

Early childhood education effect on children's vocabulary [♦]

Guilherme Issamu Hirata¹

Abstract

This paper investigates the effect of an Early Childhood Education program on children's vocabulary. Using 2016 data from Petrolina, we compare children attending school to children not enrolled at the time of data collection. To account for selection bias, information from a parents' assessment is used to create control variables associated with characteristics usually not observed by the researcher that are potentially correlated with children's enrollment status and child development. Results show positive and statistically significant impacts on expressive vocabulary. There is also evidence that the program is more effective for children with lower reading exposure at home.

Keywords

Early childhood education. Expressive vocabulary. Receptive vocabulary. Child development.

Resumo

Este artigo analisa o efeito de um programa de Educação Infantil sobre o vocabulário das crianças. Usando dados de Petrolina de 2016, o artigo compara crianças que frequentavam escola com crianças fora da escola. Para lidar com o viés de seleção, informações coletadas junto aos pais são utilizadas para criar variáveis de controle associadas a características geralmente não observáveis e que são potencialmente correlacionadas com matrícula e desenvolvimento infantil. Os resultados mostram impactos positivos sobre vocabulário expressivo. Ademais, há evidência de que o programa é mais efetivo para crianças com menos exposição à leitura em casa.

Palavras-Chave

Educação infantil. Vocabulário expressivo. Vocabulário receptivo. Desenvolvimento infantil.

Classificação JEL

I28. J13

[♦] I'm grateful to João Batista Oliveira for the ideas and comments on the research design and earlier versions of this paper. This paper also benefited from comments from Marcos Nakaguma (its editor) and one anonymous referee. For financial support, I thank Instituto Alfa e Beto. I am thankful to the local authorities and principals from Petrolina (PE) for the cooperation and assistance provided during the project. I also wish to thank Walfrido Neto and Williana Melo for conducting the field work implementation. The author was responsible for the research design and supervision of data collection.

¹ Pesquisador IDados – End.: R. Visconde de Pirajá, 470, 701 – CEP: 22410-002 – Rio de Janeiro/RJ Brasil – E-mail: guilherme@idados.id – ORCID: <https://orcid.org/0000-0003-0661-2414>.

Recebido: 18/10/2019. Aceito: 03/10/2020.

Editor Responsável: Marcos Yamada Nakaguma



Esta obra está licenciada com uma Licença Creative Commons Atribuição-Não Comercial 4.0 Internacional.

1. Introduction

Previous research presents evidence of long-term benefits from children's language development. In order to have language developed appropriately, there is a consensus that most language development should occur by the age of five, i.e., before children enter formal education (see, for instance, Durkin 1966; Hart and Risley 1995; Cunningham and Stanovich 1997; Sénéchal and LeFevre 2002; Camilli *et al.* 2010).

In the process of children becoming effective communicators, receptive and expressive vocabulary play key roles. Receptive vocabulary refers to the ability to receive, decode and interpret language. Expressive vocabulary involves the ability to use language to produce a message (McIntyre *et al.* 2017). In this sense, one should expect different interventions to have different impacts on each ability (Sénéchal, 1997).

There are many actors and ways through which children learn and develop language abilities, which means that there are many different types of intervention that could and have been implemented to better understand and foster language development. Sénéchal (1997), for instance, investigated the differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. Passive listening to stories contributes to development of receptive vocabulary, whereas interactive reading and question-answer interactions promote expressive reading. The former is more common in home and informal settings, the latter in schools with well-prepared teachers.

This paper uses data from Petrolina (PE) to analyze the effect of an early childhood education (ECE) program delivered at day-care and preschool units (henceforth "schools") on children's vocabulary. Based on data collected in mid-2016, the empirical strategy was to compare the level of expressive and receptive vocabulary of children enrolled in schools with that of children not attending school.¹ This strategy implies that we assume school attendance and exposure to the ECE program are the same treatment. Proxies for unobservable factors were used to mitigate possible selection bias. Among them are measures for parents' confidence in their own attitudes towards parenting and for cognitive stimulation at home. Three matching-based techniques are used as well, including the Inverse

¹ We use enrolled and attending interchangeably.

Probability Weighting proposed by Abadie (2005). The sample included children born in 2012.

The results show that children exposed to the ECE program have an expressive vocabulary level that is much higher than that observed for children out of school. The effect ranges from 0.20 to 0.43 of a standard deviation. Positive results ascend in all specifications tested at 10% significance level. There is also evidence that school attendance is more effective for children with lower reading exposure at home, conditional on socioeconomic status and other controls. On the other hand, there are no significant differences between the groups in terms of receptive vocabulary.

The results cannot be generalized directly, i.e., the estimated effect refers specifically to the municipality of Petrolina and its socioeconomic and institutional context. Because there is a lot of variation in school curricula across municipalities in Brazil, the school effect can vary as well. To give an idea about the context, at the time of data collection, the day-care and preschool programs in Petrolina were part of a broader ECE program aimed at 0-6-year-old children. Called Nova Semente (New Seed), the program, which started in 2010, had about 150 units attending approximately 60 children each, covering 9,000 children in 2016 (the data collection year). The program offered care for up to 10 hours a day, with part of the curricula focusing on adult-child interactions and language and executive functions development. Another distinct aspect of the Nova Semente initiative was that unit implementation followed the initiative of the local community; the school managers were usually appointed by the community, while the local government supervised the pedagogical program implementation.

This paper adds to the literature that investigates the impact of early childhood education on children's outcomes. For instance, Berlinski, Galiani, and Gertler (2009) show that an expansion of universal pre-primary education in Argentina improved students' outcomes later in third grade, while Felfe, Nollenberger, and Rodríguez-Planas (2015) found an improvement in reading skill at age 15 for children exposed to an expansion of public full-time childcare in Spain in the 1990s. Curi and Menezes-Filho (2009) estimated that preschool attendance in Brazil is associated with an increase of 1.5 in years of schooling and 16% in income. Carneiro and Ginja (2014) estimated that participation in the Head Start program lowers depression among adolescents and reduces engagement in criminal

activities among young adults. Heckman, Pinto, and Savelyev (2013), by analyzing data from an evaluation of the Perry Preschool program, argued that the program induces changes in personality skills that impact adult outcomes.

This paper also speaks to a broader literature on early childhood investments. Currie and Almond (2011) showed that school entry characteristics may be as important as years of schooling to explain outcomes in the future and that damages caused by shocks early in life can be remediated. Cunha *et al.* (2006) argue that when children are stimulated correctly, they are able to better perform in primary school because they are ready to be taught the subjects they are exposed to in this school period.

This paper has five sections besides this Introduction and the Final Remarks. Section 2 explains the methodology while section 3 explains the procedures adopted for the sample selection. Section 4 presents the instruments used in the data collection and section 5 presents a descriptive analysis. Finally, section 6 presents and discusses the results.

2. Methodology

In order to evaluate the ECE effect on vocabulary in this case, the ideal scenario would be the one in which participation (i.e. school enrollment) was decided by lottery. This would ensure that there are neither observed nor unobserved differences, on average, between children attending and not attending school. However, the school enrollment in Petrolina follows the first-come-first-served rule. Therefore, if, for example, the parents who were able to enroll their children are the same ones with more financial resources, it is likely that the children of these parents perform better in cognitive tests, since these parents may have spent more resources since child birth on activities that stimulate their child's cognitive development. In this case, the difference in vocabulary between children enrolled and those not attending school is not necessarily a school effect. If a higher parental expenditure is related to a child's cognitive development (which may be untrue), the school effect will be confounded with the effect of the previous investments made.

In the present case, we may face an omitted variable bias caused by self-selection. This will occur if there is an omitted variable that is related to the fact that the child is attending school and at the same time impacts the development of the child's vocabulary (socioeconomic level, for instance). In order to mitigate this possible bias, the strategy was to conduct a detailed interview with the child's parents in order to obtain proxies for the factors usually not observed by the researcher, such as parents' involvement in the child's education, and quality of the home environment, among others (the full list is discussed in Section 4). The proxy variables will be used to control for possible confounders, as explained below.

To estimate the effect of ECE exposure on vocabulary we first estimate the following equation by Ordinary Least Squares (OLS),

$$voc_i = \beta_0 + \beta_1 ECE_i + \beta_2 X_i + \beta_3 Z_i + \epsilon_i \quad (1)$$

where voc_i is a vocabulary measure for child i , ECE_i is a binary variable indicating whether the child attends school, X_i is a vector of the child's characteristics (gender, age in months, residence area (urban or rural)), Z_i is a vector of the parents' characteristics, and ϵ_i is the error term that captures all the other unobservable effects that were not included in the equation. The parameter of interest is β_1 which indicates the average difference in vocabulary between the treatment groups, conditional on the control variables.

If the parents' assessment provides proxies that control for unobservable characteristics correlated with school attendance, then $E[\epsilon|ECE, X, Z] = 0$, which means that the OLS estimator will produce unbiased estimated coefficients. In this case, β_1 recovers the school attendance effect on vocabulary. Unfortunately, in practice, it is not possible to test if $E[\epsilon|ECE, X, Z] = 0$, so we are relying on the previous assumption.

As alternative methods, we implement three matching-based estimators: Nearest Neighbor matching (a non-parametric approach where each treatment unit is matched to the most similar control unit according to the vector of covariates), matching using Propensity Score (in which each treatment unit is matched to the most similar control unit according to the estimated propensity score), and Inverse Probability Weighting (in which the control units are reweighted using the propensity score so that units with a higher probability of being treated receive a larger weight).

At this point, all techniques have their properties well documented and are largely used in the literature when the selection is on observables.² As such, all techniques rely on the so-called conditional independence assumption (CIA), which means that conditional on the vector of covariates included in the matching procedure, the outcome is orthogonal to the treatment assignment. Therefore, the main limitation of these methods is the fact that they do not control for a possible bias coming from selection on unobservables. Nevertheless, as highlighted before, this paper takes advantage of a rich dataset covering several characteristics that are used as proxy for unobserved characteristics that may be correlated to the treatment assignment. We describe each variable in section 4.

3. Sample

To carry out the proposed study, two groups of children born in 2012 were selected. The first group was composed of children enrolled in Nova Semente units. The second was composed of children who did not attend school at the time of the survey. Hereinafter, the first will be called the intervention (or treated) group, and the second, the control group.

Before going to the field, the control group would consist of children in the age range of interest who were on the waiting list for the Nova Semente units. Each unit had its own waiting list. For each child on the waiting list, a child of the same gender and age (in months) enrolled in the unit would be selected. However, just a few children on the waiting lists were found, because either the registration information was not updated (address or contact phone), or the child was already attending another school. In total, 31 children on the waiting lists were surveyed.

The solution to increase the sample size was to conduct an active search in the same (or next to) neighborhood of the school units (100 students in each group was the initial goal). In this search, we followed the previous criteria to obtain groups of homogeneous children: for each child attending school we had, we tried to find a similar child nearby the school unit. This

² See, for instance, Rosenbaum and Rubin (1983), Dehejia and Wahba (2002), Abadie (2005), Cameron and Trivedi (2005), Schultz and Strauss (2008).

technique is similar to a snowball sampling in which recruited individuals name potential participants. Fortunately for the society, less so for the research, it was difficult to find children out of school. So in the end of the day, during the active search, we eventually gave up on keeping the neighbor criterion and stopped the assessment phase when the number of children out of school in the sample was equal to the number of children attending school. We did not keep track of the assessment refusals (intervention or control), so we cannot estimate a refusal rate. What we can say about the field work is that, in general, we did not face major problems to assess children and their families once they were found; the biggest challenge was to find them to compose the control group.

The selected sample was composed of 174 children, 50% enrolled and 50% out of school. It was not possible to obtain a one-to-one relationship in terms of gender: there were 89 girls and 85 boys. Also, due to non-response in some questions of the parents' questionnaire, some observations were lost when estimating equation (1). There are 166 observations with complete information, 50% in each group.

4. The instruments

Children and parents were assessed between June and July, 2016. Children and parents in the intervention group were assessed at the school units in which the children were enrolled, while parents and children in the control group were interviewed at their households.

To assess children's vocabulary, we used two instruments: one measured the receptive vocabulary (Peabody Picture Vocabulary Test)³ and the other assessed the expressive vocabulary (Child Nomination Test).⁴ In Peabody, the assessor says the name of an object or animal and shows the child a picture with four drawings. The child should point out the drawing representing the word said by the assessor. In the Nomination test, the assessor shows a drawing and the child must say the name of the object or animal contained in the drawing.

³ Capovilla *et al.* (1997).

⁴ Seabra, Trevisan, and Capovilla (2012).

For the parents, instruments were chosen and designed to capture personal characteristics as well as information on their interaction with the child. Usually the variables derived from the instruments listed below are neither observable to the researcher nor found at individual level in publicly available databases. Some 80% of the interviews were carried out with the mother. The instruments were:

- i) Parent self-efficacy.⁵ This instrument assesses parents' confidence in their own attitudes towards parenting. Among the seven items of the questionnaire that has a four-level Likert scale, one example is "I know what to do when my child cries".
- ii) Cognitive stimulation at home.⁶ StimQ is the name of the instrument used to analyze the cognitive stimuli that the child receives from parents at home through play and games. Two subscales are used, where all items are "yes/no" ones. The first, called Parental Involvement in Developmental Advance, which has 10 items, evaluates parents' proactivity in teaching their children letters, numbers, and words. The second subscale, Parental Verbal Responsiveness, which has 14 items, evaluates the activities that parents perform with the child, such as storytelling or hide-and-seek games.
- iii) Quality of the home environment.⁷ This is a 15-item Likert scale instrument that assesses the quality of the home environment to which the child is exposed daily. Item examples are "No matter how hard we try, it always seems like we're late" and "The phone takes up a lot of our time at home".
- iv) Maternal depression.⁸ The Edinburgh Postnatal Depression Scale (EDPS) is used to detect depression symptoms before or after childbirth in women and men. It contains 10 Likert-scale statements, such as "I have been able to laugh and see the funny side of things".
- v) Reading books and comics. Parents were asked if they usually read books and/or comic books to the child.
- vi) Physical punishment. Parents have been asked whether and how often they have hit and/or spanked the child's hand in the last three months.

⁵ Črnčec, Barnett, and Matthey (2008).

⁶ Dreyer, Mendelsohn, and Tamis-LeMonda (1996).

⁷ Matheny Jr. *et al.* (1995).

⁸ Cox, Holden, and Sagovsky (1987).

The instruments listed in (i) to (iv) are used and validated internationally, were translated to Portuguese, adapted to the Brazilian context, and used in at least two studies in Brazil as far as we know (Weisleder et al., 2018; Mendelsohn et al., 2020). For each instrument, a two-parameter item response theory (IRT) model was estimated to generate an index.⁹ In the case of StimQ, separate models were estimated for the two subscales. Each index uses items coded to reflect better circumstances or conditions. For parent involvement and parent responsiveness, higher values indicate greater involvement and responsiveness, respectively. For the home environment, the higher the value, the less turbulent the home environment is. For parental efficacy, the higher the value, the more confident the parents are about child rearing. For maternal depression, higher index values refer to less depression symptoms. All indexes have zero mean and standard deviation equals one.

For reading, a three-value categorical variable was created, where zero indicates that the parents do not read to the child, 1 (one) indicates that they read books or comics, and 2 (two) indicates that the parents read both books and comics. For physical punishment, a three-value categorical variable was constructed: zero indicates that the parents had not hit their child in the last three months; 1 (one) indicates that the child was punished rarely or sometimes; and 2 (two) indicates that punishment occurred frequently (“often”, “almost always” or “always”).

In addition, we also investigated the socioeconomic status of the children's families using information on goods possession, such as books and smartphones.¹⁰ Again, we constructed an index using the same IRT model discussed before, where higher values refer to higher socioeconomic levels.

⁹ IRT models are used to estimate unobserved latent variables, such as ability or personality trait, using an instrument composed of items related to the characteristic of interest. These variables are assumed to be normally distributed, with $N(0,1)$. If one has a sufficient number of items, it is possible to recover the characteristic's distribution. In a two-parameter model, the item difficulty (where the item is located in the scale) and discrimination (the probability change of “success” at the difficult level) are used to estimate the characteristic. As many instruments we use contain ordered Likert-scale responses, we used a graded IRT model (even binary items are ordered in the instruments used, i.e., we don't use unordered items, as it would be the case for race (white and black), for instance).

¹⁰ The full list of items includes number of books in the home, smartphones, computers, tablets, DVD players, electronic games, and cable TV.

5. Descriptive analysis

This section presents the sample characteristics. We begin with basic statistics of the variables used in equation (1), followed by a discussion on the vocabulary outcomes. In the end of the section, we report data on enrollment period.

As mentioned earlier, the analysis does not rely on a random sample of children. Although we tried to match children using the place of residence, age and gender, the procedure was not perfect. In this section, we present a descriptive analysis to compare the groups, especially regarding the proxies for unobservable factors. Table 1 presents the mean (or proportion) and standard deviation (for continuous variables) of each variable for each group separately. In addition, the table reports the p-value of the equality of means test between groups.

In general, the groups are relatively similar. Only three out of 16 variables present different means across groups at 5% level. But the difference in these variables shows the non-random character of school attendance and, therefore, draws attention to the importance of controlling for these factors in the estimation procedure. The table shows the large set of variables that can be constructed from parents' assessment that are potentially relevant to explain the development of children's vocabulary. Under the assumption that we are able to control for unobserved factors by including these variables as independent variables in the regression or in the matching procedure, we assure that we are comparing comparable children.

The children's mean age (in months) is almost identical in both groups, but the proportion of boys in the treatment group is slightly lower than that observed in the control group, as mentioned previously. However, the equality of means test does not reject the hypothesis that the proportions are equal. In addition, as the students were paired by place of residence, the proportion of children in rural areas is the same in both groups.

Table 1 – Mean by group and equality of means test p-value

Variables	Intervention		Control		p-value
	Mean	s.d.	Mean	s.d.	
Age (in months)	48.6	2.9	48.4	3.4	0.664
Boys (proportion)	0.470		0.518		0.537
Rural area (proportion)	0.229		0.229		1.000
Person assessed: mother (proportion)	0.783		0.855		0.229
Socioeconomic status	0.116	0.737	-0.118	0.736	0.042
Read ... to the child (proportion)					
Neither books nor comics	0.193		0.241		0.454
Books or comics	0.373		0.422		0.529
Books and comics	0.434		0.337		0.204
Physical punishment					
Never	0.398		0.361		0.634
Rarely or sometimes	0.325		0.217		0.117
Often, almost always, always	0.277		0.422		0.051
Parents' involvement	-0.081	0.754	0.110	0.904	0.140
Parents' responsiveness	-0.112	0.768	0.153	0.829	0.034
Parents' efficacy	0.117	0.882	-0.131	0.872	0.071
Mother's depression	0.140	0.852	-0.081	0.831	0.092
Household environment	0.271	0.855	-0.231	0.902	0.000
Receptive vocabulary (raw score)	11.07	5.83	8.27	7.06	0.006
Expressive vocabulary (raw score)	13.32	7.34	13.24	9.15	0.951
Observations	166		166		

The other variables in the table come from the parents' questionnaire. In the intervention group, more than 78% of the people assessed are the mothers, while in the control group, the percentage is higher, 85.5% (the difference is not statistically significant). This information is relevant in the sense that the questionnaire was aimed at the child's parents; secondary sources of information, such as grandparents or uncles, can generate measurement errors. Besides the mother, 8.4% (7.2%) of the respondents are the fathers in the intervention group (control).

One can observe that children exposed to the ECE program have a higher socioeconomic status than the children not exposed. Although the difference is small (the index varies between -3 and 3), it is statistically significant. Thus, if this difference is not controlled for in the estimation,

one could mistakenly attribute to the ECE program the effect of the socioeconomic status on vocabulary. The same occurs with the quality of the home environment: larger values indicate a more turbulent environment, and this is what is observed in the control group compared to the intervention group.

The parents' interaction with the child is captured through several indicators. It turns out that about 60 to 65% of parents physically punished their children at least once in the three months prior to the interview. In the control group, parents tend to punish more often, but the differences are not significant. Likewise, the reading of books and/or comic books for the child is observed in 75 to 80% of households, again without statistically significant differences between the groups. The parent involvement indicator on cognitive stimulation shows that the groups are similar, with the mean being slightly higher in the control group. On the other hand, it is observed through the responsiveness index that parents of the control group play and talk to the children a little bit more (significant difference at 5%). This might be associated with parental confidence in "being parents" and acting as such, as suggested by the difference in the parental efficacy.

This section showed that there are important differences between groups that need to be controlled for. We are relying on the set of control variables presented above to overcome possible omitted variable bias. Nevertheless, it is somewhat surprising that 13 out of 16 variables are not different between groups, despite the non-random sample we have at hand.

In addition, at the bottom of Table 1 we report the raw scores of the outcomes (for the estimations, we standardize both outcomes so that they have zero mean and standard deviation equals one). Considering the age range, the receptive vocabulary level is low and the expressive vocabulary level is medium compared to other studies using children in public schools from medium income municipalities in the state of São Paulo (Seabra, Trevisan, and Capovilla, 2012; Capovilla and Capovilla, 1997).¹¹ As one can see, there is no difference between the groups in receptive vocabulary,

¹¹ It is worth mentioning that the cited references use different samples, so the sample used to conclude that the expressive vocabulary level is low is not the same one used to conclude that the receptive vocabulary level is medium. Evidence reported by Pazeto, León, and Seabra (2017), using data from a private school from São Paulo city, present a similar pattern in which expressive vocabulary level is higher than receptive vocabulary.

but there is a 3-point difference in expressive vocabulary. Next section shows the extent to which the (lack of) difference remains after implementing our empirical strategy.

Lastly, Figure 1 presents the distribution of enrollment period for the treatment group. The period, measured in months, was calculated using the enrollment date and the assessment date. We have consistent information for 69 children (out of 83). Unfortunately, we do not have data on attendance, so we do not know how much time each child actually spent in school.

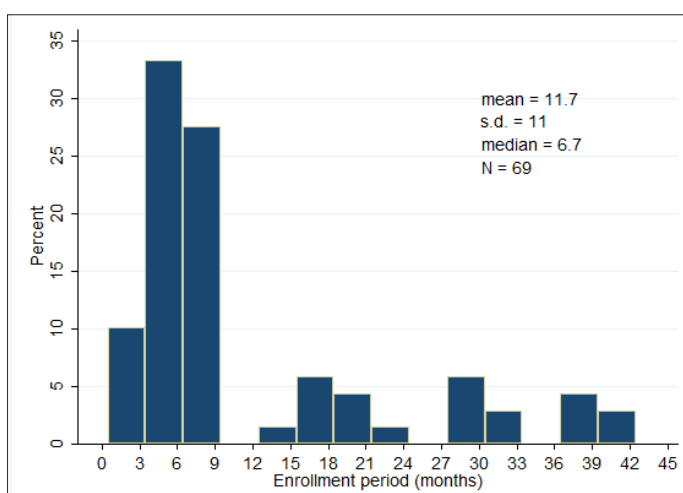


Figure 1 – Distribution of enrollment period for the treated group

Source: Authors' elaboration.

The figure shows that, at the time of data collection, most children were attending school for up to 9 months. The distribution is very asymmetric to the right, so the mean enrollment period, 11.7 months, is five months above the median period. Still, taking into account that children at this age (four years old on average) develop very fast, if the early childhood education provided is of good quality, the enrollment period is potentially enough to impact children's vocabulary.

6. Results

This section presents the estimated association between preschool attendance and children's vocabulary. There are two outcome variables measuring children's vocabulary level: one for expressive vocabulary and another for receptive vocabulary. Both variables are standardized (mean equals zero and standard deviation equals one). Figures 2 and 3 present the estimated differences between intervention and control groups (square marks) and 5% confidence intervals for receptive and expressive vocabulary, respectively. The vertical lines mark zero. Each figure reports the results for the five specifications: two using OLS (with and without control variables), Inverse Probability Weighting (IPW), Propensity Score (PS) Matching, and Nearest Neighbor (NN) Matching. The control variables included in the OLS regression are the same ones used in the three matching approaches. Appendix Figure A1 shows the propensity score distribution by treatment assignment before and after weighting observations, a procedure used to test the CIA assumption. Table 2 in next subsection reports the point estimates and standard errors for both outcomes for all specifications while Appendix Table A1 reports all estimated coefficients for the OLS regression with controls.

Figure 2 indicates that the ECE program has a positive effect on children's expressive vocabulary. The effect is always statistically significant at 10%, and significant at 5% in all specifications. The point estimate is high, varying between 0.38 and 0.43 of a standard deviation (excluding PS matching), i.e., on average, children attending school in Petrolina have an expressive vocabulary level that is 0.4 of a standard deviation higher than that observed for children not enrolled.

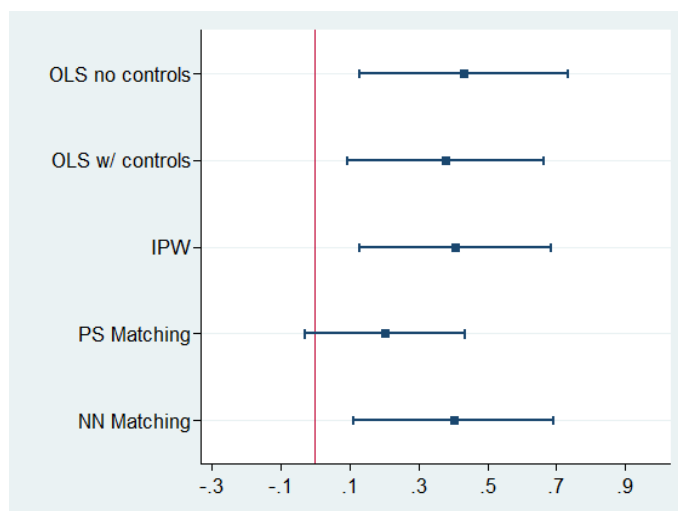


Figure 2 – Impact of Preschool on Expressive Vocabulary

Source: Authors' elaboration. 95% confidence intervals presented. Notes: The figure shows the impact estimates from five different specifications. From top to bottom: OLS; OLS controlling for child's socio-demographic characteristics as well as parent's variables associated with childbearing, depression, interaction with the child, and reading activities; Inverse Probability Weighting; Propensity Score Matching; and Nearest Neighbor Matching. The variables used in the propensity score estimations are the same ones included in the OLS with controls. All impacts are measured in standard deviation units.

Taking a closer look at the results, we uncover interesting findings. As expected, the estimated effect in the absence of control variables for observable and unobservable factors (OLS no controls) is quantitatively the largest one. This is because other factors related to the fact that the child is enrolled also influence vocabulary level. However, as one can see in the figure, controlling for these factors has only a moderate effect on the estimated effect. This is consistent with the descriptive analysis carried out before: although enrollment was not random, the differences between groups are not big (see Table 1). It is worth mentioning that there are controls for child's age and gender, socioeconomic status, place of residence, the identity of the person who answered the parents' questionnaire, and different dimensions of parents' characteristics, behavior and attitudes toward the child.

For receptive vocabulary, Figure 3 shows that there are no differences between children attending school and children not attending school. As one can see in the figure, no specification presents statistically significant coefficients. In addition, four out of five point estimates are negative (but the standard errors are larger than the point estimates).

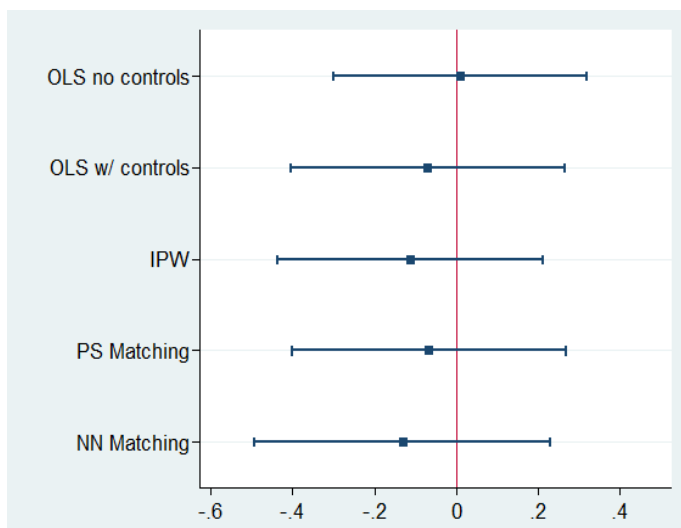


Figure 3 – Impact of Preschool on Receptive Vocabulary

Source: Authors' elaboration. Notes: 95% confidence interval presented. The figure shows the impact estimates from five different specifications. From top to bottom: OLS; OLS controlling for child's socio-demographic characteristics as well as parent's variables associated with childbearing, depression, interaction with the child, and reading activities; Inverse Probability Weighting; Propensity Score Matching; and Nearest Neighbor Matching. The variables used in the propensity score estimations are the same ones included in the OLS with controls. All impacts are measured in standard deviation units.

The distinct result on receptive and expressive vocabulary is twofold. On the one hand, as discussed earlier, passive listening to stories fosters receptive vocabulary development. As parents from both groups report they read to the child, it is not totally unexpected that there is no impact in the Peabody test. This is an evidence that receptive vocabulary can be developed at home even by vulnerable families. We present evidence to support this argument in the next section. On the other hand, as discussed earlier, in addition to passive reading, promoting expressive vocabulary requires interactive reading, an activity that demands more from parents and that was developed by the ECE program in Petrolina.

In conclusion, the main results show that preschool attendance can have substantial impacts on children's expressive vocabulary. The results are restricted to the Petrolina context and cannot be generalized without caution since the Nova Semente program was a very specific ECE program.

6.1. Robustness

As mentioned before, 31 children in the control group come from waiting lists (each unit had its own list). These children are potentially more similar to the treated children since their parents were also interested in getting them enrolled in preschool. In this section we use this restricted sample to compose the control group in an attempt to further reduce a potential selection bias. To avoid comparing children from different neighborhoods, we also restrict the treatment group: we keep in the sample only the children enrolled in the units that have children in the waiting list included in the control group. So instead of using data from 31 schools, we use data from 17. Both restrictions reduce the sample to 60 observations. Table 2 presents the results previously discussed in columns 1 (expressive) and 3 (receptive) as well as the results using the restricted sample in columns 2 and 4.

As seen, the point estimates using the restricted sample are qualitatively similar to the results using the full sample. The coefficients suggest that preschool attendance is positively associated with better expressive vocabulary and negatively associated with receptive vocabulary. However, all estimates (but one) are statistically nonsignificant.

In the case of expressive vocabulary, the coefficients are smaller, suggesting that the impact estimated before was in part a result of sample selection despite all the control variables included in the model. On the other hand, this could be a "heterogeneous" impact for the schools having a waiting list included in the analysis. In any case, the coefficients are also less precisely estimated.

Table 2 – Impact of ECE program on expressive and receptive vocabulary by estimation method and sample

	Expressive Vocabulary		Receptive Vocabulary	
	Full sample	Waiting list	Full sample	Waiting list
	(1)	(2)	(3)	(4)
OLS no controls	0.432*** (0.155)	0.318 (0.252)	0.010 (0.158)	-0.094 (0.243)
OLS with controls	0.379*** (0.146)	0.202 (0.242)	-0.070 (0.171)	-0.323 (0.232)
IPW	0.407*** (0.142)	0.203 (0.216)	-0.112 (0.166)	-0.358* (0.200)
PS Matching	0.202* (0.119)	0.113 (0.301)	-0.067 (0.171)	-0.339 (0.269)
NN Matching	0.402*** (0.148)	0.241 (0.282)	-0.131 (0.184)	-0.219 (0.238)
Observations	166	60	164	59

Standard errors in parentheses. *** 1% significant. ** 5% significant. * 10% significant. Full sample includes all children with complete information. Waiting list sample includes children from waiting lists in the control groups and children enrolled in schools with waiting lists in the intervention group.

In turn, the results for receptive vocabulary are less unstable in comparison to the full sample results. The larger negative estimates in absolute terms are consistent with an effect similar to the John Henry effect, in which the control group increases effort when they are not treated. In the case studied in this paper, parents would be trying to compensate for the lack of early childhood education while they wait for a spot for their child. This strategy would work better with receptive vocabulary since it is associated with passive learning. But, again, only one out of 10 estimates (full and restricted sample) for receptive vocabulary is statistically significant.

This section showed that comparing groups potentially more similar to each other due to parents' willingness to demand early childhood education presents results that are not different from the ones estimated using whole sample in qualitative terms. In the next section, we perform additional estimations to better understand the results.

6.2. Heterogeneity

In this section, two additional exercises are carried out. The first one estimates the effect of attending school for different points of the vocabulary distribution. The second one estimates whether the effect is associated with any subgroup (for instance, boys or parents that do not physically punish their children).

The first exercise allows one to see whether the average effect reported before is concentrated either at some parts of the distribution (for example, among those who already have a high proficiency level) or is spread throughout the distribution. Quantile regressions are estimated for five distribution percentiles: 10th, 25th, 50th (median), 75th and 90th. All specifications include all control variables previously used.¹² Observations are weighted by the inverse of the estimated propensity score. Table 3 presents the estimated coefficients as well as the standard errors for each outcome.

The results suggest that school attendance is important for most children. As shown in the table, for the expressive vocabulary, the estimated coefficients are statistically significant up to the 75th percentile. In turn, on the top of the distribution, the point estimate is smaller and statistically nonsignificant. For the receptive vocabulary, no impact was found. Either the impact is too small (up to 75th percentile) or the point estimate is very imprecisely estimated. These results can be interpreted as empirical evidence of the importance of the program to correct gaps in expressive vocabulary development.

Table 3 – Effects at different points in the distribution of vocabulary proficiency

	(1) 10 th	(2) 25 th	(3) Median	(4) 75 th	(5) 90 th
Expressive Voc.	0.471*** (0.123)	0.658*** (0.119)	0.535** (0.206)	0.478** (0.224)	0.314 (0.303)
Observations	166	166	166	166	166
Receptive Voc.	-0.014 (0.100)	-0.016 (0.112)	0.043 (0.144)	0.003 (0.291)	-0.222 (0.530)
Observations	164	164	164	164	164

Robust standard errors in parentheses. *** 1% significant. ** 5% significant. Each row and column reports estimated coefficients for the treatment dummy from quantile regressions. Observations weighted by the inverse of propensity score.

¹² Estimated coefficients for the control variables are available upon request.

The second exercise is a subgroup analysis. The idea is to analyze if the estimated effect varies according to subgroups. This would reveal the factors most associated with vocabulary development in the present context, possibly indicating whether and where there is room for other interventions. This exercise is performed in the following way. We run OLS regressions as equation (1) in which we add an interaction term of the treatment dummy and the covariate of interest. One separate regression is run for each covariate. Observations are weighted by the inverse of the propensity score. Table 4 reports the coefficients associated to these interactions for the expressive vocabulary outcome.¹³

Table 4 – Heterogeneous effects of school attendance on expressive vocabulary

	Coef. (s.e.)		Coef. (s.e.)
Age	0.023 (0.040)	Read books or comics	-0.826** (0.343)
Boy	0.165 (0.291)	Read book and comics	-0.718* (0.390)
SES	-0.213 (0.190)	Parental involvement	0.242 (0.157)
Rural	-0.304 (0.343)	Parental responsiveness	-0.160 (0.177)
Household environment	0.053 (0.183)	Punishment: sometimes	-0.221 (0.369)
Parental efficacy	0.038 (0.189)	Punishment: often	0.077 (0.330)

Standard errors in parentheses. *** 1% significant. ** 5% significant. * 10% significant. Each coefficient refers to the interaction term between the treatment dummy and the respective control variable.

In general, no distinct effects are observed. For example, it can be stated neither that girls have benefited more than boys, nor that children with higher socioeconomic status benefited more than children with lower socioeconomic status. The same is true for all the other control variables shown in the table. The only exception is the act of reading books and/or comics to the child at home. The negative coefficients for both variables

¹³ Results for receptive vocabulary show no differential impacts by subgroups and are not reported (they are available upon request).

mean that the school effect is smaller for children exposed to books/comics reading than the one observed for children whose parents do not engage in reading activities with the child at home. In addition, it does make sense that the (absolute) point estimate is bigger for reading books or comics compared to reading both, since it is reasonable to expect more reading exposure time in the second case.

Therefore, the results show that the program impacts more the vocabulary of children who have no reading exposure outside the school environment. In this sense, school attendance is an instrument to overcome the lack of reading activities. This result is consistent with the literature studying the effect of reading activities on child development:¹⁴ more exposure to books improves several cognitive dimensions, including expressive vocabulary.

7. Final remarks

This paper analyzed the effect of an ECE program on four-year-old children's vocabulary using data from Petrolina (PE) collected in 2016. Expressive and receptive vocabularies were measured through the Child Nomination Test and the Peabody Test, respectively. The results indicate positive associations between expressive vocabulary and preschool attendance, with point estimates varying between 0.20 and 0.43 of a standard deviation. But there was no impact on receptive vocabulary.

The estimation methods used do not allow us to state that the effect is a causal effect. Concerns about self-selection are present. Nevertheless, the information provided by the parents' assessment is very rich, covering several personal and interaction with the child characteristics that are usually not observed. The assessment gave birth to many proxies for unobservable factors potentially associated with both child enrollment and vocabulary development. In the regression, the inclusion of these variables reduces (or even eliminates) possible omitted variable bias.

¹⁴ See Klass, Dreyer, and Mendelsohn (2009); Needlman *et al.* (2005); Zuckerman and Khandekar (2010); Weisleder *et al.* (2018).

Last, but not least, this paper finds evidence for language development in both home and school situations. Schools may play an even more important role for students coming from relatively language-poor environments. To achieve robust results, it may require an equally robust curriculum and adequately trained teachers, since the evidence also shows that expressive vocabulary can be developed at home and particularly when parents read and talk about books with their children.

References

- Abadie, Alberto. 2005. "Semiparametric difference-in-differences estimators". *The Review of Economic Studies* 72(1): 1-19.
- Abidin, R. Richard. 1990. *Parenting Stress Index-Manual*. Charlottesville, VA: Pediatric Psychology Press.
- Berlinski, Samuel, Sebastian Galiani, and Paul Gertler. 2009. "The effect of pre-primary education on primary school performance." *Journal of Public Economics* 93, no. 1-2: 219-234.
- Bradley, Robert H., and Robert F. Corwyn. 2002. "Socioeconomic status and child development." *Annual Review of Psychology* 53, no. 1: 371-399.
- Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics: methods and applications*. Cambridge University Press.
- Camilli, Gregory, Sadako Vargas, Sharon Ryan, and W. Steven Barnett. 2010. "Meta-analysis of the effects of early education interventions on cognitive and social development." *Teachers college record* 112, no. 3: 579-620.
- Capovilla, Fernando, and Alessandra Capovilla. 1997. "Desenvolvimento linguístico na criança dos dois aos seis anos." *Ciência Cognitiva: Teoria, Pesquisa e Aplicação* 1, no. 1: 353-80.
- Capovilla, Fernando, Alessandra Capovilla, Leila Nunes, Ivânia Araújo, Débora Nunes, Daniel Nogueira, and Ana Beatriz Bernat. 1997. "Versão brasileira do Teste de Vocabulário por Imagens Peabody." *Distúrbios da Comunicação*, 8(2), 151-162.
- Carneiro, Pedro, and Rita Ginja. 2014. "Long-term impacts of compensatory preschool on health and behavior: Evidence from Head Start." *American Economic Journal: Economic Policy* 6, no. 4: 135-73.
- Cox, John L., Jeni M. Holden, and Ruth Sagovsky. 1987. "Detection of postnatal depression: development of the 10-item Edinburgh Postnatal Depression Scale." *The British Journal of Psychiatry*, 150, no. 6: 782-786.
- Črnčec, Rudi, Bryanne Barnett, and Stephen Matthey. 2008. "Development of an instrument to assess perceived self-efficacy in the parents of infants." *Research in Nursing & Health* 31, no. 5: 442-453.
- Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov. 2006. "Interpreting the evidence on life cycle skill formation." In *The Handbook of Economics of Education*, ed. Eric Hanushek and Finis Welch, cap. 12, p. 697-812. North-Holland.
- Cunningham, Anne E., and Keith E. Stanovich. 1997. "Early reading acquisition and its relation to reading experience and ability 10 years later." *Developmental Psychology* 33, no. 6: 934.
- Curi, Andréa Zaitune, and Naércio Aquino Menezes-Filho. 2009. "A relação entre educação pré-primária, salários, escolaridade e proficiência escolar no Brasil." *Estudos Econômicos (São Paulo)* 39, no. 4: 811-850.

- Currie, Janet, and Douglas Almond. 2011. "Human capital development before age five." In *Handbook of Labor Economics*, edited by David Card and Orley Ashenfelter, vol. 4, p. 1315-1486. Elsevier.
- Dehejia, Rajeev H., and Sadek Wahba. 2002. "Propensity score-matching methods for nonexperimental causal studies." *Review of Economics and Statistics* 84, no. 1: 151-161.
- Dreyer, Benard P., Alan L. Mendelsohn, and Catherine S. Tamis-LeMonda. 1996. "Assessing the child's cognitive home environment through parental report; reliability and validity." *Early Development and Parenting: An International Journal of Research and Practice* 5, no. 4: 271-287.
- Durkin, Dolores. 1966. "The achievement of pre-school readers: Two longitudinal studies." *Reading Research Quarterly*, p. 5-36.
- Felfe, Christina, Natalia Nollenberger, and Núria Rodríguez-Planas. 2015. "Can't buy mommy's love? Universal childcare and children's long-term cognitive development." *Journal of Population Economics* 28, no. 2: 393-422.
- Hart, Betty, and Todd R. Risley. 1995. *Meaningful differences in the everyday experience of young American children*. Paul H. Brookes Publishing.
- Heckman, James, Rodrigo Pinto, and Peter Savelyev. 2013. "Understanding the mechanisms through which an influential early childhood program boosted adult outcomes." *American Economic Review* 103, no. 6: 2052-86.
- Klass, Perri, Benard P. Dreyer, and Alan L. Mendelsohn. 2009. "Reach out and read: literacy promotion in pediatric primary care." *Advances in Pediatrics* 56, no. 1: 11-27.
- Matheny Jr., Adam P., Theodore D. Wachs, Jennifer L. Ludwig, and Kay Phillips. 1995. "Bringing order out of chaos: Psychometric characteristics of the confusion, hubbub, and order scale." *Journal of Applied Developmental Psychology* 16, no. 3: 429-444.
- McIntyre, Lauren J., M. Hellsten Laurie-Ann, Julia Bidonde, Catherine Boden, and Carolyn Doi. 2017. "Receptive and expressive English language assessments used for young children: a scoping review protocol." *Systematic Reviews* 6, no. 1: 70.
- Mendelsohn, Alan L., Luciane da Rosa Piccolo, João Batista Araújo Oliveira, Denise SR Mazzuchelli, Aline Sá Lopez, Carolyn Brockmeyer Cates, and Adriana Weisleder. 2020. "RCT of a reading aloud intervention in Brazil: Do impacts differ depending on parent literacy?." *Early Childhood Research Quarterly*, 53: 601-611.
- Mercy, James A., and Lala Carr Steelman. 1982. "Familial influence on the intellectual attainment of children." *American Sociological Review*, p. 532-542.
- Needlman, Robert, Karen H. Toker, Benard P. Dreyer, Perri Klass, and Alan L. Mendelsohn. 2005. "Effectiveness of a primary care intervention to support reading aloud: a multicenter evaluation." *Ambulatory Pediatrics* 5, no. 4: 209-215.
- Pazeto, Talita de Cássia Batista, Camila Barbosa Riccardi León, and Alessandra Gotuzo Seabra. 2017. "Avaliação de habilidades preliminares de leitura e escrita no início da alfabetização." *Revista Psicopedagogia*, 34, no. 104: 137-147.
- Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The central role of the propensity score in observational studies for causal effects". *Biometrika* 70, no. 1: 41-55.
- Seabra, Alessandra, Bruna Trevisan, and Fernando Capovilla. 2012. "Teste infantil de nomeação." In *Avaliação neuropsicológica cognitiva: Linguagem oral*, ed. Alessandra Seabra and Natália Dias, v. 2: 54-86. Memnon.
- Sénéchal, Monique. 1997. "The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary." *Journal of Child Language* 24: 123-138.
- Sénéchal, Monique, and Jo-Anne LeFevre. 2002. "Parental involvement in the development of children's reading skill: A five-year longitudinal study." *Child Development* 73, no. 2: 445-460.
- Schultz, T. Paul, and John Strauss. 2008. *Handbook of Development Economics*. Vol. 4. Elsevier.
- Solis, Magaly Lucia, and Richard R. Abidin. 1991. "The Spanish version parenting stress index: A psychometric study." *Journal of Clinical Child and Adolescent Psychology* 20, no. 4: 372-378.

Weisleder, Adriana, Denise S. R. Mazzuchelli, Aline Sá Lopez, Walfrido Duarte Neto, Carolyn Brockmeyer Cates, Hosana Alves Gonçalves, Rochele Paz Fonseca, João Oliveira, and Alan L. Mendelsohn. 2018. "Reading aloud and child development: a cluster-randomized trial in Brazil." *Pediatrics* 141, no. 1: e20170723. DOI: <https://doi.org/10.1542/peds.2017-0723> (acesso em 21/11/2020).

Zuckerman, Barry, and Aasma Khandekar. 2010. "Reach Out and Read: evidence-based approach to promoting early child development." *Current Opinion in Pediatrics* 22, no. 4: 539-544.

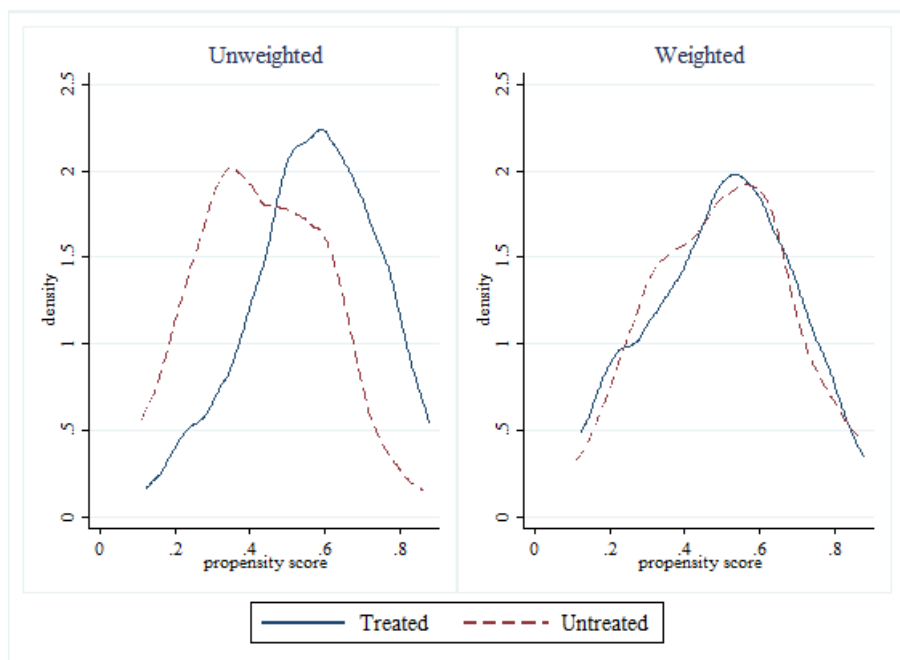
Appendix

Table A1 - Impact of ECE program on expressive and receptive vocabulary by sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Full sample				Waiting list			
	Expressive		Receptive		Expressive		Receptive	
Treated	0.432*** (0.155)	0.397*** (0.144)	0.010 (0.158)	-0.047 (0.170)	0.318 (0.252)	0.265 (0.247)	-0.094 (0.243)	-0.229 (0.235)
Age		0.101*** (0.022)		0.084*** (0.019)		0.107** (0.040)		0.086** (0.032)
Male		0.093 (0.144)		0.073 (0.158)		-0.031 (0.309)		-0.264 (0.244)
SES		0.309*** (0.087)		0.244** (0.108)		0.108 (0.148)		0.108 (0.184)
Rural		-0.152 (0.189)		-0.032 (0.173)		-0.289 (0.282)		-0.282 (0.306)
Mother assessed		0.128 (0.162)		0.110 (0.175)		0.434 (0.372)		0.555* (0.323)
Parent involvement		-0.120 (0.092)		-0.125 (0.092)		-0.161 (0.215)		0.004 (0.167)
Parent responsiveness		-0.093 (0.094)		-0.070 (0.102)		0.124 (0.226)		-0.063 (0.175)
Home environment		-0.080 (0.112)		-0.163 (0.099)		-0.243 (0.186)		-0.298 (0.183)
Parent efficacy		-0.352*** (0.114)		-0.145 (0.101)		-0.064 (0.227)		-0.065 (0.179)
Mother depression		0.025 (0.096)		0.096 (0.080)		0.160 (0.151)		0.254* (0.135)
Read books or comics		0.044 (0.186)		-0.067 (0.242)		0.093 (0.603)		0.034 (0.448)
Read books and comics		-0.014 (0.205)		-0.064 (0.239)		-0.370 (0.626)		-0.491 (0.448)
Punishment: sometimes		-0.074 (0.181)		0.050 (0.220)		0.040 (0.308)		0.176 (0.312)
Punishment: often		-0.402** (0.162)		-0.430*** (0.154)		-0.353 (0.311)		-0.152 (0.249)
Constant	-0.212* (0.119)	-4.881*** (1.046)	-0.007 (0.122)	-3.817*** (0.937)	-0.089 (0.176)	-4.975*** (1.736)	0.159 (0.168)	-3.665** (1.504)
Observations	166	166	164	164	60	60	59	59
R-squared	0.045	0.351	0.000	0.207	0.027	0.418	0.003	0.387

Robust standard errors in parentheses. *** 1% significant. ** 5% significant. * 10% significant. Full sample includes all children assessed with complete information. Waiting list includes children in the waiting lists and children enrolled in units that have a waiting list.

Figure A. 1 - Unweighted and weighted propensity score distributions by treatment group



Source: Authors' elaboration.