Evaluating the impact of participation in the Brazilian Public School Mathematical Olympiad on math scores in students' standardized tests

Roberta Loboda Biondi♣
Itaú-Unibanco

Lígia Vasconcellos♦
Itaú-Unibanco

Naercio Aquino de Menezes-Filho♠
Insper and FEA/USP

Abstract

The Brazilian Public School Mathematical Olympiad (*Olimpíada Brasileira de Matemática das Escolas Públicas,* or OBMEP) has been held yearly since 2005, promoted by the federal government together with the Institute of Pure and Applied Mathematics (IMPA), with the objectives of motivating students to study mathematics and improving basic education. In this paper we evaluate the OBMEP's impact on school performance in mathematics, as reflected in the standardized test *Prova Brasil,* and analyze the program's cost-benefit relation. We use an estimator that is doubly-robust as the method combines regression and propensity score weighting. The results show the OBMEP has a positive and statistically significant impact on school performance in mathematics among 9th-graders. This impact increases for schools that participate in repeated yearly OBMEPs, and is larger in the higher test score percentiles. Robustness checks confirmed these results. The cost-benefit analysis shows that the investment in OBMEP brings benefits in terms of the participants' future incomes that outweigh its costs.

Keywords: impact evaluation, doubly-robust estimators, school achievement, economic return.

JEL Classification: I2 Education

1 Introduction

The current educational debate in Brazil is centered on the need to raise the quality of the education offered and to universalize school attendance. In the latter respect, while 97% of children between the ages of 6 and 14 were attending school in 2007, showing universal attendance has virtually been achieved among this age range, the same cannot be said for other age groups. That year the figure for 15 to 17-year-olds was only 82.1%[1].

With respect to educational quality, the most recent national evaluations show some positive variations in proficiency tests, but the performance of Brazilian students on international evaluations such as that of the Programme for International Student Assessment (PISA) indicates the quality is still very low. Brazil occupied the last positions since the first assessment in 2000.

Various policies by the federal, state and municipal governments, along with initiatives from civil society and business, have aimed at addressing these two questions, but particular emphasis has been focused

♣ roberta.biondi@itau-unibanco.com.br
♦ ligia.vasconcellos@itau-unibanco.com.br
♠ NaercioAMF@insper.org.br
We thank the contributions of the members of the IMPA board of directors for their valuable comments, particularly Marcelo Viana and Eduardo Moacir Krieger.
[1] The attendance rates were calculated from micro-data obtained from the 2007 National Household Survey (*Pesquisa Nacional por Amostra de Domicílio* – PNAD) conducted by the Brazilian Institute of Geography and Statistics (IBGE).

on improving the quality of the education offered by public schools. This article contributes to this effort by analyzing the cost-benefit relation of the Public School Mathematical Olympiad (OBMEP), a program aimed at encouraging better mathematics learning.

The Public School Mathematical Olympiad program has been promoted yearly since 2005 by the Ministry of Education and the Ministry of Science and Technology in partnership with the Institute of Pure and Applied Mathematics (IMPA) and the Brazilian Mathematical Society (SBM), the last two responsible for its academic direction. In 2009, the OBMEP was held for the fifth time and nearly 19 million students participated. The number of schools and students participating has been steadily increasing (Table 1). The number of students participating in this program is substantial. Indeed, the program is considered one of the major competitions held among public school students in the country. The IMPA has an extensive team, organized into regional groups to operationalize the program, and rural as well as urban schools are included.

Table 1 – Schools and Students Participating in the OBMEP – 1st Phase

	2005	2006	2007	2008	2009
Number of schools	31,030	32,655	38,450	40,377	43,854
Number of students	10,520,831	14,181,705	17,341,732	18,326,029	19,198,710
% of municipalities[2]	93.50%	94.50%	98.10%	98.70%	99.1%

Source: Data from the OBMEP.

The stated objectives of the OBMEP are to stimulate and promote the study of mathematics among public school students; to contribute to the quality of basic education; to identify talented young people and encourage them to pursue careers in science and technology; to encourage the professional improvement of public school teachers; to contribute to the integration of public schools and public universities, research institutes and scientific societies; and to promote social inclusion by spreading knowledge.

Because of its objectives to improve the quality of public math education and its scope in terms of the number of participants, it is reasonable to assume that the OBMEP can have a positive influence on the average result of the large-scale public school assessments carried out by the federal government to measure educational quality, such as the *Prova Brasil*.

All schools that sign up for the Mathematical Olympiad receive a booklet of sample questions and answers called the "*Banco de Questões*", prepared by professors of the IMPA and members of the SBM. This booklet is sent to the teachers at each school whose students will participate. Its use is optional. It is intended not only to prepare students specifically for the Olympiad, but also to have a positive influence on the overall teaching of math at the participating schools and thus the performance of their students.

This paper presents the results of the impact evaluation of the OBMEP. Our aim is to quantify its effects on the education received by Brazilian student by assessing its impact on the average math scores of the schools on a standardized assessment test (the *Prova Brasil*). We also calculate the economic return of the program by comparing the current costs against the future benefits generated for the students.

In light of the objectives of the Mathematical Olympiad, the questions posed here are: Is it possible to identify in large-scale government evaluations the incentive provided by the OBMEP to the study of math in public schools? And more specifically: Does participation in the OBMEP bring a measurable improvement in average math performance at the participating schools?

We used the strategy of combining linear regression and propensity score weighting as to have a doubly robust estimator, both in level and difference in differences. The resulting estimator pointed to a

[2] The local political unit in Brazil is the municipality, which is similar to a county in the United States.

positive and statistically significant impact on the average math scores of 9th-graders on the 2007 *Prova Brasil*. Since the majority of schools participated in more than one edition of the OBMEP, we also investigated whether the impact is heterogeneous by number of participations. The results show that the impact is greater for schools participating more than once. We also carried out estimates for different percentiles of the distribution of student scores and found a significant effect, not only for better performing students, but also for those with lower scores, though the impact is relatively greater for higher achievers.

In the next section we present the data sources and the sample considered in the analysis. Then we describe the strategies to identify the impact. Finally, we present the results and analyze the economic return, considering some hypotheses on the future of students participating in the program.

2 Data and Sample

The potential participants in the OBMEP are students in the sixth to twelfth school grades[3]. The students signing up are divided into the following levels: I- students enrolled in the six and seven grades; II – students enrolled in the eight or nine grades, including adult education students; and III – high school students.

The Mathematical Olympiad occurs in two phases. In the first, all public schools can sign up voluntarily to participate with their students in any or all of the three levels, depending on the grades offered at the particular school. At this time each school receives the booklet of sample questions and answers. The tests in this phase are given and corrected at each participating school, which can send the top 5% of its students to take part in the second phase of the Olympiad. Students receiving a score of zero in the first phase are not classified for the second phase. The table below shows the number of schools signed up for the OBMEP by year and level in the first phase.

Table 2 – Number of schools signing up for the OBMEP by level and year (1st phase)

Year	Level I		Level II		Level III	
	Students	Schools	Students	Schools	Students	Schools
2005	3,655,677	27,508	3,077,481	27,383	3,787,673	13,255
2006	4,851,150	29,766	4,026,207	29,132	5,304,348	14,277
2007	5,963,883	35,260	4,917,276	34,360	6,460,573	16,321
2008	6,270,982	37,031	5,246,995	36,349	6,808,052	17,133

Source: Prepared by authors with data from the OBMEP

Most schools that participate in the OBMEP do so repeatedly. For level II, which covers students in the 8th and 9th grades during the four years from 2005 to through 2008, 43,573 signed up[4], of which 43% participated in all four editions and 66% participated in at least three.

Moreover, considering the OBMEP for 2007 at level II, the proportion of students participating in relation to students enrolled in those schools was very high. The median of this proportion among all the participating schools was 94%, and in the first quartile of the distribution this proportion was 75%. This shows that schools tend to sign up a large part of their students in the first phase.

[3] Basic education in Brazil changed its duration from 11 to 12 years. Children shall now enter school sooner, with 6 years of age.

[4] As reference, according to the 2006 School Census, there were about 170 thousand public schools in the country, 45% of which were located in urban areas.

To analyze the impact of the OBMEP we used as an indicator the score on the 2007 *Prova Brasil*, a standardized assessment test given every two years (since 2005) by National Educational Research Institute (Inep), part of the Ministry of Education (MEC), to all urban public schools in the country. This evaluation is censitary, ensuring the scores are representative for the schools. The test is given to students in the 5th and 9th grades of all urban schools with at least 20 students enrolled in each grade. The *Prova Brasil* uses the same methodology as the Basic Education Assessment Test (SAEB), based on the item response theory (IRT), which permits comparison of the scores in Portuguese and mathematics of students in different grades at the same time.

We obtained the infrastructure characteristics and educational indicators of the schools used as controls from the 2006 School Census and the municipal population and per capita income figures from the 2000 Demographic Census (IBGE-2000).

With the use of these data, the impact evaluation was restricted to urban schools with students enrolled in level II of the OBMEP and that administered the 2007 *Prova Brasil* to 9th-graders. Since it was impossible to identify the students participating in the program and their scores on the *Prova Brasil*, we used the average score received by each school. In other words, our unit of observation is the school, not the student.

The schools sign up for the OBMEP at the start to the school year and the tests in the first phase are given at the start of the second semester (middle of August). Then the second phase occurs in October or November. The 2007 *Prova Brasil*, in turn, was given to students in November that year, so its result could be influenced by the Mathematical Olympiad.

2.1 The sample

Among the 168,436 active public schools in Brazil (2006 School Census), 68,961 were potential participants in the OBMEP at level II, that is, they had students enrolled in the 8th or 9th grades, or offered adult education programs of equivalent grade level. However, only 34,360 of these signed up for the Olympiad, and not all of these participated in the *Prova Brasil*. Table 3 presents the sample of schools in the treatment group (participants in the OBMEP) .

Of the 34,360 participating schools at level II of the Olympiad in 2007, 34,222 had information available in the 2006 School Census and 22,996 participated in the *Prova Brasil*.

Table 3 – Sample of treated schools

	No. of schools
Signed up for the 2007 OBMEP (level II)	34,360
Filters	
found in the 2006 Census	34,222
urban	25,841
participants in the 2007 *Prova Brasil*	22,996
% of students signed up in relation to school enrollment > 10%	**22,703**

For our evaluation we defined a minimum threshold of 10% of students regularly enrolled in the 8th and 9th grades signed up to take the first phase of the Olympiad. This cut-off criterion eliminated only 293 schools[5]. Our analysis then was based on a sample of 22,703 schools in the treatment group. The table below shows the sample of schools in the control group.

Table 4 – Sample of control schools

	No. of schools
Potential OBMEP participants (level II)	68,961
Filters	
non-participants in the 2007 OBMEP (level II)	35,263
urban	14,571
participants in the 2007 *Prova Brasil*	4,052
never participated in the OBMEP	**1,756**

Among the over 68 thousand schools that could potentially have participated in the 2007 OBMEP, 35,263 did not do so, making them candidates for the control group. However, among them only 4,052 took part in the 2007 *Prova Brasil*. This failure to participate in the *Prova Brasil* is due to the fact that 59% of the schools not participating in the 2007 OBMEP were rural (recalling that the test is not given in rural schools), and of the remainder, most did not have more than 20 students in the 9th grade.

Table A in the Appendix shows the comparative statistics between the schools participating and not participating in the OBMEP of this sample. It can be seen that the treated schools are larger, with more students and teachers, and also have relatively better average student characteristics, such as higher percentages of students with at least one parent who completed college, that went to preschool, that do not work and that were never held back.

In short, for the majority of schools considered, those taking part in the Mathematical Olympiad had statistically significant differences in comparison with those that did not take part, showing a more advantageous profile. This is reflected in the average scores on the *Prova Brasil*: in both years analyzed and both subjects (Portuguese and math), the scores of the schools that signed up for the OBMEP were higher.

Directly comparing the average score of a group of treated and untreated schools without first ensuring that the two groups are similar with respect to other characteristics can skew the results and give the wrong sign of the effect of the OBMEP on students' performance. Below we explain the strategy to correct this problem, by weighting the schools of the control group according to their similarities in the observable characteristics in relation to the treatment schools.

3 Impact evaluation methodology

To infer the quantitative effect of the Olympiad on the average math scores of the schools that signed up for the program, we need to know what these schools' scores would have been had they not taken part in the OBMEP.

This question brings the problem of the unobserved counterfactual, because obviously we cannot observe the math scores of the students participating in the OBMEP if they had not participated. To address

[5] As mentioned before, the participating schools tend to sign up all their students eligible to take part in the OBMEP for its first phase. For this reason, in removing from the sample schools signing up fewer than 10% of their students for the first phase, few schools were eliminated from the sample.

this question we need a control group (non-participating schools) to replace the counterfactual that is similar to the treated schools (participants in the OBMEP), to avoid the problem of selection bias.

Formally stated, we can define Y_0 as the potential result of a particular school if it did not sign up for the Olympiad; Y_1 as the potential result of that school if it did sign up; T = 1 when the school signed up and T = 0 when it did not. We can observe $Y_1|T=1$ and $Y_0|T=0$, but never observe $Y_1|T=0$ and $Y_0|T=1$. We want to know the difference between the score obtained by schools that signed up for the program and the score they would have received had they not signed up. We can write this as $D = E[Y_1 \mid T = 1] - E[Y_0 \mid T = 1]$.

What we are really observing is $G = E[Y_1 \mid T = 1] - E[Y_0 \mid T = 0]$, so the difference between these terms gives us the selection bias $B = G - D = [Y_0 \mid T = 1] - E[Y_0 \mid T = 0]$. This bias arises if the control group is inadequate, such as when the schools that did not sign up for the program were very different from those that did. Since signing up for the OBMEP is voluntary, inclusion in the treated or control group is not random, case in which there would be a potential problem of selection bias.

In response to this, we use a control group with similar characteristics to the treated group, working with the hypothesis of selection on observables. This hypothesis appears reasonable since we have various databases providing a large set of descriptive data on school management, infrastructure, faculty and student body, among other characteristics of the public schools participating in the Olympiad.

Through adequate econometric methods, this rich set of characteristics can be used to predict the conditional probability of receiving the treatment among all the schools in the sample allowing us to find a control group that resolves the potential selection bias problem.

Assuming the matrix X is a set of observable characteristics that determine the participation in the treatment and its result, the key hypotheses to eliminate selection bias are:

(a) $Y_0 \perp T \mid X$, that is, independence of the potential results in relation to the treatment, given the characteristics of the observables (treatment ignorability assumption);

(b) Implicit common support hypothesis: $0 < \Pr(T=1 \mid X) = p(X) < 1$ for $\forall X \in \chi$, where χ is the support of the distribution of X. There is no value of X for which one can say for sure to which group (T=1 or T=0) it belongs.

These two hypotheses are known as strong ignorability. Rosenbaum and Rubin (1983) showed that, given (a) and (b), then the following also holds:

(c) $Y_0 \perp T \mid p(X)$, where p(X) is the probability of being treated given X, or the propensity score. This hypothesis reduces the dimension necessary to resolve the matching.

The identification assumption thus depends on there being no unobservable variables that affect the schools' results differently between the treated and control group.

The current econometric literature contains various methods based on propensity scoring to infer causality between the treatment and the result. One of the best known is propensity score matching, whereby the treated units are matched with the control units according to their estimated probabilities, assuming some hypothesis on the functional form with which X affects the treatment probabilities and the result. The use of the propensity score has the advantage of reducing the dimensional size of the covariates, facilitating their operationalization. However, the literature contains criticisms of this method. The main bone of contention is that the function p(X) is estimated and it can affect the variance of the estimator in the matching. It is impossible to know the asymptotic distribution of the propensity score, therefore the standard errors of the estimators may be unreliable.

We thus used linear regression with estimated propensity score weighting to find the estimate of the average affect of the treatment on the treated (ATT). The idea is to attribute different weights to the schools in the control group according to the characteristics and probabilities of participating in the OBMEP. According to the econometric literature, this method has advantages over others based on the propensity score, mainly with respect to the estimator's efficiency even with imposition of a functional form to estimate p(X)[6].

We implemented this method of combining regression with propensity scoring weighting in two steps. In the first step we estimated $\hat{p}(X) = pr(T_i = 1 | X_i = x)$ from a binary response model assuming a standard logistic distribution (logit) function. In the second step we used a linear regression of Y_i (math score) on T_i and X_i weighted by a function of the treatment and non-treatment probabilities resulting from the first step estimation. From there we found the average treatment effect on the treated (ATT). The weighting used was given by:

$$w_i = \frac{\hat{p}(x_i)}{p} \circ \frac{1-p}{1-\hat{p}(x_i)} \circ \frac{1-T_i}{1-p} \text{ for untreated observations; and} \qquad (1)$$

$$w_i = \frac{T_i}{p} \text{ for treated observations} \qquad (2)$$

where $p = \sum_{n=1}^{n1} \hat{p}(X) | T = 1$ and n1 = number of treated units.

The resulting estimator can be defined as doubly-robust according to the estimators developed by Robins and Rotnitzky (1995)[7]. In explaining the advantages of the method combining regression and propensity score, Imbens and Wooldridge (2008) made an analogy to the omitted variable bias problem. Suppose one's interest is to estimate the coefficient of the treatment in a linear regression of Y_i on T_i, X_i and a constant. Upon carrying out a regression of Y_i only on T_i and the constant, a bias is produced, equal to the product of the coefficient of X_i of the long regression and the coefficient of X_i in a regression of T_i on a constant and on X_i. The weighting factor can be interpreted as the factor that removes the correlation between T_i and X_i, and the linear regression can be interpreted as the factor that removes the direct effect of X_i. As a result, this estimator leads to additional robustness not found in other methods based on the estimated propensity score, by removing the correlation between the omitted covariates and by reducing the correlation between the omitted and included variables.

To choose the set of variables X included to estimate the logit and the weighted regression of Y_i on T_i and X_i, we used the method of stratification by estimated probability proposed by Dehejia and Wahba (1999), by which within each stratum we verified the balancing of each component of X between the treatment and control groups. This method ensures efficient estimation of $\hat{p}(X)$. We divided the sample into four strata according to the estimated $\hat{p}(X)$ and tested the balance of each component of X. For cases of imbalance, we performed iterations or changed the model's functional form until all the included variables were balanced.

The last column in Table A of the Appendix shows the t-statistics of the balancing test for all the variables included in the estimation. It can be seen that there are no significant differences between the treatment and control groups when the samples are divided by strata of $\hat{p}(X)$, proving that the distribution of the included variables is balanced between the two groups.

The covariates included in the model refer to the infrastructure and teaching conditions of the schools (average class size, average number of hours per day, school size), characteristics of the municipalities where

[6] For more details on the impact methodologies, see also Imbens and Wooldridge (2008).
[7] See also Scharfstein, Rotnitzky and Robins (1999).

they are located and information on the students. The student data were taken from the socioeconomic questionnaire of the *Prova Brasil* and contain important information on the profile of the students attending the treated and untreated schools. We also included dummies to identify the different regions of Brazil where the schools are located.

We included the Portuguese scores for 2005 and 2007 on the *Prova Brasil* because each school's raw score in the evaluations can bring relevant information not captured by the other school inputs considered in the model, such as school quality or management. These can influence the probability that a school participates in the Olympiad, so it can be correlated with the result of the math evaluation[8].

In the first step of the method, the explanatory power of the logit with inclusion of all the covariates mentioned was 14.6%. Most of them were significant at 10% to explain participation in the OBMEP, except the following (6 out of 32 variables): high teacher turnover, interruption of school activities; whether the school has a student entry test; and three student profile variables (race, sex and previous attendance in preschool)[9].

In the next section we present the results of the impact evaluation.

4 Results of the impact evaluation

The results presented here refer to the impact of schools' participation in the 2007 OBMEP on the average math achievement of 9th-grade students on the *Prova Brasil* for 2007.

As mentioned in the methodology section, the method of combining regression and propensity score weighting is valid to eliminate the selection bias under the assumption of selection on observables. However, because of the availability of the schools' scores on the 2005 *Prova Brasil*, we performed an additional estimate by difference in differences. This method can be used in cases where one has indicators at two moments in time, before and after the treatment. It has the advantage of enabling an additional control by unobservable characteristics that do not change over time. Hence, the estimate of the average effect of the treatment on the treated (ATT) will be more reliable.

But in the case of the OBMEP, in comparing the scores of 2005 and 2007 between the treated and control schools, they could be underestimated, because there are schools that participated since 2005 and the scores that year could have been influenced by this participation. Hence, the estimates by this method control for the potential problems of bias more adequately, but they can underestimate the impact found.

In the difference-in-difference estimates, we considered the same set of variables X to estimate $\hat{p}(X)$ and for the regression. However, we included the math score on the 2005 *Prova Brasil* (pre-treatment score) as an explanatory variable, which we can also call the initial condition of the schools. The covariate is important in estimating the impact of the OBMEP, both to explain the schools' participation in the program and the variation of the scores between the two years.

Also in the difference-in-difference model, instead of using the students' characteristics in level, we considered their variations between 2005 and 2007, to capture the changes in student profiles that could be

[8] In this form, we restricted the sample and the results found to schools that participated in the 2005 and 2007 *Prova Brasil*.

[9] We performed another test to see if the variables included fit well with the probabilities of participation in the program between the treated and untreated schools. To do this, we compared the distributions of $\hat{p}(X)$ before and after the matching. The distributions before and after matching were very similar, demonstrating that the variables included ensure good quality of the matching. The graph in the appendix compares the density kernel distributions before and after matching for the treated and control schools.

related to their performance. The table below shows the results of the average impact measured by these two methods[10].

Table 3 – ATT estimates: Impact of the 2007 OBMEP on the math scores of 9th-graders on the *Prova Brasil*[11]

	Average Impact on the Math Score		No. of Schools	
	Coeff.	t	Treated	Control
Differences of means without controls	7.44***	10.89	22,703	1,756
ATT	2.14***	4.73		
ATT (diff-in-diff)*	1.99***	5.03		

*** estimates significant at 1%.

* For this estimate we included the initial condition measured by the score in 2005 as an independent variable.

The difference between the scores of the schools that signed up for the OBMEP in 2007, without any control for their characteristics, is 7.44 points. In contrast the estimated ATT is 2.14 points, and by the difference-in-difference model the ATT is 1.99 points, both of which are statistically significant at 1%. These results demonstrate that the OBMEP had a positive and significant impact on the average math scores of 9th-graders ranging from 1.99 to 2.14 points. The difference between the two estimates is small, giving greater confidence in the signs and significance to the estimated impact.

Tables B and C in the Appendix present the results of the second step, according to equations (1) and (2), for the scores in level and differences.

We performed a robustness check to verify whether by applying the same methodology we could estimate the program's impact for a population that was not directly exposed to the treatment, although it was related to a group that did receive the treatment. For this we used the math scores of the fifth-grade students of the schools that participated in the OBMEP in 2007. These students were not directly exposed to the treatment since the OBMEP is offered only to students starting in the sixth grade.

We did not expect to find any impact. If an impact is found, it can indicate that some unobservable variable influenced (biased) the results of the 8th grade estimates. We calculated estimates both by the method of combining regression with propensity score and by matching (nearest neighbor), for the specifications in level and difference in differences. As expected, we did not find a statistically significant impact in any of them. This outcome validates the methodology and the results found for the eighth-graders.

The result in Table 3 is the average impact of participating in the 2007 OBMEP independent of the number of times schools participated in other years. As stated before, most of the schools that participated in the OBMEP in 2007 also did so in at least one previous year. The table that follows shows the distribution of schools in the sample by the number of participations in the OBMEP since 2005.

[10] All the control variables used in the estimates are presented in Table A of the Appendix, with the exception of the regional dummies.

[11] We also carried out the estimates by the propensity score matching method. The results had the same sign and significance of those presented here.

Table 4 – Schools in the sample by number of times participating in the OBMEP

Participations	Number	Percentage
Once	1,960	9%
Twice	5,104	22%
Three Times	15,639	69%
Total	22,703	100%

Based on this, we also performed estimates to distinguish the impact according to the number of participations in the three versions of the OBMEP from 2005 to 2007. For this we constructed three samples: i) the first considering as treated the schools that only participated in 2007; ii) the second considering as treated only the schools that participated twice (in 2006 and 2007 or 2005 and 2007); and iii) the third considering as treated only the schools participating in all three years. The control group in each case was composed of the schools that never participated in the OBMEP. We applied the same method and set of variables explained before, both in the in-level and difference-in-difference specifications. The table below summarizes the results.

Table 5 – ATT estimates: Impact of the 2007 OBMEP on the math scores of 8th-graders on the *Prova Brasil*, by number of participations

	Observed Difference of Means		**Impact of the 2007 score (level)**		**Impact of the 2007 score (diff-in-diff)**		**No. of Schools**	
	Difference	t	ATT	t	ATT	t	Treated	Control
Once	1.83	2.07	0.76	1.92	0.65	1.65	1,960	1,756
Twice	3.92	5.23	1.51	3.82	1.51	4.10	5,104	1,756
Three Times	8.94	12.93	2.38	4.70	2.19	5.02	15,639	1,756

All the ATT estimates indicate statistically significant and positive impacts and indicate that the greater the number of participations in the Mathematical Olympiad, the greater the effect on the school score. Under the in-level specification, the schools that only participated in 2007 obtained an average of 0.76 point higher in comparison with the control group, while schools that participated twice (2007 and one other year) obtained an average of 1.51 points more and schools that participated in all three years obtained 2.38 more points[12].

The estimates by difference in differences, a method that allows controlling for constant unobservable characteristics between 2005 and 2007, corroborate this result, showing significant and positive results for all the samples. However, the ATT point values are slightly lower for the score in level (2007).

Note that the schools that only participated once in the Mathematical Olympiad, in 2007, presents diff-in-diff estimates that are free from contamination of a participation in the 2005 Olympiad. We found a small drop in the point value of the impact, indicating that controlling for unobservable characteristics is important to find an unbiased impact. The explanation for the existence of this bias caused by unobservables could be related to school characteristics not measured, such as the abilities or involvement of the principal, the dedication of teachers, or any other aspect not included in the dataset that might have influenced the scores in a different way between the treatment and control groups and that did not change between 2005 and 2007.

[12] We did not evaluate the *Prova Brasil* for 2008 because the results were not available. The next test will be given in 2009.

For the other difference-in-difference estimates presented here, this analysis is not completely valid, because the reduction in the values in the diff-in-diff estimates might have occurred also due to contamination of the scores from 2005, as the school had participated that year.

We carried out one more exercise examining the same schools over the three years. The idea was to compare schools that did not participate in 2005 but did in the next two years with schools that did not participate in any of the three years. The difference in this new sample (ii) is that the treatment group only contains schools that participated in the Mathematical Olympiad two consecutive years (2006 and 2007), which represents 60% of the schools in the original sample (ii).

The results follow the same direction and significance (at 1%) as the estimate for the original sample (ii). The point values are very close, with an impact of 1.35 in 2007 in the in-level specification and 1.25 in the difference-in-difference specification. The advantage of this exercise is that the estimate is also free of contamination due to participation in the Mathematical Olympiad in 2005.

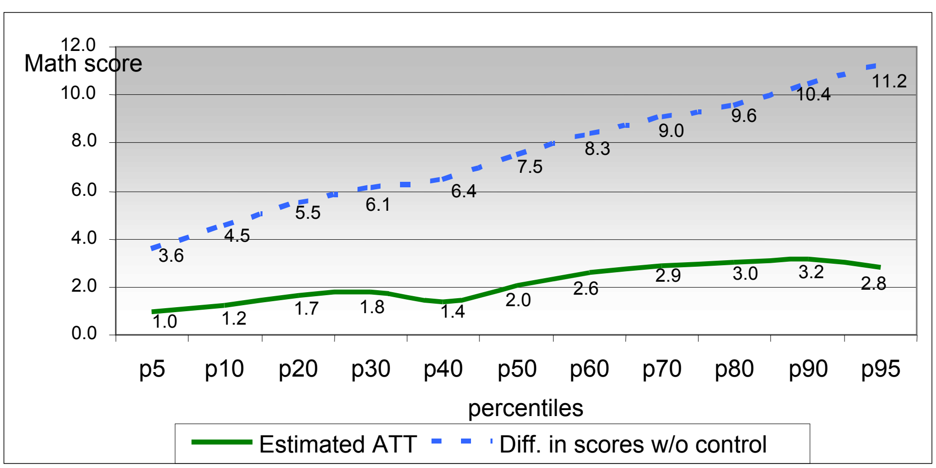
4.1 Impact estimates by percentiles of the distribution of student scores

One of the objectives of the OBMEP is to identify young people with a talent for math so as to give better opportunities to youths who may come from adverse socioeconomic conditions. Also, according to the rules only the students with scores in the top 5% continue to the second phase of the Mathematical Olympiad.

Therefore, we calculated estimates separately, considering the student scores in specific percentiles as the result variable, to shed light on whether there is a difference in the impact depending on higher or lower student scores on the *Prova Brasil*.

From the distribution of math scores of all 9th-graders, we calculated the scores in the deciles of the distribution, and also in the 5th and 95th percentiles, and used these as results in the impact estimates. For each percentile we applied the same methodology and same set of observable characteristics as described previously, considering all the schools participating in 2007, irrespective of the number of times participating. The graph below depicts the results.

Graph 1 – Estimates of the impact of the OBMEP by math score percentiles



All the estimates are significant at 5%. We used the specification in level of the scores in the percentiles.

The solid green line shows the ATT estimates in the percentiles. All the results are significant and positive for all levels, indicating the OBMEP improves the scores of all students. However, the effect is stronger for the higher percentiles, with impacts ranging from 1 to 3 points on the math score.

We stress that the curves of the impact by percentiles have different slopes when comparing the differences observed in the scores and the estimated ATT. The difference in the simple average between the scores of the treated and control schools increases for better students much more sharply than the difference in the impact of the OBMEP.

Table 6 – Math score percentiles: Average between the treated and control schools, ATT and relative impact

Percentiles	Score of the participating schools	Score of the non-participating schools	Estimated ATT	ATT/mean score of untreated schools (%)
p5	178.52	174.95	0.97	0.56
p10	190.66	186.15	1.24	0.67
p20	206.92	201.45	1.67	0.83
p30	219.11	213.04	1.77	0.83
p40	229.79	223.37	1.40	0.63
p50	239.90	232.44	2.05	0.88
p60	250.14	241.80	2.59	1.07
p70	261.28	252.26	2.91	1.15
p80	274.44	264.88	3.01	1.14
p90	292.53	282.09	3.17	1.13
p95	306.94	295.71	2.84	0.96

Table 6 shows the averages observed in the treatment and control schools in each percentile, the estimated ATT and the ratio between this and the average score of the control schools. The results here show that the relative impact is greater for students with better performance on the *Prova Brasil*, but is present for all students. This fact is important, demonstrating that the OBMEP improves the average performance of schools, and that this gain affects all students, not only the higher achievers.

5 Analysis of the Economic Return

From the estimated impacts of the OBMEP on the math scores on the *Prova Brasil* of 9th-graders, we performed an analysis comparing the costs and benefits of the program over the students' lifetimes. The idea is to translate the effect found in the previous section into monetary benefits. This required making some assumptions on how higher math scores can affect job earnings and constructing scenarios to compare the additional earnings against the amounts invested in the program.

In 2007, 4.9 million students took part in the first phase of the OBMEP at level 2, of whom 9% participated only in 2007, 23% in one other year (2005 or 2006) and 69% in all three years. All these students could have benefited from the OBMEP.

To analyze the flow of benefits we used the following hypotheses and procedures:

(i) The estimated positive impact on the average score of 9th-graders holds in absolute values for all the students signed up for the first phase of the 2007 OBMEP. Since we know the impact separately according to the number of participations, we performed three calculations of the return.

(ii) The expected monetary return from participation of the students in the OBMEP was calculated from this improvement in performance, by number of participations. A study with panel data in the United States shows this relationship exists (MURNANE et al., 2000). There are no panel data available in Brazil to follow the same individuals, but the study by Curi and Menezes-Filho (2007) evaluated whether the quality of learning measured in terms of math proficiency obtained on the SAEB (Basic Education Evaluation Test) among high school seniors of a determined generation affects the salaries of this cohort five years later. The authors showed that the performance on the educational evaluations affects future salaries with an estimated elasticity of 0.3. According to these findings, the improved performance of the 9th-graders on the *Prova Brasil* will boost their future wage income with an estimated elasticity of 0.3.

(iii) We also estimated the impact on grade-promotion and dropout rates in primary and secondary school to verify possible effects on the flow and years of schooling of the treated students, but we found no significant results.

(iv) We assumed that the returns of education on salary are constant over time. With data from the 2007 PNAD (National Household Survey), we projected the annual wage earnings of an 18-year-old with nine years of schooling when entering the job market until retirement age of 60.

From the estimated impact on the math scores of students by the number of participations in the OBMEP and the performance-income elasticity, we calculated the expected variations in annual salaries. For one participation (percentage variation of 0.32% on the average of the treated students) we assumed an increase of 0.10% in future annual salaries, with these percentages rising to 0.19% with two participations and 0.30% with three.

With respect to the costs of the OBMEP, we considered a figure of R$2.00 per student per year. Besides this, we formulated a second scenario assuming the existence of opportunity costs, with costs of R$7.60 per student per year. We considered the costs to be proportional to the number of times the schools participated in the OBMEP.

The table that follows presents the economic return of the program, broken down by number of participations, considering the two cost scenarios.

Table 7 – Economic Return of the OBMEP

		Once	Twice	Three Times
Scenario 1	Total NPV	R$ 28.3 million	R$ 136 million	R$ 736.9 million
	NPV/student	R$ 67.80	R$ 133.70	R$ 211.65
	IRR/year	39%	42%	45%
Scenario 2	Total NPV	R$ 26 million	R$ 124.6 million	R$ 678.4 million
	NPV/student	R$ 62.20	R$ 122.50	R$ 194.85
	IRR/year	22%	23%	25%

We considered a discount rate of 5% a year to calculate the IRR.
US$1 = R$2.19 (average 1st semester 2009).

In both cost scenarios, the program's return is positive and high. The greater the number of participations, the greater the return is, reaching a net present value per student of R$211.65, a total NPV of R$ 736 million and an internal rate of return of 45% a year in scenario 1. In scenario 2, even with higher costs, the IRR is 25%. Note that the incremental value of each participation per student is relatively constant, with respective gains of R$67.80 in present value for the first participation, R$65.90 more for the second participation and R$77.95 more for the third participation (scenario 1).

By calculating the average return weighted by the percentage of participating schools, we obtained a NPV per student of R$181.70 and an IRR of 45% a year. The total NPV, that is, the sum of the gains of all students participating, is R$901 million. This indicates that the Mathematical Olympiad appears to be a good investment in terms of public policy because the costs are relatively low and the number of beneficiaries is very high.

Therefore, according to this impact evaluation, the OBMEP has a positive influence on the quality of public school education, increasing the average math score of the treated schools in national educational assessments. This result becomes more pronounced as the number of times a school participates increases and for students with better school performance. Our calculation of the economic return shows the OBMEP has a high rate of return and will generate future earnings benefits for the participants, without considering other possible positive externalities for the students and for the society in general.

6 Conclusion

We carried out an economic evaluation of the Brazilian Public School Mathematical Olympiad (OBMEP), a program promoted since 2005 by the Institute of Pure and Applied Mathematics (IMPA) in partnership with the Ministry of Education and the Ministry of Science and Technology. One of the program's aims is to encourage the study of mathematics at public schools and to increase the quality of public education.

Each year since its inception the OBMEP has attracted an increasing number of participants from schools and students in the sixth through 12th grades. In 2009 the number of students participating was over 19 million, from over 40 thousand schools. It is currently considered the leading school academic competition in the country. To see whether it is living up to its goals, we evaluated the impact of schools' participation in the 2007 OBMEP on the average math scores obtained in the 2007 *Prova Brasil* (INEP/MEC) for 9th-graders.

We used a two-step estimation, combining linear regression with propensity score weighting. The resulting estimator is asymptotically more efficient than other methods based on this estimated probability and thus is considered doubly-robust. We showed that the OBMEP has a positive and statistically significant effect of 2.14 points on the average math scores of the 9th-graders of schools on the *Prova Brasil* (2007). This impact rises as the number times the school participates in the program increases, and is greater in the higher student score percentiles.

The analysis of the economic return also brought positive results, showing that the OBMEP, by enhancing the quality of public school education in the country, will generate future gains in terms of earnings of the participants.

References

Dehejia, R. and S. Wahba. Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association*, 94, pp. 1053-1062, 1999.

Franco, A. M. P.; Menezes-Filho, N. A. Uma análise de rankings de escolas brasileiras com dados do SAEB. Artigo apresentado no XXXVI Encontro Nacional de Economia, Salvador (BA), Dec. 2008.

Heckman, J., Navarro-Lozano. S. Using matching, instrumental variables and control functions to estimate economic choice models. *NBER* Working Paper 9497, 2003.

Holland, P. Statistics and Causal Inference (with discussion). *Journal of the American Statistical Association*, 81, pp. 945-970, 1986.

IBGE – Pesquisa Nacional por Amostra de Domicílio, 2007.

IBGE – Censo Demográfico, 2000.

Inep/MEC – Microdados da Prova Brasil 2005 e 2007 e do Censo Escolar 2006.

Imbens, G. M. and Wooldridge, J. M. Recent developments in the econometrics of program evaluation, *NBER working paper series,* WP 14251, Cambridge, 2008.

IMPA Regulamento da OBMEP, available at www.obmep.org.br.

Mizala, A.; Romaguera, P.; Urquiola, M. Socioeconomic status or noise? Tradeoffs in the generation of school quality information. *Journal of Development Economics*, vol. 84, pp. 61-75, Sept. 2007.

Murnane, R. J.; Willett, J. B.; Duhaldeborde, Y.; Tyler, J. H. How important are the cognitive skills of teenagers in predicting subsequent earnings? *Journal of Policy Analysis and Management*, vol. 19, n. 4, pp. 547-568, 2000.

Robins, J. M. and Rotnitzky, A. Semiparametric Efficiency in Multivariate Regression Models with Missing data, *Journal of the American Statistical Association*, 90, pp. 106-121, 2005.

APPENDIX

Table A – Profile of schools participating in the OBMEP

			Difference of means test	
	Participants	Non-participants	t-statistic before balancing	t-statistic after balancing ***
1. Information from the *Prova Brasil**				
Score_9th_MAT 05	239.7	233.1	-12.15	-
Score_9th_MAT 07	239.8	231.7	-17.25	-
Score_9th_Portugues 05	224.4	220.8	-7.35	1.23
Scrore_9th_Portuguese 07	227.8	222.4	-12.68	0.36
1.1 Questionnaire on Principals				
% of principals with postgraduate degree	71.2	62.5	-7.53	-0.53
% of principals with 11 to 15 years in heading the school	4.7	6.9	4.00	-1.12
% of principals with more than 15 years heading the school	3.2	4.1	2.07	0.35
% of principals between the ages of 30 and 39	24.1	22.0	-1.97	-0.63
% of principals who were chosen by a competitive exam or election process	17.5	13.4	-4.20	0.35
% of schools receiving state government funding	68.0	58.1	-7.44	0.35
% of schools with student entry test	0.8	1.0	1.12	0.17
% of schools with high staff turnover	38.7	32.9	-4.67	-1.06
% of schools with interruption in school activities	19.0	20.8	1.81	0.44
1.2 Questionnaire on Students (9th grade)				
% of students starting with preschool	81.7	78.1	-10.02	-0.33
% of boys	45.7	46.0	0.86	0.85
% of students with parents who attend parent-teacher meetings	91.6	89.1	-13.63	-1.88
% of students with at least one parent who completed high school	7.4	5.2	-12.71	-0.37
Average number of cars owned by the household	1.7	1.7	7.94	1.18
% of white students	35.7	34.2	-2.96	-0.21
Average 8th-grade enrollment	92.7	63.2	-18.09	-1.64
% of municipal schools (as opposed to state schools)	35.0	44.6	8.16	-1.21
2. 2006 School Census*				
Average number of teachers in 1st-9th grades	28.2	26.0	-6.40	0.01
% of schools with Internet access	57.0	44.5	-10.25	1.17
% of schools that have computers for use by 8th and 9th graders	39.0	25.5	-11.21	0.79
% of teachers with college diplomas	88.3	84.6	-6.73	-1.69
% of schools with a "cycle system"[13]	36.7	32.2	-3.74	1.76
Average class size in the 9th grade	32.3	30.2	-9.45	-0.28

[13] A system where schooling is organized into cycles lasting 2 to 4 years, where students can only be held back one year at the end of a cycle.

Average number of hours in the 9th grade school day	4.4	4.3	-10.92	0.19
3. 2000 Demographic Census** Average population of the municipalities	635,183	1,414,768	15.57	-1.26
Average per capita income of the municipalities (R$)	263.4	311.8	12.31	-1.73

Source: *Inep/MEC **IBGE.

*** For the balancing test, we divided the sample into four p-score strata, to provide the statistics for the first quartile, with the results for the other quartiles following the same scheme.

Table B – LS regression weighted by the propensity score (doubly robust)

Y = math score in the 9th grade	coefficient	robust standard error	t	p-value
Participated_2007 OBMEP (ATT)	2.14	0.45	4.73	0.00
Dummy_Midwest Region	0.24	0.89	0.27	0.79
Dummy_Southeast Region	2.33	0.91	2.56	0.01
Dummy_South Region	0.57	1.12	0.51	0.61
Score_Prova Brasil 2005_PORT 9th	0.06	0.02	3.41	0.00
Score_Prova Brasil 2007_PORT 9th	0.86	0.03	33.31	0.00
Log_9th grade enrollment	-0.28	0.40	-0.70	0.49
Municipal_school	-0.04	0.54	-0.08	0.94
No of teachers 1st-9th grades	0.03	0.02	1.78	0.08
School has Internet access	0.18	0.43	0.42	0.67
Use_computers by students	-0.44	0.46	-0.95	0.34
School with cycle system	0.51	0.66	0.77	0.44
Cycle*municipal school (interaction)	-0.50	0.76	-0.66	0.51
Prop. of teachers with college degrees (squared)	0.77	0.78	0.99	0.32
Average class size in 9th grade	0.01	0.03	0.49	0.63
Average hours in school day	-1.39	0.47	-2.97	0.00
Log_municipal population	-1.07	0.22	-4.82	0.00
Log_per capita municipal income	0.24	0.71	0.34	0.73
% of principals with postgraduate degrees	0.46	0.47	0.96	0.34
% of schools with high staff turnover	0.41	0.45	0.92	0.36
% of schools with interruption in school activities	-0.83	0.38	-2.18	0.03
Age_principal (between 30 and 39)	0.26	0.50	0.52	0.60
Time heading school (more than 10 years)	-1.65	0.61	-2.71	0.01
Principal chosen by competitive exam or election	-0.38	0.68	-0.56	0.58
School has student entrance exam	0.37	1.64	0.22	0.82
School receives state government funding	-0.25	0.38	-0.66	0.51
% of white students	4.95	2.09	2.37	0.02
% of parents who attend parent-teacher meetings	0.83	2.64	0.32	0.75
% of students with a car in the household	-8.61	2.19	-3.92	0.00
% of students who started with preschool	-3.72	2.67	-1.40	0.16
% of boys	13.14	2.19	6.00	0.00
% of students with at least one parent who completed high school	7.48	3.56	2.10	0.04
Constant	50.53	6.71	7.53	0.00

R2 = 0.86
Number of observations: 13,217

Table C – Diff-in diff: LS regression weighted by the propensity score (doubly robust)

Diff_in_diff (Y2007 - Y2005)	coefficient	robust standard error	t-stat	p-value
Participated_2007 OBMEP	1.99	0.39	5.03	0.00
Dummy_Midwest Region	1.23	0.83	1.48	0.14
Dummy_Southeast Region	3.12	1.01	3.08	0.00
Dummy_South Region	3.66	1.08	3.40	0.00
Score_Prova Brasil 2005_MAT 9th	-0.83	0.02	-38.71	0.00
Score_Prova Brasil 2007_PORT 9^{th}	0.81	0.03	31.14	0.00
Log_9th-grade enrollment	0.14	0.36	0.39	0.70
Municipal_school	0.01	0.51	0.02	0.98
No. of teachers in 1st-9th grades	0.01	0.02	0.64	0.52
School has Internet access	0.39	0.39	0.99	0.32
Use_computers by students	-0.31	0.39	-0.79	0.43
School has cycle system	1.37	0.59	2.32	0.02
Cycle*municipal school (interaction)	-0.74	0.77	-0.96	0.34
Prop. of teachers with college degrees (squared)	0.56	0.73	0.76	0.45
Average class size in 9th grade	0.02	0.03	0.68	0.50
Average hours in school day	-1.16	0.39	-2.97	0.00
Log_municipal population	-1.12	0.15	-7.31	0.00
Log_per capita municipal income	0.70	0.76	0.92	0.36
% of principals with postgraduate degrees	0.20	0.46	0.44	0.66
% of schools with high staff turnover	0.39	0.37	1.06	0.29
% of schools with interruption in school activities	-0.93	0.39	-2.41	0.02
Age_principal (between 30 and 39)	0.02	0.41	0.04	0.97
Time heading school (more than 10 years)	-1.51	0.57	-2.64	0.01
Principal chosen by competitive exam or election	-0.84	0.58	-1.45	0.15
School has student entrance exam	0.68	1.64	0.41	0.68
School received state government funding	-0.26	0.40	-0.65	0.52
Var. in % of white students	1.76	1.56	1.13	0.26
Var. in % of boys	8.64	1.85	4.66	0.00
Var. in % of students who live with their parents	2.22	1.38	1.61	0.11
Var in % of students starting in preschool	-0.25	2.51	-0.10	0.92
Var in % of students who work	-0.65	1.65	-0.39	0.70
Var. in % of students with washing machine at home	-3.36	2.00	-1.67	0.09
Var. in % of students with at least one parent who completed high school	5.44	2.92	1.86	0.06
Var. in % of students who did not respond to question on mother's schooling level	1.61	1.92	0.83	0.40
Var. in % of students with Internet access at home	0.76	2.51	0.30	0.76
Constant	20.28	3.63	5.58	0.00

R2 = 0.7103
Number of observations: 13,217

Graph A – Quality of matching (nearest neighbor): Density kernel of the probabilities of treatment before and after matching