Universidad de San Andrés

Maestría en Economía

No Child Left Behind and the effectiveness of Professional Development Programs on Students' Performance. Evidence from Schools in Missouri.

Autor: Nicolás Matías Ajzenman

Mentor: Sebastián Galiani

Buenos Aires, 7 de agosto de 2013

No Child Left Behind and the effectiveness of Professional Development Programs on Students' Performance. Evidence from Schools in Missouri.

Nicolas Ajzenman[1]
Universidad de San Andres

Advisor: Sebastián Galiani

ABSTRACT

Among other reforms, the No Child Left Behind Act (NCLB) included special funding for Professional Development Programs, with the particular objective of enhancing the performance of the neediest schools. In this paper, we evaluate for the first time a state-wide Professional Development program in Elementary and High Schools in the State of Missouri. Using matching and Diff-in-Diff techniques, we find that the Program seemed to have a positive and significant effect in the Math scores (annual standardized tests) for almost all the evaluated grades in the High School level (+2% to 4.7%) during 2011-12. However, no significant effects were found in Math or Science in Primary School. The results are particularly relevant because the NCLB Act clearly establishes automatic mechanisms to punish or reward schools using our outcome as the decision variable. The Professional Development Programs aim to improve the children scores in the standardized tests to fulfill the NCLB requirements.

[1] This paper is my Master's Thesis (Universidad de San Andrés) and is based on a Research Project on Education ("Missouri Department of Higher Education - Improving Teacher Quality Grant Program Evaluation") with Sebastián Galiani to whom I am very grateful for his guidance. Some of the results exhibited in this paper can also be found in a Technical Report published by M.A. Henry Consulting, LLC, titled "Missouri Department of Higher Education- Improving Teacher Quality Grant Program Evaluation", where I worked in the Quantitative Analysis. Suggestions or comments can be sent to nicolas.ajzenman@gmail.com. Any remaining errors are mine.

1. Introduction

The No Child Left Behind (NCLB, 2002) Act is undoubtedly the most far-reaching education policy initiative in the USA since the Elementary and Secondary Education Act (ESEA, 1965). The trademark of this legislation required states to conduct annual student evaluations to identify schools that are failing to make "adequate yearly progress" (AYP) towards the explicit goal of having all students achieve proficiency in reading and math by 2013-14. The notion of "Accountability" is remarkable independently of the specific programs and policies designed to improve the students' performance. A vast literature has been written in this respect, theoretical and empirical, to support or to reject the apparent benefits of NCLB's accountability improvements. Critics usually argue that accountability has several negative consequences for the broad cognitive development of children. Nichols and Berliner (2007) argue that NCLB and other test-based accountability policies cause educators to shift resources away from non-tested subjects (like arts) and to focus instruction in math and reading on the set of topics that are most surely represented on the standardized tests (Rothstein et al. 2008, Koretz, 2008). On the other hand, Dee and Jacob (2011) found generally positive impacts in some especial groups: a large impact on Math, 4th graders, more concentrated among white and Hispanic students, among students who were eligible for subsidized lunch, and among students at all levels of performance; and a moderate positive effect in 8th grade for Math tests.

Beyond the potential "accountability effects", the NCLB Act has triggered new Professional Development programs for teachers (work-shops, summer schools, especial courses, etc), funded by the Federal Government and aimed at improving the performance of the neediest schools. These programs, which are not new, became notably more common since the sanction of the Act, as it includes specific funding for that matter, with specific goals. Although there is not a single theoretical framework to analyze Professional Development, as there is not one particular method, the theoretical links among Professional Development, teacher learning and practice, and student learning are straightforward (Cohen & Hill, 2000; Fishman, Marx, Best, & Tal, 2003; Garet et al., 2001; Guskey & Sparks, 2004; Kennedy, 1998; Loucks-Horsley & Matsumoto, 1999).

Professional development is supposed to affect student achievement through three steps: (a) it enhances teacher knowledge and motivation, which (b) improves classroom teaching and, finally, (c) raises student achievement. In any case, the empirical evidence of the effectiveness of these programs is scarce. Moreover, the literature of Math and Science is usually focused on High-School, but not Elementary grades (Slavin & Lake, 2008; Slavin, Lake, & Groff, 2009) and in many cases, even if the applied statistical methods are acceptable, the papers don't fulfill the methodological requirements of the specialized literature in Education (mainly, the duration of the experiment which is usually too short, or the external validity of the results if the experiment is focused in just a few schools or grades, Slavin et al., 2012). What is more, students' scores in standardized tests aren't usually taken as the main outcome, and they should, as the score is the main policy tool derived from the NCLB Act.

In this paper, we evaluate the effects of a state-wide educational policy in Math and Science applied to Primary and Secondary Schools on the standardized tests scores of the students: the Missouri's Cycle-9 programs of the Improving Teacher Quality State Grants. The programs were offered in more than 65 schools in almost all grades in the state of Missouri. This is the first time a rigorous evaluation is performed - although with usual the limitations of the non-experimental techniques - to evaluate the effectiveness of a key aspect of the No Child Left Behind Act: the Professional Development programs. The evaluation of these programs is particularly relevant because it is the final goal of the NCLB Act to improve the performance of the neediest children and performance is measured by policy-makers exclusively using our outcome of interest.

The ideal method to answer our research question is a Randomized Controlled Trial with a random assignment of the treatment and control group. If that would have been the case, we could have been sure of the balance in the pre-treatment characteristics of the control and treatment groups and, thus, of the unbiasedness of our estimations. However, the program was not assigned randomly and thus we had to use non-experimental techniques (matching and Diff-in-Diff) to identify a causal effect with the available data.

Using matching and Diff-in-Diff techniques, we find that the Professional Development Program seems to have a positive and significant effect in Math for almost all the evaluated grades in the High School level (+2% to 4.7%). However, no significant effects were found in Math or Science in Primary School.

The remainder of the paper proceeds as follows. First we describe the No Child Left Behind Act, the Professional Development Programs in Missouri and the particular programs we evaluate in this paper. Second, we describe the data and methods we used to perform the evaluation. In that section, we show how we manage to find a proper counterfactual group – considering that an "ideal" control group was not possible to find, as the assignment of the treatment was not random – and present evidence to support our claim that the estimations are unbiased. The core of our analysis is contained in the Results section, where we discuss the effects of the Programs on the different grades. Finally, we present the conclusions and comments.

2. The No Child Left Behind Act and the Professional Development Programs

2.1. The No Child Left Behind Act

For more than four decades, the Elementary and Secondary Education Act (ESEA), enacted in 1965 under the Lyndon Johnson Administration has been the cornerstone of the national educational system in the USA. This Act, which was announced as one the major subjects of Johnson's "War on Poverty" plan, represented the most expansive federal education bill ever passed (two years after the Act was passed, the annual federal budget for Education jumped from $ 1.5 to $ 4 billion[2]). The federal government's role in education grew significantly, as well as the reporting requirements from local and state school districts.

The Act was clearly focused on assisting those primary and secondary schools serving children of low-income families. Trough a special source of funding (Title I), the law allocated large resources to meet the needs of educationally deprived children, especially through compensatory programs for the poor. More specifically, the "Title I" was thought to provide federal assistance to "local educational agencies serving areas with concentrations of children from low-income families to expand and improve their educational programs by various means (including preschool programs) which contribute to meeting the special educational needs of educationally deprived children" (Section 201, Elementary and Secondary School Act, 1965).

As it was thought in the 1965's Act, the Federal Government would determine the amount of help each state educational department needs. The local departments, in turn, would approve projects plans submitted by the local educational agencies, only if they were consistent with the compensatory spirit of the law. The criteria used to distribute the federal funds was based on more or less objective parameters like the number of schools "in need" (i.e., with children from families with an annual income less than some specific amount). As the state governments had to prove that the federal funds given under Title I would be used specifically for compensatory programs, the law implied a more evident incidence of the Federal Government. Moreover, in order to get the Title I funds, the counties, schools and states had to keep a meticulous accountability of their resources, as well as a periodic evaluation of the situation of their children and teachers, otherwise the eligibility criteria wouldn't be fulfilled and the projects would never be approved.

[2] A brief review of the Act can be found at http://www.gse.harvard.edu/news_events/features/2005/08/esea0819.html

Since the original Act was passed in 1965, it experienced seven reauthorizations and modifications by the Congress. Many changes were applied to the first version, like including specific funds for education of children with disabilities or for migrant children; or the inclusion of special programs for bilingual education; as well as other changes in the criteria used to assign the funds, the inclusion of charter schools, and other reforms.

However, the major change to the original act was undoubtedly the No Child Left Behind Act (NCLB) of 2001, which reauthorized the ESEA including some structural changes. As it stated in the official Act, the NCLB aims "To close the achievement gap with accountability, flexibility, and choice, so that no child is left behind", and this objective is supposed to be achieved by different reforms to the original Act.

2.2. Accountability in the No Child Left Behind Act

One of the main changes brought up by the NCLB is related to the concept of Accountability. The law requires all public schools that receive Federal funds to administer standardized tests (designed by each state) annually to all the students in grades 3rd to 8th and once in high-school, in reading and math. The test results are required to be public; schools must report their performance, and states must exhibit their progress toward their proficiency objectives. As the objective of the law is to ensure that all groups of students are progressing at an adequate rate and, more specifically, to improve the performance of the minorities that had been "left behind", the test results must be broken out and reported according to poverty, race, ethnicity, disability, and limited English proficiency (U.S. Department of Education, 2002).

2.3. Title I funds

The Title I of the Act was - as it was in the original ESEA – designed to help to improve the academic achievement of the Disadvantaged. Title I provides flexible funding that may be used to provide additional instructional staff, professional development, extended-time programs, and other strategies for raising student achievement in high-poverty schools. Since the pass of the NCLB Act, Title I funds should be used only for effective educational practices, as the programs are required to use effective methods and instructional strategies grounded in "scientifically based research" (U.S. Department of Education, 2002). However, the local authorities have considerable flexibility to allocate the Title I resources among their programs, which can be "school-wide" (for schools with 40% or more of low-income students) or "Targeted" (to help especially the low-achieving students).

The accountability requirements are strongly linked with the Title I funds. States must develop and implement a single, statewide accountability system that will be effective in ensuring that all districts and schools make adequate progress, and hold accountable those that do not (U.S. Department of Education, 2002). All districts and schools receiving Title I[3] funds have to meet state "adequate yearly progress" (AYP) goals for their total student populations and for specified demographic subgroups, based on the compulsory standardized test. If these schools fail to meet AYP goals for two or more years, they are classified as schools "in need of improvement" and face consequences to remedy their situation. For instance, if one of these schools fails to meet their AYP target for two or more consecutive years, parents of children in that school have the choice to transfer their children to schools not identified as "in need of

[3] Schools where at least 35 percent of the children in the school attendance area are from low-income families or at least 35 percent of the enrolled students are from low-income families are eligible to receive federal Title I funds. Over half of all public schools receive funding under Title I.

improvement". When the school fails to meet AYP goals for three or more consecutive years, students are eligible for supplemental educational services (like tutoring). When (AYP) is not achieved for four consecutive years, the district must implement structural changes (which can imply the replacement of school staff or a new curriculum, among others). If the AYP is not achieved for five consecutive years, the district must prepare a plan to completely restructure the school; this may be to turn the school to a public charter school or even to enter into a contract to have an outside entity operate the school. A school is no longer considered "in need of improvement" when it meets AYP for two consecutive years.

2.4. Title II funds (Improving Teacher Quality State Grants)

Title II funds, or "The Improving Teacher Quality State Grants program" is intended to increase students' achievement by elevating teacher and principal quality. The program is flexible enough to let each state use the funds trough different interventions. In any case, the law indicates that each state education agency must develop a plan to ensure that all teachers are "highly qualified"[4]. As in the case of Title I, all activities supported with Title II funds must be based on a scientifically research. States apply to the U.S. Department of Education for funding, and funds are allocated through a formula that is a function of the school-age population and the number of children in poverty in each state. About 2.5% of allotted funds for teacher quality activities are distributed on a competitive basis through sub-grants to partnerships of high-need districts and higher education institution that prepares teachers (U.S. Department of Education, 2002).

Each State Agency for higher education may develop criteria for awarding sub-grants to eligible partnerships comprised of at least one institution of higher education, one school of arts and sciences, and one high-need Local Education Agency. The partnerships use the funds to conduct professional development activities in core academic subjects to ensure that teachers, highly qualified paraprofessionals, and (if appropriate) principals have subject-matter knowledge in the academic subjects they teach. The State Agency should identify scientifically projects based on professional development that is effective in increasing student academic achievement (U.S. Department of Education, 2006).

[4] "Highly qualified" means that the teacher must have with full certification, a bachelor's degree, and demonstrated competence in subject knowledge and teaching skills

2.5. Title II funds in the State of Missouri – The Improving Teacher Quality Grant Cycle-9

Within the Improving Teacher Quality State Grants, the state of Missouri has been implementing many programs to improve the educational outcomes of the state students. Some of them were programs funded by sub-grants awarded to partnerships which included the participation of higher education institutions, which year by year participate in competitions to get the resources and implement their programs, usually targeted to teachers of the neediest schools in the state.

In 2010, the Missouri Department of Higher Education (MDHE) published the Request for Proposals (RFP) with terms and objectives of the "Cycle-9 Grant Competition". Each cycle (there is approximately once cycle per year since the implementation of the NCLB Act) has its own RFP, which establishes the grade level and content area focus required for the distribution of awards. The Cycle-9 RFP focused on the core areas of mathematics and science at any grade level (K-12). The Cycle-9 RFP resulted in 16 project proposals requesting more than $ 2.8 million in grant funds. A panel of math and science professionals from K-12 and higher education institutions, as well as MDHE staff members, reviewed the proposals and, finally, five new programs were selected and two on-going projects got a total of near $ 1,025,000 of funding and were implemented during the years 2011-2012. [5]

The Cycle-9 FRP clearly stated the main objectives for the academic year[6]:

- Improve student achievement in targeted mathematics and/or science content areas.
- Increase teachers' knowledge and understanding of key concepts in targeted mathematics and/or science content areas
- Improve teachers' pedagogical knowledge and practices that utilize scientifically-based research findings and best practices in inquiry-based instruction.
- Improve the preparation of pre-service teachers through improvements in mathematics and/or science content and/or pedagogy courses.

Through projects' Professional Development (summer sessions, follow-up sessions and other contacts), the programs seek to improve teachers' content knowledge, classroom practice and use of data and assessment tools. Lesson enhancement, resources and support networks combine to support teachers in improving performance. Pedagogical practice – according to the Improving Quality Teacher

[5] Details can be found at http://www.dhe.mo.gov/ppc/grants/ITQGCycle-9.php
[6] See the Technical Report "ITQG – Cycle 9 Impact Report Synopsis" by M.A. Henry Consulting, 2012

Grant - is grounded in inquiry-based instruction, focusing on modeling and meaningful activities in math and science that promote students' learning.[7]

As requested in the Cycle-9 RFP, all of the seven projects awarded funding were focused on Math, Science or Both. Approximately 72% of the schools were high-need public schools. As it can be seen in the Appendix, each winning project had its own particular objective and target population. What is more, not all the projects followed the same academic method, but all of them have the same overall objective: improve students' achievement in math and or science, trough the professional development of their teachers.

All the programs were based on workshops (or courses). In many cases, the workshops would strengthen the participants' knowledge, but also enhance their teaching strategies and improve the use of technology in their particular courses. In other cases, the courses are aimed at teaching how to link the science (or math) content to the real-world problems or environmental issues. Although the pedagogy specific strategy may vary in each project, the inquiry-based approach and the inclusion of technology is common among all the projects.

2.6 How effective are Professional Development Programs?

The Professional Development Programs have standards of quality developed by many organizations (see Cocoran, 2007, Hawley et al, 1999). Moreover, there is a great consensus about what constitutes effective Professional Development (Loucks-Horsley et al., 2003) in Math and Sciences classes.

Experiments evaluating practical applications of alternative science programs and practices are rare at all grade levels. Most of the few experiments are brief laboratory-type studies, not evaluations of practical programs. Moreover, there is few evidence of the effect of Professional Development programs on students' performance in science using a proper methodology[8]. In addition, the literature is usually focused on High-School, but not Elementary (Slavin & Lake, 2008 and Slavin, Lake, & Groff, 2009). What is more, the authors show that the majority of the studies on Professional

[7] See the Complete Technical Report "ITQG – Cycle 9 Impact Report" by M.A. Henry Consulting, 2012

[8] The criteria to define a proper methodology in this literature doesn't refer only to the unbiasedness of a randomized experiment (or a properly matched control group); the authors also emphasize the duration of the experiments (as there is evidence that year-long evaluations usually evaporate the initial effect of brief programs) and the validity of the "performance measures" (which has to be exactly identical in control and treatment groups. This problem is known as the "treatment-inherent measures problem", see Slavin et. al, 2012). What is more, the authors identify many studies where the unit of assignment is the school, district or town, the unit of observation is the student and there is no clustering of the Standard Errors.

Development are focused on "cooperative learning" strategies among students, instead of purely teacher training.

One example of a rigorously estimated effect of a Professional Development program on the students' performance in science can be found in Mant, J., Wilson, H., & Coates, D. (2007). The authors designed a randomized experiment of Professional Development for Elementary Science courses and found a significant effect on the children' performance using standardized tests in England. In the experiment, treatment schools were provided with extensive professional development intended to increase engagement and conceptual challenge in science lessons Teachers learned to use thinking skills strategies and were encouraged to emphasize higher-order thinking, practical work, investigations, and purposeful, focused recording. This experiment, although used statistically acceptable methods, was exclusively for Science of Elementary Schools and included only sixteen treated schools.

A year-long evaluation of a Professional Development program for Math courses in Elementary courses (Dynamic Pedagogy[9]) was carried out by Armour–Thomas, Walker, Dixon-Roman, Mejia, & Gordon (2006) in two majority African-American K-3 elementary schools. The authors, who matched control and treatment using pretest scores, found a significant improvement in the performance of the treated students. This study was performed exclusively in Mathematics of Elementary schools. Only 60 students were "treated".

The Talent Development Middle School Mathematics Program (a part of a comprehensive school reform model that included extensive professional development and on-site coaching) was evaluated in three inner-city Philadelphia middle schools by Balfanz, Mac Iver, & Byrnes (2006). Two were majority African American and one majority Hispanic. The schools were matched on demographics and test scores with three control schools. After three years of implementation, eighth graders were compared on district-administered SAT-9 scores, controlling for their fourth grade SAT-9 scores, there were no differences in Math Procedures, but there were significant differences in Math Problem Solving. Only three schools participated in this experiment, all of them were High Schools and the outcomes were not the standardized tests used by the NCLB Act.

The empirical literature on Professional Development, even if it properly identifies a causal effect, is usually focused on specific programs and small experiments with a few participant schools. The papers usually focus on one particular program in Math

[9] Dynamic Pedagogy is a professional development program in which teachers learn to prepare and deliver lessons appropriate to students' current knowledge, misconceptions, and past errors in Mathematics.

or Science and in just a few particular Primary or Secondary schools. What is more, students' scores in standardized tests are not always taken as the main outcome.

In this paper, we evaluate the effects of a state-wide educational policy in Math and Science applied to Primary and Secondary Schools on the standardized tests scores of the students. Although for methodologically reasons – that we detail in the Methods section – we don't use a sample including all the treated teachers, the programs were offered in more than 65 schools in almost all grades in the state of Missouri. Moreover, this is the first time a rigorous evaluation (although using non-experimental techniques, with its usual limitations which are detailed in the Methods section) is performed to evaluate the effectiveness of a key aspect of the No Child Left Behind Act: the Improving Teacher Quality State Grants. The evaluation of these programs is particularly relevant because it is the final goal of the NCLB Act to improve the performance of the neediest children and that performance is measured by our outcome of interest: the scores in the standardized tests. The performance of each school is officially measured by the "Adequate Yearly Progress" indicator, which is based on the standardized test scores, and that's our key variable to identify the schools that need structural reforms.

3. Data & Evaluation Methods

Original data was provided at student level by the Missouri Department of Elementary and Secondary Education for years 2009 to 2012 and includes Elementary and High School data.[10] The datasets contain student information (grade, school, score in the standardized tests of Mathematics and Science) and teacher information (treatment status, gender, ethnicity, salary, extra duty salary, highest degree earned, years of experience in the district, the state and in public schools). Data includes students' results in the standardized evaluations in mathematics and sciences at two different stages: Elementary/Middle grades (Missouri Assessment Program Test - MAP) and High School grades (End-of-Course Test - EOC). The data is presented at the student level and it separated by MAP and EOC.

[10] Official Data was provided by the Missouri Department of Elementary and Secondary Education. However, we are not able to determine if the sample is complete (that is, if it contains all the students and teachers of the system in all years for treated and control schools). If the full sample was somewhat trimmed and the restriction was imposed randomly, then our results would be biased by the criteria used to restrict the data we received from the Department.

3.1 Methods for MAP and EOC teachers

The ideal design to evaluate the impact of the programs would be a Randomized Controlled Trial. Unfortunately, the programs were not assigned randomly; instead, the schools that considered as the "neediest" were included as treated. Therefore, as a second best, we had to construct a control group using non-experimental (matching) techniques to identify the causal effect of the programs on the academic result of the students. The selected teachers for control and treatment were then used to perform a Diff-in-Diff analysis. By comparing changes instead of levels, we are able to control for observed and unobserved time-invariant teacher characteristics that might be correlated with the program implementation as well as with the outcomes studied. The change in the control group is an estimate of the true counterfactual—i.e. what would have happened to the treatment group if there had been no intervention.

At this point, it is important to emphasize that the matching technique has certain limitations. In particular, the matching is based on the observables (in this case, pre-treatment characteristics of the teachers and pre-treatment trends), but this doesn't guarantee that the unobservable characteristics are also balanced.

Formally, our difference-in-differences model can be specified by this equation:

$$Y_{it} = \alpha(D_t * T_i) + \beta T_i + \gamma D_t + G_{it} + S_{it} + \mu_i$$

where y_{it} is the outcome of interest in student i in year t, T_i is an indicator variable that takes on the value one if student i´s corresponds to a treated teacher, D_t is an indicator variable that takes on the value one if the year is 2012 (treatment) and 0 if it is the pre-treatment year (2011), the interaction ($D_{t*} T_i$) is our variable of interest and captures the effect of the difference in difference model and, finally, Gi_t and S_{it} are grade and school fixed effect, respectively.

In this model, α is the difference in difference estimate of the (average) effect of the program on the outcomes of interest. The key identifying assumption for this interpretation is that the change in the outcomes of interest in control teachers is an unbiased estimate of the counterfactual. While we cannot directly test this assumption, we can test whether the secular time trends in the control and treatment teachers were the same in the pre-intervention periods (Heckman and Hotz, 1989). If the secular trends are the same in the pre-intervention periods, then it is likely that they would have been the same in the post intervention period if the treated teachers had not been treated.

If it turns out that the pre-treatment secular trends are not statistically the same between treatment and control teachers, we need to match teachers on their pre-treatment trends in order to select a valid control group that satisfies this key identification assumption so we can estimate the average treatment effect on the treated teachers. One methodology to implement this strategy is the generalized difference in difference estimator proposed by Heckman et al. (1997) and Heckman et al. (1998). The idea is to include as matching variables, in addition to observable teacher characteristics, pre-intervention trends of the outcomes of interest. The objective is to obtain a subset of control school that satisfies the Heckman and Robb (1985) criteria with respect to pre-intervention trends.

We are particularly interested in testing the balance of pre-treatment ***trends*** in the outcomes variable we analyze (students scores) between treatment and control groups as the method we implement rely on the assumption that the pre-treatment outcome trends are independent of the treatment status. For that reason, we keep only the observations that correspond to those teachers for whom we have information in the four years (2009, 2010, 2011 and 2012). If a teacher participated in the program in 2012 but he wasn't in the system in 2011, then we can't consider him because we don't know how his performance in the years previous to the treatment was. Even if he was in the system in 2011 but no in 2010, we cannot include him in the analysis, because we are not able to determine how the evolution of his students was throughout the pre-treatment years. This substantially reduces the sample size and diminishes the statistical power of our analysis, but it is the only way to test the balance in the pre-treatment groups and thus consider that estimations are not biased by pre-treatment differences. Although we do know that three years of data is not the ideal number, the pre-treatment analysis is necessary because, if we found significant differences in the three-year trend, we had to rule out our impact analysis.

3.2 Selection of MAP Controls

Ideally, the unit of observation should be the teacher. In the first place, teachers –and not students - are the ones that can be treated or not treated. Besides, **we have only teacher's pre-treatment characteristics to do the matching on the observables**. Nonetheless, the original dataset is presented at the student level: we have math and science scores for each student in the tests taken in 2009, 2010, 2011 and 2012. Doing a collapse of the information to obtain average scores by teachers would imply a drastic reduction of our sample size and would certainly reduce the variability of the available data[11]. An intermediate solution could have been to collapse the data at school-grade-teacher level (that is, each teacher would have as many observations as grades-schools she/he works in), but the problem would be that many teachers moved from one grade to another (or even from one school to another) throughout

[11] If, for example, each teacher has a course of thirty students, the collapse would leave us just one observation for that teacher instead of thirty.

the years and we need each unit of observation to be present in **each** of the four years of the sample.

To solve this problem we implemented a two step approach: first, we used information at teacher's level just to do the matching; second, once we selected the treated and control teachers we merged them with the student dataset to perform the analysis at the student level. Thus, we kept only those students that corresponded to the teachers we selected when we did the matching on the teachers' observable characteristics.

For the first step (matching on pre-treatment characteristics) we had to collapse the information to the teachers' level. Throughout the years of the sample, teachers worked in different grades and even different schools simultaneously and, moreover, many teachers moved from one grade to another (or even from one school to another). Therefore, to make a suitable comparison of teachers among treatment and control groups and analyze if the pre-treatment trends are balanced it is necessary to calculate the average scores by teacher net of the school and grade effect. Thus, we first ran a regression using the average scores by teacher as the dependent variable and the school and grade as the independent variables and then predicted the residuals, which represent what we are looking for: the average scores by teacher net of grade and school effects ("adjusted scores"). [12] We are interested in testing the balance of the adjusted scores trends between the treatment and control groups.

Once we got the adjusted scores we proceed to do the matching to select "similar" treatment and control teachers. For each of the Mathematics and Science groups, we estimated a Propensity Score (the probability of being treated, according to pre-treatment characteristics), using the following Probit model:

$$I_i = \beta Var_{2011-10i} + \gamma Var_{2010-09i} + \gamma Var_{2011-09i} + X_i + \mu_i$$

Where I_i indicates if the observation is a control or treated teacher. $Var_{2011-10i}$ is the difference between the adjusted score of teacher i in 2011 and the adjusted score of corresponding to the same teacher one year before $Var_{2010-09i}$ is the difference between the adjusted score of teacher i in 2010 and the adjusted score of corresponding to the same teacher one year before and $Var_{2011-09i}$ is the difference between the adjusted score of teacher i in 2011 and the adjusted score of

[12] Formally, we ran the following regression for each year: $Y_i = \beta Grade_i + \gamma School_i + \mu_i$; where Y is score of each student, Grade is the grade of the student and School is the school of the student. We are interested in $adj\hat{Y}_i = \hat{Y}_i - \hat{\beta} Grade_i + \hat{\gamma} School_i$, which is the score of the student net of the grade and school effect. Note that the main regression controls for school and grade using fixed effects of each variable. Therefore, as the estimation of the treatment effect will be "net" of school and grade effects, it is enough to analyze the pre-treatment balance trends of the adjusted scores.

corresponding to the same teacher two years before. The variable X_i is a set of pre-treatment characteristics (2011) of the teacher *i*: Years of Experience Teaching in the District, Experience Teaching in Public Schools, Experience Teacher in the State, Annual Salary, Dummy variables for Teacher's highest degree earned and Dummy variables for Gender.[13]

The method used to find a suitable control group was "radius matching": basically, once the Propensity Score is estimated for each observation, each variable within the "treated" group is matched with one or more observations within the "non-treated" group, the condition is that the difference in the propensity scores of the two observations is smaller than an arbitrary radius. The value that takes the radius was chosen in order to maximize the similarity in the pre-treatment characteristics of each treatment observation with its control (especially the pre-treatment trends), but without losing too many observations in the treatment group, which are relatively scarce. We kept the teachers that are included in the arbitrary radius; thus, we kept just those control observations that have a propensity score similar to at least one treatment observation.

3.3 Selection of EOC Controls

As in the case of the MAP tests, we first calculated students adjusted scores (net of grade and school effects), then collapsed the information at teacher's level to test if the pre-treatment characteristics and adjusted scores trends were balanced between treatment and control groups and, finally, merged those teachers with the student's information to do the impact evaluation.

Unfortunately, in the case of the EOC, the number of teachers that simultaneously have information for 2009, 2010, 2011 and 2012 is scarce. To begin with, the number of treated teachers in the original sample is smaller than in the case of MAP, as the Professional Development programs were more focused on primary teachers. Moreover, as there was a high mobility of teachers through schools-years-districts (an expectable consequence of the NCLB Act), merging the Teacher ID Codes of the period 2009-2012 generated a considerable shrunk in the database (that is, many of the teachers that were in the system in 2012, were not in one or more of the previous years)[14], which ended up with only 41 and 53 teachers for Science and Mathematics,

[13] The inclusion of other variables would have caused a rejection in the balancing property. In the case of MAP- Math, the gender was not included for the same reason.

[14] This happened because our database doesn't include all the districts in the State, but only those were the program was carried on. Thus, if a teacher moved from one district to another which wasn't affected by the programs, we lost him/her. The mobility across schools and districts is much more frequent in secondary schools were teachers give classes of one particular subject. Moreover, with the NCLB act, many schools had to renew their programs and/or staff, as a consequence of not reaching the Adequate Yearly Process.

respectively. In these conditions doing a matching (and keeping just the observations within the common support) would reduce the sample size too much, to the point that it would make the analysis impossible. The only choice we had was to test the balance of pre-treatment characteristics and trends between treatment and controls unmatched groups and proceed with the difference in difference analysis only if the adjusted score trends were balanced.

3.4 Pre-treatment balance for MAP and EOC teachers

Tables 1-2 show the pre-treatment balance in different variables for MAP tests: teachers' characteristics (like experience, salary, etc), which were the variables used to do the matching. It is important to emphasize that, as the impact evaluation is estimated using a Diff-in-Diff model, the unbalance in the pre-treatment levels of the outcomes does not imply that the estimation will be biased. However, a good balance in the pre-treatment trends is fundamental: if they are not significantly different in the pre-intervention periods, then it is likely that they would have been the same in the post intervention period if the treated schools had not been treated. Pre-treatment trends analysis was studied separately and was tested formally. The results can be seen in Table 3.[15]

Tables 1-2 show the MAP pre-treatment characteristics for treatment and control groups for the sample without matching and the sample with matching. There are significant differences between pre-treatment characteristics among treatment and control groups in the unmatched sample (like gender or salary), but almost all of these difference disappear once we restrict the sample using our matching criteria. This happens for Math and Science samples. Probably the reduction of the sample size (and the consequent increase in the standard deviation of the estimations) had something to do with this, especially in the variables that are measured at the teachers' level (basically, teachers characteristics) instead of students' level (pre-treatment trends); however, not only the estimated standard deviations changed: the estimations of the means of control and treatment groups are considerably more similar in the matched sample for almost every variable.

Moreover, Table 5 shows that there are no significant differences between the trends of the scores among treated and control groups in the "matching sample", which is the one that we use to perform the impact evaluation.

[15] Unfortunately we don't have data for too many pre-treatment years (just 2009, 2010 and 2011). To formally test if the trends in those years was similar between treated and control teachers, we used the following model: $Y_i = \beta Year_{2011-09i} + \gamma T_{2011-09i} + \gamma (T * Year)_{2011-09i} + School_{FE} + +Grade_{FE} + \mu_{2011-09i}$ where Y is the outcome of interest (math or science scores), Year goes from 1 to 3 (2009 to 2011). The unit of observation in this case is the student. Therefore, variable T takes a 1 when the student is in a class with a "potentially treated" teacher and 0 otherwise. The model includes school and grade fixed effects and, as in the principal model, the errors are clustered at the Year-School-Grade Level. We are interested in the estimation of γ. If it is significantly different from 0, then we cannot assume that the trends are balanced.

Tables 3-4 show the pre-treatment balance in different variables of the EOC observations (no matching): teacher's characteristics (experience, salary, etc). The tables show the average of pre-treatment characteristics and outcomes for treatment and control groups using the unmatched sample. Pre-treatment characteristics (teachers' salaries, degrees, gender, ethnicity, experience in the district, state and in public schools) are balanced.

Although the apparent balance in the teachers' characteristics could be a consequence of the small sample, the most important to assure the unbiasedness of our impact estimations is not the balance in the static teachers' variables, but the balance in the pre-treatment trends of the outcome of interest (scores, which are measured at the students' level). When we test the pre-treatment trends balance (Table 5), we find that, in the case of Math the trends are clearly balanced but, in the case of science, they are significantly different. Therefore, we are able to continue with the impact analysis using the MAP Science and Math "matching" samples and with the EOC Math sample with no matching. We have to disregard the EOC Science sample, as we could not be sure that the results would be unbiased. [16]

[16] The estimated impact for EOC Science scores, which in this case we cannot be sure that is unbiased, was not significant. Detailed results are available upon request.

4. Results

Tables 6 to 9 show the estimated effects of the program on the students' scores. The results are mixed; in the case of the MAP tests, the average effects for all the grades together are not significant in Science and Mathematics. In the case of Science, the effects remain not significant when the analysis is performed analyzing the sample by grade. Curiously, when we analyze the Mathematics scores by grade, we find only one significant effect at 5% in the sixth grade and this effect is negative. Nevertheless, although significant, the number represents a decrease of near 1.9% and, what is more important, the effect turns statistically insignificant when controlling for school fixed effects.

In the case of the EOC tests, the situation is completely different. Even with a considerably smaller sample size, the results are mostly significant. In the case of the Mathematics scores, we find positive and significant effects for all grades together and each of the grades, except for 12th grade, where there is not significant effect. The increase varies in a range of 2% to 4.7% depending on the grade – 11th grade shows the larger effect and 10th grade shows the smaller and the specification (with or without school fixed effects). EOC Science scores show positive and significant effect in each of the grades; however, as we stated in the Methods section, we cannot be sure about the unbiasedness of this estimator, as we found significant differences in the pre-treatment trends.

Broadly speaking, our main findings show that the program was effective at the High School level (EOC tests) and ineffective at the Elementary Level (MAP tests). The positive and significant effects in the EOC tests can be seen for Mathematics in almost all of the analyzed grades.

The programs seem to be effective, at least at the High School level in Mathematics, although ineffective at the Primary level. For the first time, the Professional Development programs designed for the Title II of the No Child Left Behind Act show an impact on the fundamental variable that the law is supposed to affect. As the programs were focused on the neediest schools, which are those that have to show some improvement in their students' scores to reach the adequate yearly progress status, this evaluation provides an especially useful information for the policy makers in Education.

5. Concluding Comments

Professional Development programs for teachers are a fundamental component of the No Child Left Behind policies to reduce the performance gap among the students of different schools. In this paper, we have evaluated for the first time the effectiveness of this component in the State of Missouri.

Our results show that these programs showed a significant and positive impact on High-School students, specifically in the case of Math. However, no significant effect was found for Elementary students. These results are very important, as Math and Science are the subjects were students show the most striking performance gaps.

A relevant question is why the programs were successful for High School students and not for Elementary students. Is it a problem of the teachers, the students or the measurement?

Through projects' Professional Development, the programs seek to improve teachers' content knowledge, classroom practice and use of data and assessment tools in order to improve the teachers performance. Pedagogical practice, if is effective in teachers, is supposed to promote students' learning. In this paper we focus on the "total effect" (the impact on the children scores) and we avoid measuring the intermediate effects (on the teachers behavior and knowledge) because there is not enough available data. However, the technical report of the program (M.A. Henry, 2012) presents some qualitative evidence (class-room observations, teachers surveys, focus groups, etc) that shows an apparent improvement in the teachers' knowledge and teaching methods (which include the design of the class with the new contents and tools and the implementation of the new materials, among others) in both elementary and high school levels. This evidence, although not conclusive, support that idea that the "first step" of the program was effective in Elementary and Secondary teachers: all of them seemed to improved their instruction methods.

What about the transition from improved teacher to improved student? Why it was so different in Elementary and High School students? A plausible explanation could be that the content range of the Primary courses is considerably broader than the content of the Secondary courses and, thus, the Professional Development Programs for High School students were able to focus on their **tested** abilities. For example, according to M.A. Henry (2012), the Algebra I and Biology I tests have a specific content focus, unlike the broad range of science and mathematics questions on the MAP test. When target concepts are taught and tested, the chances for resulting significant change are increased, as with the High School tests. In other words, our results don't necessarily imply that the Programs were not effective to improve the knowledge of Primary students. Instead, there were ineffective to improve the specific knowledge measured by the standardized tests.

In any case, improving the academic performance of the neediest students is not an automatic process. It needs a push in the right direction. Scarce funds should be used in programs that have been proven effective. Therefore, the constant evaluation of the programs implemented with the Improving Teacher Quality Grants is fundamental in order to find the best and most cost-effective practices in Professional Development. This evaluation is the first step.

6. References

Armour-Thomas, E., Walker, E., Dixon-Roman, E., Mejia, B., & Gordon, E., “An evaluation of dynamic pedagogy on third-grade math achievement of ethnic minority children,” Unpublished Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA, 2006, April.

Balfanz, R., Mac Iver, D. and Byrnes, V., ”The implementation and impact of evidence based mathematics reforms in high poverty middle schools: Multi-school, multi-year studies,” Journal for Research in Mathematics Education, 37, 33–64, 2006.

Cohen, D., & Hill, H., “Instructional policy and classroom performance: The mathematics reform in California,” Teachers College Record, 102(2), 294–343, 2000.

Corcoran, T., “Teaching matters: How state and local policymakers can improve the quality of teachers and teaching,” CPRE Policy Brief No. 48. Philadelphia: CPRE. of North America, 2012.

Dee, T. and Jacob, B., "The impact of no Child Left Behind on student achievement," Journal of Policy Analysis and Management, John Wiley & Sons, Ltd., vol. 30(3), pages 418-446, Summer, 2011.

Fishman, B., Marx, R., Best. S. and Tal, R., “Linking teacher and student learning to improve professional development in systemic reform. Teaching and Teacher Education,” 19, 643–658, 2003.

Garet, M. S., Porter, A. C., Desimone, L. M., Birman, B. F. and Yoon, K. S, “What makes professional development effective? Results from a national sample of teachers,” American Educational Research Journal, 38(4), 915–945, 2001.

Guskey, T. and Sparks, D, “Linking professional development to improvements in student learning,“ in E. M. Guyton & J. R. Dangel (Eds.), Research linking teacher preparation and student performance: Teacher education yearbook XII. . Lanham, MD: Rowman & Littlefield Education, 2004.

Hawley, W. and Valli, L., “The essentials of effective professional development: A new consensus,“ in L. Darling-Hammond & G. Sykes (Eds.), Teaching as the learning profession: Handbook of policy and practice (pp. 127-150). San Francisco, CA: Jossey-Bass Publishers, 1999.

Heckman, J. and Hotz, J., “Choosing among alternative non-experimental methods for estimating the impact of social programs: The case of manpower training,“ American Statistical Association, 84(408): 862–880, 1989.

Heckman, J., Ichimura, H., Smith, J. and Todd, P., "Matching as an econometric valuation estimator: Evidence from evaluating a job training program," Review of Economic Studies, 64(4), 605-54, 1997.

Heckman, J., Ichimura, H., Smith, J. and Todd, P., "Characterizing selection bias using experimental data," Econometrica 66(5), 1017-98, 1998.

Heckman J. and Robb R. Jr, "Alternative Methods for Evaluating the Impact of Interventions: An Overview," *Journal of Econometrics*, *30*(1-2), pp. 239-67, 1985.

M.A. Henry Consulting, Technical Report, "Missouri Department of Higher Education-Improving Teacher Quality Grant Program Evaluation", 2012.

Kennedy, M, "Form and substance of in service teacher education," Madison, WI: National Institute for Science Education, University of Wisconsin–Madison, 1998.

Koretz, D., "Measuring Up: What Educational Testing Really Tells Us," Harvard University Press, 2008.

Loucks-Horsley, S. and Matsumoto, C., "Research on professional development for teachers of mathematics and science: The state of the scene," School Science and Mathematics, 99(5), 258–271, 1999.

Loucks-Horsley, S., Love, N., Stiles, K., Mundry, S., & Hewson. P., "Designing professional development for teachers of science and mathematics," Thousand Oaks, CA: Corwin Press, 2003.

Mant, J., Wilson, H. and Coates, D., "The effect of increasing conceptual challenge in primary science lessons on pupils' achievement and engagement," International Journal of Science Education, 29(14), 1707–1719, 2007.

Slavin, R. and Lake, C., "Effective programs in elementary mathematics: A best-evidence synthesis," Review of Educational Research, 78 (3), 427-515, 2008.

Slavin, R., Lake, C. and Groff, C., "Effective programs in middle and high school mathematics: A best-evidence synthesis," Review of Educational Research, 79 (2), 839-911002E, 2009.

Slavin, R., Lake, C., Hanley, P. and Thurston, A., "*Effective Programs for Elementary Science: A Best-Evidence Synthesis*, "Baltimore, MD: Johns Hopkins University, Center for Research and Reform in Education, 2012, May.

U.S. Department of Education, Office of Elementary and Secondary Education, "No Child Left Behind: A Desktop Reference," Washington, D.C., 2002.

U.S. Department of Education, Office of Elementary and Secondary Education, "Improving Teacher Quality State Grants, ESEA Title II, Part A, Revised," Washington D.C., 2006, Oct.

Rothstein, R., Jacobsen R. and Wilder, T., "Grading Education: Getting Accountability Right," Teachers College Press, New York, 2008.

7. Tables

Table 1 – Pre-Treatment Characteristics Balance (MAP Sample – Mathematics)

	No Matching (N=2653 teachers)			Matching (N=78 teachers)		
	Control	Treatment	Diff	Control	Treatment	Diff
Experience in the District (years)	10.01	6.25	-3.76**	7.23	6.95	-0.28
	(0.12)	(0.73)		(0.67)	(0.87)	
Experience in Missouri (years)	12.25	8.85	-3.4**	9.64	9.36	-0.28
	(0.14)	(0.92)		(0.879)	(0.97)	
Experience in Public School (years)	12.90	9.03	-3.87**	9.71	9.59	-0.12
	(0.14)	(0.943)		(0.88)	(0.99)	
Annual Salary ($)	45326.00	35852.00	-9474.00	39601.00	36995.00	-2606.00
	(245.48)	(938.14)		(995.48)	(1039.09)	

SD in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2 - Pre-Treatment Characteristics Balance (MAP Sample – Science)

	No Matching (N=2653 teachers)			Matching (N=259 teachers)		
	Control	Treatment	Diff	Control	Treatment	Diff
Experience in the District (years)	10.16	5.72	-4.44***	7.92	5.72	-2.20*
	(0.205)	(0.78)		0.290	(0.78)	
Experience in Missouri (years)	12.52	8.61	-3.91**	10.46	8.70	-1.76
	(0.22)	(1.18)		0.390	1.250	
Experience in Public School (years)	13.20	8.83	-4.37**	11.05	8.94	-2.11
	(0.23)	(1.2)		0.410	1.270	
Annual Salary ($)	44045.00	34915.00	-9130**	39885.00	35082.00	-4803**
	(371)	(1101)		(420)	(1154)	

SD in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3 - Pre-Treatment Characteristics Balance (EOC Sample – Mathematics)

	No Matching (N=43 Teachers)		
	Control	**Treatment**	**Diff**
Experience in the District (years)	11.50	7.39	-4.11
	(2.5)	(0.6)	
Experience in Missouri (years)	12.00	10.03	-1.97
	(3)	(0.84)	
Experience in Public School (years)	12.00	12.43	0.43
	(3)	(1.03)	
Annual Salary	42957.00	41593.00	-1364.00
	(1185)	(2268)	

SD in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4 - Pre-Treatment Characteristics Balance (EOC Sample – Science)

	No Matching (N=32 Teachers)		
	Control	**Treatment**	**Diff**
Experience in the District (years)	7.69	12.00	4.31
	(0.916)	(1.52)	
Experience in Missouri (years)	9.72	13.00	3.28
	(1.09)	(2)	
Experience in Public School (years)	12.03	13.00	0.97
	(1.4)	(2)	
Annual Salary ($)	43042.00	42097.00	-945.00
	(1602)	(1403)	

SD in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5 – Pre-Treatment Trends Balance (MAP and EOC Samples – Math & Science)

Students Scores	MAP Math	MAP Science	EOC Math	EOC Science
Temporal Trend * Potential Treatment (Interaction)	**1.4**	**-0.27**	**-3.9**	**3.3****
Temporal Trend (2009-2011)	Yes	Yes	Yes	Yes
Potential Treatment Dummy	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes
Model with Matching Sample	Yes	Yes	No	No
Mean of Control	695.36	699.87	208,06	205.55
N (Students)	20,931	54,713	3,205	2,487

SD in parentheses

*** p<0.01, ** p<0.05, * p<0.1

SD clustered at School-Grade-Year level

Table 6 – Results MAP Mathematics

Math MAP Scores	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment Effect	**-2.26**	**-3.6**	**2.02**	**5.02**	**-8.27**	**-13.8****	**-2.6**	**-2.92**	**0.41**	**0.145**
	(2.377)	**(2.28)**	**(7.21)**	**(9.22)**	**(5.67)**	**(6.75)**	**(3.55)**	**(3.93)**	**(3.03)**	**(3.29)**
						[-1.9%]				
School FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Grade FE	Yes	No	No	No	No	No	No	No	No	No
Grade 5			Yes	Yes						
Grade 6					Yes	Yes				
Grade 7							Yes	Yes		
Grade 8									Yes	Yes
ALL GRADES	Yes	Yes								
Mean of Control	701.08	701.08	655.03	655.03	692.92	692.92	695.06	695.06	713.06	713.06
N	14,557	14,577	328	328	2,779	2,779	5,885	5,885	5,565	5,565

SD in parentheses
*** p<0.01, ** p<0.05, * p<0.1
SD clustered at School-Grade-Year level for "all grades" and School-Year for the individual grades
Numbers in the row "Treatment Effect" are the estimations of the treatment effect. If the estimation is significant at least at 10% level of confidence, we include the variation in percentage terms of the average scores respect to the mean of the control group (number in brackets). Models can include School Fixed Effects ("Yes" in "School FE") and Grade Fixed Effects ("Yes" in "Grade FE"); each model is restricted to a particular grade or includes data of all the grades together ("ALL GRADES").

Table 7 – Results MAP Science

Science MAP Scores	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Effect	**-1.345**	**-0.641**	**2.160**	**3.816**	**-1.423**	**-1.562**
	(2.173)	**(2.004)**	**(5.276)**	**(6.819)**	**(2.395)**	**(2.338)**
School FE	Yes	No	Yes	No	Yes	No
Grade FE	Yes	No	No	No	No	No
Grade 5			Yes	Yes		
Grade 8					Yes	Yes
ALL GRADES	Yes	Yes				
Mean of Control	701.54	701.54	666.83	666.83	703.93	703.93
N	37,703	37,703	2,679	2,679	35,024	35,024

SD in parentheses
*** p<0.01, ** p<0.05, * p<0.1
SD clustered at School-Grade-Year level for "all grades" and School-Year for the individual grades
Numbers in the row "Treatment Effect" are the estimations of the treatment effect. If the estimation is significant at least at 10% level of confidence, we include the variation in percentage terms of the average scores respect to the mean of the control group (number in brackets). Models can include School Fixed Effects ("Yes" in "School FE") and Grade Fixed Effects ("Yes" in "Grade FE"); each model is restricted to a particular grade or includes data of all the grades together ("ALL GRADES").

Table 8 – Results EOC Mathematics

Math EOC Scores	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment Effect	**7.299***	**4.277***	**6.187*****	**7.188***	**4.205*****	**4.486***	**8.556*****	**8.787*****	**1.033**	**-3.022**
	(3.667)	**(2.472)**	**(0.485)**	**(3.461)**	**(1.290)**	**(2.500)**	**(2.026)**	**(1.941)**	**(1.183)**	**(3.197)**
	[3.6%]	**[2%]**	**[2.9%]**	**[3.4%]**	**[2.1%]**	**[2.3%]**	**[4.4%]**	**[4.7%]**		
School FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Grade FE	Yes	No	No	No	No	No	No	No	No	No
Grade 9			Yes	Yes						
Grade 10					Yes	Yes				
Grade 11							Yes	Yes		
Grade 12									Yes	Yes
ALL GRADES	Yes	Yes								
Mean of Control	210.99	210.99	212.34	212.34	196.99	196.99	189.69	189.69	185.61	185.61
N	3,464	3,464	2,230	2,230	860	860	227	227	147	147

SD in parentheses
*** p<0.01, ** p<0.05, * p<0.1
SD clustered at School-Grade-Year level for "all grades" and School-Year for the individual grades
Numbers in the row "Treatment Effect" are the estimations of the treatment effect. If the estimation is significant at least at 10% level of confidence, we include the variation in percentage terms of the average scores respect to the mean of the control group (number in brackets). Models can include School Fixed Effects ("Yes" in "School FE") and Grade Fixed Effects ("Yes" in "Grade FE"); each model is restricted to a particular grade or includes data of all the grades together ("ALL GRADES").

Table 9 – Results EOC Science

Science EOC Scores^	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Effect	**3.425*****	**3.318*****	**1.748*****	**1.531**	**4.236*****	**4.774*****
	(0.811)	**(0.855)**	**(0.344)**	**(1.602)**	**(0.156)**	**(1.104)**
	[1.2%]	**[1.3%]**	**[0.8%]**		**[2.1%]**	**[2.2%]**
School FE	Yes	No	Yes	No	Yes	No
Grade FE	Yes	No	No	No	No	No
Grade 10			Yes	Yes		
Grade 11					Yes	Yes
ALL GRADES	Yes	Yes				
Mean of Control	205.44	205.44	203.73	203.73	199.44	199.44
N	3,511	3,511	2,063	2,063	1,448	1,448

SD in parentheses
*** p<0.01, ** p<0.05, * p<0.1
SD clustered at School-Grade-Year level for "all grades" and School-Year for the individual grades
Numbers in the row "Treatment Effect" are the estimations of the treatment effect. If the estimation is significant at least at 10% level of confidence, we include the variation in percentage terms of the average scores respect to the mean of the control group (number in brackets). Models can include School Fixed Effects ("Yes" in "School FE") and Grade Fixed Effects ("Yes" in "Grade FE"); each model is restricted to a particular grade or includes data of all the grades together ("ALL GRADES").
^ As explained in the Methods Section, the results of EOC Science are not reliable, as we cannot guarantee that tha pre-treatment trends are similar between Treatment and Control

Appendix (Funded Programs of Cycle-9)

Cycle-9 Awarded Funding Projects

Lead Institution	Amount	Grade Level	Focus	Project Title
Columbia College	$ 205,570.73	K-6	Math & Science	Tomorrow's Hope for Renewal, Innovation, and Vision in Education (THRIVE)
Maryville University	$ 165,588.15	K-3	Science	Constructivist Early Childhood Science: Building Inquiring Minds
Missouri State University	$ 179,292.82	9-12	Math	Build and Connect Math Concepts Through In-Depth and Technology-Rich Explorations
Missouri University of Science & Technology	$ 222,153.60	5,6,7	Math & Science	Science Education and Quantitative Literacy: An Inquiry-based Approach
Southeast Missouri State University	$ 153,230.08	4-8	Math	Boosting Bootheel Mathematics
University of Central Missouri	$ 201,439.15	6-12	Math	Inquiry + Tech = Middle School Math Mastery
University of Missouri-Columbia	$ 206,867.91	K-6	Science	QUEST: Quality Elementary Science Teaching

Source: Missouri Department of Elementary and Secondary Education

As requested in the Cycle-9 RFP, all of the seven projects awarded funding are focused on Math, Science or Both. Approximately 72% of the schools were high-need public schools. As it can be seen in the Appendix Table and in the description of the programs, each winning project has its own particular objective and target population. What is more, not all the projects follow the same academic method, but all of them have the same overall objective: improve students' achievement in math and or science, trough the professional development of their teachers.

All the programs were based on workshops (or courses). In many cases, the workshops would strengthen the participants' knowledge, but also enhance their teaching strategies and improve the use of technology in their particular courses. In other cases, the courses are aimed at teaching how to link the science (or math) content to the real-world problems or environmental issues. Although the pedagogy specific strategy may vary in each project, the inquiry-based approach and the inclusion of technology is common among all the projects.

Project Summaries[17]

Project: Tomorrow's Hope for Renewal, Innovation, and Vision in Education (THRIVE)

The Tomorrow's Hope for Renewal, Innovation, and Vision in Education (THRIVE) Project will establish a collaborative partnership among Columbia College faculty and the Eldon R-I School District in Miller County. An additional valuable partner is the Big Muddy National Fish and Wildlife Refuge. The primary goal of the partnership is to create a dynamic Professional Learning Community (PLC) for elementary school educators (pre-service and in-service grades K-6) devoted to the use of environmental science as a conceptual anchor for integrative and innovative practices in science and mathematics education. Within the THRIVE PLC, K-6 teachers and administrators (n = 31) will participate in a minimum of 121 hours of professional development (PD) in environmental sciences with mathematics integration. Other participants include pre-service elementary teachers from Columbia College (n = 5/year) and science student "teaching cadets" from Eldon R-I High School (n=10/year)

Project Title: Constructivist Early Childhood Science: Building Inquiring Minds

Maryville University's School of Education and College of Arts and Sciences, in partnership with the St. Louis Public Schools, the Missouri Botanical Garden, and non-public school partners, seeks support to implement a two-year Improving Teacher Quality grant. The proposed project will build understanding, use, and assessment of constructivist science teaching strategies by early childhood teachers in the two district early childhood centers, grades 3-5 continuation school and selected nonpublic schools, and improve student learning outcomes in science.

Project activities will include:

• Hold yearly 56-hour summer institutes for K-4 teachers and administrators from the SLPS early childhood centers (Wilkinson, Stix, Humboldt) and selected nonpublic schools on implementing constructivist practices in urban classrooms, building science content knowledge, and developing hands-on minds-on environmental education teaching. Content will focus on integrated environmental science concepts, science inquiry skills, constructivist approaches to inquiry learning, identifying misconceptions in science, performance assessment, and integrating literacy and mathematics in science. Maryville education and science faculty will serve as institute faculty with assistance from consultants from the Missouri Botanical Garden.

[17] All the descriptions can be found in the site of Missouri Department of Higher Education (http://www.dhe.mo.gov/ppc/grants/ITQGCycle-9.php)

• Provide 64 hours yearly of follow-up ongoing instruction along with pedagogical coaching and lesson analysis for teachers to help them design and integrate constructivist inquiry based science instruction. Maryville faculty and district curriculum supervisors will serve as faculty and coaches and involve school administrators in implementing constructivist science strategies.

• Provide coaching and assistance for participating teachers in the design and implementation of performance-based assessment strategies and the use of data teams.

• Provide participating teachers with hands-on environmental education learning resources designed to support inquiry-based instruction

• Infuse Maryville University teachers-in-training into Stix, Wilkinson, and Humboldt classrooms, focused on integrated approaches to science curriculum and instruction. Incorporate additional research-based teaching strategies on addressing the needs of culturally diverse students into science methods courses for pre-service teachers

Project Title: Build and Connect Math Concepts Through In-Depth and Technology-Rich Explorations

This 2-year project will be a continuation of an effort to bring together high school mathematics teachers from rural high needs school districts in western Southwest Missouri, to continue to study the use of in-depth and technology-rich explorations for improved math learning. The project will build on work done in year 1 with participants to embrace inquiry-learning as an effective math pedagogy strategy. Staff from Missouri State University and the Southwest Center for Educational Excellence, rural teachers, administrators, and members of business and industry will work together to develop a comprehensive 1-week immersion summer institute with follow-up workshops and observations during each academic year to address increasing needs of these math teachers at the secondary level. Getting quality professional development into these rural schools will also help their staff reach the ultimate goal of improved student achievement. Year 1 of this project continuation focused on geometry and measurement. Year 2 will focus on probability and statistics. Connections to other math topics will also be incorporated to serve the needs of partner schools using integrated curriculums. Paramount to this project is a continued effort to move math instruction from an emphasis on procedural fluency to a more inclusive design that interweaves conceptual and procedural understanding, the ability to formulate, represent and solve mathematical problems, and the capacity for logical thought, reflection, explanation, justification, and sense making. Problem-based instruction, use of current technology for student exploration (software, interactive computer sites, calculators, data-collection probes), and development of curriculum and assessment will be employed in the project to encourage this movement. Beliefs about how students learn, effective feedback, and student engagements strategies will also be examined over the course of the project.

Project Title: Science Education and Quantitative Literacy: An Integrated, InquiryBased Approach

This three-year project will provide professional development in inquiry-based methods to 40 mathematics and science teachers from the south-central region of Missouri, with 75% from high-needs schools. An integrated approach will be employed that highlights the synergy between science experiments and use of mathematics to understand experimental data. Participants will conduct hands-on activities to explore scientific and mathematical concepts and develop their own problem solving strategies so that their students will apply mathematical concepts when learning science and use scientific data in learning mathematics. A learning cycle approach based on the 5E instructional model will form the foundation of these instructional activities. The same set of teachers will attend all three years of the project, allowing for an in-depth and sustained professional development experience. Any vacancies will be filled by teachers from the same school, thus maintaining an uninterrupted partnership with the participating schools. An environmental theme focused on renewable energy and conservation will form an integral part of the project. Missouri Department of Conservation, a project partner, will collaborate in developing conservation related activities. Presentations by Missouri S&T solar car and solar house design teams will provide a segue to renewable energy themed learning modules. Mathematics and science GLEs identified as problematic areas for students in the participating schools via recent MAP performance data will form the core of the first year program. The target GLEs will be revised on a yearly basis based on new student performance data as well as feedback and assessment data from participants. Instructional materials will be developed through a cooperative effort between the Director of the S&T Teacher Education Program, faculty members from S&T's Biology, Mathematics, and Physics Departments, education consultants from the Missouri Department of Conservation, school coordinators from participating schools, and master teachers from our previous ITQG projects.

Project Title: Boosting Bootheel Mathematics

"Boosting Bootheel Mathematics" will serve the following Missouri Bootheel high-need school districts: Bernie R-XIII, Gideon 37, Malden R-1, Risco R-II, Scott County Central, and Lift for Life Academy a High-Need Charter School supported by Southeast Missouri State University. South Pemiscot County R-V will also be included in this project. St. Teresa School will be the project's sole private school partner. The need for math teacher professional development is supported by:

• MAP Data (2006-2008) revealing these schools experienced a downward trend in mathematics proficiency (grades 4 to 8)

• School leaders wanting to implement intervention strategies to improve both student proficiency and motivation.

• Needs assessment surveys indicating a desire for more professional development.

A cohort of grades 4-8 math teachers will participate in a five-day summer institute that will focus on strategies to boost student engagement/motivation, which in turn will increase student achievement. These strategies will include using an inquiry-based approach, problem solving strategies, assessment tools for both formative and summative assessments, data analysis for the purpose of driving instruction, and a deeper level of teachers' content knowledge and pedagogical strategies. All strategies will be linked to Grade Level Expectations, Show Me Standards, and the NCTM Process and Content Standards. Participants will develop math lessons that integrate literacy and writing skills. The summer institute will be followed by:

• Six follow-up workshops during the academic year to collaborate and reflect on practice.

• Project instructor visits to school districts to coach and assist participants in implementing effective teaching practices. Project instructors will also model lessons and effective teaching practices for pre-service teachers. This project will deliberately pursue administrator participation by offering an incentive to the schools for follow-up activities and/or purchase of materials to provide additional math resources.

Project Title: Inquiry + Tech = Middle School Math Mastery

This project is a professional development program at the University of Central Missouri (UCM) in collaboration with the school districts in Bates, Benton, Henry, Hickory, Jackson, Johnson, Lafayette, Pettis, and St. Clair that will positively impact middle school mathematics instruction in these counties (80% of the school districts in these counties are considered to be high need). The program will consist of a 90-hour workshop in July-August 2011 (2012) with 30 middle school mathematics teachers from the above counties. Five follow-up sessions (six hours per session) will be conducted for the workshop participants during the academic years 2011-2012 (2012-2013). A total of 130 teachers will be trained to use technology and effectively teach mathematics with the inquiry-based learning by the end of the program. UCM will partner with these school districts by forming peer-support groups with the participants during the July-August workshops, sustaining the project throughout the year through the follow-up sessions, maintaining continuous communication through an online discussion forum, and publishing a newsletter semi-annually. The July-August workshops will strengthen the participants' mathematics knowledge, as well as enhance their teaching strategies through inquiry-based, cooperative learning, integration of writing into mathematics, and the use of technology. It has been shown that inquiry-based learning is a very effective method to enhance the student learning. Various activities aligned with the Missouri Grade Level Expectations are designed to show how inquiry-based learning complemented with technology can work together to increase students' performance in mathematics. Five follow-up sessions are planned for additional support and enrichment activities for the teachers to re-enforce and assess what they have learned in the summer workshops. Pre-service mathematics teachers at UCM will be invited to attend the follow-up sessions so that they can be better prepared for their teaching career in mathematics.

Project Title: QUEST: Quality Elementary Science Teaching

This 3-year project will provide professional development in inquiry-based teaching to 30 K-6 teachers and special educators, as well as 6 pre-service teachers annually. The program promotes success for all learners in science through applying principles of Universal Design for Learning (UDL) to instruction. Building on our experience from a previous ITQ project, we will utilize an innovative format that combines teacher professional development with a summer academy for students. Participating teachers attend a 2-week full-time summer institute at the University of Missouri. During week one teachers develop their understanding of physical science content and real-world applications of physics concepts to environmental and social issues. Specific topics are selected to address teachers' greatest area of need in terms of subject matter knowledge, and GLEs they report most difficult for students to master. In week two, participants refine their instruction and assessment skills and are guided in implementing the knowledge and pedagogical skills they develop in the program and applying UDL to teach the content to elementary students. 6 of the participants will serve as 'mentors' during this week. Mentor participants will have received additional training on data team processes, so that they can guide teachers in using assessment data to inform instruction. Following the summer institute, teachers continue their professional development through (1) implementation of administrator-supported action plans targeting improved student achievement in science, (2) individualized support visits by project staff, (3) a series of face-to-face follow-up sessions, and (4) an interactive project website that enables teachers to share instructional materials they develop and discuss their implementation of these materials with colleagues in other buildings and school districts. This project represents a strong collaboration between the Department of Physics & Astronomy, the MU Science Education Center, the Department of Special Education, Heart of Missouri Regional Professional Development Center, and partner schools.