Received: 10 June 2020 | Revised: 28 July 2022 | Accepted: 30 August 2022

DOI: 10.1111/boer.12373

BULLETIN OF ECONOMIC RESEARCH

RESEARCH ARTICLE

Good students, avid readers: The cost of academic success

Antonello E. Scorcu[1] | **Laura Vici**[2] | **Roberto Zanola**[3]

[1]Department of Economics, University of Bologna, Bologna, Italy

[2]Department of Statistical Sciences, University of Bologna, Bologna, Italy

[3]Institute of Public Policy and Public Choice, University of Eastern Piedmont, Alessandria, Italy

Correspondence
R. Zanola, Institute of Public Policy and Public Choice, University of Eastern Piedmont, Via Cavour 84, 15100 Alessandria, Italy.
Email: roberto.zanola@uniupo.it

Abstract

The effects of reading habits on academic performances have been carefully investigated, but little is known about the effects of academic achievements on students' leisure reading. This paper investigates this issue by estimating the effects of academic achievements, proxied by the number of exams passed, on leisure reading, measured by the number of leisure books read in a year. Using an online survey submitted to the students at the University of Bologna, Italy, we adopt a two-step control-function technique to control for endogeneity. The empirical evidence suggests the existence of a negative relationship between students' academic achievements and the time devoted to leisure reading. This result holds for students of different fields of study and is stronger for male students. The Blinder–Oaxaca decomposition supports the existence of a gender-specific idiosyncratic effect.

KEYWORDS

academic performances, control function, decomposition, gender, leisure reading

JEL CLASSIFICATION

C3, I2, Z1

1 | INTRODUCTION

Reading skills are a crucial determinant of individual human capital accumulation. Higher reading skills enhance capabilities and competencies, improve the long-run prospects of school and academic success (Tomul & Polat, 2013), and lead to permanent higher incomes and better health prospects (Brunello et al., 2017; McIntosh & Vignoles, 2001).

The literature distinguishes two basic motivations for the accumulation of reading skills (McGeown et al., 2012; Suárez-Fernández & Boto-García, 2019). *Intrinsic* motivations are linked to leisure reading and induce a strong positive effect on reading skill accumulation. Reading driven by *extrinsic* motivations refer to educational and professional subjects and lead to weaker benefits, even if they are specifically related to human capital accumulation (Troyer et al., 2019). Although there exists a number of contributions focusing on intrinsic or extrinsic motivations, the way in which these motivations are connected has not been explicitly investigated. An exception is Suárez-Fernández and Boto-García (2019) who used a sample representative of the Spanish population and showed that extrinsic motivations increase the likelihood of leisure reading, but reduce the time devoted to it. Whereas several studies focused on the impact of leisure reading on overall academic achievements (Clark & Rumbold, 2006; Whitten et al., 2016), a reverse link could emerge.

In fact, students could devote more time to educational reading (extrinsic motivations), and this could jeopardize the time devoted to leisure reading (intrinsic motivations). However, disregarding that studying for a longer time could reduce leisure reading, it might lead to an underestimation of the positive overall effects of reading habits. This could have far-reaching consequences, as any reduction in leisure reading during the university period could potentially affect the previously acquired reading habits, mining the lifelong benefits of leisure reading (Bavishi et al., 2016).

Our approach focuses on this overlooked reverse link, suggesting a novel perspective in the reading debate. We investigate how extrinsic reading motivations could influence intrinsic motivations among university students, taking into account their characteristics and attitudes.

To this aim, in 2013 we conducted an online survey among students enrolled at the University of Bologna (Italy). We collected 20,772 answers and performed a two-step control-function technique to investigate the relationship between academic achievement, measured by the average number of passed exams within one academic year (which proxies the time devoted to educational reading), and leisure reading, measured by the number of non-academic books read in the last year. Based on our sample, the empirical evidence suggests that stronger and longer involvement in studying, associated with better academic performances, are likely to reduce leisure reading, regardless of whether students are avid or reluctant readers.

As the literature highlights gender-related differences in reading habits (Cobb-Clark & Moschion, 2017; Hochweber & Vieluf, 2018), we investigate this issue by also performing separate gender regressions.

In order to evaluate whether reading differences are due to the different distribution of the characteristics in the two subsamples or to gender idiosyncratic effects, we also apply an extension of the Blinder–Oaxaca decomposition method. The decomposition analysis shows the occurrence of an idiosyncratic gender effect due to the different impacts of the reading determinants on male and female students.

The rest of the paper is organized as follows. The next section shortly illustrates the empirical strategy, while Section 3 presents the data. Section 4 displays the results of the two-step

control-function approach applied to different samples and the Blinder-Oaxaca decomposition. Section 5 discusses the main empirical findings and concludes.

2 | EMPIRICAL STRATEGY

Academic achievements are time-consuming and require students to dedicate less time to other activities, such as leisure reading. Let Y be the number of non-academic books read in the last year, our measure of leisure reading. We assume that reading non-academic books in the last year follows a Poisson distribution:

$$P(Y = y/\mu) = e^{-\mu}\frac{\mu^{y}}{y!} \tag{1}$$

where $\mu = E(Y) = \text{var}(Y)$, $\mu > 0$, and $y_i = 0,1,2,3,\ldots$.

The individual realizations depend on both academic achievements,[1] measured by the average number of passed exams within an academic year, *exams*, and a vector of control covariates, $\boldsymbol{x}$:

$$\mu_i = E(Y_i) = \left(\beta_i exams_i + x_i'\beta_x\right) \tag{2}$$

where βs are unknown parameters estimated from data. Hence, the Poisson regression for individual i is:

$$Y_i = \frac{\mu^{Y}e^{-\mu}}{Y!} + \varepsilon_i \tag{3}$$

As Y and *exams* are likely to be jointly determined, we control for endogeneity to avoid the bias arising from the standard single-equation estimator. We thus use a control-function approach, a two-step procedure that postulates the existence of a structural relationship between the endogenous variable and other exogenous regressors (Wooldridge, 2010). In the first stage of the procedure, a linear model is specified and tested for the endogenous variable:

$$exams_i = x_i'\gamma_x + z_i'\gamma_z + u_i \tag{4}$$

where z indicates the instruments, a vector of exogenous variables affecting *exams* and not affecting Y by any other means than the endogenous variable, either directly or indirectly.

We specify the error term as follows:

$$\varepsilon_i = \rho u_i + \nu_i \tag{5}$$

with $\nu_i \sim [0,\sigma^2{}_{\nu}]$, $\varepsilon_i \sim [0,\sigma^2{}_{\varepsilon}]$. ν_i is independent of ε_i. *exams* is exogenous if $\rho = 0$, otherwise it is endogenous, being correlated with ε.

Together with the exogenous covariates, both *exams* and the residuals of the auxiliary regression [4] are used as regressors in the second-step equation [3]. If, as expected, ρ is statistically different

[1] Our survey mainly focuses on reading habits and includes a limited number of questions. Although a relationship between studying and other leisure activities (e.g., sport practices, listening to or playing music, etc.) is expected to occur, the specific nature of the survey prevents us to explore this issue.

from zero, the variance-covariance matrix (VCE) must be estimated through bootstrap to get bias-corrected confidence intervals.[2]

3 | DATASET

Our empirical investigation relies upon a survey we conducted among students enrolled at the University of Bologna, Italy (UniBo) in May–July 2013, made available only recently, due to Privacy Statement.[3] After having designed the survey questionnaire, we pretested it and did a pilot. We collected information on 27,616 students, more than 12% of the UniBo student population, a remarkable figure compared to similar surveys. Participation in the survey was voluntary, and this prevented the adoption of any predetermined sample design. Hence, caution is required for any generalization of our results. A number of cases were discarded due to incomplete or inconsistent answers, and a total of 20,772 respondents were considered in this study.

The survey focused on studying and reading practices. It collected information about the cultural, individual, and socio-demographic characteristics of students (age, gender, nationality, parents' education level, professional status, size of the family library, place of residence), educational records (high school final mark), cultural attitudes (non-academic reading habits) and study practices (importance of different learning tools, peers' studying behavior and practices, etc.). The survey also collected information about academic characteristics, such as the field of study, the year of enrollment and the number of passed exams.

As discussed above, our goal is to capture the effect that the time allocated to extrinsic reading to get better academic performances has on intrinsic reading habits. We use the number of books read for leisure in one year as a proxy for students' reading attitude, a measure frequently adopted in the literature (Canoy et al., 2006). Several variables could be used to measure academic achievements, each of them capturing different aspects of academic performances, for example, the average exam mark shows the quality of academic performances; the number of exams passed, controlled for the year of enrollment, depicts efficiency performances; etc. However, in order to avoid sensitive questions like the average exam mark, which could negatively affect the survey participation rate, induce response biases, and increase the cases of non-responses (De Schrijver, 2012; Tourangeau & Yan, 2007), we only collected information on the following academic performance measure: the number of exams passed in one academic year. The anonymity of the questionnaires impeded us from merging the survey data with private academic information collected by the university. Moreover, we avoided asking any sensitive information that could bias answers and affect the outcome. Therefore, the choice of the number of exams passed per year as an academic performance variable, together with the adoption of an anonymous online survey, allows for partially controlling for the social desirability bias (Larson, 2019).

As discussed in the previous section, the first step of the control-function approach requires the identification of a vector of instrumental variables, which define the learning technique of the student.

We have identified three instruments that reflect the perception of the importance that certain study tools have in raising academic achievements: *lect_imp*, the usefulness of attending lectures;

[2] The sandwich estimator is valid only with independent observations.

[3] All the students enrolled at the UniBo received an email by the University staff, inviting them to fill in an anonymous questionnaire. The email contained a link to the survey. With intervals of 2 weeks, two recalls were sent to the students who did not completed the survey. There was full compliance with privacy requirements.

notes_imp, the usefulness of studying on notes students took during lectures; and *hand_imp*, the usefulness of studying on handouts teachers provided as course material. These five-level Likert scale instruments define students' study strategies, which are linked to academic performance. The perceived importance of these instruments reflects the tools students use to study, which are supposed to increase the effort spent in the learning process and the probability of passing exams (thus, affecting the variable exam). However, these tools are not expected to influence leisure reading by any other means than the endogenous variable exam, which can proxy the effort and time spent studying. In the next section, we check for the non-exogeneity of these instruments through the Sargan-Hansen's *J* test.

Previous studies suggest that several other determinants affect leisure reading and should be included in the vector $\boldsymbol{x}$ of exogenous regressors. We consider four sets of independent variables: demographic characteristics, university context, student characteristics, and cultural capital.

3.1 | Demographic characteristics

Previous studies have unanimously pointed out that more females indulge in leisure reading than males (Swalander & Taube, 2007). Additionally, females show a more positive attitude to reading (Logan & Johnston, 2009) and a preference for reading a larger variety of book genres compared to males (Clark et al., 2008). Hence, we include *gender*, a dummy variable that is equal to 1 for female students, and 0 otherwise. We also perform two separate regressions for the male and female subsamples. Other variables usually considered important are *age* and its square, *age_sq*, because of the (possibly non-linear) influence of the time assigned to specific educational, cultural, leisure activities and reading preferences (Stockmans, 1999). Also, the distinction between domestic and foreign students is expected to affect both the attitudes to read for academic purposes and for leisure and to pass exams. Moreover, in another country, the literature readily available is mostly available in the hosting country's language. However, often students use books also to learn a foreign language. Moreover, the innovation in publishing allows reading content on alternative digital devices and in several languages.

3.2 | University context

The faculty of enrollment tends to identify a specific set of preferences related to academic and cultural subjects and approaches to studying. For example, a student enrolled in a BA in Classical Studies is a priori more likely to spend more time reading, for both academic and leisure purposes, than a student enrolled in a BA in chemistry, because of the different learning techniques and use of laboratories. To capture these differences, we use the dummy variable *soc_hum*, which is equal to 1 if the student is enrolled in social science, humanities, classical studies, foreign languages, and literature degrees, and zero otherwise. We also consider specific dummies for the different fields of studies (social sciences; technology and applied sciences; mathematical, physical and natural sciences; humanities and foreign languages; medicine, pharmacy and sport sciences) to introduce interaction terms with gender. The distance between the university buildings and students' residences may also affect the time devoted to reading and studying. For example, daily commuters spend time moving to and from the university. During this time, they can do only a few activities,

that is, talking with friends and mates, playing with electronic devices or, possibly, reading.[4] We introduce a dummy variable, *commuter*, which is equal to 1 if the student is a commuter, and zero otherwise. Similarly, non-resident students living in university residences or renting flats spend part of their time cooking and cleaning, differently from residents and commuters, who often still live with their parents. We, therefore, introduce another dummy variable, *non_resid*, equal to 1 for non-resident students and 0 otherwise.

3.3 | Student characteristics

Proficient students tend to read more for leisure and, following a virtuous circle, they are more likely to perform better and reach higher educational achievements (van Schooten & de Glopper, 2003). Hence, we use the high-school final mark, *mark*, as a proxy for student ability. As this variable refers to the high school performance, it avoids any potential endogenous effect related to undergraduate years. We expect *mark* to positively affect both the number of exams passed and the number of leisure books read in a year. Furthermore, the attitude to study on textbooks or complementary academic books often captures a higher propensity to read books, even outside the university context. Hence, two 5-level Likert scale variables are introduced: the attitude to study on textbooks, *text*, and complementary academic books, *other_books*.

3.4 | Cultural capital

In the literature, the stock of personal cultural capital is usually linked to parental education and family social status. Parents can influence students' attitudes toward reading, being a model for their children (Tramonte & Willms, 2010). We measure parental education by introducing a set of dummy variables capturing the maximum education level attained by students' parents (primary school or less, middle school, high school diploma and university, master of doctoral degree). Moreover, the attitude to read for leisure is also influenced by the cultural environment (Sieben & Lechner, 2019), proxied by the number of books (divided by 100) in the family house, *library* (OECD, 2002).

Table 1 displays the descriptive statistics. Female students constitute 59.6% of our sample. Respondents are, on average, 22 years old. Only 45% of respondents live in the same province of the University Campus where they study, 37% live in a different province and/or region and commute to attend lectures, whereas the remaining part rented a flat in the city of the University campus where they are expected to attend lectures. Only 4% of respondents are foreign students,[5] and most of them are enrolled in international degree programs taught in English. 57% of respondents are enrolled in social sciences, humanities, and foreign languages degree programs. The rest are enrolled in degree programs concerning technology, applied sciences, mathematics, physics, natural sciences, medicine, etc. On average, students pass between four and five exams per year,

[4] The statistical office of the Bologna municipality shows that more than 75% of students commuting to/from Bologna use trains or buses (http://www.comune.bologna.it/iperbole).

[5] In the following analysis, we omit this variable from our specification since it is never statistically significant. On request, we can provide three regression results: one in which we introduce a dummy variable for foreigner students; other two regressions where we split the sample in domestic and foreign students. These results highlight that no interesting result emerges.

TABLE 1 Descriptive statistics

Variable name	Variable description	Type	*N*	%	Mean	Min-max	SD
Dependent variable							
Y	Number of non-academic books read in the last year	C	20,772		4.421	0.5–19	2.000
Endogenous variable							
exams	Number of passed exams per academic year	C	20,772		4.465	0.5–19	1.990
Instruments							
lect_imp (IV)	Perceived usefulness of attending lectures to increase the number of passed exams (5-level Likert scale)	O	20,772		3.984	1–5	0.982
notes_imp (IV)	Perceived usefulness of studying on notes to increase the number of passed exams (5-level Likert scale)	O	20,772		4.086	1–5	1.020
hand_imp (IV)	Perceived usefulness of studying on handouts to increase the number of passed exams (5-level Likert scale)	O	20,772		4.152	1–5	0.918
Demographic characteristics							
gender	Dummy variable for gender: female = 1; male = 0	D	20,772		0.596	0–1	0.491
age	Age of students, measured in years	C	20.772		22.254	18–28	2.288
age_sq	Students' age squared	C	20.772		500.475	324–784	104.884
University context							
soc_hum	Dummy variable for fields of study: Social sciences, Humanities, Foreign Languages = 1; otherwise (Technology, Applied Sc., Mathematical, Physical and Natural Sc., Medicine, Pharmacy, Sport Sc.) = 0	D	20.772		0.573	0–1	0.495
comm	Dummy variable for commuter student = 1; otherwise = 0	D	20.772		0.373	0–1	0.484
non_resid	Dummy variable nonresident student = 1; otherwise = 0	D	20.772		0.176	0–1	0.380
Student characteristics							
mark	High-school final mark	C	20.772		84.568	60–100	12.356
text	Attitude to study on textbooks (5-level Likert scale)	O	20.772		4.116	0–5	0.988
other_books	Attitude to study on complementary academic books (5-level Likert scale)	O	20.772		2.847	0–5	1.175
cultural capital							
mother	Education level of student's mother	D	20.772			0–1	
	• elementary school (reference)			6.72			

(Continues)

TABLE 1 (Continued)

Variable name	Variable description	Type	*N*	%	Mean	Min-max	SD
	• middle school			22.12			
	• high school			47.23			
	• university			23.94			
father	Education level of student's father	D	20.772			0–1	
	• elementary school (reference)			7.30			
	• middle school			25.88			
	• high school			41.89			
	• university			24.93			
library	Size of home library (number of books/100)	C	20.772		3903	0–100	7.941

Note: Variable type: O = ordinal variable, C = continuous variable; D = dummy variable.

and the tools more often used to pass exams are studying on handouts provided by teachers, notes taken while attending lectures, and recommended textbooks.

4 | RESULTS

4.1 | Full sample

Table 2 displays the estimated coefficients of the auxiliary regression (step I) and the Poisson regression (step II). The Sargan–Hansen's *J* test supports the validity of the instruments chosen in the auxiliary regression ($\chi^2(2) = 0.926$; *p*-value = 0.629).

In step I, we show that female students pass a larger number of exams every year than male students and that the number of exams passed per year rises with age at a decreasing rate. The overall non-linear effect of age is probably due to the fact that basic and core exams, which often require more effort, concentration, and time to prepare, are usually concentrated in the first years of the academic programs. Additionally, freshmen often need some time to learn how to organize time and study at the university. Clever students (with higher high school marks) pass more exams per academic year. Since the education level of the parents and the cultural capital of the family (proxied by the size of the home library) positively affect student performances, high school final grades capture the effect of these cultural environment variables.

Rather than disperse one's effort in reading other textbooks (*other_books*), students consider the attendance of lectures and the use of lecture handouts and notes (lect_imp, notes_imp, hand_imp) essential tools to pass exams.

Turning to the Poisson regression, step II corroborates our main hypothesis that better academic performance reduces leisure reading, as the *exams* coefficient is negative and statistically significant. The results are consistent with the possibility that, in the short run, academic ambition may induce successful students to sacrifice leisure reading to get better academic achievements.[6]

[6] Ignoring the endogenous effect that the time devoted to study has on the time dedicated to leisure reading leads to underestimate the effects that leisure reading has on the final academic performances. In fact, when we estimate a Poisson

TABLE 2 Count-data two-step control-function regression (full sample)

VARIABLES	Auxiliary regression (step I)Dependent variable: exams			Poisson regression (step II)Dependent variable: Y		
Exams				−0.193	**	(0.047)
gender	0.079	***	(0.029)	0.135	***	(0.019)
age	0.045	***	(0.011)	0.035	***	(0.007)
age_sq	−0.000	***	(0.000)	−0.000	***	(0.000)
soc_hum	0.014		(0.031)	0.160	***	(0.019)
comm	0.106	***	(0.039)	0.063	***	(0.024)
non_resid	−0.022		(0.035)	0.058	**	(0.023)
mark	0.013	***	(0.001)	0.007	***	(0.001)
text	0.012		(0.015)	0.016		(0.010)
other_books	−0.086	***	(0.013)	0.031	***	(0.009)
mother						
middle school	−0.102		(0.071)	0.109	***	(0.042)
high school	−0.012		(0.070)	0.157	***	(0.040)
university	−0.132	*	(0.077)	0.138	***	(0.046)
father						
middle school	−0.113	*	(0.069)	−0.075	*	(0.042)
high school	−0.062		(0.069)	−0.002		(0.044)
university	−0.126	*	(0.074)	0.031		(0.047)
library	0.000		(0.002)	0.031	***	(0.002)
constant	1.906	***	(0.258)	0.924	***	(0.198)
lect_imp (IV)	0.127	***	(0.016)			
notes_imp (IV)	0.056	***	(0.017)			
hand_imp (IV)	0.042	***	(0.016)			
rho	0.189	***	(0.047)			
No. of observations	20,772					
Sargan—Hansen's *J* test	$\chi^2(2) = 0.864$ ($p = 0.649$)					

Note: The first-step auxiliary regression models the effect of regressors, included instrumental variables, on the endogenous variable, *exams*; the second-step Poisson regression models the effect of regressors, included *exams*, on the number of books read for leisure, *Y*. Bootstrapped standard errors in parentheses.

***, **, *Significance at 1, 5, and 10% confidence level, respectively.

The age of students positively affects reading at a decreasing rate. Students enrolled in tech and science programs tend to read less for leisure than students enrolled in social sciences and humanities programs.[7] Both commuters and non-resident students read for pleasure more than residents. Moreover, students with higher final marks in secondary school read more than less proficient students, triggering positive virtuous feedback between academic performances and

regression without controlling for endogeneity, the number of exams passed per year does not significantly affect the time devoted to read for pleasure.

[7] Students enrolled in different fields of study may categorize specific books as either leisure or academic reading. The inclusion of field-of-study dummies control for this potential confounding effect.

reading. Studying more intensively additional academic books signals a positive attitude toward leisure reading. The socioeconomic status and cultural capital, measured by parents' education and the home library size, positively affect leisure reading.

4.2 | Gender analysis

The importance of gender in shaping reading habits is widely recognized since females read more than males (Cobb-Clark & Moschion, 2017; Hochweber & Vieluf, 2018). Hence, in order to specifically investigate potential gender differences, we split the data into two subsamples and estimate separate regressions.

Columns (1) and (2) display the estimated coefficients of the auxiliary (Table 3A) and Poisson (Table 3B) regressions for both the male and female samples. The Sargan–Hansen's J test still supports the validity of the instruments chosen in both auxiliary regressions ($\chi^2(2) = 1.039(p = 0.595)$ for the female sample and $\chi^2(2) = 5.011(p = 0.082)$ for the male sample).

Older male students pass a larger number of exams per year. Moreover, the number of passed exams per year is higher for male students enrolled in humanities programs. The ability of students, proxied by the secondary school final mark (*mark*), positively affects academic achievements, whereas the propensity to read other academic books negatively impacts the number of passed exams per year, regardless of gender. The cultural environment (the education of parents and the size of the home library) are also not significant in gender-specific regressions.

Female and male students adopt different studying strategies. In fact, while both groups attach large importance to attending lectures, males prefer to study on lecture notes, whereas females are more likely to study on handouts.

A number of noticeable differences also emerge in the second step of the analysis. Although the effort spent to pass exams by female students does not affect the time devoted to leisure reading, a clear trade-off between the time devoted to studying and leisure reading emerges for male students. The time males spend studying in part substitutes the time dedicated to non-academic readings. As in the auxiliary regression, age positively affects the propensity to read books, particularly for male students. Female resident students dedicate less time to reading with respect to commuters and non-residents, whereas the distinction is not statistically significant for males. Regarding cultural capital, the home library size has a similar effect on all the students, while parents' education levels differently affect reading habits. The higher the level of mothers' education the stronger the impact on leisure reading of female students, whereas this effect is almost not significant for male students. The role of fathers' education levels is almost irrelevant for both male and female students.

Clear differences emerge for students enrolled in humanities or social science faculties, with respect to students enrolled in science, technology, engineering, and mathematics degrees (STEM). The positive influence of reading for leisure is stronger when male students are enrolled in humanities and social science programs.

Given that gender differences play a crucial role in affecting reading behavior and that student reading habits may strongly depend on the field of study, we estimate a further model where gender and fields of study are interacted (Tables 3A and 3B, columns 3). While there are no substantial differences with respect to the full sample model, some peculiarities emerge from the introduction of the interaction terms. Male students enrolled in humanities (*mal_hum*), read more for leisure than any other category of students (being males enrolled in Medicine programs the reference class, *mal_med*), followed by female students enrolled in Humanities (*fem_hum)* and

TABLE 3a First-step auxiliary regressions (female and male subsamples and full sample with interaction terms)

Auxiliary regression (first step)	Female			Male			Interactions		
Dependent variable: exams	Coef.		SD	Coef.		SD	Coef.		SD
age	0.025	*	(0.015)	0.077	***	(0.016)	0.049	***	(0.011)
agesq	−0.000	*	(0.000)	−0.001	***	(0.000)	−0.000	***	(0.000)
soc_hum	−0.089	*	(0.045)	0.159	***	(0.046)			
comm	0.062		(0.053)	0.167	***	(0.058)	0.115	***	(0.038)
non_resid	−0.063		(0.052)	0.028		(0.057)	−0.001		(0.035)
mark	0.014	***	(0.002)	0.013	***	(0.002)	0.014	***	(0.001)
text	−0.008		(0.023)	0.034		(0.021)	0.000		(0.015)
other_books	−0.075	***	(0.016)	−0.101	***	(0.019)	−0.078	***	(0.013)
mother									
middle school	−0.081		(0.097)	−0.157		(0.107)	−0.100		(0.071)
high school	0.051		(0.097)	−0.137		(0.108)	−0.008		(0.070)
university	−-0.105		(0.102)	−0.211	*	(0.116)	−0.132	*	(0.077)
father									
middle school	−0.058		(0.086)	−0.210	**	(0.104)	−0.092		(0.068)
high school	−0.015		(0.086)	−0.143		(0.102)	−0.037		(0.069)
university	−0.151		(0.095)	−0.111		(0.112)	−0.118		(0.074)
library	0.001		(0.002)	−0.001		(0.002)	0.001		(0.002)
lect_imp (IV)	0.159	***	(0.022)	0.092	***	(0.022)	0.145	***	(0.016)
notes_imp (IV)	0.016		(0.024)	0.104	***	(0.023)	0.045	***	(0.017)
hand_imp (IV)	0.049	**	(0.022)	0.041	*	(0.023)	0.047	***	(0.016)
constant	2.387	***	(0.353)	1.299	***	(0.395)	2.040	***	(0.267)
fem_socs							−0.158	**	(0.078)
fem_tech							−0.473	***	(0.090)
fem_math							−0.514	***	(0.092)
fem_hum							−0.450	***	(0.078)
fem_med							0.205	**	(0.092)
mal_socs							−0.194	**	(0.083)
mal_tech							-0.425	***	(0.080)
mal_math							-0.661	***	(0.088)
mal_hum							-0.323	***	(0.087)
mal_med (reference)									
rho	0.021		(0.058)	0.366	***	(0.092)	0.154	***	(0.043)
No. observations	12,362			8410			20,772		
Sargan–Hansen's *J* test	$\chi^2(2) = 1.132$ ($p = 0.568$)			$\chi^2(2) = 4.822$ ($p = 0.090$)			$\chi^2(2) = 0.802$ ($p = 0.670$)		

Note: The first-step auxiliary regression models the effect of regressors, included instrumental variables, on the endogenous variable, *exams*. The interaction terms refer to gender and different fields of study. Bootstrapped standard errors in parentheses. ***, **, *Significance at 1, 5, and 10% confidence level, respectively.

TABLE 3b Second-step Poisson regressions (female and male subsamples and full sample with interaction terms)

Poisson regression (second step)	Female(1)			Male(2)			Interactions(3)		
Dependent variable: Y	**Coef.**		**SD**	**Coef.**		**SD**	**Coef.**		**SD**
exams	−0.021		(0.058)	−0.381	***	(0.093)	−0.157	***	(0.043)
age	0.013		(0.008)	0.081	***	(0.014)	0.035	***	(0.007)
agesq	−0.000		(0.000)	−0.001	***	(0.000)	−0.000	***	(0.000)
soc_hum	0.059	**	(0.024)	0.346	***	(0.033)			
comm.	0.096	***	(0.031)	0.037		(0.043)	0.057	**	(0.023)
non_resid	0.102	***	(0.031)	0.021		(0.042)	0.048	**	(0.023)
mark	0.005	***	(0.001)	0.009	***	(0.002)	0.006	***	(0.001)
text	0.017		(0.013)	0.014		(0.016)	0.012		(0.010)
other_books	0.027	***	(0.010)	0.038	**	(0.016)	0.025	***	(0.008)
mother									
middle school	0.131	***	(0.047)	0.064		(0.082)	0.108	***	(0.040)
high school	0.134	***	(0.047)	0.144	*	(0.078)	0.156	***	(0.039)
university	0.170	***	(0.053)	0.071		(0.086)	0.138	***	(0.044)
father									
middle school	−0.085	**	(0.041)	−0.108		(0.084)	−0.080	*	(0.042)
high school	−0.050		(0.041)	0.036		(0.085)	−0.014		(0.043)
university	−0.029		(0.048)	0.124		(0.090)	0.028		(0.046)
library	0.032	***	(0.003)	0.028	***	(0.002)	0.030	***	(0.002)
constant	0.888	***	(0.261)	0.640	*	(0.339)	0.794	***	(0.200)
fem_socs							0.246	***	(0.050)
fem_tech							0.166	***	(0.060)
fem_math							0.346	***	(0.071)
fem_hum							0.380	***	(0.055)
fem_med							0.291	***	(0.056)
mal_socs							0.200	***	(0.056)
mal_tech							−0.045		(0.057)
mal_math							−0.013		(0.063)
mal_hum							0.459	***	(0.059)
mal_med (reference)									
No. observations	12,362			8410			20,772		

Note: The second-step Poisson regression models the effect of regressors, included *exams*, on the number of books read for leisure, *Y*. The interaction terms refer to gender and different fields of study. Bootstrapped standard errors in parentheses.
***, **, *Significance at 1, 5, and 10% confidence level, respectively.

Mathematics (*fem_math*). Females studying in the fields of Medicine (*fem_med*) and Social Sciences (*fem_socs*) exhibit lower intensity in leisure reading. No difference in reading habits emerges when males enrolled in mathematical (*mal_math*) or technological (*mal_tech*) fields of study are compared to the reference group. These peculiar results support the use of the dummy *soc_hum*

in the full-sample model. Additionally, the results highlight how males and females behave differently when enrolled in the same field of study.

4.3 | Decomposition analysis

The previous empirical evidence suggests a further investigation of gender differences in leisure reading. In particular, we explore whether these differences in leisure reading emerge because of the specific distribution of characteristics in the two samples or are due to any gender idiosyncratic effect.

To this aim, we apply an extension of the Blinder-Oaxaca method to our count-data control-function model to distinguish and quantify how much gender differences in reading are due to the differences in the distribution of characteristics between male and female students (*endowment component, E*) or to the differences in the effects these characteristics exert on leisure reading (*coefficient component, C*).

In detail, the mean difference in the dependent variable, *Y*, between females *(f)* and males *(m)*, can be decomposed as the sum of an endowment effect, *E*, and a coefficient effect, *C* as follows:[8]

$$\overline{y_f} - \overline{y_m} \begin{cases} = \overline{e^{X_f\beta_f}} - \overline{e^{X_m\beta_m}} \\ = \left\{ \dfrac{\overline{e^{X_f\beta_f}} - \overline{e^{X_m\beta_f}}}{E} \right\} - \left(\dfrac{\overline{e^{X_m\beta_f}} - \overline{e^{X_m\beta_m}}}{C} \right) \\ = \sum_{k=1}^{K} W_{\Delta X_k} E + \sum_{k=1}^{K} W_{\Delta\beta_k} C = \sum_{k=1}^{K} E_k + \sum_{k=1}^{K} C_k \end{cases} \tag{6}$$

The two groups differ widely in the way covariates influence leisure reading. Results show that the endowment effect accounts for 10.9% of the difference in leisure reading between groups, while the coefficient effect accounts for 89.1% of this difference. Hence, the idiosyncratic component dominates in explaining the reading gap among genders. Table 4 shows the decomposition into the E and C components for each variable.

Two results neatly emerge. First, female students with better academic performances are associated with higher levels of leisure reading. Since most of the female–male gap is attributed to differences in coefficients, it is likely that idiosyncratic gender characteristics contribute to defining different attitudes to leisure reading. There are gender differences in reading attitudes even at early ages (Van Ours, 2008). Differences are found to be persistent over time and our results seem to corroborate these findings also in the case of academic students.

Second, the decomposition confirms the crucial role of the coefficient effect on leisure reading for each covariate. Almost all covariates are statistically significant, but the weight of the *C* component is often much larger than the weight of the *E* component, suggesting the dominance of a gender idiosyncratic effect.

[8] Multivariate decomposition is performed using the *mvdcmp* Stata package (Powers et al., 2011). This package is based on the methods proposed by Powers and Yun (2009), and extends the Oaxaca–Blinder decomposition technique to nonlinear responses. In case of Poisson regression models, *mvdcmp* exactly decomposes the difference in the average observed outcomes (Wooldridge, 2010).

TABLE 4 Blinder–Oaxaca decomposition results. Differences in the female and male academic performances due to characteristics (E) or coefficients (C) effects

Differences due to	Characteristics (E component)	Coefficients (C component)
	Coeff. (S.E.)	Coeff. (S.E.)
exams	−0.037*** (0.006)	3.197*** (0.366)
age	−0.024*** (0.004)	−8.569*** (0.676)
age_sq	0.018*** (0.003)	3.385*** (0.273)
soc_hum	0.082*** (0.013)	−0.608*** (0.029)
comm.	0.016*** (0.002)	0.183*** (0.032)
non_resid	−0.003*** (0.000)	0.253*** (0.037)
mark	0.072*** (0.006)	−0.938*** (0.228)
text	0.020*** (0.006)	0.040 (0.133)
other_books	0.030*** (0.004)	−0.459*** (0.077)
mother		
middle school	0.009*** (0.001)	−0.008 (0.0.34)
high school	0.002*** (0.000)	−0.208*** (0.075)
university	−0.024*** (0.003)	0.020 (0.044)
father		
middle school	−0.015*** (0.002)	−0.049 (0.034)
high school	0.001* (0.001)	−0.230*** (0.062)
university	−0.023 (0.002)	−0.206*** (0.042)
library	−0.060*** (0.004)	0.030*** (0.011)
constant[a]		4.878*** (0.574)

Note: The E component captures the differences in the distribution of characteristics between male and female students; the C component captures the idiosyncratic gender effect. Male students constitute the reference group, which is also the low-outcome group

[a]The estimates of the endowment effect do not include the constant.

***, **, *Significance at 1, 5, and 10% confidence level, respectively.

5 | CONCLUSIONS

Most of the studies that analyze leisure reading patterns stress the importance of reading activity for both educational purposes and personal development (Clark & Rumbold, 2006). However, relatively little is known about the effect of academic achievements on the students' time allocation, being leisure reading part of it (Fernandez-Blanco et al., 2017).

Following this perspective, we propose a new approach to analyzing the effects of leisure reading, suggesting how previous works might have underestimated the positive effects of reading habits. We investigate the role of academic achievements on students' leisure reading by using a large dataset collected through an online survey conducted among students enrolled at the University of Bologna, Italy. The empirical evidence is consistent with the possibility that successful students tend to sacrifice leisure reading to get better academic achievements. Additionally, and in line with the existing literature, our results confirm the crucial role of gender among the main factors affecting leisure reading habits. In particular, the Blinder–Oaxaca decomposition focuses

on the persistent differences between how female and male students devote their time to leisure reading, highlighting the gender idiosyncratic nature of reading attitudes.

Due to the specific nature of the sample, any generalization of our findings must be made with caution. However, three policy implications can be drawn. First, the policies currently aimed at promoting leisure reading often emphasize the importance of having and choosing books to read, focusing on a favorable family and school environment (Clark & Rumbold, 2006). This necessary condition implicitly deemphasizes the role exerted by the availability of leisure time. Hence, policies aimed at effectively sustaining life-long reading should promote leisure reading also during academic years, by softening the trade-off between studying and leisure reading. A second issue, complementary to the previous one, concerns the long-run effect of leisure reading (Becker & Murphy, 1988), as the reduction of leisure reading during academic years could negatively impact future reading habits. The empirical evaluation of these intertemporal effects is well beyond the scope of this paper, based on sectional data. However, the far-reaching benefits accruing from leisure reading are well-known and we must be aware of the importance of preserving and sustaining lifelong leisure reading habits, as reading for leisure improves well-being, the quality of relationships with others and increases personal income (Brunello et al., 2017). From a public policy perspective, this study suggests the importance to adopt strategies to promote leisure reading during academic years. Finally, with the adoption of standard strategies aimed at promoting leisure reading among academic students, the gender reading gap is likely to persist. For this reason, gender-specific strategies are strongly recommended to promote leisure reading among (academic) students.

We are aware of the limitations of our study, mainly due to the nature of the data. To better explore reading choices during University years, additional information on the volume of academic reading would be useful. However, any attempt to proxy it might be biased since students are asked to read other materials (handouts, lecture notes, etc.) beyond textbooks. Hence, a future research direction could explore the complex nature of academic reading in greater detail, which can offer an insightful perspective on present and future reading habits.

ACKNOWLEDGMENT

We gratefully acknowledge constructive and insightful comments provided by the Editor and two anonymous Referees. R. Zanola thanks University of Eastern Piedmont for funding his partecipation to this study.

Open Access Funding provided by Universita degli Studi del Piemonte Orientale Amedeo Avogadro within the CRUI-CARE Agreement.

ORCID

Roberto Zanola https://orcid.org/0000-0002-2056-3324

REFERENCES

Bavishi, A., Slade, M. D., & Lvy, B. R. (2016). A chapter a day: Association of book reading with longevity. *Social Science and Medicine*, *164*, 44–48.

Becker, G. S., & Murphy, K. M. (1988). A theory of rational addiction. *Journal of Political Economy*, *96*(4), 675–700.

Brunello, G., Weber, G., & Weiss, C. T. (2017). Books are forever: Early life conditions, education and lifetime earnings in Europe. *Economic Journal*, *127*, 271–296.

Canoy, M. F. M., van Ours, J. C., & van der Ploeg, F. (2006). The economics of books. In V. A. Ginsburgh & D. Throsby (Eds.), *Handbook of the economics of art and culture* (Vol., 1, 721–761). Elsevier.

Clark, C., Osborne, S., & Akerman, R. (2008). *Young people's self-perceptions as readers: An investigation including family, peer and school influences*. National Literacy Trust.

Clark, C., & Rumbold, K. (2006). *Reading for pleasure: A research overview*, National Literacy Trust.

Cobb-Clark, D. A., & Moschion, J. (2017). Gender gaps in early educational achievement. *Journal of Population Economics*, *30*(4), 1093–1134.

De Schrijver, A. (2012). Sample survey on sensitive topics: Investigating respondents' understanding and trust in alternative versions of the randomized response technique. *Journal of Research Practice*, *8*(1), M1.

Fernandez-Blanco, V., Prieto-Rodriguez, J., & Suarez-Pandiello, J. (2017). A quantitative analysis of reading habits: Some evidence from Spain. *International Journal of Arts Management*, *19*(3), 19–32.

Hochweber, J., & Vieluf, S. (2018). Gender differences in reading achievement and enjoyment of reading: The role of perceived teaching quality. *The Journal of Educational Research*, *111*(3), 268–283.

Larson, R. B. (2019). Controlling social desirability bias. *International Journal of Market Research*, *61*(5), 534–547.

Logan, S., & Johnston, R. (2009). Gender differences in reading ability and attitudes: examining where these differences lie. *Journal of Research in Reading*, *32*(2), 199–214.

McGeown, S. P., Norgate, R., & Warhurst, A. (2012). Exploring intrinsic and extrinsic reading motivation among very good and very poor readers. *Educational Research*, *54*(3), 309–22.

McIntosh, S., & Vignoles, A. (2001). Measuring and assessing the impact of basic skills on labor market outcomes. *Oxford Economic Papers*, *53*(3), 453–81.

OECD. (2002). *Reading for change: Performance and engagement across countries*. Results from PISA.2000. OECD.

Powers, D. A., Hirotoshi, U., & Yun, M. S (2011). mvdcmp: Multivariate decomposition for nonlinear response models. *Stata Journal*, *11*(4), 556–576.

Powers, D. A., & Yun, M. S. (2009). Multivariate decomposition for hazard rate models. *Sociological Methodology*, *39*, 233–263.

Sieben, S., & Lechner, C. M. (2019). Measuring cultural capital through the number of books in the household. *Measurement Instruments for the Social Sciences*, *2*(1), 1–6.

Suárez-Fernández, S., & Boto-García, D. (2019). Unraveling the effect of extrinsic reading on reading with intrinsic motivation. *Journal of Cultural Economics*, *43*(4), 579–605.

Swalander, L., & Taube, K. (2007). Influences of family based prerequisites, reading attitude, and self-regulation on reading ability. *Contemporary Educational Psychology*, *32*, 206–230.

Stockmans, M. J. W. (1999). Reading attitude and its effect on leisure time reading. *Poetics*, *26*, 245–261.

Tomul, E., & Polat, G. (2013). The effects of socioeconomic characteristics of students on their academic achievement in higher education. *American Journal of Educational Research*, *1*(10), 449–455.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, *133*(5), 859–883.

Tramonte, L., & Willms, J. D. (2010). Cultural capital and its effects on education outcomes. *Economics of Education Review*, *29*(2), 200–213.

Troyer, M., Kim, J. S., Hale, E., Wantchekon, K. A., & Armstrong, C. (2019). Relationship among intrinsic and extrinsic reading motivation, reading amount and comprehension: A conceptual replication. *Reading and Writing*, *32*, 1197–1218.

Van Ours, J. (2008). When do children read books? *Education Economics*, *16*(4), 313–328.

Van Schooten, E., & de Glopper, K. (2003). The development of literary response in secondary education. *Poetics*, *31*, 155–187.

Whitten, C., Labby, S., & Sullivan, S. L. (2016). The impact of pleasure reading on academic success. *The Journal of Multidisciplinary Graduate Research*, *2*(4), 48–64.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT Press.

How to cite this article: Scorcu, A. E., Vici, L., & Zanola, R. (2023). Good students, avid readers: The cost of academic success. *Bulletin of Economic Research*, *75,* 609–624. https://doi.org/10.1111/boer.12373