Repetition undermines social inclusion at school

La repetición perjudica la inclusión social en la escuela

Pablo Brañas-Garza
https://orcid.org/0000-0001-8456-6009
Universidad Loyola Andalucía

Diego Jorrat
https://orcid.org/000-0002-50
Universidad Loyola Andalucía

Abstract

Regardless of other possible effects, grade repetition forces students to disconnect from their friends in class and connect with their new classmates. This study quantifies how grade retention affects students' social integration. To analyze short-term effects, we use a propensity score matching to compare retained students with their "statistical twins". For long-term effects, we compare current repeaters with those who repeated in the past. The results are not optimistic. In the short term, retained students are less popular, have more enemies and fewer "good" friends in the classroom. They are also more likely to appear in hate networks. In the long term, 'former' retained students are slightly more popular than current students, but in all other respects they remain the same. We conclude that grade retention has a strong negative impact on students' social relationships, and that this effect hardly diminishes over time.

Keywords: Adolescents, social networks, grade retention, school vulnerability, social inclusion, social behaviour, education system.

Resumen

Independientemente de otros posibles efectos, la repetición de curso obliga a los estudiantes a desconectar de sus amigos de la clase y a conectar con sus nuevos compañeros de clase. Este trabajo cuantifica como la repetición tiene implicaciones en la integración social de los estudiantes. Para mirar los efectos de corto plazo, empleamos el método de emparejamiento por puntuaciones de propensión para comparar al repetidor con un "gemelo estadístico". Para el largo plazo, comparamos repetidores actuales con otros que lo hicieron

anteriormente. Los resultados no son optimistas. En el corto plazo, los repetidores son menos populares, tiene mayor número de enemigos y menos "buenos" amigos dentro de la clase. Además, aparecen con mayor frecuencia en las redes de odio. En el largo plazo, los repetidores "antiguos" son algo más populares que los actuales, pero en todo lo demás son iguales. Podemos concluir que la repetición tiene un impacto muy negativo en las relaciones sociales de los estudiantes y que dicho efecto apenas se aminora con el paso del tiempo.

Introduction

One of the most pressing issues in the Spanish education system is the high rate of grade repetition. According to the Education at a Glance 2024 report, Spain has a repetition rate of 7.8% in lower secondary and 6.5% in upper secondary education, whereas the average for OECD countries is less than half those figures (2.2% and 3.2%, respectively). Grade repetition refers to a situation in which a student who has completed an academic year must remain at the same level for an additional year. It is important to note that the decision to repeat a grade rarely originates from the student's environment but is instead made collectively by the school's teaching staff. According to Royal Decree 984/2021, repetition is considered an exceptional measure, permitted at most twice during the stage, and must always be based on an evaluation of the student's progress and their ability to acquire essential learning outcomes.

According to statistics published by the Ministry of Education, Vocational Training and Sports (MEFPD, 2024), based on data from the 2022–2023 academic year, 7.3% of first-year students in compulsory secondary education (ESO) repeated the grade, along with 6.8% in the second year, 7.3% in the third, and 6.7% in the fourth. In every year, the percentage of male students repeating is consistently higher than that of female students. Particularly striking is the proportion of boys repeating the first year of ESO, which reaches 8.7%, compared to 5.8% among girls.

Although the number of students repeating a grade is high, there are many reasons to question whether this policy has any real benefits. On the one hand, staying in the same grade for an additional year seems to have, at best,

uncertain effects on academic performance. One might expect that grade repetition would promote learning - consolidating knowledge and allowing for a better match between students' abilities and the level of instruction. However, the evidence suggests that this is not the case and that it may even reduce academic achievement (García Pérez et al., 2014).

Second, these potential benefits appear to come at a high personal cost for the student: stigma from teachers or classmates, a decrease in self-confidence, and difficulty adjusting to a new peer group (see Manacorda, 2012). Indeed, there is causal evidence that grade repetition increases the likelihood of early school dropout (Jacob and Lefgren, 2009; Manacorda, 2012; De Witte et al., 2013; Freeman and Simonsen, 2015; González-Rodríguez et al., 2019).

Finally, grade repetition imposes a significant financial cost on institutions, which must fund an extra year of schooling for the repeating student. This additional year also imposes an economic cost on the student, as it delays entry into the labor market and thus the start of earning labor income (Tafreschi and Thiemann, 2016).

Research on grade repetition has focused almost exclusively on academic aspects, such as the educational performance of repeaters or the factors associated with a higher likelihood of repeating a grade (see, for example, González-Rodríguez et al., 2016; González-Betancor et al., 2019; López et al., 2023; and Nieto-Isidro et al., 2023). It is somewhat surprising that little effort has been made to study how repetition affects students' social integration. After all, when students are required to repeat a grade, they are separated from their classroom friends and forced to interact with a new group of peers. This is unlikely to come at no cost.

In this study, we focus precisely on this issue: the cost of grade repetition in terms of students' social integration. By social integration, we refer to the number of friends (popularity), centrality, and clustering of each student (whether they are repeaters or not). To conduct this analysis, we use the TeensLab dataset (Vasco et al., 2025), which contains information on more than 5,000 secondary school students. This dataset covers multiple dimensions of students' cognitive and non-cognitive skills, as well as measures of academic outcomes—such as grades—and other essential aspects such as the students' future orientation (patience) and their tolerance or preference

for risk. In some of the TeensLab schools (although not all), students were also asked whether they were currently repeating a grade or had repeated one in the past. In fact, the proportion of students currently repeating stands at 9.54%, a figure consistent with that reported by MEFPD (2024).

In addition, TeensLab provides the individual network measures mentioned earlier and, more importantly, contains information from over 200 independent classroom networks. As we will see later, the network measures are computational calculations with no subjective component; they simply assign a numerical value to each variable of interest. In other words, network analysis tells us not whether a student feels more or less isolated, but whether they are actually more isolated, regardless of whether they perceive it or are even aware of it.

This study aims to answer two research questions. First, we measure the effect of grade repetition on students' social integration. Second, we examine whether this effect persists over time or fades after a few years.

To address the first research question, we compare network measures between repeaters and non-repeaters. To avoid obvious endogeneity problems—since repeaters are inherently different from non-repeaters by virtue of repeating —we use a statistical technique known as *propensity score matching* (hereafter, PSM). This method allows us to search within the TeensLab dataset for statistical "twins" of the repeaters— that is, students who are statistically similar to repeaters in certain characteristics except for the fact that they have not repeated a grade. The matching is based on variables that are both observable and unobservable to teachers (such as patience, risk aversion, and cognitive skills), all measured in TeensLab. This approach enables us to isolate the "causal" effect of repeating on the outcome variables, as it compares comparable samples. It is also important to note that, by comparing repeaters to their classmates, we are capturing the immediate (short-term) effect of repetition.

To address the second research question, we compare students who are currently repeating a grade with students who repeated a grade in the past. In other words, both groups share the stigma of having repeated—only that some are experiencing it now, while others experienced it previously. As we will see throughout the paper, these two groups exhibit similar observable

characteristics, allowing us to assume they are comparable and that any differences in social integration measures can be attributed to the effect of currently repeating. This second analysis enables us to assess whether the effects of grade repetition persist over time.

Method

### Social Integration and Network Metrics

The study of relationships between students in the classroom is not a new topic. Sociograms were first used in the United States in the 1930s and have since been widely employed, particularly to identify patterns of interaction, detect conflicts, and improve classroom dynamics. As we will see throughout this section, the unidirectional metric (which refers to one student nominating another) is relatively simple. However, the metric that captures interactions between students—for example, the shortest (or longest) path between two individuals—can be quite cumbersome. Interested readers are encouraged to consult the book by Jackson (2010) and the collection of papers edited by Bramoullé et al. (2016). They may also refer to the work of Ruiz-García et al. (2023) on the creation of an index for triadic relationships, that is, how the third friend of a pair of friends becomes connected to the other.

  What is a network? Let us consider any given classroom where many students have relationships with one another, but not necessarily with everyone. Suppose that each student i declares which classmates from the set C are their friends. We will call A_i the set of friends of student i, and the "super" set of all declared friendships will be referred to as the *friendship network*: $(\{A_i\}_{i \in C}, C)$, where $a_i = |A_i|$ represents the number of friends that student i has within the class. It is important to note that, unless a student is friends with the entire class, we will necessarily have $a_i < C - 1$.

  This setup implies that there will be classmates who are "strangers" to student i—that is, they do not belong to their set of friends: $E_i = C \backslash A_i$. With these two definitions—friends and strangers—in hand, we can illustrate an example of a network. Imagine a small class with only six students, where

each of them has their own set of friends: $A_1 = \{2, 3\}$, $A_2 = \{1\}$, $A_3 = \{1, 4\}$, $A_4 = \{3\}$, $A_5 = \emptyset$, $A_6 = \{2, 3\}$.

FIGURE I. Example of a Network with Six Students
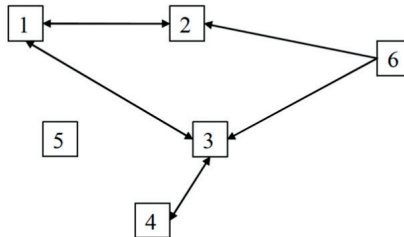
Source: Compiled by the authors.

Figure I illustrates a common phenomenon in networks. Students 1 and 2 nominated each other - $A_1 = \{2, 3\}$, $A_2 = \{1\}$ - so student 1's set of friends includes 2, and student 2's set includes 1. However, this is not the case with student 6 - $A_6 = \{2, 3\}$ – who declared 2 as a friend, but was not listed as a friend by student 2. This *lack of reciprocity* in friendships is quite common: one student may consider another a friend, while the feeling is not reciprocated.

When we examine networks such as those in TeensLab, we find that many relationships are not reciprocal but are instead declared by only one of the two individuals. In other words, *h* considers *j* to be a friend, while *j* sees the relationship differently. Thus, the sets A_i represent the friends declared by each student *i*, but these relationships are not necessarily reciprocal. In fact, student *i* may appear in the friendship sets of other students who are not part of their own set of friends (as in the previous example between students 2 and 6).

Out-degree and *Popularity (in-degree)*. When we refer to *out-degree*, we mean the number of elements in the set of friends declared by each student, $a_i = |A_i|$. In contrast, *in-degree*, denoted as d_i, refers to the number of students (other than i) in the class *C* who include subject *i* in their set of friends. In other words, *out-degree* indicates how many friends a student nominates, while *in-degree* indicates how many students nominate him. It is important to note

that d_i can range from a minimum value of 0 (no one nominated the student as a friend—as with student 5 in Figure II) to a maximum value of *C–1* (the entire class except themselves).

In the network literature, *in-degree* is commonly referred to as *popularity*. We focus on *in-degree* because it is a variable that the student cannot manipulate, as it depends on others rather than on themselves. This is particularly relevant because it helps eliminate endogeneity concerns.

What is centrality? Centrality refers to connectivity. Some students in a class are capable of linking subgroups (*clusters*) that, in their absence, would remain independent networks. In Figure II, we revisit the network from Figure I and add a new cluster composed of three students, all of whom are friends with one another: $A_7 = \{8, 9\}$, $A_8 = \{7,9\}$, $A_9 = \{7,8\}$. If student 6 becomes friends with student 7, the two previously separate networks merge into a single network. It is important to emphasize that the role of student 6 is critical: if, for any reason, their connection with student 7 is broken, the network will split back into two disconnected groups. For this reason, individuals who occupy central positions are referred to as *key players* (see Ballester et al., 2006).

FIGURE II. Example of a central player
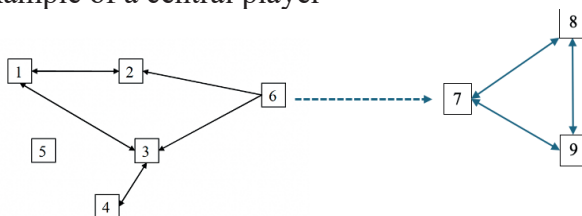
Source: Compiled by the authors.

Unfortunately, there is no a single definition of centrality; on the contrary, there are many—such as centrality, eigenvector, rank, and others. In this study, we use a very common metric: *betweenness*. For simplicity, we will say that a person is more central if they have a greater ability to connect others, and less central if their capacity to connect others is lower. If we think about all the possible connections in the network from Figure II -for example, between 1 and 2, 1 and 3, ..., 1 and 6; 2 and 1, 2 and 3, ..., 2 and 6; and so on-

the most central student will be the one who appears most frequently along those paths. Another way to explain it is that one person is more central in the network than another if the removal of the former causes greater disruption to the network than the removal of the latter.

The TeensLab Database

TeensLab is a consortium formed by the Universities of Barcelona, Carlos III, Granada, Loyola, and the Basque Country, aimed at studying economic behavior among adolescents in Spain. There is evidence showing that these skills (both cognitive and non-cognitive) are important determinants of real-life decision-making in adulthood and are correlated with variables typically associated with "positive outcomes" such as education, savings, and others (see Dohmen et al., 2011; Golsteyn et al., 2014; Falk et al., 2018; Angerer et al., 2023).

The TeensLab project gathered data from 5,890 students across 33 educational centers located in two Spanish regions: Andalusia and Catalonia. Data collection was conducted with the consent of school principals and followed strict standards of anonymity and confidentiality. The methodology combined surveys with *lab-in-the-field* experiments. The main dimensions captured in the dataset include: (i) economic preferences (regarding risk and time), (ii) cognitive skills, (iii) strategic thinking, and (iv) network metrics at the classroom level. In addition, the dataset contains variables related to students' sociodemographic characteristics, as well as a range of complementary measures on physical appearance, mood (happiness), expectations, and other aspects.

The design of the experiment included several specific features. First, it was implemented as a classroom activity in each school to maximize response rates (see Alfonso et al., 2023) and was conducted through an online platform, SAND (Social Analysis and Network Data), to guarantee data protection. Second, students completed the experiment using tablets, which allowed them to read the instructions independently, proceed through the questionnaire sequentially without the possibility of returning to previous screens, and respond to the survey.

Third, the entire questionnaire was administered in Spanish, and due to restrictive school policies, hypothetical (rather than real) incentives were used in the experimental tasks. However, prior evidence shows that the behavior of both adolescents and adults does not differ between real and hypothetical payment schemes, at least for risk and time preferences, supporting the reliability of the results (Brañas-Garza et al., 2021, 2023; Alfonso et al., 2023).

Finally, the sample includes students aged between 10 and 23 years (Mean = 14.10, SD = 1.94), covering various educational levels: primary education (8.62%), lower and upper secondary education (84.94%), high school (1.90%), and vocational training (4.53%). The sample is balanced by gender: 49.68% identify as female, 49.68% as male, and the remaining 0.64% as "other" or "prefer not to say".

Vasco et al. (2025) provides a detailed description of the database. The complete dataset is publicly available at the following link: https://github.com/teenslab/datateenslab.

This study uses data from 1,821 students. This number is smaller than the full TeensLab dataset for two reasons. First, we focus exclusively on students enrolled in compulsory secondary education (ESO), which yields a sample of 5,003 students. Second, the question regarding whether a student is currently repeating or had previously repeated a grade was included in only 11 schools, resulting in a subsample of 2,155 ESO students for whom this information is available. As explained below, we ultimately work with 1,821 observations after excluding 129 students who repeated more than a year ago and 205 students with missing values in key variables. The next section presents the empirical strategy and the data used in this article.

Variables definition, empirical strategy and estimation method

To answer the research questions posed in this study and provide some causal evidence, it is necessary to apply different empirical strategies. Endogeneity arises when an explanatory variable is correlated with the error term in a regression model. This leads to biased and inconsistent estimates of the causal effect of grade repetition on the outcome variables.

Endogeneity can arise for several reasons, but in our case two stand out: (i) omitted variables, where other factors - such as family background, personal motivation, or school quality - may affect both the likelihood of repeating a grade and the student's social integration; and (ii) simultaneity or reverse causality, where social integration itself may affect the likelihood of repeating (students with lower levels of social integration may be more likely to repeat). To avoid these problems, we use different subsamples that are comparable and allow us to draw more reliable conclusions.

We perform two separate analyses. First, we compare current repeaters with non-repeaters who have a similar profile. To do this, we use a statistical technique known as propensity score matching (PSM) with a kernel-based matching approach (KPSM). The basic idea behind this method is that for each observation in the treatment group (students who are currently repeating a grade), we find a "statistical twin" in the control group (students who are neither currently repeating nor have repeated in the past). This allows us to create a counterfactual that helps us estimate the effect of grade repetition on the relevant outcome variables.

Second, we compare students who are currently repeating to those who have repeated in the past. Obviously, this sample is smaller, but it remains highly informative because we are comparing individuals who effectively share a key characteristic: having repeated a grade.

Before presenting the empirical strategy, we describe the variables used in this study. The TeensLab dataset provides information on students' age, gender, academic performance, and whether the student or their parents were born abroad (migrant status). It also includes measures of cognitive skills (using the Cognitive Reflection Test, CRT; see Brañas-Garza et al., 2019a; Frederick, 2005; and Thomson and Oppenheimer, 2016), patience, and risk tolerance. The patience measure is based on Alfonso et al. (2023), and the risk tolerance measure on Vasco and Vázquez (2025). Appendix B describes these tasks in detail.

Patience and risk tolerance are included as control variables because they help capture unobservable characteristics. There is strong evidence that patience is associated with perseverance and tends to correlate with better academic outcomes (see Brañas-Garza et al., 2019b, for a review). Risk toler-

ance, in turn, is associated with a wide range of behaviors, from entrepreneurship to alcohol consumption (see Dohmen et al., 2011, for a review).

Table I presents a summary of the variables. It is important to note that the variables for academic performance, patience, and risk tolerance were standardized using the min-max method, transforming them into a range from 0 to 1, where higher values indicate a higher level in each characteristic.

TABLE I. Summary Statistics of Matching and Outcome Variables

	N	Mean	SD	Min	Max
Control and matching variables					
Repeaters	1821	0.04	0.20	0	1
Female	1821	0.49	0.50	0	1
CRT	1821	0.50	0.27	0	1
GPA	1821	0.63	0.40	0	1
Patience	1821	0.48	0.35	0	1
Risk	1821	0.60	0.16	0	1
Age	1821	14.24	1.14	10	18
Migrant	1821	0.21	0.41	0	1
Outcome variables					
In-degree friends	1821	8.12	3.85	0	21
In-degree best friends	1821	2.93	2.01	0	11
In-degree enemies	1821	2.34	2.50	0	22
In-degree worst enemies	1821	0.70	1.27	0	12
Out-degree friends	1821	9.00	6.30	0	30
Out-degree best friends	1821	3.20	2.90	0	28
Out-degree enemies	1821	2.80	3.61	0	29
Out-degree worst enemies	1821	0.82	1.63	0	29
Betweenness friends	1821	16.86	19.22	0	63.73
Betweenness best friends	1821	16.35	24.67	0	80.14
Betweenness enemies	1821	10.55	18.68	0	62.67
Betweenness worst enemies	1821	0.76	2.11	0	8
Clustering friends	1821	0.68	0.19	0	1
Clustering best friends	1821	0.52	0.34	0	1

Clustering enemies	1821	0.23	0.29	0	1
Clustering worst enemies	1821	0.06	0.19	0	1

Source: Compiled by the authors using data from TeensLab.

Additionally, the bottom panel of Table I presents the outcome variables used to measure social integration, covering four dimensions: *In-degree*, *Out-degree*, *Betweenness*, and *Clustering*. *In-degree* is defined as the number of classmates who identify the student as a friend (i.e., a measure of the student's popularity), whereas *Out-degree* corresponds to the number of peers the student names as friends. *Betweenness* indicates how central (or important) the student is within the class network, and *Clustering* measures the extent to which the student's friends are also friends with one another. We conduct all analyses separately for each type of relationship: friends, best friends, enemies, and worst enemies. Appendix A shows the interface that each student uses to indicate their friends (and enemies), providing the information necessary to calculate all these variables.

We now proceed to explain the empirical strategy used to measure the short- and long-term effects of grade repetition.

Comparison between current repeaters and non-repeaters

The main challenge in comparing repeaters and non-repeaters is that these groups may differ in aspects such as academic performance, patience, risk tolerance, and other characteristics. If these differences are not controlled for, any direct comparison could lead to spurious conclusions, as observed effects on social integration could be driven by other factors rather than by repetition itself. To address this issue, propensity score matching (PSM) estimates the probability that a student will repeat a grade (the propensity score) based on a set of individual characteristics. Each repeater is then matched to non-repeaters with similar propensity scores, ensuring that the groups are comparable.

To estimate the propensity score, we use the variables listed at the top of Table I. Additionally, we apply exact matching on migrant status, CRT score, age, and average grade (GPA). This ensures that each repeater is only compared with non-repeaters who match exactly on these key characteristics.

There are different ways to perform matching. A common method is to assign each repeater a single non-repeater with the closest propensity score (1-to-1 matching). However, this approach can be inefficient, as it discards information and may produce less stable estimates. Instead, we use Kernel matching, which weights multiple non-repeaters with similar propensity scores, assigning greater weight to those who are closer. This reduces variance and improves the precision of the estimates. All analyses were conducted using the *kmatch* package in Stata (see Jann, 2017, for a detailed description).

To verify that the matching procedure has produced a valid comparison group, we perform three key diagnostic tests:

- #1 – *Mean differences before and after matching* (Figure III, left panel). Before matching, repeaters and non-repeaters show significant differences across several characteristics. These differences, represented by circles, refer to the raw data. However, after matching, these differences disappear or are greatly reduced in the matched data, represented by triangles, indicating that the groups have been properly balanced.

FIGURE III. Mean differences and variance ratio: Repeaters vs. non-repeaters.
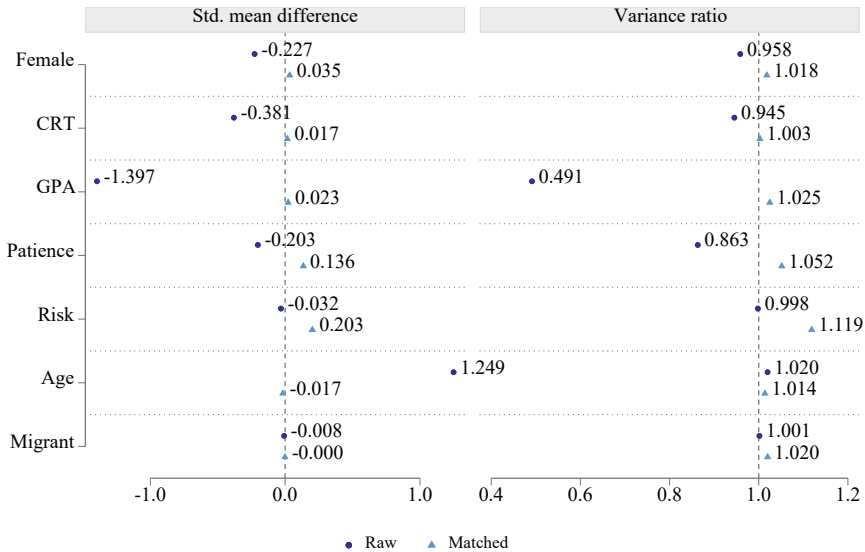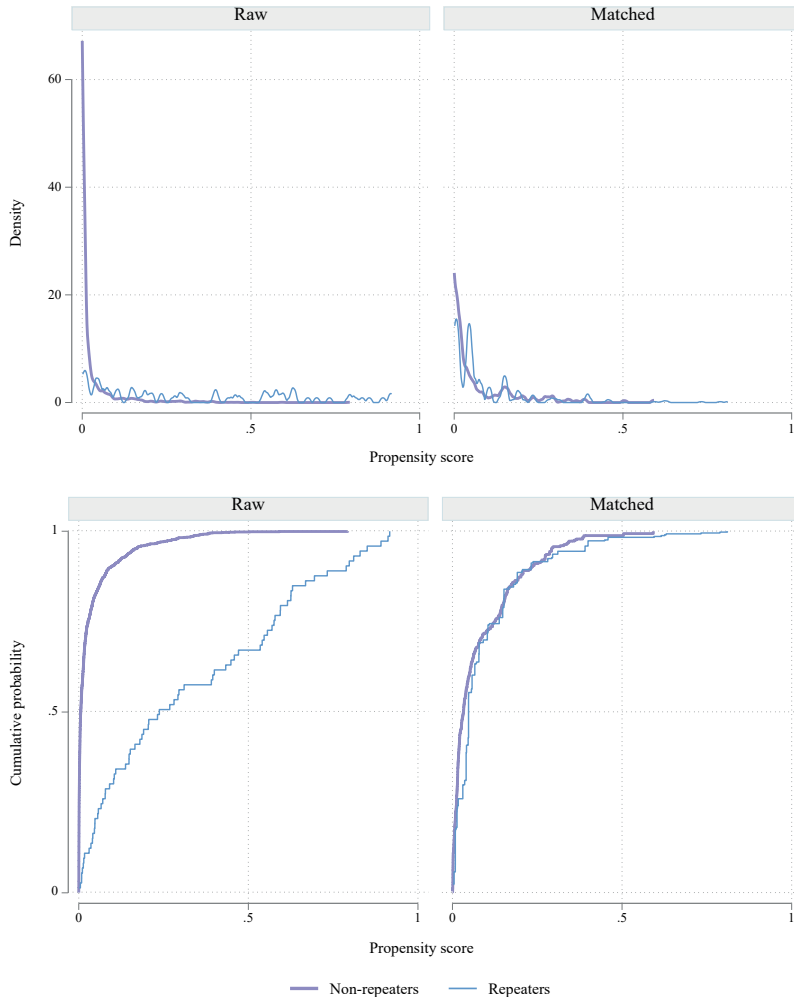

Source: Compiled by the authors based on data from TeensLab.

- ▪ *#2 – Variance ratio between repeaters and non-repeaters* (Figure III, right panel). Effective matching requires balancing not only means but also the variance of covariates. The right panel of Figure III shows that, prior to matching, the variance of key characteristics differed substantially between repeaters and non-repeaters (represented by circles). After matching, most variance ratios converge toward 1 (represented by triangles), indicating that the distributions of covariates are now similar across groups.
- ▪ *#3 - Distribution of the propensity score before and after matching* (Figure IV). The top panel displays the distribution of the propensity score, while the bottom panel shows its cumulative distribution. Both panels illustrate that, prior to matching (left panel, "Raw Data"), the distributions of repeaters and non-repeaters were quite different. After matching (right panel, "Matched Data"), the distributions become much more similar, indicating adequate com-

mon support. This suggests that there are sufficient non-repeaters with comparable characteristics to the repeaters, allowing for valid comparisons.

FIGURE IV. Distribution (top) and cumulative distribution (bottom) of the propensity score.

Source: Compiled by the authors based on data from TeensLab.

Since the diagnostic tests indicate good balance in the means of the control variables, variance ratios close to 1 (indicating balanced dispersion), and adequate common support, we conclude that the Kernel-based PSM procedure has successfully produced a comparable control group. This allows us to credibly estimate the causal effect of grade repetition on social integration outcomes by comparing repeaters with a well-defined group of non-repeaters.

To estimate the effect, we compute the difference between the mean outcome for repeaters and a weighted average outcome for non-repeaters, where weights are based on the estimated propensity scores. This procedure identifies the Average Treatment Effect on the Treated (ATT)—the causal impact of grade repetition on students who actually repeated—rather than the Average Treatment Effect (ATE), which would reflect the effect if repetition were applied to all students. The estimated immediate effect of repetition on each outcome variable is thus:

$$\widehat{ATT} = \frac{1}{N_{D=1}} \sum_{i|D=1} \left[Y_i - \sum_{j|D=0} w_{ij} Y_j \right] \quad (1)$$

where $N_{D=1}$ is the total number of individuals currently repeating a grade *(D=1)*; Y_i denotes the outcome for individual *i* in the treatment group (repeaters) and Y_j denotes the outcome for individual *j* in the control group, that is non-repeaters *(D=0)*; and w_{ij} is the weight assigned to individual *j* in the control group to match him or her to treated individual *i*. In addition, when estimating the ATT, we include school fixed effects to control for idiosyncratic differences across schools (e.g., differences in grade repetition policies.

Finally, although Kernel-based PSM significantly reduces selection bias, it does not guarantee perfect causal identification, as there may still be unobserved factors that influence both grade repetition and social integration (e.g., personal motivation or family support). Nevertheless, within the set of observational methods, Kernel PSM provides a robust strategy for generating a valid comparison between repeaters and non-repeaters.

Comparison between current repeaters and former repeaters

To estimate whether grade repetition has a lasting effect over time, we focus on the sample of repeaters. However, unlike the previous section, we do not compare them to matched synthetic twins but rather to other students who also repeated, distinguishing between those who are currently repeating and those who repeated in prior years. As shown in Table II, the sample includes 203 repeaters: 74 students who were repeating at the time of the experiment and 129 students who had repeated in the past. Table II also shows that there are no significant differences between these groups in terms of gender (female), GPA, CRT score, patience, risk tolerance, age, migrant status, or whether they repeated more than once.

The only significant difference (*p<0.01*) between the two samples is found in academic performance (GPA), where current repeaters have lower average grades than those who repeated a year or more ago. This result is expected, as the questions on academic performance refer to the previous school year, and current repeaters, by definition, had lower grades in that period.

In summary, the two subsamples are comparable across almost all observable variables, allowing us to assume similarity in unobservable characteristics as well. Under this identification assumption, we can estimate the lasting effect of grade repetition using a multiple linear regression model. Specifically, we estimate the following model by Ordinary Least Squares (OLS) for each outcome variable *q= {In-degree, Out-degree, Betweenness, Clustering}*:

$$Y_{is}^q = \beta_0 + \beta_1 * R_i + \theta^c * X_i^c + \tau_s + \epsilon_i \quad (2)$$

where Y_{is}^q denotes the outcome variable q for each individual i in school s; R_i equals 1 if student i is currently repeating and 0 if he/she repeated in a previous year; X_i is a vector of control variables $c=$ {*Female, CRT, GPA, Patience, Risk, Migrant*}; τ_s represents school fixed effects; and ϵ_i is the error term. Robust standard errors are used to account for heteroskedasticity. The estimated long-term effect of grade repetition is captured by the coefficient $\widehat{\beta_1}$.

TABLE II. Differences in observable variables: Current vs. former repeaters.

	Repeat		Difference
	≥1 year	Now	
	(1)	(2)	(2) – (1)
Female	0.442	0.378	-0.063
	(0.499)	(0.488)	(0.072)
CRT	0.360	0.405	0.045
	(0.262)	(0.266)	(0.039)
GPA	0.311	0.180	-0.131***
	(0.354)	(0.273)	(0.048)
Patience	0.483	0.414	-0.068
	(0.334)	(0.321)	(0.048)
Risk	0.618	0.592	-0.026
	(0.178)	(0.164)	(0.025)
Age	15.411	15.575	0.164
	(1.275)	(1.117)	(0.180)
Migrant	0.198	0.216	0.018
	(0.400)	(0.414)	(0.059)
Multiple repetition	0.132	0.162	0.030
	(0.340)	(0.371)	(0.051)
Observations	129	74	203

Source: Compiled by the authors based on data from TeensLab. Note: Standard errors in parentheses. The *Difference* column tests for equality of means and asterisks denote significance levels: ***$p<0.01$; **$p<0.05$ y *$p<0.10$.

Results

Following the structure of the paper, we divide the results into two sections. First, we compare current repeaters with non-repeaters who share similar

characteristics—that is, their "statistical twins." Second, we compare current repeaters with students who repeated in the past. All the analyses were conducted using Stata 18.

Comparison between current repeaters and non-repeaters

In this section, we compare students who are currently repeating a grade with the rest of their classmates who are not repeaters. Figure V presents the ATT estimates for each outcome variable, using the Kernel PSM methodology described earlier.

Panel A shows that repeaters are less popular than their classmates ($p < 0.05$), and the negative effect is even stronger ($p < 0.01$) when considering best friends. Interestingly, they are more popular in enemy networks ($p < 0.01$), although no differences are observed for worst enemies.

Panel B shows that repeaters nominate a similar number of friends as their peers, but they report significantly fewer best friends ($p < 0.01$). No significant differences are found in the number of enemies nominated. In addition, panel C reveals no substantial differences in overall centrality—repeaters are neither more nor less central in the classroom network. However, they do appear to be more central in enemy networks ($p < 0.05$).

Regarding *clustering*, panel D highlights two important findings. First, repeaters are less likely to be embedded in cohesive groups of best friends ($p < 0.05$)—that is, their best friends are less likely to be connected to each other. Second, repeaters are more likely to appear in tightly connected clusters of enemies, where all members report considering them an enemy ($p < 0.05$). All estimates and robust standard errors are reported in Table A1 in Appendix C.

We summarize the findings as follows:

Result 1: Repeaters are less popular, have more enemies, and report fewer best friends in the class. Moreover, they appear more frequently in enemy networks and have enemies who are friends with one another.

These results are particularly relevant because the comparison group

for repeaters consists of statistical twins -that is, individuals who are identical in all observed characteristics except for having repeated a grade. While the Kernel PSM method does not provide definitive causal identification -since those who repeated actually did so, and those who did not were never exposed to the treatment - we can assume that the matched twins are so similar that the decision to repeat can be regarded as quasi-random.

In this context, we can think of repetition as the result of a process that is, to some extent, random: among a set of very similar students, some ended up repeating due to bad luck, while others did not. Under this assumption, we interpret the results as showing that grade repetition significantly damages students' social capital—in other words, their social integration is substantially reduced.

FIGURE V. ATT of grade repetition based on Kernel PSM, with 95% confidence intervals.
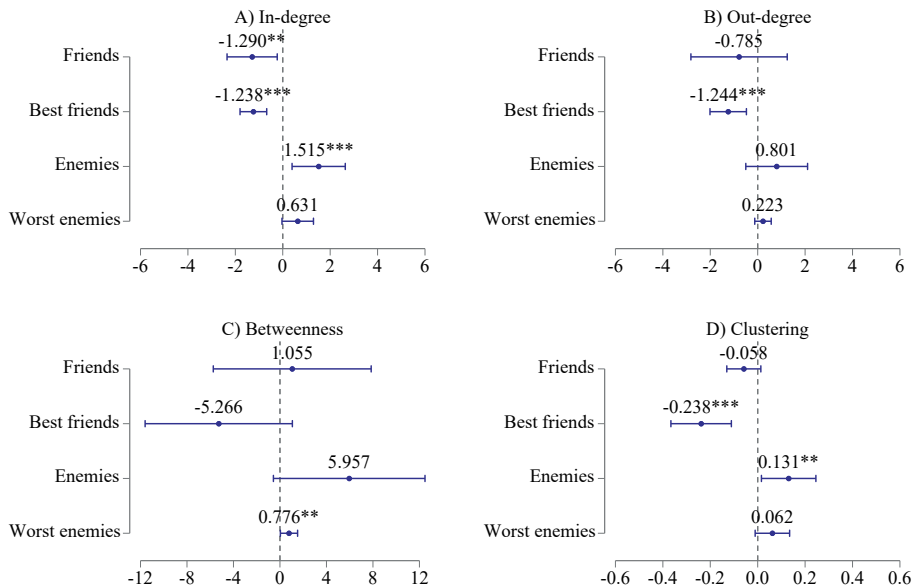
Source: Compiled by the authors based on data from TeensLab. School fixed effects are included in the estimation.

Comparison between current repeaters and former repeaters

In the previous section, we showed that grade repetition has a negative impact on social integration, but we do not yet know how long this impact lasts—in other words, whether students recover their social capital one or more years after repeating. To address this question, we examine differences between students who are currently repeating and those who repeated in a previous year. Importantly, the measure of social inclusion refers to the time at which the data were collected. That is, all individuals in this analysis carry the stigma of having repeated a grade, but some are experiencing it now, while others experienced it in the past.

Figure VI displays the estimated coefficient $\widehat{\beta_1}$ for each outcome variable, based on the multiple linear regression model described above. In Panel A, we find that students who are currently repeating are less popular ($p < 0.05$) than those who repeated in the past. A similar effect is observed for best friends ($p < 0.05$). No significant differences are found in enemy networks.

We also find no effect on out-degree (Panel B): students who are currently repeating do not nominate more or fewer friends (or enemies) than those who repeated in the past. Likewise, we observe no substantial differences in centrality (Panel C), nor any significant impact on clustering (Panel D). For a more detailed presentation of the results, Tables A2 through A5 in Appendix C report the full regression estimates, both with and without controls, for each outcome variable.

In summary:

Result 2: Compared to students who repeated in the past, current repeaters are less popular in friendship networks. No significant differences are found in out-degree, centrality, or clustering.

Although we cannot -and should not- draw strong conclusions from a relatively small sample (n = 203), the evidence is nonetheless concerning**.** Result 2 shows that the only meaningful difference between former and current repeaters is that the latter are less popular. This suggests that the negative

effect of repetition on popularity is not permanent and tends to fade over time. Apart from this, former and current repeaters are virtually identical across all other outcome variables.

FIGURE VI. Current vs former repeaters, multiple linear regression model with 95% CI.
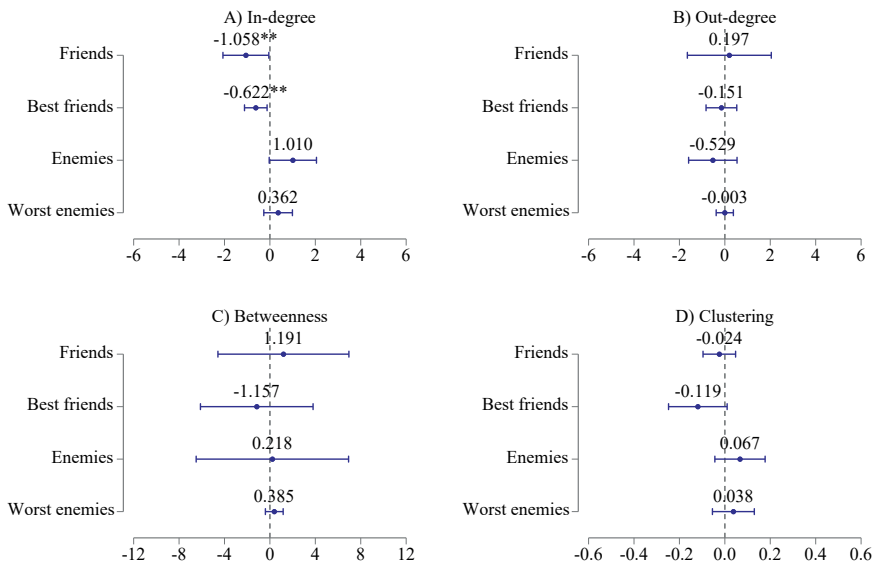
Source: Compiled by the authors based on data from TeensLab. School fixed effects are included.

This implies that, except for (un)popularity—which appears to recover one or more years after repeating—all other patterns described in Result 1 remain unchanged: repeaters continue to have more enemies, fewer close friends, and are more frequently embedded in hostile networks. In short, the negative effects of grade repetition on students' social integration persist over time.

Conclusions

Grade repetition in Spain is a longstanding issue which, although it has improved slightly, remains unresolved. The country continues to show alarmingly high rates compared to other OECD countries.

Existing research raises serious concerns about the actual benefits of grade repetition for students, as there is little evidence of improvements in academic performance (García Pérez et al., 2014). By contrast, there is clear evidence of direct costs, such as stigmatization—both by peers and sometimes by teachers—and reduced self-confidence (see Manacorda, 2012). There is even causal evidence suggesting that repetition increases the likelihood of school dropout (Jacob and Lefgren, 2009; Manacorda, 2012; De Witte et al., 2013; Freeman and Simonsen, 2015; González-Rodríguez et al., 2019).

It is also important to remember that maintaining a repetition rate of around 10% imposes a substantial cost on the education system. These costs are not only economic but also logistical, as schools must accommodate more students in classrooms that, by design, lack flexibility. In addition, repetition imposes a financial burden on students themselves, as it delays their entry into the labor market and, consequently, the point at which they begin earning income (Tafreschi and Thiemann, 2016).

This study explores a previously understudied consequence of grade repetition. Using network metrics and data from the TeensLab project (Vasco et al., 2025), we examine how repetition affects the social capital -or social integration- of students who repeat a grade. The analysis follows two complementary approaches: first, we compare repeaters with non-repeaters who share similar characteristics (their "statistical twins"); second, we compare them with other students who repeated in previous years.

To measure differences in network metrics between repeaters and non-repeaters, we use a statistical technique known as propensity score matching, which allows us to construct "statistical twins" and isolate the quasi-causal effect of repetition on social integration. Because we compare repeaters with their classmates, this analysis captures the immediate—or short-term—effect of repeating a grade.

The results are concerning. Repeaters are less popular, have more enemies, fewer close friends, appear more frequently in hostile networks, and have enemies who are friends with one another. In short, the short-term impact of repetition on students' social capital is severe: they not only lose social

ties but also become targets within enemy networks.

To assess the long-term effects of grade repetition, we compare students who repeated in the past with those who are currently repeating. This comparison is meaningful because both groups share the stigma of having repeated a grade. The only difference is timing—some are repeating now, while others did so previously—which allows us to identify which effects persist over time.

When comparing former repeaters with current ones, we find a single difference: current repeaters are less popular. In all other outcome variables, the two groups are virtually identical. This suggests that, years after repeating, students may recover their popularity, but no substantial improvements are observed in other dimensions. In short, they still have fewer friends, more enemies, appear central in networks of hostility, and so on. That is, their relational capital does not significantly recover beyond the dimension of popularity.

Taken together, our findings suggest that grade repetition severely undermines students' social capital: repeaters tend to lose friends, accumulate more enemies, and occupy more central positions within enemy networks. Importantly, these effects appear to persist over time.

Acknowledgement

Bibliographic references

Alfonso, A., Brañas-Garza, P., Jorrat, D., Lomas, P., Prissé, B., Vasco, M., & Vázquez-De Francisco, M. J. (2023). The adventure of running experiments with teenagers. *Journal of Behavioral and Experimental Economics, 106*, 102048.

Angerer, S., Bolvashenkova, J., Glätzle-Rützler, D., Lergetporer, P., & Sutter, M. (2023). Children's patience and school-track choices several years later: Linking experimental and field data. *Journal of Public Economics, 220*, 104837.

Ballester, C., Calvó-Armengol, A., & Zenou, Y. (2006). Who's who in networks. Wanted: The key player. *Econometrica*, *74*(5), 1403-1417.

Bramoullé, Y., Galeotti, A. & Rogers, B. (2016). The Oxford Handbook of the Economics of Networks, Oxford: Oxford University Press.

Brañas-Garza, P., Estepa-Mohedano, L., Jorrat, D., Orozco, V., & Rascón-Ramírez, E. (2021). To pay or not to pay: Measuring risk preferences in lab and field. *Judgment and Decision Making, 16*(5), 1290-1313.

Brañas-Garza, P., Jorrat, D., Espín, A. M., & Sánchez, A. (2023). Paid and hypothetical time preferences are the same: Lab, field, and online evidence. *Experimental Economics, 26*(2), 412-434.

Brañas-Garza, P., Kujal, P., & Lenkei, B. (2019a). Cognitive reflection test: Whom, how, when. *Journal of Behavioral and Experimental Economics, 82*, 101455.

Brañas Garza, P. E., Espín, A. M., & Jorrat, D. (2019b). Midiendo la paciencia. *Economía Industrial*, *413*, 21-31.

de Witte, K., Cabus, S., Thyssen, G., Groot, W., & van den Brink, H. M. (2013). A critical review of the literature on school dropout. *Educational Research Review*, *10*, 13-28.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association, 9*(3), 522-550.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of*

Economics, 133(4), 1645-1692.

Freeman, J., & Simonsen, B. (2015). Examining the impact of policy and practice interventions on high school dropout and school completion rates: A systematic review of the literature. *Review of Educational Research*, *85*(2), 205-248.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25-42.

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

Golsteyn, B. H., Grönqvist, H., & Lindahl, L. (2014). Adolescent time preferences predict lifetime outcomes. *The Economic Journal, 124*(580), 39-61.

González-Betancor, S. M., & López-Puig, A. J. (2016). Grade retention in primary education is associated with quarter of birth and socioeconomic status. *PLoS ONE, 11*(11), 1-19.

González-Rodríguez, D., Vieira, M. J., & Vidal, J. (2019). Factors that influence early school leaving: a comprehensive model. *Educational Research, 61*(2), 214-23.

Jacob, B. A., & Lefgren, L. (2009) The effect of grade retention on high school completion. *American Economic Journal: Applied Economics*, 1, 33–58.

Jackson, M. (2019). Social and Economic Networks. Princeton: Princeton University Press.

Jann, B. (2017). kmatch: Kernel matching with automatic bandwidth selection. Stata Users' Group Meetings 2017 11, UK.

López, L., González-Rodríguez, D., & Vieira, M-J. (2023). Variables que afectan la repetición en la educación obligatoria en España. *Revista Electrónica de Investigación Educativa, 25*, e17, 1-15.

Manacorda, M. (2012). The cost of grade retention. *Review of Economics and Statistics*, *94*(2), 596-606.

MEyFP, Ministerio de Educación y Formación Profesional (2024). *Las cifras de la educación en España. Curso 2022-2023.* Recuperado de https://www.educacionfpydeportes.gob.es/servicios-al-ciudadano/estadis-

ticas/indicadores/cifras-educacion-espana/2022-2023.htmlNieto-Isidro, S., & Martínez-Abad, F. (2023). Repetición de curso y su relación con variables socioeconómicas y educativas en España. *Revista de Educación*, 402, 207-236.

OCDE (2024). *Panorama de la Educación: Indicadores de la OCDE 2024.* OECD Publishing, Paris. Recuperado de https://www.libreria.educacion.gob.es/libro/panorama-de-la-educacion-indicadores-de-la-ocde-2024-informe-espanol_184584/.

Ruiz-García , M., Ozaita, J. Pereda, M., Alfonso, A., Brañas-Garza, P., Cuesta, J.A., & Sanchez, A. (2023). Triadic influence as a proxy for compatibility in social relationships, *Proceedings of the National Academy of Sciences*, 120 (13) e2215041120

Tafreschi, D., & Thiemann, P. (2016). Doing it twice, getting it right? The effects of grade retention and course repetition in higher education. *Economics of Education Review*, *55*, 198-219.

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making*, *11*(1), 99-113.

Vasco, M., Alfonso, A., Arenas, A., Cabrales, A., Cuesta, J. A., Espín, A. M., . & Brañas Garza, P. (2025). Economic preferences and cognitive abilities among teenagers in Spain. *Scientific Data 12*, 7.

Vasco, M., & Vazquez, MJ. (2025). The Gumball machine. *PLoS ONE*, *20(4)*, e0320576.

Contact address: Pablo Brañas-Garza. Universidad Loyola Andalucía, Loyola Behavioral Lab. E-mail: branasgarza@gmail.com

Appendix

Appendix A: Network elicitation protocol in TeensLab

Appendix B: Self-reported grade point average (GPA), CRT, Patience, and Risk elicitation.

Self-reported GPA

To measure GPA (grade point average), students were asked how many "sobresalientes" (equivalent to A+) and "notables" (equivalent to A) they had received in their three main subjects—mathematics, language, and English—during the previous academic year. A "sobresaliente" was assigned a value of 2 points and a "notable" to 1 point, and 0 otherwise, resulting in a GPA variable with a maximum of 6 points. To ensure comparability and avoid scale issues, the GPA was standardized using the min-max method, rescaling it to range from 0 to 1. Higher values on this scale indicate a greater number of top grades and, accordingly, stronger academic performance.

A continuación, vamos a preguntarte sobre tus **CALIFICACIONES** durante el **año pasado**.

¿Has obtenido algún SOBRESALIENTE (9 o 10) en el curso pasado?

☐	Si
☐	No

¿En qué asignatura/s obtuviste sobresaliente? Puedes marcar más de una.

☐	Inglés
☐	Lengua Castellana y Literatura
☐	Matemáticas
☐	Otra. Indica cuál: :

¿Has obtenido algún NOTABLE (7 u 8) en el curso pasado?

☐	Si
☐	No

¿En qué asignatura/s obtuviste notable? Puedes marcar más de una.

☐	Inglés
☐	Lengua Castellana y Literatura
☐	Matemáticas
☐	Otra. Indica cuál: :

Cognitive Reflection Test (CRT)

The CRT refers to the Cognitive Reflection Test developed by Frederick (2005) and later adapted for non-adult populations by Thomson and Oppenheimer (2016). The test consists of three questions designed to elicit both intuitive and reflective responses. Each item offers an intuitive but incorrect answer, and a correct one that requires analytical reasoning. Based on this task, we compute the number of reflective responses, with higher scores indicating greater cognitive reflection (see Brañas-Garza et al., 2019b, for a review).

Por favor, responde las preguntas siguientes. En estas preguntas sí hay respuestas correctas e incorrectas.

En una biblioteca, todos los meses se duplica el número de libros. Si la biblioteca tarda 48 meses en llenarse, ¿cuánto tardaría en llenarse a la mitad? Indica con un número.

> Introducir valor

Introduzca el valor de la respuesta.

Si estás corriendo una carrera y pasas a la persona que está en segunda posición, ¿en qué posición estás? Indica con un número. Por ejemplo: 1 (primero), 2 (segundo), etc

> Introducir valor

Introduzca el valor de la respuesta.

El padre de Emilia tiene 3 hijas. Las dos primeras se llaman Abril y Mayo. ¿Cuál es el nombre de la tercera hija?

> Introducir texto

2000 caracteres restantes.

Patience and risk elicitation

Time Preferences (Patience). Patience is measured using a task developed by Alfonso et al. (2023), in which students make a series of six sequential choices between receiving a smaller amount of money today or a larger amount in the future. This variable is also standardized using the min-max method, such that values closer to 1 represent higher levels of patience. Below, we present screenshots of the task interface.

Time discount

Time discounting

A lo largo de las siguientes 6 pantallas, tendrás que tomar 6 decisiones (una por pantalla) sobre **cómo quieres recibir una cantidad de dinero** hipotética.

Tu tarea es **elegir qué opción prefieres**, sabiendo que, si eliges una recibirás el dinero **mañana**, y si eliges la otra lo recibirás **la semana que viene**.

Conforme vayas avanzando en las decisiones, la cantidad de dinero que recibirás por esperar será cada vez mayor.

Decision #1

Decision #2

Decision #3

Risk preferences. Risk attitudes are measured using a task developed by Vasco and Vázquez (2025), in which adolescents make six sequential decisions. In each decision, they choose between two options (A and B), each represented by a gumball machine with different payouts and probabilities. Option A is the safer choice, while Option B is riskier, offering a wider range of possible outcomes. A student's risk preference is measured as the number of times they choose Option B (ranging from 0 to 6). As with previous variables, the measure is standardized using the min-max method, so that values closer to 1 indicate a greater willingness to take risks. Screenshots of the task interface are shown below.

Risk elicitation

Risk task

En esta tarea vas a tomar decisiones sobre **probabilidades**. Probabilidad es la posibilidad de que algo pase, es decir, cómo alguien está seguro de qué va a ocurrir.

En cada decisión **tendrás que elegir entre dos opciones** donde podrás ganar dinero ficticio: en ambas vas a ganar, pero puedes tener más suerte y ganar más, o puedes tener menos suerte y ganar menos. **Las probabilidades de ganar van cambiando** de una decisión a otra.

A continuación, se te presentan **6 preguntas diferentes**. Tu tarea consiste en elegir la opción **A)** o la opción **B)** en **todas** las preguntas.

Para realizar la tarea, pasa a la **siguiente pantalla**.

Decision #4

¿Qué prefieres?

A)

8€
10€ 10€
10€ 8€

B)

2€
20€ 20€
20€ 2€
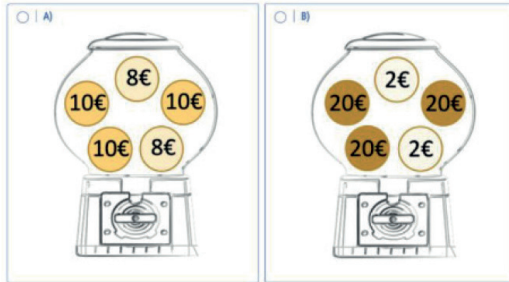
Decision #5

¿Qué prefieres?

A)

10€
10€ 10€
10€ 8€

B)

20€
20€ 20€
20€ 2€
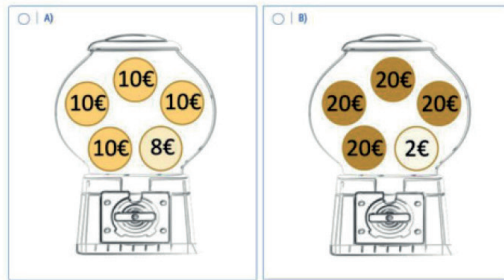
Decision #6

¿Qué prefieres?

A)

10€
10€ 10€
10€ 10€

B)

20€
20€ 20€
20€ 20€
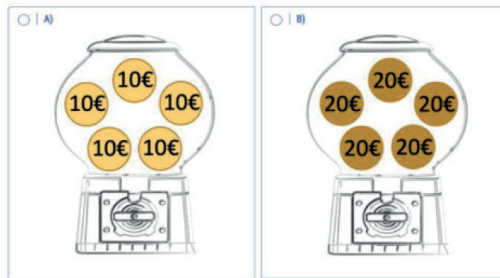
Appendix C: Estimations results

TABLE A1. Estimation of the short-term effect of grade repetition: ATT estimates using kernel PSM.

	(1)	(2)	(3)	(4)
	Friends	Best friends	Enemies	Worst enemies
a) In-degree				
ATT	-1.290**	-1.238***	1.515***	0.631*
	(0.540)	(0.287)	(0.572)	(0.339)
Observations	1,821	1,821	1,821	1,821
b) Out-degree				
ATT	-0.785	-1.244***	0.801	0.223
	(1.036)	(0.392)	(0.664)	(0.177)
Observations	1,821	1,821	1,821	1,821
c) Betweenness				
ATT	1.055	-5.266	5.957*	0.776**
	(3.463)	(3.230)	(3.325)	(0.378)
Observations	1,821	1,821	1,821	1,821
d) Clustering				
ATT	-0.0584	-0.238***	0.131**	0.0620*
	(0.0366)	(0.0650)	(0.0586)	(0.0371)
Observations	1,821	1,821	1,821	1,821

Note: Standard errors in parentheses. Asterisks indicate statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

TABLE A2. Estimation of the long-term effect of grade repetition on in-degree, based on multiple regression analysis.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Friends	Friends	Best friends	Best friends	Enemies	Enemies	Worst enemies	Worst enemies
Repeater	-0.486	-1.058**	-0.387*	-0.622**	0.235	1.010*	0.158	0.362
	(0.495)	(0.511)	(0.234)	(0.252)	(0.565)	(0.528)	(0.316)	(0.319)
Female		-0.281		-0.251		0.672		0.138
		(0.465)		(0.212)		(0.414)		(0.232)
CRT		0.520		-0.357		-2.390***		-1.237***
		(0.892)		(0.403)		(0.767)		(0.447)
GPA		-0.636		-0.261		0.550		-0.340
		(0.738)		(0.351)		(0.654)		(0.375)
Patience		-0.915		-0.722**		1.428**		0.720*
		(0.745)		(0.303)		(0.698)		(0.422)
Risk		0.508		0.163		-0.859		-0.662
		(1.463)		(0.693)		(1.075)		(0.567)
Age		0.035		0.110		-0.541**		-0.341**
		(0.206)		(0.101)		(0.230)		(0.140)
Multiple repetition		0.816		0.210		0.614		0.430
		(0.669)		(0.289)		(0.780)		(0.517)
Constant	7.629***	7.665**	2.396***	1.319	2.322***	10.257***	0.947**	6.571***
	(0.740)	(3.419)	(0.363)	(1.674)	(0.593)	(3.840)	(0.460)	(2.264)
Observations	203	190	203	190	203	190	203	190
	0.161	0.222	0.135	0.199	0.118	0.222	0.047	0.124
School Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses. Asterisks indicate statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

TABLE A3. Estimation of the long-term effect of grade repetition on out-degree, based on multiple regression analysis.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Friends	Friends	Best friends	Best friends	Ene-mies	Enemies	Worst enemies	Worst enemies
Repeater	-0.200	0.197	-0.352	-0.151	-0.394	-0.529	-0.080	-0.003
	(0.946)	(0.936)	(0.395)	(0.343)	(0.484)	(0.540)	(0.202)	(0.192)
Female		-1.761**		-0.123		1.124**		0.094
		(0.822)		(0.305)		(0.544)		(0.247)
CRT		-1.279		0.091		2.023*		0.423
		(1.656)		(0.533)		(1.164)		(0.488)
GPA		1.916		-0.375		1.466		0.763
		(1.260)		(0.471)		(0.974)		(0.538)
Patience		0.168		-0.151		-0.267		0.199
		(1.101)		(0.424)		(0.689)		(0.307)
Risk		-2.033		-0.746		-1.596		0.542
		(2.394)		(0.828)		(1.267)		(0.515)
Age		0.453		0.206		-0.372		-0.036
		(0.428)		(0.152)		(0.247)		(0.138)
Multiple repetition		0.484		-0.086		1.721**		0.228
		(1.477)		(0.410)		(0.683)		(0.224)
Constant	6.267***	0.718	2.584***	-0.026	2.731*	7.977*	0.827***	0.514
	(1.554)	(6.722)	(0.606)	(2.482)	(1.470)	(4.573)	(0.234)	(2.089)
Observa-tions	203	190	203	190	203	190	203	190
	0.081	0.111	0.073	0.049	0.057	0.139	0.028	0.035
School Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses. Asterisks indicate statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

TABLE A4. Estimation of the long-term effect of grade repetition on betweenness, based on multiple regression analysis.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Friends	Friends	Best friends	Best friends	Enemies	Enemies	Worst enemies	Worst enemies
Repeater	1.678	1.191	-1.410	-1.157	-1.312	0.218	0.093	0.385
	(2.922)	(2.922)	(2.313)	(2.513)	(3.053)	(3.400)	(0.390)	(0.397)
Female		-4.719*		2.810		7.355**		0.916**
		(2.519)		(2.433)		(3.034)		(0.353)
CRT		-2.674		-8.275		0.267		-0.887
		(4.412)		(5.199)		(5.407)		(0.660)
GPA		4.414		-5.443**		5.298		0.160
		(4.111)		(2.557)		(5.560)		(0.666)
Patience		4.448		-4.771		1.187		0.425
		(3.784)		(3.083)		(4.408)		(0.575)
Risk		-9.688		2.880		-4.962		-0.205
		(6.202)		(8.841)		(9.744)		(1.215)
Age		2.259*		1.325		-0.995		-0.054
		(1.315)		(1.207)		(1.496)		(0.172)
Multiple repetition		2.792		-0.438		1.640		0.068
		(4.328)		(2.772)		(3.794)		(0.481)
Constant	7.656**	-22.829	9.108**	-9.000	15.167**	26.217	1.702**	1.938
	(3.782)	(19.600)	(4.023)	(16.659)	(6.329)	(27.330)	(0.849)	(3.217)
Observations	203	190	203	190	203	190	203	190
	0.022	0.052	0.038	0.063	0.042	0.054	0.065	0.089
School Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses. Asterisks indicate statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

TABLE A5. Estimation of the long-term effect of grade repetition on clustering, based on multiple regression analysis.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Friends	Friends	Best friends	Best friends	Enemies	Enemies	Worst enemies	Worst enemies
Repeater	0.018	-0.024	-0.099	-0.119*	0.069	0.067	0.039	0.038
	(0.035)	(0.036)	(0.062)	(0.065)	(0.052)	(0.056)	(0.041)	(0.047)
Female		0.023		0.012		-0.005		-0.023
		(0.031)		(0.057)		(0.045)		(0.034)
CRT		0.112*		0.002		-0.121		-0.042
		(0.067)		(0.110)		(0.074)		(0.051)
GPA		-0.112**		-0.156*		0.036		0.069
		(0.047)		(0.091)		(0.064)		(0.063)
Patience		-0.097*		0.050		-0.047		-0.021
		(0.051)		(0.093)		(0.067)		(0.038)
Risk		0.134		0.173		-0.198		-0.108
		(0.121)		(0.170)		(0.128)		(0.102)
Age		-0.027**		-0.023		0.004		-0.002
		(0.014)		(0.028)		(0.020)		(0.016)
Multiple repetition		0.065		-0.052		-0.003		0.048
		(0.046)		(0.095)		(0.066)		(0.058)
Constant	0.753***	1.168***	0.687***	0.982**	0.154**	0.254	0.014	0.114
	(0.039)	(0.239)	(0.107)	(0.455)	(0.069)	(0.325)	(0.028)	(0.255)
Observations	203	190	203	190	203	190	203	190
	0.167	0.213	0.069	0.090	-0.009	-0.034	-0.013	-0.036
School Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses. Asterisks indicate statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.