The Impact of Child Labor on his/her Performance in School.

Ana L Kassouf; Marcos Garcias; Camila Rossi; Ida Bojicic Ono

SUMMARY

In Brazil, close to 12% or more than three million children and youths from 10 to 17 years old combine work with study, which can harm their school achievement by restricting time spent on assignments or by not allowing children to make efficient use of the time in school as the work make them tired. Understanding the behaviour of this group is crucial for understanding to what extent child labor interferes with children's school performance.

In this study we intend to analyze the impact of child labor on his or her academic performance using "Prova Brasil" school achievement tests from 2007 and 2011. Prova Brasil is a census data set with information on 5^{th} and 9^{th} grade students in urban public schools. From the available data we created a large panel data set with students in 5^{th} grade in 2007 and in 9^{th} grade in 2011.

The employment of children is a polemic issue. A group claims that the work performed by kids may bring benefits as they could learn from their jobs and be better off in the future. On the other hand, a young person who works cannot study as much as one dedicating full time in school and this could bring negative effects for his or her future. This study helps understand the consequences a child face if simultaneously study and work.

To measure the impact of kids working in the market and or in their households on Portuguese and Mathematics test scores we estimated fixed effect models controlling for year, states and students effects invariable in time and we weighted the models by the inverse probability weights to account for possible attrition bias. An instrumental variable approach proposed by Lewbel (2002) was also applied to the models to account for the endogeneity of child labor. Individual, parents, teachers and principal characteristics as well as school infrastructure were used as control variables.

The models were also estimated separately by gender. Results show that the work performed by children either in the household or in the market is detrimental to their academic performance. When using the whole sample, that is, boys and girls together, the largest negative effect occur when working in both outside and inside the house. However, when estimating only for girls the largest negative impact on test scores was when they were working only in the market and when estimating only for boys the largest impact appeared when they were working in both places.

It seems from these analyses that younger 5^{th} grade kids suffer larger harm from work compared to older 9^{th} grade kids and that girls working only in the market present the worst scenario in terms of lowering school achievement.

1. Introduction

In recent years, Brazil has experienced an impressive decline in children and youths labor. According to the national household survey (PNAD), in 1992, about 23 percent of Brazilian children and youths aged 10 to 15 worked, compared to 7 percent in 2014 (IBGE 2014). The Brazilian law prohibits the work of children under the age of 16, except for apprentices, in which case the minimum age is 14.

With respect to educational indicators such as illiteracy rates and years of schooling, Brazil still lags behind other Latin American countries. However, during the 1990s, school attendance increased, mainly in primary school and for students aged 7 to 14. In 1992, 87 percent of the children aged 7 to 14 attended schools. By 2014, this percentage reached 98 percent.

One of the reasons why Brazil continues to lag other countries in school achievement despite the increases in school attendance may be that a high percentage of students work while they attend school. According to the 2014 PNAD data, of more than 26.5 million Brazilian kids aged 10 to 17, 81 percent only study, 3.2 percent work outside their houses and do not study, 11.8 percent combine work with study, and 4.1 percent neither work nor study. This statistic shows that there are more than three million children and adolescents who continue to divide their time between working and studying, which could harm their school achievement by restricting time spent on assignments or by not allowing children to make efficient use of the time in school as the work make them tired.

This study contributes to the literature by analyzing the direct impact of child labor on the academic progress of students as measured by standardized achievement tests in Portuguese and Mathematics. The analyses are also performed separately by boys and girls to observe eventual differences and discriminations by gender.

Authors such as Gunnarsson et al. (2004), Psacharopoulos (1997), Heady (2003), Akabayashi & Psacharopoulos (1999), Stinebrickner and Stinebrickner (2003), Bezerra et al (2009), Dumas (2012), Emerson et al (2013) among others, studied the effect of early child labor on student achievement test scores in different countries. However, this present study differs from previous ones as we use a richer and more updated census data being able to create a panel and better control for observed and unobserved effects. Moreover, we investigate not only if children work outside their houses but also the number of hours spent in household activities performed by them. Specifically, students´ performance may be affected differently by the work conducted inside their household compared to the work performed outside their houses or in the labor market.

There are some studies in the literature looking at the effect of the work performed by children on their attendance rate rather than on performance, such as Ravallion and Wodon, (2000), Assaad et al. (2001), Canals-Cerda and Ridao-Cano (2004), Beegle et al. (2008 and 2006) and Edmonds (2008). However, in Brazil is common to see kids combining work and school and therefore looking at the effect of child labor on learning is even more important than on school enrollment.

The impact of child labor on learning may be negative if children divide their time between studying and working many hours in jobs that require lots of efforts, which could harm their school achievement. On the other hand, the impact of child labor on learning may be positive if the job involves tasks that result in learning and skills improvement. So, the direction of the expected impact of child labor on learning is unclear.

The main objective of this study is then to measure the impact of children labor force participation (here treated as domestic work, market work and hours spent in domestic work) on their learning outcomes. The estimates are based on students from urban public schools coming from Prova Brasil census data collected in 2007 and 2011. Kids in 5th grade in 2007 are merged with students in 9th grade in 2011 to create a panel data.

The richness of the data allows us to control for individual, parents, teachers and principal characteristics as well as school infrastructure. Moreover, having two years panel data, allows us to control for unobserved effects that do not vary in time.

We also correct for attrition bias using the *inverse probability weights* proposed by Wooldridge (2002) as weaker students that may repeat the year or abandon school dropout from 2007 to 2011 sample. We also address the endogeneity of child labor using an instrumental variable approach proposed by Lewbel (2012).

The results show that child labor reduce school performance in Portuguese and Mathematics and the larger the number of hours working in domestic tasks, the lower the kids´ achievement in school.

2. Bibliography Review

Research focusing on the impact of child labor on school performance in developing economies is scarce. In this section we present a short summary of the main studies analyzing the effect of children´s work on their performance in school in recent years. We focused on research using data from Brazil, but we also present relevant studies from other countries.

One of the most recent studies in this area is due to Emerson, Ponczek and Souza (2013), who investigated the impact of child labor on Portuguese and Mathematics standardized tests for children in grades 2 and 8 of elementary public schools in São Paulo from 2007 to 2010, using difference in difference approach. The annually dataset used - Prova São Paulo - allowed the authors to create a panel and to explore the causality between the work performed by children and its effect on their test scores. The authors observed that performing some sort of work while studying has a harmful effect on students´ scores in Portuguese and Mathematics.

Mavrokonstantis (2011) investigates the impact of child labour on mathematics test scores over a three year horizon in Vietnam. Using instrumental variables strategy he shows that the impact of child labour is negligible in rural areas, but in urban areas, child labour significantly impedes educational attainment. The instruments used were rice prices, assets and area of land owned by the household.

Gunnarsson et. al (2006) studied the impact of child labor on children´s test scores (Mathematics and Language) at 3rd and 4th grades of Elementary School in 9 Latin American countries. The results show negative impacts of child labor on Languages (Portuguese for Brazil and Spanish for other 8 countries) and Mathematics test scores. The "work" variable was treated as endogenous and exogenous in the model. To predict child labor they use as instruments the country's school starting age or truancy age. They claim that most children are involved in unpaid jobs and, because of that, market wages would not adequately capture the value of time outside school even if such information were available.
Edmonds (2008) analyzing school attendance and children´s work environments found that the school attendance of children working outside their houses is lower when compared to those working inside their houses. Also, children who work in both places in the house and outside have higher school attendance than those who work only outside their houses. The author claims that children who work outside their houses tend to spend more hours working than those who help in household tasks. Moreover, school attendance declines gradually with the increase in working hours and it becomes dramatically lower for children who work between 35 and 45 hours per week. Causal studies of the impact of child labor on schooling face the challenge of isolating some factor that affects child labor without simultaneously affecting schooling. This is difficult, as Edmonds claims, because child labor, schooling and leisure are not decisions that are mutually exclusive. According to the author, it is hard to imagine how one can be affected without all other decisions being affected. "Panel data on child labor histories is rarely available, so studies typically compare current labor supply to current attainment. … this is hard, because current work status necessarily depends on past education and work histories as these affect the value of child time and whether it's valuable for the child to work. This makes interpretation difficult, but studies typically find that attainment is lower for working children …" (page 3646).

Using data from the Monthly Employment Survey in six metropolitan regions of Brazil, in the period 1984-1997, Cavalieri (2002) evaluated the impact of child work on school performance, measured by repetition and dropout rates. To estimate such effects, the author uses propensity score matching method and estimates the difference in the average probability of approval in the series between the control and treatment groups (ie, children who did not work compared to children working). The results show that child labor increases the dropout rates and decreases approval rates in school for 10 to 14 years old children.

Beegle, Dehejia, Gatti and Krutikova (2009) using panel data from Tanzania evaluated the impact of child labor on education outcomes. They used the occurrence of crop and rainfall shocks as instruments for child labor. They found negative effects of child labor on school years and on the probability of completing primary school.

Dumas (2012) estimated the impact of working during childhood on proficiency tests between 1995 and 2003 in Senegal. The author found that children working under 17 hours a week had slightly better performance than the others. However, work had a detrimental effect when was over 17 hours a week. The author also found that a child being an employee when he started to work was highly detrimental to the learning process compared to a child being apprentice or being involved in family activities.

Bezerra et al. (2009) analyzed the impact of child labor on school achievement using Brazilian school achievement test data from the 2003 (SAEB). The authors tried to control for the endogeneity of child labor using as instruments the average wage for unskilled male labor in the state. The results showed that children and adolescents who do not work have better school performance than students who work. Up to two hours of work per day do not have a statistically significant effect on school performance, but additional hours decrease student's achievement. Moreover, working in the market has higher negative effect on school performance than working in the household.

The studies just described emphasize the importance of having good instruments to identify the models due to the endogeneity of child labor. At the same time, most researchers raise the issue of being extremely difficult to have good instruments given the available data. Recently, panel data analyses have provided tools to improve those results, but there are still concerns related to attrition and reverse causality that need to be addressed. Using a large data set that covers the whole Brazil and taking advantage of the panel structure, we add to this literature by accounting for attrition bias using inverse probability weights and by accounting for the endogeneity of child labor using an instrumental variable approach proposed by Lewbel (2012).

3. Methodology

In this study we create a ~~pseudo~~ panel using data from 2007 and 2011. We selected children in the 5th grade in 2007 and then we tried to find a similar child in the same school who is 4 years older in 2011, using variables such as gender, month and year of a child´s birth. Although we cannot know for sure if he or she is the same child, we believe that there is a large chance that we choose an equivalent child.

It is known in the literature that there is a reverse causality between child labor and schooling as students who do not perform well in school can lead to families deciding that they should invest more time in work. At the same time, children spending time in work activities in their own houses or in the market might do poorly in school.

Creating a two years panel with students information and using fixed effect model allows us to, at certain degree, control for both the endogeneity of child labor and the presence of other unobservables (e.g. kids ability and parental preferences) that are potentially correlated with the decision to work and the performance in school. We tried also to control for students, parents, teacher and principal characteristics as well as school infrastructure and students' motivation, using a rich set of variables available from Prova Brasil data.

Consider the variable of interest y, such as children´s performance in school, then,

$$y_{it} = \alpha_i + \gamma_t + \delta S_i T_t + \beta' x_{it} + \varepsilon_{it}$$

Where α is the individual fixed effect, γ is the time fixed effect, δ is the vector of parameters of the 27 states (S) time trends (T), β is the vector of parameters of the exogenous variables x and ε is the error term. The x variables include children´s, parents´, teacher´s, principals´ characteristics and school infrastructure, children´s work status, and others.

Our identification rests on the assumption that – after controlling for individual and year fixed effects, state time trends, observable child, parents, teacher, principal and school characteristics – the presence of children working in a given school in a given year is unlikely to be correlated with unobserved factors that determine the education outcomes examined in this paper (test scores in Portuguese and Mathematics).

A problem that may arise in estimating the fixed effect model in our case is that there are weaker students that may repeat the year or abandon school, resulting in non-random sample attrition. The resulting panel data obtained will then be formed by the best students that passed through the grades 5^{th} to 9^{th} without repetitions or dropouts. In this case we would incur in sample selection bias. To deal with this problem we use the inverse probability weights described in Wooldridge (2002). Briefly, the approach consists on estimating a probit model where the dependent variable assumes value one if the student is re-interviewed in 2011 and zero otherwise. The exogenous variable in the probit model is the lagged scores. Based on the estimated coefficient the probability of being re-interviewed is estimated and the inverse probability weights (IPW) are calculated as the inverse of the estimated probabilities. The bias vanishes when the IPW are used as weights in the fixed effect model.

Another problem we may have is that child labor is endogenous in the model and we would have to use an instrumental variable approach to estimate consistently the coefficient in the structural equation. However, as often happen in these kinds of analyses, we do not have a good outside instrument. To deal with this problem we apply the methodology proposed by Lewbel (2012), relying on the hetoroskedasticity of the error term and consisting on creating instrumental variables from the own model.

Details of the Wooldridge´s approach and the Lewbel´s method are given below.

3.1 Inverse Probability Weights

The dataset we are using suffer from attrition, as many individuals dropout of school, repeat the school year or change schools. Due to that, the remaining students are not representative of the original population and the results may be affected by attrition bias. The reason is that the individuals who drop out of a panel differ systematically from those who stay in it.

Consider a panel data having N individuals surveyed into two different years (t = 1, 2). Let s_{it} denote the selection indicator for each time period, where s_{it} = 1 if both y_{i1} and y_{i2} are observed, and zero when y_{i2} is not observed. Consider (x_{it}, y_{it}) are observed.

Wooldridge (2002) states that the sequential nature of attrition makes first differencing a natural choice to remove the unobserved effect:

$$\Delta y_{it} = \Delta x_{it}\beta + \Delta u_{it} \qquad t = 2,\ldots, T \qquad (1)$$

In our case, t = 2 (year 2007 and 2011). Let the score for individual i in the second year be y_{i2}, and in the first year y_{i1}, as well as the exogenous variables in the first year be x_{i1}, and in the second year be x_{i2}. Then, y_{i2} is observed only if there is no attrition. Considering z_{i1} an instrumental variable that only affect attrition, we can write a selection equation:

$$s_{it}{}^{*} = x_{i1}\gamma + z_{i1}\delta + v_i \qquad (2)$$

We do not observe $s_{it}{}^{*}$, but we do observe s_{it}, which takes the value 1 when both y_{i1} and y_{i2} are observed, and zero when y_{i2} is not observed.

Following Wooldridge (2002), ideally, at each t we would observe (y_{it}, x_{it}) for any unit that was in the random sample at t = 1. Instead, we observe (y_{it}, x_{it}) only if s_{it} = 1. According to Wooldridge (2012) "we can easily solve the attrition problem if we assume that, conditional on observables in the first time period, say, z_{i1}, (y_{it}, x_{it}) is independent of s_{it}", that is

$$\Pr(s_{it} = 1 | y_{it}, x_{it}, z_{it}) = \Pr(s_{it} = 1 | z_{it}) \qquad (3)$$

This assumption is called "selection on observables" because we assume that conditional on z_{i1}, selection is independent of (y_{it}, x_{it}) or that the distribution of s_{it} given [z_{i1}, (y_{it}, x_{it})] does not depend on (y_{it}, x_{it}).

Under attrition on observables, we can estimate the "Inverse Probability Weights" (IPW) to solve the problem of sample attrition. This method relies on an auxiliary variable that could be related to attrition and to the outcome variable, i.e., needs a much weaker assumption for the z_{i1} variable – u_{it} and v_{it} are independent [Fitzgerald et al. (1998)]. In this case, all variables that might affect the selection equation and the main equation are observables.[1]

There are two steps to obtain the Inverse Probability Weights. First we estimate a probit model of s_{it} on z_{i1} and let $\hat{p}_{it}$ be the fitted probabilities from this model. In the second step the score function in year 2 is weighted by $\left(1/\hat{p}_{it}\right)$, while in year 1the weight is one (for t = 1, $s_{it} = \hat{p}_{it} = 1$ for all i).

The intuition behind this procedure is that it gives more weight to individuals who have similar initial characteristics to individuals that subsequently attrit than to individuals with characteristics that make them more likely to remain in the panel.

The most frequent choice of the auxiliary variable in panel data is a lagged value of y according to Wooldridge (2002) and Fitzgerald et al., (1998). According to Moffit et al. (1999), who also studied sample attrition in panel data, "assuming serial correlation in

[1] A different way to solve the problem of sample attrition is through selection models, often referred to selection on unobservables. This case relies on a set of instrumental variables z_{i1} that is correlated with attrition but does not affect the outcome variable – u_{it} and v_{it} are correlated and z_{i1} is independent of u_{it} [Heckman (1979)]. Here z_{i1} cannot be part of x_{it}. An example is the unobservable variable such as individuals´ motivation affecting both the selection into the model and the outcome variable of interest.

the y process, such lagged variables will be related to current values of y conditional on x. If attrition is related to lagged y, least squares projection of y on x using the non-attriting sample will yield biased and inconsistent coefficient estimates. Estimation of attrition probabilities and subsequent weighted least square estimation yields consistent estimation instead"(page 136). In this study we use the score in Portuguese and Mathematics in 2007 as the z_{i1} variable.

3.2 Random Attrition Tests

The bias that attrition may cause comes from the fact that the observations lost follow a certain pattern, for example, weaker students who repeat the school year or dropout of school are not in the sample in the following period. Therefore, the weaker students are out of the sample and we are analyzing only the best students, which introduce bias in our results. Because of that, it is important to test if attrition is random. If the hypothesis of random attrition is not rejected, the results are not biased and no further action is needed. On the other hand, if we reject the hypothesis, attrition is not random and the inverse probability weights need to be used in the analyses.

To test the hypothesis of random attrition we look at the significance of the lagged variable coefficient used in the probit model. Statistically significant coefficients suggest that attrition bias might be present when estimating the score models. Another test of random attrition is the BGLW test due to Becketti, Gould, Lillard and Welch (1988). This test involves regressing the scores variable from the first wave of a survey (2007) on the explanatory variables, an attrition dummy, and the attrition dummy interacted with the other explanatory variables. An F-test of the joint significance of the attrition dummy and the interaction variables is then conducted to determine whether the coefficients from the explanatory variables differ between individuals who are retain or attrit from the panel.

3.3 The Lewbel´s Approach

Estimating the relationship between child labor and schooling is complicated because students who work might do poorly in school, but poor performance in school can also lead to families deciding that children should invest more time in work. School characteristics, family characteristics, and individual characteristics all affect both child labor and school achievement. A way to correct for the endogeneity of child labor as a right hand side variable is to use an instrumental variable approach. However, we do not have a good outside instrument, correlated with child labor and not correlated with the outcome variable, test scores.[2] To circumvent this problem we used the Lewbel (2012) approach, consisting on creating instrumental variables from the model and on identifying the coefficients in the model based on heteroskedasticity.

Following Lewbel (2012), consider the structural equation

$$y = x'\beta_1 + w\mu + \varepsilon_1 \qquad (4)$$

where

$$w = x'\beta_2 + \varepsilon_2 \qquad (5)$$

In our study, y represents test scores; x are exogenous variables, such as children, parents, teacher and school´s characteristics; w measures working status of children and ε_1 and ε_2 are unobserved errors.

[2] We did try to include as instruments the wage rates by sex, states, educational level etc from the National Household surveys (PNADs). However, the tests and results showed that those were weak instruments and we decided to discard them from this analysis.

If we have exclusion restrictions, that is, one or more elements of β_1 equal zero and the corresponding elements of β_2 nonzero, we could identify the model using two stage least squares, in which we estimate equation (5) to obtain the fitted values $\widehat{w}$ and then we estimate equation (4) on $\widehat{w}$ and on the subset of *x* that has nonzero coefficients.

However, very often we do not have exclusion restrictions and therefore instruments to identify the model. In general, variables affecting *y* also affect *w*.

Let *z* be a vector of observed exogenous variables, possibly being a subvector of *x* or even equal to *x*. In this case, Lewbel (2012) shows that under the assumptions,

$$E(x\varepsilon_1) = 0,\ E(x\varepsilon_2) = 0,\ cov(z, \varepsilon_1\varepsilon_2) = 0 \qquad (6)$$

along with some heteroskedasticity of the errors ε_1e ε_2 , the structural equation can be identified.

Defining matrices Ψ_{zx} and Ψ_{zz} by

$$\Psi_{zx} = E\left[\begin{pmatrix} x \\ [z - E(z)]\varepsilon_2 \end{pmatrix}\begin{pmatrix} x \\ w \end{pmatrix}'\right],\ \Psi_{zz} = E\left[\begin{pmatrix} x \\ [z - E(z)]\varepsilon_2 \end{pmatrix}\begin{pmatrix} x \\ [z - E(z)]\varepsilon_2 \end{pmatrix}'\right]$$

and let Ψ be any positive definite matrix that has the same dimension as Ψ_{zz}, Lewbel shows that,

$$\beta_2 = E(xx')^{-1}E(xw)$$

$$\begin{pmatrix} \beta_1 \\ \mu \end{pmatrix} = (\Psi'_{zx}\Psi\Psi_{zx})^{-1}\Psi'_{zx}\Psi\left[\mathrm{E}\begin{pmatrix} x \\ [z - E(z)]\varepsilon_2 \end{pmatrix} y\right]$$

This result means that β_2 and μ can be obtained by two stage least squares regression of *y* on *x* and *w* using *x* and [z - E(z)]ε_2 as instruments.

The estimation procedure is as follows. The coefficient β_2 is estimated by linearly regressing *w* on *x* to obtain the residuals $\hat{\varepsilon}_2$. Then β_1 and μ can be estimated by regressing *y* on *x* and *w* using *x* and $(z - \bar{z})\hat{\varepsilon}_2$ as instruments, where $\bar{z}$ is the sample mean of *z*. Let over bars denote sample averages, the resulting estimators are

$$\hat{\beta}_2 = (\overline{xx'})^{-1}\overline{xw}, \quad \hat{\varepsilon}_2 = w - x'\hat{\beta}_2$$

and

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\mu} \end{pmatrix} = (\widehat{\Psi}'_{zx}\widehat{\Psi}_{zz}^{-1}\widehat{\Psi}_{zx})^{-1}\widehat{\Psi}'_{zx}\widehat{\Psi}_{zz}^{-1}\begin{pmatrix} \overline{xy} \\ \overline{(z - \bar{z})\hat{\varepsilon}_2 y} \end{pmatrix}$$

4. Data

4.1 Microdata of Prova Brasil.

Prova Brasil is a census dataset for students in their 5th and 9th grades in urban public schools, collected by the Ministry of Education every 2 years. The great majority of students in Brazil, and mainly from lower income families, is in public schools. Students are expected to start school at the age of 6, so they would be 10 or 11 years old at grade 5 and 14 or 15 years old at grade 9.

Prova Brasil data set contains information on test scores and employment of students. The exams administered are standardized, multiple-choice designed to measure students' abilities and capacities in Portuguese (with a focus on reading comprehension) and Mathematics. As a way to compare different students throughout the years the reading and Mathematics tests administered to students use the Item Response Theory. The scores are mapped into cumulative performance scales, which means that students who are placed at a given level are competent at the skills required at the previous levels of the scale. Based on percentage scales, test scores classify students into levels of achievement in Portuguese and in Mathematics. The highest is the level reached, the best is the student´s performance. Level zero is a critical point where students are not able to read, calculate or understand the contents of the test. The scales are different for Portuguese and Mathematics as well as for 5^{th} and 9^{th} grades.

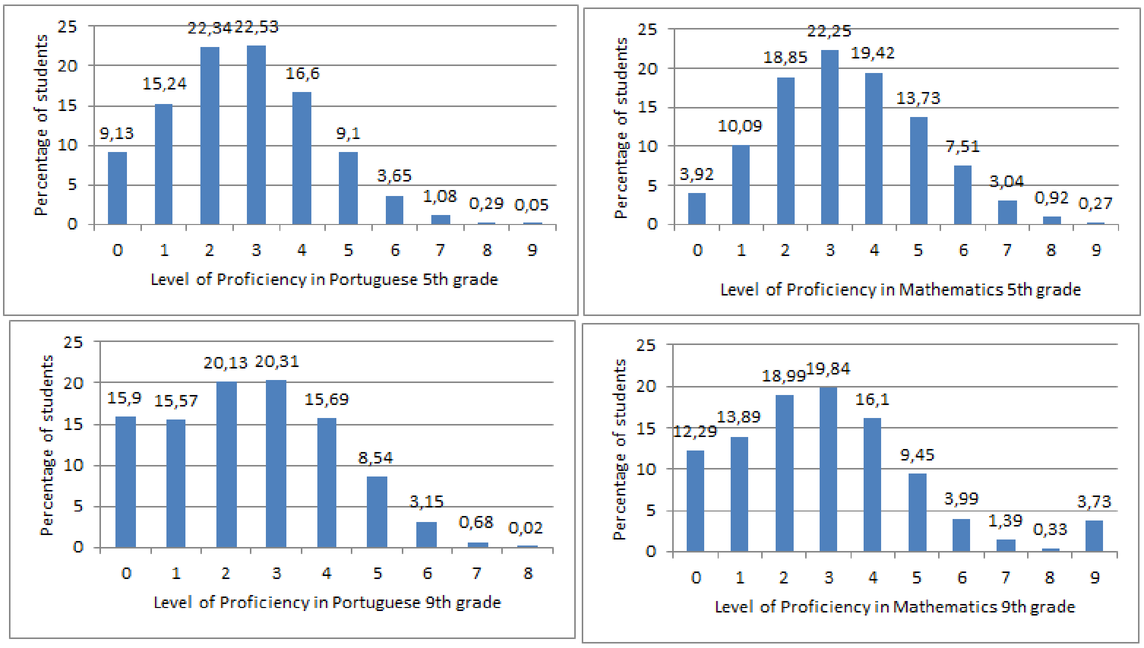
In the 5^{th} grade of primary education, the scale in the Portuguese test consists of ten possible levels, according to the score obtained: Level 0 (below 125), level 1 (between 125 and 150), level 2 (between 150 and 175), level 3 (between 175 and 200), level 4 (between 200 and 225), level 5 (between 225 and 250), level 6 (between 250 and 275), level 7 (between 275 and 300), level 8 (between 300 and 325), level 9 (between 325 and 350).

Similarly, the scale in the Mathematics test consists of 10 possible levels: Level 0 (below 125), level 1 (between 125 and 150), level 2 (between 150 and 175),), level 3 (between 175 and 200), level 4 (between 200 and 225), level 5 (between 225 and 250), level 6 (between 250 and 275), level 7 (between 275 and 300), level 8 (between 300 and 325), level 9 (between 325 and 375).

In the 9^{th} grade of primary education, the scale in the Portuguese test consists of 9 possible levels, according to the score obtained: Level 0 (below 200), level 1 (between 200 and 225), level 2 (between 225 and 250), level 3 (between 250 and 275), level 4 (between 275 and 300), level 5 (between 300 and 325), level 6 (between 325 and 350), level 7 (between 350 and 375), level 8 (between 375 and 400).

Similarly, the scale in the Mathematics test consists of 10 possible levels: Level 0 (below 200), level 1 (between 200 and 225), level 2 (between 225 and 250), level 3 (between 250 and 275), level 4 (between 275 and 300), level 5 (between 300 and 325), level 6 (between 325 and 350), level 7 (between 350 and 375), level 8 (between 375 and 400), level 9 (between 400 and 425).

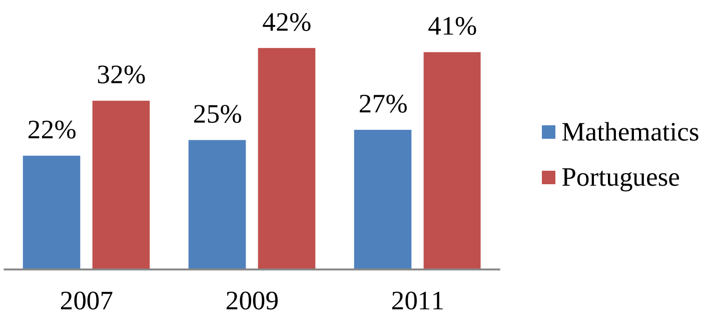
Figure 3.1 shows the percentage of students in 5^{th} and 9^{th} grade by each level of proficiency in Portuguese and Mathematics described above. Observe that more students are concentrated in lower levels of proficiency and a small percentage of kids are in higher levels of proficiency.



Source: Microdata of Prova Brasil

Figure 3.1 - Percentage of students in 5th and 9th grade in urban public schools in Brazil by the level of proficiency in Portuguese and Mathematics.

In the beginning of 2000 the National Institute of educational studies and analyses (INEP), from the Ministry of Education, created a classification of the proficiency levels dividing into four stages: very critical, critical, fair or basic and adequate. Figure 3.2 shows the percentage of students in the 9th grade with test score in levels basic or adequate in Portuguese and Mathematics. In 2007 32% of the students were in level suitable in Portuguese, increasing to 42% in 2009 and remaining almost the same in 2011 (41.3%). This means almost 60% of students in critical and very critical levels. In 2007 22% of the students had acceptable level in Mathematics, increasing to 25% in 2009 and 27% in 2011. Mathematics is even worse than language, more than 70% of the students are in critical levels.



Source: Microdata of Prova Brasil

Figure 3.2 - Percentage of students in the 9th grade of primary education in urban public schools in Brazil with Portuguese and Mathematics test scores in levels basic or adequate.

4.2 Creating a Panel Data Set.

In 2007 there were 48,745 schools in the data set, from which 16,121 had both 5th and 9th grade. Many schools had only primary level (1st to 5th grade) while others had only lower secondary level (6th to 9th grade). Since we had to merge 5th grade students in 2007 with 9th grade students in 2011, we did work only with schools that had both 5th and 9th grades. There were 1,815,010 students in the 16,121 schools, but only 975,065 students were in 5th grade in 2007. Similarly, in 2011 there were 16,586 schools with 5th and 9th grade from a total of 56,222 schools. There were 2,161,355 students in the 16,586 schools, but only 1,092,237 were in 9th grade in 2011. We merged 5th grade students in 2007 with 9th grade students in 2011 using year and month of birth, gender and school codes. Therefore, when one of these "merge variables" presented missing values we were not able to merge them and they were excluded from the data. After excluding those missing values we ended up with 835,175 students in 2007 and 841,977 students in 2011 to be merged. After merging the students we observed duplicated observations, that is, students with certain characteristics in 2007 who matched with more than one student in 2011 or vice versa. For example, a male student who was born in May 1996 and was in 5th grade in 2007 could have been merged with three 9th grade students in 2011 that were also male and were also born in May of 1996 in a given school. Therefore, from 835,175 students in 2007, there were 524,799 merged, including single pairs and duplicated merges. In duplicated cases we selected only the first observation, that is, in the final panel we have only one student in 2007 merged with only one student in 2011, resulting in 342,574 students. A summary of all the steps in merging the data as well as the number of observations in each step is presented in table 4.1. As a robustness check we selected different students in 2007 and 2011 when the merge was not perfect, not the first one. With the new panel data we estimated the same models presented in section 5, obtaining very similar results in all of the choices made.

Table 4.1 –Construction of the panel dataset.

Steps	2007	2011
Total number of schools	48,745	56,222
Schools with 5th and 9th grades	16,121	16,586
All students in schools with 5th and 9th grades	1,815,010	2,161,355
Students in 5th grade in 2007 and in 9th grade in 2011	975,065	1,092,237
Same as above without missing values in merge variables	835,175	841,977
After merging including duplicated students	524,799	524,799
After dropping duplicated students	342,574	342,574
After dropping all missing values in the variables	187,699	187,699

Source: Microdata of Prova Brasil.

From the above discussion on the creation of the panel data set in 2007 and 2011 we conclude that there was a large drop in the number of observations in the merge process, from more than eight hundred thousand to less than three hundred and fifty thousand. Students who repeated the school year or dropped out of school were not found in the subsequent survey year.

The data shows that 31.5% of the 5th grade students and 34.6% of the 9th grade students claimed that they have repeated at least one school year, while around 7.3% of the 5th grade student and 5.7% of the 9th grade students have dropped out of school at least once. Table 4.2 shows the number and percentage of students who have repeated or dropped out of school at least once in the past. The total number of observations in table 4.2 was supposed to be 835,175, but due to missing values in the analyzed variables, this number is lower.

Moreover, since one of the identifiers in the merge process is the school code, students changing schools in the analyzed period were also lost in the merge process. Also, any mistake a student had possibly made in answering his or her month and year of birth or his or her gender in one of the periods impairs his or her identification when merging. Taken into account all of these possible problems, we ended up with 524,799 students from a total of 835,175 students, losing therefore 310,376, possibly due to repetitions, drop outs or transfer schools in the period. Besides that, remained finally in the merge process 342,574 students from 524,799, losing therefore 182,225 observations (86,270 in 2007 and 95,955 in 2011) due to data duplications, i.e., we left in the final data base observations with one to one merge, all the others duplicated ones were excluded from the data base.

Table 4.2 - Number and percentage of 5th and 9th grade students who repeated at least one grade in the past or dropped out from school*.

Grade Failure	2007 5th grade		2011 9th grade	
	number	%	number	%
No	539,493	68.54	542,263	65.41
Yes	247,632	31.46	286,794	34.60
Drop out				
No	730,108	92.73	783,106	97.07
Yes	57,267	7.27	47,647	5.73

Source: Microdata of Prova Brasil.
* The observations can vary from variable to variable because of missing values.

Besides data on test scores and work performed by children, Prova Brasil has information on students, teacher, principal and school characteristics. Some of the available information is: children´s age, gender and race, mothers´ education, fathers´ education, family size, family possession of goods (TV, washing machine, refrigerator, computer etc), reading habits, encouragement of the family towards children´s education, teacher´s age, gender, education , wage, number of years of experience, number of hours teaching per week, use of didactic equipment, principal´s age, gender, education, number of years of experience, school maintenance status, school infrastructure, presence of computer, internet, libraries, sport facilities, music, science labs etc.

Finally, the drop from 342,574 to 187,699, as presented in the last line of table 4.1, is due to missing data in the variables used in the econometric models and due to the use of a balanced panel, the number of observations actually used in the econometric analyses is then 375,398 (187,699 in 2007 and 187,699 in 2011). Since we have only two years panel data, singleton groups are not used in the estimations of the Fixed Effect Models and every missing value in any specific variable results in the elimination of the whole line. Hence, all the subsequent analyses use two years panel with 375,398 observations.

Table 4.3 has the description of all the variables used in the econometric models as well as the mean and standard deviation of each variable calculated based on the 375,398 observations.

The minimum required level of proficiency (basic level) differs from 5^{th} to 9^{th} grade and from Portuguese to Mathematics. A child supposedly knows Portuguese basic level if he/she gets 200 points at grade 5 and 225 at grade 9. Similarly, in Mathematics the minimum score to reach a basic level is 275 at grade 5 and 300 at grade 9. In 5^{th} grade, the averages scores in Portuguese is 185 and in Mathematics is 201.5 while in 9^{th} grade is 250.5 in Portuguese and 256.5 in Mathematics. These numbers show that on average students are below the basic level. Although presenting the mean and standard deviation of the scores, in the econometric models we use standardized test scores (mean zero and standard deviation one). Since we are comparing students´ scores in grades 5 and 9, we decided to use standardized scores for full comparability across grades.

Table 4.3 - Description of the variables, mean and standard deviation in 2007 and 2011.

Variables	Description of the Variables	2007			2011		
		obs	mean	s.d.	obs	Mean	s.d.
Scores_Portuguese	Portuguese test Score	187699	185.03	41.1	187699	250.5	45.1
Scores_Mathematics	Mathematics test Score	187699	201.48	43.3	187699	256.5	45.9
Hours_work_hh	Hours in hh activities/day	187699	1.43	1.12	187699	1.47	1.06
Work_market	=1 if child works outside hh	187699	0.09	0.29	187699	0.15	0.36
Not_working	don't work	187699	0.59	0.49	187699	0.50	0.50
Work_hh	Work only in the hh	187699	0.32	0.47	187699	0.35	0.48
Work_market	Work only in the market	187699	0.05	0.22	187699	0.09	0.29
Work_both	Work in both	187699	0.04	0.20	187699	0.06	0.24
White	Whites + Asians	187699	0.40	0.49	187699	0.40	0.49
Grade_failure	N. years repeat school year	187699	0.23	0.51	187699	0.23	0.51
Hh_member	Number of people in hh	187699	4.91	1.47	187699	5.91	1.65
Car	Number of cars in hh	187699	0.57	0.72	187699	0.66	0.75
floor_sch	=1 if floor in school	187699	0.59	0.49	187699	0.64	0.48
Start_maternal	=1 if child start 2 to 4 years	187699	0.39	0.49	187699	0.33	0.47
Start_preschool	=1 if child start 4 to 6 years	187699	0.37	0.48	187699	0.46	0.50
Start_grade1	=1 if start school at 6 or 7	187699	0.18	0.39	187699	0.19	0.39
Age_teacher	Age of the teacher	187699	40.84	8.71	187699	42.43	8.55
Experience_teacher	Number of years teaching	187699	14.35	6.71	187699	14.84	6.95

Source: Microdata of Prova Brasil.

Children respond if they work or not outside their house and if they work in household tasks. Moreover, if they perform household tasks, they report the number of hours spent on those activities. Table 4.4 shows the number and percentage of students in the 5^{th} and 9^{th} grade, according to their work status in 2007 and 2011.

Table 4.4 - Number and percentage of 5th and 9th grade students, according to their work status°.

Work Status	2007 5th grade number	%	Average hours/day spent in domestic work	2011 – 9th grade number	%	Average hours/day spent in domestic work
Do not work*	110,306	58.77	0.71	92,946	49.52	0.74
Work only in the hh	59,695	31.8	2.67 (20.29)	66,626	35.5	2.51 (12.35)
Work only in the market	9,631	5.13	0.74	17,051	9.08	0.64
Work in both	8,067	4.3	2.84 (28.66)	11,076	5.9	2.68 (21.58)
Total	187,699	100	1.42 (7.68)	187,699	100	1.47 (5.66)

Source: Microdata of Prova Brasil.

*Considered not working if worked 1 hour or less in the household per day.

º Numbers between brackets show the share of observations spending 4 hours or more per week in domestic work.

Close to 59% of the students in 5th grade neither work in the household nor in the market. Close to 32% work only in the household, 5% work only in the market or outside their houses and 4% work in both, outside and inside their houses. The average hours spent in domestic work per day is larger when the kids work in both household and market, spending 2.84 hours/day. When they work only in the household they spend 2.67 hours/day. The percentage of students working increase with age (or grade) as it can be observed in Table 4.4 for 9th grade students where 9.1% work in the market and 6% in both.

Table 4.5 has the number and percentage of boys and girls working in the market or in the household. In 5th grade, 40% of girls work in the household against 31% of boys or close to 10 percentage points more girls perform household tasks. However, in 9th grade the difference is larger, 53% of girls work in the household and 27% of boys or 26 percentage points more girls compared to boys. The figure is different when looking at labor performed outside the house or in the market, where the number of boys is more than twice the number of girls. In 5th grade, 14% of boys work in the market compared to 6% of girls, while in 9th grade, 21.5% of boys work in the market compared to 10% of girls. The values shown in Table 4.5 differ from those in Table 4.4. Table 4.5 shows number of kids working in the market or in the household, independent if they are working in both or only in one of the categories, while Table 4.4 displays the number of kids working only in the market, working only in the household or in both places.

Table 4.5 - Number and percentage of 5th and 9th grade students working in the market or in their household.

Work	2007 – 5th grade				2011 – 9th grade			
	Boys		girls		boys		girls	
	number	%	number	%	Number	%	number	%
work_market	11,580	13.91	6,118	5.86	17,871	21.47	10,256	9.82
work_house*	25,768	30.95	41,994	40.20	22,396	26.90	55,306	52.95

Source: Microdata of Prova Brasil.

*Considered not working if worked 1 hour or less in the household per day.

The number of hours a child spends working in their own household per day is presented in Table 4.6. Around 82% of the children spend time in household activities. Some studies show that giving children household chores helps to form accountability and self-confidence and that they are more likely to succeed in adulthood [Rossmann (2002)]. However, if a child is overloaded with household chores, working a large number of hours per day can harm his or her future life as less time is allocated to studying and doing homework. Because of that, when a child claimed to be performing household tasks one hour or less per day, we considered that he/she was not working.

Table 4.6 - Number and percentage of 5th and 9th grade students, according to the number of hours they work in their household per day.

	2007 5th grade		2011 – 9th grade	
Hours working hh/day	number	%	number	%
Do not work in the hh	34,578	18.42	30,681	16.35
1 or less hr/day	85,359	45.48	79,316	42.26
2 hours/day	35,479	18.9	46,745	24.9
3 hours/day	17,861	9.52	20,338	10.84
4 or more hrs/day	14,422	7.68	10,619	5.66
Total	187,699		187,699	

Source: Microdata of Prova Brasil.

Table 4.7 has the same information as Table 4.6 but stratified by gender. Observe that girls not only work more in the household than boys but they spend more hours doing household tasks. In the 9th grade 23% of girls work 3 or more hours a day in household activities compared to 9% of boys.

Table 4.7 - Number and percentage of 5th and 9th grade students, according to the number of hours per day they work in their household by gender.

Hours working hh/day	2007 – 5th grade				2011 – 9th grade			
	Boys		Girls		Boys		Girls	
	number	%	number	%	number	%	number	%
not work in hh	20,095	24.14	14,483	13.87	22,109	26.56	8,572	8.21
1 or less hr/day	37,382	44.91	47,977	45.93	38,740	46.54	40,576	38.85
2 hours/day	13,966	16.78	21,513	20.6	15,150	18.2	31,595	30.25
3 hours/day	6,403	7.69	11,458	10.97	4,526	5.44	15,812	15.14
4 or more hr/day	5,399	6.49	9,023	8.64	2,720	3.27	7,899	7.56
Total	83,245		104,454		83,245		104,454	

Source: Microdata of Prova Brasil.

5. Results

In 2007 there are 835,175 students in 5th grade that we try to merge with "similar" students in 2011, when they were supposed to be in 9th grade. If every 5th grade student pass the year and stay in the same school, he or she would be in the 9th grade in 2011. However, many students may change schools, may repeat the year or may drop out from school, causing attrition bias in our sample. In our data, only 342,574 students are found in 2011. To deal with this problem we use the Inverse Probability Weight (IPW) approach presented by Wooldridge (2002). The procedure consists on estimating a reduced form probit models using the whole sample of students in 2007 who are in schools with both 5th and 9th grades, with the dependent variable equals 1 if the student in 2007 was re-interviewed in 2011 and 0 otherwise. The regressor is the scores in 2007. After estimating the coefficients of the probit model we obtain predicted probabilities of being re-interviewed. The Inverse probability weights are calculated as the inverse of the predicted probabilities. These weights are used in the fixed effect panel and in the instrumental variable panel models taking value 1 in the first period (2007) and taking the inverse of the predicted probabilities in the second period (2011).

The estimated coefficients of the probit model that has as a dependent variable value 1 if the student is in the sample in 2007 and in 2011 and 0 otherwise are presented in Table 5.1. Observe that the coefficients of the score variables are highly statistically significant, indicating that attrition bias might be present when estimating children´s school performance models. Also, the largest the kids´ test scores the largest the probability to stay in school and therefore in the sample, as we expect. For the BGLW attrition test we estimate the score models using both the whole sample of students in 2007 in schools with 5th and 9th grades and the one-to-one merged sample (non-attriting) in 2007, testing the equality of the coefficients joinly. The *F*-tests are highly significant and reject the equality of coefficients, that is, the joint impact of the predetermined variables for the whole sample and the non-attriting sample are statistically different. From these results we conclude that attrition is not random and we need to use the inverse probability weights to correct for sample attrition bias.

Table 5.1 – Coefficients of the Probit Model.

Variables	Coefficients of the Probit Model Dependent variable is equal to 1 if student is in 2007 and in 2011 and 0 otherwise	
Scores_portuguese	0.202***	-
	(0.0014)	-
Scores_mathematics	-	0.184***
	-	(0.0014)
	-0.228***	-0.229***
Constant	(0.0014)	(0.0014)
Pseudo R^2	1.54%	1.85%
Observations	833,278	833,278

Based on the estimated coefficients of table 5.1, the probability of the students being re-interviewed was estimated and the inverse probability weights were obtained to be used in the fixed effect models having test scores as dependent variables.

Tables 5.2 and 5.3 present individual level estimates of the impact of child labor on students´ performance at school measured by standardized test scores in Portuguese and Mathematics, respectively, using fixed effect models.

All regressions have year fixed effects, individual fixed effects, and state level time trends. The first column is the simplest specification, excluding even direct effects of child, parents, teacher and school variables. The estimated impacts for the two outcomes are highly significant, with the expected signs. Students working in the household and in the market (*work in both*) have 0.23 standard deviation (s.d.) lower test scores in Portuguese, while those working only in the market have their score reduced by 0.24 s.d. and those working only in the household have their score reduced by 0.06. Similarly the reduction in Mathematics test scores is .20 s.d. (table 5.3) when working in the household and in the market, 0.16 when working only in the market and 0.055 s.d. when working only in the household.

In the third column we also excluded the direct effects of child, parents, teacher and school variables, including however a binary variable if a person works in the market or not, independent if a child is working in the household or not and the number of hours working in the household per day. Again the estimated impacts for the two outcomes are highly significant, with the expected signs. Students working in the market have their score reduced by .21 s.d. in Portuguese and .15 s.d. in Mathematics. Moreover, each additional hour working in household tasks reduce their score in .03 for both Portuguese and Mathematics.

A possible problem that may arise is omitted variable bias. Despite using individual and year fixed effects, and state time trends, child labor may be correlated with trends in students, parents, teachers and school characteristics that directly affect test scores. In the first and third columns of Tables 5.2 and 5.3 we did not include control variables, while in the second and fourth columns we added 14 control variables. Adding control variables yields estimates of child labor impacts a little smaller in magnitude, as expected, but still very similar to the first and third columns. Most of these variables are significant, with the expected signs. One can claim that there are other important variables affecting test scores not included in the model, such as parents´ and teachers´ education. However, we chose to use only the 14 variables presented in Tables 5.2 and 5.3 to avoid dropping even more observations due to missing values observed in the variables. In table A1 of the appendix we present the same panel including 33 exogenous variables instead of 14 and the coefficients of the children´s work variables were very similar to the ones presented in columns (2) and (4) of Tables 5.2 and 5.3. Table A2 in the appendix has the description of the 33 variables, mean and standard deviation.

Estimates in columns (1) to (4) of tables 5.2 and 5.3 may be biased due to attrition as described before. Columns (5) and (6) present new estimates using the inverse probability weights (IPW) obtained from the probit model (Table 5.1) as proposed by Wooldridge (2002). The new estimates for the working variables are practically the same as the ones presented in columns (2) and (4). Despite of the fact that we included individual fixed effect, control variables and we controlled for attrition bias, one can still claim that the estimates in columns (5) and (6) are biased due to the endogeneity of child labor. Trying to take the endogeneity problem into account we use an instrumental variable (IV) approach developed by Lewbel (2002), whose results are presented in columns (7) to (10)[3]. For comparison reasons, results in columns (7) and (8) include IV and exclude IPW, while in columns (9) and (10) include both IV and IPW. Again, there is almost no difference in the magnitude of the coefficients between columns (7)/(8) and (9)/(10). However, when including IV, all the kids´ work coefficients are negative and significant as before but the magnitudes of the coefficients change. The larger negative impact is when kids work in both the market and the household (-0.33 for Portuguese scores and -0.31 for Mathematics scores) followed by when they work only in the market (-0.28 for Portuguese scores and -0.19 for Mathematics scores) and lastly when

[3] We implemented tests for hetorskedaticity of the errors as required in Lewbel´s method together with other model assumptions showed in equation (6) of the methodology and they were satisfied.

they work only in the household (-0.24 for Portuguese scores and -0.13 for Mathematics scores). Similar results were observed in the other columns.

It is reasonable to assume that children working in the market are more exposed to harm and to more demanding activities when compared to those working in their households and close to their families. Also, they probably face more working hours on average as they need extra traveling time to go to work. Working in the market and in the household translates into a heavier workload. Because of that we expect the negative impact of working in both market and household on children´s school performance to be higher than the impact of working only in the market and this to be larger than working only in the household, as our results showed.

The change in the magnitude of the work coefficients when using IV approach reflects the endogeneity or the reverse causality between child labor and the test scores. The largest increase in the magnitude of the coefficients is observed for the variable "work only at home", in which case the coefficient increased almost five times for Portuguese and almost three times for Mathematics test scores, comparing the columns with and without IV. We can say that there is more flexibility in household tasks as children work in their own household and under the supervision of their parents. So, parents aware of the fact that their children are not doing well in school can immediately impede their work or reduce their hours of work at home. On the other hand, when working in the market, besides the need of income, there are contracts and third parties involved, making changes less flexible. Based on these arguments we can say that working in the household is more flexible~~endogenous~~ than working in the market, the reason why there was a larger change in the magnitude of the work in the household coefficient, reflecting endogeneity.

As it follows we quickly analyze each of the control variables included in the models.

Race white include white and Asian children, while blacks and mulattos were omitted. This variable reflects, besides cultural and ethnic aspects, socioeconomic level of the population, since a large number of blacks and mulattos are poor. Prova Brasil data show that the average test scores in Portuguese and Mathematics are above average for white and Asian children and below average for black and mulatto children, reflecting differences in ethnic groups, including education outcomes. The coefficients were positive and significant for the whites and Asians, ~~reflecting~~ corroborating with the fact that those kids perform better in school than blacks and mulattos.

The repetition variable (grade*_failure*) reflects the number of times a child repeated the school year in the past, once or twice or more. Those variables negatively affect test scores, since they reflect students that have the worst academic performance.

The larger the number of people in the student´s household (*hh_member*) the lower is his/her test scores. Studies show that in large families the provision of goods may be scarce and often older kids work to help the family budget [Emerson & Souza (2002)].

Students in higher income families are expected to have a better performance in school. However, the data used in this study do not have individual or family income information. To circumvent this problem, we used the number of cars owned by families as proxy for family income. Three binary variables represent families possessing one car, two cars or three or more cars. There are other family possessions available in the data set, such as: VCRs, color TV, computers, etc. as well as the number of bathrooms and if the household has a housekeeper. We tried to use principal components to create an income index with all those variables. However, due to a large number of missing values, the coefficients were not significant at 10% level or less and so we chose to use only the variable cars as a proxy of income. Close to 50% of the children do not have car in their households, but the average scores for those with car is higher than for those without car. As expected the coefficients for those owning more cars are positive.

Children starting to attend school before the age of 6 or 7 (*start_preschool and start_maternal*) had better academic performance than those that started after 7 years old (omitted). In Brazil children from 2 to 4 years old are supposed to be in maternal schools while kids from 4 to 6 years old are in preschool. According to the law, a six years old kid has to be in the first grade. However, many kids actually start school after completing 7 years of age. It is not mandatory to start school before first grade but the effect of starting school from 4 to 6 years old on test scores was the largest one. Similarly, although having less effect than in preschool, those that started from 2 to 4 years old (*start_maternal*) and those that started at 6 or 7 years old (*start_grade1*) had better test scores than those that started later in school. Paes de Barros et al (2001) analyzed the impact of pre-school and kindergarten on different outcomes. They cited a large number of studies in Colombia, Peru, Jamaica, Turkey etc. showing that early childhood development has positive effects on physical, mental and economic wellbeing of a person. Specifically, they cited that intervention in children between 0 and 6 years old improve health, nutrition and cognitive development, increase enrollment and decrease dropout.

School infrastructure is measured by the variable *floor_sch*, if there is good floor condition in the school. The coefficient is positive and significant as expected.

It seems plausible that teachers have an important task in improving students´ academic performance. Trying to measure these effects, we included in the regressions teachers´ age and experience. The inclusion of age squared allows a parabolic curve and the signs of the coefficients show a concave downward parabolic curve. Experience had positive impact on test scores. Glewwe et al (2011) review 43 higher quality studies to investigate which specific school and teacher characteristics have strong positive impacts on learning. According to the authors, “there is little evidence that teachers’ level of education has any impact on student test scores”. Also, teacher gender had an ambiguous impact. On the other hand, they concluded that teacher experience showed a positive effect on test scores. Hanushek & Rivkin (2006) also reviewed studies on teacher quality and concluded that teachers´ advanced degrees did not have systematic relationship to student outcomes. Glewwe & Kremer (2006), after reviewing retrospective studies of the determinants of learning, concluded that “there are no general results regarding which teacher and school variables raise learning in developing countries”.

Table 5.4 shows the effect of child labor on test scores in Portuguese and Mathematics stratified by gender using the inverse probability weights and the instrumental variable approach. Columns (1) to (4) show the coefficients for boys and columns (5) to (8) for girls. Observe that the effect of boys working in both the market and the household on test scores was almost double of this effect for girls. On the other hand, when working only in the market the effect for girls was higher than for boys. Columns (3), (4), (7) and (8) show the coefficients of working in the market and number of hours performing household tasks. In both cases the effects were larger for boys.

Table 5.5 shows a summary of the impacts of the working variables on test scores in Portuguese and Mathematics for the whole sample and separately by boys and girls. The impacts were calculated based on the coefficients of columns (9) and (10) of table 5.2 and 5.3 and on the coefficients of table 5.4 with the average and standard deviation of the scores in table 4.3. For example, in column (9) of table 5.2, the impact of working only in the household on Portuguese test scores was 0.24 s.d.. This value was multiplied by the standard deviation of 41.1 and then divided by the mean of the score in 5^{th} grade 185.03(see table 4.3) resulting in a Portuguese score decrease of 5.24%.

The largest observed impact was for 5th grade girls who suffered a reduction of 10% in their Portuguese test scores when they worked only in the market. The largest impact for boys was a reduction of 8.3% in their 5th grade Portuguese test scores when they worked in both outside and inside their home. The largest impact for the whole sample, including boys and girls, was a reduction of 7.4% in their 5th grade Portuguese test scores when they worked in both outside and inside their home, while the smallest impact was on Mathematics test scores (2.4%) for 9^{th} grade students working in the household. Table 5.5 also shows that for each hour involved in household tasks, kids test scores decrease from 0.9% up to 2.4%.

Table 5.2 –Coefficients of the fixed effect models with and without IPW and IV for test scores in Portuguese in the 2007/ 2011 panel.

Variables	Portuguese									
	Without IPW and without IV				With IPW and Without IV		Without IPW and With IV		With IPW and With IV	
	Col (1)	Col (2)	Col (3)	Col (4)	Col (5)	Col (6)	Col (7)	Col (8)	Col (9)	Col (10)
Work only at home	-0.057***	-0.051***	-	-	-0.050***	-	-0.237***	-	-0.236***	-
	(0.004)	(0.004)	-	-	(0.005)	-	(0.011)	-	(0.012)	-
Work only in the market	-0.239***	-0.222***	-	-	-0.222***	-	-0.278***	-	-0.276***	-
	(0.008)	(0.008)	-	-	(0.009)	-	(0.015)	-	(0.016)	-
Work in both	-0.228***	-0.207***	-	-	-0.208***	-	-0.337***	-	-0.331***	-
	(0.009)	(0.009)	-	-	(0.010)	-	(0.023)	-	(0.025)	-
Work in the market	-	-	-0.209***	-0.193***	-	-0.194***	-	-0.215***	-	-0.210***
	-	-	(0.006)	(0.006)	-	(0.007)	-	(0.015)	-	(0.017)
Hours worked at home	-	-	-0.030***	-0.027***	-	-0.026***	-	-0.113***	-	-0.110***
	-	-	(0.002)	(0.002)	-	(0.002)	-	(0.005)	-	(0.005)
White	-	0.057***	-	0.057***	0.056***	0.055***	0.054***	0.052***	0.052***	0.051***
	-	(0.005)	-	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
grade_failure_1	-	0.270***	-	0.271***	0.266***	0.267***	0.271***	0.273***	0.267***	0.269***
	-	(0.013)	-	(0.013)	(0.014)	(0.014)	(0.013)	(0.013)	(0.014)	(0.014)
grade_failure_2	-	-0.014	-	-0.013	-0.013	-0.011	-0.010	-0.007	-0.009	-0.005
	-	(0.012)	-	(0.012)	(0.014)	(0.014)	(0.012)	(0.012)	(0.014)	(0.014)
car_1	-	0.018	-	0.019	0.019	0.020	0.018	0.021	0.020	0.022
	-	(0.013)	-	(0.013)	(0.015)	(0.015)	(0.013)	(0.013)	(0.015)	(0.015)
car_2	-	0.050***	-	0.051***	0.050***	0.051***	0.051***	0.053***	0.051***	0.052***
	-	(0.013)	-	(0.013)	(0.015)	(0.015)	(0.013)	(0.013)	(0.015)	(0.015)
car_3	-	0.046***	-	0.047***	0.046***	0.047***	0.046***	0.048***	0.046***	0.047***
	-	(0.014)	-	(0.014)	(0.016)	(0.016)	(0.014)	(0.014)	(0.016)	(0.016)
Hh_member	-	-0.053***	-	-0.053***	-0.053***	-0.053***	-0.052***	-0.051***	-0.052***	-0.051***
	-	(0.001)	-	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)

Table 5.2 – continue.

Variables	Portuguese									
	Without IPW and without IV				With IPW and Without IV		Without IPW and With IV		With IPW and With IV	
	Col (1)	Col (2)	Col (3)	Col (4)	Col (5)	Col (6)	Col (7)	Col (8)	Col (9)	Col (10)
Start_maternal	-	0.232***	-	0.232***	0.230***	0.229***	0.229***	0.227***	0.226***	0.225***
	-	(0.010)	-	(0.010)	(0.011)	(0.011)	(0.010)	(0.010)	(0.011)	(0.011)
Start_pre_school	-	0.269***	-	0.268***	0.267***	0.267***	0.266***	0.265***	0.265***	0.264***
	-	(0.010)	-	(0.010)	(0.011)	(0.011)	(0.010)	(0.010)	(0.011)	(0.011)
Start_grade_1	-	0.143***	-	0.143***	0.143***	0.142***	0.141***	0.141***	0.141***	0.141***
	-	(0.011)	-	(0.011)	(0.012)	(0.012)	(0.011)	(0.011)	(0.012)	(0.012)
floor_sch	-	0.008*	-	0.008*	0.008	0.008	0.007*	0.007*	0.007	0.007
	-	(0.004)	-	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
Age_teacher	-	0.004*	-	0.004*	0.004*	0.004*	0.004**	0.004**	0.004*	0.004*
	-	(0.002)	-	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
age_teacher2	-	-0.000**	-	-0.000**	-0.000*	-0.000*	-0.000**	-0.000**	-0.000*	-0.000**
	-	(0.000)	-	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Experience_teacher	-	0.001***	-	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
	-	(0.000)	-	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
States x Trend	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Individual fixed effect	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year fixed effect	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	375,398	375,398	375,398	375,398	375,398	375,398	375,398	375,398	375,398	375,398

Standard errors in parentheses
*** significant at 1% level
** significant at 5% level
* significant at 10% level

Table 5.3 –Coefficients of the fixed Effect models with and without IPW and IV for test scores in Mathematics in the 2007/ 2011 panel.

	Mathematics									
Variables	Without IPW and without IV				With IPW and Without IV		Without IPW and With IV		With IPW and With IV	
	Col (1)	Col (2)	Col (3)	Col (4)	Col (5)	Col (6)	Col (7)	Col (8)	Col (9)	Col (10)
Work only at home	-0.055***	-0.049***	-	-	-0.050***	-	-0.136***	-	-0.132***	-
	(0.004)	(0.004)	-	-	(0.005)	-	(0.011)	-	(0.012)	-
Work only in the market	-0.156***	-0.141***	-	-	-0.141***	-	-0.194***	-	-0.192***	-
	(0.008)	(0.008)	-	-	(0.009)	-	(0.014)	-	(0.016)	-
Work in both	-0.195***	-0.176***	-	-	-0.177***	-	-0.312***	-	-0.306***	-
	(0.009)	(0.009)	-	-	(0.010)	-	(0.023)	-	(0.025)	-
Work in the market	-	-	-0.147***	-0.133***	-	-0.133***	-	-0.190***	-	-0.189***
	-	-	(0.006)	(0.006)	-	(0.007)	-	(0.015)	-	(0.016)
Hours worked at home	-	-	-0.033***	-0.030***	-	-0.030***	-	-0.072***	-	-0.069***
	-	-	(0.002)	(0.002)	-	(0.002)	-	(0.005)	-	(0.005)
White	-	0.054***	-	0.054***	0.052***	0.052***	0.052***	0.051***	0.050***	0.049***
	-	(0.005)	-	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
grade_failure_1	-	0.216***	-	0.217***	0.211***	0.211***	0.215***	0.216***	0.210***	0.210***
	-	(0.012)	-	(0.012)	(0.014)	(0.014)	(0.012)	(0.013)	(0.014)	(0.014)
grade_failure_2	-	-0.049***	-	-0.048***	-0.049***	-0.048***	-0.047***	-0.045***	-0.047***	-0.045***
	-	(0.012)	-	(0.012)	(0.014)	(0.014)	(0.012)	(0.012)	(0.014)	(0.014)
car_1	-	-0.025*	-	-0.024*	-0.024	-0.023	-0.028**	-0.026*	-0.027*	-0.025*
	-	(0.013)	-	(0.013)	(0.015)	(0.015)	(0.013)	(0.013)	(0.015)	(0.015)
car_2	-	0.024*	-	0.024*	0.024	0.025*	0.021	0.023*	0.021	0.023
	-	(0.013)	-	(0.013)	(0.015)	(0.015)	(0.013)	(0.013)	(0.015)	(0.015)
car_3	-	0.023	-	0.023*	0.023	0.023	0.022	0.022	0.021	0.022
	-	(0.014)	-	(0.014)	(0.016)	(0.016)	(0.014)	(0.014)	(0.016)	(0.016)
Hh_member	-	-0.045***	-	-0.045***	-0.045***	-0.045***	-0.045***	-0.044***	-0.045***	-0.044***
	-	(0.001)	-	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)

Table 5.3 – continue.

Variables	Mathematics									
	Without IPW and without IV				With IPW and Without IV		Without IPW and With IV		With IPW and With IV	
	Col (1)	Col (2)	Col (3)	Col (4)	Col (5)	Col (6)	Col (7)	Col (8)	Col (9)	Col (10)
Start_maternal	-	0.200***	-	0.199***	0.197***	0.196***	0.197***	0.196***	0.194***	0.194***
	-	(0.010)	-	(0.010)	(0.011)	(0.011)	(0.010)	(0.010)	(0.011)	(0.011)
Start_pre_school	-	0.240***	-	0.240***	0.238***	0.237***	0.237***	0.237***	0.235***	0.234***
	-	(0.010)	-	(0.010)	(0.011)	(0.011)	(0.010)	(0.010)	(0.011)	(0.011)
Start_grade_1	-	0.112***	-	0.112***	0.111***	0.111***	0.110***	0.110***	0.109***	0.110***
	-	(0.011)	-	(0.011)	(0.012)	(0.012)	(0.011)	(0.011)	(0.012)	(0.012)
floor_sch	-	0.017***	-	0.017***	0.017***	0.017***	0.017***	0.017***	0.016***	0.017***
	-	(0.004)	-	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
Age_teacher	-	0.001	-	0.002	0.001	0.001	0.002	0.002	0.001	0.002
	-	(0.002)	-	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
age_teacher2	-	-0.000	-	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	-	(0.000)	-	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Experience_teacher	-	0.002***	-	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***
	-	(0.000)	-	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
States x Trend	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Individual fixed effect	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year fixed effect	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	375,398	375,398	375,398	375,398	375,398	375,398	375,398	375,398	375,398	375,398

Standard errors in parentheses
*** significant at 1% level
** significant at 5% level
* significant at 10% level

Table 5.4 – Coefficients of the fixed effect models with IPW and with IV for test scores in Portuguese and Mathematics in 2007 and 2011 panel by gender

Variables	Boys				Girls			
	Portuguese		Mathematics		Portuguese		Mathematics	
	Col (1)	Col (2)	Col (3)	Col (4)	Col (5)	Col (6)	Col (7)	Col (8)
Work only at home	-0.226***	-	-0.217***	-	-0.192***	-	-0.123***	-
	(0.018)	-	(0.018)	-	(0.016)	-	(0.016)	-
Work only in the market	-0.267***	-	-0.226***	-	-0.464***	-	-0.346***	-
	(0.038)	-	(0.040)	-	(0.047)	-	(0.045)	-
Work in both	-0.362***	-	-0.345***	-	-0.200***	-	-0.110***	-
	(0.032)	-	(0.032)	-	(0.039)	-	(0.039)	-
Work in the market	-	-0.202***	-	-0.171***	-	-0.202***	-	-0.145***
	-	(0.028)	-	(0.029)	-	(0.033)	-	(0.032)
Hours worked at home	-	-0.098***	-	-0.103***	-	-0.078***	-	-0.053***
	-	(0.007)	-	(0.007)	-	(0.006)	-	(0.006)
White	0.051***	0.050***	0.051***	0.050***	0.054***	0.053***	0.051***	0.050***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
grade_failure_1	0.233***	0.235***	0.189***	0.191***	0.276***	0.280***	0.248***	0.250***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
grade_failure_2	-0.025	-0.024	-0.068***	-0.067***	-0.010	-0.005	-0.010	-0.006
	(0.019)	(0.019)	(0.019)	(0.019)	(0.020)	(0.020)	(0.019)	(0.019)
car_1	0.065***	0.070***	-0.006	-0.000	-0.044**	-0.042*	-0.046**	-0.044**
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.021)
car_2	0.080***	0.084***	0.025	0.029	0.003	0.003	0.018	0.018
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
car_3	0.062***	0.063***	0.011	0.012	0.018	0.019	0.030	0.031
	(0.023)	(0.023)	(0.022)	(0.022)	(0.023)	(0.023)	(0.023)	(0.023)
Hh_member	-0.049***	-0.049***	-0.045***	-0.044***	-0.054***	-0.054***	-0.044***	-0.044***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)

Table 5.4 – continue.

Variables	Boys				Girls			
	Portuguese		Mathematics		Portuguese		Mathematics	
	Col (1)	Col (2)	Col (3)	Col (4)	Col (5)	Col (6)	Col (7)	Col (8)
Start_maternal	0.220***	0.220***	0.187***	0.188***	0.231***	0.230***	0.203***	0.202***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.015)	(0.015)	(0.015)	(0.015)
Start_pre_school	0.279***	0.280***	0.246***	0.247***	0.245***	0.245***	0.230***	0.230***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.015)	(0.015)	(0.015)	(0.015)
Start_grade_1	0.159***	0.158***	0.126***	0.125***	0.124***	0.125***	0.100***	0.100***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.016)	(0.016)	(0.015)	(0.015)
floor_sch	0.006	0.006	0.013*	0.013*	0.010	0.009	0.020***	0.020***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)
Age_teacher	0.007**	0.007**	0.006*	0.006*	0.001	0.002	-0.002	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
age_teacher2	-0.000**	-0.000**	-0.000*	-0.000*	-0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Experience_teacher	0.002**	0.002**	0.001**	0.001**	0.001**	0.001**	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
States x Trend	yes	yes	yes	yes	yes	yes	yes	yes
Individual fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
Year fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
Observations	166,490	166,490	166,490	166,490	208,908	208,908	208,908	208,908

Standard errors in parentheses
*** significant at 1% level
** significant at 5% level
* significant at 10% level

Table 5.5 – The impact of the working variables on test scores in Portuguese and Mathematics in percentage for the whole sample and for girls and boys.

	Portuguese 5th grade	Mathematics 5th grade	Portuguese 9th grade	Mathematics 9th grade
		ALL		
Work only at home	- 5.24	- 2.83	- 4.24	- 2.36
Work only in the market	- 6.13	- 4.12	- 4.96	- 3.43
Work in both	- 7.35	- 6.57	- 5.95	- 5.47
Work in the market	- 4.66	- 4.06	- 3.78	- 3.38
Hours worked at home	- 2.44	- 1.48	- 1.98	- 1.23
		GIRLS		
Work only at home	- 4,13	- 2,57	- 3,21	- 2,17
Work only in the market	- 9,98	- 7,22	- 7,77	- 6,10
Work in both	- 4,30	- 2,30	- 3,35	- 1,94
Work in the market	- 4,35	- 3,03	- 3,38	- 2,56
Hours worked at home	- 1,68	- 1,11	- 1,31	- 0,93
		BOYS		
Work only at home	- 5,17	- 4,79	- 4,32	- 3,91
Work only in the market	- 6,11	- 4,99	- 5,11	- 4,07
Work in both	- 8,28	- 7,62	- 6,92	- 6,22
Work in the market	- 4,62	- 3,78	- 3,86	- 3,08
Hours worked at home	- 2,24	- 2,28	- 1,87	- 1,86

6. Conclusions.

Using Prova Brasil census data collected in 2007 and 2011 we created a panel data of 5^{th} and 9^{th} grade students to measure the impact of children labor force participation on their learning outcomes measured by Portuguese and Mathematics test scores.

The fixed effect models controlling for year, states and students effects, were weighted by the inverse probability weights to account for possible attrition bias and an instrumental variable approach proposed by Lewbel (2002) was used to account for the endogeneity of child labor in the test score models. Individual, parents, teachers and principal characteristics as well as school infrastructure were used as control variables.

The estimated parameters were, in general, statistically significant and revealed a negative effect of child labor on school achievement. Students who work inside the home only experienced a negative impact on their achievement test scores, but the negative impact was greater for students who only worked outside the house and for those who worked in both inside and outside their house. Students who work outside the house have a heavy work load, possibly tire themselves physically, encounter greater difficulty coming to class regularly, are more tired during class and have less time and energy to devote to their studies than students who do not work or who only work in the house. Moreover, each additional hour that a student performs household tasks lowers school achievement.

When estimating only for girls, the largest negative impact on test scores was when they were working only in the market and when estimating only for boys the largest impact appeared when they were working in both places. Since household tasks are usually attributed to girls, when girls work only in the market, they probably have a very heavy load of work.

It seems from these analyses that younger 5^{th} grade kids suffer larger harm from work compared to older 9^{th} grade kids and that girls working only in the market present the worst scenario in terms of lowering school achievement.

Our results indicate that domestic work, which is often not counted in social statistics and not considered dangerous, should be included in policies designed to combat child labor. Contrary to ILO vision, our research showed that domestic work has negative effect on children´s school performance. According to ILO "Household chores undertaken by children in their own homes, in reasonable conditions, and under the supervision of those close to them are an integral part of family life and of growing up, therefore something positive." As a policy to reduce the amount of time children spend in household activities we can suggest extending schooling day.

Child labor, whether it occurs inside or outside the home, causes a decrease in school achievement. Policy makers will have to make efforts to prohibit child labor, through social programs, enforcement of the law and labor inspections or by raising awareness about the importance of education and the hazard of early entrance into the job market. A difficult issue for policymakers who would like to eradicate child labor is that families might rely upon the earnings of children and adolescents to meet basic needs.

In this case, conditional cash transfer programs, such as Bolsa Familia and PETI are important sources of income allowing kids to stop working.

Our results also suggest that Brazilian students might benefit from early entrance into school, from better school infrastructure and from more experienced teachers. Delays in school are responsible for a great deal of the weak performance of students. To solve these problems requires educational policies that address the issues of school repetition and drop out, late entry into schools, incentives to improve school quality, and the poor school infrastructure that is found in some regions of the country.

Bibliography

Akabayashi, H.; Psacharopoulos, G. The trade-off between child labour and human capital formation: a Tanzanian case study. **Journal of Development Studies**, London, v. 35, n. 5, p. 120-140, June 1999.

Assaad, Ragui and Nadia Zibani. The Effect of Child Work on School Enrollment in Egypt., Working Papers 0111, **Economic Research Forum**. April 2001.

Basu, K. & Z. Tzannatos. Child Labor and Development : An Introdution.**The World Bank Economic Review**. v.17, n.2. 2003a.

Basu, K. & Z. Tzannatos. The Global Child Labor Problem: What do we know and what can we do ? **The World Bank Economic Review.** v.17, n. 2. 2003b.

Becketti, Gould, Lillard, Welch. "The Panel Study of Income Dynamics after Fourteen Years: An Evaluation", **Journal of Labor Economics**, 6. 1988.

Beegle, K, R Dehejia, and R Gatti, "Why Should We Care About Child Labor?: The Education, Labor Market, and Health Consequences of Child Labor," Journal of Human Resources, 2009, 44 (4).

Bezerra, M., A L Kassouf, and M Arends-Kuenning. The impact of child labor and school quality on academic achievement in Brazil," IZA Discussion Paper 4062, 2009.

Bryson, A., R. Dorsett and S. Purdon. The use of propensity score matching in the evaluation of active labour market policies. Working paper n. 4. Policy Studies Institute and National Center for Social Research.

Canals-Cerda, Jose and Cristobal Ridao-Cano, The dynamics of school and work in rural Bangladesh, 2004.

Cavalieri, C. O impacto do trabalho infantil sobre o desempenho escolar. Tese de doutorado. FGV, SP. 2002

Coleman, J. et al. **Equality of Educational Opportunity**. Washington, D.C., U.S. GPO. 1966.

Deaton, Angus. **The Analysis of Household Survey. A Microeconometric approach to development policy**. John Hopkins University Press. 1997.

Dehejia and Wahba. Propensity Score-matching Methods for Nonexperimental causal Studies. **The Review of Economics and Statistics**, 84(1): 151–161. 2002.

Dumas, Christelle. Does Work Impede Child Learning? The Case of Senegal**. Economic Development and Cultural Change**, 2012, 60 (4), 773 - 793.

Edmonds, Eric V., “Child Labor,” in T. Paul Schultz and John A. Strauss, eds., Handbook of Development Economics, Vol. 4 of **Handbook of Development Economics**, Elsevier, December 2008, chapter 57, pp. 3607–3709.

Emerson, P. and A. Souza. “Birth Order, Child Labor and School Attendance in Brazil.” **World Development**. 2008, v. 36, n. 9.

Emerson, P., V. Ponczek and A. Souza. Child labor and Learning. Rede de Economia Aplicada REAP. Working paper 61. 2013.

Fitzgerald, J., P. Gottschalk and R. Moffitt (1998). An analysis of sample attrition in panel data: the Michigan Panel Study of Income Dynamics. The Journal of Human Resources 33(2), 251- 299.

Glewwe, P. M. Kremer. Schools, teaches and education outcomes in developing countries. In **Handbook of the Economics of Education**, vol 2. Chapter 16. Ed. by E. Hanushek and F. Welch. 2006.

Glewwe, P., E. Hanushek, S. Humpage and R. Ravina. 2011. School Resources and Educational Outcomes in Developing Countries: A Review of the Literature from 1990 to 2010. NBER Working Paper Series, Working Paper 17554.

Gunnarsson, V.; Orazem, P.F.; Sanchez, M.A. Child labor and school achievement in Latin America. Iowa State University: Department of Economics, 2004. 37 p. (Working Papers Series, 03023).

Hanushek, Eric A. and D. D. Kimko, Schooling, Labor Force Quality and the Growth of Nations, **American Economic Review**. 2000, 90 (5), 1184-1208.

Hanushek, E. & S. G. Rivkin. Teacher Quality. In **Handbook of the Economics of Education**, vol 2. Chapter 18. Ed. by E. Hanushek and F. Welch. 2006.

Hanushek, Eric A. and L. Zhang, Quality consistent estimates of international schooling and skill gradients, **Journal of Human Capital**, 2009, 3 (2), 107-143.

Heady, C. The effect of child labor on learning achievement. **World Development**, Amsterdam, v. 31, n. 2, p. 385-398, 2003.

IBGE. Brazilian National Institute of Statistics. Nataional Brazilian Household Survey (PNAD). 2014.

Lewbel, A. Using Heteroskedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models. **Journal of Business and Economic Statistics.** v. 30. 2012.

Moffit, R., J. Fitzgeral, P. Gottschalk. Sample Attrition in Panel Data: The role os selection on observables. **Annales Déconomie et de Statistique**, n. 55-56. 1999.

Mavrokonstantis, P. The Impact of Child Labour on Educational Attainment: Evidence from Vietnam. Young Lives Sudent Paper. 2011.

Orazem & Gunnarsson. Child Labor, school attendance and performance : A Review. ILO working paper. 2003.

Paes de Barros. Brasil Desenvolvimento da Primeira Infância: Foco sobre o Impacto das Pré-Escolas. 2001 Relatório Nº: 22841-BR. World Bank.

Psacharopoulos, G. Child labor versus educational attainment: some evidence from Latin America. **Journal of Population Economics,** Berlim, v. 10, n. 4, p. 377-386, Aug. 1997.

Ravallion, M and Q Wodon, Does Child Labour Displace Schooling? Evidence on Behavioural Responses to an Enrollment Subsidy**. Economic Journal**, March 2000, 110 (462), C158-75.

Rossmann, Marty. Involving Children in Household Tasks: Is It Worth The Effort? University of Minnesota. 2002.

Stinebrickner, R.; Stinebrickner, T.R. Working during school and academic performance. **Journal of Labor Economics**, Chicago, v. 21, n. 2, p. 449-472, 2003.

Wooldridge, J. **Econometric Analysis of Cross Section and Panel Data**. The MIT Press. 2002.

Appendix

Table A1 – Coefficients of the fixed Effect models including 33 exogenous variables, without IPW and without IV for test scores in Portuguese and Mathematics in the 2007/ 2011 panel.

Variable	Without IPW and without IV			
	Score in Portuguese		Score in Mathematics	
Work only at home	-0.046***	-	-0.046***	-
	(0.006)		(0.006)	
Work only in the market	-0.199***	-	-0.123***	-
	(0.011)		(0.011)	
Work in both	-0.182***	-	-0.149***	-
	(0.012)		(0.012)	
Work in the market	-	-0.171***	-	-0.113***
		(0.008)		(0.008)
Hours worked at home	-	-0.024***	-	-0.027***
		(0.002)		(0.002)
white	0.055***	0.054***	0.051***	0.051***
	(0.006)	(0.006)	(0.006)	(0.006)
grade_failure_1	0.248***	0.249***	0.191***	0.192***
	(0.018)	(0.018)	(0.018)	(0.018)
grade_failure_2	-0.039**	-0.038**	-0.074***	-0.073***
	(0.017)	(0.017)	(0.017)	(0.017)
car_1	0.057***	0.058***	0.008	0.009
	(0.018)	(0.018)	(0.018)	(0.018)
car_2	0.060***	0.061***	0.032*	0.032*
	(0.017)	(0.017)	(0.017)	(0.017)
car_3	0.054***	0.055***	0.021	0.022
	(0.018)	(0.018)	(0.018)	(0.018)
Hh_ member	-0.051***	-0.051***	-0.045***	-0.045***
	(0.002)	(0.002)	(0.002)	(0.002)
Maternal	0.193***	0.193***	0.161***	0.161***
	(0.014)	(0.014)	(0.014)	(0.014)
pre_school	0.244***	0.244***	0.214***	0.214***
	(0.014)	(0.014)	(0.014)	(0.014)
grade_1	0.131***	0.131***	0.096***	0.096***
	(0.015)	(0.015)	(0.015)	(0.015)
floor_sch	0.006	0.006	0.010	0.010
	(0.006)	(0.006)	(0.006)	(0.006)
Age_teacher	0.002	0.002	0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.003)
age_teacher2	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Experience_teacher	0.001**	0.001**	0.001***	0.001***
	(0.001)	(0.001)	(0.001)	(0.001)
drop_out_1	0.130***	0.129***	0.138***	0.136***
	(0.037)	(0.037)	(0.037)	(0.037)
drop_out_2	0.077*	0.077*	0.089**	0.089**
	(0.040)	(0.040)	(0.040)	(0.040)

Table A1 – continue.

Variable	Without IPW and without IV			
	Score in Portuguese		Score in Mathematics	
classroom_sch	0.011*	0.011*	0.016***	0.017***
	(0.006)	(0.006)	(0.006)	(0.006)
colorTV_1	-0.167***	-0.166***	-0.153***	-0.151***
	(0.020)	(0.020)	(0.020)	(0.020)
colorTV_2	-0.054***	-0.053***	-0.062***	-0.061***
	(0.008)	(0.008)	(0.008)	(0.008)
colorTV_3	-0.013*	-0.013*	-0.028***	-0.028***
	(0.007)	(0.007)	(0.007)	(0.007)
computer_sch	0.002	0.002	0.003	0.003
	(0.007)	(0.007)	(0.007)	(0.007)
Video	0.055***	0.055***	0.027***	0.027***
	(0.009)	(0.009)	(0.009)	(0.009)
Computer	0.091***	0.090***	0.059***	0.059***
	(0.007)	(0.007)	(0.007)	(0.007)
Bathroom	0.018***	0.018***	0.029***	0.029***
	(0.004)	(0.004)	(0.004)	(0.004)
moth_read	0.107***	0.107***	0.095***	0.095***
	(0.016)	(0.016)	(0.016)	(0.016)
fath_read	0.063***	0.063***	0.062***	0.062***
	(0.012)	(0.012)	(0.012)	(0.012)
go_sch_meet	0.043***	0.043***	0.040***	0.040***
	(0.006)	(0.006)	(0.006)	(0.006)
motivate_study	0.135***	0.135***	0.054**	0.055**
	(0.022)	(0.022)	(0.022)	(0.022)
motivate_school	0.268***	0.267***	0.204***	0.203***
	(0.019)	(0.019)	(0.019)	(0.019)
dohomew_port	0.138***	0.139***	0.026***	0.026***
	(0.007)	(0.007)	(0.007)	(0.007)
dohomew_math	0.054***	0.054***	0.225***	0.225***
	(0.007)	(0.007)	(0.007)	(0.007)
Sex_teacher	0.015*	0.014*	-0.007	-0.007
	(0.008)	(0.008)	(0.008)	(0.008)
Sch_teacher_college	0.013	0.012	0.025**	0.024**
	(0.010)	(0.010)	(0.010)	(0.010)
States x Trend	yes	yes	yes	yes
Individual fixed effect	yes	yes	yes	yes
Year fixed effect	yes	yes	yes	yes
Observations	276,425	276,425	276,425	276,425

Standard errors in parentheses
*** significant at 1% level
** significant at 5% level
* significant at 10% level

Table A2 - Description of the variables, mean and standard deviation in 2007 and 2011.

Variables	Description of the Variables	2007			2011		
		obs	mean	s.d.	obs	Mean	s.d.
Scores_Portuguese	Portuguese test Score	187699	185.03	41.07	187699	250.49	45.05
Scores_Mathematics	Mathematics test Score	187699	201.48	43.27	187699	256.51	45.87
Work_in_hh	=1 if child works in hh	187699	0.36	0.48	187699	0.41	0.49
Hours_work_hh	Hours in hh activities per day	187699	1.43	1.12	187699	1.47	1.06
Work_market	=1 if child works outside hh	187699	0.09	0.29	187699	0.15	0.36
don't work	don't work	187699	0.59	0.49	187699	0.50	0.50
Work onky hh	Work onky in the hh	187699	0.32	0.47	187699	0.35	0.48
Work onky market	Work onky in the market	187699	0.05	0.22	187699	0.09	0.29
Work in both	Work in both	187699	0.04	0.20	187699	0.06	0.24
White	White+Asians	187699	0.40	0.49	187699	0.40	0.49
Repetition	N. years repeat school year	187699	0.23	0.51	187699	0.23	0.51
Car	Number of cars in hh	187699	0.57	0.72	187699	0.66	0.75
Hh_member	Number of people in hh	187699	4.91	1.47	187699	5.91	1.65
Start_maternal	=1 if child start 2 to 4 years	187699	0.39	0.49	187699	0.33	0.47
Start_preschool	=1 if child start 4 to 6 years	187699	0.37	0.48	187699	0.46	0.50
Start_grade1	=1 if started sch. in grade 1	187699	0.18	0.39	187699	0.19	0.39
Drop out	N. years dropping out	184281	0.05	0.24	187263	0.03	0.19
classroom_sch	Classroom in good condition	182481	0.62	0.48	185420	0.62	0.49
Color_TV	Number of color TVs in hh	183039	1.68	0.81	185810	1.89	0.79
Computer_sch	Computer in good condition	185020	0.95	0.23	185097	0.71	0.45
Video	Number of vcr in hh	185007	0.82	0.39	186269	0.92	0.27
Computer	Number of computer in hh	182476	0.35	0.48	187349	0.68	0.47
Bathroom	Number of bathrooms in hh	185590	1.36	0.67	187352	1.36	0.68
Mother_read	=1 if mother reads and writes	181357	0.96	0.20	181793	0.96	0.20
Father_read	=1 if father reads and writes	180494	0.93	0.25	158129	0.94	0.24
Go_sch_meetings	=1 if parents often go to school meetings	181317	0.66	0.47	186727	0.61	0.49
Motivate_study	=1 if parents encourage kids to study	179558	0.98	0.15	184216	0.99	0.10
Motivate_school	=1 if parents encourage kids to go to school	183491	0.97	0.18	186265	0.99	0.11
Do_port_homework	=1 if a child often does Port. homework	180991	0.79	0.41	186572	0.60	0.49
Do_math_homework	=1 if a child often does Math. homework	180526	0.82	0.38	186606	0.59	0.49
Sex_teacher	=1 if teacher is a man	187004	0.16	0.37	187551	0.10	0.29
Education_teacher	=1 if teacher has college	185260	0.90	0.30	187381	0.95	0.22
Age_teacher	Age of the teacher	187699	40.84	8.71	187699	42.43	8.55
Experience_teacher	Number of years teaching	187699	14.35	6.71	187699	14.84	6.95
floor_sch	=1 if floor in school	187699	0.59	0.49	187699	0.64	0.48