Manuscript Details

Manuscript number	ECOEDU_2020_257
Title	Early Parental Reading and Reading for Enjoyment: how it matters for boys and girls?
Article type	Research Paper

Abstract

This study examines the combined effect of two literacy practices on reading achievement scores at ages 8-9: the frequency of parental book reading to 4/5 year-olds, here treated as a multivalued treatment, and subsequent reading for enjoyment. Using the Longitudinal Study of Australian Children, we test whether reading for enjoyment impacts reading achievement over and above parental book reading and whether this impact is gender-specific. To account for endogeneity existing IV methods are extended to a nonlinear regression setting and an Average Treatment Effect Specification Test (ATESET) is proposed. Results from the reading achievement equation indicate that while for boys reading for enjoyment is a transmission mechanism of early parental reading, for girls the two inputs act as unrelated sources

Keywords	Human capital; reading skills; reading to children; reading for enjoyment; gender differences
Manuscript category	Full Length Research Article

Corresponding Author	Montezuma Dumangane
Corresponding Author's Institution	European Commission - Joint Research Center
Order of Authors	Montezuma Dumangane, Luisa Araujo, Patricia Dinis Mota da Costa, Nuno Crato, Nils Picker

Submission Files Included in this PDF

File Name [File Type]

CoverLetter_EPR_RE.pdf [Cover Letter]

EPR_RE_Highlights.pdf [Highlights]

EPR_RE_TitlePage.pdf [Title Page (with Author Details)]

EPR_RE_Manuscript.pdf [Manuscript File]

Submission Files Not Included in this PDF

File Name [File Type]

Manuscript_LatexSourceFiles.zip [LaTeX Source File]

To view all the submission files, including those not included in the PDF, click on the manuscript title on your EVISE Homepage, then click 'Download zip file'.

Research Data Related to this Submission

There are no linked research data sets for this submission. The following reason is given:
The authors do not have permission to share data

Dr. Celeste K. Carruthers
Editor-in-chief
Economics of Education Review

April 29, 2020

Dear Dr. Carruthers,

I would like to submit the manuscript entitled "Early Parental Reading and Reading for Enjoyment: how it matters for boys and girls?" by Luisa Araújo, Patricia Costa, Nuno Crato, Montezuma Dumangane and Nils Picker, to be considered for publication as an original article in the *Economics of Education Review*.

The paper addresses how gender differences explain the relation between early parental reading and reading for enjoyment and its final impact in reading achievement, and provides a discussion on the relevance of other forms of human capital investment. I believe the findings will be of interest to the readers of your journal.

We declare that this manuscript is original, has not been published before and is not currently being considered for publication elsewhere.

We know of no conflicts of interest associated with this publication, and there has been no significant financial support for this work that could have influenced its outcome. As Corresponding Author, I confirm that the manuscript has been read and approved for submission by all the named authors.

We hope you find our manuscript suitable for publication and look forward to hearing from you in due course.

Sincerely,

(Montezuma Dumangane)

Montezuma Dumangane
European Commission - Joint Research Center
Unit I.1 - Monitoring, Indicators & Impact Evaluation
Competence Centre on Microeconomic Evaluation
Via Fermi, 2749, TP 361
21027 Ispra (VA), Italy
Tel./Fax: +39 0332 7837787
Email: montezuma.DUMANGANE@ec.europa.eu

Early Parental Reading and Reading for Enjoyment: how it matters for boys and girls?

Highlights

- Parental reading and enjoyment reading impact boys' and girls' reading skills uniquely.
- Boys are dependent on early parental reading to later engage in reading for enjoyment.
- Girls engage in reading for enjoyment independently of this early exposure.
- Library visits from childhood have gender specific effects on reading for enjoyment.
- Unobserved parental human capital investment determine the child performance.

Early Parental Reading and Reading for Enjoyment: how it matters for boys and girls?

Luisa Araújo[a], Patricia Costa[b], Nuno Crato[b,d,c], Montezuma Dumangane[b,*], Nils Picker[b]

[a]*Instituto Superior de Educação e Ciências, Alameda das Linhas de Torres 179, 1750-142 Lisboa, Portugal*

[b]*European Commission, Joint Research Centre (JRC), Ispra, Italy*

[c]*REM - Research in Economics and Mathematics, CEMAPRE, Rua do Quelhas 6, 1200 - 781 Lisbon, Portugal*

[d]*ISEG, Technical University of Lisbon, Rua do Quelhas 6, 1200 - 781 Lisbon, Portugal*

Funding: Nuno Crato was partially supported by the Project CEMAPRE/REM – UIDB/05069/2020 – financed by FCT/MCTES through national funds.

Declarations of interests: None.

Disclaimer: The findings and views reported in this paper, are those of the authors and should not be attributed to Department of Social Services, the Australian Institute of Family Studies or the Australian Bureau of Statistics or an official position of the European Commission.

*Corresponding author at: European Commission - Joint Research Centre (JRC), Unit I.1 - Monitoring, Indicators & Impact Evaluation, Competence Centre on Microeconomic Evaluation (CC-ME), Via E. Fermi 2749, TP 361, 21027 Ispra (VA), Italy, Tel: +39 0332 783787

Email addresses: luisa.araujo@iseclisboa.pt (Luisa Araújo), Patricia.COSTA@ec.europa.eu (Patricia Costa), ncrato@iseg.ulisboa.pt (Nuno Crato), Montezuma.DUMANGANE@ec.europa.eu (Montezuma Dumangane), nils.picker@gmail.com (Nils Picker)

Early Parental Reading and Reading for Enjoyment: how it matters for boys and girls?

Abstract

This study examines the combined effect of two literacy practices on reading achievement scores at ages 8-9: the frequency of parental book reading to 4/5 year-olds, here treated as a multivalued treatment, and subsequent reading for enjoyment. Using the Longitudinal Study of Australian Children, we test whether reading for enjoyment impacts reading achievement over and above parental book reading and whether this impact is gender-specific. To account for endogeneity existing IV methods are extended to a nonlinear regression setting and an Average Treatment Effect Specification Test (ATESET) is proposed. Results from the reading achievement equation indicate that while for boys reading for enjoyment is a transmission mechanism of early parental reading, for girls the two inputs act as unrelated sources.

Keywords: Human capital, reading skills, reading to children, reading for enjoyment, gender differences

JEL: C21, C26, I21, J16, J24

1. Introduction

Numerous studies have shown that parental reading to young children is positively associated with an increase in reading performance in the later primary school years. This evidence has been gathered in Canada (Sénéchal and Young, 2008), in Europe (Araújo and Costa, 2015) and in the United States (Mol and Bus, 2011; Whitehurst et al., 2001). More recently, Kalb and Ours (2014) used data from a longitudinal Australian survey to investigate the effect of parental reading to children at age 4/5 on their reading skills at ages 8-9 and 9-10. They conclude that there is a positive effect, after controlling for a range of endogenous factors, and that parental book reading gives children a head-start in life. This type of parental investment in the early life of children supports their later cognitive development and should thus be encouraged (Kalb and Ours, 2014). Other studies corroborate the notion that to invest in cognitive skills, namely early reading and numeracy, both at home and in school, affects the economic fortune of nations and brings economic returns to individuals (Hanushek and Woessmann, 2015; Heckman and Jacobs, 2011). With respect to literacy attainment, research indicates that children who fail to develop basic reading skills to comprehend what they read by the fourth grade are likely to face reduced educational opportunities (Chall and Jacobs, 2003; Adams, 2009). Conversely, primary school children who like to read and read for enjoyment develop good reading skills (Stanovich, 2000) and those in secondary school that read for enjoyment outside of school have higher reading scores (OECD, 2010).

Reading literacy literature suggests that there is a relationship between the two, specifically, that parental book reading at a young age might increase children's motivation to read, which in turn will result in more frequent reading for enjoyment (see Neuman and Dickinson, 2001, p. 901). Indeed, Sénéchal et al. (2006) have shown that parents reports of shared reading in kindergarten predict children's reports of reading for enjoyment in Grade 4, after controlling for parent education, child vocabulary and reading skill. Data from the Longitudinal Study of Australian Students (LSAC) also shows that children who have been read to at ages 4/5 are more likely to report that they enjoy reading at ages 10-11 (AIFS, 2015). Thus, existing studies suggest that reading to young children positively affects their reading comprehension in later grades, that those children who have been read to are more likely to enjoy reading later, and that those that enjoy reading and engage in frequent reading outside of school have higher reading achievement. However, to our knowledge, there are no studies that look at whether there is a cumulative effect of being read to and enjoying reading in reading achievement. More specifically, no longitudinal research exists (Sénéchal and Young, 2008) to determine if there is an additional advantage of reading for enjoyment in the reading achievement of children who have been exposed to book reading early in life.

Our paper investigates whether there is an additional reading gain related with reading for enjoyment in the reading achievement of children at ages 8-9 who have been exposed to parental book reading at ages 4-5. Thus, we extend the results of Kalb and Ours (2014) by introducing an additional factor - reading for enjoyment - that is known to be associated with higher reading achievement in primary and secondary school. We use LSAC, a longitudinal data set, to investigate if there is such a direct causal effect, whether reading for enjoyment is a determinant of reading achievement at ages 8-9 or if a transmission mechanism of early parental reading at ages 4/5. Furthermore, we test if the determinants of reading achievement are gender specific. Our analyses allow for the estimation of those causal effects by using an IV identification strategy that controls for several possible sources of endogeneity. If recreational reading is found to be an independent determinant of reading skills it becomes relevant to investigate other sources of early parental human capital investment that influence a child's reading habits.

In short, our contribution to the literature offers a comprehensive view of the reading development of boys and girls, considering both family literacy practices and individual reading habits and their role in determining reading achievement.

The paper is structured as follows. Section 2 presents relevant literature on parental reading and reading for enjoyment influence on children's reading achievement, as well as educational achievement gender differences. Section 3 shortly describes the data used in the paper and present the relevant descriptive statistics for our study. Section 4, outlines the modelling strategy, which includes the rationale for the empirical model, and in particular the discussion on the endogeneity sources. Section 5 presents the estimation methods. These are further discussed in Appendix B, where an extension of the IV strategy in Wooldridge (2010) is provided, and an Average Treatment Effect Specification Test (ATESET) is proposed. Section 6 presents the results of both the reading for enjoyment and the reading achievement equations. Section 7 provides several robustness checks, and in particular, it implements the Inoue and Rossi (2011) strong identification test. Section 8 concludes.

2. Review of the Literature

Sénéchal (2012) has shown that family literacy practices related to teaching preschool children about literacy skills, such as naming alphabet letters, and reading to them increase reading comprehension in fourth grade. Not all parents engage in teaching literacy skills while they read to their children, but even the informal activity of book reading boost children's test. Children who enter primary school with a good knowledge of ABC letters and whose parents read to them frequently tend to score the highest. But high frequency of parental book reading alone is itself associated with higher reading performance in European countries (Araújo and Costa, 2015). The study by Kalb and Ours (2014) with LSAC data suggests that this association is causal in nature and that there are differences between boys and girls.

Literacy scholars contend that parental book reading boosts children's test scores in primary school mainly because during this joint activity they acquire new vocabulary (Araújo and Costa, 2015). The same vocabulary building occurs later when children are able to read by themselves and choose to read outside of school (Stanovich, 2000). Several studies in education have found a positive relation between reading for enjoyment outside of school and reading outcomes (Cunningham and Stanovich, 1997; OECD, 2010) and research has clearly established a reciprocal causality effect known as the Mathew effect (Stanovich, 2000), or the rich get richer and the poor get poorer. Young students who learn to read with ease and have good reading skills read more for enjoyment throughout the primary school years and greater enjoyment leads to better reading skills (Cox and Guthrie, 2001). Thus, whereas those that have better reading skills are more likely to engage in reading for enjoyment, (Canoy et al., 2006) poor readers avoid reading and the less they read the further behind they get.

Frequent reading supports the acquisition of new vocabulary from context (Nagy et al., 1985; Verhoeven et al., 2011) and this, in turn, supports reading comprehension (Stanovich et al., 1998; Duff et al., 2015; Sénéchal et al., 2006), for example, showed that frequency of reading for enjoyment was associated with reading comprehension in grade 4, after controlling for parental education. Children's reported frequency of reading for enjoyment has also been found to account for unique variance in writing quality (Sénéchal et al., 2018). Thus, fourth grade students who read more for enjoyment develop reading comprehension and are also more likely to write more coherent narratives.

The literature also suggests that girls are more likely than boys to engage in reading activities at early ages. Girls choose to read more for enjoyment than boys and they consistently score above boys in reading tests. In the US, The National Assessment of Educational Progress (NAEP) shows that girls have always scored higher than boys at the 4th, 8th and 12th grades (Loveless, 2015). International large-scale surveys such as PIRLS, the Program in International Reading Literacy Study, provide a similar picture. In 2016, girls outperformed boys in 48 out of 50 participant countries at the 4th grade level (Mullis et al., 2017). And on the 2015 Program for International Student Assessment (PISA), 15-year-old females outperformed males in reading in all sixty five participating countries (OECD, 2016).

Since the gender gap in reading is nearly universally observed, some contend that it originates in biological or developmental differences between boys and girls (Loveless, 2015). Indeed, as early as 4 and 5 years of age, vocabulary measures and early literacy knowledge have been found to already favour girls in Canada and the US (Baker and Milligan, 2016), but this may be related to another hypothesis which explains the reading gap with gender-specific practices. In Canada, the US and the UK time inputs from parents, including reading, storytelling, and teaching letters and numbers are larger for girls than for boys and the additional time parents spend in these activities with 2-5 year-old girls positively impacts girls' language and mathematics scores (Baker and Milligan, 2016). In addition to gendered practices, expectations associated with gender roles may also influence the construction of a reader from an early age. For instance, parental differences in reading habits may lead girls to value literacy more and to develop a positive self-concept of themselves as readers (Sullivan and Brown, 2015).

The National Literacy Trust defines reading for enjoyment as reading that we do of our own free will, anticipating the satisfaction that we will get from the act of reading. It also refers to reading that having begun at someone else's request we continue because we are interested in it (Clark and Rumbold, 2006). The terms reading for pleasure and reading for enjoyment are used interchangeably in the literature and independent reading, voluntary reading, and recreational reading are often adopted designations in the United States (Clark and Rumbold, 2006). Interestingly, in a UK Reading Agency survey of 1,110 children aged 4 -11 over a fifth (22%) selected visiting the library as the action most likely to make them want to read, compared to less than 2% saying that reading a book on an iPad or Kindle was the thing most likely to make them want to read. Only one in ten said that buying a book would be most likely to get them reading. Visiting the library was the second most popular action with discovering an author that they like coming out top. Over a quarter of children selected this option.[1]

In large scale surveys, such as PIRLS and PISA, the frequency of reading for enjoyment outside of school appears as a categorical question in the student questionnaire. Similarly, with respect to family literacy practices, specifically frequency of book reading to young children, the question appears in the parental questionnaire.

3. Data and descriptive statistics

We used data from Growing Up in Australia, the Longitudinal Study of Australian Children (LSAC). LSAC started in 2004 and collects information on family and social issues, children's development and well being. The unit of observation in LSAC is the "study child" which is the child who is being studied. The information is collected from parents, caregivers, teachers and the study children themselves (see Appendix A). The LSAC can be used to assess how reading for enjoyment is related to forms of early parental human capital investment (i.e. frequency of parental reading and frequency of visits to libraries) and whether they are related to a child's reading skills. In addition, it allows us to explore how these relationships can differ by gender.

If reading for enjoyment is a potential determinant of reading achievement, then it is of relevance to consider which other forms of early parental human capital investment, beyond parental reading, can

[1] https://readingagency.org.uk/news/media/children-say-going-to-the-library-encourages-them-to-read-more-than-buying-them-new-books-or-ipads.html.

impact achievement. The frequency and continuity of visits to libraries with children, from a child's early ages, is a parental activity likely to induce an independent interest in reading activities later on. Although these two forms of human capital investment may not be independent, their potential impact on the child has a distinct nature. On the one hand, parental reading is a joint activity that requires the physical presence of the parent, on the other hand, the library set up often encourages the child to interact with the books and related environment independently.

Figure 1 shows histograms exploring the relationship between early parental reading, frequency of visits to libraries, and reading for enjoyment in both the boys and girls sub-samples. The frequency of early parental reading is a categorical variable representing the "number of weekdays a child was read to at home at the age of 4/5". It that assumes three categories: "less than three days", "3-5 days" and "6-7 days".[2] The regularity of visits to libraries since the age of 2/3 is measured by the variable "age at last regular library visits". The following categories are considered: "2/3 or never", "4/5", "6/7 with interruptions" and "6/7 without interruptions". For example, children regularly visiting at 4/5, but not before, fall in the same category as children visiting at both 2/3 and 4/5. The last two categories assume a qualitative difference between a child that consistently visited a library at all ages, from those that regularly visited at 6/7 but missed regularity at some earlier ages.

The first row of Figure 1 shows that parents report being equally likely to have read to boys as to girls, as well as to have taken them to library visits. However, reading to the child at home at the age of 4/5 is a more popular form of human capital investment than regularly visiting a library. Most children (around 60%), were read to 6-7 days a week (the maximum frequency) at the age of 4/5, while only 16% regularly visited a library at all ages.

Since the two forms of early human capital investment may be correlated, the second row of Figure 1 explores how the frequency of parental reading at the age of 4/5 distribution varies with the reported frequency of visits to libraries. As expected, the histograms show distributions with increasing frequency of parental reading as the regularity of visits to library increases. In particular, the last two figures reveal a significant difference in the parental reading frequencies of children who last regularly visited libraries at the age of 6/7 with and without interruptions. However, combining the frequencies in the figures reveal that children who were read to 6-7 days are equally likely to have last visited a library periodically at the age of "2/3 or never", or at "6/7 without interruptions". This indicates that the two forms of human capital investment - parental reading and visits to the library - are not unequivocally correlated. In other words, the two variables are not measuring the same literacy

[2]The early parental reading variable is recorded from the parents' questionnaire. It answers the question, "In the past week, on how many days have you or an adult in your family, read to child from a book?" and distinguishes from four categories: None; 1 or 2 days; 3-5 days and; 6-7 days. We considered only three categories aggregating the two lowest groups (reading to the child of 0 times and 1-2 times per week), as most parents reported to have read at least once per week.

practices.

The last row of Figure 1 shows how reading for enjoyment at the age of 7/8, represented as a dummy variable, is associated with both forms of parental human capital investment.[3] The histograms shows that girls read more autonomously than boys. However, in the boys' sub-sample, while the relation with parental reading frequency is smoothly increasing, the positive relationship with visits to libraries only arises when moving to the group with the highest visits regularity. On the contrary, in the girls' sub-sample, parental reading has a weaker correlation with reading for enjoyment, while increasing the age of last regular visit to a library to 4/5, significantly and permanently increases the frequency of reading for enjoyment.

Overall this set of descriptive statistics suggests that: the way parents invest in the two forms of early parental human capital investment here considered does not depend on the child's gender; there is a positive association between parental reading frequency at the age of 4/5 and the frequency and regularity with which the child visited a library since the age of 2/3 and; early parental reading is more strongly associated to reading for enjoyment for boys than for girls, while the positive association of the latter with library visits is present at an earlier age for girls.

The second set of statistics examines how early parental reading and reading for enjoyment is associated with the performance of boys and girls in the NAPLAN test at the age of 8-9. The NAPLAN score is a standardised measure that allows the comparison of children's performance over time.[4] The top left panel of Figure 2 shows a positive relationship between the NAPLAN score distribution and the frequency of early parental reading. This association is stronger in the boys distribution, as shown by the shift in the score distribution of boys who were read to 6-7 days/week. For this gender, and in particular at lower reading frequencies, the score distribution exhibits more dispersion than the girls' suggesting a less homogeneous relationship between reading achievement and parental reading. The top right panel shows that children who read for enjoyment have higher scores in the NAPLAN test, but there is more dispersion.

These two variables ultimately reflect past parental behaviour and the current level of reading as an activity performed by the child. However, being a form of human capital investment, reading to the child is likely to stimulate and develop the interest in reading activities later on, as suggested by the first set of descriptive statistics discussed above. To further investigate how they might interact and determine the children's reading achievement, the bottom row of Figure 2 examines the NAPLAN

[3]The reading for enjoyment indicator was constructed from a categorical variable recording whether children "enjoy reading at home that is not part of school work?". Two categories are considered: "Yes", and the aggregation of "Sometimes" and "No".

[4]NAPLAN (National Assessment Program) reading tests assesses literacy proficiency across the English learning area in line with the English Australian Curriculum. In particular, in terms of knowledge and interpretation of written language. NAPLAN is a standardised direct assessment of child's academic skills

score distribution conditional on the parental reading frequency in the sub-samples of children who read and do not read for enjoyment. If the two variables were unrelated, the location of the score distribution would shift according to parental reading frequencies in the same fashion. The comparison between the two panels reveals a clear positive correlation between parental reading frequency and the NAPLAN scores. The difference in the performance across the two sub-groups is particularly evident in boys, which suggests that for this group, reading for enjoyment and early parental reading are not independent determinants of reading achievement.

In the boys' sub-sample, the frequency of early parental reading is always negatively associated with the dispersion of the scores' distribution. In particular, this distribution seems to be bimodal at the lowest frequency of parental reading, suggesting that there is a sub-population of high achieving boys that were not exposed to this form of parental input. Despite the limitations of the univariate analysis, the set of descriptive statistics above shed some light on the gender differences found regarding the relation between the NAPLAN scores and the two variables of interest. Both variables are positively associated with the reading scores, although the relationship is more heterogeneous in the boys sub-sample. This highlights the importance of conditioning the analysis on both observable variables that explain this heterogeneity and, on unobservables determining reading achievement, that may be correlated with both early parental reading and reading for enjoyment.

Furthermore, and most importantly, the statistics suggest a higher degree of dependence between reading for enjoyment and early parental reading in boys. These are the two main issues to be addressed in the model specification that follows – the degree of dependence between the two variables and the endogeneity problem.

4. Methods

This paper addresses the following research questions: What is the impact of early parental reading and reading for enjoyment in the children reading achievement? Does parental reading impacts reading achievement directly or indirectly through increased reading for enjoyment? Are there other forms of human capital investment that affect reading achievement through increased reading for enjoyment? Are the direct and indirect nature of these effects gender specific?

4.1. Empirical Strategy

The estimation strategy to answer the research questions is based on the idea that if a child develops reading for enjoyment habits mainly due to early parental reading then, conditionally on this variable, reading for enjoyment should be redundant to explain reading achievement. On the other hand, if the relevant determinants of reading for enjoyment exclude parental reading, but consider other forms of parental human capital investment, then both variables might explain the child's reading achievement with a reduced degree of substitution.

Several sources of endogeneity may arise in the specification of the reading outcomes equation. Both early parental reading and reading for enjoyment may be correlated with individual ability which is known to be a non-negligible determinant of performance. While parental reading alone has already been discussed in the literature and IV strategies have been suggested (see Kalb and Ours, 2014), the potential endogeneity of reading for enjoyment is taken into account in this paper.

Parental reading at the age 4-5 is a treatment to which the child was exposed according to different frequency levels. As such, first we need to identify the causal impact and the nature of its transmission mechanism when endogeneity causes the ignorability of treatment assumption to fail. To address this, the IV methods in Wooldridge (2010) chapter 21, are extended in Appendix B to a nonlinear regression setup, with a categorical treatment and endogenous controls. The procedure explores the nonlinearity of the first stage regressions to generate internal instrumental variables that provided over-identifying restrictions. The identification power of these instruments is tested with the strong identification test of Inoue and Rossi (2011) which is robust against the alternative of weak instruments.

Since the outcome variable - the NAPLAN score test - is bounded and continuous, the analysis is conducted under the estimation methods proposed in Ramalho and Ramalho (2017) for cross-sectional fractional regression models. The nonlinearity of the outcome equation implies that although the treatment effect parameter is fixed, the treatment effect itself is heterogeneous. The functional form specification is investigated through the implementation of two forms of the RESET test: the standard Ramsey (1969) form and a restricted version. The later, denoted by ATESET (Average Treatment Effect Specification Test), addresses a necessary condition required for the identification of the Average Treatment Effect (ATE) in the Wooldridge (2010) chapter 21 framework.

4.2. The empirical model

The outcome of interest is the students NAPLAN (National Assessment Program) reading achievement score at the age of 8-9. We start by assuming that the child's i reading achievement (Y_i) is ultimately the result of reading skills (S_i), i.e. the child's overall reading ability, the inherited ability (A_i) and, by factors measuring the effort (E_i) put on academic activities:

$$Y_i = G(A_i, S_i, E_i) \tag{1}$$

The variables in the right hand side of (1) are unobserved but can be expressed as functions of observable variables and error terms. In particular the child's reading skills equation is given by

$$S_i = f(R'_{4/5,i}\boldsymbol{\theta_1} + \theta_2 J_i + v_i) \tag{2}$$

where J_i is the dummy variable representing reading for enjoyment, $R_{4/5,i}$ is the categorical variable measuring the frequency of parental reading at the ages of 4 and 5 (as in Kalb and Ours, 2014) and,

v_i represents the effect of contextual factors affecting the child's reading skills - parents educational level, household income etc..

We assume that reading for enjoyment is an important driver of the children's reading skills, since time spent reading for enjoyment is related to higher achievement as discussed in the literature section. Furthermore, it is important to note that J_i in equation (2) is likely to be an outcome of early parental human capital investment. Two forms of early human capital parental investment are considered relevant in this context: (i) the frequency of early parental reading, requiring the parents' direct involvement and; (ii) the frequency of exposure to activities where reading is encouraged and perceived as a pleasant activity. Libraries are able to provide such an experience and therefore stimulate the children's reading habits.

This motivates the third and final equation. Let L_i measure how regularly the child attended a library between the ages of "2 to 3" and "6 to 7", as defined in Section 3. Then the equation for reading for enjoyment can be written as

$$J_i = h(R'_{4/5,i}\alpha_{\mathbf{1}} + L'_i\alpha_{\mathbf{2}} + u_i) \tag{3}$$

where u_i represents the reading environment surrounding the child. Equation 3 establishes that early parental reading can also have an indirect effect on the child reading skills through the development of reading for enjoyment.

In a reduced form approach derived by replacing both equations (3) and (2) in (1), the child's reading performance would be determined by the two forms of early parental human capital investment. However, this would exclude reading for enjoyment from the specification, which is observable and whose impact is of interest.

Thus, a reduced form based on replacing equation (2) in equation (1) will be considered but where the estimated coefficients need to be interpreted in light of the restrictions implicit in the estimates of equation (3). Assuming a linear approximation of the argument arising from replacing (2) in (1), the equation of interest becomes

$$Y_i = G(A_i + R'_{4/5,i}\beta_{\mathbf{1}} + \beta_2 J_i + X'_i\beta_{\mathbf{3}} + \epsilon_i) \tag{4}$$

Early parental reading can be seen as a treatment whose causal impact on the reading achievement is of interest, while controlling for the effect of reading for enjoyment, which may depend on the treatment, and of other variables. These measure the effect of the observable components of the academic effort (E_i in equation 1), the contextual variables (v_i in equation 2) and, the reading environment, (u_i in equation 3).

This three equation model embodies two alternative transmission mechanisms of early parental reading by imposing restrictions on $\theta_{\mathbf{1}}$ and $\alpha_{\mathbf{1}}$ in equations (2) and (3).

4.3. Autonomous versus non-autonomous children

Consider first the case where reading for enjoyment depends on both forms of early human capital investment and in particular on the parents' direct involvement ($\alpha_1 \neq 0$ in equation 3). If in addition $\theta_1 = 0$ in equation (2), then early parental reading only indirectly determines the child's reading skills (S_i). As a consequence, conditioning the reading test score equation on reading for enjoyment only, would suffice to capture the effect of early parental human capital investment. However, unless $\alpha_2 = 0$, early parental reading does not capture all the dimensions of the parent's early human capital investment on the child.

Under these assumptions: (i) the estimates of β_2 in equation (4) will capture the residual effect (net of the effect of early parental reading, $R_{4/5,i}$) of reading for enjoyment and; (ii) imposing the restriction $\beta_1 = 0$ should produce a significant increase in the estimated β_2, the reading for enjoyment coefficient.[5]

An alternative transmission mechanism arises if parental reading only has a direct effect - through the reading skills equation - on the reading test score performance, but is not relevant to explain the child's acquisition of reading habits. Under this assumption, the exposure to an environment where the child can autonomously develop reading habits is what drives reading for enjoyment. In this case ($\theta_1 \neq 0$ and $\alpha_1 = 0$), β_1 and β_2 are a function of the structural parameters in equation (2) only. Furthermore, since the two effects are conditionally independent, the restricted specifications where one of these variables is left out should deliver the parameter estimates in the full specification.

The remainder of the analysis aims at exploring whether these two alternative transmission mechanisms of the effect of early parental reading may be associated with different gender profiles.

4.4. Endogeneity sources and Instrumental variables

Different sources of endogeneity can arise depending on the restrictions imposed on the model specification discussed above. Kalb and Ours (2014) note that early parental reading is likely to be correlated with unobserved factors in the reading skills equation. This may arise from the link between early parental human capital investment and current ability, since the later may be associated to the development of cognitive skills that tend to be highly correlated across time. The endogeneity of reading for enjoyment may result from two arguments: reverse causality with the reading skills since, those who are better readers read more for enjoyment and those who read more for enjoyment become better readers (see Cox and Guthrie, 2001; Canoy et al., 2006) and; correlation with unobserved factors measuring motivational aspects that influence the child's performance. This source of endogeneity can,

[5] To see this note that when equation 3 is assumed (approximately) linear: $\partial Y_i/\partial R_{4/5,i} \propto (\beta_1 + \beta_2\alpha_1)$ so that the coefficient on $R_{4/5,i}$ measures its total effect (direct and indirect) and; $\partial Y_i/\partial J_i = Y_i/\partial u_i^* \propto \beta_2$ where $u_i^* = J_i - R'_{4/5,i}\boldsymbol{\alpha_1}$ and the coefficient on J_i is the coefficient of its residual effect once the effect of parental reading is discounted.

however, be undetected if J_i is a function of $R_{4/5,i}$, since its reduced form coefficient captures a net effect.

In addition to the two variables of interest, any other choice variable in the reading score equation associated with motivational or effort aspects of performance is likely to be endogenous. An example of such a variable also used by Kalb and Ours (2014) but assumed as exogenous is the number of TV hours watched during weekdays.

The results in Kalb and Ours (2014) show that early parental reading is positively correlated with the unobservables in the reading skills equation. They suggest that "...parents who observe that their children have insufficient reading skills will start to read more to them to help develop the child's skills", thus inducing a positive correlation with the error equation. In fact, this behaviour that reflects the parent's ability and willingness to develop their children's skills is an unobservable variable in the reading skills equation. Furthermore, it is likely to be correlated with both reading for enjoyment, as parents can promote it, and with children's daily exposure to TV, as parents may restrict it.

To address these sources of endogeneity the following instrumental variables are used. Following Kalb and Ours (2014) discussion, the "number of siblings" and "whether the child is the oldest child in the family at the age of 4/5" can be used as instruments for early parental reading, since these are likely to affect the time parents had available for reading to the child during the early childhood.

A valid instrumental variable for reading for enjoyment should be correlated with the child's development of autonomous reading habits, but exogenous to the reading skills. A natural candidate, suggested by the specification of equation (3), is the regularity of visits to libraries between the ages "2 to 3" and "6 to 7".[6] An instrumental variable for the number of TV hours watched in weekdays, reflect whether the child develops other interests outside school activity. Thus, two additional variables will be considered that identify whether the child regularly participates in sports or art activities during the week.

5. Estimation

5.1. Reading for enjoyment and parental reading: Joint probit and ordered probit

Testing whether J_i depends on $R_{4/5,i}$ can be performed by specifying a binary regression model for equation (3) conditionally on both forms of early parental human capital investment variables and additional contextual variables. To account for possible endogeneity of the categorical $R_{4/5,i}$ the probit equation is jointly estimated with ordered probit specification by FIML where exclusion restrictions arising from the choice of instrumental variables are imposed in the probit equation.

[6] This is categorical variable that distinguishes children that in the age intervals 2/3, 4/5 and 6/7: never visited a library regularly, visited regularly in one period, two periods or in all three periods.

5.2. Reading achievement equation

The dependent variable Y_i is the NAPLAN reading test score of children at the age of 8/9 years which takes any value in the 0 to 10 interval. Upon suitable scaling, a fractional regression model (see, e.g. Papke and Wooldridge, 1996) is the natural candidate for $G(\cdot)$ in equation (4). If $G(\cdot)$ is taken to be the logistic function then the reduced form specification becomes

$$Y_i = 1 - (1 + \exp(R'_{4/5,i}\beta_{\mathbf{1}} + \beta_2 J_i + X'_i\beta_{\mathbf{3}} + \varepsilon_i))^{-1} \tag{5}$$

In this specification the error term ε_i, includes unobserved ability and unobserved effort. The variables in X_i include: child age, child parents' education level, parents' age, child's health status, participation in extra tutoring activities, whether the child likes physical activities, average number of TV hours watched on weekdays and on weekends, number of children's books in the home, whether another language than English is spoken at home to the child, whether the child has TV in bedroom, whether the child has access to internet, logarithm of total household income.[7] As a proxy of prior achievement we included also the Peabody Picture Vocabulary Test (PPVT-III).[8] In appendix A, table A1 presents descriptive statistics of the variables used in this model.

Early parental reading is in the reading score equation an endogenous treatment. In the Appendix B, we extend the IV methods in Wooldridge (2010) chapter 21, to estimate the average treatment effect (ATE), to account for nonlinearity of the outcome regression, multivalued treatment and endogenous controls. We show that given a set of instrumental variables, IV estimation can proceed by exploiting the discrete nature of the endogenous variables, $R_{4/5,i}$, J_i and TV_i, to derive instruments from respectively, Ordered Probit, Probit and Logit Fractional first stage regressions fitted values. [9]. An issue arises from the categorical nature of the treatment indicator, since two instruments are required for parameter identification. However using the "as good as randomly assigned" property (see Angrist and Pischke, 2009, chapter 4) allows to derive an additional instrument, given by the product between a category of $R_{4/5,i}$ - the dummy indicator for the 3-5 days a week frequency of parental

[7] The definition of parents does not necessarily refer to the biological parent. Instead the concepts of "Parent 1" and "Parent 2" are used to denote, respectively, the person who knows the most about the child and, another person in the household with a parental relationship to the child, or the partner of "Parent 1".

[8] In LSAC a short and adapted form of PPVT test was applied to the children. In the PPVT test the child points to (or says the number of) a picture that best represents the meaning of the word read out by the interviewer.

[9] Exploiting nonlinearity of the first stage regression can be useful for parameter identification for several reasons. Firstly, efficient GMM instruments are often a function of conditional expectations of the endogenous variables. In the linear case Newey (1990) shows that, if a nonlinear model gives a better approximation of the first stage regression, the resulting IV estimator will be more efficient than the standard 2SLS. Secondly, nonlinearity allows to enrich the set of instruments providing overidentifying restrictions and therefore more efficient estimates. Thirdly, if the instrumental variables are weak exploiting nonlinearity may be crucial for parameter identification. In the limiting case of non-relevant instrumental variables, nonlinearity provides internal instruments, that allow identification by functional form (see Escanciano et al. (2016) for discussion of identification by functional form). A recent example can be found in Bun and Harrison (2019), where it is shown, the usefulness of using nonlinear functions of the exogenous regressors only, as instrumental variables, in the linear regression with endogenous interactions.

reading - and its original instrumental variable.

Let Z^* be the extended set of instrumental variables obtained by adding to the nonlinear fitted values, a sub-set of relevant instrumental variables.[10] Under the assumption 2.c in Appendix B the moment condition,

$$E\left[1 - T(y)\exp\left(- R'_{4/5,i}\beta_{\mathbf{1}} - \beta_2 J_i - \beta_3 TV - X'_{2i}\beta_{\mathbf{3}}\right)\middle| X_{2i}, Z^*\right] = 0 \quad (6)$$

defines the GMM IV estimator proposed in Ramalho and Ramalho (2017). It delivers consistent estimates, under the presence of random effects type of unobserved heterogeneity, when both the treatment indicator $R_{4/5}$, and the variables J and TV are endogenous.

In Appendix B a model specification analysis addressing the key properties of the GMM estimator is developed. First, the proposed procedure suggests the implementation of a form of the Ramsey (1969) RESET functional form specification test, here denoted by ATESET (Average Treatment Effect Specification Test), that addresses the consistent estimation of the Average Treatment Effect (ATE). Second, empirically testing the validity of the proposed product instrument is performed by a standard C-test (see Ruud, 2000; Hayashi, 2000; Eichenbaum et al., 1988). This test, reveals the subset of orthogonality conditions generated by the product instrument validity, based on the relation between the J statistics of the full IV estimator and their restricted versions. Third, relying on nonlinearity to produce overidentifying restrictions, comes at the risk of weak identification of the IV estimator, if the rank condition is not satisfied (see, e.g. Nelson and Startz, 1990a,b). The risk is increased if some instrumental variables do not satisfy the relevance condition.[11] Thus, along with the identification by functional form strategy, the strong identification test proposed by Inoue and Rossi (2011) is implemented. The test infers whether the instruments are sufficiently strong so that standard inference is reliable, while maintaining the exogeneity condition. This allows for testing the two features of the estimated instruments on which the proposed GMM IV relies, namely, nonlinearity and relevance of the instrumental variables.

6. Results

This section applies the methods proposed to provide empirical evidence on the difference between boys and girls, in the transmission mechanism of early parental reading into reading achievement.

6.1. Reading for enjoyment and parental reading

Table 1 shows the joint probit and ordered probit ML estimates, of the reading for enjoyment and early parental reading equations, for both boys and girls sub-samples. The interest lies in the probit

[10] This sub-set of original instrumental variables was chosen as to minimise the variance of the GMM estimator

[11] Note that this condition may not suffice to guarantee desirable properties of the IV estimators. The "weak instruments" problem can arise even when the first stage regression coefficients tests are significant at the conventional levels (see Staiger and Stock, 1997).

estimates of the early parental human capital investment variables, namely, the regularity of library visits between the ages "2 to 3", "4 to 5" and "6 to 7" and, the early parental reading intensities at the age "4/5".

The results show a significant difference between the two genders. While in the boys sub-sample the probability of reading for enjoyment increases with exposure to a (high) frequency of parental reading at the age of 4-5 (6-7 times a week), in the girls sub-sample there is no evidence of such dependence. Furthermore, in the boys sub-sample, conditional on the included regressors, there is a significant positive correlation between the unobservables of both equations. This suggests that boys that have been read to frequently have a higher propensity to enjoy reading because, for example, parental book reading increased their motivation to read.

These results are consistent with the behaviour described in Section 4.3, where the development of reading for enjoyment depends on the parents' direct involvement, through either the frequency of parental reading or other unobservables related to the parents' willingness to read.

The regularity of library visits also exhibits a distinct effect on the probabilities of reading for enjoyment of boys and girls. For boys, the effect is significant only when the child visited uninterruptedly since the age of 2/3, while for girls, the effect is significant at all frequency levels. In particular, the results suggest that the effect on girls' reading for enjoyment is most significant on children who regularly attend libraries until the age of 4-5. For this gender, the presence of children's books at home has a positive and statistically significant effect (at the 5% significance level), on the development of reading habits. Contrarily to boys, this suggests that the exposure of the child to a stimulating reading environment contributes to the development of reading for enjoyment in girls. This is consistent with the behaviour described in Section 4.3 and is further emphasised by the absence of correlation between the unobservables of the two equations.

6.2. Reading scores model: the transmission mechanism of early parental reading

The previous results showed that boys benefited from early parental reading to develop reading for enjoyment, while girls developed this habit without the parents' direct intervention. We now apply the strategy proposed in Section 4.3 to confirm these findings, while estimating the causal effect of $R_{4/5,i}$ and J_i in the outcome equation. The model selection, and in particular the exogeneity properties of the variables $R_{4/5,i}$, J_i and TV_i, and the robustness analysis are discussed in detail in Section 7.

The first column of Table 2 shows the GMM IV estimates of the preferred specification for the boys sub-sample, where $R_{4/5,i}$ and TV_i were the variables identified as endogenous. This result is consistent with J_i being functionally dependent of $R_{4/5,i}$. It implies that it suffices to instrument $R_{4/5,i}$ to account for the source of endogeneity not related to effort (see discussion in Section 4.4). In this specification, both $R_{4/5,i}$ coefficients are significant at a 1% significant level and exhibit the expected sign. The higher the frequency of early parental reading the higher the expected reading test score. The coefficient on J_i is significant at a 5% significant level, but small in magnitude. However, because in this sub-sample the two variables are functionally related, this coefficient captures the effect of J_i net of its dependence on $R_{4/5,i}$. Consequently, as discussed in Section 4.3 a specification that omits this variable should deliver a substantially larger estimate. The results of the restricted specification presented in the second column of the table, obtained by omitting $R_{4/5,i}$, confirm this argument.[12] The reading for enjoyment estimated coefficient is now 10 times larger and statistically significant at a 1% significance level.

The two equations are informative on the causal effect of both early parental reading and reading for enjoyment on the reading score equation: the full specification provides the average treatment effect of $R_{4/5,i}$, while the restricted version allows inferring on the effect of J_i. The bottom of Table 2 shows the two forms of the RESET functional form specification test. In both cases the null hypothesis of correct functional form is not rejected. In particular, the restricted version (ATESET) does not reject the omission from the equation of individual specific treatment effects that would render inconsistent the estimates of the average treatment effect of $R_{4/5,i}$.

We now discuss the results for the girls sub-sample, to bring further evidence on the independence between $R_{4/5,i}$ and J_i in the reading test score equation. Table 3 shows three sets of GMM IV estimates: the full and preferred specification where both $R_{4/5,i}$ and J_i were identified as endogenous, and two restricted versions derived by omitting one of these variables. The absence of a functional relation between $R_{4/5,i}$ and J_i in the girls sub-sample, explains why these are two independent sources of endogeneity. Furthermore, as discussed in Section 4.3, under this assumption, the coefficients on J_i and $R_{4/5,i}$ in the reading score equation can be estimated separately. The results show that when $R_{4/5,i}$ is omitted the coefficient of J_i is 70% larger. However at a 5% significance level the equality of the coefficients cannot be rejected, and the 95% confidence intervals are respectively $(0.097, 0.555)$ and $(0.323, 0.820)$ for the full and the restricted specification, therefore overlapping for around 50% of their range. On the other hand, when J_i is omitted, the coefficients of $R_{4/5,i}$ in the restricted and full specification are very similar, providing further evidence that for girls the two effects are independent, and therefore their impact in the reading test scores is cumulative.

In contrast with the boys equation, the full specification provides all information needed to infer on both the early parental reading treatment effect and the effect of reading for enjoyment. Interestingly, the ATESET test in the J_i restricted specification, rejects at a 5% significance level the null hypothesis under which the model identifies correctly the average treatment effect. Indeed, since the parametric specification is nonlinear all partial effects are heterogeneous, as such, omission of relevant variables

[12] As expected, in this specification, the variable J_i becomes endogenous, reflecting its dependence on $R_{4/5,i}$.

results in their misspecification, and in particular of the average treatment effect associated to early parental reading.

Low performing boys are more likely to be affected by TV watching. This may be relate to the type of programs they watch, to weather they prioritise doing homework before watching TV, etc.

The equation shows that frequency of parental book reading at the age of 4/5 is associated with TV viewing. This suggests that the parents that read frequently to their children valued a literacy-rich educational environment from early on that may have repercussions later in life, namely at the age 8/9.

Tables 4 and 5 show estimates of the (conditional) partial effects under the assumption that the dependence between observables and unobservables is restricted to the conditional mean (see Ramalho and Ramalho, 2017).[13] All partial effects are standardised with respect to the distribution of reading test score. This allows to compare the magnitude of the effects in the boys and girls distribution.

Table 4 shows the conditional average partial effects of $R_{4/5,i}$ for both boys and girls computed from the full specifications. The point estimates show that early parental reading has a larger impact on the distribution of the boys reading score tests, and that reading for enjoyment only slightly affects this partial effect in the girls' specification. For boys, this is a consequence of the small coefficient of J_i, whereas for girls, it reflects the independence of the two effects. The partial effect of being read to "3-5 days a week", represents a shift in the boys' distribution of the reading score tests from the percentile .275 to the median or a shift from the median to the percentile .725. For the girls' distribution, these are respectively .35 and .725. The higher parental reading frequency partial effect represents a shift on test score distribution, from the percentile .20 to the median for both boys and girls, or a shift from the median to the percentiles .825 for boys and girls if $J_i = 0$ and of .875 if $J_i = 1$.

Table 5 shows the partial effects of J_i. The boys' estimate is computed from the restricted specification and therefore is not conditional on $R_{4/5,i}$. The second row of the table shows the joint effects of J_i and $R_{4/5,i}$ on the girls' tests scores. The boys' figure represents a shift in the score test distribution from the 5% percentile to the median or from the median to the 95% percentile. This is only comparable to the cumulative effect on girls, of reading for enjoyment and having been read to at early ages: "if 3-5 days a week" those percentiles are 10% and 92.5% and in 6-7 days a week they become 5% and 97.5%.

In short, our results show that the effect of both early parental reading and reading for enjoyment on reading achievement is different for boys and girls. The former seems to develop reading for enjoyment via early exposure to parental book reading, while girls develop this independently from the parents'

[13]This estimator is a natural extension of the smearing technique suggested in Duan (1983) and consists of a two-step procedure: (i) estimation of the errors from the residuals implicit in moment condition (6) and; (ii) averaging across the residuals the derivative of the fractional regression at specified values of the covariates (sample means).

early investment in reading to their children. Thus, frequent parental book reading at the age of 4-5 is a transmission mechanism for boys: it impacts on reading for enjoyment, which in turn affects their future reading achievement at age 8-9. For girls, parental reading impacts directly on future reading achievement at age 8-9. Thus, for girls, we have two independent factors that add to each other: parental reading and reading for enjoyment. This suggests that human capital investment other than parental book reading is enough for girls to engage in reading for enjoyment at the age of 8-9, while for boys parental reading is of great importance.

Additionally, our results in Table 1 suggest that human capital investment, at least through library visits, acts differently at different ages for boys and girls. For boys, a continuous investment in library visits is necessary for it to have a significant impact. For girls, a moderate investment is enough to create a lasting impact.

7. Robustness analysis

This Section describes the model selection and, in particular, the endogeneity assumptions, and provides the robustness analysis. It discusses the instrumental variables strategy by testing the crucial identification assumptions of the GMM IV estimator, and provides estimates under alternative GMM estimators.

7.1. Model selection: endogeneity pattern

Consistent estimation of equation (5) can be attained by instrumenting all potential endogenous variables. However, this would ignore that the nature of the transmission mechanism of early parental reading may determine the exogeneity properties of reading for enjoyment in the outcome equation. Furthermore, efficiency considerations require that the set of endogenous variables is correctly identified.

The first pane of Tables 6 and 7 show the GMM IV estimates of the three potential endogenous variables under the possible combinations of exogeneity assumptions. In addition to the first stage fitted values, the estimators use the set of original instrumental variables that minimises the estimated variance of the endogenous coefficients. The first column shows the GMM coefficients under the exogeneity assumption, while the last column assumes all three variables are endogenous. A comparison of these two specifications reveals a significant bias in all coefficients of the exogenous specification. Furthermore, results show that the magnitude and significance of the estimated coefficients are very sensitive to the exogeneity assumptions.

The diagonal entries in the second pane of Tables 6 and 7 show p-values for the J-statistic of Hansen (1982). The test is informative on the orthogonality conditions validity, implying both the exogeneity and the exclusion restriction assumptions on the instruments. However, it is well acknowledged that the validity of the overidentifying restrictions is neither a sufficient nor necessary condition for the validity of the moment conditions(see e.g., Newey, 1985). The OIR may be satisfied under the presence of endogenous regressors, in which case the estimator is asymptotically biased (see Parente and Silva, 2012). Thus, given the OIR test lack of power, its results are complemented with Durbin-Wu-Hausmann (DWH) exogeneity tests to identify the specification with the subset of variables correctly assumed as endogenous (see Durbin, 1954; Wu, 1973; Hausman, 1978). The $\{i, j\}$ entries refers to the exogeneity test of the regressor in row i, not included as endogenous in the specification of column j.

Results for the boys' sub-sample show that the J-statistic is unable to identify a unique specification satisfying the orthogonality conditions, but it rejects the null hypothesis whenever early parental reading is considered to be exogenous (columns 3, 4 and 6). However, the DWH test shows that the specification where $R_{4/5,i}$ and TV_i are assumed endogenous (column 7) is the only one that does not reject the exogeneity of the variable left out of the endogenous set (J_i).[14]

The second panel of Table 7 shows the results in the girls sub-sample. It is evident the OIR test inability to uniquely identify a set of valid moment conditions since, in none of the specifications, the null hypothesis is rejected[15]. However, only when both $R_{4/5,i}$ and J_i are considered endogenous (column 5) the exogeneity of the variable left out of the potential endogenous set is not rejected by the DHW test. In all other specifications, exogeneity of the variable left out is rejected, in particular, when $R_{4/5,i}$ and J_i are individually assumed as endogenous (columns 2 and 3), thus suggesting that these are two independent sources of endogeneity.

The implicit bias of the GMM estimates obtained under the exogeneity assumption is informative on the correlation between the variables identified as endogenous and the unobservables in the reading achievement equation. The $R_{4/5,i}$ in both boys and girls sub-samples, and the J_i in the girls' sub-sample estimated coefficients, imply a negative correlation with the error term. This suggests that, children whit a worst reading achievement performance were more read to at the age of 4/5, and are more likely to read more outside the school duties. The TV at weekdays estimate in the boys' equation implies a positive correlation, thus suggesting that low achieving boys watch less frequently TV during weekdays. These results are similar to those reported by Kalb and Ours (2014) and interpreted as describing the parent's efforts to develop the child's skills by reading when detected as insufficient. The same reasoning applies to parents encouraging independent reading and restricting access to TV as a means to increase the academic performance of low achievers children.

[14] Further insight into the OIR test results that do not reject the null hypothesis (validity of the orthogonality conditions) can be gained from the estimated coefficients of the potential endogenous variables. When early parental reading is assumed endogenous and TV watched in weekdays is (wrongly) assumed as exogenous (columns 2 and 5), its coefficient is very close to zero in value and statistically not significant, suggesting that the GMM IV orthogonality conditions are artificially satisfied by excluding this variable from the conditioning set.

[15] Again the GMM estimated coefficients of the assumed exogenous variables tend to be very close to zero.

7.2. Instrumental variables and strong identification

The identification strategy exploits both the "as good as random" property to the generate an additional product instrument for $R_{4/5}$ and, the nonlinearity of the first stage regressions. We now investigate the properties of the GMM IV estimators with respect to these instrumental variables.

The C-test performed on the product instruments for $R_{4/5}$ "3-5 days a week" in all specifications where this variable is assumed endogenous do not reject the null hypothesis, therefore confirming the validity of these instrumental variables.

The third pane of Tables 6 and 7 show the results of the Inoue and Rossi (2011) strong identification test. The test statistic is based on the distance between two GMM IV shrinkage estimators with different asymptotic distributions. This is achieved by considering the just-identified estimator that uses only the fitted first stage regressions as instruments. Implementation of the test statistic requires choosing a value for the shrinkage parameter. The cross-validation procedure proposed in Inoue and Rossi (2011) that minimises a function (trace and determinant) of the mean-squared error of the shrinkage estimator was adopted.[16]

The results show that in all specifications the null hypothesis of strong identification of GMM IV estimates is not rejected. This implies that the relevance condition is satisfied for which a necessary condition is the validity of the rank condition. In particular the test results validate the use of the extended set of instruments that arise from exploiting the nonlinearity of the first stage regressions.[17]

Table 8 shows the first stage (nonlinear) regressions estimates of the preferred full specifications. The results for the boys sub-sample show that while the ordered probit coefficients of the instrumental variables in the equation for $R_{4/5,i}$ are statistically significant, the only significant variables in the fractional regression for TV_i are the included exogenous variables. As such, the (strong) identification result of the GMM IV estimator, relies on the nonlinearity of the estimated instrument for TV_i obtained from the logistic regression.[18]

The third and fourth columns of Table 8 show the first stage nonlinear regressions estimates of the preferred full specification in the girls sub-sample. The instrumental variables in both nonlinear first stage regressions are statistically significant, and thus the quality (relevance) of the estimated instruments exploits both the exogenous source of variation and the nonlinearity of the econometric specification.

---

[16] Results from a response surface analysis (see Davidson and Mackinnon, 2004) of the Monte Carlo experiments in Inoue and Rossi (2011) that exploit a nonlinear relationship between the empirical size of the test, the sample size and the shrinkage parameter were also used and the results without significant differences

[17] Note that the test statistic assumes the validity of the instruments. Furthermore when all additional instrumental variables were added the test statistic rejected the null hypothesis therefore identifying non relevant/weak instruments.

[18] Another source of identification arises by the existence of data driven exclusion restrictions, i.e., exogenous variables that are statistically significant in the first stage regression but not in the outcome equation.

7.3. Alternative estimators: GMM CUE and IVs

This section performs additional robustness checks that exploit implications of weak identification in GMM estimators. The first relies on the robustness of the Continuous Updating Estimator (GMM CUE) to the weak instruments problem (see, e.g. Stock et al., 2002). The second, acknowledges that if identification is weak GMM estimates can be sensitive to the addition of instruments.

Table 9 shows the two step GMM IV and the GMM CUE estimates of the endogenous variables in the full and restricted specifications. For both sub-samples the estimates are very similar, providing evidence that the 2SGMM is indistinguishable from the weak instrument robust GMM CUE.

Tables 10 and 11 shows the sensitivity of the endogenous variables estimated coefficients to different combinations of the instrumental variables. The variables that were found to be weak instruments were excluded from this set.[19] Two sets of estimates were considered. The first uses the original instrumental variables only (inefficient) and the second includes the fitted nonlinear regression as additional instruments (efficient)[20]. The estimates of the preferred specification are presented in the last row of the table. The GMM IV efficient estimates are very similar and tend to converge and to become more efficient as the number of instruments increase. The coefficient of TV at weekdays exhibits a worse behaviour, in terms of precision, when being the oldest child is excluded from the instrument set. Comparing the inefficient and efficient estimates is revealing of the fitted values from the first stage regression identification power. In its absence none of the GMM IV estimators delivers statistically significant estimates of the TV at weekdays coefficient in the boys sub-sample and of the reading for enjoyment coefficient in the girls sub-sample.

8. Conclusion

The development of reading skills is an essential determinant of academic success and later on in life in terms of earnings. Parents have a crucial role in stimulating reading habits even before their children start education at school. Previous work has identified the impact of reading at the age of 4-5 years of age on children's reading skills (Kalb and Ours, 2014). In this paper, early parental reading is considered as one form of early human capital investment, together with the regularity of visits to libraries during childhood. We investigate the role of both in the development of autonomous reading habits and ultimately in a child's reading achievement. Importantly, we also add to the existing evidence by identifying gender-specific transmission mechanisms.

While the results show that early parental reading is an essential determinant of reading achievement at ages 8-9, the nature of its transmission mechanism differs by gender. For boys, the parents'

[19] The Inoue and Rossi (2011) strong identification test rejected the null hypothesis when this variables were included.
[20] The term efficient is used here in relation to the two proposed estimator only

direct intervention through reading to them at early stages develops their reading for enjoyment. The results show that this dependence arises also from unobserved factors related to a higher parental propensity to read. Conversely, girls are less dependent on parents' direct involvement to develop reading habits. For them, reading for enjoyment is not directed linked to parental reading, nor to unobserved factors in this propensity, once we condition on the early frequency of library visits.

In both groups, reading for enjoyment at ages 8-9 was stimulated by the frequency of visits to the library at early ages. This is true for girls, regardless of the continuity of these visits in time. At the same time, it is only relevant for boys who visited libraries regularly throughout the early years, since the ages of 2-3. In any case, this result highlights the importance of other forms of early human capital investment that can stimulate a child's interest in reading activities.

The analysis is based on GMM IV estimates of a reading achievement reduced form equation that embeds a sequence of behavioural relationships between reading skills, early parental reading, reading for enjoyment and regularity of visits to libraries during childhood. In the full specification, early parental reading is an endogenous multivalued treatment, while the effect of reading for enjoyment depends on whether it depends on that form of parental intervention. The nature of the transmission mechanism of early parental reading is inferred from restricted versions of the full equation. These revealed and confirmed the gender-specific (in)dependence between reading for enjoyment and early parental reading, as estimated by an auxiliary FIML equation.

In addition to the two variables of interest, the specification identifies the TV exposure during weekdays as endogenous in the boys' specification. This further increases the set of instrumental variables required to identify the GMM IV estimator. Estimation of the average treatment effect of early parental reading extends the methods in Wooldridge (2010) chapter 21 to consider: multivalued endogenous treatment, other endogenous regressors in the context of a fractional regression model. In addition, the procedure suggested a form of the functional misspecification test that addressed the consistent estimation of the average treatment effect.

Exploiting the nonlinearity of the first stage regressions of the endogenous variables provided a valuable source of identification, in particular where the quality of the instruments was low. The estimator robustness to the risk of a weak instruments problem was assessed by the implementation of the Inoue and Rossi (2011) strong identification test.

An important feature of the analysis is the identification of the different endogeneity properties of those variables in the boys and girls specification. According to the empirical model, they are associated with the nature of the transmission mechanism of early parental reading. The results confirm its gender-specific nature and the (in)dependence between early parental reading and reading for enjoyment. For boys, endogeneity in the reading achievement equation arises from early parental reading, whereas in girls, the two variables are independent sources. The estimates obtained from a specification ignoring endogeneity are severely biased. The bias direction suggests that parents with children who show less (cognitive) skills are more read to. A similar result is found in reading for enjoyment, which indicates that in such cases, the variable should be interpreted as independent reading. Again, children with fewer skills may be encouraged to read outside the school duties to improve.

These findings are similar to those reported in Kalb and Ours (2014) and suggest that parents are proactive in helping their children overcome difficulties in their early acquisition of reading abilities. As the authors state "...children who have better reading skills are less likely to be read to, or similarly, children with worse reading skills are more likely to be read to. It could be the case that parents who observe that their children have insufficient reading skills will start to read more to them to help develop the child's skills".

Our study is unique in that it considers aspects of the home literacy environment early in life, its relation with reading engagement during primary school and the impact of these two factors on reading achievement. The reading studies reviewed indicate that parental book reading has an impact on later reading habits (Sénéchal et al., 2006; Sénéchal and Young, 2008) and that reading for enjoyment (OECD, 2010; Sénéchal, 2012) impacts reading performance. But the evidence was lacking regarding the influence of the two factors on reading achievement. For the first factor, we show this influence considering children of different genders and how their reading performance is differently affected by direct human capital investment in the form of book reading and by the indirect effect of visits to the library. Considering the second factor, we show not only that it influences reading performance, but also that it is more dependent on parental reading for the boys.

The finding that boys are more dependent on parental reading during early childhood to later engage in reading for enjoyment, suggest that this activity should be viewed by parents as a crucial one to help their sons develop reading skills, since research shows that those who read more become better readers (Cox and Guthrie, 2001; OECD, 2010; Stanovich, 2000). This is not to say that parents do not need to read to their daughters, but simply that girls are more prone to develop reading for enjoyment, even without this direct intervention from parents. In fact, visits to the library are likely parent-initiated and may be, in and by themselves, sufficient for girls to develop reading for enjoyment. Additionally, we must acknowledge that children's response to the frequency of engagement in reading for enjoyment does not necessarily reflect only self-selection or choice. In other words, it may be that parents and/or teachers assign home readings. The students' questionnaire does not distinguish between these two possibilities. Nonetheless, our data shows that girls engage in this activity more often than boys. While it is unlikely to conceive that teachers differentiate between boys and girls as far as the amount of assigned readings at home, parents might do it based on gender preferences.

Clearly, children are off to a better start in life if they have been exposed to parental reading in the early years. However, our findings also suggest that visits to the library per se may influence reading achievement. As such, societal environments that promote reading and reading-related activities have an important role in motivating students to read and consequently in making them better readers.

We used the NAPLAN reading score as the outcome measure at ages 8-9 to examine the determinants of the performance of boys and girls, because it is the first moment Australian students are evaluated at the national level. Subsequent NAPLAN test scores are likely to be determined by students' previous reading achievement.

Acknowledgements: The authors thank Nicoletta Rosati, Massimiliano Bratti and seminar participants at the 5th LEER Workshop on Education Economics at the University of Leuven, at the 2018 Budapest International Conference on Education Economics at the Centre for Economic and Regional Studies of the Hungarian Academy of Sciences, at the 5th Lisbon Research Workshop on Economics, Statistics and Econometrics of Education at ISEG, and at the 12th ARLE Conference, for their comments and suggestions.

References

Adams, M.J., 2009. The challenge of advanced texts: The interdependence of reading and learning.. in e. hiebert (ed.) ed.. Guilford Press, New York, NY. chapter Reading more, reading better: Are American students reading enough of the right stuff? pp. 163–189.

AIFS, 2015. Longitudinal Study of Australian Children Data User Guide. Australian Institute of Family Studies, Melbourne.

Angrist, J.D., Krueger, A.B., 1991. Does compulsory school attendance affect schooling and earnings? The Quarterly Journal of Economics 106, 979–1014.

Angrist, J.D., Pischke, J.S., 2009. Mostly Harmless Econometrics. Princeton Univers. Press.

Araújo, L., Costa, P., 2015. Home book reading and reading achievement in eu countries: the progress in international reading literacy study 2011 (pirls). Educational Research and Evaluation: from Theory to Practice 21, 422–438. doi:10.1080/13803611.2015.1111803.

Baker, M., Milligan, K., 2016. Boy-girl differences in parental time investments: Evidence from three countries. Journal of Human Capital 10, 399–441. doi:10.1086/688899.

Bun, M.J.G., Harrison, T.D., 2019. OLS and IV estimation of regression models including endogenous interaction terms. Econometric Reviews 38, 814–827. doi:10.1080/07474938.2018.1427486.

Canoy, M., van Ours, J.C., van der Ploeg, F., 2006. The economics of books. victor a. ginsburgh and david throsby (eds.) ed.. Handbook of the Economics of Art and Culture, Amsterdam, Elsevier.

Chall, J., Jacobs, V., 2003. Poor children's fourth-grade slump. American Educator 27, 14–15.

Clark, C., Rumbold, K., 2006. Reading for pleasure: A research overview. Technical Report. United Kingdom: National Literacy Trust.

Cox, K.E., Guthrie, J.T., 2001. Motivational and cognitive contributions to students' amount of reading. Contemporary Educational Psychology 26, 116–131. doi:10.1006/ceps.1999.1044.

Cunningham, A.E., Stanovich, K.E., 1997. Early reading acquisition and its relation to reading experience and ability 10 years later. Developmental Psychology 33, 934–945. doi:10.1037/0012-1649.33.6.934.

Davidson, R., Mackinnon, J.G., 2004. Econometric theory and methods. Oxford University Press, New York.

Duan, N., 1983. Smearing estimate: A nonparametric retransformation method. Journal of the American Statistical Association 78, 605–610. doi:10.1080/01621459.1983.10478017.

Duff, D., Tomblin, J.B., Catts, H., 2015. The influence of reading on vocabulary growth: A case for a matthew effect. Journal of Speech, Language, and Hearing Research 58, 853–864.

Durbin, J., 1954. Errors in variables. Review of the International Statistical Institute 22, 23. doi:10.2307/1401917.

Eichenbaum, M.S., Hansen, L.P., Singleton, K.J., 1988. A time series analysis of representative agent models of consumption and leisure choice under uncertainty. Quarterly Journal of Economics 103, 51–78. doi:https://doi.org/10.2307/1882642.

Escanciano, J.C., Jacho-Chávez, D., Lewbel, A., 2016. Identification and estimation of semiparametric two-step models. Quantitative Economics 7, 561–589. doi:10.3982/QE328.

Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029. doi:10.2307/1912775.

Hanushek, E.A., Woessmann, L., 2015. The Knowledge Capital of Nations. The MIT Press.

Hausman, J.A., 1978. Specification tests in econometrics. Econometrica 46, 1251. doi:10.2307/1913827.

Hayashi, F., 2000. Econometrics. Princeton University Press.

Heckman, J., 1997. Instrumental variables: A study of implicit behavioral assumptions used in making program evaluations. The Journal of Human Resources 32, 441. doi:10.2307/146178.

Heckman, J., Jacobs, B., 2011. Policies to create and destroy human capital in europe , 253–322doi:10.7551/mitpress/9780262015318.003.0009.

Inoue, A., Rossi, B., 2011. Testing for weak identification in possibly nonlinear models. Journal of Econometrics 161, 246–261. doi:10.1016/j.jeconom.2010.12.012.

Kalb, G., Ours, J.C.V., 2014. Reading to young children: A head-start in life? Economics of Education Review 40, 1–24. doi:10.1016/j.econedurev.2014.01.002.

Loveless, T., 2015. Brown Center Report on American Education: How Well Are American Students Learning? Technical Report. Brookings Institute.

Mol, S.E., Bus, A.G., 2011. To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. Psychological Bulletin 137, 267–296. doi:10.1037/a0021890.

Mullis, I.V.S., Martin, M.O., Foy, P., Hooper, M., 2017. PIRLS 2016 International Results in Reading. Technical Report. Boston College, TIMSS & PIRLS International Study Center. URL: http://timssandpirls.bc.edu/pirls2016/international-results/.

Nagy, W.E., Herman, P.A., Anderson, R.C., 1985. Learning words from context. Reading Research Quarterly 20, 233. doi:10.2307/747758.

Nelson, C.R., Startz, R., 1990a. The distribution of the instrumental variables estimator and its t-ratio when the instrument is a poor one. The Journal of Business 63, S125–S140.

Nelson, C.R., Startz, R., 1990b. Some further results on the exact small sample properties of the instrumental variable estimator. Econometrica 58, 967. doi:10.2307/2938359.

Neuman, S., Dickinson, D.K., 2001. Handbook of early literacy research. Guilford Press, New York.

Newey, W.K., 1985. Generalized method of moments specification testing. Journal of Econometrics 29, 229–256. doi:10.1016/0304-4076(85)90154-X.

Newey, W.K., 1990. Efficient instrumental variables estimation of nonlinear models. Econometrica 58, 809. doi:10.2307/2938351.

OECD, 2010. PISA 2009 Results (Volume III): Learning to Learn : Student Engagement, Strategies and Practices. OECD Publishing, Paris, PISA.

OECD, 2016. PISA 2015 Results (Volume I): Excellence and Equity in Education. OECD Publishing, Paris, PISA. doi:10.1787/9789264266490-en.

Pagan, A.R., Hall, A.D., 1983. Diagnostic tests as residual analysis. Econometric Reviews 2, 159–218. doi:10.1080/07311768308800039.

Papke, L.E., Wooldridge, J.M., 1996. Econometric methods for fractional response variables with an application to 401(k) plan participation rates. Journal of Applied Econometrics 11, 619–632. doi:10.1002/(sici)1099-1255(199611)11:6¡619::aid-jae418¿3.0.co;2-1.

Parente, P.M., Silva, J.S., 2012. A cautionary note on tests of overidentifying restrictions. Economics Letters 115, 314–317. doi:10.1016/j.econlet.2011.12.047.

Pesaran, M.H., Taylor, L.W., 1999. Diagnostics for IV regressions. Oxford Bulletin of Economics and Statistics 61, 255–281. doi:10.1111/1468-0084.00128.

Ramalho, E.A., Ramalho, J.J., 2017. Moment-based estimation of nonlinear regression models with boundary outcomes and endogeneity, with applications to nonnegative and fractional responses. Econometric Reviews 36, 397–420. doi:10.1080/07474938.2014.976531.

Ramsey, J.B., 1969. Tests for specification errors in classical linear least-squares regression analysis. Journal of the Royal Statistical Society. Series B 31, 350–371.

Ruud, P., 2000. An introduction to classical econometric theory. Oxford University Press, New York.

Sénéchal, M., 2012. Child language and literacy development at home. London: Taylor and Francis. pp. 38–50. doi:10.1002/9780470757642.ch13.

Sénéchal, M., Hill, S., Malette, M., 2018. Individual differences in grade 4 children's written compositions: The role of online planning and revising, oral storytelling, and reading for pleasure. Cognitive Development 45, 92–104. doi:10.1016/j.cogdev.2017.12.004.

Sénéchal, M., Ouellette, G., Rodney, D., 2006. The misunderstood giant: On the predictive role of early vocabulary to future reading. Handbook of early literacy research 2, 173–182.

Sénéchal, M., Young, L., 2008. The effect of family literacy interventions on children's acquisition of reading from kindergarten to grade 3: A meta-analytic review. Review of Educational Research 78, 880–907. doi:10.3102/0034654308320319.

Staiger, D., Stock, J.H., 1997. Instrumental variables regression with weak instruments. Econometrica 65, 557. doi:10.2307/2171753.

Stanovich, K.E., 2000. Progress in understanding reading : scientific foundations and new frontiers. Guilford Press, New York.

Stanovich, K.E., Cunningham, A.E., West, R.F., 1998. Literacy experiences and the shaping of cognition., in: Global prospects for education: Development, culture, and schooling.. American Psychological Association, pp. 253–288. doi:10.1037/10294-009.

Stock, J., Yogo, M., Wright, J., 2002. A survey of weak instruments and weak identification in generalized method of moments. Journal of Business and Economic Statistics 20, 518–529.

Sullivan, A., Brown, M., 2015. Reading for pleasure and progress in vocabulary and mathematics. British Educational Research Journal 41, 971–991. doi:10.1002/berj.3180.

Verhoeven, L., van Leeuwe, J., Vermeer, A., 2011. Vocabulary growth and reading development across the elementary school years. Scientific Studies of Reading 15, 8–25. doi:10.1080/10888438.2011.536125.

Whitehurst, G., Lonigan, C., Christopher, B., Lonigan, J., 2001. Emergent literacy: Development from pre-readers to readers .

Wooldridge, J.M., 1997. On two stage least squares estimation of the average treatment effect in a random coefficient model. Economics Letters 56, 129–133. doi:10.1016/s0165-1765(97)81890-3.

Wooldridge, J.M., 2010. Econometric Analysis of Cross Section and Panel Data. The MIT Press, London, England.

Wu, D.M., 1973. Alternative tests of independence between stochastic regressors and disturbances. Econometrica 41, 733. doi:10.2307/1914093.

Table 1: Reading for enjoyment: Joint Probit and Ordered Probit

	Boys				Girls			
	Probit		Ordered Probit		Probit		Ordered Probit	
	Reads for enjoyment		Parental reading		Reads for enjoyment		Parental reading	
Constant	-0.199	(0.830)	—	—	-0.401	(0.950)	—	
Age	0.021	(0.070)	-0.102	(0.065)	0.001	(0.077)	-0.015	(0.072)
Poor health	0.230	(0.265)	0.057	(0.227)	0.864	(0.432)**	-0.906	(0.378)**
Likes sports activities	-0.017	(0.034)	0.044	(0.033)	-0.023	(0.037)	0.017	(0.034)
Non-english speaking	0.102	(0.113)	-0.290	(0.102)***	-0.003	(0.122)	-0.490	(0.110)***
Children's books home	0.089	(0.051)*	0.197	(0.046)***	0.201	(0.062)***	0.260	(0.054)***
TV in bedroom	-0.106	(0.101)	-0.219	(0.076)***	-0.088	(0.086)	-0.160	(0.078)**
TV at weekdays	-0.057	(0.044)	-0.070	(0.042)*	-0.103	(0.052)**	-0.110	(0.045)**
TV at weekend	-0.014	(0.039)	0.004	(0.036)	-0.040	(0.043)	-0.077	(0.040)*
Internet access	-0.415	(0.296)	0.126	(0.229)	0.356	(0.308)	-0.524	(0.277)*
Household income	-0.014	(0.046)	0.124	(0.039)***	-0.060	(0.062)	0.128	(0.045)***
Education Parent 1	-0.010	(0.047)	0.142	(0.043)***	0.054	(0.047)	0.148	(0.042)***
Education Parent 2	0.008	(0.042)	0.101	(0.035)***	0.020	(0.040)	0.078	(0.036)**
Age Parent 1	0.124	(0.100)	0.130	(0.089)	-0.144	(0.105)	0.309	(0.099)***
Age Parent 2	0.009	(0.082)	0.132	(0.074)*	0.200	(0.090)**	0.168	(0.084)**
Library until 4/5	-0.059	(0.094)	0.324	(0.087)***	0.332	(0.105)***	0.280	(0.097)***
Library until 6/7	-0.054	(0.089)	0.171	(0.080)**	0.193	(0.097)**	0.187	(0.085)**
Library from 2/3 to 6/7	0.207	(0.102)**	0.453	(0.100)***	0.206	(0.117)*	0.708	(0.121)***
Read to 4/5 (3-5 days)	0.070	(0.089)	—	—	0.064	(0.128)	—	—
Read to 4/5 (6-7 days)	0.309	(0.090)***	—	—	0.019	(0.179)	—	—
Oldest child	—	—	0.337	(0.077)***	—	—	0.420	(0.082)***
Number of siblings	—	—	-0.078	(0.040)**	—	—	-0.084	(0.041)**
μ_1	—	—	1.838	(0.782)**	—	—	2.869	(0.878)***
μ_2	—	—	2.708	(0.797)***	—	—	3.767	(0.878)***
ρ	—	—	0.546	(0.129)***	—	—	-0.101	(0.208)
N	1275				1282			

Robust standard errors in parenthesis. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Parameters, μ_1, μ_2 and ρ are respectively, thresholds of the Ordered Probit and correlation coefficient of the Normal distribution.

Table 2: GMM IV estimates: Boys' equation

	Full specification		Without Parental Reading	
Constant	-0.262	(0.843)	-1.730	(0.964)*
Age	-0.031	(0.029)	-0.031	(0.029)
Extra tutoring	-0.132	(0.048)***	-0.118	(0.040)***
PPVT	0.156	(0.028)***	0.164	(0.030)***
Poor health	-0.107	(0.085)	-0.222	(0.094)**
Likes sports activities	-0.032	(0.014)**	-0.010	(0.013)
Non-english speaking	0.070	(0.046)	0.072	(0.046)
Children's books home	-0.036	(0.026)	-0.007	(0.027)
TV in bedroom	-0.056	(0.035)	-0.042	(0.037)
TV at weekends	0.116	(0.059)*	0.031	(0.082)
Internet access	-0.114	(0.127)	0.070	(0.102)
Household income	0.018	(0.020)	0.023	(0.019)
Education Parent 1	-0.007	(0.027)	0.042	(0.028)
Education Parent 2	0.020	(0.019)	0.045	(0.02)**
Age Parent 1	0.052	(0.037)	0.052	(0.041)
Age Parent 2	-0.024	(0.031)	-0.020	(0.031)
TV at weekdays	-0.410	(0.189)**	-0.106	(0.261)
Reads for enjoyment	0.062	(0.030)**	0.653	(0.186)***
Read to 4/5 (3-5 days)	0.250	(0.087)***	—	—
Read to 4/5 (6-7 days)	0.393	(0.121)***	—	—
N	1275	—	1275	—
R^2	0.32	—	0.37	—
J-Hansen OIR test (p-value)	0.76	—	0.15	—
RESET test (p-value)	0.90	—	0.72	—
ATESET test (p-value)	0.76	—	—	—

Robust standard errors in parenthesis. Significance levels: $^{***}p < 0.01$ $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 3: GMM IV estimates: Girls' equation

	Full specification		Without Parental Reading		Without Reads for enjoyment	
Constant	-1.625	(0.261)***	-0.828	(0.470)*	-0.933	(0.482)*
Age	-0.030	(0.020)	-0.033	(0.026)	-0.033	(0.021)
Extra tutoring	-0.173	(0.039)***	-0.194	(0.042)***	-0.181	(0.044)***
PPVT	0.180	(0.020)***	0.153	(0.031)***	0.178	(0.023)***
Poor health	-0.074	(0.111)	-0.252	(0.092)***	0.042	(0.102)
Likes sports activities	-0.002	(0.011)	0.001	(0.013)	-0.002	(0.012)
Non-english speaking	0.199	(0.039)***	0.180	(0.046)***	0.208	(0.040)***
Children's books home	-0.028	(0.019)	-0.026	(0.023)	-0.014	(0.020)
TV in bedroom	0.002	(0.023)	0.016	(0.032)	0	(0.025)
TV at weekends	-0.010	(0.012)	0.068	(0.030)**	0.043	(0.041)
Internet access	-0.023	(0.092)	-0.097	(0.105)	0.031	(0.100)
Household income	0.022	(0.013)*	0.023	(0.016)	0.002	(0.014)
Education Parent 1	0.011	(0.014)	0.015	(0.018)	0.009	(0.016)
Education Parent 2	0.009	(0.010)	-0.002	(0.014)	0.002	(0.012)
Age Parent 1	0.016	(0.029)	0.066	(0.036)*	0.024	(0.038)
Age Parent 2	0.015	(0.026)	-0.012	(0.034)	0.021	(0.028)
TV at weekdays	0.006	(0.013)	-0.267	(0.090)***	-0.200	(0.136)
Reads for enjoyment	0.326	(0.117)***	0.552	(0.137)***	—	—
Read to 4/5 (3-5 days)	0.184	(0.080)**	—	—	0.197	(0.093)**
Read to 4/5 (6-7 days)	0.320	(0.109)***	—	—	0.318	(0.139)**
N	1282	—	1282	—	1282	—
R^2	0.44	—	0.31	—	0.38	—
J-Hansen OIR test (p-value)	0.56	—	0.89	—	0.56	—
RESET test (p-value)	0.21	—	0.56	—	0.32	—
ATESET test (p-value)	0.46	—	—	—	0.05	—

Robust standard errors in parenthesis. Significance levels: $^{***}p < 0.01$,$^{**}p < 0.05$, $^{*}p < 0.1$.

Table 4: Partial effects: Early Parental Reading

	Boys		Girls	
	$J = 1$	$J = 0$	$J = 1$	$J = 0$
Read to 4/5 (3-5 days)	0.61	0.60	0.52	0.48
	(0.20,1.01)	(0.20,1.0)	(0.07,0.81)	(0.06,0.76)
Read to 4/5 (6-7 days)	0.96	0.95	0.91	0.86
	(0.40,1.53)	(0.39,1.52)	(0.31,1.49)	(0.28,1.43)

Partial effects in the first two rows are differences with respect to read to 0-2 days.
95% confidence intervals in parenthesis computed from delta method standard errors.

Table 5: Partial effects: Reading for Enjoyment

	Boys	Girls		
		$R_{4/5}$ (0-2 days)	$R_{4/5}$ (3-5 days)	$R_{4/5}$ (6-7 days)
Reads for enjoyment	1.60	0.88	0.91	0.92
	(0.75,2.45)	(0.25,1.49)	(0.28,1.53)	(0.28,1.56)
Reads for enjoyment and Parental reading	—	—	1.39	1.78
			(0.85,1.94)	(1.20,2.35)

Partial effects from boys specification estimated from restricted model without early parental reading. 95% confidence intervals in parenthesis computed from delta method standard errors.

Table 6: GMM coefficients of the potential endogenous variables - $R_{4/5}$, P and TV - under different exogeneity assumptions: Boys sub-sample

	GMM IV Estimators							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	—	$R_{4/5}$	J	TV	$R_{4/5}$, J	J, TV	$R_{4/5}$, TV	$R_{4/5}$, J, TV
TV at weekdays	-0.028**	-0.009	-0.013	-0.370*	-0.002	0.058	-0.410**	-0.322
Reads for enjoyment	0.112***	0.076***	0.624***	0.083***	0.466***	0.678***	0.062**	0.199
R4/5 (3-5 days)	-0.003	0.411***	-0.022	-0.061	0.354***	-0.005	0.250***	0.274***
R4/5 (6-7 days)	0.035	0.536***	-0.040	-0.024	0.411***	-0.026	0.393***	0.413***
R^2	0.52	0.41	0.38	0.34	0.39	0.36	0.32	0.36
Overidentifying Restrictions and Hausmann Test p-values[(1)]								
$R_{4/5}$	0	0.326	—	—	—	—	—	—
J	0	—	0.006	—	—	—	—	—
TV	0.082	—	—	0.004	—	—	—	—
$R_{4/5}$, J	0	0.005	0.001	—	0.730	—	—	—
J, TV	0.085	—	0.455	—	—	0.041	—	—
$R_{4/5}$, TV	0	0.002	—	0	—	0	0.773	—
$R_{4/5}$, J, TV	0.001	0.021	0.017	0.001	0.048	0.002	0.248	0.594
C-Test (p-values)	0	0.927	—	—	0.253	—	0.763	0.424
I&R Strong Identification test (p-values)								
LOOCV Trace		0.917	0.960	0.996	0.994	0.971	0.998	0.896
LOOCV Det		0.424	0.966	0.989	0.313	0.305	0.710	0.140

$R_{4/5}$, J, TV represent respectively, early parental reading, reading for enjoyment and TV watched during weekdays. Significance levels: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. All GMM IV estimators include a set of socio-demographic variables and other controls. All GMM IV estimators are instrumented by the predicted values from the nonlinear regressions of the endogenous variable and additional(s) instrumental variables to generate overidentifying restrictions.
[(1)] Diagonal and off-diagonal entries show, respectively, p-values from Qui-squared distributions for the OIR Hansen's J-test and DWH test statistics, with degrees of freedom given by, respectively, the number of overidentifying restrictions and the number of regressors tested for endogeneity.
The shrinkage parameter in the Inoue and Rossi (2011) Strong Identification (I&R) was computed from the Leave One Out Cross-Validation (LOOCV) procedure with both the trace and determinant criteria.

Table 7: GMM coefficients of the potential endogenous variables - $R_{4/5}$, P and TV - under different exogeneity assumptions: Girls sub-sample

	GMM IV Estimators							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	—	$R_{4/5}$	J	TV	$R_{4/5}$, J	J, TV	$R_{4/5}$, TV	$R_{4/5}$, J, TV
TV at weekdays	-0.013	0	0.010	-0.373***	0.006	-0.284***	-0.079	0.067
Reads for enjoyment	0.118***	0.114***	0.689***	0.081***	0.326***	0.494***	0.099***	0.302***
R4/5 (3-5 days)	0.004	0.252***	-0.006	-0.013	0.184**	-0.027	0.226**	0.223**
R4/5 (6-7 days)	0.065**	0.423***	0.055*	0.002	0.320***	0	0.374***	0.384**
R^2	0.51	0.44	0.36	0.31	0.44	0.35	0.44	0.42
Overidentifying Restrictions and Hausmann Test p-values[(1)]								
$R_{4/5}$	0	0.838	—	—	—	—	—	—
J	0	—	0.721	—	—	—	—	—
TV	0.004	—	—	0.196	—	—	—	—
$R_{4/5}$, J	0	0.010	0.039	—	0.545	—	—	—
J, TV	0	—	0.001	0	—	0.814	—	—
$R_{4/5}$, TV	0.001	0.204	—	0.017	—	—	0.802	—
$R_{4/5}$, J, TV	0.003	0.095	0.065	0.002	0.317	0.007	0.011	0.426
C-Test (p-values)	—	0.833	—	—	0.606	—	0.889	0.930
I&R Strong Identification test (p-values)								
LOOCV Trace		0.950	0.973	0.901	0.850	0.977	0.700	0.538
LOOCV Det		0.671	0.846	0.712	0.889	0.987	0.809	0.812

$R_{4/5}$, J, TV represent respectively, early parental reading, reading for enjoyment and TV watched during weekdays. Significance levels: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. All GMM IV estimators include a set of socio-demographic variables and other controls. All GMM IV estimators are instrumented by the predicted values from the nonlinear regressions of the endogenous variable and additional(s) instrumental variables to generate overidentifying restrictions.
[(1)] Diagonal and off-diagonal entries show, respectively, p-values from Qui-squared distributions for the OIR Hansen's J-test and DWH test statistics, with degrees of freedom given by, respectively, the number of overidentifying restrictions and the number of regressors tested for endogeneity.
The shrinkage parameter in the Inoue and Rossi (2011) Strong Identification (I&R) was computed from the Leave One Out Cross-Validation (LOOCV) procedure with both the trace and determinant criteria.

Table 8: First stage nonlinear regressions

	Boys				Girls			
	Parental Reading		TV Weekdays		Parental Reading		Reads for enjoyment	
	Ordered Probit		Fractional		Ordered probit		Probit	
	Exogenous variables							
Constant	—	—	0.899	(0.587)	—	—	-0.478	(1.012)
Age	-0.118	(0.072)	-0.025	(0.046)	-0.019	(0.073)	-0.006	(0.078)
Extra tutoring	0.105	(0.110)	-0.020	(0.080)	-0.090	(0.121)	-0.122	(0.125)
PPVT	0.072	(0.058)	-0.081	(0.037)**	0.144	(0.069)**	0.097	(0.073)
Poor health	0.008	(0.252)	0.076	(0.153)	-0.961	(0.396)**	0.778	(0.424)*
Likes sports activities	0.058	(0.036)	-0.037	(0.023)	0.017	(0.035)	-0.025	(0.038)
Non-english speaking	-0.283	(0.115)**	-0.087	(0.069)	-0.413	(0.115)***	0.034	(0.125)
Children's books home	0.190	(0.049)***	-0.097	(0.030)***	0.239	(0.054)***	0.173	(0.059)***
TV in bedroom	-0.196	(0.083)**	0.045	(0.053)	-0.129	(0.079)	-0.052	(0.087)
TV at weekends	-0.008	(0.038)	0.317	(0.029)***	-0.068	(0.041)*	-0.040	(0.042)
Internet access	0.130	(0.255)	-0.161	(0.287)	-0.629	(0.282)**	0.249	(0.321)
Household income	0.101	(0.042)**	-0.008	(0.039)	0.100	(0.045)**	-0.096	(0.064)
Education Parent 1	0.156	(0.047)***	-0.080	(0.029)***	0.122	(0.043)***	0.030	(0.047)
Education Parent 2	0.104	(0.039)***	-0.054	(0.024)**	0.066	(0.037)*	0.014	(0.039)
Age Parent 1	0.112	(0.097)	0.010	(0.060)	0.262	(0.100)***	-0.162	(0.110)
Age Parent 2	0.144	(0.083)*	0	(0.051)	0.148	(0.084)*	0.222	(0.090)**
Reads for enjoyment	0.230	(0.070)***	-0.041	(0.044)	—	—	—	—
TV at weekdays	—	—	—	—	-0.098	(0.045)**	-0.082	(0.048)*
	Instrumental variables							
Oldest child	0.346	(0.083)***	-0.069	(0.054)	0.386	(0.082)***	0.211	(0.092)**
No. siblings	-0.081	(0.044)*	-0.041	(0.028)	-0.069	(0.042)*	-0.039	(0.049)
Visit library 1/3 periods	0.217	(0.089)**	0.036	(0.053)	0.159	(0.086)*	0.191	(0.096)**
Visit library 2/3 periods	0.290	(0.093)***	0.034	(0.062)	0.414	(0.093)***	0.255	(0.102)**
Visit library 3/3 periods	0.486	(0.107)***	0.030	(0.07)	0.739	(0.125)***	0.194	(0.119)
Arts activities	0.218	(0.074)***	-0.064	(0.048)	0.152	(0.075)**	0.152	(0.081)*
Sports activities	0.186	(0.072)**	-0.063	(0.046)	0.140	(0.072)*	0.134	(0.08)*
μ_1	3.020	(0.848)***	—	—	3.065	(0.926)***	—	—
μ_2	3.987	(0.853)***	—	—	3.977	(0.928)***	—	—

Robust standard errors in parenthesis. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Parameters, μ_1, μ_2 are the thresholds of the Ordered Probit model.

Table 9: Two step GMM and GMM-CUE estimates

	Full specification		Without Parental Reading		Without Reads for enjoyment	
	2SGMM	GMMCUE	2SGMM	GMMCUE	2SGMM	GMMCUE
			Boys			
TV at weekdays	-0.410**	-0.416**	-0.106	-0.108	—	—
Reads for enjoyment	0.062**	0.062*	0.653***	0.730***	—	—
Read to 4/5 (3-5 days)	0.250***	0.264***	—	—	—	—
Read to 4/5 (6-7 days)	0.393***	0.414***	—	—	—	—
			Girls			
TV at weekdays	0.006	0.006	-0.267***	-0.290***	-0.200	-0.202
Reads for enjoyment	0.326***	0.328***	0.552***	0.581***	—	—
Read to 4/5 (3-5 days)	0.184**	0.194**	—	—	0.197**	0.199**
Read to 4/5 (6-7 days)	0.320***	0.334***	—	—	0.318**	0.320**

Robust standard errors in parenthesis. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 10: Boys IV estimates: Endogenous variables

	TV at weekday		Read to 4/5 (3-5 days)		Read to 4/5 (6-7 days)	
Instr. Variables	Ineff	Eff	Ineff	Eff	Ineff	Eff
	—	-0.587**	—	0.570	—	0.513***
	—	(0.238)	—	(0.238)	—	(0.025)
Oldest child	—	-0.418**	—	0.310***	—	0.461***
	—	(0.171)	—	(0.171)	—	(0.026)
No siblings	—	-0.411	—	0.269**	—	0.442***
	—	(0.252)	—	(0.252)	—	(0.038)
Visit library	—	-0.458	—	0.437	—	0.485***
	—	(0.279)	—	(0.279)	—	(0.03)
Oldest child; No siblings	-0.353	-0.426**	0.136	0.275***	0.228	0.439***
	(0.345)	(0.207)	(0.345)	(0.207)	(0.038)	(0.033)
Oldest child; Visit lib.	0.120	-0.392**	0.353**	0.290***	0.441***	0.432***
	(0.517)	(0.162)	(0.517)	(0.162)	(0.046)	(0.026)
Visit library; No siblings	-0.395	-0.384*	0.275	0.237**	0.447	0.38***
	(1.479)	(0.228)	(1.479)	(0.228)	(0.133)	(0.034)
Oldest; No sibl.;	-0.251	-0.410**	0.222*	0.250***	0.340*	0.393***
Visit library	(0.369)	(0.189)	(0.369)	(0.189)	(0.045)	(0.03)

Visit library instrumental variables are the vector of indicators: Visit library 1/3 periods, Visit library 2/3 periods and,Visit library 3/3 periods.
Robust standard errors in parenthesis. Significance levels: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.
Inefficient (Ineff) and Efficient (Eff) estimator use as instruments, respectively, none and the fitted values, $\hat{p}_{TV}$, $\hat{p}_{R_{4/5}}$ and $\hat{p}_J$.

Table 11: Girls IV estimates: Endogenous variables

	Reads for enjoyment		Read to 4/5 (3-5 days)		Read to 4/5 (6-7 days)	
Instr. Variables	Ineff	Eff	Ineff	Eff	Ineff	Eff
	—	0.394**	—	0.361*	—	0.336**
	—	(0.168)	—	(0.014)	—	(0.168)
Oldest child	—	0.305**	—	0.184	—	0.338**
	—	(0.149)	—	(0.014)	—	(0.149)
No siblings	—	0.338*	—	0.183*	—	0.312**
	—	(0.175)	—	(0.015)	—	(0.175)
Visit library	—	0.449***	—	0.323	—	0.285**
	—	(0.160)	—	(0.014)	—	(0.160)
Oldest child; No siblings	0.252	0.284**	0.277**	0.208**	0.456**	0.357***
	(0.214)	(0.122)	(0.015)	(0.014)	(0.214)	(0.122)
Oldest child; Visit library	0.082	0.366**	0.317	0.138	0.500	0.282**
	(0.628)	(0.146)	(0.016)	(0.014)	(0.628)	(0.146)
Visit library; No siblings	0.132	0.400***	0.221	0.154*	0.371	0.266**
	(0.455)	(0.148)	(0.015)	(0.013)	(0.455)	(0.148)
Oldest child; No siblings;	0.198	0.326***	0.249**	0.184**	0.418***	0.320***
Visit lib.	(0.196)	(0.117)	(0.014)	(0.013)	(0.196)	(0.117)

Visit library instrumental variables are the vector of indicators: Visit library 1/3 periods, Visit library 2/3 periods and,Visit library 3/3 periods.
Robust standard errors in parenthesis. Significance levels: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.
Inefficient (Ineff) and Efficient (Eff) estimator use as instruments, respectively, none and both fitted values, $\hat{p}_{R_{4/5}}$ and $\hat{p}_J$.

Figure 1: Frequency of early human capital investment activities and of reading for enjoyment for boys (blue) and girls (pink).

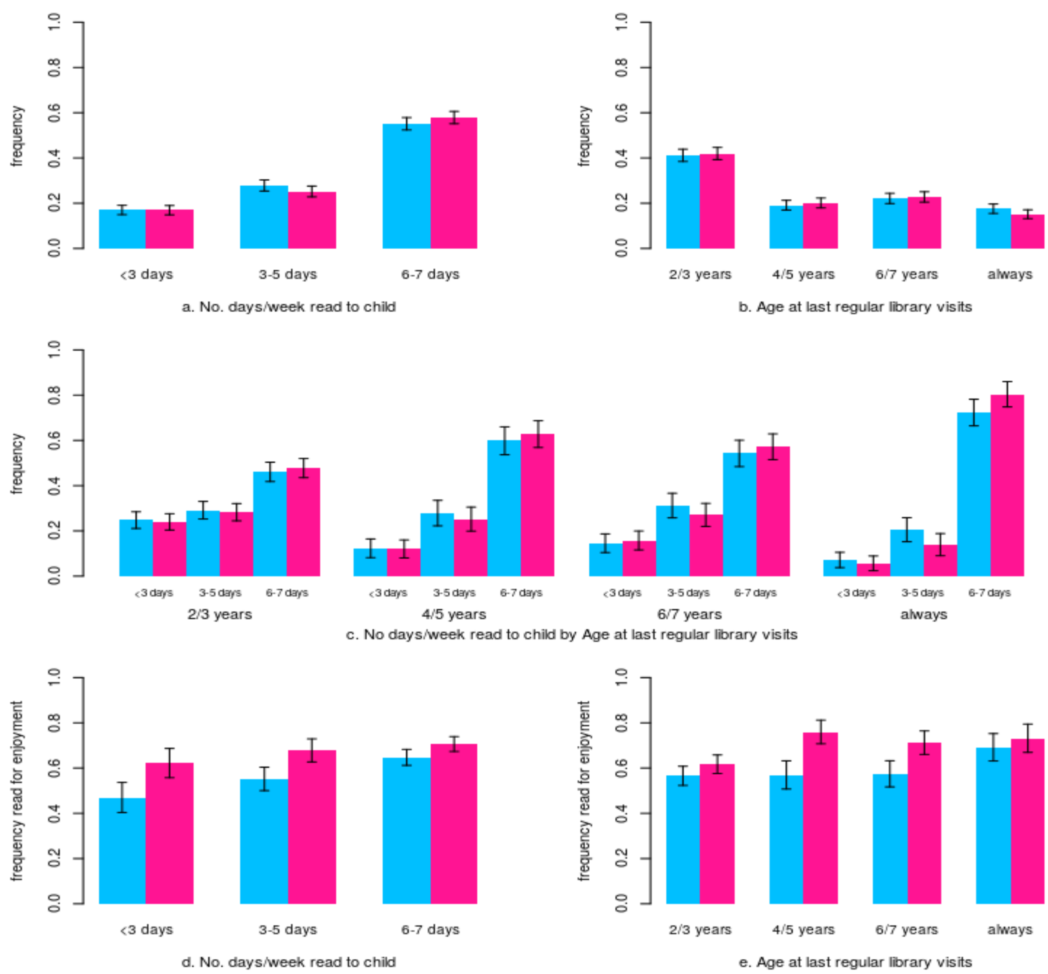
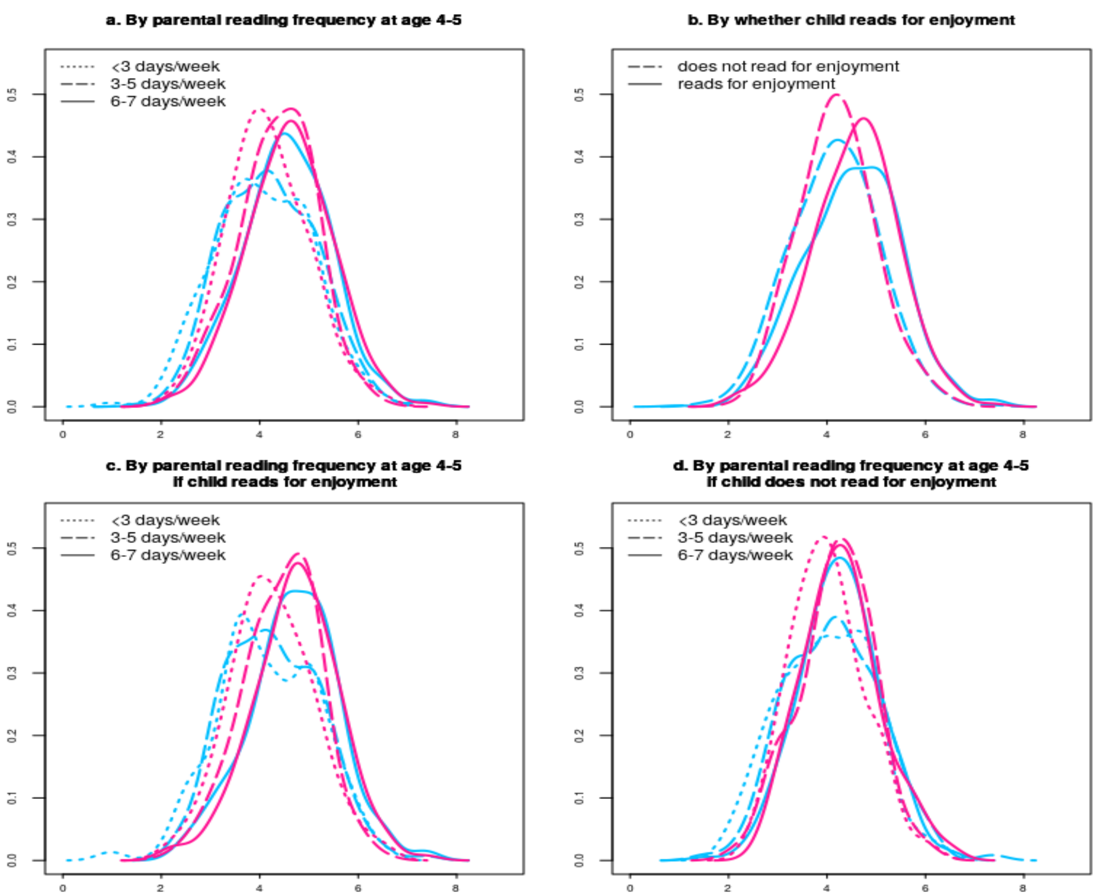


Figure 2: NAPLAN reading score at age 8-9 density for boys (blue) and girls (pink).



Appendix A. The LSAC dataset: Descriptive statistics of the variables in the model

LSAC is a longitudinal study on children in Australia jointly conducted by the Department of Social Services (DSS), the Australian Institute of Family Studies (AIFS) and the Australian Bureau of Statistics (ABS). LSAC collects information on family and social issues, children's development and well being. In particular, this survey includes information on children's physical and mental health, education and social, cognitive and emotional development. The study aims at understanding "child development, inform social policy debate and identify opportunities for intervention and prevention strategies in policy areas concerning children and their families". [21] The unit of observation in LSAC is the study child. It collects information from different sources: children (study child), parents, child carers, pre-school and schoolteachers. The first data collection was carried out in 2004 and the survey is a biennial cohort-based panel dataset.

The Birth (B) cohort, the one we use, participating children were born in a 12-month period from March 2003 to February 2004 (in the first wave children are aged between 0 and 1) and around 4000 children are followed. The sample is drawn in two stages: first a number of postcodes was randomly selected and then children were randomly selected within the relevant postcodes[22]. Stratification was used in order to ensure proportional geographic representation for states/territories and capital city statistical division/rest of state areas[23].

[21] https://growingupinaustralia.gov.au/data-and-documentation/data-user-guide and Kalb and Ours (2014).
[22] A few remote areas were excluded from the first stage of the sampling design.
[23] More information can be found at http://data.growingupinaustralia.gov.au/data/docs/userguide/11-surveymethodology.html

Table A.1: Means of variables in the regressions: cohort B

Variables	Description of the variable	Boys	Girls
N	Sample size	1275	1282
NAPLAN	Child's NAPLAN score in third grade (scale 0-10)	4.29	4.39
PPVT	PPVT computer test score divided by 10 (scale 0-10)	6.48	6.57
Age	Child's Age(8-9 years)	8.39	8.39
Poor health	Child's current health state (=1 fair or poor)	0.02	0.02
Likes sports activities	How much child enjoys physical activity or exercise (scale 1-5)	4.40	4.35
Non-english speaking	Child speaks a language other than English at home (=1 if yes)	0.36	0.36
Children's books home	Number of children books child has at home including library books (scaled to 0-4)	3.62	3.74
TV in bedroom	Child has capacity to watch TV in the bedroom (=1 if yes)	0.26	0.26
TV at weekends	Number of hours on a typical weekend child watches TV programs or movies at home (1-5 hours)	3.34	3.33
TV at weekdays	Number of hours on a typical weekday child watches TV programs or movies at home (1-5 hours)	2.77	2.70
Internet access	Internet access at home (=1 if yes)	0.96	0.97
Household income	Total usual household income before taxes (log monetary units)	7.62	7.63
Education parent 1	Highest year of primary or secondary school completed by Parent 1 (scaled to 0-5)	4.47	4.47
Education parent 2	Highest year of primary or secondary school completed by Parent 2 (scaled to 0-5)	4.30	4.26
Age parent 1	Age of Parent 1 (age/10)	3.96	3.98
Age parent 2	Age of Parent 2 (age/10)	0.11	0.10
Extra tutoring	Whether child participated regularly in e.g. remedial reading or extra tutoring in the last 12 months (=1 if yes)	0.11	0.10
Reads for enjoyment	Whether child enjoys reading at home that is not part of school work (=1 if yes)	0.56	0.67
Parental Reading	Number of days the parents or other adult at home read a book to the child at the age of 4-5 (0-7 days)	2.30	2.36
No Siblings	Number of siblings living in the household	1.50	1.48
Oldest child	Whether he/she is the oldest child (=1 if yes)	0.42	0.41
Library until 2/3	Last regular visits to library at age 2/3 or never (=1 if yes)	0.41	0.42
Library until 4/5	Last regular visits to library at age 4/5 (=1 if yes)	0.19	0.20
Library until 6/7	Last regular visits to library at age 6/7 but not always (=1 if yes)	0.22	0.23
Library from 2/3 to 6/7	Regular visits to library from 2/3 to 6/7 (=1 if yes)	0.16	0.15
Sports activities	Whether child participates frequently in sports activities outside school (=1 if yes)	0.62	0.61
Arts activities	Whether child participates frequently in arts activities outside school (=1 if yes)	0.35	0.61

Appendix B. GMM IV estimation of the average treatment effect in exponential and fractional regression models with categorical treatment and endogenous controls

Let y be the outcome variable satisfying $0 \leq y < 1$ and define $T(y) = y/(1-y)$[24]. If y_g with $g = 0, 1, 2$ are, respectively, the potential outcomes without and with two intensity levels of the treatment $d_g = 0, 1$, the observed transformed outcome for a fractional variable can be written in the exponential switch regression form as

$$T(y) = \exp\left(\mu_0 + \sum_{g=1}^{2}(\mu_g - \mu_0)d_g + v_0 + \sum_{g=1}^{2}(v_g - v_0)d_g\right) \tag{B.1}$$

where $\mu_g = E(T(y_g))$ and $v_g = T(y_g) - \mu_g$, $g = 1, 2$. To proceed with IV estimation as in Wooldridge (2010) chapter 21, when x_1 and x_2 are, respectively, $(k_1 \times 1)$ and $(k_2 \times 1)$ vectors of endogenous and exogenous variables, z is a $(l \times 1)$ vector of instruments for both d_g and x_1, with $l \geq (k_1 + 2)$, the following assumptions are required:

Assumption 1. *The individual specific treatment effects are homogeneous, i.e.,:*

$v_g = v_0, \quad g = 1, 2.$

Assumption 2. *Let $L(v_0|x_1, x_2)$ represent the linear projection of v_0 on x_1 and x_2 with an intercept, and define $\eta_0 \equiv e^{v_0 - L(v_0|x_1,x_2)}$. The conditional expectation of η_0 satisfies:* **(a)** $E(\eta_0|x_1, x_2, z) = E(\eta_0|x_1, x_2) \neq E(\eta_0|x_2)$; **(b)** $E(\eta_0 z|\mathbf{d}, x_1, x_2) = 0$, $\mathbf{d} = (d_1, d_2)$ *and;* **(c)** $E(\eta_0|x_2, z) = 1$.

Assumption 3. *The conditional probability of the treatment indicators d_g can be expressed as known parametric function of x_2 and z:* **(a)** $P(d_g = 1|x_2, z) = H_g(x_2, z, \gamma_d)$; **(b)** $P(d_g = 1|x_2, z) \neq P(d_g = 1|x_2)$.

Assumption 1 entails a functional form restriction. If it does not hold, IV estimation of the average treatment effect will not be consistent, since the last term in equation (B.1) becomes $\sum_{g=1}^{2} h_g(x)d_g$ where $h_g(x) = v_g - v_0$. This introduces an unknown form of nonlinearity that can be tested with a standard Ramsey's RESET test.[25] However, if $h_g(x)$ can be arbitrarily approximated by polynomials in $x\prime\hat{\beta}$, a restricted version of the RESET test can be implemented by testing the significance of $d_g(x\prime\hat{\beta})^p$, $g = 1, 2$ in an augmented regression, with a standard qui-squared distributed Wald test. The test will be referred to as ATESET (Regression Specification Average Treatment Effect) since under the null hypothesis, the GMM IV estimator consistently estimates the average treatment effect. Since x includes endogenous regressors IV estimation of the augmented regression needs to account for the additional endogenous regressors or alternatively use the forms of the test suggested in Pagan and Hall (1983) or in Pesaran and Taylor (1999).

[24] These methods also apply for non-negative outcomes in which case $T(y) = y$ and an exponential regression model applies

[25] Wooldridge (2010) synthesis estimation results in Angrist and Krueger (1991), Heckman (1997), and Wooldridge (1997) that impose additional functional form restrictions in the form of functions of the regressors.

The first equality in assumption 2.a entails an exclusion restriction for z, while the second allows for x_1 to be endogenous once v_0 is replaced by its linear projection. Assumption 2.b is a consequence of z being an instrumental variable, and therefore, as good as randomly assigned, i.e., each z_j is independent of potential outcomes, conditional on covariates (see Angrist and Pischke (2009) chapter 4). It follows that the product $d_g z_j$, can be added to the set of instrumental variables since

$$Cov(z_j d_g, \eta_0) = E(\eta_0 z_j d_g) = E[d_g E(\eta_0 z_j | \mathbf{d}, x_1, x_2)] = 0 \quad \text{(B.2)}$$

Thus, each instrumental variable for the categorical treatment indicator can be used to derive an additional valid instrument. Condition B.2 can be empirically tested with a C-test (see Ruud (2000), Hayashi (2000) and Eichenbaum et al. (1988)) addressing the validity of the orthogonality conditions defined by these product instruments.

Similarly to the Wooldridge (2010) procedure, assumption 3.a allows the fitted probabilities, $\hat{H}_{g,i} \equiv H_g(x_{2,i}, z_i, \hat{\gamma}_d)$ obtained from a qualitative ordered dependent variable model first stage regressions, to be used as instrumental variables for d_g. This instrumental variables deliver consistent estimates even if those regressions are misspecified.[26] Assumption 3.b ensures that the instrumental variables z are relevant. However, that condition is not necessary for identification of the IV estimator if nonlinearity of the first stage regressions can be exploited to provide a form of functional form identification.

Let z^* be the extended set of instrumental variables obtained by adding to z the instruments implied by assumptions 2.b and 3.a. Under assumption 2.c the moment condition

$$E\left[1 - T(y) \exp\left(-\delta_0 - \sum_{g=1}^{2} \tau_g d_g - x_1' b_1 - x_2' b_2 \right) \bigg| x_2, z^* \right] = 0 \quad \text{(B.3)}$$

defines the GMM IV estimator proposed in Ramalho and Ramalho (2017). It delivers consistent estimates, under the presence of random effects type of unobserved heterogeneity, when both the treatment indicator $\mathbf{d}$ and x_1 are endogenous.

[26] A similar argument applies to the elements of x_1 whose conditional expectations are appropriately represented by nonlinear regression functions (e.g. if the variables are binary, categorical, discrete, bounded in etc.)