

Unraveling the effect of extrinsic reading on reading with intrinsic motivation

Sara Suárez-Fernández*

suarezfsara@uniovi.es

David Boto-García

botodavid@uniovi.es

Department of Economics, University of Oviedo, Asturias, Spain

ABSTRACT

In this paper, we analyze the effect of the time spent on reading for leisure (when the motivation is intrinsic) on the time devoted to reading for job-related or educational purposes (when the motivation is extrinsic). To do so, we use the Cultural Habits and Practices Survey conducted by the Ministry of Education and Culture of Spain in 2014-2015. As the main determinants of the time spent on intrinsic reading, we consider the time devoted to extrinsic reading, sociodemographic characteristics, labor situation, participation in other leisure activities and an index of cultural capital at home. We estimate a Heckman model that allows us to control for self-selection. Results show that the time devoted to intrinsic reading mainly depends on the time spent on reading with extrinsic motivation, human capital and cultural background. Reading with extrinsic motivation increases the likelihood of intrinsic reading but reduces the time allocated to it. Therefore, our results suggest a substitution pattern between time spent reading for leisure and reading with extrinsic motivation.

JEL classification: **C24, J22, L82, Z1**

Keywords: **leisure, intrinsic motivation, extrinsic motivation, reading, Heckman, cultural economics**

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. All errors in fact or interpretation are ours. We would like to thank Víctor Fernández-Blanco and Juan Prieto-Rodríguez, the editor and two anonymous referees for their constructive comments

*Corresponding author: Sara Suárez-Fernández suarezfsara@uniovi.es
+34677029628 +34985104885
ORCID ID orcid.org/0000-0003-2894-3732

1. Introduction

Reading plays an important role in people's future. Through the acquisition of human capital (Becker, 1964), it has plenty of benefits. It improves reading abilities, creativity and language and cognitive skills (Kloosterman et al. 2011; Mol and Bus 2011), which determine later socio-economic success (Heckman and Masterov 2007). The purpose of reading is the comprehension of what it is being read. In this regard, the motivation for reading has been shown to play an important role, since it affects both reading amount and comprehension (Morgan and Fuchs 2007; Schiefele et al. 2016; Schaffner et al. 2013).

According to the literature, reading motivation can be intrinsic, when it is internally satisfying to you, or extrinsic, when you read to pass an exam or just because it is required in your job (Clark and Zoysa 2011; Rosenzweig and Wigfield 2017). Reading as an entertainment, that is, when the motivation is intrinsic, enhances academic and social success (Gottfried et al. 2015; Mol and Bus 2011; Notten 2011; Nielen et al. 2016). This can be explained by the richness of the contents and the possibility of choosing the type of reading (an article, the newspaper, or a novel). Conversely, when reading becomes an imposition (i.e. the motivation is extrinsic), and especially when it involves plain, technical or shallow contents, reading benefits are notably reduced (Soemer and Schiefele 2019). Therefore, the benefits of reading are bigger when the motivation is intrinsic, but they are substantially reduced when the motivation is extrinsic.

In this context, we examine the effect that the amount of extrinsic reading (for work, study, job searching...) has on the allocation of time to intrinsic reading. Our empirical model is based on a theoretical two-step model for time allocation. Given that readers might not be randomly selected from the population, we account for possible sample selection bias by estimating a Heckman model. We firstly estimate the probability of reading with intrinsic motivation. Then, we model the demand for intrinsic reading conditional on being reader. Our results show that extrinsic motivation for reading is positively associated with the likelihood of becoming a reader with an intrinsic motivation. However, the more time spent on extrinsic reading, the lower the amount of time allocated to intrinsic reading.

The paper is structured as follows. Section 2 resumes the related literature. Section 3 develops the theoretical framework. The database is described in Section 4. Section 5 describes the econometric model. Following, Section 6 reports estimation results. Lastly, Section 7 summarizes the main conclusions.

2. Literature review

Reading has been shown to produce several long-term positive effects in people's lives. For instance, reading during childhood and adolescence is positively related to educational achievement, occupational standing, enhanced vocabulary and adult literacy skills, among others (Allington and McGill-Franzen 2003; Andreassen and Braten 2010; Suggate et al. 2018; Sikora et al. 2019). What is more, reading has been shown to protect individuals against cognitive deterioration and dementia (Verghese et al. 2003), and to be associated with longer life expectancy (Bravishi et al. 2016).

However, not all types of reading are equally beneficial. The literature has found that the acquisition of reading ability is closely associated with perceiving it as an enjoyable activity (Stokmans 1999; Becker et al. 2010; McGeown et al. 2012b). This is especially true for children (Guthrie et al. 2007; Retelsdorf et al. 2011). Therefore, reading motivation plays an important role (Morgan and Fuchs 2007; Schaffner et al. 2013; Schiefele et al. 2016). According to the literature, reading motivation can be broadly classified into intrinsic (when it is internally satisfying to you) and extrinsic (when it is externally imposed to you). Intrinsic motivation is positively related to reading amount and comprehension, whereas extrinsic motivation is not (Logan et al. 2011; McGeown et al. 2012b; Wang and Eccles 2013; Stutz et al. 2016; Troyer et al. 2019). Moreover, individuals' low intrinsic motivation has important negative effects, since it is related, for example, to lower educational aspirations (Metsäpelto et al. 2017). The study by Schutte and Malouff (2004) shows that the engagement in recreational reading is mainly driven by intrinsic motivation.

Given that the main purpose of reading is the comprehension of what it is being read, the linkages between reading motivation and comprehension have received substantial attention (Wang and Guthrie 2004; Schiefele et al. 2012; Stutz et al. 2016). Reading comprehension determines academic performance (Wang and Eccles 2013), and it is positively related to earnings and employment (Murnane et al. 2000; McIntosh and Vignoles 2001; Carbonaro 2007). Perceiving reading as pleasant activity fosters comprehension. Conversely, difficult texts increase mind wandering, leading to a negative correlation between text difficulty and comprehension (Soemer and Schiefele 2019).

When it comes to enhance intrinsic motivation for reading, scholars agree to note that parents play a crucial role in their children's behavior (Gil-Flores 2009; Gottfried et al. 2015). Based on the cultural reproduction hypothesis (Bourdieu and Passeron 1990), parents transmit their tastes in highbrow activities at home, being the family

responsible for much of their sons and daughters' habits (Goux et al. 2017). Also, family cultural capital has been shown to affect adolescent motivation for reading (Chiu and Chow 2010). Accordingly, there is a positive effect of parents' leisure reading practices and the reading attitude of their offspring (Verboord and Van Rees 2003; Notten et al. 2012; Pfof et al. 2016). Interestingly, Wollscheid (2014) find that sons are strongly influenced by the reading behavior of their fathers, whereas daughters are more affected by their mothers. Similar results are reported by Mullan (2010). Furthermore, parent's reading to young children has positive effects that show up during adulthood (Kalb and van Ours 2014).

Teachers also have a great deal of responsibility in developing their students' reading taste (Guthrie and Klauda 2012). Nielen et al. (2016) note that many students perceive reading as a threatening activity, which explains their low reading motivation and frequency. The way teachers introduce reading to students becomes crucial for later recreational reading engagement (Guthrie and Davis 2003).

Another stream of research has paid attention to the sociodemographic profile of an avid reader. The literature agrees to note that reading habits differ depending on gender, age, education and income. We now proceed to discuss the main empirical findings about the reader profile.

In general terms, females read more than males (Coles and Hall 2002; Clark et al. 2008; Logan and Johnston 2009; Clark 2011; Brozko et al. 2014). Females also exhibit a higher intrinsic motivation for reading (McGeown et al. 2012a; Schaffner et al. 2013; McGeown 2015) which can explain, for example, their better performance on PISA reading exams (Torppa et al. 2018). However, this pattern does not hold for all genres of reading. For example, Elvestad and Blekesaume (2008) note that males spend more time reading the newspaper. Also, females tend to opt for novels, whereas males seem to prefer short stories, reports and comics (Scales and Rhee 2001; Clark and Foster 2005). Interestingly, Bortolussi et al. (2010) find that both genres appreciate more books with a male protagonist. Childhood socialization and gender-role stereotypes can partially explain these findings (Tepper 2000).

Reading is more prevalent among elderly people (Roe and Taube 2012). In spite of this, each new generation reads less than the generation it replaces in the case of newspapers (Lauf, 2001). In addition, there has been a decline in books' reading frequency, probably due to the low influx of avid readers among the recent birth cohorts (Knulst and van der Broek, 2003).

Another stylized fact is the existence of a positive relationship between reading and education level (Scales and Rhee 2001), which is even stronger as formal education includes more literacy components (Verboord 2005). Time spent on reading crucially depends on reading ability (McQuillan and Au 2001; Ivey and Broaddus 2001), which is fostered by education and learning-by-doing. The more the individual reads, the greater his/her reading ability and, therefore, the higher his/her interest for reading. As said by Stanovich (2000), "The rich get richer" (p.184), which is consistent with the process of the cultivation of tastes described by McCain (1979). The ability to appreciate culture depends on the accumulated cultural capital, as individuals read more, their reading ability improves, which increases their probability of reading (Becker 1996; Stigler and Becker 1977; Becker and Murphy 1988).

As for the effect of income, while some studies find it positively related to reading (Elvestad and Blekesaune 2008), others find no significant effects (Stokmans 1999). Time availability can be more important than income level, since reading has been shown to be positively related to unemployment rates (Siddiqi et al. 2007). In this sense, Molina et al. (2016) show that being self-employed has a negative effect on the time spent reading in Spain.

When studying the decision to read, it is important to consider the competition for time allocation between reading and other leisure activities. Knulst and Kraaykamp (1998) find that television viewing is one of its main substitutes, reducing both academic and recreational reading. Similar findings are reported in Mokhtari et al. (2009). In line with this, Koolstra and van der Voort (1996) show how children's reading concentration decreases as time spent on television viewing increases. Additionally, doing several leisure activities at the same time, such as listening to music or the radio while reading, produces a more superficial reading (Loan 2012).

As for the empirical findings about reading patterns in Spain, Fernández-Blanco and Prieto-Rodríguez (2009) scrutinized both the determinants of reading frequency and the number of books read. Their results show that both dimensions are positively related with education, physical cultural capital at home and participation in other cultural activities. More recently, Fernández-Blanco et al. (2017) studied the total number of books read, distinguishing between readers and non-readers. They find that females and elderly people read more books. Interestingly, visiting monuments and museums is positively associated with the number of books read. Furthermore, unemployed and retired individuals are less likely to read.

3. A two-stage decision model for leisure time allocation

Our empirical estimations are based on a theoretical two-stage decision model for time allocation. Individuals are assumed to derive utility from both the time devoted to leisure and goods consuming, as in the classical time allocation model proposed by Becker (1965). Taking into account monetary and time constraints, the decision of time allocation is the result of the maximization of the following utility function:

$$\text{Max } U = f(X, L)$$

subject to

$$\left. \begin{array}{l} X = Y + wH \\ L + H = T \end{array} \right\} \quad (1)$$

where L indicates the number of leisure hours; X is the Hicksian composite good representing the consumption of all goods other than leisure, whose relative prices are assumed to be the same for all individuals (Hicks 1939); H is the number of working hours; w is the real wage for each working hour; T is the total available time and Y is the income from other sources but work. In the optimum, each individual chooses the number of hours devoted to work (H^*) and leisure (L^*) so that the marginal utility of leisure equals the product of the marginal utility of consumption and the real wage.

In the short run L^* is fixed, as the time allocated to work (H^*) can be only adjusted in the long-run. Based on this, Kooreman and Kapteyn (1987) further disaggregated the residual leisure time into multiple unpaid activities. Therefore, our framework assumes that individuals allocate their total available time to work and leisure in a first stage (long-run equilibrium), and then they divide the chosen leisure time (L^*) into a set of N leisure alternatives (short-run equilibrium). In the short run, the individual maximizes a second leisure utility function:

$$\text{Max } V = f(L_1, L_2, \dots, L_N)$$

$$\text{subject to } \sum_{j=1}^N L_j = L^* \quad (2)$$

where L_j denotes the time spent in leisure activity j .

Since only the second decision is modelled here, it is important to note that our model is valid just for the short-run. Under the assumption that the utility function with leisure and consumption as the basic arguments is weakly separable (Deaton and Muellbauer 1980), we only need the total available amount of time left for leisure taking pre-fixed

values of the selected working hours for consumption (H^*), the total available time (T), the real wage (w) and the non-labor income (Y)¹. The conditional leisure time “demand” function (Pollak, 1969) for activity j is given by:

$$L_j = f(P_j, L^*, Y + wH, Pr) \quad (3)$$

Where P_j indicates the price for each period of leisure j , L^* is the total amount of time for leisure selected in the first stage, $Y + wH$ is the total income and Pr is a composite term which denotes individual’s preferences for the activity j . This term is a vector of taste shifters with observable characteristics of the individual (Z) that controls for observable heterogeneity in preferences. A random term (e) for non-observable characteristics is also included.

4. Data

Our dataset is drawn from the 2014-15 wave of the Cultural Habits and Practices Survey (*Encuesta de Hábitos y Prácticas Culturales*) conducted by the Ministry of Education and Culture of Spain. This survey gathers information on Spanish households’ cultural habits, as well as sociodemographic and labor-related data for a representative sample of the Spanish population. A total of 15,154 individuals were surveyed along 2014 and 2015. Since they were not followed over time, our data has a cross-sectional structure.

In the survey, people were asked about several aspects of their habits concerning cultural practices, paying special attention to the intensity of their consumption. Furthermore, several socioeconomic characteristics such as age, gender, the number of house members, the labor status or the educational level were also gathered.

It is important to note that we are interested in reading not only with intrinsic motivation but also with extrinsic motivation. Therefore, we only consider students, employed and unemployed people in our sample. Students spend time reading due to an extrinsic motivation, to achieve their goals in their studies. In a similar way, employed people are sometimes required to read at their jobs. As for unemployed people, they might allocate time to formative courses, job interviews or job searching, which might also involve extrinsic reading. Housewives (househusbands) and those who are retired or disabled were thus removed from the sample. Our final sample comprises a total of 10,319 individuals.

¹ This implies that the MRS between leisure activities j and k within leisure decision is independent of working hours.

Among the surveyed questions, respondents were asked the following: *'If you read within the last three months, how much time do you usually spend on reading on a working-day due to professional or educational reasons? And due to other reasons (entertainment)?'* This question was specified both for week days and weekends, separately. Here we only analyze week days. The reason is that people usually have to read more with extrinsic motivation during the week. At the same time, time availability for reading with intrinsic motivation during week days is lower. Therefore, it is more interesting to study the relation between them during the week. Importantly, our dataset does not allow us to determine the type of material which is being read.

Our dependent variable is *'Intrinsic reading'* (expressed in logarithms and denoted as Ln_IR), that is, the declared average daily time (minutes) spent on reading with an intrinsic motivation from Monday to Friday by each individual². As our main interest is to examine the relationship between reading with intrinsic and extrinsic motivation, our main explanatory variable is the exogenous time spent on reading with extrinsic reasons, *'Extrinsic reading'*, (also expressed in logarithms, hereafter Ln_ER). In order to account for the possible non-linearity of this relation, we included the square term of Ln_ER (denoted as Ln_ER_sq) in both equations.

The following explanatory variables are defined as controls:

- $Csoc$ is a vector of sociodemographic characteristics. It comprises gender, *'Male'* (*'Female'*) when the individual is a male; *'Age'* (in years); *'Household members'* (number of members living in the household) and *'Youths'* (number of household members younger than 18 years old).
- $Creg$ is a vector of regional dummies which comprises the region where the individual lives (NUTS 2) and the population size of his/her place of residence.
- $Clab$ refers to labor status, including *'Self-employed'*, *'Employee'*, *'Unemployed'* and *'Student'*.
- $Cedu$ stands for the educational level, which contains: *'Primary'* when individual's higher level of education is primary education or lower, *'Secondary'* when he/she reached secondary education, *'GCE'* (General Certificate of Education) when the individual finished two additional years of education after compulsory education, *'Vocational'* which characterizes both Vocational Education and Training and

² As our dependent variable, *Intrinsic Reading* (IR) is heavily skewed and has considerable non-normal kurtosis, we firstly estimate the λ parameter of the Box-Cox transformation in order to test if the regression model for the time devoted to reading is better in logs than in levels. Given that $\hat{\lambda} = 0,13$, we have greater support for expressing this variable in logs (Cameron and Trivedi 2009). The same holds for *Extrinsic Reading* (ER) so we also take its natural logarithm (Ln_ER)

Certificate of Higher Education, and *'University'* meaning that the individual reached different grades of University degrees or higher education.

- Concerning habits, we define the vector Z. It includes dummy variables to account for the daily consumption of the following: *'Television'*, *'Radio'*, *'Computer'*, *'Gaming'* (for playing video games), and *'Sport'* (for those who practice sport more than three times a week). We also considered *'Library'* as another dummy variable that controls for the habit of going to the library with any purpose at least once a week.
- To account for cultural interests, we aggregated the frequency of participation in cultural visits and events in the variable *'Cultural Participation'*, which is measured as the number of times the individual declares having participated in cultural activities within the last three months. The activities considered are the following: visiting monuments, archeological sites, museums, art exhibitions, art galleries or file sites and attending to the ballet, opera, Spanish operetta (zarzuela), theatre, circus, popular music concerts, classic music concerts or cinema.
- We lack household or individual income. Following the practice of Fernández-Blanco and Prieto-Rodríguez (2009), who use the same survey for 2010-11, a Principal Component Analysis (PCA) was conducted to proxy households' income³. PCA generates an index for the number of electronic devices and physical cultural items that people have at home⁴. For further information about the PCA, see Annex 1.

Descriptive statistics for each variable defined above are presented in Table 1.

³ We assume that physical cultural capital is highly correlated with household income.

⁴ Several attempts to split PCA into two or three groups of variables were made, in order to check if different types of cultural capital could affect reading habits differently. However, the use of a single index of cultural capital provides the best model fit, so we it was the preferred option.

Table 1. Summary of descriptive statistics (N=10,319)

Variable	Description	Mean	SD	Min	Max
ER	Reading with an extrinsic motivation during week days (minutes)	38.52	94.08	0	1,200
IR	Reading with an intrinsic motivation during week days (minutes)	35.85	56.89	0	540
Male	Dummy variable for Male=1	0.52	0.49	0	1
Age	Respondent's age (in years)	40.04	12.93	16	90
House members	Number of individuals living in the household	3.30	1.23	1	13
Youths	Number of household members younger than 18 years old	1.64	2.30	0	27
Primary	Primary education (5-12 years old)	0.09	0.28	0	1
Secondary	Secondary education (12-16 years old)	0.3	0.46	0	1
GCE	General Certificate of Education (16-18 years old)	0.17	0.37	0	1
Vocational	Vocational Education and Training (VET) and Certificate of Higher Education (HND)	0.19	0.39	0	1
University	Tertiary education	0.24	0.42	0	1
Student	The respondent is a student	0.13	0.34	0	1
Self-employed	The respondent is self-employed	0.12	0.32	0	1
Employee	The respondent is employee	0.53	0.49	0	1
Unemployed	The respondent is unemployed	0.20	0.40	0	1
Library	Dummy variable for going to the library once a week	0.18	0.39	0	1
Television	Dummy variable for watching TV daily	0.88	0.31	0	1
Radio	Dummy variable for listening to the radio daily	0.64	0.47	0	1
Computer	Dummy variable for using computer daily	0.52	0.49	0	1
Gaming	Dummy variable for playing video games daily	0.05	0.22	0	1
Sports	Dummy variable for practising sport daily	0.21	0.40	0	1
Cultural participation	Number of attendances to cultural events in the last three months	4.68	9.58	0	344
PCA	Index of cultural capital at home	0.32	0.73	-1.89	3.31

To test whether the effect of ER on IR differs depending on education and gender, two interaction terms were included: one between the logarithm of extrinsic reading and university education (denoted by $Ln_ER_University$) and the other between the logarithm of extrinsic reading and the dummy variable that controls for gender (denoted by Ln_ER_Male).

Table 2 displays the average time (minutes) spent reading with an intrinsic motivation during week days depending on gender, age and educational level, both for the whole sample and the subsample of readers. We provide t-tests for equality of means regarding gender. Similarly, we provide one-way ANOVA tests to compare differences in intrinsic reading depending on age and education. This table compares unconditional

average values. We later propose a regression model to analyze the effect of each characteristic on the likelihood and intensity of reading with an intrinsic motivation.

Regarding gender, females tend to read significantly more than males. Average reading also depends on age, with those aged between 30 and 44 years old spending less time on reading. As for the education level, differences arise when considering the full sample, but not among readers.

Table 2. Average minutes spent reading depending on gender, age and education

		INTRINSIC READING (from Monday to Friday, in minutes)					
		ALL SAMPLE			READERS		
		n	mean	Comparison	n	mean	Comparison
GENDER	MALE	5463	28.27	t-test	2157	71.61	t-test
	FEMALE	4856	44.39	t=14.5 (p<0.001)	2814	76.60	t=2.81 (p=0.004)
AGE	UNDER 30	2505	37.68	ANOVA	1209	78.07	ANOVA
	30-44	3806	33.28	F=6.23	1802	70.30	F=6.7
	OVER 44	4008	37.16	(p=0.002)	1960	75.98	(p=0.001)
EDUCATION	PRIMARY	4120	23.71	ANOVA	1319	73.26	ANOVA
	GCE	1755	42.30	F=138.01	955	77.73	F=1.22
	VOCATIONAL	1964	35.76	(p<0.001)	966	72.71	(p=0.030)
	UNIVERSITY	2480	51.53		1731	73.83	
		10319	35.86		4971	74.43	

5. Estimation Method

In this section, we specify the econometric approach employed in the analysis. Our dependent variable, the time (minutes) devoted to reading with intrinsic motivation, is a mixture of zero and positive observations. Since it is only observed over certain value, we have a censoring problem, which requires special modelling issues. Ordinary Least Squares (OLS) regression will not yield consistent parameter estimates (Maddala 1999). The impact of censoring on the OLS depends on the proportion of censored observations which, in our sample, is around 57%. Although we could consider estimating a Tobit model, one restrictive assumption is that any variable specified in the model equally affects both the probability of observing a positive outcome and the conditional expected value. Therefore, Tobit cannot address the possible two-step nature of reading habits. Probably positive values of the dependent variable are not randomly selected from the population, so the results of the second stage regression may suffer from selection bias: a covariate can increase the probability of reading but reduce its intensity, and vice versa. To separately analyze the decision of whether to

read or not (participation) and how much time to allocate on it (intensity), we estimate a Heckman model.

The Heckman model

To overcome the possible selection bias, Heckman (1979) proposed a simple consistent estimation method that eliminates the specification error in the case of censored samples. Instead of assuming that each individual has a latent (unobserved) demand for reading with intrinsic motivation (IR_i^*) which is only observed when IR_i^* exceeds a certain threshold ξ , the Heckman model assumes that the dependent variable depends on a second latent variable ($Read_i^*$). Therefore, the following participation equation is defined:

$$Read_i^* = z_i\gamma + u_i \quad u_i \sim N(0,1)$$

$$Read_i = \begin{cases} 1 & \text{if } Read_i^* > 0 \\ 0 & \text{Otherwise} \end{cases} \quad (4)$$

The intensity equation (time spent on intrinsic reading) is given by:

$$\ln IR_i = \begin{cases} \ln IR_i^* & \text{if } Read_i = 1 \\ - & \text{Otherwise} \end{cases} \quad (5)$$

where $IR_i^* = \exp(X_i'\beta + \varepsilon_i)$, X_i is a vector of exogenous regressors and ε_i is a random error term with zero mean and σ standard deviation. Due to lognormality, it holds that $\ln T = \xi \neq 0^5$.

While the former estimates the probability of observing a positive outcome (participation equation), the latter estimates the quantity of time devoted to reading conditional on observing positive values (intensity equation). To estimate the model, it is necessary to assume that ε_i and u_i follow a normal bivariate distribution⁶ where ρ is the correlation between them (Cameron and Trivedi 2005).

$$\begin{pmatrix} \varepsilon_i \\ u_i \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & 1 \end{pmatrix} \right] \quad (6)$$

⁵ This transformation sets the dependent variable to missing as $T=0$ so it is necessary to substitute the censoring point T by another small value close to zero but different from it. This allows $\ln T$ to exist.

⁶ There have been some attempts to move away from the normality assumption (Martins 2001). Nonetheless, the empirical literature is dominated by the joint normal distribution assumption.

The observation mechanism assigns $\ln IR_i = \ln IR_i^*$ only if $Read_i = 1$, meanwhile $\ln IR_i$ is non-observable in case $Read_i = 0$. This means that we only observe positive values of the dependent variable if the individual has previously been cataloged as participant in the first stage⁷.

In this model, the decision whether or not to read and the amount of time devoted to it are assumed to be two separate choices that may be explained by a different set of determinants. The decision to read is assumed to be related to social and cultural factors, whereas the decision of how much to read depends more on individual's leisure time availability. Although most empirical studies use the same variables in both equations, it is appropriate to look for exclusion restrictions⁸, so that a regressor could generate nontrivial variation in the selection equation but does not affect the outcome variable (Cameron and Trivedi 2009). Accordingly, the variables contained in the vector Z are only considered in the participation equation.

6. Results

In this section we present the results obtained from the Heckman model. Table 3 reports the results of the parameter estimates⁹. We employed the one-step Full Information Maximum Likelihood estimator (FIML), as it has been proved to be more efficient than the alternative two-step procedure usually known as Heckit¹⁰. Standard errors are robust to heteroskedasticity. Assuming that reading is a two-step decision process, the rejection of the null hypothesis of no correlation between the error terms of the participation and intensity equations proves evidence of the preferability of the Heckman model compared to an alternative two-step independent model. Indeed, the statistical significance of the rho parameter (ρ) indicates that the two equations are interdependent.

Even though we model a "leisure-time demand" function, we do not include price among our independent variables. Since our database is a cross-section, an index of prices would be the same for the whole sample (there is no variation in the price offered to different groups of consumers). Anyway, we analyze the time spent on

⁷ As individuals were questioned about their reading habits within the last three months, we consider that when a respondent declares that he/she has not read during the considered period, we cannot deem him as an actual reader. This validates the election of the Heckman model which assumes that all the zeros are generated by the non-participation decision.

⁸ As the error terms are assumed to be correlated and the structure of the model is recursive, it is recommended that at least one variable in Z does not appear in X .

⁹ Estimates were conducted using the *heckman* module in Stata 14.

¹⁰ When the error assumptions are met, the FIML estimator will always be more efficient than the Heckman two-step alternative (Puhani 2000).

reading, not the demand for reading-materials. We assume that the price is not a key determinant of the time devoted to intrinsic reading. Interested readers have several (even free-access) alternatives, such as libraries, the internet, borrowing books from friends or family, and second-hand bookshops.

We consider three model specifications. Model A is a baseline model which controls for sociodemographic characteristics, labor situation, education level and regional dummies. Model B adds variables regarding leisure daily habits (only to the participation equation) as well as cultural participation and cultural capital at home (in the two equations). Finally, Model C incorporates interaction terms between extrinsic reading and university education and being a male (*Ln_ER_University* and *Ln_ER_Male*, respectively)¹¹.

Parameter estimates remain largely unchanged both in magnitude and significance across the three model specifications, indicating that our model is robust. Based on the log likelihood and the AIC and BIC criteria, model fit increases as we move from Model A to Model C, so we select Model C as our preferred model. In the models, we controlled for regional characteristics. These results are not reported, but available under request. Now we proceed to explain the results obtained.

We find that when people's obligations involve reading, that is, when they have an extrinsic motivation to read, they show a higher probability of being free-time readers, that is, to read with intrinsic motivation (column 5), although in a decreasing rate. However, extrinsic reading negatively affects the amount of time devoted to intrinsic reading, also in an attenuated path (column 6). These results point to a substitution pattern between extrinsic and intrinsic reading during week days. The more time reading at work or studying, the more probable is to be a reader during free time, but the less time is allocated to it. Spending a lot of time working or studying among texts might cause eyestrain, headache or reading-saturation, which might lead individuals to choose other sources of entertainment within free time.

Looking at the interaction terms, we find that extrinsic reading reduces the probability of intrinsic reading more for males than for females (column 5). In addition, extrinsic reading negatively affects the probability to read with intrinsic motivation for individuals with higher education (column 5). However, neither the interaction of gender nor that of

¹¹ To examine the potential existence of multicollinearity, we have computed the Variance Inflation Factor (VIF) after an OLS regression. All the values are below 10, which is normally taken as the threshold point, so we consider that there is no such problem in our model.

education with extrinsic reading are significant to explain the quantity of time allocated to intrinsic reading (column 6).

Regarding sociodemographic features, the probability of intrinsic reading increases with age (column 5), but no differences are found in the amount of time devoted to it¹². This result is contrary to Fernández-Blanco and Prieto-Rodríguez (2009) and Roe and Taube (2012), who agree that young people read less¹³. However, the mentioned authors analyze the number of books read whereas we consider all kind of reading, so results are not directly comparable. The fact that women are more likely to read is in line to previous findings (Logan and Johnston 2009; Clark 2011; Brozko et al. 2014). Conditional on being a reader, men read more, which is consistent with Elvestad and Blekesaume (2008).

Concerning house structure, the more members living at home, the less likely it is to read (column 5) but, within readers, they read more (column 6). Probably larger families lead to less time availability, especially when formed by elders to take care of. However, larger families could share reading materials or books recommendations, thus enhancing reading intensity. The presence of youths at home exerts a negative effect both on the probability and on the time spent on intrinsic reading, in line with Liu et al. (2017).

Continuing with labor situation, we set unemployed as the omitted category. Being employed (either self-employed or as an employee) has a negative effect on intrinsic reading-intensity (column 6) in comparison to that of unemployed peers. What is more, as in Molina et al. (2016), self-employed people show less likelihood of being readers (column 5). Being a student is not statistically significant in terms of the likelihood of being an intrinsic reader (column 5). However, it affects negatively on the amount of time on intrinsic reading (column 6).

¹² To allow for possible non-linearities, squared term was introduced. Since it was not statistically significant, it was finally omitted. We also defined a set of age groups. The results of this alternative specification are consistent: age does not explain the amount of time reading with intrinsic motivation. Results are available upon request.

¹³ Given that we have not considered neither retired nor those younger than 13 years old in our sample, age range is substantially narrower than usually, which could partially explain why age issues do not arise.

Table 3. Estimation results for intrinsic reading (Heckman) (N=10,319)

Dependent variables	Model A		Model B		Model C	
	(1) Participation Prob (Read=1)	(2) Intensity Ln_IR	(3) Participation Prob (Read=1)	(4) Intensity Ln_IR	(5) Participation Prob (Read=1)	(6) Intensity Ln_IR
Explanatory variables						
Ln_ER	0.187 ^c (0.031)	-0.352 ^c (0.025)	0.151 ^c (0.032)	-0.337 ^c (0.025)	0.157 ^c (0.037)	-0.337 ^c (0.029)
Ln_ER_Square	-0.023 ^c (0.006)	0.074 ^c (0.005)	-0.022 ^c (0.007)	0.073 ^c (0.005)	-0.024 ^c (0.008)	0.073 ^c (0.006)
Ln_ER_University					-5.6e-04 ^b (2.7e-04)	2.2e-04 (2e-04)
Ln_ER_Male					8.0e-04 ^b (2.8e-04)	-3.0e-04 (2.6e-04)
Male	-0.417 ^c (0.026)	0.082 ^c (0.028)	-0.468 ^c (0.027)	0.058 ^b (0.029)	-0.499 ^c (0.029)	0.073 ^b (0.033)
Age	0.005 ^c (0.001)	0.001 (0.001)	0.007 ^c (0.001)	0.001 (0.001)	0.007 ^c (0.001)	0.001 (0.001)
House members	-0.005 (0.013)	0.015 (0.011)	-0.019 (0.013)	0.017 (0.011)	-0.019 (0.013)	0.017 (0.011)
Youths	-0.019 ^c (0.007)	-0.018 ^c (0.006)	-0.012 ^a (0.007)	-0.019 ^c (0.006)	-0.013 ^a (0.007)	-0.018 ^c (0.006)
Student	0.191 ^c (0.057)	-0.169 ^c (0.049)	0.020 (0.059)	-0.120 ^b (0.049)	0.006 (0.060)	-0.113 ^b (0.048)
Selfemployed	-0.115 ^b (0.048)	-0.227 ^c (0.044)	-0.103 ^b (0.049)	-0.233 ^c (0.043)	-0.101 ^b (0.049)	-0.233 ^c (0.043)
Employee	0.020 (0.034)	-0.143 ^c (0.032)	0.024 (0.035)	-0.137 ^c (0.031)	0.022 (0.035)	-0.137 ^c (0.031)
GCE	0.455 ^c (0.038)	-0.152 ^c (0.039)	0.311 ^c (0.039)	-0.116 ^c (0.039)	0.304 ^c (0.040)	-0.113 ^c (0.038)
Vocational	0.393 ^c (0.037)	-0.136 ^c (0.039)	0.277 ^c (0.037)	-0.103 ^c (0.038)	0.271 ^c (0.037)	-0.101 ^c (0.038)
University	0.761 ^c (0.038)	-0.259 ^c (0.044)	0.515 ^c (0.041)	-0.199 ^c (0.043)	0.541 ^c (0.042)	-0.211 ^c (0.045)
Library			0.387 ^c (0.038)	-0.089 ^c (0.031)	0.391 ^c (0.038)	-0.091 ^c (0.030)
Television			-0.119 ^c (0.041)		-0.120 ^c (0.041)	
Radio			0.081 ^c (0.028)		0.082 ^c (0.028)	
Computer			0.271 ^c (0.030)		0.270 ^c (0.030)	
Gaming			0.058 (0.055)		0.054 (0.055)	
Sport			0.149 ^c (0.031)		0.146 ^c (0.031)	
PCA Cultural capital			0.186 ^c (0.027)	-0.023 ^b (0.011)	0.185 ^c (0.027)	-0.023 ^b (0.011)

Cultural participation			0.012 ^c (0.003)	-1.5e-04 (0.001)	0.012 ^c (0.003)	-2.1e-04 (0.001)
Constant	-0.563 ^c (0.085)	4.733 ^c (0.103)	-0.609 ^c (0.099)	4.625 ^c (0.109)	-0.589 ^c (0.100)	4.623 ^c (0.107)
<i>Rho (ρ)</i>	-0.680 ^c (0.053)		-0.628 ^c (0.071)		-0.633 ^c (0.067)	
Observations	10,319		10,319		10,319	
Censored	5,348		5,348		5,348	
Uncensored	4,971		4,971		4,971	
Regional dummies	YES		YES		YES	
Population size	YES		YES		YES	
Log Likelihood	-11,744.00		-11,471.29		-11,465.00	
BIC	24,116.44		23,672.68		23,697.07	
AIC	23,624.00		23,100.58		23,096.01	

Note: Robust standard errors in parentheses

^a, ^b and ^c denote statistical significance at 10, 5 and 1 percentage level

Results in italics are not statistically significant

With regard to educational level, all coefficients are statistically significant, both in the participation and intensity equations. The higher is the education level, the more probable to read with intrinsic motivation (column 5), in line with previous research (Scales and Rhee 2001). Nevertheless, higher educational levels are negatively related to intrinsic reading's intensity (column 6). Conditional on being a reader, lower educated individuals are more prone to spend more time reading than those with higher education. This seems to be quite counterintuitive but it could be explained through a lifelong reading-saturation. The accumulated extrinsic reading of academic years might link reading with an imposed duty instead of a source of amusement (Nielen et al. 2016).

With regard to the relation between reading and other leisure activities, our empirical results show that, when an individual declares that he/she watches TV every day, the likelihood of intrinsic reading decreases (column 5). This finding is consistent with previous studies (Knulst and Kraaykamp 1998; Mokhtari et al. 2009; Fernández-Blanco et al. 2017). Conversely, listening to the radio or using the computer act as complements, showing a positive effect on the probability of intrinsic reading. Similarly those who go to the library once a week are more likely to read with intrinsic motivation, in contrast to Fernández-Blanco and Prieto-Rodríguez (2009), who indicate that libraries are mainly a place where people go for studying. When the individual declares that practices sport, it is more prone to intrinsic reading. Conversely, playing video games does not affect reading's habits.

Another interesting result is the significant and positive effect of the total number of attendances to cultural visits and cultural events (*Cultural Participation*) on the

probability of intrinsic reading in the participation equation (column 5). Nonetheless, once the individual reads with intrinsic motivation, the accumulated cultural capital does not affect the amount of time devoted to it.

Lastly, the Principal Component Analysis for cultural capital of the household (*PCACAPITAL*) indicates that the more TVs, computers, musical instruments, books or electronic devices the individual has at home, the higher the likelihood of being an intrinsic reader, in line with Chiu and Chow (2010). By contrast, the effect of this variable on reading's intensity for those who actually read is negative.

To gain a better understanding of the importance of the different covariates in explaining participation and intensity decisions, we compute the corresponding marginal effects. Table 4 presents the average marginal effects (AME) on the probability of reading with intrinsic motivation (1) and on the expected value for those who actually read (2)¹⁴.

Table 4. Average Marginal Effects (N=10,319).

	(1) Participation equation (%)	(2) Intensity equation (%)
In_ER	2.94	-10.13
Male	-16.28	-9.04
Age	0.23	0.33
Cultural participation	0.40	0.36
PCA Cultural capital	6.22	-3.64

Note: All coefficients are statistically significant at the 1 % significance level

We find that when time spent in extrinsic reading rises 1 per cent, the probability of being an intrinsic reader increases by 2.9 per cent (column 1), whereas the amount of time spent reading for those who actually read decreases by 10.1 per cent (column 2). The conditional expected time of extrinsic reading for those who are readers is 9 per cent higher for women (column 2), who also display a higher probability of being readers (16.3 per cent according to column 1). A marginal increase in age and cultural participation translates into a 0.23 and 0.40 higher probability of being a reader, respectively (column 1). Finally, a one per cent increase in cultural capital at home increases the probability of being a reader by 6.2 per cent but reduces the amount of time spent on reading by 3.6 per cent.

¹⁴ For the sake of parsimony, we only report the AME for In PRW, Male, Age, PCA Cultural Capital and Cultural Participation. The rest of AME are available from the authors upon request.

To analyze whether the effects of extrinsic reading on intrinsic reading differ depending on the source of the motivation for the extrinsic reading, we run separate regressions for students, employed and unemployed people (see Annex 2). The most notable findings are the following. First, time spent on extrinsic reading only affects the probability of becoming a reader for employed people. Second, the magnitude of the negative effect of extrinsic reading on the amount of time devoted to intrinsic reading is higher for employed individuals, followed by unemployed and students. Third, whereas in Model C (Table 3) males read more time with intrinsic motivation than females, in the separate regressions we find no statistical gender differences. Fourth, the presence of youths at home and the practice of sports only affect the time spent on intrinsic reading for employed individuals. In sum, there are some differences in the relation between intrinsic and extrinsic behaviors depending on the labor situation.

7. Conclusions

This paper analyzes the relation between reading with intrinsic and extrinsic motivations. The analysis of reading habits is justified by its importance on the acquisition and improvement of cognitive and linguistic skills, as well as its influence on the development of human capital. Our database is drawn from the Cultural Habits and Practices Survey, conducted by the Ministry of Education and Culture in Spain in the period 2014-2015. The decisions of whether to read with intrinsic motivation or not, and the quantity of time devoted to it, can be understood as two separate and sequential choices. Therefore, we use a two-step Heckman model. As main determinants, we consider the time devoted to reading with extrinsic motivation, a set of socioeconomic characteristics, the practice of other leisure activities (which complement or substitute reading) and the quantity of cultural capital at home.

First, we study the determinants of the likelihood of being an intrinsic reader. Our findings show that being a woman, age, the educational level, the physical cultural capital at home, cultural participation, going to the library, listening to the radio, practicing sport and using the computer are positively related to the likelihood of reading with intrinsic motivation. On the contrary, the presence of youths at home and watching TV negatively affect the probability of intrinsic reading. Concerning labor status, self-employed people have a lower propensity to read than their unemployed peers.

Our estimations indicate that the relation between the exogenous need to read because one's job or studying (extrinsic reading) and the choice to read as an entertainment (intrinsic reading) is positive, although in a decreasing path. An individual who does extrinsic reading shows a higher probability of reading within his free time, maybe because those whose obligations require reading are more likely to be white collars, so their environment (workmates, job assignments) could probably generate conditions that encourage the intrinsic motivation for reading. In this line, those who are studying and have to read at high school or university are probably more exposed to different types of recreational reading too.

Regarding how much time to read for those catalogued as readers in the participation equation, the more extrinsic reading, the less time allocated to reading with intrinsic motivation. Possibly, extrinsic reading could imply perceiving reading as a humdrum activity, thus producing disutility.

Ongoing with sociodemographics, males tend to read more for intrinsic reasons, whereas the effect of age is not significantly different from zero. Besides that, self-employed and employees devote less time to intrinsic reading than unemployed people, maybe due to their lower time availability compared to those searching for a job. Surprisingly, highly educated people allocate less time to read with intrinsic motivation, perhaps because higher education implies lots of years of extrinsic reading. Finally, physical cultural capital negatively affects reading time, maybe because of the competition for time among different cultural activities at home.

The findings of this paper suggest that reading with extrinsic motivation exerts two different types of effects on reading with intrinsic motivation. On the one hand, there might be a "daily-saturation" phenomenon, which could explain the negative relation between extrinsic reading and the amount of time devoted to intrinsic reading. On the other hand, there might be a "vital-saturation", which could explain the negative relation between highly educated individuals and their lower amount of time allocated to intrinsic reading.

Limitations and further research

Considering all types of reading allows us to analyze reading from a wider approach than prior research, mainly focused on reading books or scholar texts. Nonetheless, additional work is needed to determine how different kinds of extrinsic reading affect intrinsic reading to obtain more precise findings and conclusions. In the same vein, another limitation is the fact that we do not know the kind of job or the area of

knowledge to which the individual belongs. Such information could be helpful to how different extrinsic reading affects reading with intrinsic motivation. Although the Cultural Habits and Practices Survey (CHPS) used in this study conveys information on individuals' characteristics, it lacks information on other people in the same household. Consequently, another limitation of the paper is the lack of information about the educational background and the reading habits of the household. Additionally, we also lack information on income. Although we proxy it by the number of devices at home, we acknowledge the limitations of this measure. Similarly, our database does not provide information on the time spent in other cultural activities, which could help defining complement or substitution patterns among cultural activities. Further research is needed to get to know the optimal balance between extrinsic and intrinsic reading.

REFERENCES

- Allington, R., and McGill-Franzen, A. (2003). The impact of summer setback on the reading achievement gap. *Phi Delta Kappan*, 85(1), 68–75.
- Andreassen, R. and Bråten, I. (2010). Examining the prediction of reading comprehension on different multiple-choice tests. *Journal of Research in Reading*, 33, 263–283.
- Becker, G.S. (1964). *Human Capital*. New York: Columbia University Press for the National Bureau of Economic Research.
- Becker, G.S. (1965). A Theory of the Allocation of Time. *Economic Journal*, 75, 493-517.
- Becker, G.S. (1996). *Accounting for Tastes*. Cambridge, Massachusetts. Harvard University.
- Becker, G.S. and Murphy, K.M. (1988). A Theory of Rational Addiction. *Journal of Political Economy*, 96(4), 675-700.
- Becker, M., McElvany, N. and Kortenbruck, M. (2010). Intrinsic and extrinsic reading motivation as predictors of reading literacy: A longitudinal study. *Journal of Educational Psychology*, 102(4), 773–785.
- Bortolussi, M., Dixon, P. and Sopcak, P. (2010). Gender and reading. *Poetics*, 38, 299-318.
- Bourdieu, P., and Passeron, J.C. (1990). *Reproduction in Education, Society and Culture*. 2nd ed. Sage, London.
- Bravishi, A., Slade, M.D. and Levy, B.R. (2016). A chapter a day: association of book reading with longevity. *Social Science and Medicine*, 164, 44-48.
- Brozko, W.C., Sulkunen, S., Shiel, G., Garbe, C., Pandian, A. and Valtin, R. (2014). Reading, Gender and Engagement. Lessons from five PISA countries. *Journal of Adolescent and Adult Literacy*, 57(7), 584-593.
- Cameron, A.C. and Trivedi, P.K. (2005). *Microeconometrics. Methods and applications*. Cambridge University Press.
- Cameron, A.C. and Trivedi, P.K. (2009). *Microeconometrics using Stata*. Stata Press.
- Carbonaro, W. (2007). The effects of education and cognitive skill on earnings: How much do occupations and jobs matter? *Research in Social Stratification and Mobility*, 25, 57-71.
- Chiu, M.M. and Chow, B.W.Y. (2010). Culture, motivation, and reading achievement: High school students in 41 countries. *Learning and Individual Differences*, 20, 579-592.
- Clark, C. (2011). *Setting the baseline: The National Literacy Trust's first annual survey into young people's reading 2010*. London: National Literacy Trust.
- Clark, C. and Foster, A. (2005). *Children's and young people's reading habits and preferences: The who, what, why, where and when*. London: National Literacy Trust.
- Clark, C., Osborne, S. and Akerman, R. (2008). *Young people's self-perceptions as readers: An investigation including family, peer and school influences*. London: National Literacy Trust.
- Clark, C. and Zoysa, S. (2011). *Mapping the interrelationships of reading enjoyment, attitudes, behaviour and attainment: an exploratory investigation*. National Literacy Trust, London.

- Coles, M. and Hall, C. (2002). Gendered readings: learning from children's reading choices. *Journal of Research in Reading*, 25(1), 96-108.
- Deaton, A.S. and Muellbauer, J. (1980). *Economics and consumer behavior*. Cambridge: Cambridge University Press.
- Elvestad, E. and Blekesaume, A. (2008). Newspaper Readers in Europe. A Multilevel Study of Individual and National Differences. *European Journal of Communication*, 23(4), 425-447.
- Fernández-Blanco, V. and Prieto-Rodríguez, J. (2009). Análisis de los hábitos de lectura como una decisión económica. *Estudios de Economía Aplicada*, 27(1), 235-260.
- Fernández-Blanco, V., Prieto-Rodríguez, J. and Suárez-Pandiello, J. (2017). A quantitative analysis of reading habits in Spain. *International Journal of Arts Management*, 19(3), 19-32.
- Gil-Flores, J. (2009). Hábitos y Actitudes de las Familias Hacia la Lectura y Competencias Básicas del Alumnado. *Revista de Educación*, 350, 301-322.
- Gottfried, A.W., Schlackman, J., Gottfried, A.E., and Boutin-Martinez, A.S. (2015). Parental provision of early literacy environment as related to Reading and educational outcomes across the academic lifespan. *Parenting*, 15(1), 24-38.
- Goux, D., Gurgand, M. and Maurin, E. (2017). Reading enjoyment and reading skills: Lessons from an experiment with first grade children. *Labour Economics*, 45, 17-25.
- Guthrie, J.T. and Davis, M.H. (2003). Motivating struggling readers in middle school through an engagement model of classroom practice. *Reading & Writing Quarterly*, 19(1), 59-85.
- Guthrie, J.T., Hoa, A.L.W., Wigfield, A., Tonks, S.M., Humenick, N. M., and Littles, E. (2007). Reading motivation and reading comprehension growth in the later elementary years. *Contemporary Educational Psychology*, 32(3), 282-313.
- Guthrie, J.T. and Klauda, S.L. (2012). Making textbook reading meaningful. *Educational Leadership*, 69(6), 64-68.
- Heckman, J.J. (1979). Sample selection bias as a specification error. *Econometrica*, 47, 153-161.
- Heckman, J.J. and Masterov, D.V. (2007). The Productivity Argument for Investing in Young Children. *Review of Agricultural Economics*, 29(3), 446-493.
- Hicks, J.R. (1939). Value and capital. *The Economic Journal*, 49(194), 294-300.
- Ivey, G. and Broaddus, K. (2001). 'Just Plain Reading': A survey of what makes students want to read in middle school classrooms. *Reading Research Quarterly*, 36, 350-377.
- Kalb, G. and Van Ours, J.C. (2015). Reading to young children: A head-start in life? *Economics of Education Review*, 40, 1-24.
- Kloosterman, R., Notten, N., Tolsma, J., Kraaykamp, G. (2011). The Effects of Parental Reading Socialization and Early School Involvement on Children's Academic Performance: A Panel Study of Primary School Pupils in the Netherlands. *European Sociological Review*, 26(2), 1-16.
- Knulst, W. and Kraaykamp, G. (1998). Trends in Leisure Reading: Forty Years of Research on Reading in the Netherlands. *Poetics*, 26(1), 21-41.
- Knulst, W. and Van den Broek, A. (2003). The readership of books in times of de-reading. *Poetics*, 31, 213-233.
- Koolstra, C.M. and van der Voort, T.H.A. (1996). Longitudinal effects of television on children's leisure-time reading: a test of three explanatory models. *Human Communication Research*, 23, 4-35.
- Kooreman P. and Kapteyn A. (1987). A disaggregated Analysis of the Allocation of Time within the Household. *Journal of Political Economy*, 95(2), 223-249.
- Kraaykamp, G. (2003). Literary socialization and reading preferences. Effects of parents, the library, and the school. *Poetics*, 31, 235-257.
- Lauf, E. (2001). Research Note: The Vanishing Young Reader. Sociodemographic Determinants of Newspaper Use as Source of Political Information in Europe, 1980-98. *European Journal of Communication*, 16(2), 233-243.
- Liu, T., Lin, C., Huang, K. and Chen. Y. (2017). Effects of noise type, noise intensity, and illumination intensity on reading performance. *Applied Acoustics*, 120, 70-74.
- Loan, F.A. (2012). Impact of the Internet Surfing on Reading Practices and Choices. *Webology*, 9(1), Article 94. Available at: <http://www.webology.org/2012/v9n1/a94.html>
- Logan, S. and Johnston, R.S. (2009). Gender differences in reading: Examining where these differences lie. *Journal of Research in Reading*, 32(2), 199-214.
- Logan, S., Medford, E. and Hughes, N. (2011). The importance of intrinsic motivation for high and low ability readers' reading comprehension performance. *Learning and Individual Differences*, 21(1), 124-128.

- Maddala, G.S. (1999). *Limited-dependent and qualitative variables in econometrics*. Cambridge, UK: Cambridge University Press.
- Martins, M.F.O. (2001). Parametric and semiparametric estimation of sample selection models: an empirical application to the female labour force in Portugal. *Journal of Applied Econometrics*, 16(1), 23–39.
- McCain, R. (1979). Reflection on cultivation of taste. *Journal of Cultural Economics*, 3(1), 30-52.
- McGeown, S.P., Goodwin, H., Henderson, N. and Wright, P. (2012a). Gender differences in reading motivation: Does sex or gender identity provide a better account? *Journal of Research in Reading*, 35(3), 328–336.
- McGeown, S.P., Norgate, R. and Warhurst, A. (2012b). Exploring intrinsic and extrinsic reading motivation among very good and very poor readers. *Educational Research*, 54(3), 309–322.
- McGeown, S.P. (2015). Sex or gender identity? Understanding children’s reading choices and motivation. *Journal of Research in Reading*, 38(1), 35-46.
- McIntosh, S. and Vignoles, A. (2001). Measuring and assessing the impact of basic skills on labour market outcomes. *Oxford Economics Papers*, (3), 453-481.
- McQuillan, J. and Au, J. (2001). The effect of print access on reading frequency. *Reading Psychology*, 22, 225–248.
- Metsäpelto, R., Silinskas, G., Kiuru, N., Poikkeus, A., Pakarinen, E., Vasalampi, K., Lerkkanen, M. and Nurmi, J. (2017). Externalizing behavior problems and interest in reading as predictors of later reading skills and educational aspirations. *Contemporary Educational Psychology*, 49, 324-336.
- Mokhtari, K., Reichard, C.A. and Gardner, A. (2009). The Impact of Internet and Television Use on the Reading Habits and Practices of College Students. *Journal of Adolescent and Adult Literacy*, 52(7), 609-619.
- Mol, S.E. and 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.
- Molina, J.A., Campaña, J.C. and Ortega, R. (2016). What do you prefer for a relaxing time at home: reading, watching TV or listening to the radio? *Applied Economic Letters*, 23(18), 1278-1284.
- Morgan, P.L., and Fuchs, D. (2007). Is There a Bidirectional Relationship between Children's Reading Skills and Reading Motivation? *Exceptional children*, 73(2), 165-183.
- Mullan, K. (2010). Families that read: a time diary analysis of young people’s and parents’ reading. *Journal of Research in Reading*, 33, 414–430.
- Murnane, R.J., Willett, J.B., Duhaldeborde, Y. and Tyler, J.H. (2000). How Important Are the Cognitive Skills of Teenagers in Predicting Subsequent Earnings? *Journal of Policy Analysis and Management*, 19(4), 547-568
- Nielen, T.M.J., Mol, S.E., Sikkema-de Jong, M.T. and Bus, A.G. (2016). Attentional bias toward reading in reluctant readers. *Contemporary Educational Psychology*, 46, 263-271.
- Notten, N. (2011). *Parents and the media. Causes and consequences of parental media socialization*. Doctoral dissertation. Nijmegen. The Netherlands: Radboud University Nijmegen.
- Notten, N., Kraaykamp, G., Konig, R.P. (2012). Family media matters: unraveling the intergenerational transmission of reading and television tastes. *Sociological Perspectives*, 55, 683–706.
- Pfost, M., Schiefer, I. M. and Artelt, C. (2016). Intergenerational continuity in attitudes toward reading and reading behavior. *Learning and Individual Differences*, 51, 179-188.
- Puhani, P.A. (2000). The Heckman correction for sample selection and its critique. *Journal of Economic Surveys*, 14, 53–68.
- Retelsdorf, J., Köller, O., and Möller, J. (2011). On the effects of motivation on reading performance growth in secondary school. *Learning and Instruction*, 21(4), 550–559.
- Roe, A. and Taube, K. (2012). *To read or not to read – That is the question. Reading Engagement and Reading Habits in a Gender Perspective*. In: Egelund, N. (Ed.), Northern Lights on PISA 2009 – Focus on Reading. Nordic Council of Ministers, Copenhagen.
- Rosenzweig, E.Q., and Wigfield, A. (2017). What if reading is easy but unimportant? How students’ patterns of affirming and undermining motivation for reading information texts predict different reading outcomes. *Contemporary Educational Psychology*, 48, 133-148.
- Scales, A.M., and Rhee, O. (2001). Adult reading habits and patterns. *Reading Psychology*, 22, 175–203.
- Schaffner, E., Schiefele, U. and Ulferts, H. (2013). Reading Amount as a Mediator of the Effects of Intrinsic and Extrinsic Reading Motivation on Reading Comprehension. *Reading Research Quarterly*, 48(4), 369-385.

- Schiefele, U., Schaffner, E., Möller, J. and Wigfield, A. (2012). Dimensions of Reading Motivation and Their Relation to Reading Behavior and Competence. *Reading Research Quarterly*, 47(4), 427-463.
- Schiefele, U., Stutz, F. and Schaffner, E. (2016). Longitudinal relations between reading motivation and reading comprehension in the early elementary grades. *Learning and Individual Differences*, 51, 49-58.
- Schutte, N.S. and Malouff, J.M. (2004). University student reading preferences in relation to the big five personality dimensions. *Reading Psychology*, 25(4), 273-295.
- Sikora, J., Evans, M.D.R. and Kelley, J. (2019). Scholarly culture: How books in adolescence enhance adult literacy, numeracy and technology skills in 31 societies. *Social Science Research*, (77), 1-15.
- Siddiqi, A., Subramanian, S.V., Berkman, L., Hertzman, C. and Kawachi, I. (2007). The welfare state as a context for children's development: a study of the effects of unemployment and unemployment protection on reading literacy scores. *International Journal of Social Welfare*, 16, 314-325.
- Soemer, A. and Schiefele, U. (2019). Text difficulty, topic interest, and mind wandering during reading. *Learning and Instruction*, 61, 12-22.
- Stanovich, K.E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford.
- Stigler, G.J. and Becker, G.S. (1977). De Gustibus Non Est Disputandum. *American Economic Review*, 67(2), 76-90.
- Stokmans, M.J.W. (1999). Reading attitude and its effect on leisure time reading. *Poetics*, 26, 245-261.
- Stutz, F., Schaffner, E. and Schiefele, U. (2016). Relations among reading motivation, reading amount, and reading comprehension in the early elementary grades. *Learning and Individual Differences*, 45, 101-113.
- Suggate, S., Schaughency, E., McAnally and H., Reeseb, E. (2018). From infancy to adolescence: The longitudinal links between vocabulary, early literacy skills, oral narrative, and reading. *Cognitive Development*, 47, 82-95
- Tepper, S. (2000). Fiction reading in America: explaining the gender gap. *Poetics*, 27, 255-275.
- Torppa, M., Eklund, K., Sulkunen, S., Niemi, P. and Ahonen, T. (2018). Why do boys and girls perform differently on PISA Reading in Finland? The effects of reading fluency, achievement behaviour, leisure Reading and homework activity. *Journal of Research in Reading*, 41(1), 122-139.
- Troyer, M., Kim, J.S., Hale, E., Wantchekon, K.A. and Armstrong, C. (2019). Relations among intrinsic and extrinsic reading motivation, reading amount, and comprehension: a conceptual replication. *Reading and Writing*, 32, 1197-1218.
- Verboord, M. (2005). Long-term effects of literary education on book-reading frequency: an analysis of Dutch student cohorts 1975-1998. *Poetics*, 33, 320-342.
- Verboord, M. and Van Rees, K. (2003). Do changes in socialization lead to a decline in reading level? How parents, literary education, and popular culture affect the level of books read. *Poetics*, 31, 283-300.
- Vergheze, J., Lipton, R.B., Katz, M.J., Hall, C.B., Derby, C.A., Kuslansky, G., et al. (2003). Leisure activities and the risk of dementia in the elderly. *The New England Journal of Medicine*, 348, 2508-2516.
- Wang, J.H.Y., and Guthrie, J.T. (2004). Modeling the effects of intrinsic motivation, extrinsic motivation, amount of reading, and past reading achievement on text comprehension between US and Chinese students. *Reading Research Quarterly*, 39, 162-186.
- Wang, M.T., and Eccles, J.S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction*, 28, 12-23.
- Wollscheid, S. (2014). The impact of the leisure reading behaviors of both parents on children's reading behavior: investigating differences between sons and daughters. *Poetics*, 45, 36-54.

ANNEX 1: PRINCIPAL COMPONENT ANALYSIS FOR CULTURAL CAPITAL

Following the practice of Fernández-Blanco and Prieto-Rodríguez (2009), we conducted a PCA for cultural capital in order to proxy household income. We included the number of books, e-books, encyclopedias, vinyl CDs, DVDs, CDs and Blue Rays, and other audio devices, as well as the number of computers. Descriptive statistics of each variable can be seen on Table A.

The coefficient for the first factor is positive for all considered variables and, consequently, first factor analysis predicts that the cultural capital is positively correlated with the variables that we contemplate. The eigenvalue of the first factor is 1.025 and it explains the 62.68 per cent of the total variance.

Table A1. Principal Component Analysis (Cultural capital) (N=10,319)

Variable	Mean	Std. Dev.	Factor1
Number of books	144.94	375.19	0.4365
Number of eBooks	92.30	1369.67	0.1327
Number of encyclopedias	3.02	8.10	0.2176
Number of vinyls	18.80	134.75	0.4388
Number of DVD, CD	66.09	162.94	0.6368
Number of other audio devices	95.35	905.23	0.1462
Number of computers	1.40	1.24	0.3873

Number of params = 21

Retained factors = 6

LR test: independent vs. saturated: Prob>chi2(325) = 0

ANNEX 2: HECKMAN ESTIMATES BY LABOR STATUS

Table A2. Estimation results for intrinsic reading (Heckman) by labor status

Dependent variable	Students (N=1,397)		Employed (N=6,770)		Unemployed (N=2,152)	
	(1)	(2)	(3)	(4)	(5)	(6)
	Participation Prob (Read=1)	Intensity Ln_IR	Participation Prob (Read=1)	Intensity Ln_IR	Participation Prob (Read=1)	Intensity Ln_IR
Explanatory variables						
In PRW	0.0330 (0.0660)	-0.2655 ^c (0.0580)	0.1836 ^c (0.0543)	-0.4158 ^c (0.0391)	0.1879 (0.1205)	-0.3595 ^c (0.0835)
Ln_ER_Square	0.0018 (0.0130)	0.0547 ^c (0.0113)	-0.0326 ^b (0.0127)	0.0938 ^c (0.0088)	-0.0326 (0.0265)	0.0794 ^c (0.0183)
Ln_ER_University	-0.0014 ^a (0.0007)	-0.0001 (0.0005)	-0.0005 (0.0003)	-0.0003 (0.0002)	0.0002 (0.0007)	-0.0004 (0.0005)
Ln_ER_Male	0.0003 (0.0005)	-0.0004 (0.0005)	0.0013 ^c (0.0003)	-0.0004 (0.0002)	-0.0003 (0.0008)	-0.0000 (0.0009)
Male	-0.3363 ^c (0.0915)	0.0724 (0.0885)	-0.5264 ^c (0.0354)	0.0595 (0.0374)	-0.5011 ^c (0.0650)	-0.0121 (0.0870)
Age	0.0080 (0.0081)	0.0036 (0.0077)	0.0071 ^c (0.0016)	0.0009 (0.0012)	0.0081 ^c (0.0027)	0.0012 (0.0023)
House members	-0.0613 (0.0395)	0.0684 ^a (0.0352)	-0.0049 (0.0166)	0.0139 (0.0131)	-0.0269 (0.0295)	-0.0036 (0.0252)
Youths	-0.0091 (0.0268)	-0.0033 (0.0249)	-0.0163 ^a (0.0086)	-0.0215 ^c (0.0068)	-0.0197 (0.0162)	-0.0134 (0.0140)
GCE	0.1388 (0.0898)	-0.1301 ^a (0.0749)	0.3433 ^c (0.0510)	-0.0477 (0.0518)	0.3681 ^c (0.0973)	-0.0575 (0.0977)
Vocational	0.0091 (0.1242)	-0.0913 (0.1131)	0.2883 ^c (0.0462)	-0.0499 (0.0454)	0.3747 ^c (0.0853)	-0.0618 (0.0953)
University	0.1822 (0.1924)	-0.0857 (0.1493)	0.6197 ^c (0.0507)	-0.1857 ^c (0.0537)	0.2815 ^c (0.1062)	0.0252 (0.0979)
Library	0.1714 ^b (0.0709)	-0.0906 (0.0592)	0.4770 ^c (0.0528)	-0.1079 ^c (0.0395)	0.5313 ^c (0.0904)	0.0575 (0.0821)
Television	-0.1534 ^a (0.0827)		-0.0923 ^a (0.0511)		-0.2780 ^b (0.1291)	
Radio	0.1031 (0.0634)		0.0628 ^a (0.0351)		0.1690 ^c (0.0625)	
Computer	0.1765 ^b (0.0708)		0.2449 ^c (0.0343)		0.4475 ^c (0.0741)	
Gaming	0.0762 (0.0913)		0.0359 (0.0843)		-0.0019 (0.1279)	
Sport	-0.0805 (0.0659)		0.2080 ^c (0.0401)		0.1141 (0.0759)	
PCA Cultural capital	0.1117 ^a (0.0623)	-0.0529 (0.0364)	0.1738 ^c (0.0310)	-0.0085 (0.0119)	0.2494 ^c (0.0868)	-0.0205 (0.0498)
Cultural participation	0.0066 ^a (0.0034)	0.0018 (0.0026)	0.0112 ^c (0.0042)	-0.0002 (0.0019)	0.0341 ^c (0.0077)	0.0006 (0.0025)
Constant	-0.1367 (0.2837)	4.5181 ^c (0.2287)	-0.6241 ^c (0.1215)	4.3946 ^c (0.1210)	-0.6524 ^c (0.2276)	4.3601 ^c (0.2725)
<i>Rho</i> (ρ)	-0.7955 ^c (0.1759)		-0.6237 ^c (0.1260)		-0.1479 0.2639	
Observations	1,397		6,770		2,152	
Censored	651		3,441		1,256	
Uncensored	746		3,329		896	
Regional dummies	YES		YES		YES	

<i>Population size</i>	YES	YES	YES
<i>Log Likelihood</i>	-1,679.512	-7,494.952	-2,179.148
<i>BIC</i>	3,513.025	15,143.9	4,512.295
<i>AIC</i>	3,916.665	15,669.06	4,949.205

Note: Robust standard errors in parentheses

^a, ^b and ^c denote statistical significance at 10, 5 and 1 percentage level

Results in italics are not statistically significant