Direct and Indirect Impacts of Pre-School on Student Proficiency[1]

Cristine Pinto[+], Daniel Santos[†] and André Portela Souza[◇]

Abstract

Este estudo estima o efeito da pré-escola na proficiência dos alunos, separando este efeito em dois componentes, um efeito direto e um indireto. O efeito direto é o efeito da pré-escola na proficiência em determinada série, enquanto o efeito indireto se refere ao efeito da pré-escola via repetência. Durante a pré-escola, o aluno se torna mais familiarizado com o ambiente escolar, e desenvolve habilidades como cooperação, atenção que são importantes para a permanância do mesmo na escola. Estas habilidades desenvolvidas na pré-escola ajudam o processo de aprendizado nas séries subsequentes, e consequentemente influenciam a probabilidade de um aluno repetir em uma série subsequente. Como repetência também afeta a proficiência dos alunos nos anos futuros na escola, a pré-escola tem um impacto indireto na proficiência via repetência. Usando os dados da Prova Brasil 2005, encontramos que frequentar a pré-escola decresce a probabilidade de repetir uma série em mais ou menos 7% e tem um efeito positivo na proficiência em matemática. O impacto direto da pré-escola varia de 86% a 51%, enquanto o impacto indireto através da repetência varia de 14% to 49%.

This study develops an empirical selection model to estimate the impact of pre-school attendance on student proficiency disentangling it into two, a direct and an indirect impacts. The direct impact is the effect on the proficiency at a grade level in itself. Attending a pre-school, a student may develop abilities that are important to learning such as attention, cooperation, curiosity to learn, etc. The indirect impact is the impact through retention reduction in the previous grades. Having attended a pre-school, a student accumulate initial knowledge that can help the learning process in the first grade. The knowledge accumulated in the first grade have an influence in the learning in the next grade, and so on. Because of lack of knowledge from previous grades, student may be eventually delayed in a given grade. Since the delay in itself affects the learning, the pre-school may affect the grade repetition which in turn, through delay, may affect the proficiency later on. It finds that attending a pre-school decrease the probability to have repeated a grade at least once by seven to ten percentage points. It also finds that attending a pre-school increases Math proficiency at Fourth grade. The direct impact ranges from 86% to 51% and the indirect impact through grade repetition reduction from 14% to 49%.

JEL Classification: C1, C5, I21, I28

Key Words: Pre-School, Grade Repetition, Proficiency, Selecion Models

Palavras chaves: Pré-escola, proficiência, repetência, Modelos de seleção

Área: Microeconomia Aplicada

[1] [+]Escola de Economia de São Paulo, FGV; E-mail: cristine.pinto@fgv.br.

[†] FEA-RP/USP, Ribeirão Preto; E-mail: ddsantos@gmail.com

[◇]Escola de Economia de São Paulo, FGV; E-mail:andre.portela.souza@fgv.br

1 Introduction

Investments in early childhood education has been emphasized in the literature as a key element to the human capital formation of the individuals. For instance, Heckman and Cunha (2007) developed a model in which investments in different stages of the life cycle of a child may produce different returns. There are two fundamental periods for the development of a child´s skill. The first is the "sensitive period" in which investment for the production of certain skills are more productive. And the second is the "critical period", the only time a child's life when certain skills can be developed. The interpretation of the development of skills according to a multi-stage technology may imply that the investments made in a stage allow for "self-productivity", since the skills acquired amplify the effects of other skills. Moreover, this technology is complementary, as previous accumulated skills help the assimilation of other skills in future stages, increasing the productivity of investments in the subsequent stage. This approach is fundamental to guide public policies, since investments for the development of skills in disadvantaged adolescents tend to be far less productive than investments in the education of poor children in pre-school age.

The approach of Cunha and Heckman gained more prominence because it also finds root in neuroscience. Knudsen et al. (2006) present a review of the literature on the subject and concluded that human skills are acquired through a sequence of "sensitive periods". There would be times when the development of certain neural circuits would be more plastic, i.e. beings humans possess a hierarchical sequence of great moments in their lives to receive environmental influences. For the first years of life, the authors state that the cognitive, linguistic, social and emotional are strongly shaped by early life experiences of the developing child, and all contribute significantly to the success professional of future adult.

The effects of some intervention programs in early childhood are supported by several studies and various cost-benefit[2]. The conclusion in general is that the program participants had higher scores on proficiency exams, completed high school in greater proportions, and had lower rates of school delay and involvement in crime and delinquency. For some programs Knudsen et al. (2006) and Barnett (2008) report rates of return in the order of 17% and raising the IQ of the subjects with persistent effects to long term.

The literature on early childhood education also includes work on Latin America. Berlinski et al. (2006) use the expansion of the network of preschools in Argentina as a treatment to increase enrollment rates in kindergarten, and conclude that one year of preschool raised the note on proficiency exams in 3rd grade by 8% on average (equivalent to 23% standard deviation), and positively affected the attention and participation in classroom and school discipline and effort. In Uruguay, Berlinski et al. (2008) evaluate the impact of early childhood education among siblings of the same family who attended preschool with those who did not attend, thus controlling for specific purposes and any hidden background of the family. Their results indicate that the age of 15 students who attended preschool are 27 percentage points higher chance of being in school, and encourage authors to defend pre-school as an appropriate policy to prevent a school year repetition in the early school years.

Brazilian students perform poorly compared to students of the same cohort in other countries. According to 2009 PISA exam, more than 70% of Brazilian Fifteen-year-old students have less than the basic skills in Mathematics. The top ten percentile of the Math grade distribution of Brazilian students is roughly equal to the median of Math grade distribution among the OECD country students (World Bank (2010)). Not only the Brazilian students do not learn the required knowledge for their ages but lag behind in their primary grades or even drop out from the primary school. Of all students that start primary education, around 70% of them do not finish primary school (World Bank (2010)).

Improving the performance of Brazilian students in the primary education certainly

[2] See, for instance, Cunha, Heckman, et al. (2005); Krueger (2002); Barnett (2008); Knudsen et al. (2006); e Curie (2001).

requires improvement in the primary education itself. However, as the early childhood education literature suggests, investments in pre-school may be a good way to improve the educational outcomes in the primary school years. Finding the right mix of investments in these different levels of education is crucial for public policy aimed to improve educational outcomes. One important step for that is to obtain credible impacts of pre-school on student outcomes.

The World Bank (2001) reports cases in Brazil of informal programs of early childhood development, such as PROAP (Food Program for Pre-School), with assistance from health, nutrition and education, which resulted, according to the World Bank, in an increase in school performance. However, in the same work, the World Bank (2001) cites research by IPEA8, based on 1997 data, which concludes that the public investment for early childhood development is uneven, being designed for the most part to wealthier families. Corroborating the work of the World Bank, the survey by the Working Group established by Ministerial Decree No. 3219 (2006) based on microdata from 2004 Brazilian household survey that the poorest classes of the population are those benefit less from education for children. These results suggests that there is a perverse preschool model in terms of reducing social inequality, since by depriving the poorest children of access to childcare and preschool, they have their ability of developing skills in the correct age deprived. Thus, the empirical question of the real impact of preschool education in the future of Brazilian students outcomes would have to be considered to support a policy recommendation for expansion of public investments for early childhood development.

There are some studies that estimate the impact of preschool on children´s outcomes in Brazil. Curi and Menezes-Filho (2006) used the Living Standards Survey (LSMS), conducted between 1996 and 1997 in the Northeast and Southeast beside the results of the 2003 National Assessment of Education (SAEB) to investigate the effect of preschool education on completed years of education, salaries and school proficiency. Through a Logit discrete choice model the authors found that the fact that pre-school has a positive and significant impact on the completion of the four educational levels (Primary, Gymnasium, High School and University), while day care centers have a positive and significant effect in completing high school and university education. In addition, Ordinary Least Squares (OLS) regressions, it was found that preschool has a positive and significant effect on the income of that individual as an adult, and on the results of proficiency exams in the 4th, 8th, and 11th grade. Almeida and Pazzelo (2001) explore the variation of the proportion of preschool attendees within states and along the years in Brazil and find that a cohort in the states that experienced a greater increase in the pre-school attendees also experienced greater improvement in primary education proficiency. Clearly, both studies suffer from potential endogeneity problems of the preschool attendance variable. If pre-school attendance, a family choice variable, is correlated to unobservable variables such as child ability and family tastes for education, these unobservable variables are also correlated to proficiency later on, making any naive estimation less credible.

This study contributes to the literature in two ways. First, it develops an empirical model to estimate the impact of pre-school attendance on student proficiency disentangling it into two channels, a direct impact and an indirect one. The direct impact is the impact in the proficiency at the grade level considered in itself. Attending a pre-school, a student may develop abilities that are important to learning such as attention, cooperation, curiosity to learn, etc.. The indirect impact is the impact through retention reduction in the previous grades. Having attended a pre-school, a student learn better at a given grade that influences the learning of the next grades. This process is summarized in the fact that a student may be eventually delayed in a given grade. Since the delay in itself affects the learning, the pre-school may affect the grade repetition which in turn, through delay, may affect the proficiency later on. Second, it estimates credibly the impact of pre-school for an environment where retention and delay is pervasive, that is, the Brazilian educational system, which helps to inform policy makers in developing countries about the allocation decisions of the educational resources.

The study estimates the impacts of attending a pre-school on: (i) probability of student grade retention in primary school; and (ii) student Math proficiency at Fourth grade of primary school in Brazil in 2005. Moreover, it disentangles the effect on proficiency between direct effect (developed skills to learn at 4th grade) and indirect effect through grade repetition or delay (developed skills to learn before 4th grade). It finds that attending a pre-school decrease the probability to have repeated, and increases Math proficiency at Fourth grade. The direct impact ranges from 86% to 91% and the direct impact through grade repetition reduction from 8% to 14%.

The paper in divided in five sections. The second section presents the empirical model to be estimated. The third section describes the data sources and the sample selection. The fourth section presents the results. The fifth section concludes.

2 Empirical Model

The estimation of pre-school on student outcome has to deal with the potential endogeneity of pre-school attendance. Pre-school attendance, a family choice variable, is correlated to unobservable variables such as child ability and family tastes for education. These unobservable variables are also correlated to proficiency later on. Also, retention and school delay is associated with proficiency, pre-school and child ability, motivation and family tastes for education. In this context, we work with system of equations:

$$Y_{it} = X_i\beta + \lambda p_i + \delta A_{it} + \upsilon_{it}$$

$$p_i = X_i\tau + Z_{it}^{p}\alpha + \varepsilon_{it}$$

$$A_{it} = \mathbf{I}\left\{X_i\beta^A + \lambda^A p_i + Z_{it}^{A}\alpha^A + \eta_{it} > 0\right\}$$

where Y_{it} is the student i proficiency at year t, X_i is a vector of student, family, teacher and school characteristics, p_i is an indicator variable if a student attended pre-school, A_{it}

is an indicator variable if a student repeated a grade, Z_{it}^{p} are the instrumental variables that affects pre-school attendance but do not affect directly the other two outcomes, and Z_{it}^{A} is the instrumental variable that affects the grade repetition but do not affect directly pre-school and proficiency. Note that pre-school affects grade repetition and proficiency. Since grade repetition affects proficiency as well, pre-school impacts proficiency through two channels, one directly and another through grade repetition.

Clearly, the error terms of the three equations are correlated. In order to account with the endogeneity of grade repetition and pre-school, we will proceed with a model with control functions[3]. The assumptions for the error distributions are $(\upsilon_{it}, e_{it}) \sim N(\mathbf{0}, \Sigma)$, where $e_{it} = (\varepsilon_{it}, \eta_{it})$

$$\Sigma = \begin{bmatrix} \sigma_{\upsilon}^{2} & \begin{bmatrix} \sigma_{\varepsilon\upsilon} \\ \sigma_{\eta\upsilon} \end{bmatrix}^{T} \\ \begin{bmatrix} \sigma_{\varepsilon\upsilon} \\ \sigma_{\eta\upsilon} \end{bmatrix} & \sigma_{e}^{2} \end{bmatrix}$$

In the first moment, the parents decide to enroll or not their kids at preschool. At school, given their motivation, effort and initial ability, the kids pass or do not pass each grade. First, we look at the probability of being approved given that the children started (or did not start) school at preschool,

$$\mathbb{E}\left[A_{it} \middle| X_i, p_i, Z_{it}^{A}\right] = \Pr\left[\eta_{it} > -X_i\beta^{A} - \lambda^{A} p_i - Z_{it}^{A}\alpha^{A} \middle| X_i, p_i, Z_{it}^{A}\right]$$

where $(\eta_{it}, \varepsilon_{it}) \sim N(\mathbf{0}, \Omega)$. Using the properties of a bivariate normal,

$$\eta_{it} = \gamma\varepsilon_{it} + w_{it}, \text{ where } w_{it} \perp \varepsilon_{it} \text{ and } w_{it} \sim N\left(0, \sigma_{w}^{2}\right)$$

with

$$\gamma = \frac{\sigma_{\varepsilon\eta}}{\sigma_{\varepsilon}^{2}}$$

[3] The full develpement of the model can be found in the appendix.

and

$$\sigma_w^2 = \sigma_\eta^2 - \frac{\sigma_{\eta\varepsilon}^2}{\sigma_\varepsilon^2}$$

$$\begin{aligned}
\mathbb{E}\left[A_{it}\middle|\, X_i, p_i, Z_{it}^A\right] &= \Pr\left[\eta_{it} > -X_i\beta^A - \lambda^A p_i - Z_{it}^A\alpha^A \middle|\, X_i, p_i, Z_{it}^A\right] \\
&= \Pr\left[\gamma\varepsilon_{it} + w_{it} > -X_i\beta^A - \lambda^A p_i - Z_{it}^A\alpha^A \middle|\, X_i, p_i, Z_{it}^A\right] \\
&= \Pr\left[w_{it} > -X_i\beta^A - \lambda^A - Z_{it}^A\alpha^A - \gamma\varepsilon_{it} \middle|\, X_i, p_i = 1, Z_{it}^A\right] \cdot p_i \\
&\quad + \Pr\left[w_{it} > -X_i\beta^A - Z_{it}^A\alpha^A - \gamma\varepsilon_{it} \middle|\, X_i, p_i = 0, Z_{it}^A\right] \cdot (1 - p_i)
\end{aligned}$$

Then we have

$$\begin{aligned}
\mathbb{E}\left[A_{it}\middle|\, X_i, p_i, Z_{it}^A\right] &= \Pr\left[w_{it} > -X_i\beta^A - \lambda^A p_i - Z_{it}^A\alpha^A - \gamma\varepsilon_{it} \middle|\, X_i, p_i, Z_{it}^A\right] \\
&= 1 - \Phi\left(\frac{-X_i\beta^A - \lambda^A p_i - Z_{it}^A\alpha^A - \gamma\varepsilon_{it}}{\sigma_w}\right) \\
&= \Phi\left(\frac{X_i\beta^A + \lambda^A p_i + Z_{it}^A\alpha^A + \gamma\varepsilon_{it}}{\sigma_w}\right)
\end{aligned}$$

with $\varepsilon_{it} = p_i - X_i\tau + Z_{it}^p\alpha$. This approach was proposed by Smith and Blundell (1986). If we normalized $\sigma_\varepsilon^2 = \sigma_\varepsilon^2 = 1$,we have

$$\mathbb{E}\left[A_{it}\middle|\, X_i, p_i, Z_{it}^A\right] = \Phi\left(\frac{X_i\beta^A + \lambda^A p_i + Z_{it}^A\alpha^A + \rho\varepsilon_{it}}{\sqrt{1-\rho^2}}\right)$$

where ρ is the correlation between ε and η.

Now, we look at the conditional expectation of test scores,

$$\mathbb{E}\left[Y_{it}\middle|\, X_i, p_i, A_{it}\right] = X_i\beta + \lambda p_i + \delta A_{it} + \mathbb{E}\left[\upsilon_{it}\middle|\, X_i, d_i, A_{it}\right]$$

Working with the last term,

$$\begin{aligned}\mathbb{E}\left[\upsilon_{it}\right|X_i,p_i,A_{it}] &= \mathbb{E}\left[\upsilon_{it}\right|X_i,p_i=1,A_{it}=1]\cdot p_i\cdot A_{it}\\ &+\mathbb{E}\left[\upsilon_{it}\right|X_i,p_i=1,A_{it}=0]\cdot p_i\cdot(1-A_{it})\\ &+\mathbb{E}\left[\upsilon_{it}\right|X_i,p_i=0,A_{it}=1]\cdot(1-p_i)\cdot A_{it}\\ &+\mathbb{E}\left[\upsilon_{it}\right|X_i,p_i=0,A_{it}=0]\cdot(1-p_i)\cdot(1-A_{it})\end{aligned}$$

We can write

$$\upsilon_{it}=\gamma_0\varepsilon_{it}+\gamma_1\eta_{it}+\varpi_{it}$$

where $\gamma_0=\frac{\sigma_{\varepsilon\upsilon}}{\sigma_\upsilon^2}$,$\gamma_1=\frac{\sigma_{\eta\upsilon}}{\sigma_\upsilon^2}$ and $\varpi_{it}\perp(\varepsilon_{it},\eta_{it})$.After some manipulations, we end up with[4]:

$$\begin{aligned}&\mathbb{E}\left[\upsilon_{it}\right|X_i,p_i,A_{it}]\\
=\ &(\gamma_0+\rho\gamma_1)\Bigg[\phi\left(X_i\tau+Z_{it}^p\alpha\right)\Phi\left(\frac{X_i\beta^A+\lambda^Ap_i+Z_{it}^A\alpha^A-\rho\left(X_i\tau+Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right)p_iA_{it}\\
&+\phi\left(X_i\tau+Z_{it}^p\alpha\right)\left(1-\Phi\left(\frac{X_i\beta^A+\lambda^Ap_i+Z_{it}^A\alpha^A-\rho\left(X_i\tau+Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right)\right)p_i\left(1-A_{it}\right)\\
&-\phi\left(X_i\tau+Z_{it}^p\alpha\right)\Phi\left(\frac{X_i\beta^A+\lambda^Ap_i+Z_{it}^A\alpha^A-\rho\left(X_i\tau+Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right)\left(1-p_i\right)A_{it}\\
&-\phi\left(X_i\tau+Z_{it}^p\alpha\right)\left(1-\Phi\left(\frac{X_i\beta^A+\lambda^Ap_i+Z_{it}^A\alpha^A-\rho\left(X_i\tau+Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right)\right)\left(1-p_i\right)\left(1-A_{it}\right)\Bigg]\\
&+(\gamma_1+\rho\gamma_0)\Bigg[\phi\left(X_i\beta^A+\lambda^Ap_i+Z_{it}^A\alpha^A\right)\Phi\left(\frac{X_i\tau+Z_{it}^p\alpha-\rho\left(X_i\beta^A+\lambda^Ap_i+Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)p_iA_{it}\\
&-\phi\left(X_i\beta^A+\lambda^Ap_i+Z_{it}^A\alpha^A\right)\Phi\left(\frac{X_i\tau+Z_{it}^p\alpha-\rho\left(X_i\beta^A+\lambda^Ap_i+Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)p_i\left(1-A_{it}\right)\\
&+\phi\left(X_i\beta^A+\lambda^Ap_i+Z_{it}^A\alpha^A\right)\left(1-\Phi\left(\frac{X_i\tau+Z_{it}^p\alpha-\rho\left(X_i\beta^A+\lambda^Ap_i+Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)\right)\left(1-p_i\right)A_{it}\\
&-\phi\left(X_i\beta^A+\lambda^Ap_i+Z_{it}^A\alpha^A\right)\left(1-\Phi\left(\frac{X_i\tau+Z_{it}^p\alpha-\rho\left(X_i\beta^A+\lambda^Ap_i+Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)\right)\left(1-p_i\right)\left(1-A_{it}\right)\Bigg]\end{aligned}$$

[4]The full description of the calculations are presented in the appendix A.

We have a 3 step procedure:

1. STEP 1: Linear regression of p_i on the explanatory variables and instruments

2. STEP 2: Probit of A_{it} on the explanatory variables, instruments and the residual from step 1

3. STEP 3: Regression of Y_{it} on the explanatory variables and the control functions.

3 Data Sources and Sample Selection

Two data sources are used. The first dataset used are the 2005 SAEB dataset. It is collected by the Anísio Teixeira National Institute of Studies and Research (INEP) from the Ministry of Education of Brazil. SAEB dataset (Sistema de Avaliação de Educação Básica) consists of microdata on students proficiency test-scores on Portuguese and Mathematics for students on the 4th and 8th grades of the primary education, and on 3th grade of the secondary education. It also includes two questionnaires that elicit information on the family background of the student and on the school characteristics. It runs every two years since 1995 and its sample is stratified to be representative at different levels (with some variations across the years), e.g., grade levels, state and national levels, public and private schools, metropolitan and non-metropolitan areas. It is important to note that it is not a panel data. Rather, it is a series of cross-section data where the students take only one of the two exams, Portuguese or Mathematics. The 2005 SAEB is used to constructed the outcomes variables (student Math proficiency and student repetition indicator variable) as well as the pre-school attendance indicator variable and additional controls (individual and family characteristics and school and teacher characteristics).

The second dataset used are the 1998 to 2003 School Censuses. The School Censuses are also collected by he Anísio Teixeira National Institute of Studies and Research (INEP) from the Ministry of Education of Brazil. They are yearly census on the schools and includes information on number of schools, number of students in each school, number of teachers by school, etc., disaggregated by municipalities, educational levels (pre-school, primary, secondary, and professional education), and systems (public or private). The School Census are used to constructed the instrumental variables. The instrumental variables for the pre-school attendance indicator variable are the number of pre-schools per four-to-six-year-old population in each municipality and the number of pre-school teachers per four-to-six-year-old population in each municipality.

The instrumental variable for the repetition indicator variable is an indicator variable if the municipality educational system is organized in cycles or not. An educational system can be pedagogically organized in cycles or series. If it is organized in series, a student can repeat any given grade or serie depending on her performance on that particular grade. If it is organized in cycles, first to fourth grades encompass the first cycle of the primary education and fifth to eighth grades forms the second cycle. A student never repeats a grade from first to third grades. In the fourth grade, a student may repeat it or not depending on her performance in the fourth grade. Similarly, a student is allowed to move on along the second cycle until eighth grade when she can fail it or not depending on her performance on the eight grade. The decision to adopt a serie or cycle system hinges on the municipality authority discretion. The instrumental variable is an indicator variable that assumes value of one if the municipality adopted a cycle system in 2003 and zero otherwise.

The direct and indirect impacts of pre-school on student proficiency are estimated for fourth graders in 2005. The proficiency measures are Math test scores from Math exam taken by the students at the end of the fourth grade and handled by external evaluators. The SAEB exams are based on the item response theory (ITR) so that they can be comparable across students and years. The student questionnaire is a set of questions about the student and her family. Particularly interest to this study, there is a retrospective question that asks when the student entered school first time: day-care or kindergarten, or first grade of primary education. Thus, the pre-school indicator variable assumes value equals one if she answers day-care or kindergarten, and zero otherwise. It also asks wether the student has repeat a grade and if yes, how many times. A student can only repeat a grade in the primary or secondary education. A repetition indicator variable is created so that it assumes value equals one if a student have repeated at least one grade in the primary education, and zero otherwise.

It is important to note that the SAEB dataset are at the student level. However, the instruments are at the municipality level. Since there is no information where the student was four or five years before, the instruments are assigned to the student current municipality. For each student municipality in 2005, the average number of pre-school and pre-school teachers (by pre-school age population) from 1998 to 2001 are calculated and assigned to each student. The average across 1998 to 2001 years are used for two reasons. First, the age distribution of fourth graders ranges from Eight to Fifteen years old so these years capture the time when these student would attend pre-school potentially. Second, it reduces potential measurement errors. Finally, the cycle indicator variable follows a similar procedure. For each student municipality in 2005, an indicator variable is created to assign whether that municipality had a cycle system in 2003. Note that a regular student would be in the Second grade in 2003.

The Table 1 below shows the age distribution of the 2005 SAEB Fourth graders along with some descriptive information.

Table 1: Age distribuiton

Age	8	9	10	11	12	13	14	>15
Frequency	0,19	7,42	49,94	24,95	9,07	4,50	2,17	1,77
Proficiency	172,24	199,03	201,78	182,41	161,87	157,43	154,25	153,68
Grade Repetition	21,52	5,66	7,27	38,98	72,53	76,01	76,54	73,59
Pre-school	0,66	0,81	0,84	0,76	0,61	0,54	0,50	0,43

Four things are worth mentioning. First, almost half of the student are not in the corrected age. Less than 8% of the fourth graders have less than ten years old and more than 17% of them have more than eleven years old. Second, among these delayed student, around 3/4 of them have repeated at least one grade and 1/4 have repeated none. It is most likely that this last group of students entered late in school. Third, the corrected-age students have better performance in Math proficiency than the delayed students. And fourth, there are greater shares of pre-school attendees among corrected age students compared to delayed students. All these regularities taken together suggest that the age itself may be an endogenous variable correlated with pre-school and proficiency.

Although school entry age choice is not modeled in this version of the paper, we deal with this potential endogeneity by comparing of two samples. The first sample encompasses all fourth grade students with valid information. The pre-school and grade repetition indicator variables are treated as endogenous variables and two specifications are estimated. One that controls for age (treating age as exogenous) and another one that does not control for age. Since grade repetition is positively correlated with age, the grade repetition variable captures the effects of age and repetition itself in this second specification. In order to have the effect of age through repetition only, a second sample is restricted to include those student at correct age with no repetition and those delayed students with repetition only.

The variables used in the estimations with both samples are: individual characteristics (Math proficiency, grade repetition indicator variable, pre-school attendance indicator variable, age, and male and white indicator variables); family characteristics (mother´s education and father´s education); school characteristics (public or private, and region indicator variables); teacher characteristics (white indicator variable, education, experience, and salary). The descriptive statistics of the overall sample are in the TABLE B1 in the appendix. The average proficiency is 189 and its standard deviation is 49 (SAEB scale). 27% of fourth graders have repeated a grade, 69% have attended pre-school, 45% are male, and 70% are in public schools.

The variation of the instruments come from the cross-section of municipalities. There are around one thousand municipalities in the SAEB sample. There are around 38 pre-schools per municipality and 278 pre-school teachers per municipality. There are on average around 0,81 pre-schools per 100 children in the municipalities, ranging from 0,04 to 9,8. There are on average 10,57 teachers per pre-school, ranging from lowest one teacher to maximum of 72 teachers per pre-school. The descriptive statistics of the instruments can also be found in the appendix.

4 Results

The Figure 1 below depicts the age distribution among fourth graders in Brazil in 2005 by pre-school attendance status. It also shows the percentage of students that have repeated at least one grade by age and pre-school attendance status.

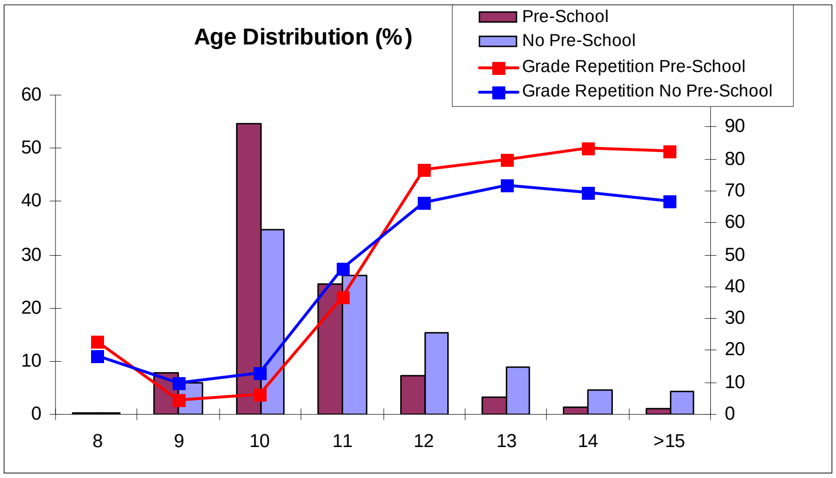

Figure 1: Grade Repetition vs Age Distribution

First, one can see that those that attended pre-school are more likely to be in the corrected age compared to those that are delayed at fourth grade. Second, among those in the corrected age, having attended pre-school is positively associated with lower probability of grade repetition. On the other hand, among those delayed students, having attended pre-school is positively associated with higher probability of grade repetition. These patterns suggest that pre-school is associated with earlier entry in primary school.

The Figure 2 below shows the age distribution among Fourth graders and their average Math proficiency by pre-school attendance status.

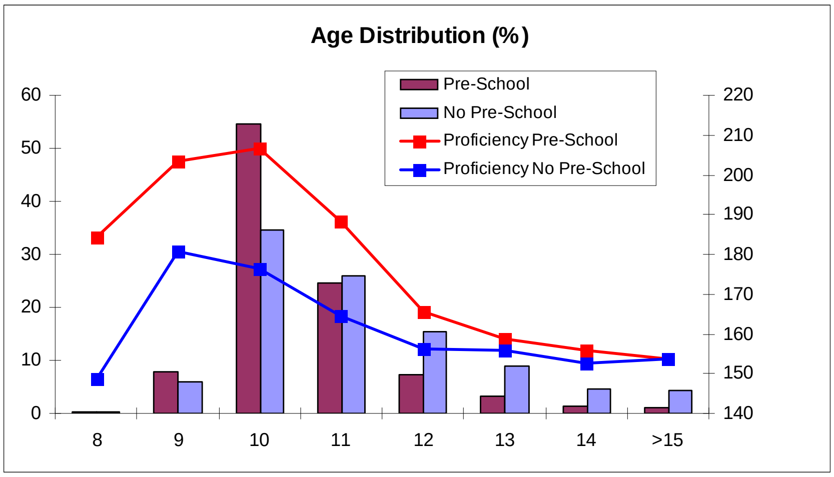

Figure 2: Proficiency vs Age Distribution

The Figure 2 reveals that having attended pre-school is positively associated with higher test scores but these differences reduces with age among those at fourth grade. The regularities of the Figures 1 and 2 suggest that (i) pre-school is associated with proficiency, (ii) pre-school is associated with age at fourth grade, and (iii) age at fourth grade is associated with grade repetition and late entry at primary school. Since age at fourth grade is associated with proficiency, it is possible that pre-school can affect proficiency later in life through direct and indirect channels. The direct channel would be the one which through pre-school the student learn cognitive and non-congnitive abilities that are important in themselves to retain knowledge in the fourth grade. The indirect channels would be the learning accumulated in the previous grades that are reflected in the corrected flow through the primary school grades. Pre-school can affect repetition that it in turn affects proficiency.

Two samples are used. An overall sample and a restricted sample. The overall sample includes all fourth grade students with valid information. The restricted sample includes students at corrected age with no grade repetition and delayed students with at least one repetition only. Two different specifications are estimated for the overall sample. One that includes age and its quadratic as controls and another that does not control for age. The former specification treats age as exogenous and the latter leaves free its correlation with grade repetition.

Two different sets of instruments are used in all samples. In one specification the instruments for pre-school attendance are the 1998 to 2001 municipality average number of pre-school teachers per one hundred children (four to six year old) and the 1998 to 2001 municipality average number of pre-schools per one hundred children (four to six year old). In other specification, we only use the 1998 to 2001 municipality average number of pre-school teachers per one hundred children (four to six year old). In all specifications the instrument for grade repetition is the 2003 cycle indicator variable.

The exclusion assumptions of the empirical model are: (i) the instrumental variables that are proxies for quantity and quality of pre-school available at the municipality when the student was at the pre-school age affects directly the probability to attend a pre-school but do not affect directly the proficiency at fourth grade (once controlled by family, school and regional characteristics). In other words, it is assumed that the pre-school inputs affects later proficiency through pre-school attendance only. Differences in pre-school availability are associated with differences in costs to attend a pre-school that are exogenous to the family choice; and (ii) the existence of a cycle system in the primary school affects the probability to repeat a grade in the primary school but does not affect directly the proficiency at fourth grade (once controlled by family, school and regional characteristics). Note that at fourth grade a student is evaluated and can failed it under the cycle and series systems. There is no repetition policy under the cycle system operates from first to third grades only.

4.1 Pre-School Attendance

The Table 2 presents the results for the probability to attend a pre-school. The results of the specifications (1) to (2) are for the overall sample and the results of the specification (3) and (4) are for the restricted sample.

Table 2: Preschool Attendance (Step 1)

Variables	(1)	(2)	(3)	(4)
Professors/Pop4a6 (98-01)	0.0100***	0.0102***	0.00909***	0.00896***
	(0.00113)	(0.000924)	(0.00129)	(0.00105)
Preschools/Pop4a6 (98-01)	0.00105		-0.00110	
	(0.00557)		(0.00644)	
Observations	35,041	35,041	26,140	26,140
F-Test	0.0000	0.0000	0.0000	0.0000

(i) *** significant at 1%; ** at 5%; * at 10%;
(ii) Additionall controls:individual characteristics (age, and male and white indicator variables); family characteristics (mother´s education and father´s education); school characteristics (public or private, and region indicator variables); teacher characteristics (white indicator variable, education, experience, and salary).

In all the regressions, we do no include age as a control variable since we have evidence that this variable may be endogenous The grade repetition indicator variable collapses the age effect in itself since by construction the students that have repeated a grade are the delayed ones. In almost all the specifications the instruments for quantity and quality of pre-schools are positively associated with the likelihood to attend a pre-school. They are all jointly significant different from zero.

The estimations of the probabilities to have attended a pre-school are the first stage regressions of the three equation empirical model described above. It allows to predict the value of $X_i\tau + Z_{it}^p\alpha$ and the residual in the first stage. The residual enters as an additional explanatory variable in the estimation of the probability to repeat a grade. Under the exclusion restrictions discussed above, the addition of this term allows one to control for the endogeneity of pre-school attendance on the grade repetition equation.

4.2 Grade Repetition

The Table 3 below presents the results for the estimation of the probability to repeat a grade in the primary education (up to grade four). The regressions follow the same specifications as the pre-school probability model with three additional terms: the cycle indicator variable, the pre-school indicator variable and the residual of the first stage (Preschool (Residual)). These are the results of the second stage regressions of the three equation empirical model.

Table 3: Grade Repetition (Step 2)

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Ciclo (2003)	-0.0952***	-0.0941***	-0.0941***	-0.218***	-0.219***	-0.219***
	(0.0200)	(0.0202)	(0.0202)	(0.0267)	(0.0268)	(0.0268)
PreSchool	-0.241***	-0.0268	-0.0357	-0.446***	-0.644	-0.621
	(0.0175)	(0.314)	(0.320)	(0.0227)	(0.470)	(0.470)
PreSchool (Residual)		-0.215	-0.206		0.199	0.175
		(0.323)	(0.319)		(0.471)	(0.471)
Observations	35,041	35,041	35,041	26,140	26,140	26,140

(i) *** significant at 1%; ** at 5%; * at 10%;
(mother´s education and father´s education); school characteristics (public or private, and region indicator variables);

The identifying assumption is that the cycle system affects the probability to repeat a grade but does not affect directly the proficiency at fourth grade (after controlling for the additional variables). We have evidence that the students in municipalities with cycle system are less likely to repeat a grade as expected.

Having attended a pre-school makes a student less likely to repeat a grade. The coefficient of the control function Preschool is the estimation of $\frac{\rho}{\sqrt{1-\rho^2}}$ which implies a coefficient of correlation between the error terms of the pre-school and grade repetition equations of - 0,21 (specification (2)). If the error term captures the individual ability, this negative correlation means that higher ability child is more likely to attend a pre-school and less likely to repeat a grade. Columns (4) - (6) presents the results for the restricted sample, in which we repeated students are the ones with age higher or equal than 11. In this model, the correlation between the error terms of the pre-school and grade repetition is positive. This result is counterintuitive since we do no expect the low ability children to be more likely to attend pre-school and more likely to repeat a grade.

Although these results are interesting in themselves, that is, pre-school attendance indeed affects the educational flow along the primary education, they also serve to construct the control function to be used as controls in the proficiency equation.

4.3 Math Proficiency

The Table 4 below presents the results for the regressions for Math proficiency. They present the results for the third stage regression of the three equation empirical model. The Table 4 also presents the results for OLS regressions.

The variables of interest are the pre-school and grade repetition indicator variables. Additional variables are the control functions for pre-school and grade repetition that controls for the endogeneity of these two variables. The OLS results show positive correlation between pre-school and Math proficiency at fourth grade and a negative correlation between grade repetition and proficiency. The specifications (1) to (2) and (3) to (4) of the three equation empirical model show weaker positive impacts of pre-school and stronger negative impacts of repetition. This weaker effect is obtained after controlling for selection. If the students that are going to be preschools are the ones with high initial ability, the coefficient of preschool has to decrease after controlling for selection.

Taking the point estimates of the Table 4, one can calculate the direct and indirect impacts of pre-school on student proficiency. These impacts are shown in Table 5 below. Considering the results of specification (4), the total impact of pre-school is 9,15 points at the SAEB scale, representing 0,19 standard deviation of the Math test scores. The direct impact and indirect impacts represent 51% and 49% of them, respectively. Indeed, both channels of pre-school attendance affect the proficiency at fourth grade at the same proportion.

Table 4: Proficiency (Step 3)

Variables	OLS	OLS	(1)	(2)	OLS	OLS	(3)	(4)
PreSchool	14.73***	13.14***	8.789***	8.858***	16.52***	13.99***	4.780**	4.678**
	(0.532)	(0.523)	(1.798)	(1.800)	(0.643)	(0.641)	(2.005)	(2.002)
Grade Repetition		-18.93***	-23.06***	-22.89***		-19.75***	-22.85***	-22.77***
		(0.508)	(3.040)	(3.049)		(0.716)	(1.582)	(1.580)
Control Function (Pre)			62.41***	62.64***			49.49***	49.19***
			(5.365)	(5.358)			(4.947)	(4.932)
Control Function (Rep)			20.96***	20.61***			24.77***	24.23***
			(3.541)	(3.574)			(4.694)	(4.674)
Observations	35,041	35,041	35,041	35,041	26,140	26,140	26,140	26,140

(i) *** significant at 1%; ** at 5%; * at 10%;

and father´s education); school characteristics (public or private, and region indicator variables); teacher characteristics (white indicator variable, education, experience, and salary).

Table 5: The Direct and Indirect Effect of PreSchool

	(1)		(2)		(3)		(4)	
Direct Effect	8,789	86%	8,858	85%	4,78	51%	4,68	51%
Indiret Effect	1,48	14%	1,57	15%	4,55	49%	4,47	49%
Coefficient	-23,06		-22,89		-22,85		-22,77	
Effect on Repetition	-0,0642		-0,0687		-0,20		-0,20	
Total Effect	10,27		10,43		9,33		9,15	
Total Effect (std. Dev.)	0,21		0,21		0,19		0,19	

5 Conclusions

This study developed an empirical self-selection model to estimate the impact of pre-school attendance on student proficiency disentangling it into two channels. An direct and an indirect impacts. The direct impact is the impact in the proficiency at the grade level considered in itself. Attending a pre-school, a student may develop abilities that are important to learning such as attention, cooperation, curiosity to learn, etc.. The indirect impact is the impact through retention reduction in the previous grades. Having attended a pre-school, a student learn better at a given grade that influences the learning of the next grades. This process is summarized in the fact that a student may be eventually delayed in a given grade. Since the delay in itself affects the learning, the pre-school may affect the grade repetition which in turn, through delay, may affect the proficiency later on. Second, it estimates the impact of pre-school for an environment where retention and delay is pervasive, that is, the Brazilian educational system, which helps to inform policy makers in developing countries about the allocation decisions of the educational resources. It finds that attending a pre-school decrease the probability to have repeated a grade at least once and increases Math proficiency at fourth grade. The direct impact ranges from 51% to 86% and the direct impact through grade repetition reduction from 14% to 49%.

The results shed some light on the debate of how much to invest in pre-primary and primary education in contexts where retention and delay are pervasive. Improvements on the quantity and quality of pre-schools may enhance the proficiency of primary school students through many channels that improve not only the benefits of primary education (proficiency) but also reduce its costs through decrease in the retention rates. Any cost and benefit analysis of pre-school and primary school investment decisions should take these channels into account.

References

[1] Almeida„ R. B., Pazello, E. T. (2010). O efeito da pré-escola sobre o desempenho escolar futuro dos indivíduos. In: 38 Encontro Nacional de Economia - ANPEC, 2010, Salvador - BA. Anais do 38 Encontro Nacional de Economia - ANPEC.

[2] Banco Mundial - Departamento de Desenvolvimento Humano. (2001). Desenvolvimento da Primeira Infância: Foco sobre o Impacto das Pré-Escolas. Brasil: Banco Mundial.

[3] Barnett, W.S. (2008). Preschool education and its lasting effects: Research and policy implications. Education and the Public Interest Center & Education Policy Research Unit: Arizona State University and University of Colorado at Boulder.

[4] Becker, G. S., & Tomes, N. (1986). Human Capital and the Rise and Fall of Families. Journal of Labor Economics , Vol. 4 (3), 1-39.

[5] Berlinski, S., Galiani, S., & Gertler, P.J. (2006). The Effect of Preprimary Education on Primary School Performance. William Davidson Institute. Working Paper No. 838.

[6] Berlinski, S., Galiani, S., & Manacorda, M. (2008). Giving children a better start: Preschool attendance and school-age profiles. Journal of Public Economics, 92, 1416-1440.

[7] Cunha, F., & Heckman, J. (2007). The Technology of Skill Formation. American Economic Review, Vol. 97 (2), 31-47.

[8] Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2005). Interpreting the Evidence on Life Cycle Skill Formation. National Bureau Of Economic Research , Working Paper 11331.

[9] Curi, A. Z., & Menezes-Filho, N. A. (2006). Os Efeitos da Pré-escola sobre os Salários, a Escolaridade e a Proficiência Escolar. XXXIV Encontro Nacional de Economia. Salvador, BA: ANPEC.

[10] Currie, J. (2001). Early Childhood Education Programs. Journal of Economic Perspectives, 15 (2), 213-238.

[11] Feinstein, L. (2003). Inequality in the Early Cognitive Development of British Children in the 1970 Cohort. Economica, 70 (277), 73-97.

[12] Griliches, Z. (1977). Estimating The Returns To Schooling: Some Econometric Problems. Econometrica , Vol. 45 (1), 1-22.

[13] Griliches, Z., Hall, B., & Hausman, J. (1978). Missing data and self-selection in large panels. Annales de I'INSEE , 137-176.

[14] Grupo de Trabalho instituído pela Portaria Interministerial n^{o}3.219. (2006). Relatório do Grupo de Trabalho Interministerial. Brasília: Ministério da Educação - Secretaria de Educação Básica.

[15] Knudsen, E., Heckman, J.J., Cameron, J.L., & Shonkoff., J.P. (2006). Economic, Neurobiological, and Behavioral Perspectives on Building America's Future Workforce. Proceedings of the National Academy of Sciences, 103, no. 27: 10155-10162.

[16] Krueger, A.B. (2002). Inequality, Too Much of a Good Thing. Paper to the Princeton University Industrial Relations Section. Princeton, NJ: Princeton University Press, 12-23.

[17] Schady, N. (2006). Early Childhood Development in Latin America and the Caribbean. Journal of the Latin American and Caribbean Ecoonomic Association, Vol. 6 (2), 185-213.

[18] World Bank. (2010). Achieving World Class Education in Brazil: The Next Agenda. Human Development Sector Management Unit, Latin America and the Caribbean Regional Office.

A Appendix A: Empirical Model

In appendix A, we present the derivation of the control functions used in the empirical model.

We work with system of equations:

$$Y_{it} = X_i\beta + \lambda p_i + \delta A_{it} + \upsilon_{it}$$

$$p_i = X_i\tau + Z_{it}^{p}\alpha + \varepsilon_{it}$$

$$A_{it} = \mathbf{I}\left\{X_i\beta^A + \lambda^A p_i + Z_{it}^{A}\alpha^A + \eta_{it} > 0\right\}$$

Assumptions:

- $(\upsilon_{it}, e_{it}) \sim N(\mathbf{0}, \Sigma)$, where $e_{it} = (\varepsilon_{it}, \eta_{it})$

$$\Sigma = \begin{bmatrix} \sigma_{\upsilon}^2 & \left[\begin{smallmatrix}\sigma_{\varepsilon\upsilon} \\ \sigma_{\eta\upsilon}\end{smallmatrix}\right]^T \\ \left[\begin{smallmatrix}\sigma_{\varepsilon\upsilon} \\ \sigma_{\eta\upsilon}\end{smallmatrix}\right] & \sigma_e^2 \end{bmatrix}$$

In the first moment, the parents decide to enroll or not their kids at preschool. At school, given their motivation, effort and initial hability, the kids pass or do not pass each grade. First, we look at the probability of being approvaded given that the children started (or did not start) schoold at preschool,

$$\mathbb{E}\left[A_{it}\middle| X_i, p_i, Z_{it}^A\right] = \Pr\left[\eta_{it} > -X_i\beta^A - \lambda^A p_i - Z_{it}^A\alpha^A \middle| X_i, p_i, Z_{it}^A\right]$$

where $(\eta_{it}, \varepsilon_{it}) \sim N(\mathbf{0}, \Omega)$. Using the properties of a bivariate normal,

$$\eta_{it} = \gamma\varepsilon_{it} + w_{it}, \text{ where } w_{it} \perp \varepsilon_{it} \text{ and } w_{it} \sim N\left(0, \sigma_w^2\right)$$

with

$$\gamma = \frac{\sigma_{\varepsilon\eta}}{\sigma_{\varepsilon}^{2}}$$

and

$$\sigma_{w}^{2} = \sigma_{\eta}^{2} - \frac{\sigma_{\eta\varepsilon}^{2}}{\sigma_{\varepsilon}^{2}}$$

$$\begin{aligned}
\mathbb{E}\left[A_{it} \middle| X_i, p_i, Z_{it}^{A} \right] &= \Pr\left[\eta_{it} > -X_i\beta^{A} - \lambda^{A} p_i - Z_{it}^{A}\alpha^{A} \middle| X_i, p_i, Z_{it}^{A} \right] \\
&= \Pr\left[\gamma\varepsilon_{it} + w_{it} > -X_i\beta^{A} - \lambda^{A} p_i - Z_{it}^{A}\alpha^{A} \middle| X_i, p_i, Z_{it}^{A} \right] \\
&= \Pr\left[w_{it} > -X_i\beta^{A} - \lambda^{A} - Z_{it}^{A}\alpha^{A} - \gamma\varepsilon_{it} \middle| X_i, p_i = 1, Z_{it}^{A} \right] \cdot p_i \\
&\quad + \Pr\left[w_{it} > -X_i\beta^{A} - Z_{it}^{A}\alpha^{A} - \gamma\varepsilon_{it} \middle| X_i, p_i = 0, Z_{it}^{A} \right] \cdot (1 - p_i)
\end{aligned}$$

Then we have (since $w_{it} \perp \left(X_i, p_i, Z_{it}^{A} \right)$):

$$\begin{aligned}
\mathbb{E}\left[A_{it} \middle| X_i, p_i, Z_{it}^{A} \right] &= \Pr\left[w_{it} > -X_i\beta^{A} - \lambda^{A} p_i - Z_{it}^{A}\alpha^{A} - \gamma\varepsilon_{it} \middle| X_i, p_i, Z_{it}^{A} \right] \\
&= 1 - \Phi\left(\frac{-X_i\beta^{A} - \lambda^{A} p_i - Z_{it}^{A}\alpha^{A} - \gamma\varepsilon_{it}}{\sigma_w} \right) \\
&= \Phi\left(\frac{X_i\beta^{A} + \lambda^{A} p_i + Z_{it}^{A}\alpha^{A} + \gamma\varepsilon_{it}}{\sigma_w} \right)
\end{aligned}$$

with $\varepsilon_{it} = p_i - X_i\tau + Z_{it}^{p}\alpha$. This approach was proposed by Smith and Blundell (1986). If we normalized $\sigma_{\eta}^{2} = \sigma_{\varepsilon}^{2} = 1$,we have

$$\mathbb{E}\left[A_{it} \middle| X_i, p_i, Z_{it}^{A} \right] = \Phi\left(\frac{X_i\beta^{A} + \lambda^{A} p_i + Z_{it}^{A}\alpha^{A} + \rho\varepsilon_{it}}{\sqrt{1-\rho^{2}}} \right)$$

where ρ is the correlation between ε and η.

Now, we look at the conditional expectation of test scores,

$$\mathbb{E}\left[Y_{it} \middle| X_i, p_i, A_{it} \right] = X_i\beta + \lambda p_i + \delta A_{it} + \mathbb{E}\left[\upsilon_{it} \middle| X_i, d_i, A_{it} \right]$$

Working with the last term,

$$\begin{aligned}\mathbb{E}\left[\upsilon_{it}\right| X_i, p_i, A_{it}] &= \mathbb{E}\left[\upsilon_{it}\right| X_i, p_i = 1, A_{it} = 1] \cdot p_i \cdot A_{it} \\ &+\mathbb{E}\left[\upsilon_{it}\right| X_i, p_i = 1, A_{it} = 0] \cdot p_i \cdot (1 - A_{it}) \\ &+\mathbb{E}\left[\upsilon_{it}\right| X_i, p_i = 0, A_{it} = 1] \cdot (1 - p_i) \cdot A_{it} \\ &+\mathbb{E}\left[\upsilon_{it}\right| X_i, p_i = 0, A_{it} = 0] \cdot (1 - p_i) \cdot (1 - A_{it})\end{aligned}$$

We can write

$$\upsilon_{it} = \gamma_0 \varepsilon_{it} + \gamma_1 \eta_{it} + \varpi_{it}$$

where $\gamma_0 = \frac{\sigma_{\varepsilon\upsilon}}{\sigma_\upsilon^2}, \gamma_1 = \frac{\sigma_{\eta\upsilon}}{\sigma_\upsilon^2}$ and $\varpi_{it} \perp (\varepsilon_{it}, \eta_{it})$

$$\begin{aligned}\mathbb{E}\left[\upsilon_{it}\right| X_i, p_i = 1, A_{it} = 1] &= \mathbb{E}\left[\gamma_0 \varepsilon_{it} + \gamma_1 \eta_{it} + \varpi_{it}\right| X_i, p_i = 1, A_{it} = 1] \\ &= \gamma_0 \mathbb{E}\left[\varepsilon_{it}\right| X_i, p_i = 1, A_{it} = 1] \\ &+\gamma_1 \mathbb{E}\left[\eta_{it}\right| X_i, p_i = 1, A_{it} = 1]\end{aligned}$$

$$\begin{aligned}&\mathbb{E}\left[\varepsilon_{it}\right| X_i, p_i = 1, A_{it} = 1] \\ =\ & \mathbb{E}\left[\varepsilon_{it}\right| X_i, \eta_{it} > -X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A, \varepsilon_{it} > -X_i\tau - Z_{it}^p\alpha] \\ =\ & \phi\left(-X_i\tau - Z_{it}^p\alpha\right)\left[1 - \Phi\left(\frac{-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A - \rho\left(-X_i\tau - Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right)\right] \\ & +\rho\phi\left(-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A\right)\left[1 - \Phi\left(\frac{-X_i\tau - Z_{it}^p\alpha - \rho\left(-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)\right] \\ =\ & \phi\left(X_i\tau + Z_{it}^p\alpha\right) \cdot \Phi\left(\frac{X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A - \rho\left(X_i\tau + Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right) \\ & +\rho\phi\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right) \cdot \Phi\left(\frac{X_i\tau + Z_{it}^p\alpha - \rho\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)\end{aligned}$$

$$
\begin{aligned}
& \mathbb{E}\left[\eta_{it}\right| X_i, p_i=1, A_{it}=1] \\
= \; & \mathbb{E}\left[\eta_{it}\right| X_i, \eta_{it} > -X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A, \varepsilon_{it} > -X_i\tau - Z_{it}^p\alpha] \\
= \; & \phi\left(-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A\right)\left(1-\Phi\left(\frac{-X_i\tau - Z_{it}^p\alpha - \rho\left(-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)\right) \\
& +\rho\phi\left(-X_i\tau - Z_{it}^p\alpha\right)\left(1-\Phi\left(\frac{-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A - \rho\left(-X_i\tau - Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right)\right) \\
= \; & \phi\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)\Phi\left(\frac{X_i\tau + Z_{it}^p\alpha - \rho\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right) \\
& +\rho\phi\left(X_i\tau + Z_{it}^p\alpha\right)\Phi\left(\frac{X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A - \rho\left(X_i\tau + Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right)
\end{aligned}
$$

At the end,

$$
\begin{aligned}
& \mathbb{E}\left[\upsilon_{it}\right| X_i, d_i=1, A_{it}=1] \\
= \; & \left(\gamma_0+\rho\gamma_1\right)\phi\left(X_i\tau + Z_{it}^p\alpha\right)\Phi\left(\frac{X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A - \rho\left(X_i\tau + Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right) \\
& +\left(\gamma_1+\rho\gamma_0\right)\phi\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)\Phi\left(\frac{X_i\tau + Z_{it}^p\alpha - \rho\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)
\end{aligned}
$$

By analogy,

$$
\begin{aligned}
\mathbb{E}\left[\upsilon_{it}\right| X_i, p_i=1, A_{it}=0] \; = \; & \gamma_0\mathbb{E}\left[\varepsilon_{it}\right| X_i, p_i=1, A_{it}=0] \\
& +\gamma_1\mathbb{E}\left[\eta_{it}\right| X_i, p_i=1, A_{it}=0]
\end{aligned}
$$

$$
\begin{aligned}
& \mathbb{E}\left[\eta_{it}\right| X_i, p_i = 1, A_{it} = 0] \\
= \ & \mathbb{E}\left[\eta_{it}\right| X_i, \eta_{it} < -X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A, \varepsilon_{it} > -X_i\tau - Z_{it}^p\alpha] \\
= \ & -\phi\left(-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A\right)\left(1 - \Phi\left(\frac{-X_i\tau - Z_{it}^p\alpha - \rho\left(-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)\right) \\
& +\rho\phi\left(-X_i\tau - Z_{it}^p\alpha\right)\left(\Phi\left(\frac{-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A - \rho\left(-X_i\tau - Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right)\right) \\
= \ & -\phi\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)\cdot\Phi\left(\frac{X_i\tau + Z_{it}^p\alpha - \rho\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right) \\
& +\rho\phi\left(X_i\tau + Z_{it}^p\alpha\right)\left(1 - \Phi\left(\frac{X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A - \rho\left(X_i\tau + Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
& \mathbb{E}\left[\varepsilon_{it}\right| X_i, p_i = 1, A_{it} = 0] \\
= \ & \phi\left(-X_i\tau - Z_{it}^p\alpha\right)\Phi\left(\frac{-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A - \rho\left(-X_i\tau - Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right) \\
& -\rho\phi\left(-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A\right)\left(1 - \Phi\left(\frac{-X_i\tau - Z_{it}^p\alpha - \rho\left(-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)\right) \\
= \ & \phi\left(X_i\tau + Z_{it}^p\alpha\right)\left(1 - \Phi\left(\frac{X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A - \rho\left(X_i\tau + Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right)\right) \\
& -\rho\phi\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)\Phi\left(\frac{X_i\tau + Z_{it}^p\alpha - \rho\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)
\end{aligned}
$$

$$
\begin{aligned}
& \mathbb{E}\left[\upsilon_{it}\right| X_i, p_i = 1, A_{it} = 0] \\
= \ & -\left(\gamma_1 + \rho\gamma_0\right)\phi\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)\Phi\left(\frac{X_i\tau + Z_{it}^p\alpha - \rho\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right) \\
& +\left(\gamma_0 + \rho\gamma_1\right)\phi\left(X_i\tau + Z_{it}^p\alpha\right)\left(1 - \Phi\left(\frac{X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A - \rho\left(X_i\tau + Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}\left[\upsilon_{it}\right| X_i, p_i = 0, A_{it} = 1] \ = \ & \gamma_0\mathbb{E}\left[\varepsilon_{it}\right| X_i, p_i = 0, A_{it} = 1] \\
& +\gamma_1\mathbb{E}\left[\eta_{it}\right| X_i, p_i = 1, A_{it} = 0]
\end{aligned}
$$

$$
\begin{aligned}
& \mathbb{E}\left[\varepsilon_{it}\right| X_i, p_i = 0, A_{it} = 1] \\
= \quad & -\phi\left(-X_i\tau - Z_{it}^p\alpha\right)\left(1 - \Phi\left(\frac{-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A - \rho\left(-X_i\tau - Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right)\right) \\
& +\rho\phi\left(-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A\right)\left(\Phi\left(\frac{-X_i\tau - Z_{it}^p\alpha - \rho\left(-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)\right) \\
= \quad & -\phi\left(X_i\tau + Z_{it}^p\alpha\right)\Phi\left(\frac{X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A - \rho\left(X_i\tau + Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right) \\
& +\rho\phi\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)\left(1 - \Phi\left(\frac{X_i\tau + Z_{it}^p\alpha - \rho\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
& \mathbb{E}\left[\eta_{it}\right| X_i, d_i = 0, A_{it} = 1] \\
= \quad & \mathbb{E}\left[\eta_{it}\right| X_i, \eta_{it} > -X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A, \varepsilon_{it} < -X_i\tau - Z_{it}^p\alpha] \\
= \quad & \phi\left(-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A\right)\Phi\left(\frac{-X_i\tau - Z_{it}^p\alpha - \rho\left(-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right) \\
& -\rho\phi\left(-X_i\tau - Z_{it}^p\alpha\right)\left(1 - \Phi\left(\frac{-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A - \rho\left(-X_i\tau - Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right)\right) \\
= \quad & \phi\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)\left(1 - \Phi\left(\frac{X_i\tau + Z_{it}^p\alpha - \rho\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)\right) \\
& -\rho\phi\left(X_i\tau + Z_{it}^p\alpha\right)\Phi\left(\frac{X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A - \rho\left(X_i\tau + Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right)
\end{aligned}
$$

$$
\begin{aligned}
& \mathbb{E}\left[\upsilon_{it}\right| X_i, p_i = 0, A_{it} = 1] \\
= \quad & -\left(\gamma_0 + \rho\gamma_1\right)\phi\left(X_i\tau + Z_{it}^p\alpha\right)\Phi\left(\frac{X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A - \rho\left(X_i\tau + Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right) \\
& +\left(\gamma_1 + \rho\gamma_0\right)\phi\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)\left(1 - \Phi\left(\frac{X_i\tau + Z_{it}^p\alpha - \rho\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}\left[\upsilon_{it}\right| X_i, p_i = 0, A_{it} = 0] \quad = \quad & \gamma_0\mathbb{E}\left[\varepsilon_{it}\right| X_i, p_i = 0, A_{it} = 0] \\
& +\gamma_1\mathbb{E}\left[\eta_{it}\right| X_i, p_i = 0, A_{it} = 0]
\end{aligned}
$$

$$
\begin{aligned}
& \mathbb{E}\left[\varepsilon_{it}\right| X_i, p_i = 0, A_{it} = 0] \\
= \quad & -\phi\left(-X_i\tau - Z_{it}^p\alpha\right)\Phi\left(\frac{-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A - \rho\left(-X_i\tau - Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right) \\
& -\rho\phi\left(-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A\right)\Phi\left(\frac{-X_i\tau - Z_{it}^p\alpha - \rho\left(-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right) \\
= \quad & -\phi\left(X_i\tau + Z_{it}^p\alpha\right)\left(1 - \Phi\left(\frac{X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A - \rho\left(X_i\tau + Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right)\right) \\
& -\rho\phi\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)\Phi\left(\frac{X_i\tau + Z_{it}^p\alpha - \rho\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)
\end{aligned}
$$

$$
\begin{aligned}
& \mathbb{E}\left[\eta_{it}\right| X_i, p_i = 0, A_{it} = 0] \\
= \quad & -\phi\left(-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A\right)\Phi\left(\frac{-X_i\tau - Z_{it}^p\alpha - \rho\left(-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right) \\
& -\rho\phi\left(-X_i\tau - Z_{it}^p\alpha\right)\Phi\left(\frac{-X_i\beta^A - \lambda^A d_i - Z_{it}^A\alpha^A - \rho\left(-X_i\tau - Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right) \\
= \quad & -\phi\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)\left(1 - \Phi\left(\frac{X_i\tau + Z_{it}^p\alpha - \rho\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)\right) \\
& -\rho\phi\left(X_i\tau + Z_{it}^p\alpha\right)\Phi\left(\frac{X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A - \rho\left(X_i\tau + Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right)
\end{aligned}
$$

$$
\begin{aligned}
& \mathbb{E}\left[\varepsilon_{it}\right| X_i, p_i = 0, A_{it} = 0] \\
= \quad & -\left(\gamma_0 + \rho\gamma_1\right)\phi\left(X_i\tau + Z_{it}^p\alpha\right)\Phi\left(\frac{X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A - \rho\left(X_i\tau + Z_{it}^p\alpha\right)}{\sqrt{1-\rho^2}}\right) \\
& -(\gamma_1 + \rho\gamma_0)\phi\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)\left(1 - \Phi\left(\frac{X_i\tau + Z_{it}^p\alpha - \rho\left(X_i\beta^A + \lambda^A d_i + Z_{it}^A\alpha^A\right)}{\sqrt{1-\rho^2}}\right)\right)
\end{aligned}
$$

At the end,

$$
\begin{aligned}
& \mathbb{E}\left[\left. \upsilon_{it} \right| X_i, p_i, A_{it}\right] \\
= \; & \left(\gamma_0 + \rho\gamma_1\right) \left[\phi\left(X_i\tau + Z_{it}^{p}\alpha\right) \Phi\left(\frac{X_i\beta^A + \lambda^A p_i + Z_{it}^{A}\alpha^A - \rho\left(X_i\tau + Z_{it}^{p}\alpha\right)}{\sqrt{1-\rho^2}}\right) p_i A_{it} \right. \\
& + \phi\left(X_i\tau + Z_{it}^{p}\alpha\right) \left(1 - \Phi\left(\frac{X_i\beta^A + \lambda^A p_i + Z_{it}^{A}\alpha^A - \rho\left(X_i\tau + Z_{it}^{p}\alpha\right)}{\sqrt{1-\rho^2}}\right)\right) p_i \left(1 - A_{it}\right) \\
& - \phi\left(X_i\tau + Z_{it}^{p}\alpha\right) \Phi\left(\frac{X_i\beta^A + \lambda^A p_i + Z_{it}^{A}\alpha^A - \rho\left(X_i\tau + Z_{it}^{p}\alpha\right)}{\sqrt{1-\rho^2}}\right) \left(1 - p_i\right) A_{it} \\
& \left. - \phi\left(X_i\tau + Z_{it}^{p}\alpha\right) \left(1 - \Phi\left(\frac{X_i\beta^A + \lambda^A p_i + Z_{it}^{A}\alpha^A - \rho\left(X_i\tau + Z_{it}^{p}\alpha\right)}{\sqrt{1-\rho^2}}\right)\right) \left(1 - p_i\right) \left(1 - A_{it}\right) \right] \\
& + \left(\gamma_1 + \rho\gamma_0\right) \left[\phi\left(X_i\beta^A + \lambda^A p_i + Z_{it}^{A}\alpha^A\right) \Phi\left(\frac{X_i\tau + Z_{it}^{p}\alpha - \rho\left(X_i\beta^A + \lambda^A p_i + Z_{it}^{A}\alpha^A\right)}{\sqrt{1-\rho^2}}\right) p_i A_{it} \right. \\
& - \phi\left(X_i\beta^A + \lambda^A p_i + Z_{it}^{A}\alpha^A\right) \Phi\left(\frac{X_i\tau + Z_{it}^{p}\alpha - \rho\left(X_i\beta^A + \lambda^A p_i + Z_{it}^{A}\alpha^A\right)}{\sqrt{1-\rho^2}}\right) p_i \left(1 - A_{it}\right) \\
& + \phi\left(X_i\beta^A + \lambda^A p_i + Z_{it}^{A}\alpha^A\right) \left(1 - \Phi\left(\frac{X_i\tau + Z_{it}^{p}\alpha - \rho\left(X_i\beta^A + \lambda^A p_i + Z_{it}^{A}\alpha^A\right)}{\sqrt{1-\rho^2}}\right)\right) \left(1 - p_i\right) A_{it} \\
& \left. - \phi\left(X_i\beta^A + \lambda^A p_i + Z_{it}^{A}\alpha^A\right) \left(1 - \Phi\left(\frac{X_i\tau + Z_{it}^{p}\alpha - \rho\left(X_i\beta^A + \lambda^A p_i + Z_{it}^{A}\alpha^A\right)}{\sqrt{1-\rho^2}}\right)\right) \left(1 - p_i\right) \left(1 - A_{it}\right) \right]
\end{aligned}
$$

We have a 3 step procedure:

1. STEP 1: Linear regression of p_i on the explanatory variables and instruments
2. STEP 2: Probit of A_{it} on the explanatory variables, instruments and the residual from step 1
3. STEP 3: Regression of Y_{it} on the explanatory variables and the control functions.

B Appendix B: Descriptive Analysis

In this section, we present a descriptive analysis of the variable used in our estimation, including the instruments. Table A1 presents the main statistics of the student´s characteristics. This table show that 69% of the students attended Preschool and 27% repeated at least once before 4th grade. In addition, 30% of the students have parents with high education, college or more. Table A2 shows the summary statistics of the school´s characteristics. 70% of the students are in public schools, 33% live in the Northeast. Table A3 shows the summary statistics of the professor´s characteristics. On average, the professors have about 13 years of experience.

Table B1: Descriptive Analysis - Student´s Characteristics

Variables	N Observations	Mean	Std. Dev.	Min	Max
Proficiency	41783	188,90	48,94	65,43	373,44
Preschool	45872	0,69	0,46	0	1
Grade Repetition	41783	0,27	0,44	0	1
Male	45872	0,45	0,50	0	1
White	45872	0,33	0,47	0	1
Age	41170	10,66	1,20	8	15
Mother´s Education					
No education	45872	0,02	0,15	0	1
Elementary School	45872	0,16	0,37	0	1
Middle School	45872	0,12	0,32	0	1
High School or more	45872	0,30	0,46	0	1
Missing	45872	0,40	0,49	0	1
Father´s Education					
No education	45872	0,03	0,18	0	1
Elementary School	45872	0,12	0,33	0	1
Middle School	45872	0,10	0,30	0	1
High School or more	45872	0,26	0,44	0	1
Missing	45872	0,48	0,50	0	1

TableB2: Descriptive Analysis - School´s Characteristics

Variables	N Observations	Mean	Std. Dev.	Min	Max
Public School	41783	0,70	0,46	0	1
Region					
North	45872	0,19	0,39	0	1
Northeast	45872	0,33	0,47	0	1
Southeast	45872	0,15	0,35	0	1
South	45872	0,12	0,33	0	1
Center	45872	0,13	0,33	0	1

Table B3: Descriptive Analysis - Professor´s Characteristics

Variables	N Observations	Mean	Std. Dev.	Min	Max
White	45872	0,38	0,48	0	1
Education	37329	14,30	1,91	8	17
Experince as a professor	37759	13,25	6,11	1	20
Salary	36953	1022,61	626,27	300	3100

Table A4 presents the summary statistics for the instruments at the student level, and table A5 shows the same statistics at the municipality level. These table shows that we have 1060 municipalities in our data. On average, these municipalities have 8 professors per preschool, and has 0,25 preschools per children with 4, 5 or 6 years old.

Table B4: Descriptive Analysis - Instruments (1998- 2001)

Variables	N Observations	Mean	Std. Dev.	Min	Max
Preschools	45585	311,98	536,65	1	2640
Professors at Preschool	45802	2.132,52	384,18	1	21179
Professors/Pop4a6	41731	5,71	2,48	0,20	26
Professors/Preschools	45585	8,62	5,73	1	91
Preschools/Pop4a6	41650	0,79	0,52	0,04	9
Cicle (2003)	45872	0,22	0,42	0	1

Table B5:Descriptive Analysis - Instruments (1998- 2001)

Variables	N Observations	Mean	Std. Dev.	Min	Max
Preschools	1050	37,55	147,70	1	2.639,75
Professors at Preschool	1057	274,80	1.045,72	1.75	21.178,75
Professors/Pop4a6	1056	5,66	3,25	0,20	26,33
Professors/Preschools	1050	10,14	8,00	1	72,43
Preschools/Pop4a6	1049	0,81	0,75	0,04	9,18
Cicle (2003)	1060	0,20	0,38	0	1

C Appendix C: Results

In this section, we present the table of the results, including all the covariates. Table C1 shows the results of step 1 that relates Preschool attendance with student´s characteristics, school´s characteristics, professor´s characteristics and the instruments. This table shows that mother´s education has an important effect on the probability of Preschool attendance.

Table C 1: Preschool Attendance (Step 1)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Male	-0.0230***	-0.0234***	-0.0234***	-0.0192***	-0.0195***	-0.0195***
	(0.00433)	(0.00432)	(0.00432)	(0.00483)	(0.00482)	(0.00482)
White	0.0217***	0.0208***	0.0208***	0.0194***	0.0186***	0.0186***
	(0.00469)	(0.00468)	(0.00468)	(0.00522)	(0.00521)	(0.00521)
Mother´s Education						
Elementary School	0.128***	0.127***	0.127***	0.140***	0.139***	0.139***
	(0.0148)	(0.0148)	(0.0148)	(0.0175)	(0.0175)	(0.0175)
Middle School	0.190***	0.187***	0.187***	0.195***	0.192***	0.192***
	(0.0153)	(0.0153)	(0.0153)	(0.0181)	(0.0180)	(0.0180)
High School or more	0.242***	0.239***	0.239***	0.248***	0.245***	0.245***
	(0.0150)	(0.0150)	(0.0150)	(0.0176)	(0.0176)	(0.0176)
Missing	0.160***	0.156***	0.156***	0.170***	0.168***	0.168***
	(0.0147)	(0.0147)	(0.0147)	(0.0174)	(0.0173)	(0.0173)
Father´s Education						
Elementary School	0.0798***	0.0809***	0.0809***	0.0920***	0.0924***	0.0924***
	(0.0131)	(0.0131)	(0.0131)	(0.0156)	(0.0155)	(0.0155)
Middle School	0.105***	0.105***	0.105***	0.122***	0.122***	0.122***
	(0.0136)	(0.0136)	(0.0136)	(0.0161)	(0.0161)	(0.0161)
High School or more	0.132***	0.131***	0.131***	0.141***	0.140***	0.140***
	(0.0131)	(0.0130)	(0.0130)	(0.0154)	(0.0154)	(0.0154)
Missing	0.0984***	0.0985***	0.0985***	0.112***	0.111***	0.111***
	(0.0125)	(0.0124)	(0.0124)	(0.0148)	(0.0148)	(0.0148)
Public School	-0.151***	-0.152***	-0.151***	-0.139***	-0.140***	-0.140***
	(0.00531)	(0.00531)	(0.00530)	(0.00584)	(0.00584)	(0.00584)
Region						
Northeast	0.0350***	0.0235***	0.0237***	0.0354***	0.0259***	0.0257***
	(0.00634)	(0.00651)	(0.00641)	(0.00721)	(0.00741)	(0.00729)
Southeast	0.0697***	0.0669***	0.0669***	0.0727***	0.0701***	0.0700***
	(0.00764)	(0.00763)	(0.00763)	(0.00835)	(0.00835)	(0.00835)
South	0.00565	-0.0170**	-0.0171**	0.0132	-0.00707	-0.00696
	(0.00837)	(0.00863)	(0.00861)	(0.00938)	(0.00968)	(0.00966)
Center	-0.00295	-0.00914	-0.00921	0.00657	0.000895	0.000961
	(0.00767)	(0.00769)	(0.00768)	(0.00861)	(0.00864)	(0.00863)
Teacher's race	0.00761	0.00643	0.00642	0.00428	0.00392	0.00393
	(0.00479)	(0.00478)	(0.00478)	(0.00534)	(0.00533)	(0.00533)
Teacher's education	0.00378***	0.00269**	0.00269**	0.00272**	0.00174	0.00174
	(0.00121)	(0.00122)	(0.00122)	(0.00137)	(0.00137)	(0.00137)
Teacher's experience	0.00167***	0.00169***	0.00170***	0.00161***	0.00164***	0.00163***
	(0.000384)	(0.000384)	(0.000384)	(0.000433)	(0.000432)	(0.000432)
Teacher's salary	5.52e-06	5.23e-06	5.21e-06	4.68e-06	4.53e-06	4.55e-06
	(4.03e-06)	(4.03e-06)	(4.03e-06)	(4.42e-06)	(4.41e-06)	(4.41e-06)
Professors/Pop4a6		0.0100***	0.0102***		0.00909***	0.00896***
		(0.00113)	(0.000924)		(0.00129)	(0.00105)
Preschools/Pop4a6		0.00105			-0.00110	
		(0.00557)			(0.00644)	
Constant	0.482***	0.452***	0.452***	0.481***	0.456***	0.456***
	(0.0237)	(0.0238)	(0.0238)	(0.0274)	(0.0275)	(0.0275)
Test de F *	-	0.0000	-	-	0.0000	-
Observations	35,041	35,041	35,041	26,140	26,140	26,140
R-squared	0.089	0.089	0.089	0.085	0.087	0.087

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
*p-value

Table C2 presents the results for the model of the probability of repeating a grade. This table shows that mother´s and father´s education are important determinants of the probability of repeating a grade. In addition, being in a public school increases the probability of repeating a grade.

Table C2: Grade Repetition (Step 2)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Male	0.240***	0.244***	0.244***	0.328***	0.324***	0.324***
	(0.0153)	(0.0170)	(0.0169)	(0.0207)	(0.0226)	(0.0226)
White	-0.112***	-0.117***	-0.117***	-0.149***	-0.145***	-0.145***
	(0.0168)	(0.0183)	(0.0183)	(0.0230)	(0.0246)	(0.0246)
Mother´s Education						
Elementary School	-0.168***	-0.196***	-0.194***	-0.411***	-0.383***	-0.386***
	(0.0465)	(0.0633)	(0.0630)	(0.0587)	(0.0881)	(0.0880)
Middle School	-0.289***	-0.330***	-0.328***	-0.584***	-0.545***	-0.550***
	(0.0487)	(0.0792)	(0.0787)	(0.0616)	(0.110)	(0.110)
High School or more	-0.571***	-0.623***	-0.621***	-0.953***	-0.903***	-0.909***
	(0.0481)	(0.0925)	(0.0919)	(0.0616)	(0.132)	(0.132)
Missing	-0.349***	-0.383***	-0.382***	-0.626***	-0.593***	-0.597***
	(0.0464)	(0.0702)	(0.0698)	(0.0587)	(0.0993)	(0.0993)
Father´s Education						
Elementary School	-0.177***	-0.194***	-0.194***	-0.204***	-0.186***	-0.188***
	(0.0417)	(0.0494)	(0.0492)	(0.0535)	(0.0686)	(0.0686)
Middle School	-0.223***	-0.245***	-0.244***	-0.312***	-0.288***	-0.291***
	(0.0438)	(0.0559)	(0.0557)	(0.0565)	(0.0804)	(0.0803)
High School or more	-0.339***	-0.367***	-0.366***	-0.408***	-0.380***	-0.384***
	(0.0424)	(0.0604)	(0.0600)	(0.0551)	(0.0860)	(0.0860)
Missing	-0.228***	-0.249***	-0.248***	-0.298***	-0.276***	-0.279***
	(0.0397)	(0.0513)	(0.0511)	(0.0510)	(0.0730)	(0.0730)
Public School	0.696***	0.728***	0.726***	0.965***	0.938***	0.941***
	(0.0218)	(0.0529)	(0.0525)	(0.0326)	(0.0722)	(0.0722)
Region						
Northeast	0.102***	0.0947***	0.0950***	0.134***	0.141***	0.140***
	(0.0218)	(0.0248)	(0.0246)	(0.0287)	(0.0336)	(0.0336)
Southeast	-0.265***	-0.280***	-0.280***	-0.281***	-0.266***	-0.268***
	(0.0297)	(0.0380)	(0.0379)	(0.0398)	(0.0533)	(0.0533)
South	-0.156***	-0.157***	-0.157***	-0.301***	-0.298***	-0.298***
	(0.0299)	(0.0303)	(0.0300)	(0.0410)	(0.0416)	(0.0416)
Center	-0.0866***	-0.0862***	-0.0862***	-0.109***	-0.108***	-0.108***
	(0.0274)	(0.0273)	(0.0274)	(0.0366)	(0.0368)	(0.0368)
Teacher's race	-0.0253	-0.0271	-0.0270	-0.0389*	-0.0379*	-0.0380*
	(0.0169)	(0.0170)	(0.0170)	(0.0227)	(0.0229)	(0.0229)
Teacher's education	-0.0123***	-0.0132***	-0.0131***	-0.0177***	-0.0171***	-0.0172***
	(0.00416)	(0.0044)	(0.0044)	(0.00550)	(0.00565)	(0.00565)
Teacher's experience	-0.00154	-0.00191	-0.00189	-0.00101	-0.000679	-0.000717
	(0.00134)	(0.0015)	(0.0015)	(0.00180)	(0.00196)	(0.00196)
Teacher's salary	-4.73e-05***	-4.86e-05***	-4.85e-05***	-7.03e-05***	-6.92e-05***	-6.93e-05***
	(1.47e-05)	(1.48e-05)	(1.48e-05)	(1.99e-05)	(2.01e-05)	(2.01e-05)
Cicle (2003)	-0.0952***	-0.0941***	-0.0941***	-0.218***	-0.219***	-0.219***
	(0.0200)	(0.0202)	(0.0202)	(0.0267)	(0.0268)	(0.0268)
Preschool	-0.241***	-0.0268	-0.0357	-0.446***	-0.644	-0.621
	(0.0175)	(0.314)	(0.320)	(0.0227)	(0.470)	(0.470)
Rediual 1st stage		-0.215	-0.206		0.199	0.175
		(0.323)	(0.319)		(0.471)	(0.471)
Constant	-0.158**	-0.261	-0.257	-0.121	-0.0265	-0.0376
	(0.0796)	(0.175)	(0.173)	(0.105)	(0.248)	(0.248)
Observations	35,041	35,041	35,041	26,140	26,140	26,140

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table C3 shows the results of the third step. These results show that not only mother´s and father´s education are important determinants, but also if the students are in a public school and teacher´s characteristics have a significant impact on Math proficiency.

Table C3: Proficiency (Step 3)

VARIABLES	OLS	OLS	(1)	(2)	OLS	OLS	(3)	(4)
Male	4.252***	5.552***	6.140***	6.135***	4.647***	5.932***	6.750***	6.743***
	(0.431)	(0.424)	(0.432)	(0.432)	(0.502)	(0.497)	(0.513)	(0.512)
White	4.358***	3.814***	3.538***	3.541***	4.980***	4.505***	4.193***	4.196***
	(0.467)	(0.458)	(0.472)	(0.472)	(0.542)	(0.535)	(0.537)	(0.537)
Mother´s Education								
Elementary School	4.626***	3.257**	3.662**	3.660**	4.276**	0.939	1.147	1.144
	(1.473)	(1.445)	(1.334)	(1.334)	(1.820)	(1.798)	(1.795)	(1.795)
Middle School	4.958***	2.688*	3.320**	3.319**	5.076***	0.569	0.783	0.776
	(1.530)	(1.502)	(1.399)	(1.399)	(1.881)	(1.862)	(1.858)	(1.858)
High School or more	16.77***	13.09***	12.88***	12.89***	16.64***	10.98***	10.20***	10.20***
	(1.497)	(1.471)	(1.385)	(1.385)	(1.839)	(1.825)	(1.830)	(1.830)
Missing	6.329***	3.687**	3.956***	3.954***	5.886***	1.157	0.996	0.992
	(1.463)	(1.437)	(1.329)	(1.329)	(1.807)	(1.790)	(1.787)	(1.787)
Father´s Education								
Elementary School	7.597***	6.154***	6.460***	6.458***	7.550***	5.749***	5.823***	5.818***
	(1.305)	(1.281)	(1.190)	(1.190)	(1.619)	(1.597)	(1.594)	(1.594)
Middle School	5.234***	3.468***	3.821***	3.819***	5.457***	3.014*	2.981*	2.975*
	(1.359)	(1.333)	(1.246)	(1.246)	(1.673)	(1.651)	(1.648)	(1.648)
High School or more	13.04***	10.81***	10.78***	10.79***	12.97***	10.42***	10.21***	10.20***
	(1.302)	(1.278)	(1.203)	(1.203)	(1.603)	(1.583)	(1.581)	(1.581)
Missing	9.566***	7.777***	8.015***	8.013***	9.606***	7.256***	7.149***	7.143***
	(1.241)	(1.218)	(1.129)	(1.129)	(1.543)	(1.524)	(1.521)	(1.521)
Public School	-38.26***	-35.12***	-33.06***	-33.09***	-37.56***	-34.96***	-32.70***	-32.72***
	(0.535)	(0.531)	(0.775)	(0.776)	(0.614)	(0.613)	(0.751)	(0.751)
Region								
Northeast	-0.920	-0.455	-0.0270	-0.0273	0.201	0.531	1.063	1.056
	(0.631)	(0.619)	(0.589)	(0.589)	(0.749)	(0.739)	(0.740)	(0.740)
Southeast	21.63***	19.99***	19.26***	19.27***	22.59***	21.02***	20.28***	20.28***
	(0.761)	(0.748)	(0.778)	(0.778)	(0.869)	(0.859)	(0.870)	(0.870)
South	16.95***	15.95***	15.36***	15.36***	18.06***	16.59***	15.65***	15.66***
	(0.833)	(0.818)	(0.808)	(0.808)	(0.975)	(0.963)	(0.972)	(0.972)
Center	10.38***	9.713***	9.247***	9.248***	10.74***	9.882***	9.458***	9.460***
	(0.763)	(0.749)	(0.730)	(0.730)	(0.895)	(0.883)	(0.884)	(0.884)
Teacher's race	2.395***	2.317***	2.218***	2.219***	2.226***	2.137***	2.085***	2.086***
	(0.477)	(0.467)	(0.474)	(0.474)	(0.555)	(0.547)	(0.546)	(0.546)
Teacher's education	0.999***	0.918***	0.907***	0.907***	1.056***	0.965***	0.910***	0.910***
	(0.121)	(0.119)	(0.117)	(0.117)	(0.142)	(0.140)	(0.141)	(0.141)
Teacher's experience	0.308***	0.295***	0.299***	0.300***	0.289***	0.278***	0.282***	0.282***
	(0.0383)	(0.0375)	(0.037)	(0.037)	(0.0450)	(0.0444)	(0.0443)	(0.0443)
Teacher's salary	0.00554***	0.00537***	0.00522***	0.00523***	0.00628***	0.00611***	0.00603***	0.00604***
	(0.000401)	(0.000394)	(0.000393)	(0.000393)	(0.000459)	(0.000453)	(0.000452)	(0.000452)
pre	14.73***	13.14***	8.789***	8.858***	16.52***	13.99***	4.780**	4.678**
	(0.532)	(0.523)	(1.798)	(1.800)	(0.643)	(0.641)	(2.005)	(2.002)
reprovado		-18.93***	-23.06***	-22.89***		-19.75***	-22.85***	-22.77***
		(0.508)	(3.040)	(3.049)		(0.716)	(1.582)	(1.580)
Control Function Pre			62.41***	62.64***			49.49***	49.19***
			(5.365)	(5.358)			(4.947)	(4.932)
Control Function Rep			20.96***	20.61***			24.77***	24.23***
			(3.541)	(3.574)			(4.694)	(4.674)
Constant	150.7***	160.6***	171.5***	171.5***	149.1***	161.4***	170.9***	170.9***
	(2.368)	(2.338)	(2.504)	(2.511)	(2.863)	(2.857)	(3.020)	(3.018)
Observations	35,041	35,041	35,041	35,041	26,140	26,140	26,140	26,140
R-squared	0.331	0.357	0.361	0.361	0.339	0.358	0.361	0.361

Standard errors in parentheses