WPS4167

IMPACT EVALUATION SERIES NO. 13

More Time Is Better:

An Evaluation of the Full-Time School Program in Uruguay

Pedro Cerdan-Infantes and Christel Vermeersch*

Abstract

This paper estimates the impact of the full-time school program in Uruguay on standardized test scores of 6th grade students. The program lengthened the school day from a half day to a full day, and provided additional inputs to schools to make this possible, such as additional teachers and construction of classrooms. The program was not randomly placed, but targeted poor urban schools. Using propensity score matching, we construct a comparable group of schools, and show that students in very disadvantaged schools improved in their test scores by 0.07 of a standard deviation per year of participation in the full-time program in mathematics, and 0.04 in language. While the program is expensive, it may, if well targeted, help address inequalities in education in Uruguay, at an increase in cost per student not larger than the current deficit in spending between Uruguay and the rest of the region.

World Bank Policy Research Working Paper 4167, March 2007

*Human Development Department, Latin America and the Caribbean Region; Email: pcerdaninfantes@worldbank.org and cvermeersch@worldbank.org. We are thankful to Emiliana Vegas for her comments and guidance. The paper benefitted from the comments of Andres Peri and seminar participants in Montevideo and Washington, DC.

1 Introduction

The length of the school day varies greatly among countries, as does the number of school days per academic year. Even though the cross-country correlation between hours spent in school and achievement is not clearly significant, many countries, including the United States, Chile or Colombia in Latin America, have chosen to lengthen the school day as a way of improving student learning. With a longer school day, it is expected that students learn more because they spend more time with teachers and devote more time to school tasks. In addition, after-school programs in the United States have been used to prevent at-risk students from engaging in harmful activities, especially in high school.

Whether or not more time in school results in better learning outcomes depends broadly on two factors: what happens in the school in the extra time, and what would the beneficiaries do were they not in school during the extra time. Since these two factors depend greatly on the characteristics of the existing education system, the implementation of the program and the beneficiaries, the effects of these programs will also depend on these characteristics. As a consequence, cross-country regressions will shed little light on the issue and careful impact evaluations would be better suited for estimating these effects. Unfortunately, the existing impact evaluations are scarce and their results, though generally favorable, are not robust. A general concern with the existing literature is the identification issues arising from the non-random assignment of students to schools or classes with longer school days. For example, Walston and West (2004) compare students who attended kindergarten full-time to students who attended part-time, within the same school, and show that full-time students had significantly higher test scores in both math and reading than part-time students. However, the self-selection of children into full-time and part-time kindergarten makes it difficult to attribute the differences in achievement to differences in kindergarten attendance. The authors cannot rule out that these differences in achievement are driven by inherent differences in the students that cannot be observed. In Latin America, two recent papers, Valenzuela (2005) and Bellei (2005), use exogenous variation in the expansion of the full-time school program in Chile to estimate the impact of longer school days on student achievement and show positive though small effects, with larger effects in language than in mathematics.

The Uruguay Full-Time School (FTS) Program is a prominent recent case of extension of school time in a middle-income country. The program lengthened the school day from a half day to a full day, and provided additional inputs to schools to make this possible, such as materials, teachers, teacher training, and construction or rehabilitation of classrooms. The program was a relatively intense intervention, targeting mid-sized, disadvantaged urban schools, and has been in place since the early 1990s. Since then, the Ministry of Education has commissioned two evaluations of the program. The first evaluation, a qualitative study of the performance and inner workings of Full Time Schools, was carried out by Equipos/Mori for MECAEP (Spanish acronym for Program for the Improvement of Quality in Primary Education, financed by the World Bank) in 2001 (Estudio de la Evaluación Social de las Escuelas de Tiempo Completo) and showed that, while opinions about the program were generally positive, implementation was incomplete at the time of the study, especially regarding the provision of resources to complement the extension of the school day. This study, however, did not attempt to measure the impact of the program on learning outcomes. The second evaluation was carried out by the Administración Nacional de Educación Pública (ANEP) and showed that the average test score in full-time schools was higher than non-full time schools with similiarly unfavorable socioeconomic characteristics, using the National Assessment of Student Learning (Evaluación Nacional de Aprendizajes), a nation-wide standardized test. While the restriction of the comparison group to schools in similar contexts does control for some socioeconomic characteristics, the analysis did not use the richness of data available to minimize biases from unobservable differences between schools that did and did not participate in the program.

In this paper, we evaluate the impact of the FTS program on student oucomes, trying to shed more light on the effects of the program using student level data and controlling for characteristics of the school that may be driving these differences. The paper is organized as follows: Section 2 describes the main characteristics of the Full-Time School program in Uruguay since its inception, emphasizing implementation issues that may affect our estimation. Section 3 describes the data, Section 4 explains the estimation strategy, Section 5 presents the main results, and Section 6 shows some cost-benefit considerations. Finally, Section 7 concludes.

2 Program Description[1]

The FTS program in Uruguay was introduced in the early 1990s as a means to increase student achievement in disadvantaged schools. The conversion of schools to the FTS was effectively done in three phases, in which both the content of the program and the characteristics of the beneficiary schools varied notably: before 1996, 1997 to 1998 and 1999 to 2005. The 52 schools that converted to FTS before 1996 served a very disadvantaged part of the population; 25 of them were "open-air" schools (that is, schools with no buildings or with buildings in very bad condition).

[1] **Sections 2 and 3 draw on three documents: Evaluacion Nacional de Aprendizajes, Resultados en Escuelas de Tiempo Completo y Areas Integradas, and the Propuesta Metodologica para Escuelas de Tiempo Completo. These documents and more information are available MECAEP's website (www.mecaep.edu.uy).**

Between 1996 and 1999, the program only incorporated 4 new schools, while the Ministry was working on a new pedagogical model for the full-time schools. From 1999 onwards, the National Public Education Administration (ANEP) implemented this new pedagogical plan and 48 more schools were converted to FTS.

The new pedagogical model came amid concern about the effectiveness of public education. In 1996, ANEP applied the National Census of Student Learning, which showed a number of alarming trends. First, overall student achievement was low: 43 percent of sixth grade pupils did not reach the "sufficient" score in language, and 65 percent failed to do so in math. Second, pupil performance was highly unequal and strongly correlated with the socioeconomic characteristics of the parents and the school. This evidence, combined with qualitative indications that the existing full-time schools had performed better than schools with similar characteristics, prompted the government to turn to the FTS program for solutions to low achievement in disadvantaged schools. The plans for expansion resulted in a special and comprehensive pedagogical plan (Propuesta Pedagógica), which was implemented in both existing and new full-time schools starting in 1999, and was supported by the World Bank through MECAEP.

The FTS program as implemented starting in 1996 encompassed not only an expansion of the school day, but also a combination of other interventions targeting students and teachers, additional classrooms and equipment, and community participation. The program includes the following interventions:

1. Construction of new classrooms;
2. A reduction of the recommended number of pupils per classroom to 25 in grades 1-3, and 28 in grades 4 to 6.
3. Lengthening of class time from 3.5 to 7 hours per day, 5 days per week, and an additional of 3 hours per week of complementary attention to students with special needs and/or community service activities, and 2 hours of teacher meetings;
4. Introduction of collective, complementary and classroom activities;
5. Constitution of a teacher committee;
6. Provision of nutritional and health care support for students;
7. Increased participation of parents, accountability/ community involvement;
8. From 1999 onwards: Teacher training in the FTS pedagogical model ("Curso I") and follow-on subject related courses ("Curso II", math, Spanish, science);
9. Provision of a set of teaching materials, such as maps, books or dictionaries.

Even though a qualitative evaluation of program implementation in 2001 indicated that provision of the different sub-components was unequal across participating schools, the implementation of the extension of the school day, school building and teacher training was practically universal in participating schools. Other components, like additional teacher training and increased community participation were voluntary. These two characteristics of the sub-components of the program will be relevant when trying to assess their impacts, as they prevent the robust causal identification of the effects of separate interventions.

3 Data

In the analysis, we use three sources of data: the National Evaluation of Learning Achievements for the Sixth Grade of Primary School, the National Census of Schools and specific program information gathered from ANEP.

3.1 National Evaluation of Learning Achievements for the Sixth Grade

The National Evaluation of Learning Achievements for the Sixth Grade of Primary School[2] (hereafter "tests") was carried out in 1996, 1999 and 2002 in the context of the Program for the Improvement of Quality in Primary Education (MECAEP hereafter), which was financed by the World Bank. It consisted of Math and Spanish assessments, and it was supposedly applied to a census of primary schools in 1996 and nationally representative samples in 1999, 2002 and 2005 (the 2005 data are not yet available). The 1996, 1999 and 2002 tests are directly comparable between years. The 2005 used item-response theory for the fist time, which may limit the comparison in the future. While the 1996 test was meant to be a census, it excluded very small schools with fewer than 6 students in 6th grade. As a result, about 1,200 schools (approximately half of the total number of schools) were not included in the sample, but they likely represent a minority of students. (Cf. Table 1) The achievement tests consist of two sets of 24 questions each in mathematics and language. The mathematics test comprises 7 questions on comprehension of mathematical concepts, 12 questions that involve the resolution of a problem, and 5 questions involving an algorithm. The language test comprises 8 questions about language, 8 argumentative questions and 8 narrative questions. All questions are multiple choice with 4 possible options. The overall score is the number of correctly answered questions out of a possible 24. There is no penalty for giving the wrong answer, so that a child who guesses on every single question of the test is expected to obtain 6 points out of 24 on average.

Information about FTS was gathered both from administrative data and data collected

[2] Evaluación Nacional de Aprendizajes, 6to año Enseñanza Primaria, Gerencia de Investigación y Evaluación ANEP-CODICEN

Table 1: National Evaluation of Learning Achievements for Sixth Grade: Sampling of Schools

Year of Conversion	Number of Public Primary Schools (excluding special schools)				
to FTS	Total	Tested in 1996	Tested in 1999	Tested in 2002	Never Tested
1996 or before	52	49	39	52	0
1997	2	2	2	2	0
1998	2	2	2	2	0
1999	10	8	8	10	0
2000	7	3	1	7	0
2001	15	12	2	15	0
2002	3	1	0	2	1
2003	5	4	0	1	1
2004	8	7	0	2	1
Never a FTS	2382	1196	189	205	1162
Total	2486	1284	243	298	1165

from ANEP. Table 1 shows the sample of schools tested in each year broken down by year of conversion to the FTS program (including those not converted). It was ANEP's policy to apply the achievement tests in all full-time schools. However, schools that converted to full time status after the year of the test were only tested if the random selection of a nationally representative sample included those schools. For example, there were 24 schools that converted between 1999 and 2002, but only 2 of them were included in the sample of the 1999 achievement tests. For this reason, it is not possible to construct a panel of schools for 1996, 1999 and 2002. Instead, we use the 1996 test data as baseline data, and use the 2002 data as follow-up testing. Since very few schools converted between 1996 and 1999, the 1999 data are of little use for evaluating the impact of the full-time schools program.

For the purpose of the analysis, we restrict the sample of schools in the following way: we exclude those schools that were not tested in 1996 (i.e. the schools for which we do not have a baseline), the schools that converted to the full-time model by 1996, and the schools that did not participate in either of the post-baseline measurement in 2002. These restrictions lead to the sample presented in Table 2. There are 28 schools that converted between 1997 and 2002, and which were tested both in 1996 and 2002. In addition, there are 190 schools that were also tested in 2002, but that either did not convert to full time status or converted after 2002.

The achievement tests were complemented with questionnaires addressed to each child, the head of household of each child, the class teacher and the school director. While the additional questionnaires are slightly different in the three years, it is possible to extract a common set of questions, which cover: (1) in the pupil questionnaire: classroom activities and motivation for

Table 2: Test sample of FTS

Year of conversion to FTS	Tested in 1996 and 2002
1997	2
1998	2
1999	8
2000	3
2001	12
2002	1
2003	1
2004	2
Never a FTS	187
Total	218

school and learning; (2) in the head of household questionnaire: household demographics, family assets, parental education and occupation, preschool attendance; (3) in the teacher questionnaire: training, experience, opinions on pedagogy, and interactions with parents; (4) in the director's questionnaire: training and experience of the school director, school infrastructure and equipment, school problems, interactions with parents, director's opinion on pedagogy.

3.2 National Census of Schools

The National Census of Schools[3] is a registry that contains yearly information on all public school in Uruguay, from 1996 to 2004. The registry contains data on pupil enrollment, repetition, grade promotion, repetition, drop-outs and attendance at each grade level. We use the 1996 data as controls for the initial conditions at the school. As these variables are potentially endogenous to participation in the full-time model, we use only the 1996 data. The Census of Schools also contains a classification of the sociocultural context of the school on a scale of 1 through 5, which is available for 1996 and 2002. In 1996, the classification was computed on the basis of data from the additional questionnaires for the National Evaluation of Learning Achievements for Sixth Grade. (Cf. 3.1) In 2002, it was computed on the basis of data reported by the school director in the newly introduced Educational Monitor instrument[4].

[3] Hiperbase Primaria, Gerencia de Investigación y Evaluación
[4] Monitor Educativo, Gerencia de Investigación y Evaluacion ANEP-CODICEN

3.3 Specific Program Information

In addition to administrative data, we gathered information from the Ministry on implementation and characteristics of each school participating. The dataset includes information for each full-time school on the year of conversion to FTS (or creation for new schools), the department and school number, as well as school characteristics and the main reason why the school was chosen for conversion. This information was complemented with informal conversations with ANEP's staff involved in the implementation of the program.

3.4 Summary Statistics

As explained above, the full-time school program targeted disadvantaged schools, which implies that full-time and non-full time schools had different characteristics at baseline. Table 3 presents summary statistics for key variables in 1996, broken down for schools that did not convert to FTS (columns 1-3), and schools that converted to FTS after 1996 (columns 4-6). The variables include test scores, teacher experience, school characteristics and household characteristics including size and parental education. Indices were constructed for "school equipment" and "buildings" using factor components analysis.[5] Column 7 presents t-statistics for the null hypotheses that the mean of the variables for the full-time schools is the same as for non-full time schools.

The two groups of schools differ on a number of different dimensions. First, full-time schools scored significantly lower than other schools on the 1996 achievement tests, and the difference is substantial (about 0.3 of a standard deviation). Second, full-time schools have significantly worse socio-economic indicators than non-full time schools, including parental education and household size. Third, there are notable differences at the school level: full-time schools had worse infrastructure, but better school equipment and smaller class sizes at baseline. Teacher experience does not seem to differ significantly. These differences were expected given that the program was targeted towards worse-off schools. These differences persist if we restrict the sample to schools that were tested both in 1996 and 2002.

[5] The underlying variables were extracted from the 1996 director's questionnaire, and include the availability of teaching materials (television, video, overhead projector, music) for the factor "equipment" and auxiliary buildings (library, lab, audio-visual room and computer room) for the factor "buildings". Finally, the variables household size, household number of rooms and parental education try to capture the socioeconomic characteristics of the household and were constructed directly from parental questionnaires.

Table 3: Summary Statistics, all tested schools, 1996

Variable description	Non-FTS Obs	Non-FTS Mean	Non-FTS St.Dev.	FTS Obs	FTS Mean	FTS St.Dev.	t-stat
Score in language, out of 24	42616	14.41	4.72	984	12.45	4.60	-4.25
Score in math, out of 24	43854	11.97	4.60	999	10.49	3.86	-3.91
Student is female =1	44633	0.50	0.50	1022	0.50	0.50	-0.15
Student repeated at least one grade =1	43667	0.29	0.45	999	0.42	0.49	6.06
Student repeated two or more grades =1	43667	0.08	0.28	999	0.14	0.35	3.51
Student's nr of years of preschool	41558	1.31	1.15	933	1.55	0.93	3.90
Household size	43847	5.16	1.89	1003	5.64	2.36	3.73
Household's number of rooms	43827	3.77	1.55	1004	3.46	1.37	-3.63
Mother has no education =1	41762	0.14	0.35	960	0.26	0.44	5.24
Mother has at least primary education =1	41762	0.86	0.35	960	0.74	0.44	-5.24
Mother has done at least some secondary =1	41762	0.62	0.48	960	0.43	0.50	-6.25
Mother has at least finished secondary =1	41762	0.30	0.46	960	0.14	0.35	-7.93
Mother has completed tertiary education =1	41762	0.19	0.39	960	0.07	0.26	-9.08
Father has no education =1	38892	0.17	0.37	895	0.29	0.46	4.87
Father has at least primary education =1	38892	0.83	0.37	895	0.71	0.46	-4.87
Father has done at least some secondary =1	38892	0.59	0.49	895	0.40	0.49	-5.89
Father has at least finished secondary =1	38892	0.29	0.45	895	0.14	0.35	-7.49
Father has completed tertiary education =1	38892	0.17	0.38	895	0.08	0.27	-7.06
Teacher has only one year of experience	1753	0.02	0.15	47	0.06	0.25	1.17
Teacher has 2-3 years of experience	1753	0.05	0.22	47	0.04	0.20	-0.22
Teacher has 4-5 years of experience	1753	0.05	0.22	47	0.17	0.38	1.80
Teacher has 6-9 years of experience	1753	0.15	0.35	47	0.13	0.34	-0.37
Teacher has at least 10 years of experience	1753	0.73	0.44	47	0.60	0.50	-1.82
Average Class Size	1196	22.89	10.17	39	20.61	6.98	-1.39
School equipement (PC factor 1)	1196	0.19	0.83	39	0.10	0.71	-0.65
School auxiliary buildings (PC factor 1)	1196	-0.26	0.67	39	-0.47	0.26	-1.89
School socio-economic quintile	1196	3.54	1.43	39	4.44	0.82	3.88

Notes: The statistics are based on the test scores and additional information that were collected at the time of the 1996 achievement tests. The non-full-time schools (columns 1-3) are those schools which were tested in 1996 and who had not converted to full-time status as of 2002. The full-time schools (colums 4-6) are the schools that were tested in both 1996 and 2002, and had converted to full-time status as of 2002. The potential comparison schools (columns 8-10) are those schools that were tested in both 1996 and 2002, and had not converted to full-time status as of 2002. t-statistics in column (7) are for a test of the null hypothesis that full-time and non-full-time have identical values. t-statistics in column (11) are for a test of the null hypothesis that full-time and comparison schools have identical values. The school socio-economic quintile was calculated by ANEP in basis of the 1996 data, and used for assigning schools to the full-time program. Schools in lower quintiles have better socio-economic environments. The t-tests at the student level and at the teacher level are clustered at the school level. * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

4 Estimation Strategy

The goal of this analysis is to identify the effect of being converted to a FTS on student achievement. In particular, and since exposure to the program varies notably between 0 and 6 years, we are interested in knowing the change in test scores, if any, associated with each year of exposure to the FTS model. A first approach to identifying this effect would be to compare the test scores of students in schools that were converted to FTS with those that were never converted:

$$Y_{ijt} = \alpha + X_{ijt} \cdot \beta + W_{jt} \cdot \gamma + E_{jt} \cdot \delta + \varepsilon_{ijt} \quad (1)$$

where Y_{ijt} is the test score of student i in school j at time t, X_{ijt} is a set of student-level controls, E_{jt} is the length of exposure to the full-time school model in school j as of time t (in years), and ε_{ijt} is a stochastic error term. In this specification, δ would provide an unbiased estimate of the effect of one additional year of exposure to FTS only if E_{jt} is not correlated with the error term, that is, only if there are no omitted variables that are correlated with both the exposure variable and the dependent variable. Since the selection of schools was non-random, and in fact targeted to schools with students from disadvantaged backgrounds, it is likely that the δ coefficient in 1 would be biased. Since more disadvantaged schools were targeted for participating in the program, school characteristics, both observable and unobservable, are surely correlated with program participation. In addition, these characteristics are also likely to affect test scores, and therefore their omission from equation 1 results in a biased estimate of program effect. This estimation would therefore result in negative bias, or an underestimation of the program effect.

Controlling for omitted variables in this model is challenging. As we saw in section 2, schools that were selected for conversion to full-time schools had to meet certain requirements like size, having one shift and having physical space to build additional classrooms or rehabilitate them. While these characteristics were necessary for conversion, they were not sufficient: there are many schools that met those requirements that were not converted. Including these variables in the regression will not eliminate all the bias since the selection among schools that met the requirements of the program, used additional criteria that are not easily measureable. The specification we use attempts to minimize the potential bias introduced by the selection of schools based on unobservable school characteristics. We use school fixed effects and different propensity score matching methodologies to control for time-invariant observable and unobservable school characteristics, and minimize the differences in the distribution of participating schools and their comparisons. In addition, the targeting of the program does allow some refining of the sample. Using information that was collected with the 1996 test, ANEP constructed a socio-economic index, and excluded schools with belonged to quintiles 1 and 2 (the most advantaged ones) from participation in the full-time program. This rule was adhered to in all cases, which means that the 66 schools from quintiles 1 and 2 can be excluded from the sample.

4.1 Fixed Effects Model

In the fixed effects model, we control for time invariant observable and unobservable school characteristics using school-level fixed effects. We include also time dummies, to control for unobservable changes common to all the observations in a particular year, student characteristics, and teacher characteristics. The effect of the full-time program is captured by the parameter of a variable that measures the number of years a school was exposed to the program before the 2002 follow-up achievement test. Since the sample excludes schools that were already full time at baseline in 1996, the highest number of years of exposure to the program is six. Since primary school has six years, and the achievement tests were taken in sixth grade, this is also the maximum number of grades that any child would have been exposed to the full-time model. The only potential bias left in this specification comes from time-variant school characteristics. The fixed effects model specification is as follows:

$$Y_{ijt} = \alpha + X_{ijt} \cdot \beta + W_{jt} \cdot \gamma + E_{jt} \cdot \delta + \sum_{k=1}^{J} D_{jk} \cdot \xi_k + T_t \cdot \tau + \varepsilon_{ijt} \tag{2}$$

where all variables from 1 are included, D_{jk} is a categorical variable that takes value 1 if $j = k$, and 0 otherwise, T_t is a categorical variable that takes value 1 if $t = 2002$ and 0 otherwise, and ε_{ijt} is a stochastic error term. t can take values 0 (for 1996) or 1 (for 2002). The length of exposure is calculated as follows:

$$E_{jt} = Year\ of\ test \text{ - } Year\ of\ conversion + 1 \tag{3}$$

4.2 Propensity Score Matching Model

Finding a comparison group for participating schools in the absence of a clear targeting mechanism is a challenge. As mentioned above, schools that converted to the full time system were quite different from other schools at baseline. Even controlling for observable characteristics separately in the regression, the distribution of the two groups will be different, and the results from a full-sample fixed effects model may be biased if there are differential trends between the full-time schools and the other schools which are not related to the full-time program. To address this potential bias, we construct control groups using different matching methods in an attempt to restrict the comparison group to schools that are as similar as possible to the full-time schools. Ideally, we would try to replicate the targeting mechanism of the program in the matching model, in order to choose schools with the same characteristics to those participating in the program (i.e. we would include the variables used in the targeting to find suitable control schools). In the absence of explicit rules for the assignment of schools to the program, we use the observable characteristics of the program schools to identify those comparison schools that are similar. In particular, we will use a propensity score model to identify schools that have a similar range of probability of participating in the full-time model, given their observable characteristics. We use the following model of participation:

$$\Pr(P_j = 1) = \Phi(\kappa + W_{j,1996} \cdot \lambda) \tag{4}$$

where P_j is a dichotomous variable that takes value 1 if school j participates in the full-time model, and $W_{j,1996}$ is a vector of characteristics of school j in 1996. This vector includes the key variables described in the summary statistics above plus "director's motivation" and the "number of shifts in the school". The latter was an explicit rule included in the election of the schools, while the former is an attempt to control for the unobservable "interest on the part of the school", which was mentioned by ANEP to be a factor in the decision to transfer schools.[6]

Using this model of participation, we estimate the propensity score - i.e. the probability of participating in the program - for both full-time and non-full-time schools. We then use the predicted probability to form three alternative comparison groups. For each of the proposed comparison groups, we fit a population-average model similar to the fixed effects model, using probability weights. The first comparison group is the set of the "closest neighbor" to each program school: for each program school, we identify the non-participating school with the closest propensity score. The second comparison group is the set of the "closest 5 neighbors" to each program school, by identifying for each program school the 5 non-participant schools with closest propensity scores. Note that in both methods, a non-program school can potentially be the closest to several program schools. We will take this into account in the second stage of the estimation by introducing probability weights on the observations.

Finally, we form a third comparison group by identifying those schools for which we can predict participation in the program with a probability above a cutoff point. This is motivated by the following observation: using the propensity score model, we find that a number of program schools have zero predicted probability of participating in the program. This means that

[6] Note that the results from the matching are robust to the inclusion of different combinations of these variables, as long as the variables for socio-economic characteristics of the students are included.

their baseline characteristics are very different from the baseline characteristics of other program schools. In particular, these schools were better off in socio-economic terms at baseline than their fellow program schools, in addition to being better equipped. It thus seems that some schools were included in the program even though they did not have disadvantaged socio-economic characteristics. Conversations with ANEP officials confirmed that certain program rules related to geographic allocation resulted in conversion of a number of better-off schools. For the purpose of the analysis, it is useful to exclude those "outliers" and re-calculate the effect of the program among the real target group of the program. In practice, we selected all schools that had a predicted probability of participation above .15, and compared outcomes between program and non-program schools within that restricted sample. By applying the same cutoff point to both program and comparison schools, we thus exclude schools with a-typical FTS characteristics.

5 Results

5.1 Full Sample Fixed Effects Estimates

Table 4 presents the results from the fixed-effects model on the full sample of schools, and the full sample of schools in the third, fourth and fifth socio-economic context quintiles (Equation 2). The dependent variable is the score on the Mathematics or Spanish language test, which ranges from 0 to 24, and the model includes parental, class and school level controls, in addition to school fixed effects and a year dummy. Columns 1 and 3 present the results for Spanish language, and columns 2 and 4 present the results for Mathematics. In columns 3 and 4, we restrict the estimation to those schools that belonged to socio-economic quintiles 3, 4 and 5, as none of the schools in quintiles 1 and 2 ever participated in the program. The results are positive for Mathematics, but are not statistically significantly different from 0 in Language. The full-sample estimates, which are subject to the largest potential bias, show that one year of exposure to the full-time program is associated with an increase of 0.21 points in Mathematics. The program is also associated with 0.12 increase in Language test scores, though the coefficient is only marginally significant. Restricting the sample to those schools in the third, fourth and fifth quintiles reduces the coefficient in Math (to 0.16), and the Language coefficient is not statististically significant. The size of the coefficient is relatively small: one year of exposure represents an increase of 3.5 - 4.6 percent of a standard deviation of student test scores, which means a full 6 year exposure would lead to a increase in test scores of 0.21 - 0.28 of a standard deviation in Math.

The coefficients on all controls have the expected sign and are significantly different from 0 in most cases. As expected, maternal education plays a very important role: other things equal, on

Table 4: Table Fixed Effects, Full Sample

	(1) Language out of 24	(2) Math out of 24	(3) Language out of 24	(4) Math out of 24
Years of exposure to FTS	0.12	0.21	0.08	0.16
	(0.06)*	(0.07)***	(0.07)	(0.07)**
Pupil is female =1	1.04	0.11	1.15	0.16
	(0.06)***	(0.06)*	(0.08)***	(0.08)**
Nr of people living in pupil's household	-0.13	-0.08	-0.13	-0.07
	(0.02)***	(0.02)***	(0.02)***	(0.02)***
Nr of rooms in pupil's home	0.11	0.10	0.06	0.04
	(0.02)***	(0.02)***	(0.03)*	(0.03)
Home equipm. modern appliances PCF1	0.52	0.58	0.48	0.52
	(0.04)***	(0.04)***	(0.05)***	(0.05)***
Mother completed primary education	0.33	0.33	0.35	1.00
	(0.11)***	(0.11)***	(0.12)***	(0.11)***
Mother has some secondary education	0.47	0.46	0.45	0.49
	(0.08)***	(0.08)***	(0.10)***	(0.10)***
Mother completed secondary education	0.63	0.62	0.60	0.49
	(0.09)***	(0.09)***	(0.13)***	(0.13)***
Mother completed tertiary education	0.53	0.63	0.52	0.59
	(0.12)***	(0.12)***	(0.21)**	(0.21)***
Teacher Experience 1 year or less =1	-0.83	-1.33	-0.89	-1.13
	(0.22)***	(0.23)***	(0.25)***	(0.25)***
Teacher Experience in years	0.00	0.03	0.02	0.05
	(0.01)	(0.01)***	(0.01)***	(0.01)***
Class size	0.01	0.02	0.01	0.00
	(0.01)	(0.01)**	(0.01)	(0.01)
This is an observation in 2002 =1	0.78	0.68	0.74	0.82
	(0.10)***	(0.10)***	(0.12)***	(0.12)***
Constant	13.12	10.59	12.46	10.38
	(0.28)***	(0.29)***	(0.35)***	(0.35)***
Observations	17249	17360	11061	11115
Number of Unique school identifiers	218	218	152	152
R-squared	0.08	0.08	0.08	0.08
Socioecon. context quintiles in sample	All	All	3, 4 and 5	3, 4 and 5
Number of full time schools	31	31	31	31
Number of comparison schools	187	187	121	121

Notes: The dependent variable is the raw test score on a scale of 0 to 20. Columns 1 and 2 include all 218 schools for which information was available in both 1996 and 2002, and who were not full time at baseline. Columns 3 and 4 are restricted to those of the 218 schools that pertain to the 3 least favored socio-economic quintiles. Those quintiles were calculated by ANEP in basis of the information that was collected in 1996. None of the schools pertaining to socio-economic quintiles 1 and 2 ever participated in the full time program. Standard errors are reported in parentheses. * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

average, a student whose mother has tertiary education scores 1.6 standard deviations higher than a student whose mother only completed primary. Increased teacher experience is associated with higher test scores, while pupils learning with very inexperienced teachers score significantly lower. The effect of class size is ambiguous.

5.2 Propensity Score Matching Estimates

5.2.1 First stage: matching

As described in the methodological section, we match the program schools with a set of comparison schools using propensity score matching on a set of observable school characteristics. First note that schools in socio-economic quintiles 1 and 2 do not belong to the support of the propensity score model, because none of them ever participated in the program. Of the 152 schools in quintiles 3, 4 and 5, 141 had sufficient information to estimate a propensity score model with the following explanatory variables: directors enthusisam and experience, average household size, number of rooms in homes and household equipment, mothers' education level, teacher experience, economic quintile, school equipment and buildings class size and type of shift system. The propensity score regression results are reported in Table 5.

The goal is to construct a comparison group of schools whose characterisitcs resemble those of the program schools. We verify the appropriateness of the methodology by comparing the baseline observable characteristics between the program group and the various comparison groups of schools (Tables 6 and 7). As evidenced in Table 6, stronger restrictions lead to the selection of a comparison group whose characteristics get closer to those of the full-time schools, though even with the closest neighbor matching, comparison schools seem to have better educated mothers and less repetition than full time schools. However, the best match of characteristics occurs when we exclude both the treatment and comparison schools that have a p-score below 0.15. As evidenced in Table 7, none of the key variables is statistically different between program and comparison schools when we restrict the sample to those schools for which we can predict participation in the program.

5.2.2 Second stage: Results using matched/restricted samples

Table 8 presents the summary of the results from the two first matching specifications. Table 1, Column 1 reports the estimated treatment effect using the full sample of schools in the 3, 4 and 5th socio-economic context quintile.[7] Columns 2 and 3 report the results when the sample is restricted to program schools and their 5 nearest neighbors and their closest neighbor, respectively.

[7] Note that this sample restriction is similar to the one in Table 4, Columns 3 and 4.

Table 5: Propensity score matching results

Dependent variable:	Full time school=1	
Director's enthusiasm level	-0.12	(0.30)
Director's experience	-0.04	(0.03)
Avg. household size	0.41	(0.29)
Avg. nr. of rooms in homes	0.63	(0.47)
Avg. household equipment p.c.	-0.82	(0.88)
Perc. of mothers with at least secondary education	3.23	(4.77)
Perc. of mothers with tertiary education	-9.75	(6.96)
Total teacher experience	0.05	(0.03)
School in economic quintile 4	-0.02	(0.57)
School in economic quintile 5	-0.30	(0.73)
School equipment p.c.	0.03	(0.28)
Auxiliary buildings p.c.	-0.26	(0.35)
Class size	0.01	(0.03)
School has morning shift	-2.13	(0.71)**
School has afternoon shift	-2.25	(0.72)**
School has 2 shifts	-1.29	(0.49)**
Constant	-4.28	(2.43)
Observations	141	
Pseudo R2	0.27	
ChiSq	37.97	
Number of full time schools	28	
Number of comparison schools	113	

Notes: The dependent variable is a categorical variable that takes value 1 if the school became a full time school between 1996 and 2002. The explanatory variables are measured at baseline in 1996. The sample only includes schools that were identified in 1996 as belonging to socioeconomic context quintiles 3, 4 and 5, excluding the most advantages quintiles 1 and 2. This is because none of the schools from quintiles 1 and 2 were transformed into full time schools, and hence the model has no statistical support for those quintiles. The omitted category is quintile 3. Household equipment, school equipment, and school auxiliary buildings are calculated using principal components estimation. (cf. section 3.4) The model is probit at the school level. Variables concerning mothers/homes/households were collected for students who took the 6th grade exam, and were aggregated at the school level. Standard errors are in parentheses. * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

Table 6: Summary statistics by propensity score group

	Full-time schools			Comparison schools					
				Full Sample		Closest 5		Closest	
Variable description	Obs	Mean	St.Dev.	Obs	Mean	Obs	Mean	Obs	Mean
Score in language, out of 24	710	12.48	4.62	4641	13.26	2260	12.81	809	12.87
Score in math, out of 24	710	10.25	3.68	4641	10.95*	2260	10.59	809	10.55
Student is female	710	0.51	0.50	4641	0.51	2260	0.51	809	0.52
Student repeated >=1 grade	706	0.41	0.49	4618	0.35***	2246	0.38	804	0.36
Student repeated >=2 grades	706	0.13	0.34	4618	0.11	2246	0.11	804	0.08**
Student's years of preschool	673	1.59	0.91	4495	1.57	2183	1.57	786	1.59
Household size	710	5.64	2.34	4641	5.37**	2260	5.59	809	5.53
Household's nr of rooms	710	3.50	1.37	4641	3.47	2260	3.44	809	3.44
Mother has no education	710	0.27	0.45	4641	0.20***	2260	0.23*	809	0.20**
Mother completed primary	710	0.73	0.45	4641	0.80***	2260	0.77*	809	0.80**
Mother attended secondary	710	0.42	0.49	4641	0.52***	2260	0.48*	809	0.51**
Mother finished secondary	710	0.14	0.35	4641	0.17	2260	0.15	809	0.16
Mother completed tertiary	710	0.08	0.27	4641	0.09	2260	0.07	809	0.07
Father has no education	657	0.30	0.46	4244	0.22***	2072	0.26	749	0.27
Father completed primary	657	0.70	0.46	4244	0.78***	2072	0.74	749	0.73
Father attended secondary	657	0.37	0.48	4244	0.49**	2072	0.45**	749	0.44
Father finished secondary	657	0.13	0.34	4244	0.18	2072	0.15	749	0.14
Father completed tertiary	657	0.09	0.28	4244	0.09	2072	0.07	749	0.06
Teacher experience <=1 yr	35	0.09	0.28	183	0.03	93	0.04	32	0.03
Teacher experience = 2-3 yr	35	0.06	0.24	183	0.05	93	0.08	32	0.09
Teacher experience= 4-5 yr	35	0.20	0.41	183	0.04**	93	0.04*	32	0.06
Teacher experience = 6-9 yr	35	0.09	0.28	183	0.16	93	0.17	32	0.13
Teacher experience >= 10 yr	35	0.57	0.50	183	0.72	93	0.67	32	0.69
Average Class Size	27	21.57	6.50	110	27.19***	57	25.99**	20	26.31**
School equipement	27	0.18	0.61	110	0.37	57	0.38	20	0.39
School auxiliary Buildings	27	-0.49	0.25	110	-0.40	57	-0.43	20	-0.50
School socio-econ. quintile	27	4.48	0.70	110	4.17*	57	4.39	20	4.50

Notes: This table reports summary statistics on key observables at the time of the baseline, for full-time schools and for three different sets of comparison schools: the full sample, the set of 5 closest schools in terms of propensity scores, and the set of closest schools. The table excludes schools that are missing any of the key variables used in the second stage analysis. School Equipment and School Auxiliary Buildings are estimated using Principal Component Factors. A t-test was performed for the equality of means between each comparison group and the group of full time schools. The t-test results are reported using star coding with * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

Table 7: Summary statistics for schools with high p scores

Variable description	Non-FTS Obs	Non-FTS Mean	Non-FTS Std. Dev.	FTS Obs	FTS Mean	FTS Std. Dev.	t-statistic
Score in language, out of 24	1554	12.70	4.32	544	12.74	4.62	0.05
Score in math, out of 24	1554	10.56	3.89	544	10.38	3.83	-0.32
Student is female	1554	0.51	0.50	544	0.53	0.50	0.66
Student repeated >=1 grade	1544	0.38	0.49	542	0.39	0.49	0.38
Student repeated >=2 grades	1544	0.11	0.32	542	0.11	0.32	0.11
Student's years of preschool	1503	1.62	0.92	513	1.58	0.92	-0.63
Household size	1554	5.65	2.15	544	5.63	2.33	-0.06
Household's nr of rooms	1554	3.45	1.36	544	3.57	1.41	1.01
Mother has no education	1554	0.24	0.43	544	0.28	0.45	1.00
Mother completed primary	1554	0.76	0.43	544	0.72	0.45	-1.00
Mother attended secondary	1554	0.44	0.50	544	0.42	0.49	-0.59
Mother finished secondary	1554	0.12	0.33	544	0.15	0.35	0.80
Mother completed tertiary	1554	0.06	0.23	544	0.08	0.27	1.31
Father has no education	1436	0.28	0.45	505	0.31	0.46	0.92
Father completed primary	1436	0.72	0.45	505	0.69	0.46	-0.92
Father attended secondary	1436	0.41	0.49	505	0.36	0.48	-1.26
Father finished secondary	1436	0.15	0.35	505	0.15	0.36	0.31
Father completed tertiary	1436	0.07	0.25	505	0.10	0.30	1.77
Teacher experience <=1 yr	65	0.05	0.21	27	0.11	0.32	0.99
Teacher experience = 2-3 yr	65	0.11	0.31	27	0.07	0.27	-0.51
Teacher experience= 4-5 yr	65	0.06	0.24	27	0.19	0.40	1.35
Teacher experience = 6-9 yr	65	0.15	0.36	27	0.07	0.27	-1.12
Teacher experience >= 10 yr	65	0.63	0.49	27	0.56	0.51	-0.65
Average Class Size	39	25.04	8.89	22	21.78	6.01	-1.53
School equipement	39	0.31	0.63	22	0.14	0.66	-1.05
School auxiliary Buildings	39	-0.53	0.33	22	-0.47	0.25	0.77
School socio-econ. quintile	39	4.56	0.64	22	4.50	0.67	-0.37

Notes: This table reports summary statistics on key observables at the time of the baseline, for full-time schools and comparison schools with a p-score above .15. The table excludes schools that are missing any of the key variables used in the second stage analysis. School Equipment and School Auxiliary Buildings are estimated using Principal Component Factors. A t-test was performed for the equality of means between the comparison group and the group of full time schools. The t-test results are reported using star coding with * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

Table 8: Summary results

	(1)	(2)	(3)	(4)
Sample	All sch. in quintiles 3, 4, 5	Matched Sample		
Methodology	Fixed-Effects	Closest 5	Closest	p-score>15
	Language			
Number of Years FTS	0.08	0.08	0.14	0.21
	(0.07)	(0.07)	(0.09)*	(0.08)**
Constant	12.43	11.56	12.52	11.89
	(0.35)***	(0.40)***	(0.56)***	(0.46)***
No obs.	11061	5910	3431	4313
No schools	152	83	49	63
R square	0.08			
	Mathematics			
Number of Years FTS	0.16	0.2	0.2	0.29
	(0.07)**	(0.07)***	(0.08)**	(0.08)***
Constant	10.35	9.74	10.73	10.26
	(0.34)***	(0.40)***	(0.55)***	(0.45)***
No obs.	11115	5910	3431	4313
No schools	152	83	49	63
R square	0.08			

Notes: This table reports the estimated program effects under different sample restrictions. Column 1 reports the estimated treatment effect using the full sample of schools in the 3, 4 and 5th socio-economic context quintile. Columns 2 and 3 report the results when the the sample is restricted to program schools and their 5 nearest neighbors and their closest neighbor, respectively. Finally, in Column 4 the sample is restricted to those schools with a propensity score higher than 0.15. All coefficients and standard errors in columns 2, 3 and 4 are bootstrapped using 300 replications, and observations that do not belong to the support of the p-score function are excluded. Full results are reported in the next table. * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

Finally, in Column 4 the sample is restricted to those schools with a propensity score higher than 0.15. The estimated impact of the program is positive, and more so in Math than in Language. The use of different sample restriction strategies points to an interesting fact: the more restricted the sample of control schools is, the larger is the effect of the full-time school program. In Math, the full-sample estimate of the program impact is 0.16 points per year of participation in FTS. Limiting the sample to the 5 closest neighbors, or to the closest neighbor of treatment schools, we find that the estimated program impact is 0.20 of a standard deviation. (Top panel, Columns 2 and 3) Program effect estimates for Language are positive but smaller than the estimated effects for Math, and they are only significantly different from zero when we match using the closest neighbor.

Column 5 reports results when the sample is restricted to schools for which we have a predicted probability of participating higher than 0.15. In this approach, a number of better-off schools which participated are excluded, ensuring more uniform characteristics among the schools included in the regression. Under that specification, we find large and significant effects of the program, for Math (0.29 points per year of participation) and Language (0.21 points per year of participation). The larger estimates in this specification can be interpreted as follows. Assuming that the effect of the program depends on the characteristics of the school, and that more disadvantaged schools benefit more from the program, we expect that the estimated program effect will be larger if we restrict the sample to more disadvantaged schools. Unfortunately, the limited sample size prevents us from testing this theory by estimating interaction effects between the treatment variable and school characteristics. Nevertheless, the results from the different specifications and the characteristics of excluded schools do seem to support this explanation.

The effect of the program is sizeable, especially when considering disadvantatged schools. Our largest estimated effect in Math is 0.29 points per year, or about 0.063 standard deviations. Thus, a full 6 year cycle may result in an increase of close to 0.38 of a standard deviation in these schools. Since the average score in Math in disadvantaged schools is slightly above 12 (the minimum passing grade), and 65 percent of their students score below passing grade, the impact of the program in these schools is substantial: in an average disadvantaged school 6 years of participation in the program would bring close to 10 percent additional students to the minimum score for passing. Though the improvement in Language scores is more modest than the improvement in Math scores (0.21 points, or about 0.044 of a standard deviation per year of exposure to the program), it is still substantial. We would expect this improvement in learning outcomes to have an effect on completion and transition to secondary in low achieving schools, increasing schooling and helping address the substantial inequalities in the Uruguayan education system.

A quick review of the other coefficients in the matching regressions, Table 9, shows that

Table 9: Propensity score matching full results

	(1)	(2)	(3)	(4)	(5)	(6)
	closest 5 neighbors		closest neighbor		pscore<.15	
	Language	Math	Language	Math	Language	Math
Years of exposure to FTS	0.08	0.2	0.14	0.2	0.21	0.29
	(0.07)	(0.07)***	(0.09)*	(0.08)**	(0.08)**	(0.08)***
Pupil is female =1	1.15	0.32	1.18	0.45	1.16	0.21
	(0.10)***	(0.11)***	(0.14)***	(0.14)***	(0.13)***	(0.12)*
Nr of people living in pupil's hh.	-0.14	-0.09	-0.15	-0.1	-0.12	-0.07
	(0.02)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***	(0.03)**
Nr of rooms in pupil's home	0.03	0	-0.01	-0.02	0.03	-0.03
	(0.04)	(0.04)	(0.06)	(0.05)	(0.05)	(0.05)
Home equipm. modern appliances	0.57	0.53	0.57	0.53	0.51	0.46
	(0.08)***	(0.08)***	(0.10)***	(0.10)***	(0.08)***	(0.09)***
Mother completed primary educ.	0.33	0.41	0.14	0.45	0.4	0.49
	(0.14)**	(0.14)***	(0.19)	(0.22)**	(0.16)**	(0.16)***
Mother has some secondary educ.	0.45	0.52	0.66	0.61	0.46	0.59
	(0.13)***	(0.12)***	(0.17)***	(0.17)***	(0.15)***	(0.15)***
Mother completed secondary educ.	0.33	0.22	-0.01	0.06	0.38	0.34
	(0.21)	(0.18)	(0.28)	(0.26)	(0.26)	(0.24)
Mother completed tertiary educ.	0.58	0.65	0.42	0.52	0.43	0.55
	(0.29)**	(0.30)**	(0.38)	(0.35)	(0.38)	(0.36)
Teacher exper. 1 year or less=1	-0.89	-1.21	-0.97	-1.44	-0.73	-0.81
	(0.35)**	(0.33)***	(0.43)**	(0.46)***	(0.37)*	(0.31)***
Teacher experience in years	0.05	0.08	0.08	0.1	0.03	0.07
	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***
Class size	0.03	0.02	0	-0.02	0.01	-0.01
	(0.01)**	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
This is an observation in 2002 =1	0.38	0.31	-0.03	0.24	0.33	0.21
	(0.15)**	(0.17)*	(0.25)	(0.25)	(0.19)*	(0.18)
P-score	0.64	-0.04	0.55	-0.47	0.7	0.37
	(0.35)*	(0.31)	(0.40)	(0.36)	(0.42)*	(0.36)
Constant	11.56	9.74	12.52	10.73	11.89	10.26
	(0.40)***	(0.40)***	(0.56)***	(0.55)***	(0.46)***	(0.45)***
Observations	5910	5910	3431	3431	4313	4313
Number of Unique school identifiers	83	83	49	49	63	63
Number of full time schools	28	28	28	28	23	23
Number of comparison schools	55	55	21	21	40	40

Notes: This table reports the estimated program effects under different sample restrictions. Columns 2 and 3 report the results when the sample is restricted to program schools and their 5 nearest neighbors in terms of p score. Columns 4 and 5 report the results when the sample is restricted to program schools and their closest neighbor in terms of p score. Finally, in Column 5 and 6, the sample is restricted to those schools with a propensity score higher than 0.15. All coefficients and standard errors are bootstrapped using 300 replications. In all columns, observations that do not belong to the support of the p-score function are excluded. * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

all estimates are stable, except for the coefficient on mother's tertiary education and the 2002 year dummy. Given that the percentage of mothers with tertiary education is much lower in more disadvantaged schools, it is not surprising that the coefficient estimates on this variable are much less precise in the restricted sample. Regarding the year dummy, we interpret the change in coefficients on the 2002 dummies and on years of exposure as follows: between 1996 and 2002, test scores increased significantly for the country as a whole. However, the time effect did not occur in poorer schools. These are likely to have suffered disproportionately from the crisis in the early 2000's, which may have resulted in differential trends between poor and better-off schools. When we include better-off schools in the comparison group for the full-time schools (column 1), we underestimate the effect of the full time school program because we do not account for the differential trend in learning achievement.

In summary, we find that the full time schools program seems to have raised student achievement, expecially in the most disadvantaged schools. It is important to remember that the goal of restricting the sample is to create a control group of schools that resembles the treatment group as much as possible. Given that the program targeted particularly disadvantaged schools, this implies that we are restricting the control group to disadvantaged schools. It is in this comparison, when we reduce the potential bias arising from different characteristics in the treatment and control groups, that we find the larger effects. The size of the effect is substantial: using our preferred estimates, we find a 0.044 of a standard deviation improvement in language, and a 0.063 percent of a standard deviation improvement in Mathematics per year of exposure to the program. For a child who spends 6 years in the program, this would add up to a 0.26 of a standard deviation improvement in Language, and a 0.38 of a standard deviation in Mathematics.

6 A Note About Cost

Even though the program appears to have yielded positive effects on learning outcomes of participating schools, the full-time school program is expensive. As currently designed, it requires the provision of new infrastructure, ranging from construction to teaching materials, and a substantial increase in recurrent costs, especially teacher's salaries. Additional activities, teacher training and parallel nutritional and health interventions in these schools contribute to the higher cost of these schools. The question of whether an expansion of the program is worthwhile would ideally involve a careful exploration of the impacts of each of the components of the program and an exact estimation of the benefits that would go farther than learning outcomes. As we discussed above, the break-down of the impacts of the major components of the program (teacher training, school time, community participation) is not possible since participation in those sub-components was either universal or voluntary, which causes identification problems that we cannot solve with the data available. Nevertheless, in this section we use estimates of costing data from ANEP to try to shed some light on the cost and benefit implications of an expansion of the program.

According to ANEP, FTS have on average 50 percent more recurrent costs (excluding teacher training) than regular urban schools,[8] mostly due to the increase in the cost of teacher salaries, which explains 71.2 percent of the difference in costs. Nutrition interventions account for an additional 19 percent, while the remaining cost increase is accounted for by the extra-curricular activities that are associated with the implementation of the new padagogical approach in FTS schools. If we also include teacher training in the costing, FTS turn out to cost 60 percent more than the average regular urban schools. Assessing whether the increase in learning compensates this extra cost requires some degree of subjectivity, given the analysis we are able to perform. However, it is important to note the following: First, even though FTS are more expensive to run than regular urban schools, this does not necessarily imply that they are too expensive. On the contrary, it may indicate underspending in the rest of the education system. In fact, by international comparisons, Uruguay underspends in the education sector given its level of income. The average per-student expenditure in primary and secondary schools is 7.2 percent of GDP per capita, well below the regional average of 13.1 percent.[9] Thus, the additional cost of FTS, while substantial when compared the cost of regular urban schools, would only bring Uruguay closer to the regional average. This by no means means that the program should be scaled up universally, but provides a relative measure of the increase in costs from the program. The second important point is that the program has a large effect in schools with students from very disadvantaged backgrounds. These are the students for whom the benefit of a marginal increase in learning would be the largest, assuming that this increase in learning results in a higher probability of completing secondary and post-secondary education. Thus, efficiency and equity arguments would recommend a targeting mechanism that ensures that the participating schools serve disadvantaged populations.

7 Conclusions

Our preferred estimates show that the full-time school program increased student test scores in third grade by 0.063 standard deviations per year in mathematics and 0.044 standard deviations

[8] MCAEP ANEP, "Estimacion del gasto por alumno en Educación Primaria Pública y costos comparados de la educacion comun y de tiempo completo". These estimates control for the size of the school.

[9] Source: The World Bank (SIMA) and EdStats, 2004.

per year in Language in schools with students from disadvantaged backgrounds. We estimate that the program leads to an additional 10 percent of students reaching the passing score on the third grade test in these schools. The program is relatively expensive, increasing recurrent costs by approximately 60 percent. A careful cost-benefit analysis could not be performed due to data limitations and the characteristics of the implementation of the program. Despite the substantial cost of the program, it may be worth considering expanding the program considering the fact that Uruguay substantially underspends in primary and secondary education when compared to the region and to countries with a similiar income level. An increase in 60 percent in expenditures would only bring Uruguay closer to the region's average. The full-time school program, if well targeted, could help address inequalities in education in Uruguay, at an increase in cost per student not larger than the current deficit in spending between Uruguay and the rest of the region. When considering scaling up the program, the targeting mechanism should ensure that the program focuses on schools serving very disadvantaged populations. In addition, it would be useful to evaluate cost and impact of the different components of the program in order to optimize the interventions. A careful roll-out plan for future expansion of the program would help evaluate the impact of these sub-components.

References

[1] Administración Nacional de Educación Pública (2002), "Los Niveles de Desempeño al Inicio de la Educación Primaria: Estudio de las Competencias Lingüisticas y Matemáticas," Primer Informe, Montevideo.

[2] MECAEP (2001), "Estudio de la Evaluacion Social de las Escuelas de Tiempo Completo," Montevideo.

[3] Administración Nacional de Educación Pública (ANEP), Proyecto MECAEP - ANEP/BIRF (1997), "Evaluación Nacional de Aprendizajes en Lengua Materna y Matemática, 6to. año de Enseñanza Primaria - 1996, Segundo Informe de Difusión Pública de Resultados," Montevideo, República Oriental del Uruguay.

[4] Bellei, Cristian (2005), "Does the length of the school day have an impact on the student's academic achievement?" Harvard Graduate School of Education, Unpublished Paper.

[5] Valenzuela, Juan Pablo (2005), "Partial Evaluation of a Big Reform in the Chilean Education System: From Half Day to Full Day Schooling" in "Essays in Economics of Education," University of Michigan doctoral dissertation No.31867777.

[6] Wooldridge, Jeffrey M. (2002), Econometric Analysis of Cross Section and Panel Data, MIT Press, Cambridge, Massachusetts.

[7] Walston Jill and Jerry West (2004), "Full-Day and Half-Day Kindergarten in the United States Findings From the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99," The Education Statistics Quarterly, Vol. 6, Issues 1 & 2.