0,19]	[0,19]	[0,20]
Trajetoria_1	9,33	11,53	15,34
	[0,23]	[0,26]	[0,29]
Trajetoria_2	2,87	3,38	6,77
	[0,21]	[0,22]	[0,22]

Fonte: Elaboração das autoras

Uma boa trajetória de envolvimento com a OBMEP impacta a nota dos alunos em Matemática na Prova Brasil em 9,33 em 2007, 11,53 em 2009 e 15,34 em 2011. Estes são valores extremamente significativos: os desvios padrão dessas notas são de 43,30 em 2007, 45,21 em 2009 e 46,88 em 2011. Ou seja, o envolvimento com a OBMEP impacta a nota dos alunos em um quarto ou um terço, dependendo do ano, do desvio padrão; impacto este extremamente significativo para uma política educacional.

Cabe destacar também que uma trajetória boa de envolvimento tem um impacto bem superior a uma trajetória regular: os valores estimados para os coeficientes da variável *Trajectoria_2* são bem inferiores e indicam que a diferença entre escolas com envolvimento regular e aquelas com envolvimento ruim é inferior à diferença entre escolas com envolvimento bom e as com envolvimento ruim. Entretanto, também é importante destacar que o impacto de uma trajetória média, assim como o de uma trajetória boa, aumentou de maneira significativa entre 2007 e 2011 (166% para a trajetória boa e 235% para a trajetória média). Entre 2005 e 2007, as escolas com trajetória regular apresentavam 1 ou 2 anos de envolvimento e este envolvimento impactou em 2,87 pontos o desempenho em Matemática na Prova Brasil de 2007. Em 2011, escolas com uma trajetória regular de envolvimento somavam entre 3 e 5 anos de envolvimento e este envolvimento impactou em 6,77 pontos as notas em Matemática na Prova Brasil de 2011. Escolas com boas trajetórias, que somavam 3 anos de envolvimento em 2007 e 6 ou 7 anos em 2011, apresentaram impactos de 9,33 e 15,34 pontos, respectivamente. Fica evidente, que, mesmo as escolas que não atingiram trajetórias boas de envolvimento nos sete anos analisados, beneficiaram-se de forma positiva e significativa do aumento de anos em que estiveram envolvidas com a OBMEP. Há, assim, um efeito longitudinal de envolvimento com a Olimpíada: quanto maior o número de anos, maior o impacto sobre o desempenho em Matemática.

As outras variáveis analisadas no modelo, tanto àquelas referente a características individuais quanto as referentes à escola, apresentam os sinais esperados e coeficientes com valores significativos e similares nas três estimações. Alunas obtêm notas inferiores que alunos, assim como estudantes que se autodeclararam de cor preta em relação aos pardos, brancos, e indígenas. Alunos atrasados em relação à sua idade têm resultados bem inferiores aos colegas regulares. O nível socioeconômico

individual é pouco relevante, mas o nível socioeconômico médio da escola é muito relevante para explicar a nota dos alunos. Escolas municipais apresentam notas superiores a escolas estaduais e esta diferença aumenta significativamente no ano de 2011. Por fim, as proporções de alunas e de alunos atrasados apresentam baixa capacidade explicativa neste modelo.

ENEM

A tabela 8 apresenta os resultados obtidos com a estimação do modelo explicitado na seção de Metodologia do relatório anterior. Cada linha desta tabela apresenta os coeficientes estimados, seguidos dos desvios-padrão estimados entre colchetes. Todos os coeficientes estimados foram significativos a um nível de confiança de 99%.

Tabela 8 - Coeficientes das variáveis incluídas na especificação do modelo base. Variável dependente: proficiência em matemática no ENEM dos anos de 2010, 2011 e 2012.

Variáveis	2010	2011	2012
Constante	384,85 [2,66]	348,68 [2,79]	311,75 [2,99]
Sexo	-45,30 [0,19]	-45,14 [0,18]	-39,86 [0,19]
Branca	10,16 [0,22]	9,07 [0,21]	9,65 [0,22]
Preta	-5,81 [0,33]	-5,18 [0,30]	-5,95 [0,31]
Amarela	10,93 [0,68]	9,33 [0,62]	10,89 [0,65]
Indígena	-6,18 [1,04]	-7,08 [1,05]	-7,61 [1,11]
AusenteCor	5,14 [0,52]	2,53 [0,57]	3,93 [0,72]
Atraso	-41,92 [0,23]	-36,75 [0,23]	-43,61 [0,24]
NSE	- -	12,47 [0,13]	13,84 [0,14]
Proporção	0,08	0,10	0,10

Alunas	[0,02]	[0,02]	[0,02]
Proporção	-0,12	0,08	0,11
Atraso	[0,01]	[0,01]	[0,01]
MNSE	24,45	28,20	31,27
	[0,37]	[0,50]	[0,52]
Federal	110,95	101,78	108,48
	[2,29]	[2,23]	[2,34]
Estadual	-4,49	-9,73	-9,85
	[1,32]	[1,34]	[1,50]
Trajetoria1	16,80	16,94	15,01
	[0,72]	[0,62]	[0,67]
Trajetoria2	2,59	3,50	2,50
	[0,65]	[0,54]	[0,63]

Fonte: Elaboração das autoras

As estimativas do pertencimento a uma boa trajetória de envolvimento com a OBMEP na nota dos alunos em Matemática no ENEM em 16,80 em 2010, 16,94 em 2011 e 15,01 em 2012. Para compreender melhor o tamanho deste impacto, é interessante notar que os desvios padrão das notas de Matemática são de 97,96 em 2010, 101,21 em 2011 e 106,28 em 2012. Ou seja, o envolvimento com a OBMEP impacta a nota dos alunos em um décimo e um quinto, dependendo do ano, do desvio padrão; impacto este bem significativo para uma política educacional. Entretanto, este valor é bem inferior ao encontrado na análise dos dados da Prova Brasil, indicando, talvez, que o envolvimento da escola com a OBMEP tenha resultados mais efetivos no Ensino Fundamental, que no Ensino Médio.

Cabe destacar também que uma trajetória boa de envolvimento tem um impacto bem superior a uma trajetória regular: os valores estimados para os coeficientes da variável *Trajetória_2* são bem inferiores e indicam que a diferença entre escolas com envolvimento regular e aquelas com envolvimento ruim é bem inferior à diferença entre escolas com envolvimento bom e as com envolvimento ruim. Nos três anos em análise, o impacto da trajetória média foi apenas entre 15%-20% do impacto da trajetória boa. Fica evidente, assim, a importância de um envolvimento contínuo, e não apenas pontual, com a OBMEP.

Também é importante notar que não há grandes variações no valor dos impactos entre os diferentes anos. Não há evidências de um efeito longitudinal de envolvimento com a Olimpíada, como ocorreu na análise dos dados da Prova Brasil.

As outras variáveis analisadas no modelo, tanto àquelas referente a características individuais quanto as referentes à escola, apresentam os sinais esperados e coeficientes com valores significativos e similares nas três estimações. Alunas obtêm notas inferiores que alunos, assim como estudantes que se autodeclararam de cor preta em relação aos pardos, brancos, e indígenas. Alunos atrasados em relação à sua idade têm resultados bem inferiores aos colegas regulares. O nível socioeconômico individual^{xvi} e o nível socioeconômico médio da escola são ambos relevantes para explicar a nota dos alunos. Os coeficientes estimados para a variável NSE Aluno atingiram valores similares aos estimados para as variáveis de trajetória, e os coeficientes para a variável MNSE foram, em todos os anos, bem superiores aos estimados para esta outra variável. Assim, o efeito da família, e das condições de vida dos alunos, são extremamente significativas para os resultados alcançados no ENEM. Este resultado, novamente, difere substancialmente do encontrado quando analisou-se os dados da Prova Brasil.

Por fim, cabe destacar o valor dos coeficientes para escola Federal: nos três anos analisados, este valor foi entorno de 110 pontos, extremamente significativo se compararmos com o desvio padrão já explicitado acima para cada um dos anos. Estudar em uma escola federal impacta em um desvio padrão a nota dos alunos em Matemática no ENEM, depois de consideradas as variáveis socioeconômicas dos alunos e das escolas. Este resultado pode ser explicado pela tradição de excelência, mas também das estratégias de seleção usadas pelas escolas federais, CEFETs e os colégios militares.

PISA

A tabela 9 apresenta o número de observações, a média, o valor mínimo e o máximo, além do desvio padrão, de cada uma das três categorias de trajetórias (boa, regular ou ruim). Esta simples análise exploratória indica como as escolas com boas trajetórias apresentam desempenho superior: a média alcançada por tais escolas é 34,11 pontos,

valor superior ao encontrado em escolas com uma trajetória ruim de envolvimento com a OBMEP. No entanto as escolas com trajetória ruim ou regular não diferem entre si.

Tabela 9 - Estatísticas Descritivas das Trajetórias Boa, Regular e Ruim - 2005-2009

	Número de Alunos	Mínimo	Máximo	Média	Desvio-Padrão
Trajetória Ruim	2390	141,15	613,97	348,67	61,88
Trajetória Média	7366	70,73	598,54	355,08	60,08
Trajetória Boa	7997	166,86	682,36	382,78	65,20

Fonte: Elaboração das autoras

Na tabela 10, são apresentados os resultados obtidos com a estimação do modelo explicitado na seção de Metodologia. Cada linha desta tabela apresenta os coeficientes estimados, seguidos dos desvios padrão estimados entre colchetes. Todos os coeficientes estimados foram significativos a um nível de confiança de 99% exceto os coeficientes estimados para as variáveis *Proporção de Alunas*, *Trajetória_2* e *Estadual*.

Tabela 10 - Coeficientes das variáveis incluídas na especificação do modelo base. Variável dependente: proficiência em matemática no PISA do ano de 2009

Variáveis	2009
Constante	365,99 [11,14]
Sexo	21,49 [0,07]
Atraso	29,33 [0,05]
ESCS	3,37 [0,038]
Proporção Alunas	0,13 [0,11]
Proporção Atraso	-0,30 [0,06]
MESCS	11,92 [2,333]
Federal	94,35

	[13,11]
Estadual	-14,98
	[9,23]
Trajectoria1	9,96
	[3,65]
Trajectoria2	-1,75
	[3,57]

Fonte: Elaboração das autoras

Uma boa trajetória de envolvimento^{xvii} com a OBMEP impacta a nota dos alunos em Matemática no PISA em 9,96 em 2009. Para compreender melhor o tamanho deste impacto, é interessante notar que o desvio padrão das notas de Matemática é de 67,45 em 2009. Ou seja, o envolvimento com a OBMEP impacta a nota dos alunos em um quinze avo do desvio padrão; impacto este bem significativo para uma política educacional. Entretanto, este valor é bem inferior ao encontrado na análise dos dados da Prova Brasil, indicando, possivelmente, que o envolvimento da escola com a OBMEP tenha resultados mais efetivos no Ensino Fundamental, e não tanto nos resultados do Ensino Médio. Esse resultado corrobora os encontrados a partir da análise dos dados do ENEM; os impactos encontrados na análise de ambos os dados são da mesma ordem de grandeza.

Cabe destacar também que uma trajetória boa de envolvimento tem um impacto bem superior a uma trajetória regular: os valores estimados para os coeficientes da variável *Trajectoria2* não são significativos, indicando não há diferença significativa entre escolas com envolvimento regular e aquelas com envolvimento ruim. Assim, comprova-se novamente a importância de um envolvimento contínuo, e não apenas pontual, com a OBMEP.

As outras variáveis analisadas no modelo, tanto àquelas referente a características individuais quanto as referentes à escola, apresentam os sinais esperados e coeficientes com valores significativos. Alunas atingem resultados piores que alunos em Matemática no PISA: as notas das meninas são inferiores em um terço do desvio padrão em relação a nota dos meninos. Já alunos atrasados em relação à sua idade têm resultados significativamente inferiores aos resultados de seus colegas regulares: o atraso afeta, negativamente, em cerca de 43,4% do desvio padrão, o resultado em

Matemática no ENEM, nos três anos analisados. O nível socioeconômico e cultural individual e o nível socioeconômico e cultural médio da escola são ambos relevantes para explicar a nota dos alunos. Os coeficientes para a variável MESCS foram, como esperado, bem superiores aos estimados para o nível socioeconômico e cultural individual.

Assim como no caso do ENEM, cabe destacar o valor do coeficiente para escola Federal: este valor foi de cerca de 94,35 pontos, extremamente significativo se compararmos com o desvio padrão já explicitado acima.

Conclusão

Os resultados encontrados mostram um impacto muito significativo do envolvimento de uma escola com a OBMEP sobre a nota de seus alunos em Matemática. Este impacto é tão maior quanto maior for o tempo de envolvimento da escola com a Olimpíada, indicando a importância do envolvimento contínuo da escola com esta iniciativa.

Este trabalho não identifica o mecanismo pedagógico que explique o impacto observado. Isso exige a realização de pesquisa com outras metodologias. No entanto, é razoável considerar que o efeito se dá porque as escolas envolvidas com a OBMEP por um período maior de tempo modificaram seu projeto pedagógico para o ensino de Matemática incluindo na sua rotina atividades que oportunizam o aprendizado dessa disciplina. Seus professores, além disso, também se envolvem mais com as atividades da OBMEP e com isso se qualificam para melhor ensino. Por fim, toda a divulgação da OBMEP e dos alunos premiados pode despertar o interesse dos próprios alunos não só pela Olimpíada, mas também pela Matemática. Tudo isso contribui de maneira direta para melhorar o ambiente de ensino de Matemática e como consequência, positivamente, o desempenho dos alunos nos testes considerados.

Apesar dos resultados extremamente positivos, cabe uma ressalva: os dados disponíveis para quantificar o envolvimento das escolas com a OBMEP são limitados. Seria interessante, para pesquisas posteriores, que fossem coletados dados que melhor caracterizassem o grau de envolvimento da escola com a OBMEP. Em particular o registro dos professores envolvidos, por exemplo, poderia ser mais bem sistematizado e outras informações, relativas às diversas interações entre a escola e a OBMEP, também deveriam ser coletadas. Tudo isso contribuiria para uma melhor qualificação de um alto envolvimento das escolas com a Olimpíada e permitiria uma estimativa ainda mais precisa do efeito desta iniciativa no desempenho dos alunos em Matemática.

No entanto, os resultados observados são importantes, principalmente pelo efeito observado ser maior do que os de outros fatores considerados em pesquisas similares. Na realidade a literatura da área é dominada pela medida do efeito de variáveis fora do controle da escola, que não podem ser modificados por políticas e práticas escolas.

Assim este tipo de pesquisa mostra sempre um quadro que não ajuda a formulação de políticas públicas. Os resultados deste trabalho identificam, entretanto, para uma política que pode ser implementada nas escolas brasileiras, como são e, portanto, mostram um caminho que precisa ser trilhado para a melhoria do aprendizado dos alunos.

Referências Bibliográficas

- ALVES, M. T. G.; SOARES, J. F. Medidas de nível socioeconômico em pesquisas sociais: uma aplicação aos dados de uma pesquisa educacional. *Opinião Pública*, v. 15, n. 1, p.1-30. Jun. 2009.
- BIONDI, R. L.; VASCONCELLOS, L.; MENEZES-FILHO, N. A. Avaliando o impacto da Olimpíada Brasileira de Matemática das Escolas Públicas (OBMEP) no desempenho de matemática nas avaliações educacionais. São Paulo: Fundação Getúlio Vargas, Escola de Economia de São Paulo. 2009.
- HOX, J.J. Applied multilevel analysis. Amsterdam: T.T. Publikaties, 1995. 126 p.
- SAMEJIMA, F. Estimation of latent ability using a response pattern of graded scores. *Psychometrika*, New York, v. 17, 1969. Monograph Supplement , n. 4, Part 2.
- SOARES, J. F. Índice de Desenvolvimento da Educação de São Paulo – IDESP: bases metodológicas. *São Paulo Perspec*, v. 23, n. 1, p. 29-41. Jan./Jun. 2009.
- SOARES, J.F.; CANDIAN, J. F. O Impacto da Olimpíada Brasileira de Matemática das Escolas Públicas (OBMEP) no desempenho dos alunos na Prova Brasil. Brasília: Centro de Gestão e Estudos Estratégicos, 2011.

Notas

A avaliação aqui empreendida tornou-se possível apenas após um contato com o Instituto Nacional de Ensino e Pesquisa Anísio Teixeira (INEP), responsável pelo exame no Brasil, e detentor de uma base que permite associar o código PISA para a escola e código dessa escola no Censo Escolar. Esta informação, crucial para o estudo aqui realizado, permite relacionar os dados do PISA com todas as informações sobre a trajetória de envolvimento das escolas com a OBMEP, apresentadas no primeiro relatório dessa consultoria.

O acesso a tais dados foi complicado: até 19 de Setembro de 2014, o INEP não possuía uma norma que regesse a cessão de dados individualizados, mas não identificáveis. A Portaria 467 definiu o procedimento para a cessão desses dados e com isso, permitiu a finalização do presente estudo. Cabe destacar que, neste processo, teve-se acesso aos

códigos das escolas participantes, mas não houve nenhum tipo de identificação dos alunos que participaram da prova.

ⁱ Estes objetivos foram adaptados do site oficial da OBMEP: <http://www.obmep.org.br>

ⁱⁱ A Coordenação Geral da OBMEP envia às escolas participantes materiais didáticos para o desenvolvimento destas atividades.

ⁱⁱⁱ Em escolas com mais de 241 alunos inscritos, foram selecionados 5% dos alunos com as maiores notas para a segunda fase da OBMEP.

^{iv} As escolas são as responsáveis por cadastrar, no site da OBMEP, o número de professores envolvidos. Em geral, tais professores são os professores de sala de aula dos alunos participantes. Em alguns casos, podem ser os professores responsáveis por atividades extracurriculares de preparo para a Olimpíada.

^v A referência clássica sobre a Teoria da Resposta ao Item é o trabalho de Hambleton (1973).

^{vi} Esta variável assume valor igual a um quando o aluno é do sexo feminino e igual a zero quando o aluno é do sexo masculino ou quando tal informação está ausente. Também utiliza-se uma variável denominada AusenteSexo, que assume o valor 1 quando a informação sobre o sexo do indivíduo está ausente e 0 quando o indivíduo é do sexo masculino. Isto é feito, neste caso e em outros ao longo do relatório, para viabilizar a inclusão na regressão dos estudantes que não forneceram a informação completa.

^{vii} Foram criadas quatro variáveis indicadoras, para descrever as cores/raças Branco, Preto, Amarelo e Indígena, que assumem valor igual a um quando o aluno se identifica como pertencente a tal cor/raça e 0 quando a raça declarada é parda, tomada como referência, por ser a mais frequente. Também criou-se uma quinta indicadora, AusenteCor, que indica se a informação sobre cor/raça está ausente ou não.

^{viii} Esta variável assume valor igual a um quando o aluno está atrasado em termos de idade/série. Nas estimações aqui apresentadas, também foi utilizada uma variável AusenteAtraso para indicar se a informação relativa a atraso escolar estava ou não ausente.

^{ix} Cabe destacar que, em Alves e Soares (2009), o NSE com valor 10 é alcançado apenas por escolas privadas com todos os indicadores (que compõe o nível socioeconômico) com valor máximo.

^x Para compreender como são calculados os resultados do PISA, e a origem dos ditos valores plausíveis, ver OCDE (2009).

^{xi} Esta variável assume valor igual a um quando o aluno é do sexo feminino e igual a dois quando o aluno é do sexo masculino.

^{xii} Esta variável é calculada pela própria OCDE e representa a diferença entre o ano frequentado pelo estudante e o ano em que este estudante deveria estar matriculado na data de aplicação do PISA. No Brasil, este ano é o Segundo do Ensino Médio.

^{xiii} A OCDE calcula essa variável a partir de um conjunto de variáveis incluindo o nível de educação dos pais, um índice de riqueza da família, um índice de posse de bens culturais, um índice de presença, no ambiente domiciliar, de insumos educacionais e por fim, um índice de status ocupacional dos pais. Para mais informações, ver OCDE (2009).

^{xiv} Juntas estas duas variáveis indicam quais escolas são estaduais (Estadual = 1 e Federal = 0), quais são federais (Estadual = 0 e Federal = 1) e quais são municipais (Estadual = 0 e Federal = 0).

^{xv} Cabe destacar como é calculado o valor de um ano de escolarização. A fórmula utilizada é: (Desempenho Adequado na 8ª série – Desempenho Adequado na 4ª série) /4. Considerando o nível adequado de desempenho, definido por Soares (2009), este valor seria de 18,75 pontos.

^{xvi} Para o ano de 2010, não haviam informações que possibilitassem o cálculo do NSE individual

^{xvii} Representada pela variável Trajetória1.