THE EFFECTS OF EARLY CHILD EDUCATION ON LITERACY SCORES USING DATA FROM A NEW BRAZILIAN ASSESSMENT TOOL

Fabiana de Felício
Studies Office of the Minister of Education – MEC

Rafael Terra de Menezes
EESP-FGV

Ana Carolina Zoghbi
CEPESP-FGV

RESUMO

Seguindo a literatura de Economia da Educação, o objetivo deste trabalho foi identificar os efeitos da Educação Infantil (EI) sobre os escores de alfabetização das crianças do 2º ano do Ensino Fundamental. Para isso, aplicamos a Provinha Brasil em um município Brasileiro (Sertãozinho no Estado de São Paulo) juntamente com um questionário sócio-econômico respondido pelos pais das crianças. Apesar dos problemas de validade externa, a avaliação dos efeitos da EI em um único município é vantajosa, já que se pode estimar os efeitos de um tipo tratamento. Vários estudos ignoram esse fato. Muitas vezes, é estimado um efeito médio de vários efeitos de tratamento e não apenas um, já que usam dados de diferentes municípios onde o EI tem diferentes níveis de qualidade. Os resultados encontrados com o uso de OLS e de *Propensity Score Matching* mostram que alunos que ingressaram com 5, 4, e 3 anos de idade, obtiveram escores de alfabetização cerca de 6% maiores do que os obtidos pelos que ingressaram na escola com 6 anos ou mais. Contudo, não se observou diferenças significativas de desempenhos entre os alunos que cursaram somente um ano de Educação Infantil e aqueles que cursaram mais tempo. Foram também realizados testes para validar a qualidade do *Matching* e conferir maior confiabilidade aos resultados.

Palavras-Chave: Educação Infantil, Escores de Alfabetização, Propensity Score Matching.
Classificação JEL: A29, I21, C21
Área Anpec: Área 11 – Economia Social e Demografia

ABSTRACT

Following the Economics of Education literature, the aim of this paper was to identify the effects of Early Child Education (ECE) on literacy scores of children attending the 2nd grade of elementary school. To do that, we applied the *Provinha Brasil* in a Brazilian municipality (Sertãozinho in State of São Paulo) in conjunction with a socioeconomic questionnaire to be answered by children's parents. Despite external validity problems, the evaluation of the effects of ECE in one municipality is advantageous, as we can estimate the effects of one kind of treatment. Several studies ignore this fact. Often, it is estimated an average effect of various treatments effects and not just one, as they use data from different municipalities where the ECE have different levels of quality. The results obtained here with the use of OLS and Propensity Score Matching show that students who started school at the ages of 5, 4, and 3 years had literacy scores approximately 6% greater than those who began school at the age of 6 years or later. However, no significant differences in performance were observed between students who attended only one and more than one year of Early Child Education. Tests were also run to validate the matching quality and to improve the reliability of results.

Keywords: Early Child Education, Literacy Scores, Propensity Score Matching.
JEL Classification: A29, I21, C21
Anpec Area: Área 11 – Economia Social e Demografia

1 INTRODUCTION

With the advent of the Human Capital Theory – especially with the work conducted by Becker (1964) – we observe the consolidation of a theoretical framework where education played a major role in the determination of poverty, of long-term growth, of per capita income, and of income inequality within and between countries. Nevertheless, the identification of this relationship between education and income per se was not enough for the implementation of public policies targeted at the improvement of people's living conditions. In this respect, we have the Economics of Education literature, whose main goal is to identify the most important factors for the development of people's skills.[1] Among these factors, family background plays a crucial role. According to these studies, the influence of family on the development of their children's skills is so important that the room for public sector action seems quite restricted. This perspective is particularly problematic for Brazil, since most adults have a low educational level and live in precarious conditions, which hinders their children's skill development and perpetuates the cycle of poverty.

Despite the limited scope of public policies aimed at improving education (in terms of quality and quantity), some factors indirectly related to family are important, as evinced by empirical studies. Improvement of school infrastructure, qualification of teachers and principals, accountability, in addition to other activities offered at schools, are some examples.[2][3][4] But another specific factorhas been given special attention in the literature: the Early Child Education.

In the last two decades, there has been a debate about the effects of Early Child Education on future educational outcomes. A considerable number of scientific studies find evidence that educational intervention early on during childhood yields significant and long-lasting results on future school performance, and even on adult life successes. Based on this literature, Cunha *et al.* (2005) introduce a model for skill development that reflects the necessity that investments in education be made in early childhood. The model also considers that later investments are important to maintain the skill level developed in early childhood, but if investments are not made in early childhood, the ability to acquire knowledge in the future will be impaired, and this cannot be offset by further investments in the future. Therefore, the authors refer to early childhood as the *Critical Period*.

Low scores in standardized assessments obtained by Brazilian students in national exams (Prova Brasil and SAEB –National Basic Education Assessment System) and also in international exams (Programme for International Student Assessment – PISA), high repetition rates and high dropout rates before accomplishing High School education, underscore the need for urgent action by the public sector, and Intervention in Early Childhood is an important alternative for Brazilian public education.

Some measures have been taken by the Brazilian federal government. To improve children's educational level and to encourage early school admission, the minimum period for accomplishment of Elementary Education has recently been extended from 8 to 9 years, which decreased the mandatory age for school admission to 6 years. Moreover, an attempt has been made to boost the supply of vacancies at day care centers and at pre-schools and to stimulate the school admission of children aged 0 to 5 years with inclusion of Early Child Education in FUNDEB (Fund for the Maintenance and Development of Basic Education and Teaching Improvement).[5] This implies the guarantee of a minimum common grant value in all states and municipalities for every child enrolled from the Early Child Education program to High School. Also, there is a congress bill in discussion that, if enacted, will make these two stages mandatory, as occurs with Elementary Education.

---

[1] Coleman *et al.* (1966), have conducted a seminal work on the determinants of school success.

[2] For educational accountability results, see Jacob (2005) and Carnoy & Loeb (2002).

[3] For further information about the effects of teachers' characteristics on school success, see Rivkin, Hanushek & Kain (2005).

[4] For the effects of some variables related to school infrastructure on students' performance in standardized exams, see Albernaz, Ferreira & Franco (2002) and Felício & Fernandes (2005).

[5] Before Fundeb, the federal fund for the transfer of grants for education was FUNDEF (Fund for the Maintenance and Development of Elementary Education and Teaching Improvement – implemented by Amendment no. 14, enacted in September 1996), which guaranteed minimum resources for students enrolled in Elementary School. Early Child Education (for children aged 0 to 6 years at the time) and High School (for teenagers aged 15 to 17 years) were financially supported only by states and municipalities. Given the evidence of the efficiency of investments in Early Child Education, this strategy of heavily investing in Elementary Education, with support from FUNDEF, has been deemed to be mistaken.

Another measure that is important for the improvement of the education system is to warrant that children can be taught to read and write at the appropriate age so that they will be ready to develop the skills that are required throughout their school life. In this regard, the Brazilian Ministry of Education developed *Provinha Brasil* in 2007, an exam that assesses the literacy of children aged 6 to 8 years.

This exam is an innovative tool used to measure the development of children's reading and writing skills, when it is still possible to correct learning deficiencies in a more effective way (at least if we compare it with the possibilities of educational intervention for older children) and to adjust, whenever necessary, the process that precedes the assessment. It can also be used to assess the effects of public policies adopted in response to deficiencies revealed by the exam results.

As *Provinha Brasil* assesses students at the very beginning of Elementary Education, the main goals of educational policies targeted at the improvement of literacy levels diagnosed by the exam should concern Early Child Education and the first years of Elementary Education. This is consistent with the model put forward by Cunha et al. Students with low literacy scores in *Provinha Brasil* might not be able to achieve the potential they would be able to if they had been properly taught to read and write. Therefore, interventions should occur when children are still young, i.e., during Early Child Education, so that students in the subsequent cohorts can be better prepared when they are assessed.

The aim of this study is to empirically determine the relationship between age at school admission (which also reflects the fact that a student attended or did not attend Early Child Education) and children's literacy score based on the results obtained in *Provinha Brasil*. However, this exam differs from other tests applied by the Ministry of Education. Its application and correction requires active participation of the local school systems. As a matter of fact, the Ministry of Education makes the test and the manuals for its application and correction and provides the public school systems all over Brazil with the exam and respective manuals in digital format, but the exam is applied and corrected by the local school systems.

There is neither a database with the exam results (as the application is decentralized), nor a socioeconomic questionnaire to be answered by the children, as occurs in SAEB, since the children who take this exam are too young to answer such a questionnaire in an accurate way. Thus, the solution we found to use the exam results in an econometric analysis was to closely follow the application of *Provinha Brasil* in a Brazilian municipality where Early Child Education was well structured and where we could also apply socioeconomic questionnaires directly to the parents. This way, we chose Sertãozinho, a municipality located in the state of São Paulo.

As this is a medium-sized town (with a projection of 109,565 inhabitants according to Brazilian Institute of Geography and Statistics – IBGE – 2008 and 2,081 students attending the 2nd grade of Elementary School in 2007), it was possible to standardize and inspect universal application properly. This means that it was also possible to apply the exam in private schools, which allowed us to draw generalized conclusions about the econometric results we obtained. Note that the cooperative attitude of private and public schools was important for the selection of this municipality.

The assessment of a single municipality, even though it has some problems relative to the external validity of the results, is desirable in an analysis using Propensity Score Matching, as the one we carried out in this study. Friedlander & Robins (1995) and Michalopoulos et al. (2004) find evidence that when treatment and control groups belong to the same place, the matching procedure yields better results. Indeed, if we consider Early Child Education an intervention, one cannot use treatment units from various localities because each place has one quality of education and thus we have a number of different treatments instead of just one.

Finally, this study is important because it verifies the effectiveness of incentive policies for the expansion of Early Child Education as a way to increase students' learning. To achieve that, we sought to identify the relationship between the attendance of Early Child Education and the literacy score attained in the 2nd grade of Elementary School, using a new assessment tool that measures an especially important set of skills, as it interferes directly with learning capacity in any area of knowledge.

This study is organized into seven sections, including this introduction. Section 2 reviews the literature on the effects of Early Child Education on children's skill development. Section 3 describes the *Provinha Brasil* in more detail, as well as the application of this exam and of the questionnaire to the

selected municipality, and also presents some descriptive statistics. Section 4 explains the whys and wherefores of using the Propensity Score Matching as estimation method. Section 5 describes the estimation results. Section 6 presents the tests used to check whether the matching was properly performed and, finally, Section 7 brings the concluding remarks.

2 LITERATURE REVIEW

The literature dealing with the effects of Early Child Education on children's skill development is quite rich and comprehensive. Most of the reported results are based on random experiments, but important studies that use non-experimental methods are also included. Some experiments conducted in the United States, which are especially noteworthy in the literature, include the following: the High/Scope Perry PreSchool Project, the Carolina Abecedarian Project and the Early Training Project (CURRIE, 2001).

The High/Scope Perry PreSchool Project was an experiment undertaken between 1962 and 1967 which assessed 123 children aged 3 to 4 years (all of the children had a vulnerable socioeconomic background). The treatment consisted of pre-school attendance for one shift (morning or afternoon) every day plus a 90-minute weekly home visit for eight months a year during two years. Students usually left the program at the age of 5 years. All of the teachers had a Master's degree and the teacher-to-student ratio was 1:6. According to Schweinhart et al. (1993), the results obtained with this program were quite positive: better performance on skill tests (at 9 and 14 years), better performance in High School, higher rates of High School accomplishment, lower rates of imprisonment (at 27 years), higher salaries (at 27 years) and lesser use of government support (at 27 years).

The Carolina Abecedarian Project was an experiment targeted at children with a vulnerable socioeconomic background and with risk for mental retardation due to the precarious situation in living (CURRIE, 2001). Admission to the experiment occurred at 6 to 13 weeks of age. The treatment consisted of an intensive child care and language development program for eight hours a day, 5 days a week, 50 weeks a year, from birth to the age of 5 years. After admission to school, the treatment group was randomized into two groups, one with a tutor who provided additional instructions at home, and another one that did not have any additional intervention. The program was concluded up to the age of 8 years. The teacher-to-student ratio was initially 1:3, rising to 1:6 as children grew older. The results were also quite positive. The treatment group had better results on proficiency tests, higher averages of school performance in High School, lower incidence of repetition (at 15 years) and lower dropout rates (at 21 years), in addition to lesser need for special education (at 15 years), and higher probability of attending college (at 21 years).

The Early Training Project was a less intensive program, targeted at children aged 4 to 5 years. It consisted of weekly home visits for 1 year, in addition to a single-shift 10-week course for two or three summers for the treatment groups. The treatment was concluded at the age of 6 years. According to Gray et al. (1983), the results showed a reduction in the need for special education for children in the treatment group.

These three experiments share a common feature: they are all small-scale projects. This may have important implications, since it is impossible to determine the external validity of these experiments. The results of these experiments if they were to be carried out on a large scale cannot be predicted. An important and more comprehensive experiment is the Head Start program, adopted by the U.S. government in 1964, and whose target public was composed of children aged 3 to 5 years. Intervention consists in providing health care, meals, snacks and child care with a higher quality standard than that which low-income parents can provide for their children. Currie & Thomas (1995) conducted a quasi-experimental study of the effects of this program on the performance obtained on the Peabody Picture Vocabulary Test (PPVT) and on the probability of never having to repeat a grade. The authors found positive effects on the performance of white children who participated of this program, whereas the results for Afro-American children, who participated in the program, were not statistically different between those who took part in the program and those who did not. After controlling for participants' ages, the authors found a positive and statistically significant effect on the performance on the PPVT and

on the probability of never having to repeat a grade, also among Afro-American children. According to the authors, this may be linked to the possibility of children from this ethnic group facing more hostile environments or opportunities after they finish the program. Thus, in order for the benefits to last longer, as occurs among white children, investments, even after the program has finished, should be made in those children enrolled in the program.

Recently, important experiments and quasi-experimental studies have been conducted in Latin America. Schady (2006) carried out a literature survey on assessment of the impact of investment programs in childhood on cognitive and non-cognitive development. Gertler & Fernald (2004, apud SCHADY, 2006), for instance, found evidence that transfers made by Mexico's Opportunities Program had a positive impact on motor skills and socioemotional problems. Behrman, Parker & Todd (2004, apud SCHADY, 2006), on the other hand, gathered evidence that these transfers had a positive impact on the probability of children enrolled in the program (aged 0 to 6 years) starting school at an earlier age. They also found evidence of higher promotion rates and higher expected schooling years among children of the treatment group.

Behrman, Cheng & Todd (2004) analyzed the results of a Bolivian pre-school program called *Proyecto Integral de Desarollo Infantil* (PIDI) using Propensity Score Matching. The program consists of intensive child care, such as full-time day care center and nutritional and educational services for children aged between 6 months and 6 years from low-income families. The authors encountered evidence of improved motor and psychosocial skills and of improved language acquisition. These results were more significant among children older than 3 years and among those who attended the program for a longer period.

Berlinski et al. (2009) assess the effects of an exogenous variation in the supply of pre-school vacancies due to a program for the construction of schools in Argentina on students' performance. The authors collected evidence that the performance in the Spanish language and in mathematics of cohorts and regions subjected to the treatment (construction of schools) was significantly higher than that obtained by children who were not exposed to this exogenous variation. They also found evidence of positive impacts on non-cognitive skills such as attention, participation and discipline in the classroom.

In Brazil, some studies were developed to assess the effects of an intervention (Early Child Education) in early childhood. Curi & Menezes-Filho (2006), for instance, demonstrate that students who attended pre-school and a day care center are more likely to finish primary education (1st to 4th grade of Elementary School), Middle School (5th to 8th grade), High School and College. The authors also observed that because individuals attended a day care center or pre-school, they have a higher average of schooling years (measured from the first year of primary education) and higher salaries as well. Finally, the authors found evidence that early school admission has positive effects on math proficiency.

This result about the effects of Early Child Education in Brazil on performance is corroborated by Felicio & Vasconcellos (2007). The authors use methods for the correction of endogeneity and self-selection bias and find positive and statistically significant effects of pre-school attendance on the performance of fourth-graders on SAEB. Depending on the region, these effects ranged from 9 to 19%.

Given the evidence of these studies and the model proposed by Cunha et al. (2005) described in the Introduction of the present study, it is fundamental to address the problem of liquidity constraints faced by socioeconomically underprivileged families. Following this line of research, the simulations of a general equilibrium model run by Restuccia & Urrutia (2004) confirm the existence of intergenerational persistence of earnings and schooling. According to the authors, this problem arises mainly from low investments in the earliest stages combined with liquidity constraints by poorer parents. Therefore, children whose parents cannot afford to invest in this initial stage will be condemned to have poor skills in the subsequent stage and, consequently, to have lower earnings in adult life.

This is particularly worrying as children with the worst family background are exactly those who take most advantage of these programs in early childhood (CURRIE, 2001). These programs would be useful to minimize adverse family conditions. This way, there is room for the government to act in order to provide equal opportunities to the children, i.e., to lessen the large disparities observed between the initial skills of economically privileged and underprivileged groups.

3 DATASET

The database used in this study was built using the results obtained from the first application of *Provinha Brasil* in Sertãozinho,[6] in the state of São Paulo, in May 2008. The exam was applied universally (to public and private education networks) with the aim of assessing literacy measured as the reading and writing skills of students attending the 2nd grade of nine-year-long Elementary School education. The choice of Sertãozinho was based on the fact that this town had a not so large number of children enrolled in this grade (2,081 students according to the 2007 School Census), thus allowing us to follow up the application of the exam and of the socioeconomic questionnaires in an appropriate fashion. The presence of a well-structured Early Child Education system is another highlight of this municipality.[7] Finally, the availability of public and private schools to participate in the study was also decisive for the selection of this municipality.

It should be underscored that *Provinha Brasil* differs from other Brazilian exams in some important aspects, besides the fact that it also assesses literacy. The first aspect concerns the responsibilities attributed to each phase of the exam. In other Brazilian exams developed by the Ministry of Education (Prova Brasil and SAEB), an external institution is hired to apply the exams, organize the data and hand the database over to the Ministry of Education. In the case of *Provinha Brasil*, the National Institute for Educational Studies and Research (INEP) is in charge of devising the material and making it available online so that schools can print it. The use of *Provinha Brasil* by Municipal Departments of Education is optional. Therefore, the application, correction and publication of results are locally assigned responsibilities[8].

Another difference lies in the assessment method. The score of *Provinha Brasil* ranges from 0 to 24 points, as opposed to the scores of Prova Brasil and SAEB, which range from 0 to 500 points. As previously mentioned, *Provinha Brasil* aims to assess children's literacy. Thus, its maximum score (24), can be achieved by children only when they are totally literate. The other exams, however, focus on determining the skill levels developed by each student during K-12 education, and in practice this means that students hardly achieve the maximum score. Nevertheless, *Provinha Brasil*, SAEB and Prova Brasil are all based on the Item Response Theory, which allows the results to be placed on the same scale and comparing the results between assessments, between the grades of the same exam, and over time. Therefore the results of these exams permit monitoring education quality.

As stated in the first section of the present study, another aspect that distinguishes *Provinha Brasil* from other Brazilian exams is the absence of questionnaires for the collection of relevant information to explain student performance. Children who take part in *Provinha Brasil* are very young, which does not allow them to answer a questionnaire in a consistent manner.

For the application of *Provinha Brasil* to public schools, the Municipal Departments of Education offer examiners a training course. To carry out this study in Sertãozinho, this was no different. We only followed up the process to guarantee that both application and correction were standardized.

However, in private schools, we endeavored to reproduce the same standard adopted for public schools. In order to set the schools ready, we relied on the help from the Board of Education, of the State Department of Education. We held a meeting with the coordinators of each school to explain the objective and importance of the study and then we trained the teachers in charge of applying the exam. The training also included instructions on how to apply the socioeconomic questionnaires to be answered by students' parents or surrogates, the same guidelines public schools were provided with.

Altogether, 14 public schools and 9 private schools were assessed in Sertãozinho. Among those students enrolled in the 2nd grade of Elementary School, 1,986 took *Provinha Brasil* exam, which is a

[6] Sertãozinho is a countryside town in the State of São Paulo whose major economic activity is the sugarcane industry.

[7] By comparing pre-school attendance between public and private schools (using data from the Basic Education Census), we noted that the average rate between the total number of enrollments in public and private schools of Sertãozinho was 6.29 between 2005 and 2006. When we made the same calculation for Brazil, the average rate was 2.69. In terms of municipal expenditures with Early Child Education per student (based on data from the National Treasury Department), we perceived that Sertãozinho spent on average R$1,761 between 2005 and 2006, while the average expenditure at the national level was R$1,196.

[8] Actually, some well known state level exams (e.g. SARESP, of the State of São Paulo) are applied and corrected by the teachers themselves, instead of an external institution.

significant share considering the estimate of approximately 2,100 students enrolled in 2008. The exam was applied on the same day in both public and private schools. With regard to the questionnaires, they were applied in order to collect information about the socioeconomic characteristics and school history of students.[9] Parents were asked to fill out the questionnaires at teacher-parent meetings, or the questionnaires were sent to those parents who did not turn up for the meetings so that they could answer and return them later.

Table 1 –Literacy Scores and characteristics of the 2nd grade students of Sertãozinho-São Paulo

	Variable	Obs	% within category	Mean Score	Std. Dev.	Min	Max
1	Literacy Scores (all students)	1986	100.0	19.4	4.0	2	24
2	Entered school at ages 7 or more	22	1.2	18.2	5.3	2	23
	Entered school at age 6	258	14.6	18.3	4.2	5	24
	Entered school at age 5	344	19.6	19.4	3.8	7	24
	Entered school at age 4	535	30.6	20.0	3.6	2	24
	Entered school at ages 3 or less	597	34.0	20.0	3.8	3	24
3	lives with mother and father	1358	73.5	19.7	3.8	2	24
	doesn't live with mother and father	486	26.5	19.0	4.1	2	24
4	Mother with College Education	101	5.6	21.4	3.1	9	24
	Mother with High School Education	373	20.3	21.0	3.2	5	24
	Mother with 8th grade completed	372	20.4	19.9	3.6	7	24
	Mother with 4th grade completed	682	37.2	19.2	4.0	5	24
	Mother without schooling	286	15.5	17.6	4.3	2	24
	Don't know mother's educational attainment	19	1.0	17.7	5.2	4	24
5	Parents don't go to school meeting	87	4.7	17.5	4.2	7	24
	Parents go to school meeting	1736	95.3	19.7	3.9	2	24
6	More than three rooms in the house	537	29.2	20.3	3.6	3	24
	Less than three rooms in the house	1309	70.8	19.2	4.0	2	24
7	Lives with 5 or more persons	818	44.3	19.0	4.2	2	24
	Lives with 4 or less persons	1026	55.7	20.0	3.7	2	24
8	Child study less than one day a week	140	7.8	17.5	4.7	2	24
	Child study less one day or more a week	1642	92.2	19.8	3.8	2	24
9	Parents see children reading 3/4 days a week	727	40.3	20.3	3.5	7	24
	Parents see it less than 3 days a week	1078	59.7	19.1	4.1	2	24
10	Parents see their children playing 3/4 days a week	1523	85.3	19.6	3.9	2	24
	Parents see their children playing < 3 days a week	265	14.7	19.3	3.9	7	24
11	Male student	989	52.4	19.2	4.0	2	24
	Female student	897	47.6	19.8	3.9	3	24
12	White	999	54.5	19.9	3.9	2	24
	Brown	729	39.8	19.2	4.0	2	24
	Asian	12	0.7	20.0	2.9	17	24
	Indian	9	0.5	20.1	4.0	12	24
	Black	84	4.6	17.9	3.8	3	24
13	There is a quiet place for studying in the house	1576	86.2	19.8	3.8	2	24
	There isn't a quiet place for studying in the house	251	13.8	18.2	4.3	3	24
14	There is a computer with access to the internet	381	20.9	21.2	3.0	7	24
	There isn't a computer with access to the internet	1461	79.1	19.1	4.0	2	24
15	There are one or more DVD devices in the house	1554	84.9	19.7	3.9	2	24
	There aren't DVD devices in the house	278	15.1	18.4	4.3	2	24
16	Family has one or more automobiles	954	53.1	20.1	3.6	4	24
	Family has n't automobiles	839	46.9	18.9	4.2	2	24

Source: Questionnaire and Literacy Test (Provinha Brazil) applied to 2nd grade students (K-12 education) of Sertãozinho-São Paulo (and their parents).

The information obtained this way was more reliable than that obtained on SAEB and on Prova Brasil from fourth-graders. Out of the 1,986 students who participated in *Provinha Brasil* in Sertãozinho,

[9] With respect to school history, we collected information about the age at which children entered school and about which school they attended at each age. The questions on socioeconomic characteristics were based on the questionnaire of Prova Brasil 2005. The questionnaire applied in this study may be obtained from the authors upon request.

1,850 questionnaires were returned. This represents a loss of only 6.85%, which is much lower than that of Prova Brasil questionnaires in 2007, in which 18.71% of fourth-graders did not answer any of the questions. Regarding the question about mother's level of education, for example, only 1% of the questionnaires filled out by the parents had "I don't know" as the answer. Conversely, in the questionnaires related to Prova Brasil in 2007, this rate was as high as 30.43%.

According to the descriptive statistics of the variables used in the estimations in the present study (Table 1), the average of correct answers on *Provinha Brasil* 2008 in Sertãozinho amounted to 19.4 points. According to the Brazilian Ministry of Education, this score is lower than what is expected in terms of literacy (21 points or more). When we assess the results considering children's age at school admission, we note that the younger the children start attending school, the higher their literacy score. Recall that after the implementation of the nine-year Elementary School education, the age for school admission is 6 years. However, we verified that 84.2% of the students from Sertãozinho started attending school at the age of 5 years or less, allowing us to conclude that a significant share of the students had attended Early Child Education.

In regard to family arrangement, those children who live with their father and mother achieved an average score of 19.7 points. Among those who did not have such a family arrangement, the average score corresponded to 19.0 points. Family size was also important. Children from big families (5 people or more) had a lower average than those from smaller families (4 people or less), with an average difference of 1 point on literacy scores. As for the the participation of parents in their children's school life, we found a difference of 2.2 points between students whose parents attended the school meetings and those whose parents did not do so. Finally, as expected, the higher the mother's level of education, the higher the student's score.

Another interesting information concerns the time children devoted to studying and reading. Those who spend more time studying or reading had better scores than those who study or read less often. The difference in score corresponded to 2.3 points.

As far as the households where children live are concerned, we perceived that higher scores are associated with better socioeconomic conditions. Thus, students who live in a house with a larger number of bedrooms, with Internet access, a DVD player, and a car, had better scores than those students from socioeconomically underprivileged families. In addition, children whose households have a quiet place for studying had on average a score 1.6 points higher than those who do not have it.

4 ESTIMATION STRATEGY ON THE PRESENCE OF SELF-SELECTION BIAS

Average performance (in terms of literacy scores of *Provinha Brasil*) of students who started school at an earlier age is certainly higher than that of those who entered school later (Table 1). Notwithstanding, there is a consensus agreement in the literature that children with a better family background tend to start school earlier than those with less favorable social conditions. They self-select to receive treatment *S*, which in this case refers to earlier school admission.

Therefore, to obtain consistent estimates of earlier school admission on the literacy scores of *Provinha Brasil* from children aged 7 to 8 years, it is necessary to find an estimation strategy that corrects the self-selection bias inherent to the difference in averages between treatment and control groups.

Ideally, the real average effect of treatment on literacy could be known if it were possible to observe the same children in two distinct situations, one in which they had been enrolled at age s (treatment designated by S), and another one in which they had started school when they were l years old (control designated by L), where $s < l$. Thus, the average treatment effect (ATE) could be obtained by:

$$E\left[Y_i^S - Y_i^L\right] \quad (1)$$

where Y_i^S is the potential result of child *i* on *Provinha Brasil* if she belongs to treatment *S*, and Y_i^L is the potential result of this same child if she belongs to control group L.[10]

Nonetheless, it is not possible to observe the same individual in these two states simultaneously. But one can observe:

$$E\left[Y_i^S|S\right] - E[Y_i^L|L] \quad (2)$$

Adding and subtracting counterfactual $E[Y_i^L|S]$ in this equation, we have

$$E\left[Y_i^S|S\right] - E[Y_i^L|L] + E[Y_i^L|S] - E[Y_i^L|S] \quad (3)$$

rearranging

$$E\left[Y_i^S - Y_i^L|S\right] + E[Y_i^L|S] - E[Y_i^L|L] \quad (4)$$

The first term in (4) is the average treatment effect on treated subjects (ATT), and the two subsequent terms stand for the self-selection bias. The interest lies in ATT, but to estimate it, it is necessary to use an estimation method that rules out the self-selection bias. A very frequent solution described in the literature consists of social experiments based on treatment randomization for a selected set of individuals, resulting in a group of treated and untreated (control) subjects. By obtaining a perfect randomization of treated individuals, potential outcomes will be independent from the treatment status, and the self-selection bias will be null, i.e., $E[Y_i^L|S] - E[Y_i^L|L] = 0$. This way, equation 4 can be rewritten as

$$E\left[Y_i^S - Y_i^L|S\right] \quad (5)$$
$$= E\left[Y_i^S|S\right] - E[Y_i^L|L] = E\left[Y_i^S - Y_i^L\right]^{11}$$

Therefore, with a perfect randomization, it is possible to estimate the ATT by comparing the average outcomes of treated and untreated groups. However, in many cases, it is the individuals' interest to receive treatment, then it might be difficult to prevent them from self-selecting to participate in the treatment, especially in social experiments. In the case proposed in the present study, this difficulty is even more evident, as the age at which children are enrolled in school is determined by the characteristics (or preferences) of their families. Additionally, besides other difficulties related to the conduct of experiments,[12] follow-up time of observation units is also a hindrance. For example, if treatment refers to school admission at the age of 3 years, while control refers to school admission at a later age, it would take us about 4 years before we could assess the effects of treatment on literacy. This time period is too long, considering the lack of Brazilian studies on the topic and also the urgent necessity for novel studies that may contribute to the recent debate about the mandatory requirement of Early Child Education in Brazil.

In view of these arguments, a more appealing alternative for estimating the effects of Early Child Education on *Provinha Brasil* literacy scores is the use of non-experimental methods based on the hypothesis of selection according to observable characteristics. This can be done if the following assumption holds true: in a given set of observable characteristics X that determine the selection for treatment, potential outcomes do not depend on treatment status (RUBIN, 1977), that is

[10] The result observed can be denoted as $Y_i = Y_i^L(1 - D) + Y_i^S D$, where D is a variable that assumes value equal to 1 if individual *i* was submitted to treatment, and 0 otherwise.

[11] This last equality requires the *Stable Unit Treatment Value Assumption* (SUTVA) to be true. This means that the potential outcome of one unit can not be related to the treatment status of another units.

[12] Duflo, Glennerster & Kremer (2006) describe several setbacks related to the conduct of random social experiments. They highlight the difficulty in avoiding the contamination of the control sample by possible treatment externalities and by social interactions. They also mention that costs may be remarkably high depending on the study design.

$$\{Y_i^S, Y_i^L \perp S_i\}|X_i \text{ (Unconfoundness Assumption)}^{13} \quad (6)$$

In fact, this is a strong assumption, but we regard it as valid for the present study.

Nevertheless, note that if there are many covariates, it might be difficult to obtain cells with treatment and control groups in a sufficient amount to estimate the treatment effect.[14] An alternative proposed by Rosenbaum & Rubin (1983) to circumvent the dimensionality problem was the use of the propensity score ($p(X_i)$), which consists of a measure that combines individual characteristics into a single indicator with the same independence property between potential outcomes and treatment assignment

$$\{Y_i^T, Y_i^C \perp T_i\}|p(X_i) \text{ (Propensity Score Unconfoundness Assumption)}, \quad (7)$$

where $p(X_i)$ gives each individual i the probability of receiving treatment based on her characteristics X_i and allows treated and untreated individuals with similar indicators to be compared. Another requirement is that observable characteristics should not fully determine treatment status, that is

$$0 < P(S_i = 1|X_i) < 1 \quad (8)$$

In this study, the propensity score was obtained using a probit regression where the dependent variable S_i is equal to 1 if the individual is treated, i.e., if she started school at age s, and 0 if she started school at age l. Explanatory variables X_i that determine treatment were chosen based on two procedures: i) statistical significance; and ii) the "hit or miss" method.[15] The first procedure consists in selecting covariates from a larger set whose coefficients are statistically significant. The second one consists in developing an indicator equal to 1 if $\hat{p}(X_i) > p$, and 0 otherwise, where $\hat{p}(X_i)$ is the estimated probability of receiving treatment and p is the percentage of treated individuals. The larger the number of correct predictions obtained with the indicator developed in relation to dependent variable S_i which designates treatment, the better the model.

The subsequent step consisted in choosing a matching algorithm based on the predicted propensity score matching (PSM) to estimate the effect of treatment on treated subjects (ATT_X). In line with Heckman, Ichimura & Todd. (1997),[16] the method used as benchmark among the different options available was the Kernel Matching (with a bandwidth of 0.06 and Epanechnikov weighting function). Using PSM has an advantage comparing to OLS conditioned on covariates, it does not suppose a linear additive functional form. For the sake of comparison and robustness check of the results, ATT_X were also estimated by: 1) Ordinary Least Squares (OLS) with covariates, 2) OLS with the reciprocal of the Propensity Score as weights;[17] 3) OLS with the Propensity Score as covariate;[18] 4) Nearest Neighbor PSM with replacement; 5) Nearest Neighbor PSM without replacement; 6) Nearest 10 Neighbors PSM with replacement; 7) Radius PSM with caliper of 0.1; 8) Radius PSM with caliper of 0.001; 9) Radius PSM with caliper of 0.0001[19]; 10) PSM within Strata with five strata; 11) PSM within Strata with 10 strata.[20]

Concomitantly with the PSM estimations, we ran tests to check whether the covariates were balanced between the treatment and control groups (i.e., to check whether both groups were alike). Two tests proposed by Rosenbaum and Rubin (1985) have such purpose. The first one consists of a *t* test to

[13] The notation used in (6) – *S* for treatment and *L* for control – is not the same one used by Rubin (1977), as it seeks to maintain the notation that was previously employed in this study.
[14] See Angrist (1998) for an application of a non-parametric matching.
[15] See Breiman et al. (1984; apud HECKMAN, ICHIMURA & TODD, 1997)
[16] In fact, the authors use a Biweight weighting function.
[17] See Imbens (2004).
[18] See Imbens (2004).
[19] See Dehejia & Wahba (2002) for an application of this method.
[20] See Rosenbaum & Rubin (1983) and Dehejia & Wahba (2002).

determine the differences between the averages of treatment and control groups for each covariate before and after the matching. The second test is based on the calculation of standardized biases of a given covariate, also before and after the matching. This test is obtained by the ratio between the difference of covariate means of the treatment and control groups and the squared root of the average of the variances of the same covariate for the treatment and control groups. A significant reduction in the bias, such that the bias indicator after the matching is lower than 5%, indicates that the explanatory variable was properly balanced.[21]

Another way to improve covariate balance was by using a trimming rule. This rule guarantees that observation units outside the common-support region will be excluded, as well as the treatment or control units within the common-support region located on a given interval (bin) of the histogram with a frequency lower than *q*%.[22]

With the results obtained from *Provinha Brasil* and the answers to the socioeconomic questionnaires applied to students' parents, nine groups were constructed for PSM implementation, each one of them referring to different treatments and controls. This is well illustrated in Table 2. In Group 1, for instance, children who started school at the age of 5 years or before were the treated subjects while those who only started school at the age of 6 or later were the controls. As previously mentioned, given that the correct age for admission to Elementary School in Sertãozinho is 6 years, treatment in Group 1 refers to the attendance of Early Child Education for at least one year, i.e., having started school at the age of 5 years or earlier; while control refers to not attending this stage, i.e., having started school at the age of 6 years or later.

Table 2 – Description of the different groups of treated and untreated units used for the Propensity Score Matching procedure

	Treated units	***Untreated units***
Group 1	*S*: Children who entered school at ages 5 or less	*L*: Children who entered school at ages 6 or more
Group 2	*S*: Children who entered school at ages 3 or less	*L*: Children who entered school at age 6
Group 3	*S*: Children who entered school at age 4	*L*: Children who entered school at age 6
Group 4	*S*: Children who entered school at age 5	*L*: Children who entered school at age 6
Group 5	*S*: Children who entered school at ages 3 or less	*L*: Children who entered school at age 5
Group 6	*S*: Children who entered school at age 4	*L* Children who entered school at age 5
Group 7	*S*: Children who entered school at ages 3 or less	*L*: Children who entered school at age 4
Group 8	S^{*}: Children who entered school at ages 5 or less in another municipality	L^{*}: Children who entered school at ages 6 or more
Group 9	S^{**}: Children who entered a private school at ages 5 or less	L^{**}: Children who entered a public school at ages 5 or less

[*] This groups were built to evaluate the quality of the Early Child Education of Sertãozinho by comparing the scores of children who enter school at an early age in this municipality with the scores of children who enter school with the same age some place else.

[**] This groups were built to evaluate the quality of the Public Early Child Education of Sertãozinho by comparing the scores of children who enter school at an early age in the public school system with the scores of children who enter private schools with the same age.

From Groups 2 through 7, treatment *S* is always associated with a specific age at admission lower than the age at admission of children from control group *L*. All possible combinations based on this rule were employed. Differently, in Group 8, treatment refers to those students who started school at the age of 5 years or earlier in another town, whereas controls are those students who started school at the age of 6 years or later in Sertãozinho. This enabled the identification of the average difference of literacy scores between students who attended Early Child Education outside Sertãozinho and those who started Elementary School there. This strategy allows assessing the role of the quality of Early Child Education on students' literacy scores.

In Group 9, on the other hand, we tried to consider treatment and control units in such a manner as to distinguish the results between those who attended Early Child Education in public schools and those who did it in private schools. Treatment in this case consisted in starting school at the age of 5 years or

[21] Caliendo & Kopeinig (2005) provide the exact formula for the test. However, the maximum acceptable percentage of bias after the matching is not precisely known (the authors assert that 5% should suffice). So, it is important to analyze the two tests jointly.

[22] See Caliendo & Kopeinig (2005).

earlier in private schools while the control group included those who started school at the same age but in public schools. This econometric exercise is important to complement the estimation of the effect of earlier school admission on literacy. As occurred in Group 8, this exercise goes beyond the idea that earlier school admission alone is sufficient for the improvement of future school performance, as the objective is to gather evidence of the role of Early Child Education quality on literacy.

5 RESULTS

In this section, we present the econometric results obtained to assess the effect of Early Child Education on children's literacy. The variables that define the treatment and control groups used to capture this effect refer to the ages at which children were enrolled in school. These variables are more suitable to achieve the intended target, as they allow assessing whether there are striking differences between children who attended Early Child Education for one year and those who did it for more than one year.

First, in Table 3, we present the OLS estimates (with robust standard errors in parenthesis) of the effects of earlier school admission. As proposed by Rubin (1977), conditional on a set of covariates that define treatment, treatment variables are independent from potential outcomes, and an OLS estimation should produce unbiased estimates. We checked whether the fact that a child attended Early Child Education, i.e., if she started Early Child Education at the age of 5 years or earlier, has a positive effect on the literacy score (OLS 1, OLS2 and OLS3 specifications), comparatively to children who only started school at the age of 6 years or later. The OLS1 estimation coefficient represents the result that is not conditional on the child's observable characteristics, and therefore, it should be biased. In this case, we observed a literacy score 9.7% greater than that obtained by students who did not attend Early Child Education.

In OLS2 specification, it was possible to assess the effect of Early Child Education when we controlled for the set of covariates used to obtain the propensity score related to the probability of each student receiving treatment, which we call a "smaller set." Note that the magnitude of the coefficient of the variable "children who entered school at ages 5 or less" was lower than in OLS1 specification, producing a positive effect of 5.0% on the literacy score.

The magnitudes of dummy coefficients related to the age at which a child started school changed when we inserted the covariates used to obtain the propensity score (smaller set). In the OLS5 specification, we observed that children who started school at the age of 5, 4 or 3 years presented literacy scores 4.4%, 4.6% and 5.9% higher than those who entered school at the age of 6 years or later, respectively. The same was observed in the OLS6 specification, in which we used more covariates (full set).

The estimates based on Propensity Score Matching are shown in Table 4.[23] As mentioned in the previous section, our main results are based on Kernel Matching,[24] which are highlighted in gray. The other matching methodologies were implemented to check the robustness of the results.

By analyzing Group 1 (treatment and control), we observed that treatment effect is positive and significant. Children who started school at the age of 5 years or less, presented literacy scores 5.7% higher compared to those who only started school at the age of 6 years or later.

In Group 2, which is composed of children enrolled in school at the age of 3 years or less (treatment group) and of children enrolled at school at the age of 6 years (control group), we found a difference of 5.8% in the literacy scores between the treatment and control groups. However, if we consider the control group as children who were enrolled in school at the age of 5 years (Group 5), or at the age of 4 years (Group 7), there was no statistically significant difference in the literacy scores between the treatment and control groups.

[23] The standard errors of these estimates were calculated conventionally and not by bootstrapping. Abadie & Imbens (2006) show that standard errors are not valid if calculated by the bootstrapping method when Nearest Neighbor Matching is implemented. For the other matching algorithms it is not clear whether it is possible or not to apply this technique. We calculated these standard errors, which can be obtained from the authors upon request. However, we highlight that no important difference was verified.

[24] Table 1 of the Appendix shows the number of observations on and off common-support region obtained by Kernel Matching.

Table 3 – OLS estimates of the effects of Early Child Education on Provinha Brasil literacy scores (in logarithm) conditioning on covariates

	OLS 1	OLS 2	OLS 3	OLS 4	OLS 5	OLS 6
Children who entered school at ages 5 or less	0.097*** (0.019)	0.050** (0.021)	0.053** (0.021)			
Children who entered school at age 5				0.074*** (0.022)	0.044* (0.023)	0.048** (0.024)
Children who entered school at age 4				0.106*** (0.021)	0.046** (0.022)	0.050** (0.023)
Children who entered school at age 3 or less				0.103*** (0.021)	0.059*** (0.022)	0.060*** (0.023)
Covariates	no	yes (smaller set)	yes (full set)	no	yes (smaller set)	yes (full set)
R-squared	0.02	0.115	0.13	0.021	0.115	0.13
N	1,756	1,599	1,528	1,756	1,599	1,528

(1) The smaller set of covariates contains dummy variables equal to 1 if the following statement is true (and 0 otherwise): lives with mother and father; a set of dummy variables indicating mother education (College Education completed, High School Education completed, 8th grade completed, and 4th grade completed - the omitted comparison group are the children whose mothers have no education or doesn't have the 4th grade completed); parents go to school meeting; more than three rooms in the house ; lives with more than 5 persons; child study less than one day a week; parents see their children reading 3 or 4 days a week; parents see their child playing 3 or 4 days a week; male student ; black student.

(2) Besides the smaller set of covariates, the full set of covariates contains the following additional variables (equal to 1 if the following statement is true and 0 otherwise): there is a quiet place for studying in the house; there is a computer with access to the internet; there are one or more DVD devices in the house; family has one or more automobiles.

(3) The omitted category refers to those students who entered school at ages 6 or more.

On the other hand, if treatment refers to school admission occurred at the age of 4 years, while control refers to school admission at the age of 6 years (group 3), we observed a positive and significant effect of 6.3% in the literacy scores of treated children. However, this effect is not perceived in Group 6, in which the treated children are exactly the same students in Group 3, and control subjects are those students enrolled in school at the age of 5 years.

Group 4 refers to students who started school at the age of 5 years (treatment) and those who entered school at the age of 6 years (control). In this case, we observed that treated children had a literacy score 5.9% higher.

It is important to underscore that we carried out two additional exercises. One to evaluate the effect of a child starting school at the age of 5 years or less (i.e., submitted to early child education) in another municipality; and another one to measure the effect of a child having attended a private school at the same age. First, Group 8 presents the comparison between students who enrolled in school at the age of 5 years or less in another municipality (treatment) and those who enrolled in school at the age of 6 years (control) in Sertãozinho. Note that treated children had a literacy score 2.5% lower, but this difference was not statistically significant. This result indicates that the literacy scores obtained by students in each group are similar. It also indicates the fact that students had Early Child Education does not guarantee that they will have higher literacy scores, and thus it favors the choice of a municipality in which Early Child Education is well structured.

Group 9 included children admitted to a private school at the age of 5 years or less (treatment), and those who had Early Child Education in a public school (control). The results do not show significant differences in literacy scores between the two groups. This result confirms that Early Child Education provided by public schools is as good as that offered at private schools, suggesting that public schools, with the largest number of vacancies in Sertãozinho, have a good structure (considering that private schools offer appropriate quality standard).

Generally, the other Propensity Score Matching procedures produce very similar results to those of Kernel Matching. The exceptions are the estimates generated by Nearest Neighbor and Nearest Neighbor Matching without reposition, whose results were mostly nonsignificant. However, both methods work more adequately when the amount of control units is much larger than treated ones, which is not the case in this paper.

Table 4 – Estimates of the Average Treatment Effect on the treated for different groups of treated and control units

	GROUPS								
	1	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***
OLS - Reciprocal of Propensity Score as weights	0.033 (0.026)	0.057*** (0.022)	0.045* (0.025)	0.051** (0.024)	0.019 (0.016)	0.000 (0.018)	0.017 (0.015)	-0.023 (0.060)	-0.086 (0.098)
OLS with the Propensity Score as covariate	0.048** (0.021)	0.061*** (0.022)	0.05** (0.025)	0.052** (0.024)	0.016 (0.017)	0.000 (0.018)	0.015 (0.014)	-0.033 (0.046)	0.004 (0.020)
Epanechnikov Kernel bandwith (0.06)	0.057** (0.026)	0.058** (0.026)	0.063** (0.028)	0.059** (0.026)	0.022 (0.017)	-0.003 (0.018)	0.015 (0.015)	-0.025 (0.050)	0.011 (0.028)
Nearest Neighbor with replacement	0.022 (0.036)	0.015 (0.032)	0.048 (0.046)	0.049 (0.037)	0.030 (0.029)	0.075** (0.034)	0.043 (0.032)	-0.015 (0.056)	0.031 (0.040)
Nearest Neighbor without replacement	-0.016 (0.029)	0.044* (0.026)	0.046* (0.026)	0.045* (0.025)	0.003 (0.019)	-0.002 (0.019)	0.007 (0.014)	0.011 (0.054)	0 .003 (0.022)
Nearest 10 Neighbors with replacement	0.053** (0.029)	0.055** (0.027)	0.076*** (0.029)	0.060** (0.027)	0.018 (0.018)	-0.005 (0.019)	0.012 (0.016)	-0.015 (0.047)	0.001 (0.023)
Radius (Caliper δ = 0.1)	0.068*** (0.024)	0.063** (0.025)	0.066** (0.027)	0.065*** (0.025)	0.021 (0.017)	0.000 (0.019)	0.014 (0.015)	-0.016 (0.049)	0.016 (0.026)
Radius (Caliper δ = 0.001)	0.051* (0.029)	0.058* (0.032)	0.078** (0.037)	0.049 (0.031)	-0.002 (0.021)	0.004 (0.022)	0.025 (0.017)	-0.026 (0.058)	0.015 (0.029)
Radius (Caliper δ = 0.0001)	0.062** (0.030)	0.042 (0.035)	0.072* (0.040)	0.065* (0.035)	0.005 (0.022)	0.005 (0.023)	0.008 (0.018)	-0.043 (0.058)	-0.010 (0.031)
Matching within stratum (5 strata)	0.052 (0.042)	0.057 (0.054)	0.057* (0.032)	0.048 (0.066)	0.031 (0.053)	0.016 (0.044)	0.014 (0.029)	-0.003 (0.036)	0.007 (0.048)
Matching within stratum (10 strata)	0.049 (0.048)	0.058 (0.068)	0.063 (0.062)	0.044 (0.080)	0.037 (0.069)	0.026 (0.053)	0.028 (0.060)	-0.015 (0.109)	0.025 (0.062)

(1) The propensity score is estimated using the probit method. The covariates included in the equation are all dummy variables equal to 1 if the following statement is true (and 0 otherwise): lives with mother and father; a set of dummy variables indicating mother education (College Education completed, High School Education completed, 8th grade completed, and 4th grade completed - the omitted comparison group are the children whose mothers have no education or doesn't have the 4th grade completed); parents go to school meeting; more than three rooms in the house ; lives with more than 5 persons; child study less than one day a week; parents see their children reading 3 or 4 days a week; parents see their child playing 3 or 4 days a week; male student ; black student. The results of the probit estimations can be obtained upon request.

(2) The group 9 includes the following additional covariates in the probit model (equal to 1 if the following statement is true and 0 otherwise): there is a quiet place for studying in the house; there is a computer with access to the internet; there are one or more DVD devices in the house; family has one or more automobiles. This was necessary because of the Treatment status in this case, given by the students enrolled at private schools at ages of 5 or less, opposed to those enrolled at public school with the same age (control units). We had to include more covariates that discriminate family income to obtain a better matching.

Finally, the results obtained show that children who attended at least one year of Early Child Education (i.e. enrolled in school at the age of 5 years or less) presented a higher literacy score than those who did not attend this stage. No differences were observed between those who attended one or more years of Early Child Education, which leads to the conclusion that one year of Early Child Education, offered by a well structured education network such as that in Sertãozinho, may be enough for children to be practically literate by the age of 7 or 8 years. This does not mean that there are no skill differences. In this paper only a limited set of skills was assessed: those skills that determine literacy. Enrollment in school at the age of 3 years or less, for instance, may contribute to the development of other important skills not contemplated by *Provinha Brasil*.

6 TESTING

In what follows, we present the tests carried out to verify matching quality. This process was done simultaneously with the propensity score estimation, so that the final specification could show an adequate covariate balance of treatment and control units.

Table 5 – Tests of difference between the covariates of treated and control groups after Matching

		GROUPS								
		1	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***
lives with mother and father	Treated	0.76	0.69	0.81	0.77	0.71	0.81	0.70	0.80	0.85
	Control	0.73	0.66	0.81	0.77	0.69	0.82	0.72	0.78	0.81
	Bias (%)	6.3	6.8	-0.7	0.9	3.1	-1.4	-4.7	4.9	11.7
	Diff p>|t|	0.11	0.29	0.91	0.91	0.62	0.81	0.46	0.78	0.27
Mother with College Education	Treated	0.00	0.00	0.00	0.00	0.10	0.04	0.10	0.04	0.33
	Control	0.00	0.00	0.00	0.00	0.08	0.04	0.11	0.04	0.34
	Bias(%)	-1.8	0.0	0.0	0.0	7.7	-0.9	-3.0	0.0	-3.2
	Diff p>|t|	0.17	.	.	.	0.27	0.89	0.69	1.00	0.84
Mother with High School Education	Treated	0.25	0.27	0.26	0.18	0.24	0.27	0.24	0.27	0.33
	Control	0.19	0.24	0.23	0.18	0.25	0.27	0.25	0.27	0.30
	Bias (%)	16.3	7.9	8.8	1.2	-1.7	-1.6	-1.0	1.0	6.1
	Diff p>|t|	0.00	0.30	0.25	0.90	0.80	0.81	0.87	0.97	0.60
Mother with 8th grade completed	Treated	0.24	0.25	0.27	0.18	0.22	0.24	0.22	0.13	0.13
	Control	0.28	0.26	0.31	0.18	0.23	0.24	0.23	0.13	0.12
	Bias (%)	-8.9	-2.5	-12.0	1.5	-1.6	-0.4	-0.7	0.4	0.7
	Diff p>|t|	0.05	0.72	0.11	0.86	0.80	0.96	0.90	0.98	0.94
Mother with 4th grade completed	Treated	0.38	0.33	0.38	0.44	0.30	0.36	0.30	0.42	0.16
	Control	0.40	0.35	0.36	0.44	0.30	0.35	0.29	0.41	0.18
	Bias (%)	-2.8	-3.6	3.8	-0.2	-0.6	2.2	1.9	0.9	-2.9
	Diff p>|t|	0.49	0.57	0.56	0.98	0.92	0.73	0.75	0.96	0.76
Parents go to school meeting	Treated	0.03	0.05	0.02	0.05	0.04	0.02	0.04	0.04	0.01
	Control	0.03	0.04	0.02	0.06	0.04	0.02	0.04	0.04	0.01
	Bias (%)	1.3	2.9	0.3	-3.4	1.6	2.7	1.7	-2.3	-0.6
	Diff p>|t|	0.60	0.55	0.93	0.61	0.78	0.59	0.81	0.88	0.93
More than three rooms in the house	Treated	0.28	0.30	0.29	0.24	0.35	0.33	0.35	0.15	0.59
	Control	0.30	0.29	0.33	0.25	0.34	0.32	0.35	0.13	0.58
	Bias (%)	-4.1	2.4	-7.2	-2.7	0.9	1.6	-1.0	3.2	2.2
	Diff p>|t|	0.34	0.73	0.31	0.75	0.89	0.81	0.88	0.86	0.86
Lives with 5 or more persons	Treated	0.42	0.39	0.44	0.46	0.37	0.43	0.37	0.36	0.22
	Control	0.43	0.39	0.43	0.46	0.37	0.40	0.36	0.36	0.24
	Bias (%)	-0.4	-0.2	2.3	1.3	-0.2	4.5	1.4	0.3	-4.2
	Diff p>|t|	0.91	0.98	0.73	0.87	0.98	0.48	0.82	0.99	0.68
Child study less than one day a week	Treated	0.05	0.06	0.06	0.05	0.06	0.05	0.06	0.04	0.02
	Control	0.07	0.07	0.09	0.06	0.06	0.04	0.06	0.05	0.02
	Bias (%)	-5.0	-1.1	-8.9	-4.2	-1.9	2.8	-1.0	-3.4	-2.8
	Diff p>|t|	0.11	0.83	0.11	0.50	0.76	0.64	0.87	0.80	0.74
Parents see their children reading 3 or 4 days a week	Treated	0.41	0.40	0.45	0.37	0.41	0.45	0.41	0.38	0.48
	Control	0.40	0.39	0.43	0.38	0.41	0.43	0.42	0.35	0.48
	Bias (%)	2.5	0.8	4.4	-3.8	1.1	3.8	-0.6	6.9	0.3
	Diff p>|t|	0.55	0.90	0.52	0.65	0.86	0.56	0.92	0.72	0.98
Parents see their children playing 3 or 4 days a week	Treated	0.87	0.86	0.87	0.88	0.86	0.88	0.86	0.82	0.93
	Control	0.84	0.84	0.83	0.88	0.88	0.89	0.86	0.82	0.91
	Bias (%)	7.6	6.2	9.8	1.6	-3.5	-2.4	-0.2	0.5	6.1
	Diff p>|t|	0.04	0.31	0.12	0.83	0.57	0.70	0.97	0.98	0.53
Male student	Treated	0.53	0.51	0.53	0.57	0.50	0.53	0.50	0.58	0.53
	Control	0.54	0.50	0.54	0.57	0.51	0.55	0.50	0.58	0.58
	Bias (%)	-1.6	1.1	-0.4	-0.5	-2.8	-3.7	-0.1	0.7	-10.8
	Diff p>|t|	0.70	0.86	0.95	0.95	0.64	0.56	0.99	0.97	0.34
Black student	Treated	0.03	0.05	0.02	0.03	0.04	0.03	0.04	0.02	0.01
	Control	0.04	0.06	0.03	0.03	0.05	0.03	0.04	0.01	0.01
	Bias (%)	-2.4	-3.0	-3.4	-0.8	-2.6	0.8	2.9	1.7	-2.3
	Diff p>|t|	0.44	0.60	0.49	0.90	0.69	0.89	0.63	0.88	0.72

(1) The dots indicate that all observations were excluded after matching because the treated and control groups could not be balanced
(2) The tests for the covariates of Group 9 include additional variables (equal to 1 if the following statement is true and 0 otherwise) such as: there is a quiet place for studying in the house; there is a computer with access to the internet; there are one or more DVD devices in the house; family has one or more automobiles. All tests show that the covariates are balanced, but we decide to omit the results for these additional variables to save space. The results of the tests can be obtained on request.

As mentioned in Section 4, two types of tests were performed to verify whether the covariates that determine the treatment (Groups 1 through 9)[25] were balanced between treated and control subjects. One of them consists of the difference in covariate means between treated and control individuals, and the other one consists of the calculation of bias reduction, both carried out after the matching. The differences between covariate means and the standardized bias calculated before the matching are not reported in Tables 5 and 6, but they may be obtained upon request. Both tests are based on Kernel PSM, defined as the main method in this paper. Figures 1 through 18 show additional tests used to analyze matching quality. These figures represent Kernel density functions of the estimated propensity scores (with Epanechnikov weighting function, to be consistent with the strategy used in this paper) before and after the matching.

The first column of Table 5 show the tests for Group 1. Considering a 5% maximum acceptable rate for standardized bias, and a p-value of at least 10% for the difference of means *t* test after the matching, four variables could not be balanced according to the two tests: Lives with mother and father; Mother with High School Education; Mother with 8th grade completed; and Parents see their children playing 3 or 4 days a week. Even after attempts to include cross variables and the application of trimming rules with many values, these variables could not be balanced.[26] However, the bias was substantially reduced for these variables (on average 54.5%).[27] In addition, Figures 1 and 2 clearly show that densities after the matching overlap, indicating that the procedure was successful, despite the fact that the tests demonstrate that the four variables mentioned above could not be balanced.

In the second column of Table 5, referring to treatment and control Group 2, three variables presented standardized bias greater than 5% after the matching: Lives with mother and father; Mother with High School Education, and Parents see their children playing 3 or 4 days a week. However, the p-value of the *t* test for all covariates was higher than 10%, and Figures 3 and 4 show that density functions overlap almost perfectly after the matching, indicating an adequate balance.

For Group 3 (third column of Table 5), the standardized bias test indicated five covariates with bias greater than 5%: Mother with High School Education; Mother with 8th grade completed; More than three rooms in the house; Child studies less than one day a week; and Parents see their children playing 3 or 4 days a week. But, in none of these cases did the *t* test reject the null hypothesis of equality between covariate averages of treatment and control groups. Moreover, Figures 5 and 6 show a practically perfect overlapping of densities after the matching.

In group 4, both the t test and the standardized bias test demonstrate adequate balance of the explanatory variables of the regression necessary to obtain the propensity scores. Figures 7 and 8 show that density largely overlaps, which allows us to conclude that the variables could be properly balanced.

The only variable that presents a standardized bias greater than 5% in Group 5 is "Mother with College Education". On the other hand, the result of the *t* test shows that the null hypothesis of equality between the mean values of the covariate of treated and control units cannot be rejected. All the other variables were balanced after the matching according to the two tests used and with figures 9 and 10.

In groups 6 and 7, the results of the tests for covariate balance were satisfactory in all cases. Figures 11 through 14 confirm improved compatibility after the matching. On the other hand, in Group 8, the variable "Parents see their children reading 3 or 4 days a week" presented a standardized bias greater than 5% (equal to 6.9%). But the *t* test did not reject the null hypothesis of equality between the means of the control and treatment units, and Figures 15 and 16 also demonstrate that matching was adequate.

In Group 9, Lives with mother and father; Mother with High School Education; Parents see their children playing 3 or 4 days a week; and Male student[28] presented a standardized bias greater than 5%,

[25] See Table 2.

[26] The final result estimated for Group 1 was based on a trimming rule of 10% without crossed variables. This yielded the best results for covariate balance.

[27] This result can obtained from the authors upon request.

[28] Two of four covariates of Group 9 not reported in Tables 5 (see note 2 in this table to know all the variables whose tests were not reported) presented standardized bias greater than 5%: There are one or more DVD players in the household (bias of 5.9%); family owns one or more cars (bias of 7.2%). However, the p-values of the *t* tests performed for these variables were all smaller than 10%, indicating no significant differences between treatment and control groups.

and therefore, they did not have an adequate balance according to this criterion. However, considering the *t* test, the hypothesis of equality between the means of the variables for the treatment and control groups could not be rejected. Figures 17 and 18 show that density functions remarkably overlapped after the matching.[29]

Figure 1 - Kernel density before matching - Group 1: Treated (Age 5 or less) vs Control (Age 6 or more)

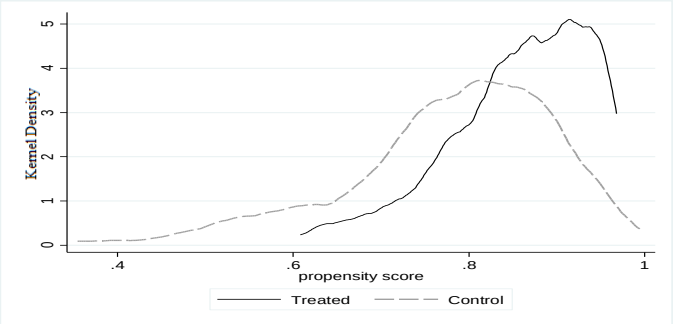


Figure 2- Kernel density after matching - Group 1: Treated (Age 5 or less) vs Control (Age 6 or more)

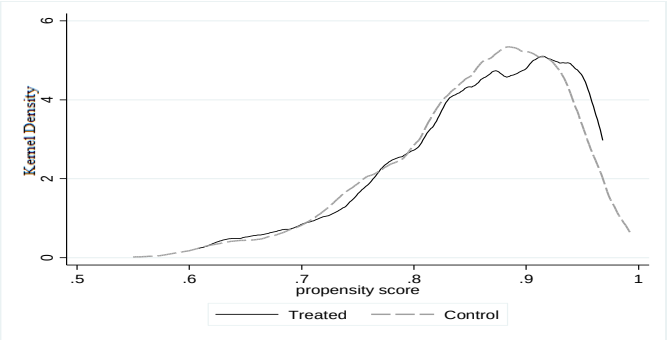


Figure 3 - Kernel density before matching - Group 2: Treated (Age 3 or less) vs Control (Age 6)

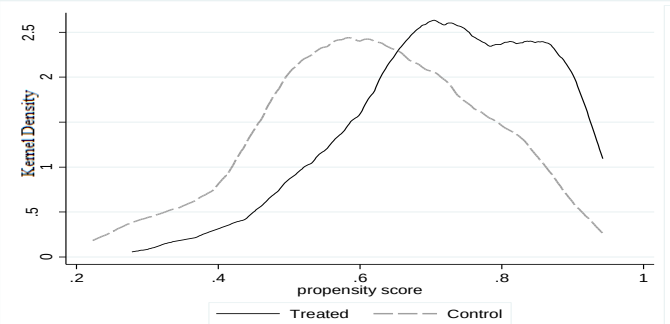


Figure 4 - Kernel density after matching - Group 2: Treated (Age 3 or less) vs Control (Age 6)

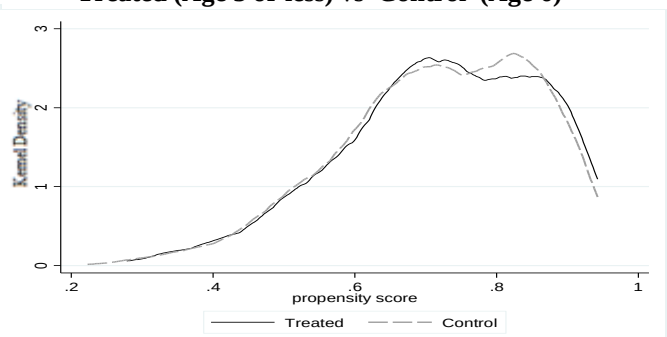


Figure 5 - Kernel density before matching - Group 3: Treated (Age 4) vs Control (Age 6)

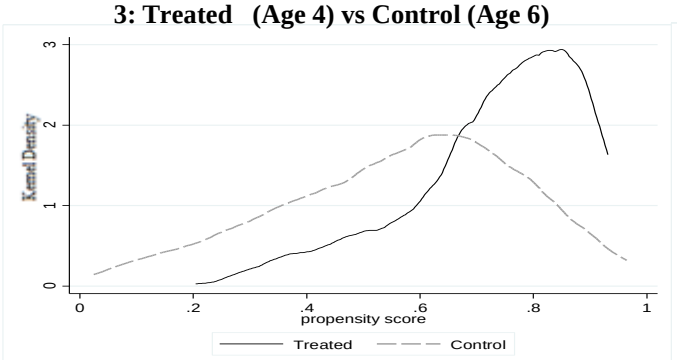


Figure 6 - Kernel density after matching - Group 3: Treated (Age 4) vs Control (Age 6)

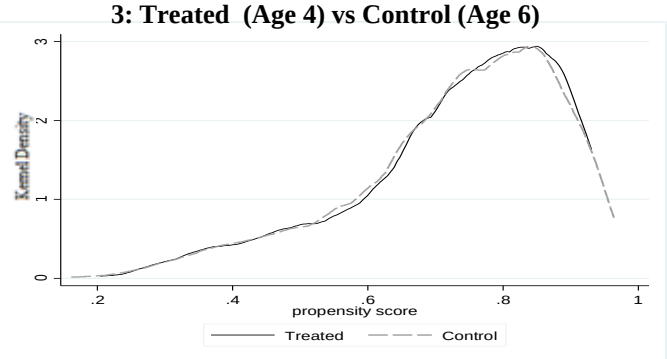


Figure 7 - Kernel density before matching - Group 4: Treated (Age 6) vs (Control: Age 5)

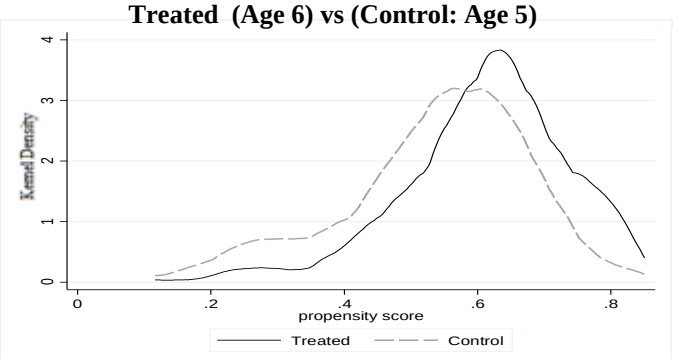
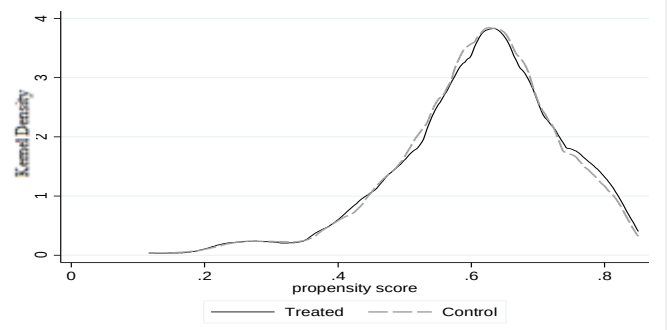


Figure 8 - Kernel density after matching - Group 4: Treated (Age 6) vs (Control: Age 5)



[29]Figures 17 and 18 have different scales. If they are placed on the same scale, it is possible to observe intense overlapping of the density function of the propensity score after the matching.

Figure 9 - Kernel density before matching - Group 5: Treated (Age 5) vs (Control: Age 3 or less)

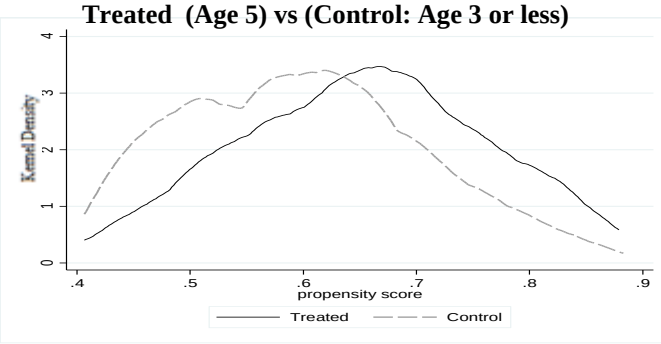


Figure 10 - Kernel density after matching - Group 5: Treated (Age 5) vs (Control: Age 3 or less)

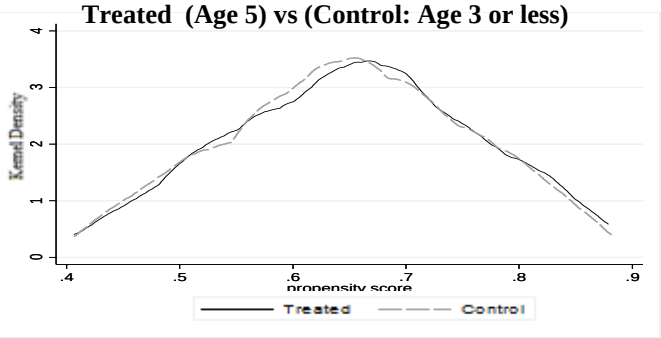


Figure 11 - Kernel density before matching - Group 6: Treated (Age 5) vs (Control: Age 4)

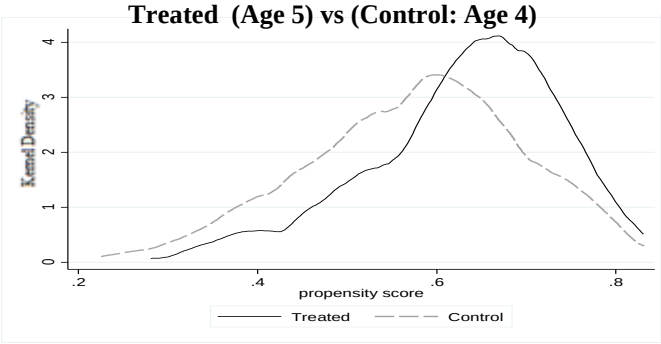


Figure 12 - Kernel density after matching - Group 6: Treated (Age 5) vs (Control: Age 4)

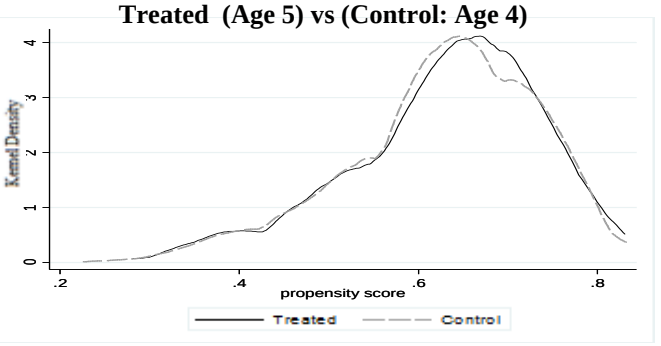


Figure 13 - Kernel density before matching - Group 7: Treated (Age 4) vs (Control: Age 3 or less)

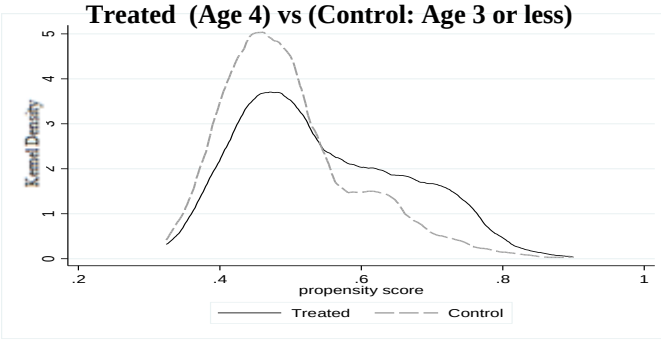


Figure 14 - Kernel density after matching - Group 7: Treated (Age 4) vs (Control: Age 3 or less)

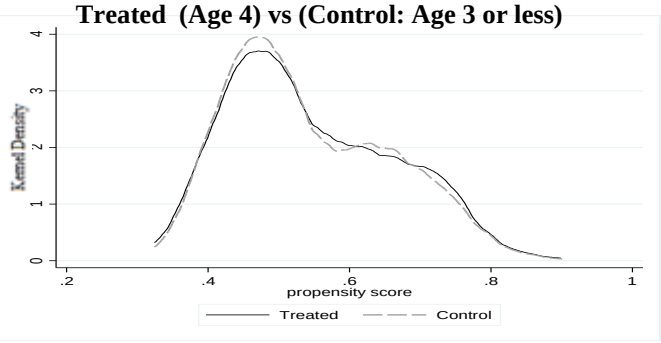


Figure 15 - Kernel density before matching - Group 8: (Treated: Age 5 - another municipaly) vs (Control: Age 6 or more in the municipality)

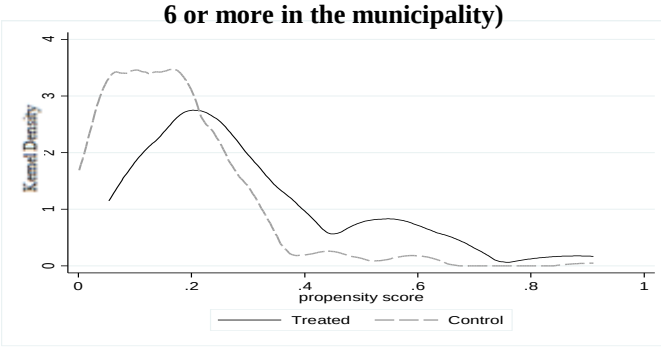
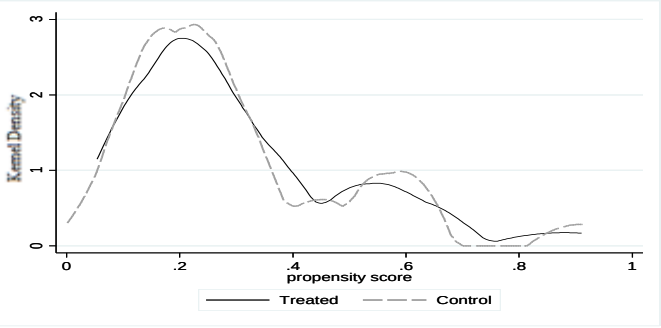


Figure 16 - Kernel density after matching - Group 8: (Treated: Age 5- another municipaly) vs (Control: Age 6+ in the municipality)



Finally, we should highlight that the test results suggest good covariate balance. This means that the treatment and control groups are very similar after the matching, attaching high reliability to the results estimated in Section 5.

Figure 17 - Kernel density after matching - Group 9: (Treated: Age 5- private school) vs (Control: Age 5- public school)

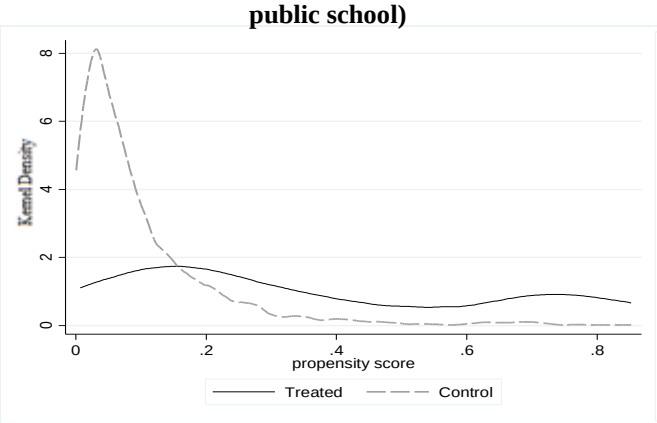
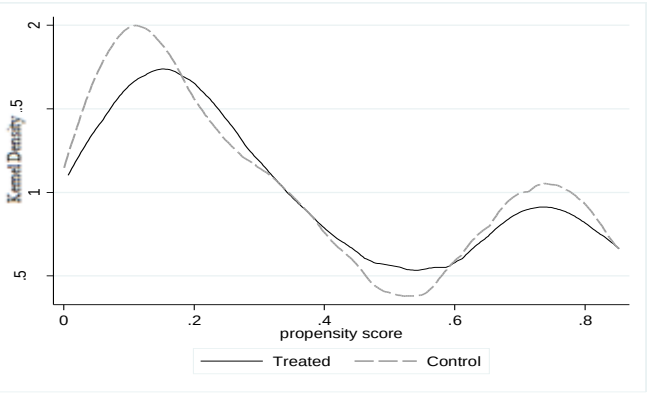


Figure 18 - Kernel density before matching - Group 9: (Treated: Age 5- private school) vs (Control: Age 5- public school)



7 FINAL REMARKS

This paper aimed to identify the effects of Early Child Education on children's literacy scores. To do that, we used data from *Provinha Brazil* applied in Sertãozinho to students attending the 2nd grade of Elementary School plus a socioeconomic questionnaire answered by the parents. The main contribution of this paper to the literature was to explore this new assessment tool proposed by the Brazilian Ministry of Education. In addition, to our knowledge, there are no studies in the literature, at least not in Brazil, that evaluate the relationship between earlier school admission and literacy scores.

The results obtained in this paper with Propensity Score Matching (and also with OLS) demonstrate that students who started school at the age of 5 years or less had higher literacy scores than those who started school at the age of 6 or later. In general, students who started school at the age of 5, 4 and 3 years or less obtained literacy scores around 6% higher than those who started school at the age of 6 or later. However, no significant differences were seen between students who attended one year of Early Child Education and those who attended for a longer period.

Indeed, there might be a problem with external validity of these results. Although the study used a non-experimental method to estimate the treatment effect, it has the same problems as experiments or non-experimental studies with treatment and control groups that belong to one locality. The effect of Early Child Education on the literacy of students all over Brazil can be even more pronounced. Therefore, its part of our research agenda to assess these effects in other municipalities in order to verify whether the results obtained in each one of them are similar. But we believe that the results obtained in this work constitute a good guidance for the implementation of public policies.

Finally, we underscore that this discussion is very relevant to the Brazilian case. If investments in education in early childhood are essential to skill development later in life, it is necessary to invest more heavily in Early Child Education (before the age of 6 years) in order to improve the quality of education in Brazil. However, school attendance rates in Brazil are not universalized yet, and expenditures per student are much lower than those verified in developed countries. This stresses that a lot more effort should be put in to expand the coverage and improve the quality of this stage of education.

REFERENCES

ABADIE, A ; IMBENS, G. W . 2006. On the Failure of the Bootstrap for Matching Estimators Issued in June *NBER Technical Working Paper* No. 325

ALBERNAZ, A.; FERREIRA, F. H. G.; FRANCO, C. 2002. Qualidade e Eqüidade na Educação Fundamental Brasileiro. *Pesquisa e Planejamento Econômico*, v. 33 No .3.

ANGRIST, J., 1998. Estimating the labor market impact of voluntary military service using Social Security data on military applicants. *Econometrica* 66, 249–88.

BECKER, G. 1964. Human Capital: a theoretical and empirical analysis, with special reference to education. Nova Iorque: NBER/Columbia University Press.

BEHRMAN, J.; CHENG, Y.; TODD; P. 2004. Evaluating Pre-school Programs when Length of Exposure to the Program Varies: A Nonparametric Approach. *Review of Economics and Statistics*. vol.86, n.1, p.108-32.

BERLINSKI, S., GALIANI, S., GERTLER, P. 2009. The effect of pre-primary education on primary school performance. Journal of Public Economics, Volume 93, Issues 1-2, February 2009, Pages 219-234.

CALIENDO, M.; KOPEINIG, S. 2005. Some practical guidance for the implementation of propensity score matching. Bonn, Germany: Institute for the Study of Labor (IZA), (IZA Discussion Papers, 1588)

CARNOY, M.; LOEB, S. 2002. Does external accountability affect student outcomes? A Cross-State Analysis. Educational Evaluation and Policy Analysis, 24(4), 305–331.

COLEMAN, J. S. et al. 1966. Equality of educational opportunity. Washington: Government Printing Office.

CUNHA, F.; HECKMAN, J.; LOCHNER, L.; MASTEROV, D. 2005. Interpreting the Evidence on Life Cycle Skill Formation. NBER Working Paper 11331, Cambridge, MA.

CURI, A., MENEZES-FILHO, N. 2006. Os efeitos da pré-escola sobre os salários, a escolaridade e a proficiência escolar. Anais do XXXIV Encontro da Anpec, Salvador, Bahia, p. 1-21.

CURRIE, J.; THOMAS, D. 1995. Does Head Start Make a Difference? The American Economic Review. vol. 85, n. 3 (Jun., 1995), p. 341-364

CURRIE, J. 2001. Early Childhood Education Programs. Journal of Economic Perspectives, vol.15, n.2, p.213-238.

DEHEJIA, R., and SADEK W. (2002). Propensity Score Matching Methods for Nonexperimental Causal Studies, National Bureau of Economics Research Working Paper No. 6829, forthcoming Review of Economics and Statistics.

DUFLO, E.; GLENNERSTER, R.; KREMER, M. 2006. Using Randomization in Development Economics Research: A Toolkit. MIT Department of Economics Working Paper No. 06-36

FELICIO, F.; FERNANDES, R. 2005. O Efeito da Qualidade da Escola sobre o Desempenho Escolar: uma avaliação do ensino fundamental no Estado de São Paulo. In: ENCONTRO NACIONAL DE ECONOMIA, 33., 2005, Salvador.

FELÍCIO, F.; VASCONCELLOS, L. 2007. O Efeito da Educação Infantil sobre o desempenho escolar medido em exames padronizados. Anais do XXXV Encontro da Anpec, Recife, Pernambuco, p. 1-22.

FRIEDLANDER, D.; PHIL, R. 1995. Evaluating Program Evaluations: New Evidence on Commonly Used Non Experimental Methods. American Economic Review, Vol.85: 923-937.

GRAY, S.W.; RAMSEY, B.; KLAUS, R. 1983. From 3 to 20: The Early Training Project in As the Twig is Bent…Lasting Effects of Preschool Programs. Consortium for Longitudinal Studies, ed. Hillsdale NJ.: Erlbaum, p. 171-200.

HECKMAN, J. J.; ICHIMURA, H.; TODD, P. E, 1997. Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme, Review of Economic Studies, Blackwell Publishing, vol. 64(4), pages 605-54, October.

IMBENS, G. 2004. Nonparametric estimation of average treatment effects under exogeneity: a review. The Review of Economics and Statistics, 86 (1) fev 2004.

JACOB, B. A., 2005. Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools, Journal of Public Economics, Elsevier, vol. 89(5-6), pages 761-796, June.

MICHALOPOULOS, C.; BLOOM, H.; HILL, C., 2004. Can Propensity Score Methods Match the Findings from a Random Assignment Evaluation of Mandatory Welfare-to-Work Programs? Review of Economics & Statistics Vol.86(1):156-179.

RESTUCCIA D.; URRUTIA C. 2004. Intergenerational Persistence of Earnings: The Role of Early and College Education. The American Economic Review, Vol. 94, No. 5 (Dec., 2004), p. 1354-1378.

RIVKIN, S. G., E. A. HANUSHEK, AND J. F. KAIN, 2005. Variable Definitions, Data, and Programs for Teachers, Students, and Academic Achievement. Econometrica Supplementary Material, 73, 2.

ROSENBAUM, P., RUBIN, D. 1983. The Central Role of the Propensity Score in Observational Studies for Casual Effects. Biometrika. n.70, p.41-55.

RUBIN, D. 1977. Assignment to a Treatment group on the basis of a Covariate. Journal of Educational Statistics. n. 2, p.1-26.

SCHADY, N. 2006. Early childhood development in Latin America and the Caribbean. World Bank Policy Research Working Paper 3869, March.