Does it matter what children read? New evidence using longitudinal census data from Spain

John Jerrim[1]

Luis Alejandro Lopez-Agudo[2]

Oscar D. Marcenaro-Gutierrez[3]

Abstract

It has long been thought that encouraging children to read is likely to be beneficial for the development of their literacy skills. However, a lot less attention has been paid to the issue of whether what students read matters for their academic progress. This paper therefore considers the association between the frequency young people read five different types of text (comics, short stories, books, newspapers and magazines) and their scores on standardised reading and mathematics tests. Drawing upon large longitudinal census data from the largest administrative region in Spain, we find that frequency of reading comics, newspapers and magazines is not associated to the development of children's cognitive skills. In contrast, there is clear and consistent evidence of a positive and increasing association between the frequency children read books and their academic achievement. We consequently conclude that recommended reading time for children should be focused upon the time they spend reading books and not other material.

Keywords: books; short novels; reading skills.

JEL Codes: I20, I21, I28, C10.

Acknowledgements: This work has been partly supported by the *Ministerio de Economía, Industria y Competitividad* under Research Project ECO2017-88883-R; the *Consejería de Innovación, Ciencia y Empresa de la Junta de Andalucía* under PAI group SEJ-532 and the postdoctoral contract from the *Plan Propio* signed by the *Universidad de Málaga*.

[1] Corresponding author. Department of Social Science, UCL Institute of Education, University College London, 20 Bedford Way, WC1H 0AL, London. E-mail: j.jerrim@ucl.ac.uk. Tel.: +44 02076126977. ORCID: 0000-0001-5705-7954

[2] Departamento de Economía Aplicada (Estadística y Econometría). Facultad de Ciencias Económicas y Empresariales. Universidad de Málaga. Plaza de El Ejido s/n, 29013, Málaga (España). E-mail: lopezagudo@uma.es. Tel.: +34 952137003. ORCID: 0000-0002-0906-3206

[3] Departamento de Economía Aplicada (Estadística y Econometría). Facultad de Ciencias Económicas y Empresariales. Universidad de Málaga. Plaza de El Ejido s/n, 29013, Málaga (España). E-mail: odmarcenaro@uma.es. Tel.: +34 952137003. ORCID: 0000-0003-0939-5064

1. Introduction

Reading is a fundamental skill that plays a key part in all our lives. Indeed, the ability to fluently read is critical to an individual being able to play an active role in modern society and is a prerequisite of labour market success. It is hence one of the main skills that children are taught to finesse at school, building upon the early literacy foundations initially laid by parents at home. It is generally expected that the main way that young people develop stronger reading skills is through practise and with engagement with increasingly longer and more challenging texts. This is supported by an extensive empirical academic literature (Kragler, 2000; Krashen, 2004; Gillespie, 2010; Shanahan, Fisher, & Frey, 2012; Westbrook, Sutherland, Oakhill, & Sullivan, 2018), which has consistently found a positive association between the frequency young people read and their scores on standardised achievement tests.

A broad consensus has thus emerged that reading is important for children's development. But does it also matter what type of text children choose to read, with engagement with certain types of text (e.g. books or novels) offering greater benefits than others (e.g. magazines or newspapers)? There are several reasons why one might suspect this to be the case. First, books are often formed of more challenging texts, increasing young people's exposure to new vocabulary and sophisticated sentence structures. The same is often not true of newspapers and magazines, which are often purposefully designed to be easily accessible amongst a wide audience. Second, books tend to require deep engagement and concentration over a sustained period of time (Moss & McDonald, 2004). Other types of texts, although they may be read just as frequently, do not take such a long time to complete and can be dipped in and out of quite superficially. Third, there is some evidence that the reading of stories can influence social cognition (e.g. Mar, 2018) and thus help children to better understand the meaning of a given text. Fourth, previous research has suggested that "deep reading" is key for children's literacy development (Wolf, 2018). Certain types of reading material (e.g. books) are likely to require greater depth of reading than others (e.g. comics) and hence having different influences upon children's cognitive outcomes. Fifth, sociological research has suggested that reading is a particular form of "cultural capital" (Bourdieu, Passeron, & Saint-Martin, 1994; Wright, 2006), with those children who demonstrate cultural attributes being rewarded by teachers with higher attainment. Hence, while books are likely to be regarded as a sophisticated type of cultural capital by those in the education system, comics and magazines are not. Finally, tests and examinations involve reading, interpreting and understanding different questions. Regular engagement with complex reading material may hence help children decipher the meaning of test questions, with Bourdieu, Passeron, and Saint-Martin (1994) noting "*Obvious in the literary disciplines, but more subtle in the sciences, the ability to manipulate academic language remains the principal factor in success in examinations*" (p. 21). If books help children to understand the formal language often used in education settings, then this may translate into superior performance in standardised tests.

Yet, despite the potential importance of this issue, few existing studies have investigated the association between how frequently young people read different text types and their literacy skills. There are, however, a handful of notable exceptions. For instance, Brozo, Shiel and Topping (2007) used data from the Programme for International Student Assessment (PISA) to show that students who read a diverse range of texts achieve higher scores on this cognitive test than those who do not. In the case of adult literacy skills, Smith (1996) found that reading a wide variety of texts (i.e., books, magazines, newspapers and six types of personal and work documents) is positively associated with higher literacy skills. However, Anderson, Wilson, and Fielding (1988) found reading books and comics to be positively associated with reading comprehension, while the with reading mail, newspapers

and magazines was negative. Similar results were found by Birr, Overby, Tysvaer, and Morris (2008), who found reading novels, short stories, picture books and plays was positively associated with students' achievement, while reading music lyrics was negatively associated with it. Mullis, Martin, Kennedy, and Foy (2007) analysed primary education students' literacy skills in PIRLS 2006 for 40 countries and found that, in most countries, students who reported reading novels and short stories more frequently performed higher than those who reported a lower frequency. They also found that reading informational texts more frequently (e.g. magazines, newspapers, directions or instructions and brochures or catalogues) did not have a clear relationship with reading achievement. Likewise, Duncan, McGeown, Griffiths, Stothard, and Dobai (2015) found that only reading fiction books was positively associated with higher reading skills, while reading magazines, newspapers, song lyrics, non-fiction books, school texts and other digital literacies were not. Hence, in summary, although most studies in the existing literature find a positive association between reading books and children's literacy skills, evidence with respect to other text types is more mixed.

A recent paper by Jerrim and Moss (2018) also explored this issue and forms the basis for the present investigation. These authors used cross-sectional data from the 2009 round of the Programme for International Student Assessment (PISA) to investigate how the frequency of 15-year-olds reading five different types of material (comics, magazines, newspapers, fiction books, non-fiction books) was linked to their functional literacy skills. Analysing data from more than 30 developed countries, they found evidence of a "fiction effect" (i.e. a sizeable association between young people reading fiction books and their scores on the PISA reading test) but little association between PISA reading scores and the frequency children read other types of texts. However, the authors noted some important limitations to their study (and much of the existing literature), most notably the cross-sectional nature of the data, the lack of high-quality controls for prior achievement and the limited information available about parental reading habits (which could confound the results). They consequently concluded that further work in this area was needed, in order to better understand whether young people's cognitive skills are particularly developed by reading specific types of material.

We take up this challenge in this paper, building directly upon the existing literature outlined above. Specifically, we use large-scale longitudinal administrative data from the largest region in Spain (Andalusia) to further explore how teenagers' reading of different material is related to the development of their cognitive skills. Our data has several advantages compared to the previous literature, including rich measures of prior achievement and a wide array of background measures that have been collected from both parents and children. Besides the usual demographic characteristics (e.g. gender, parental age and socio-economic status), the latter includes information capturing parental reading attitudes and activities, parental engagement towards their child's education (e.g. future expectations and their interaction with the child about homework and school) and children's attitudes and engagement towards school (e.g. future expectations and school enjoyment), all of which were potential unobserved confounders in previous research. Critically, the data we analyse are also longitudinal, enabling us to estimate how reading different forms of text is associated with the academic ***progress*** children make over a three-year period. Hence the primary research question we address is:

Research Question 1. *How does reading different types of material relate to the progress children make in their literacy skills between ages 10-11 and 13-14?*

Importantly, we are also able to investigate potential "spill over" effects into other subjects (most notably mathematics) which has received less attention in prior work. Why might such spill-over effects exist? There are at least three possible explanations. First, as noted by Sullivan and Brown (2015), the ability to read fluently may be important for learning in other subjects. For instance, children who are more fluent readers may be able to better navigate instructions within mathematics textbooks or by their teachers, particularly within certain mathematical tasks that require a degree of reading comprehension. Second, relatedly, reading may be an important skill in answering certain mathematics test questions – particularly those that require the application of mathematics knowledge within real-world situations. As the eighth-grade outcome measure we analyse is designed to measure real-world mathematical competencies it would not be surprising for such spill-over effects to emerge upon this particular test. Finally, previous research using data from identical twins has linked reading to improvements in general intelligence (Ritchie, Bates, & Plomin, 2014). If this relationship between reading and general intelligence is indeed causal, then it provides a clear pathway via which reading certain types of text would be associated with mathematics achievement. Consequently, our secondary research question is:

Research question 2. *Is there an association between the type of texts children read and the progress they make in mathematics achievement between ages 10-11 and 13-14?*

To preview our findings, our substantive results replicate many of those presented within Jerrim and Moss (2018). Specifically, we found no evidence that frequency of reading comics, newspapers or magazines is positively associated with young people's cognitive development. Rather, it is only the reading of more complex forms of texts (i.e. books) that is associated with academic progress. In particular, children who read books every or almost every day score 0.22 standard deviations higher on the eighth grade (age 13-14) literacy test than those who read books "almost never". There is also evidence of positive spill-overs into other subjects, with a difference of around 0.20 standard deviations in mathematics. Critically, these results refer to when a wide array of potential confounders have been controlled. We consequently conclude that children's reading time should be devoted to books and that spells looking at other material (e.g. newspapers and magazines) should not count towards their recommended reading time.

The paper now proceeds as follows. Section 2 describes the data we analyse with an overview of our empirical methodology following in section 3. Results are presented in section 4, with conclusions and directions for future research following in section 5.

2. Data

The data we use in this paper are drawn from an administrative census from the largest region within Spain (Andalusia). This is an interesting educational context for our empirical analysis as it is a relatively deprived European region with many children having poor literacy skills (Rodríguez, Álvarez, & Moreno, 2009). For instance, results from the PISA 2015 study (OECD, 2016) illustrated how almost a quarter of Andalusian 15-year-olds do not have basic functional literacy skills (defined at reaching at least level 2 in PISA), compared to a Spanish average of 16 percent and an OECD average of 20 percent. Further details about the Andalusian education system, in comparison to the whole of Spain and the average OECD economy, can be found in Table 1.

<< Table 1 >>

The data we use are based upon a census within Andalusia conducted by the Andalusian Agency of Education Assessment (*Agencia Andaluza de Evaluación Educativa* - AGAEVE). Specifically, each Andalusian child completes assessments (known as

Evaluación de Diagnóstico) in Spanish language and mathematics. The cohort of children our analysis refers to was in fifth grade (the penultimate year of primary school) in 2008/9 and eighth grade (the second year of secondary school) in 2011/12.

There were a total of 60,747 fifth grade Andalusian students in the 2008/09 academic year and we can follow 47,318 of them in 8th grade in the course 2011-12. Within our analysis, we make the following restrictions to this cohort:

- Children who repeated a grade before the fifth grade are excluded (1,993 students dropped)[4].
- Private school children in 8th grade are excluded as they were not included within the assessment programme in 5th grade (165 students dropped).

This leaves a total analytic sample of 45,160 pupils included within our analysis. The missing information in students' scores in Spanish language and mathematics reduce the sample to 43,604 and 43,833 students respectively[5]. A comparison between the non-repeater population[6] and our analysis sample is provided in Appendix A. This illustrates that differences between the characteristics of the non-repeater Andalusian population and the sample under analysis are relatively minor.

As part of this assessment programme, children and their parents complete background questionnaires in order to contextualise the results. Response rates to these questionnaires are reasonably high, standing at 78 percent for pupils and 82 percent for their parents. Several detailed questions were asked about children's and parents reading attitudes and activities that we detail below. A range of additional information was also collected about factors such as family background (e.g. parental education, occupation and household possessions), parental engagement in their child's schooling and children's attitudes and engagement towards school, among others. Full descriptive statistics about these other measures (which we use as statistical controls within some of our empirical models) can be found in Appendix B.

Measurement of children's reading activities

Our key covariate of interest is the frequency that children read different types of material outside of school. Within both the 5th and 8th grade survey, the background questionnaire children completed included the following question:

"Approximately, how much time do you spend, out of school, doing these activities?"

- I read comics
- I read tales or short novels[7]
- I read books[8]
- I read magazines
- I read newspapers

4 Students who have repeated a grade are different in many ways to the rest of the student population. As we do not believe it is possible to control for this wide array of factors in our analysis, we have chosen to focus upon only students who have not repeated a grade (García-Pérez, Hidalgo-Hidalgo, & Robles-Zurita, 2014).

[5] A missing flag to avoid dropping observations with missing data upon other covariates. We have tested the robustness of our findings to using multiple imputation instead, with little substantive change to our results.

[6] These population descriptive statistics do not include those students who were repeaters in the course 2008/09 to make them comparable with the sample descriptive statistics.

[7] Examples here could include the work of Edgar Allan Poe (e.g. The Gold-Bug) or H. P. Lovecraft (The Call of Cthulhu).

[8] For example, Harry Potter or The Lord of the Rings.

There were four possible response options: 1. never or almost never; 2. once or twice a month; 3. once or twice a week; 4. every or almost every day. This is very similar to the question used in the analysis by Jerrim and Moss (2018) and is the basis of our analysis. Specifically, we investigate whether the frequency of reading each of these different text types is associated with young people making greater progress in their Spanish language and mathematics skills.

Parental reading attitudes and activities

One of the key limitations of previous work on this topic (e.g. the recent paper by Jerrim & Moss, 2018) is that the PISA data analysed included relatively little information about parental reading behaviours. A key strength of the data we use is that parents were asked a range of questions about their own reading attitudes and activities. This includes the following questions:

- The amount of time parents spent reading per week.
- Whether parents only read because they have to.
- Whether parents enjoy talking about books with other people.
- Whether parents choose to spend their leisure time reading.
- Whether parents only read when they need information.
- Whether parents believe that reading is an important activity at home.

Further details on the wording of these questions and the response options provided can be found in Appendix C. We use parental responses to these questions as key control variables within our analysis.

Measurement of children's reading skills

A further limitation of previous work is that some studies only have access to cross-sectional data. Consequently, children's reading skills are only measured at a single time-point. A major advantage of our data compared to previous research is that we have access to longitudinal data, with children's reading skills measured at two time points. Specifically, at the end of the fifth and eighth grade, children within Andalusia complete standardised tests in Spanish language and mathematics. These tests took approximately 2 hours to complete (with a break of 30 minutes after the first hour) and were independently marked. Further details about the psychometric properties of these tests can be found at http://www.juntadeandalucia.es/educacion/agaeve/publicaciones-cuadernillos-ped.html. In Appendix D we provide histograms to illustrate the distribution of scores on these tests. There is no evidence of floor or ceiling effects, with both language scores broadly following a normal distribution. Throughout our analysis we standardise scores on these tests so that the mean is zero and standard deviation one (using each grade and subject population mean and standard deviation). All of the results presented can therefore be interpreted in terms of an effect size.

3. Methodology

A series of Ordinary Least Squares (OLS) regression models are estimated to investigate the association between frequency of reading different text types with teenagers' Spanish language and mathematics skills. These are of the form:

$$R_{ij}^{8} = \alpha + \beta.T_{ij}^{5} + \gamma.D_{ij}^{5} + \theta.R_{ij}^{5} + \tau.P_{ij}^{5} + \delta.Z_{ij}^{5} + \varphi.C_{ij}^{5} + \mu_{j}^{5,8} + \varepsilon_{ij} \qquad (1)$$

Where:

R_{ij}^{8} = Children's scores on the Spanish language/mathematics assessment in 8th grade (standardised using the mean and standard deviations of the total population in Spanish language/mathematics in 8th grade).

T_{ij}^{5} = The frequency with which young people read a given text type in 8th grade. This is entered as a set of dummy variables, with never/almost never as the reference group, in different model estimations.

D_{ij}^{5} = A vector of controls for demographic characteristics in 5th grade, including gender, parental age and socio-economic status[9].

R_{ij}^{5}= Children's scores on the 5th grade Spanish Language and mathematics assessment.

P_{ij}^{5} = A vector of control variables for parental reading attitudes and activities in 5th grade (as described in the previous section).

Z_{ij}^{5} = A vector of additional controls for parental engagement towards their child's education in 5th grade, including the level of education they expect them to complete, whether parents help with homework, ask if the child has homework and check the homework whether they ask how their school day was and whether parents encourage their children to study.

C_{ij}^{5} = A vector of controls for children's attitudes and engagement towards school in 5th grade, such as whether they enjoy school, whether they want to move to another school and the level of education they expect to complete.

$\mu_{j}^{5,8}$ = School fixed-effects in 5th and 8th grades.

ε_{ij} = Individual error term.

i = child i.

j = school j.

Model (1) has been estimated separately for students' Spanish language and mathematics skills, using each one of the five types of texts as an explanatory variable (T_{ij}^{5}). The coefficient of interest is β. This illustrates the change in Spanish language/mathematics scores depending upon how frequently the student reads each type of text outside of school (conditional upon the other factors controlled in the model).

Six specifications of this model are estimated (labelled M1 to M6) with each including a different set of controls. The first model (M1) only controls for the frequency with which young people read a given text type (variable T_{ij}^{5} in equation 1) and demographic characteristics (gender, parental age and socio-economic status; variables D_{ij}^{5} in equation 1). This is used to provide baseline estimates of the association between reading different text types and young people's reading/mathematics skills.

Specification M2 then includes controls for prior achievement in grade 5 in reading and mathematics (variables R_{ij}^{5} in equation 1). This is hence a "value-added" specification, with the estimated β coefficients now illustrating how the different text types are associated with the ***progress*** that children make in their reading/mathematics skills over this three-year

[9] A socio-economic status index (economic, social and cultural status, ESCS) was used. This was derived by the survey organizers via a principal component analysis, using the highest from mother's and father's education, the highest from mother's and father's occupation, the number of books at home and household possessions.

period. Our anticipation is that the inclusion of this control will lead to an appreciable reduction in the estimated association between reading different text types and young people's language/mathematics skills.

Of course, one reason why children who read certain text types more frequently make more progress in their reading/mathematics skills could be due to the home literacy environment. In particular, their parents may be more likely to read themselves. Model M3 consequently includes controls for parental reading attitudes and activities (variables P_{ij}^{5} in equation 1) as described in the previous section. Further controls are then added for parental engagement in their child's schooling (variables Z_{ij}^{5} in equation 1) in model M4 to establish whether this is confounding our results.

The penultimate model specification (M5) adds controls for children's attitudes and engagement towards school (variables C_{ij}^{5} in equation 1). This is to attempt to adjust for further potential omitted variable bias, with the intuition being that unobservable factors that are likely to mean children choose to read more are likely to be associated with those that lead children to enjoy or work hard at school. The final model (M6) then includes both primary and secondary school fixed-effects (variables $\mu_{j}^{5,8}$ in equation 1). These final estimates hence attempt to purge any remaining residual confounding from parents selecting particular schools for their children, which could be associated with their academic progress. In particular, note that estimates from this final model refer to differences in achievement between children from the same demographic background, who attend the same primary and secondary school, are similar in terms of their prior reading and mathematics achievement, attitudes and engagement towards school and whose parents are similarly engaged in their schooling and have similar reading activities. Although it is prudent to continue to treat such estimates as conditional associations only (rather than causal estimates) we nevertheless argue that most of the key factors likely to confound the relationship between children's reading habits and their academic achievement have been controlled.

To conclude this section, Table 2 provides some descriptive information of how frequently different genders and socio-economic groups read different text types. Girls seem to read short stories, books and newspapers more frequently than boys, while the opposite holds true for comics and magazines. In the case of students' socio-economic status (ESCS), young people from advantaged backgrounds read all the text types more frequently than those from disadvantaged homes. Finally, high-achieving students (according to their 5th grade test scores) were more likely to read tales/short novels and books compared to low-achieving students, though with little difference in terms of reading comics, newspapers and magazines.

<< **Table 2** >>

4. Results

Reading scores

Table 3 provides an overview of results with respect to the association between frequency of students reading the different text types and the age 13-14 reading test scores. These focus upon the contrast between children who read each type of material regularly (i.e. every or almost every day) in comparison to those who read the type of material hardly at all (i.e. never/almost never)[10]. As our outcome measure has been standardised, all estimates presented can be interpreted in terms of an effect size.

[10] A full set of results, including all parameter estimates for all covariates, are available upon request.

<< Table 3 >>

Starting with the results for comics, there is strong and consistent evidence of a null effect. In all six model specifications the estimated effect is basically zero and, despite the very large sample size, never statistically significant at conventional thresholds. This strongly suggests that frequently reading this type of text is unlikely to benefit young people in terms of their reading academic achievement.

Interestingly, a similar result emerges for the frequency young people read newspapers and magazines. Even in our basic model specification (M1), which controls only for basic demographic characteristics, effect sizes are small. For instance, reading a daily newspaper is associated with just a 0.05 increase in eighth grade literacy scores, with a negative association (-0.03 standard deviations) observed for magazines (although this reverts to zero from model M2 onwards – once prior achievement has been controlled). Consequently, we again reach the substantive conclusion that encouraging young people to read these types of text regularly is unlikely to benefit their literacy skills.

More positive findings emerge with respect to the two other types of material investigated – short stories and books. Regarding the former, after controlling for demographic characteristics, reading this type of material daily is associated with a 0.20 standard deviation increase in reading achievement (compared to not reading short stories at all). There is some evidence that prior achievement is to some extent confounding this result, with the estimated effect size falling to 0.14 standard deviations once fifth grade reading and mathematics scores have been controlled (model M2). However, the parameter estimate then remains broadly stable between models M2 and M6 after a wide array of potential additional confounders have been controlled. Consequently, estimates from our final model (M6) continue to suggest that regularly reading short stories is modestly associated with the progress young people make in their reading skills between ages 10-11 and 13-14 (compared to never reading such material at all).

Consistent with the findings of Jerrim and Moss (2018) the strongest results are obtained for the frequency young people read books. In our baseline model specification (M1) regular reading of books is associated with a 0.40 standard deviation increase in age 13-14 reading test scores. Again, the major confounder of this result is that children with higher levels of prior achievement tend to read books more often (as illustrated by Table 2). Once we have accounted for this fact within our analysis, the estimated effect size falls by around one-third, though is still substantial (effect size = 0.28). The inclusion of additional controls for parental reading attitudes and activities and engagement, children's attitudes towards school and engagement do little to change this result; the effect size in model M5 has barely changed, standing at 0.27. Moreover, although the inclusion of school fixed-effects does lead to some further attenuation of the estimated association (down to 0.22), it remains substantively important in terms of magnitude. We consequently conclude that there is stronger evidence of an association between the frequency of reading books and children's literacy skills than for other types of reading material.

Further evidence on this matter is presented in Table 4. Here we provide results for all of the "short stories" and "books" categories, giving more detail on the nature of this relationship. Regarding the results for short stories, note how the major difference is between the never/almost never categories and the other groups. In other words, for short stories, the magnitude of the effect size is very similar (always around 0.11 standard deviations in model M6) for reading this type of material monthly, weekly or daily. Consequently, there is *not* clear evidence of a "dose-response" relationship for this text type. Rather, it seems that reading this type of text at least once a month is potentially important for the development of

young people's reading skills. However, increasing the frequency of reading short stories beyond this (to monthly or daily) is unlikely to bring benefits for young people's academic achievement.

<< Table 4 >>

A slightly different story emerges for reading books. As noted in out discussion of Table 3, all parameter estimates for reading books are large in our base model (M1) and somewhat attenuated (although still large) once key confounders (most notably prior achievement) have been controlled. However, even in our final specification, a clear difference in the magnitude of the effect sizes can be observed across the different categories. In other words, for the reading of books, there is greater evidence of a dose-response relationship. Although the greatest difference continues to occur between "never" and "monthly" (effect size = 0.12) the association is also appreciably larger for the weekly (0.16) and daily (0.22) categories. Specifically, reading books daily rather than monthly is associated with a 0.10 standard deviation increase in age 13-14 literacy scores, conditional upon the wide array of factors controlled in the final model. Hence, out of the five text types we have considered, the strongest evidence of a positive effect upon children's reading skills emerges for the reading of books.

Mathematics scores

Table 5 turns to the results for mathematics, again focusing upon the contrast between children who never read each type of material compared to those who read it nearly every day[11].

<< Table 5 >>

Consistent with the results for reading, there is no evidence of a positive association between reading comics, newspapers and magazines and children's mathematics scores. Across all model specifications, effect sizes tend to be small and, on occasion, negative (most notably for magazines). This strengthens the substantive conclusion we previously reached that frequency of reading these types of material are unlikely to be offering substantial benefits for children's cognitive development.

For short stories and books, we identify a positive association with mathematics test scores, albeit with the effect sizes slightly smaller than those observed for children's literacy scores. Nevertheless, as illustrated by Table 6, they follow a similar broad pattern. Three particular features stand out.

<< Table 6 >>

First, the inclusion of controls for confounders (most notably prior achievement) leads to a non-trivial decline in the estimated effect size. However, even in our most detailed model specification (M6), the magnitude of the association between reading books and mathematics scores remains substantial. Second, we again observe evidence of a dose-response relationship between frequency of reading books and the progress young people make in mathematics. Note, for instance, how the daily reading of books is associated with a 0.11 standard deviation increase in age 13-14 mathematics scores compared to children who only read books on a monthly basis. Finally, the same is not true for the reading of short stories. While there is a clear difference in mathematics achievement between young people who never read short stories and those who read them at least once or twice a month (0.08

[11] Full parameter estimates for all model specifications and covariates are available upon request.

standard deviations), there is little evidence that increased frequency of reading this type of text (to weekly or daily) is then further associated with higher mathematics achievement. This can clearly be observed by the effect sizes for the "monthly", "weekly" and "daily" categories for short stories in Table 6 being very similar (standing between 0.08 and 0.10 standard deviations higher than the "never/almost never" group).

Heterogeneous effects

In additional analysis we have investigated whether there are potential heterogeneous associations in two ways. First, we have estimated our empirical models separately for boys, girls and for students from low/high socio-economic backgrounds. These results are summarised in Table 7, focusing upon estimates from our final model specification (M6). In general, there was little evidence of differential associations with reading test scores by either gender or socio-economic status for any of the text types. The one possible exception was for short-stories, where the effect size for students from disadvantaged socio-economic background (0.22) was twice that for students from the most advantaged background (0.10). Otherwise, the effect sizes estimated were similar across groups.

<< Table 7 >>

Second, we re-estimated our model using quantile regression to investigate whether the association was weaker or stronger for lower or higher achievers. The results for short-stories and books are presented in Figure 1[12]. Interestingly, on both occasions, effect sizes were larger towards the bottom end of the achievement distribution. In other words, frequently reading books or short-stories may be particularly important to raise the skills of low-achievers. For instance, reading books almost daily (compared to almost never) is estimated to increase the test scores at the bottom of the achievement distribution (i.e. at the 10^{th} or 20^{th} percentile) by around 0.35 standard deviations. In contrast, the effect size at the top of the reading distribution (i.e. the 80^{th} and 90^{th} percentile) is below 0.20. This reflects a clear pattern illustrated by Figure 1 where the effect size for reading books is particularly large towards the bottom end of the achievement distribution and then gradually declines as one moves towards the top. The pattern of results for short-stories is similar, though the magnitude of the difference across the distribution is not so stark. Nevertheless, Figure 1 provides some evidence that reading certain text types (most notably books) rather than others is likely to be particularly important in raising basic levels of reading proficiency.

<< Figure 1 >>

5. Conclusions

The ability to read is a key skill developed during childhood. Being a competent, fluent reader is vital for both participation in modern society (OECD, 2010) and for labour market success (Chiswick, Lee, & Miller, 2003; Quintini, 2014). One of the ways parents, teachers and educators try to develop young people's reading ability is by getting them to practise their skills (i.e. by getting them to sit down and read). But does it matter what young people choose to read for their cognitive development? Does reading comics, newspapers and magazines bring the same benefits as reading books? Or does it not matter what teenagers choose to read – as long as they read something? Unfortunately, there has been relatively few large-scale quantitative investigations into this issue. Hence we currently know little about

[12] There was no evidence of a positive association between frequency of reading comics, newspapers or magazines and achievement at any point along the reading test score distribution.

whether frequently reading one type of text (e.g. books) is more beneficial than another (e.g. newspapers).

The aim of this paper has been to contribute new evidence on this matter using unique longitudinal census data from the largest region within Spain. Whereas previous research in this area has mostly been based upon cross-sectional data (e.g. Jerrim & Moss 2018), our data tracks a large cohort of children over a three-year period (between ages 10-11 and 13-14) and includes rich controls for prior achievement, parental reading attitudes and activities, parental engagement towards their child's education and children's attitudes and engagement towards school. We have consequently been able to estimate a series of extensive "value-added" regression models; something that is relatively rare in this literature. This is hence one of the first studies to investigate how the reading of different text types is associated with young people's academic progress.

Our results provide further evidence that it is not only whether young people read or not that matters – but also *what* they read. As per some previous research, we find little evidence that reading newspapers, comics and magazines have positive benefits for young people's academic achievement. The association between the frequency children read these types of text and their scores on a Spanish language test is weak and often statistically insignificant once key potential confounders have been controlled. In contrast, the association between reading books/novels and young people's academic progress at school is quite strong. Even after a wide range of potential confounders have been controlled (including rich measures of prior achievement, family background, parental reading attitudes and activities, parental engagement towards their child's education and children's attitudes and engagement towards school and school fixed-effects) teenagers who read books every or almost every day continue to score around 0.22 standard deviations higher on Spanish Language tests than those who never or almost never read books. Our analysis also provides some evidence of spill-over effects, with frequently reading books (but not other text types) also associated with academic progress in mathematics.

Of course, these findings should be interpreted in the context of the limitations of this study and the need for further research. Four particular issues stand out. First, our analysis has focused upon academic progress made in reading and mathematics during the early-teenage years. At this point, reading skills are already quite well-developed. Yet we do not currently have evidence as to whether the results we have found hold for children at younger ages as well – such as when children are first starting to independently read. It is plausible that different text types (e.g. comics and magazines) have more positive (or negative) benefits at this point in children's lives. Establishing whether this is the case is therefore a key direction for future research. Second, although we have controlled for a rich array of potential confounders, our research design remains correlational, with our estimates not necessarily capturing cause and effect. Hence the next step needed within this literature is for a movement from observational to large-scale experimental research designs. This would enable causality to be more firmly established. Third, this paper has attempted to estimate the overall association between reading different text types and children's educational achievement. With the data currently available, it has not been possible to investigate the potential mechanisms by which such association occurs. Future work in this area should seek to address and identify the mechanisms through which any apparent "fiction effect" operates. Finally, our analysis has been conducted within one particular region within Spain. Although we argue that the richness of the data we have analysed is likely to mean our study has greater internal validity than much of the prior literature, it is weaker in terms of external validity – having been conducted within just one specific region of Spain. We consequently

believe that there would be significant benefits to other researchers replicating our findings using similarly rich cohort data in other national settings.

Despite these limitations, we believe that this paper has helped to fill some important gaps within the existing literature. It has provided new evidence that what children choose to read matters for their cognitive development. This has potentially important implications for parents, teachers and the wider education community. Our findings have highlighted how, in an increasingly digital world, it is important that young people are encouraged to find time to just sit down and read a good book. Other less complex and less engaging forms of reading are unlikely to bring the same benefits for their cognitive development and should not be counted as part of their "reading time". This is likely to be particularly important for low-achievers, where any effect is likely to be strongest.

References

Anderson, R. C., Wilson, P. T., & Fielding, L. G. (1988). Growth in Reading and How Children Spend Their Time Outside of School. *Reading Research Quarterly, 23*(3), 285-303.

Birr, E., Overby, M., Tysvaer, N., & Morris, K. (2008). The Complex World of Adolescent Literacy: Myths, Motivations, and Mysteries. *Harvard Educational Review, 78*(1), 107–154. doi: 10.17763/haer.78.1.54468j6204x24157

Bourdieu, P., Passeron, J.-C., & Saint-Martin, M. (1994). *Academic Discourse: Linguistic Misunderstanding and Professorial Power*. Cambridge: Polity Press.

Brozo, W. G., Shiel, G., & Topping, K. (2007). Engagement in Reading: Lessons Learned From Three PISA Countries. *Journal of Adolescent & Adult Literacy, 51*(4), 304–315. doi: 10.1598/jaal.51.4.2

Chiswick, B. R., Lee, Y. L., & Miller, P. W. (2003). Schooling, Literacy, Numeracy and Labour Market Success. *Economic Record, 79*(245), 165–181. doi: 10.1111/1475-4932.t01-1-00096

Duncan, L. G., McGeown, S. P., Griffiths, Y. M., Stothard, S. E., & Dobai, A. (2015). Adolescent reading skill and engagement with digital and traditional literacies as predictors of reading comprehension. *British Journal of Psychology, 107*(2), 209–238. doi: 10.1111/bjop.12134

García-Pérez, J. I., Hidalgo-Hidalgo, M., & Robles-Zurita, J. A. (2014). Does grade retention affect achievement? Some evidence from Spain. *Applied Economics, 46*(12), 1373–1392. doi: 10.1080/00036846.2013.872761

Gillespie, T. (2010). *Doing Literacy Criticism: The Cultivation of Thinkers in the Classroom*. Portsmouth: Stenhouse Publishers.

INE (2019). *Annual household net income per capita, by Autonomous Community*. Retrieved from http://www.ine.es/jaxiT3/Tabla.htm?t=9949 (Last time accessed April 2019).

Jerrim, J., & Moss, G. (2018). The link between fiction and teenagers' reading skills: International evidence from the OECD PISA study. *British Educational Research Journal, 45*(1), 181–200. doi: 10.1002/berj.3498

Kragler, S. (2000). Choosing Books for Reading: An Analysis of Three Types of Readers. *Journal of Research in Childhood Education, 14*(2), 133–141. doi: 10.1080/02568540009594758

Krashen, S. (2004). *The power of reading: Insights from the research. 2nd Edition*. Portsmouth, NH: Heinemann.

Moss, G., & McDonald, J. W. (2004). The borrowers: library records as unobtrusive measures of children's reading preferences. *Journal of Research in Reading, 27*(4), 401–412. doi: 10.1111/j.1467-9817.2004.00242.x

Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy, P. (2007). *PIRLS 2006 International Report*. Boston College: Lynch School of Education.

OECD (2010). *PISA 2009 Results: Learning to Learn – Student Engagement, Strategies and Practices (Volume III)*. Paris: OECD Publishing. doi: 10.1787/9789264083943-en

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

OECD (2017). *How's Life? 2017: Measuring Well-being*. Paris: OECD Publishing. doi: 10.1787/how_life-2017-en

Quintini, G. (2014). Skills at Work: How Skills and their Use Matter in the Labour Market. *OECD Social, Employment and Migration.* Working Papers, No. 158. Paris: OECD Publishing. doi: 10.1787/5jz44fdfjm7j-en

Ritchie, S. J., Bates, T. C., & Plomin, R. (2014). Does Learning to Read Improve Intelligence? A Longitudinal Multivariate Analysis in Identical Twins From Age 7 to 16. *Child development, 86*(1), 23–36. doi:10.1111/cdev.12272

Rodríguez, C., Álvarez, J., & Moreno, A. (2009). Attention to diversity programs: Psychological, academic -cognitive and social implications in the students of the autonomous community of Andalusia. *Profesorado, Revista de Currículum y Formación del Profesorado, 13*(3), 175–192.

Shanahan, T., Fisher, D., & Frey, N. (2012). The Challenge of Challenging Text. *Educational Leadership, 69*(6), 58–62.

Smith, M. C. (1996). Differences in adults' reading practices and literacy proficiencies. *Reading Research Quarterly, 31*(2), 196–219. doi: 10.1598/rrq.31.2.5

Sullivan, A., & Brown, M. (2015). Reading for pleasure and progress in vocabulary and mathematics. *British Educational Research Journal, 41*(6), 971–991. doi: 10.1002/berj.3180

Westbrook, J., Sutherland, J., Oakhill, J., & Sullivan, S. (2018). "Just reading": the impact of a faster pace of reading narratives on the comprehension of poorer adolescent readers in English classrooms. *Literacy*, in press. doi: 10.1111/lit.12141

Wright, D. (2006). Cultural capital and the literary field. *Cultural Trends, 15*(2-3), 123–139. doi: 10.1080/09548960600712934

Table 1. Comparison of Andalusian and Spanish education figures in 2015

		Andalusia	Spain	OECD
PISA 2015 Mean Scores	Reading	479	496	487
	Mathematics	466	486	478
	Sciences	473	493	488
Children who have repeated a grade by 12th grade		38%	31%	13%
Father's education	University studies	32%	43%	42%
	High school studies	15%	18%	33%
	Secondary education studies	29%	23%	15%
	Primary education studies	18%	12%	7%
	Less than primary education	7%	4%	3%
Mother's education	University studies	30%	43%	43%
	High school studies	17%	21%	33%
	Secondary education studies	30%	23%	14%
	Primary education studies	18%	11%	7%
	Less than primary education	5%	3%	3%
Annual household net income per capita, in PPPs		$16,276	$20,367	$28,443

Source: Authors' own calculations from PISA 2015, INE (2019) and OECD (2017).

Table 2. Gender, socio-economic and prior achievement differences in reading different text types

	Boys %	Girls%	Bottom SES quartile %	Top SES quartile %	Bottom grade 5 reading quartile %	Top grade 5 reading quartile
Comics						
Never/almost never	73	87	83	77	81	80
Once/twice a month	18	9	12	14	13	13
Once/twice a week	7	3	4	7	5	5
Every or almost every day	2	1	1	2	1	2
Short stories						
Never/almost never	52	38	48	40	48	40
Once/twice a month	31	38	34	35	33	36
Once/twice a week	14	17	15	18	15	17
Every or almost every day	3	7	3	7	4	7
Books						
Never/almost never	27	0.18	27	17	27	16
Once/twice a month	33	0.31	34	28	33	30
Once/twice a week	28	0.31	28	32	29	32
Every or almost every day	12	0.20	11	23	11	22
Newspapers						
Never/almost never	46	26	37	34	39	33
Once/twice a month	31	38	35	36	31	37
Once/twice a week	19	28	22	25	23	25
Every or almost every day	4	8	6	5	7	5
Magazines						
Never/almost never	48	73	63	58	61	61
Once/twice a month	24	18	20	22	20	22
Once/twice a week	19	7	12	14	13	12
Every or almost every day	9	2	5	6	6	5

Notes: Non-repeater students until 8th grade in 2011-12. Private schools in 2011-12 have not been included. Figures refer to column percentages.
Source: Authors' own calculations.

Table 3. The association between reading different text types of young people's reading scores. Effect sizes

	Comics	Short stories	Books	Magazines	Newspapers
Model 1	0	0.20***	0.40***	-0.03*	0.05**
Model 2	-0.02	0.14***	0.28***	0.00	0.04**
Model 3	-0.01	0.14***	0.28***	0.00	0.04**
Model 4	-0.01	0.14***	0.27***	0.00	0.04**
Model 5	-0.01	0.14***	0.27***	0.00	0.04**
Model 6	-0.03	0.12***	0.22***	0.00	0.02
N	**43,604**	**43,604**	**43,604**	**43,604**	**43,604**

Notes: Figures refer to the difference in eighth grade reading test scores between children who read each type of material every or almost every day compared to those who read the material never or almost never. Results presented in terms of effect sizes (standardised scores in reading using the mean and standard deviations of the total population in reading in 8th grade). Model 1 controls for demographic characteristics only (gender, socio-economic status and parental age). Model 2 controls for prior achievement in reading and mathematics. Model 3 adds parental reading attitudes and activities, while model 4 additional controls for multiple measures of parental engagement in their child's schooling. Children's attitudes and engagement towards school are added in model 5, while school fixed-effects in 5th and 8th grade are added in model 6. Standard errors are in parentheses and robust. Non-repeater students until 8th grade in 2011-12. Private schools in 2011-12 have not been included.
Coefficient: ***significant at 1%, ** significant at 5%, * significant at 10%.
Estimation method: Ordinary Least Squares (OLS).
Source: Authors' own calculations.

Table 4. The association between frequency of reading short stories/books and children's reading skills. Detailed final results. Effect sizes

	M1	M2	M3	M4	M5	M6
Short stories						
Never/almost never			Reference group			
Once/twice a month	0.14***	0.11***	0.11***	0.11***	0.11***	0.10***
Once/twice a week	0.14***	0.12***	0.12***	0.11***	0.11***	0.11***
Every or almost every day	0.20***	0.14***	0.14***	0.14***	0.14***	0.12***
Books						
Never/almost never			Reference group			
Once/twice a month	0.21***	0.15***	0.15***	0.15***	0.14***	0.12***
Once/twice a week	0.27***	0.20***	0.20***	0.20***	0.19***	0.16***
Every or almost every day	0.40***	0.28***	0.28***	0.27***	0.27***	0.22***

Notes: Figures refer to the difference in eighth grade reading test scores compared to the reference group (children who read the material never/almost never). Results presented in terms of effect sizes (standardised scores in reading using the mean and standard deviations of the total population in reading in 8th grade). Model 1 controls for demographic characteristics only (gender, socio-economic status and parental age). Model 2 controls for prior achievement in reading and mathematics. Model 3 adds parental reading attitudes and activities, while model 4 additional controls for multiple measures of parental engagement in their child's schooling. Children's attitudes and engagement towards school are added in model 5, while school fixed-effects in 5th and 8th grade are added in model 6. Standard errors are in parentheses and robust. Non-repeater students until 8th grade in 2011-12. Private schools in 2011-12 have not been included.

Coefficient: ***significant at 1%, ** significant at 5%, * significant at 10%.

Estimation method: Ordinary Least Squares (OLS).

Source: Authors' own calculations.

Table 5. The association between reading different text types of young people's mathematics scores. Effect sizes

	Comics	Short stories	Books	Magazines	Newspapers
Model 1	0.04	0.17***	0.38***	-0.14***	0.05**
Model 2	0.05	0.10***	0.23***	-0.09***	0.03*
Model 3	0.05	0.10***	0.23***	-0.09***	0.03*
Model 4	0.05	0.11***	0.23***	-0.08***	0.04**
Model 5	0.05	0.10***	0.23***	-0.08***	0.04**
Model 6	0.05*	0.10***	0.20***	-0.08***	0.03*
N	**43,833**	**43,833**	**43,833**	**43,833**	**43,833**

Notes: Figures refer to the difference in eighth grade mathematics test scores between children who read each type of material every or almost every day compared to those who read the material never or almost never. Results presented in terms of effect sizes (standardised scores in mathematics using the mean and standard deviations of the total population in mathematics in 8th grade). Model 1 controls for demographic characteristics only (gender, socio-economic status and parental age). Model 2 controls for prior achievement in reading and mathematics. Model 3 adds parental reading attitudes and activities, while model 4 additional controls for multiple measures of parental engagement in their child's schooling. Children's attitudes and engagement towards school are added in model 5, while school fixed-effects in 5th and 8th grade are added in model 6. Standard errors are in parentheses and robust. Non-repeater students until 8th grade in 2011-12. Private schools in 2011-12 have not been included.
Coefficient: ***significant at 1%, ** significant at 5%, * significant at 10%.
Estimation method: Ordinary Least Squares (OLS).
Source: Authors' own calculations.

Table 6. The association between frequency of reading short stories/books and children's mathematics skills. Detailed final results. Effect sizes

	M1	M2	M3	M4	M5	M6
Short stories						
Never/almost never			Reference group			
Once/twice a month	0.12***	0.09***	0.09***	0.08***	0.08***	0.08***
Once/twice a week	0.12***	0.10***	0.10***	0.10***	0.09***	0.10***
Every or almost every day	0.17***	0.10***	0.10***	0.11***	0.10***	0.10***
Books						
Never/almost never			Reference group			
Once/twice a month	0.19***	0.13***	0.12***	0.12***	0.12***	0.09***
Once/twice a week	0.23***	0.15***	0.15***	0.15***	0.15***	0.12***
Every or almost every day	0.38***	0.23***	0.23***	0.23***	0.23***	0.20***

Notes: Figures refer to the difference in eighth grade mathematics test scores compared to the reference group (children who read the material never/almost never). Results presented in terms of effect sizes (standardised scores in reading using the mean and standard deviations of the total population in reading in 8th grade). Model 1 controls for demographic characteristics only (gender, socio-economic status and parental age). Model 2 controls for prior achievement in reading and mathematics. Model 3 adds parental reading attitudes and activities, while model 4 additional controls for multiple measures of parental engagement in their child's schooling. Children's attitudes and engagement towards school are added in model 5, while school fixed-effects in 5th and 8th grade are added in model 6. Standard errors are in parentheses and robust. Non-repeater students until 8th grade in 2011-12. Private schools in 2011-12 have not been included.

Coefficient: ***significant at 1%, ** significant at 5%, * significant at 10%.

Estimation method: Ordinary Least Squares (OLS).

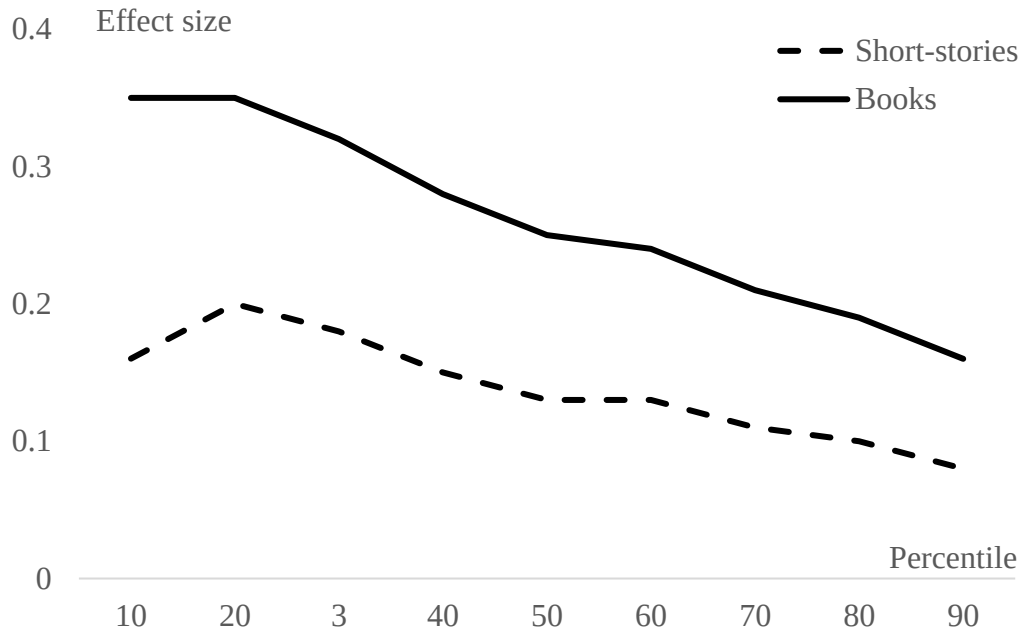
Source: Authors' own calculations.

Table 7. Investigation of heterogeneous effects. Differences by gender and socio-economic status.

	Boys	Girls	High SES	Low SES
Comics	-0.04	-0.01	-0.09	-0.12
Short Stories	0.12	0.11	0.10	0.22
Books	0.20	0.23	0.20	0.25
Magazines	0.03	-0.04	-0.03	-0.03
Newspapers	0.03	-0.01	0.06	-0.02

Notes: Estimates based upon model M6. Figures presented refer to effect sizes. High SES refers to students in the top socio-economic quartile and low SES to those in the bottom quartile. Results refer to the difference in reading test scores between children who said they read the text type 'never'/ 'almost never' compared to those who said that they read the text type every day/almost every day.

Figure 1. Quantile regression estimates of the link between reading short-stories and books and students reading test scores



Notes: Estimates based upon quantile regression using model specification M5. Results refer to the difference in reading test scores between children who said they read the text type 'never'/ 'almost never' compared to those who said that they read the text type every day/almost every day. Left-hand side of the graph (e.g. P10, P20, P30) illustrate the results for lower-achievers while the right-hand side of the graph (e.g. P70, P80 and P90) refer to the results for high-achievers.

Appendix A. Descriptive statistics in 5th grade: non-repeater population and sample

Variables		Population			Sample		
		Obs.	Mean	S.D.	Obs.	Mean	S.D.
Scores 2008-09	Reading	54,319	70.69	16.19	44,483	73.40	14.96
	Mathematics	54,089	50.82	11.83	44,303	52.84	10.74
Sex	Male	55,331	0.50	0.50	45,160	0.48	0.50
	Female	55,331	0.50	0.50	45,160	0.52	0.50
School funding	Semi-private	55,331	0.24	0.43	45,160	0.26	0.44
	Public	55,331	0.76	0.43	45,160	0.74	0.44
Socio-economic status index variables							
Parental age	Father	43,647	65.68	5.52	36,322	65.49	5.29
	Mother	46,582	68.12	5.11	38,408	67.81	4.89
Socio-economic status index		46,379	0.09	0.97	38,311	0.2	0.94
Level of education of the father	Incomplete primary education or did not attend school	43,120	0.15	0.36	36,003	0.13	0.33
	EGB or Compulsory Secondary Education	43,120	0.36	0.48	36,003	0.35	0.48
	High school, First Grade Professional Formation, Elemental Arts School and Artistic Professions, BUP, COU, Official Language School or Medium Grade Professional Formation Cycle	43,120	0.32	0.47	36,003	0.33	0.47
	Second Grade Professional Formation, Arts Speciality and Artistic Professions or High Grade Professional Formation Cycle	43,120	0.00	0.00	36,003	0.00	0.00
	University degree, PhD	43,120	0.17	0.37	36,003	0.19	0.39
Level of education of the mother	Incomplete primary education or did not attend school	45,984	0.12	0.32	38,005	0.10	0.29
	EGB or Compulsory Secondary Education	45,984	0.40	0.49	38,005	0.39	0.49
	High school, First Grade Professional Formation, Elemental Arts School and Artistic Professions, BUP, COU, Official Language School or Medium Grade Professional Formation Cycle	45,984	0.30	0.46	38,005	0.31	0.47
	Second Grade Professional Formation, Arts Speciality and Artistic Professions or High Grade Professional Formation Cycle	45,984	0.00	0.00	38,005	0.00	0.00
	University degree, PhD	45,984	0.18	0.38	38,005	0.20	0.40
Occupation of the father	Business managers or public administration	42,932	0.06	0.23	35,874	0.06	0.24
	Technicians, professionals, scientists and intellectuals. Army (officials and high ranks)	42,932	0.14	0.34	35,874	0.15	0.36
	Technicians and support professionals. Administrative employees. Little business people	42,932	0.20	0.40	35,874	0.22	0.41
	Hotel workers, personnel, protection and sellers. Army (sub-officials and low ranks)	42,932	0.15	0.36	35,874	0.15	0.36
	Agriculture and fishing qualified workers. Artisans and qualified manufacturing, construction and mining workers	42,932	0.36	0.48	35,874	0.35	0.48
	Non-qualified workers	42,932	0.06	0.24	35,874	0.05	0.22
	Performing housework	42,932	0.01	0.07	35,874	0.00	0.06
	Inactive	42,932	0.02	0.13	35,874	0.02	0.13
Occupation of the mother	Business managers or public administration	45,310	0.02	0.14	37,571	0.02	0.14
	Technicians, professionals, scientists and intellectuals. Army (officials and high ranks)	45,310	0.12	0.32	37,571	0.13	0.34
	Technicians and support professionals. Administrative employees. Little business people	45,310	0.16	0.37	37,571	0.17	0.38
	Hotel workers, personnel, protection and sellers. Army (sub-officials and low ranks)	45,310	0.14	0.35	37,571	0.14	0.35
	Agriculture and fishing qualified workers. Artisans and qualified manufacturing, construction and mining workers	45,310	0.06	0.24	37,571	0.06	0.23
	Non-qualified workers	45,310	0.11	0.32	37,571	0.11	0.31
	Performing housework	45,310	0.37	0.48	37,571	0.36	0.48
	Inactive	45,310	0.02	0.12	37,571	0.01	0.12
Number of books at home	0-10	47,220	0.10	0.29	38,841	0.08	0.27
	nov-25	47,220	0.20	0.40	38,841	0.18	0.39
	26-50	47,220	0.38	0.49	38,841	0.38	0.49
	51-100	47,220	0.16	0.37	38,841	0.18	0.38
	More than 100	47,220	0.16	0.37	38,841	0.18	0.38
The student has a place for him/her to study at home	Yes	47,512	0.88	0.32	39,033	0.90	0.30
	No	47,512	0.12	0.32	39,033	0.10	0.30
Study desk	Yes	47,512	0.90	0.30	39,033	0.92	0.28
	No	47,512	0.10	0.30	39,033	0.08	0.28
Computer	Yes	47,512	0.83	0.38	39,033	0.85	0.35

	No	47,512	0.17	0.38	39,033	0.15	0.35
Internet	**Yes**	47,512	0.60	0.49	39,033	0.63	0.48
	No	47,512	0.40	0.49	39,033	0.37	0.48
Digital TV, cable or satellite	**Yes**	47,512	0.51	0.50	39,033	0.53	0.50
	No	47,512	0.49	0.50	39,033	0.47	0.50
Video, CD or DVD player	**Yes**	47,512	0.92	0.27	39,033	0.94	0.24
	No	47,512	0.08	0.27	39,033	0.06	0.24
Casebooks and school support books (encyclopaedias, dictionaries…)	**Yes**	47,512	0.92	0.27	39,033	0.94	0.24
	No	47,512	0.08	0.27	39,033	0.06	0.24
Reading books (novels, tales, poems, comics…)	**Yes**	47,512	0.89	0.31	39,033	0.91	0.28
	No	47,512	0.11	0.31	39,033	0.09	0.28
Specialised magazines	**Yes**	47,512	0.41	0.49	39,033	0.43	0.50
	No	47,512	0.59	0.49	39,033	0.57	0.50
Daily press	**Yes**	47,512	0.32	0.47	39,033	0.33	0.47
	No	47,512	0.68	0.47	39,033	0.67	0.47

Notes: "Obs." stands for "Observations" and "S.D." indicates "standard deviation". For comparison purposes to the used sample, the population is that of non-repeater students in the course 2008/09.

Source: Authors' own calculations from DA 2008-09.

Appendix B. Descriptive statistics for the variables employed as statistical controls. Employed sample

Variables		Obs.	Mean	S.D.
The student reads comics in 8th grade	Every or almost every day	41,895	0.01	0.12
	Once or twice a week	41,895	0.05	0.21
	Once or twice a month	41,895	0.13	0.33
	Never or almost never	41,895	0.81	0.39
The student reads tales or short novels in 8th grade	Every or almost every day	41,651	0.05	0.22
	Once or twice a week	41,651	0.16	0.37
	Once or twice a month	41,651	0.34	0.47
	Never or almost never	41,651	0.45	0.50
The student reads books in 8th grade	Every or almost every day	41,643	0.16	0.37
	Once or twice a week	41,643	0.30	0.46
	Once or twice a month	41,643	0.32	0.47
	Never or almost never	41,643	0.22	0.41
The student reads magazines in 8th grade	Every or almost every day	41,713	0.06	0.24
	Once or twice a week	41,713	0.23	0.43
	Once or twice a month	41,713	0.35	0.48
	Never or almost never	41,713	0.36	0.48
The student reads newspaper in 8th grade	Every or almost every day	41,524	0.05	0.23
	Once or twice a week	41,524	0.13	0.33
	Once or twice a month	41,524	0.21	0.41
	Never or almost never	41,524	0.61	0.49
Time that parents spend reading at home, including books, magazines, newspapers and work materials in 5th grade	More than 10 hours a week	38,748	0.17	0.38
	Between 6-10 hours a week	38,748	0.30	0.46
	Between 1-5 hours a week	38,748	0.44	0.50
	Less than one hour a week	38,748	0.09	0.28
Parents read because they have to in 5th grade	Strongly agree	34,201	0.04	0.21
	Quite agree	34,201	0.05	0.22
	Quite disagree	34,201	0.21	0.41
	Strongly disagree	34,201	0.70	0.46
Parents like talking about books with other people in 5th grade	Strongly agree	34,141	0.30	0.46
	Quite agree	34,141	0.45	0.50
	Quite disagree	34,141	0.19	0.39
	Strongly disagree	34,141	0.06	0.23
Parents like spending leisure time reading in 5th grade	Strongly agree	35,389	0.31	0.46
	Quite agree	35,389	0.43	0.49
	Quite disagree	35,389	0.20	0.40
	Strongly disagree	35,389	0.06	0.23
Parents only read when they need information in 5th grade	Strongly agree	34,973	0.11	0.32
	Quite agree	34,973	0.12	0.32
	Quite disagree	34,973	0.34	0.48
	Strongly disagree	34,973	0.43	0.49
Parents say reading is an important activity at home in 5th grade	Strongly agree	35,312	0.40	0.49
	Quite agree	35,312	0.45	0.50
	Quite disagree	35,312	0.13	0.34
	Strongly disagree	35,312	0.02	0.14
Father's expected level of education to complete by his son or daughter in 5th grade	University degree	27,409	0.83	0.38
	High grade formation course	27,409	0.07	0.26
	High school	27,409	0.04	0.20
	Medium grade formation course	27,409	0.03	0.18
	Secondary education	27,409	0.03	0.17
Mother's expected level of education to complete by her son or daughter in 5th grade	University degree	35,974	0.82	0.38
	High grade formation course	35,974	0.08	0.26
	High school	35,974	0.04	0.19
	Medium grade formation course	35,974	0.03	0.18
	Secondary education	35,974	0.03	0.16
Student says parents cheer him/her to study in 5th grade	Every day	44,329	0.53	0.50
	Almost every day	44,329	0.17	0.38
	Some days	44,329	0.20	0.40
	Never	44,329	0.10	0.30
Student says parents ask if he/she have homework in 5th grade	Every day	44,625	0.85	0.35
	Almost every day	44,625	0.08	0.27
	Some days	44,625	0.05	0.21
	Never	44,625	0.02	0.13
Student says parents check that he/she does homework in 5th grade	Every day	44,370	0.62	0.48
	Almost every day	44,370	0.16	0.37
	Some days	44,370	0.13	0.33
	Never	44,370	0.09	0.28
Student says parents ask him/her about the school day in 5th grade	Every day	44,616	0.73	0.44

	Almost every day	44,616	0.16	0.37
	Some days	44,616	0.09	0.29
	Never	44,616	0.02	0.14
Student says that father, mother or any person of the family help him/her with homework in 5th grade	**Yes**	44,483	0.16	0.37
	Sometimes	44,483	0.73	0.44
	No	44,483	0.11	0.32
Student's expected level of education to complete in 5th grade	**University degree**	44,376	0.82	0.38
	High grade formation course	44,376	0.05	0.21
	High school	44,376	0.07	0.25
	Medium grade formation course	44,376	0.02	0.14
	Secondary education	44,376	0.04	0.20
Student likes going to school in 5th grade	**A lot**	44,760	0.43	0.50
	Enough	44,760	0.38	0.48
	Little	44,760	0.15	0.36
	Nothing	44,760	0.04	0.19
Student wants to move to another school in 5th grade	**Yes**	44,818	0.09	0.29
	He/she does not mind	44,818	0.02	0.14
	No	44,818	0.89	0.32

Notes: "Obs." stands for "Observations" and "S.D." indicates "standard deviation".
Source: Authors' own calculations from DA.

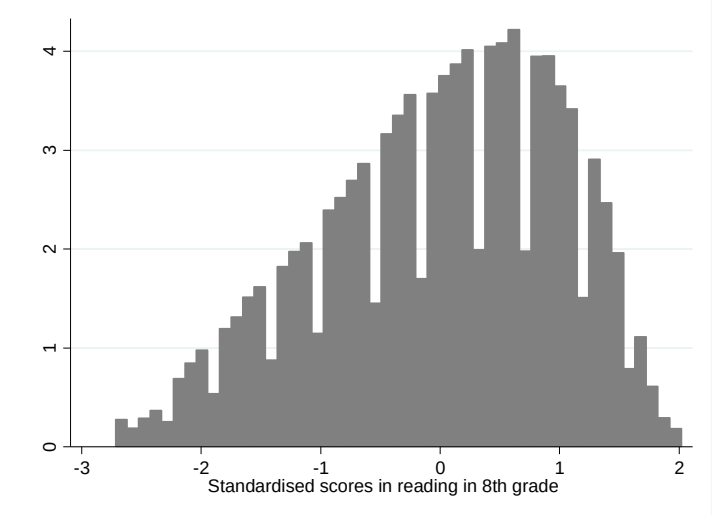
Appendix C. Wording of key background family questionnaire questions on parental reading attitudes and activities

Family questionnaire	
Question	**Possible answers**
During a normal week, how much time do you spend reading at home, including books, magazines, newspapers and work materials?	Choose one: a. Less than one hour a week. b. Between 1-5 hours a week. c. Between 6-10 hours a week. d. More than 10 hours a week.
Level of agreement with the sentence: -"I read because I have to". -"I like talking about books with other people". -"I like spending my leisure time reading". -"I only read when I need information". -"Reading is an important activity at home".	Choose one for each one of the five sentences: a. Strongly agree. b. Quite agree. c. Quite disagree. d. Strongly disagree.

Source: Authors' own translation of the family questionnaire.

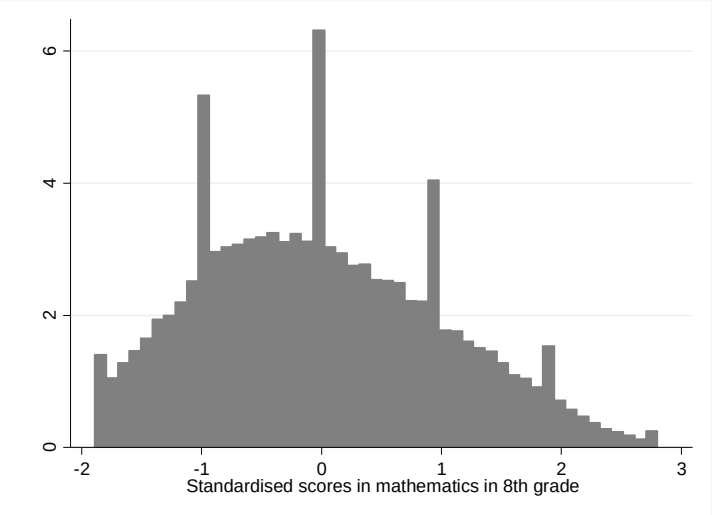
Appendix D. The distribution of grade 8 academic achievement measures

Figure D1. Histogram on students' standardised scores in reading in 8th grade



Source: Authors' own calculations.

Figure D2. Histogram on students' standardised scores in mathematics in 8th grade



Source: Authors' own calculations.