The long-term effects of grade retention: Empirical analysis on French data

Léonard Moulin et Florent Sari

DOCUMENTS DE TRAVAIL 301

The long-term effects of grade retention: Empirical analysis on French data

Léonard Moulin et Florent Sari

February 2025

ined
INSTITUT NATIONAL
D'ÉTUDES DÉMOGRAPHIQUES

https://doi.org/10.48756/ined-dt-301.0225

The long-term effects of grade retention: Empirical analysis on French data

Léonard Moulin*, Florent Sari†

Abstract

In order to estimate the causal impact of grade retention in French lower secondary school on various long-term outcomes, we conducted a longitudinal study on the trajectories of French students who entered grade 6 in the 1995-1996 academic year over a period of 17 years. The results indicate that grade retention has a significant negative impact on both obtaining a secondary school diploma and a higher education qualification. This affected students across all socioeconomic groups, both boys and girls. However, the negative effect on wages is observed only for boys and for students from low socioeconomic backgrounds. A series of robustness checks confirm the reliability of the results.

Keywords: grade retention; educational outcomes; labor market outcomes; propensity score matching; medium-term effects.

JEL code: I21; I24; I26.

*Institut national d'études démographiques (INED), F-93300 Aubervilliers, France. lenoard.moulin@ined.fr

†Université Paris-Est Créteil (Erudite) and Tepp-CNRS, FSEG, Mail des Mèches - Av. du Général de Gaulle, 94010 Créteil, France. florent.sari@u-pec.fr

1 Introduction

Educational policies on grade repetition (or retention), in which students are required to repeat a school year due to insufficient academic performance, vary widely between countries (Valbuena et al., 2021). In France, Spain, and Germany, grade repetition is frequently employed, and is widely viewed as a mechanism to support students who do not meet required performance benchmarks. These evaluation systems restrict grade promotion to those who achieve satisfactory academic results. In countries such as the United Kingdom and the Scandinavian nations, in contrast, grade repetition is rare or nearly absent. These systems favor social promotion, an egalitarian approach that automatically advances students to the next grade regardless of performance, in the aim of minimizing stigma.

Grade repetition remains a contentious issue that divides education policy experts (Contini and Salza, 2024). While it can, in some cases, provide students with additional time to acquire essential skills (Shepard and Smith, 1989), it is frequently criticized for its economic and social inefficiencies. Critics argue that the high financial costs associated with grade repetition place a significant burden on education systems, while its psychological effects (such as discouragement and social disconnection) can adversely affect students (Manacorda, 2012). Furthermore, evidence suggests that grade repetition may exacerbate educational inequalities and negatively influence the long-term academic outcomes of retained students. Its overall effectiveness and economic justification remain subjects of considerable debate within the field.

Further empirical research into the advantages and drawbacks of grade retention is thus needed. The economic literature examining its positive or negative effects on students' academic achievement is not new (Jimerson, 2001; Goos et al., 2021). Overall, studies have produced mixed evidence regarding the effectiveness of grade retention policies, depending on the institutional context, the level at which retention occurs (primary, secondary etc.), the student's gender, or their socioeconomic background. Moreover, the effects may vary depending on the outcome variable considered (risk of subsequent retention, grades in specific subjects, academic orientation choices, etc.).

While the literature examining the effects of grade retention on students' performance in the years immediately following this event is abundant, few studies explore its medium- or long-term effects. Concretely, research has primarily focused on the observable effects during schooling, with little to no attention given to the longer-term consequences, such as labor market entry. To the best of our knowledge, only the works of Eide and Showalter (2001), Babcock and Bedard (2011), and Brodaty et al. (2008) analyze the effects on careers, and particularly on earnings or wages. The first two studies on the U.S provide evidence that grade retention may increase labor market earnings, thereby benefiting students. The third study, in contrast, highlights the negative impact of educational

delays, primarily caused by grade retention, on the wages of young workers in France. It is therefore important to conduct new research to further examine these effects. Our objective in this study is to contribute to the existing literature by examining the medium- and long-term consequences of grade retention in lower secondary school on students in France. France is a particularly interesting context for new research, as retention remains understudied there despite its widespread practice. Specifically, we analyze the effects of retention on the likelihood of graduation (obtaining a secondary school diploma), pursuing higher education (obtaining a higher education qualification), and professional outcomes (employment status and monthly wage) several years after grade repetition. Our contribution is original for at least two reasons. First, while two studies have examined the medium- and long-term effects of grade retention, both are based on U.S. data. Second, although Brodaty et al. (2008) provide insights into the French case, their analysis focuses on educational delay rather than grade retention itself. Educational delay, as they define it, can result not only from grade retention but also from other factors such as changes in academic orientation, study abroad programs, interruptions in studies, or temporary disability. Our study is the first in Europe to directly analyze the effects of grade retention using a dataset that allows us to specifically identify instances.

To examine the links between grade retention during lower secondary school and academic or labor market outcomes, we use propensity score matching (PSM). This allows us to compare two groups of students: one consisting of students who repeated a grade in lower secondary school (treatment group) and the other consisting of those who did not (control group). We then compare students with the same probability of repeating a grade (same propensity scores), who either repeated or did not repeat a grade. Individuals in the two groups with equal propensity scores will tend to have the same distributions of observable characteristics. The estimated propensity score is thus used to reduce bias through matching (Rosenbaum and Rubin, 1983; Dehejia and Wahba, 2002). While these methods offer the advantage of effectively addressing selection based on observable characteristics, they do not account for selection on unobservables. However, we conduct several robustness checks to demonstrate that this issue is very limited in our case. In addition, we show that the use of such methods is particularly suitable here, given that alternative estimation techniques, notably those relying on instrumental variables, provide results that are both unstable and less precise (see section 5.1).

In this study, we use a survey conducted by the French Ministry of National Education, Panel d'élèves du second degré, recrutement 1995 (Panel of secondary school students, 1995 recruitment). This survey allows us to track a sample of students who entered grade 6 in the 1995-1996 academic year over a period of 17 years. It provides a detailed observation of individual characteristics, educational trajectories, and early labor market experiences. It thus offers a particularly rich longitudinal view of a cohort of students in France.

The results indicate that grade retention has a significant negative impact on both obtaining a secondary school diploma and pursuing a higher education qualification. Further analysis of heterogeneous effects reveals that grade retention negatively affects these outcomes for both boys and girls, with a stronger impact observed for girls. This negative effect on academic achievement is consistent across all socioeconomic status (SES) groups, with no notable variation in magnitude. As to career effects, our findings suggest that grade retention does not have a direct impact on the likelihood of being employed. However, there is a significant negative effect on monthly wages. Gender-specific analysis reveals no differences between boys and girls in terms of employment access, but a negative impact on monthly wages exclusively for boys. Lastly, our results indicate that grade retention does not affect likelihood of employment in any SES group. However, the impact on monthly wages differs: for the lowest SES group, grade retention is associated with a marginally significant wage reduction, whereas no significant effects are observed for higher SES groups.

The remainder of the paper is structured as follows: Section 2 provides a description of the theoretical mechanisms proposed to explain how grade repetition may positively or negatively affect academic and labor market outcomes, along with a review of the existing empirical literature on the topic. Section 3 presents the institutional context of grade repetition in developed countries, with a particular focus on France. Section 4 describes the data and variables used in our analysis and provides some descriptive statistics. Section 5 presents the estimation strategy. Section 6 discusses the results of our different models, while Section 7 concludes the article.

2 Literature review

2.1 Theoretical mechanisms

Grade retention is an educational practice used more or less frequently in various countries and at different levels of education (kindergarten, primary school, secondary school etc.). It is one of the most widely debated and controversial educational measures.

The first argument offered by proponents of grade retention is that it allows the retained students to acquire the basic academic skills that they did not master the previous year. Having more time to understand the important concepts and skills, on this view, enables students to achieve a sufficient level of competence to succeed at the following grade level – otherwise the learning content at that level would be too difficult for them, setting them up for failure (Piaget and Inhelder, 1962). On this view, extended teaching in the early grades (due to retention) is likely to enhance a student's likelihood of continuing to higher levels of education. A second argument is that retention provides students with additional time to mature emotionally and/or socially (Shepard and Smith,

1989). Retaining those who have had difficulties adapting to their age group or grade level, on this argument, allows them benefit from an extra year to do so. It is also argued that repeating a grade can allow students to recover self-confidence by succeeding where they have previously failed (Goos et al., 2021). And some suggest that overcoming a past difficulty can increase students' motivation to study and work. Third, some educators argue that having the lowest-performing students repeat their grade makes classes more homogeneous. This, it is argued, enables teachers to teach more effectively and enhance the achievement of all students, both repeaters and non-repeaters (Vygotsky, 1978). Finally, it is argued that repeating a grade can develop the idea of the need to maintain effort must be maintained even in case of failure. By repeating a grade, students learn that perseverance is essential to overcoming academic challenges. It is also said to serve as a strong signal that previous efforts were not enough, motivating retained students to take a more serious approach to their schoolwork (García-Pérez et al., 2014). Correspondingly, the threat of grade retention has been described as a motivational "stick", a(n anticipated) negative consequence of the failure to achieve important goals (Jacob, 2005; Ryan and Deci, 2020). The integration of the expected increase in the drive for effort and success might have broader implications for the individual's development and affect their later efforts, and ultimately their outcomes, on the labor market, such as during the job search process.

Opponents of grade retention, on the other hand, argue that retained students tend to feel discouraged and demotivated. They feel discouraged, on this account, because they experience retention as failure, and demotivated because they have to revisit a program they have already been through. This, it is argued, is likely to reduce their engagement and participation in class (Buhs and Ladd, 2001). In addition, critics of retention argue, the cost of readjusting to a new class and potentially a new teacher can worsen students' outcomes and lead to dropout. Overall, it can lead students to develop a negative attitude towards school and learning, viewing education as a source of frustration and failure (Jimerson et al., 1997). Repeating a grade may also negatively affect students' self-image, as it may leave them feeling less capable than their peers (Shepard and Smith, 1989; Manacorda, 2012). It is also argued that self-esteem can be affected by being with a group of younger students, while those of the same age are in a higher grade. Another criticism is that students who repeat a grade can be stigmatized by instructors or fellow students as a result, leading to feelings of isolation and marginalization (Jimerson, 1999; Gleason et al., 2007). The combination of low self-esteem and stigmatization, critics say, can have impacts on educational careers due to self-selection and self-fulfilling prophecies, leading these students to choose less demanding academic programs or even to drop out (Stearns et al., 2007; Andrew, 2014). Finally, it is argued that grade retention can have consequences for entry into higher education programs or the labor market (Tafreschi and Thiemann, 2016) , for two reasons. First, they argue, it sends a negative signal

to program administrators when selecting students and/or to recruiters when reviewing applications. Second, retained students spend additional time in education, forgoing labor income during this period.

There are also theoretical mechanisms to explain why the academic achievement of retained students varies with their family's social status, with grade repetition more detrimental to children from families with lower social status (Contini and Salza, 2024). One is that ambitious educational trajectories desired or planned by parents with high social status are less likely to be affected by their children repeating a grade. In such cases, these families are more likely to take effective action to ensure that their children continue their education in the most advantageous schools or tracks (Bernardi, 2014; Heiskala et al., 2020). More generally, parents with high social status are more likely to use their resources to enhance their children's chances of success and opportunities to pursue longer educational tracks; they can invest in private tutoring, transfer their children to schools offering similar tracks but with lower requirements etc. In contrast, due to a lack of resources, low-status families may be more likely to decide or be compelled to accept a reorientation of their children into less demanding (and therefore less recognized) educational tracks or shorter programs. Grade repetition may also be perceived differently depending on social status. High-status parents may see it as a temporary setback to be overcome with appropriate strategies. Low-status parents may associate it with greater stigma and have a stronger perception of failure, which can lead indirectly to decreased motivation for the child and a shift towards less valued educational tracks.

Potential arguments on how gender could interact with the effects of grade retention are also mixed. On the one hand, educational literature suggests that girls generally have better attitudes towards school, work harder, and tend to achieve better academic results than boys (Goldin, 2006; Fortin, 2015; Autor et al., 2019). Based on these factors, girls may be expected to demonstrate greater resilience in the event of grade repetition, potentially leading to some positive effects on academic achievement. On the other hand, the literature also indicates that girls are more prone to anxiety and issues with self-confidence (Jacobs et al., 2002; OECD, 2015). An experience of grade repetition as stigmatizing could exacerbating these already stronger feelings, and have a negative effect on their academic performance. Whether the effect of grade retention is predominantly positive or negative, the impact of grade repetition may differ between boys and girls.

2.2 Empirical findings

There is an extensive literature on the effects of grade retention on student outcomes (Jimerson, 2001; Xia and Kirby, 2009; Goos et al., 2021; Valbuena et al., 2021). The vast majority focuses on the consequences in terms of academic achievement, while there is little to no research examining the impact on labor market outcomes. Studies examining

the impact of retention have traditionally been conducted within the field of educational research. However, most have faced challenges related to endogeneity or selection bias, which remain the main obstacles to identifying the effects of retention (Contini and Salza, 2024).Various empirical strategies have been used to address these problems: regression discontinuity, instrumental variable methods, matching strategies, etc. In this subsection, we present a review of the literature on grade retention in the field of education economics, analyzing its effects on student outcomes, with a focus on the identification strategy.

The method used to evaluate the effect of grade retention partly depends on retention rules, which vary across countries. In this literature, many studies have used a regression discontinuity design for identification, exploiting the fact that, in some countries, grade repetition is determined by rigid rules based on standardized tests. Using such a strategy, Jacob and Lefgren (2004) examine the causal effect of grade retention on student achievement by comparing Chicago students who are close to the achievement threshold. They observe that retention increases achievement for third-grade students, but has little effect on mathematics achievement for sixth-grade students. The authors also examine the long-run effects of retention, and find that among sixth-grade students, retention does not affect the likelihood of secondary school completion (Jacob and Lefgren (2009)). However, retaining low-achieving eighth-grade students increases their probability of dropping out. Using the same sample of Chicago students, Roderick and Nagaoka (2005) examine the progress of retained students two years later and assess the short-term effects on reading achievement. The authors find no substantial positive effects for third-graders and negative effects for sixth-graders (retention being associated with lower achievement). Greene and Winters (2007); Winters and Greene (2012) exploit a discontinuity in retention probabilities under Florida's test-based promotion policy to study the effects of retention on student outcomes one or two years later. They find that students retained in the third grade outperform their promoted peers when they reach the same grade level. Using the same dataset, Schwerdt et al. (2017) confirm that third-grade retention significantly improves reading and mathematics achievement in the short term, and also reduces the probability of being retained in later grades. Manacorda (2012) measures the effect of grade failure (and thus retention) on subsequent school outcomes for upper secondary school students in Uruguay. Using a regression discontinuity design, he shows that grade failure led to significant dropout rates and lower educational attainment even four to five years later. Eren et al. (2017) examine, among other things, the potential effects of grade retention on secondary school completion. Using data from Louisiana and a regression discontinuity design, they find that grade retention in grades 4 and 8 increases the propensity to drop out of school. Most recently, Figlio and Özek (2020) use data from Florida to assess the effects of early grade retention on the short, medium, and long-term outcomes of English learners. They find significant benefits of this early grade retention policy: it improves students' English skills, reduces the time needed to achieve proficiency

and decreases the likelihood of enrolling in a remedial English course in lower secondary school.

In countries where grade retention rules are less clearly defined, alternative methods have been employed to assess its effects. For instance, some studies use instrumental variable (IV) techniques to estimate causal relationships while accounting for potential issues of omitted variables and selection bias. For example, d'Haultfoeuille (2010) estimate the short-term effects of grade retention among fifth-grade students in France. He finds that the short-term effect of grade retention is more likely positive. This result is consistent with the findings of Jacob and Lefgren (2004, 2009) for third-grade students in Chicago, but more positive than their findings for sixth-graders. In another study using French data, Alet et al. (2013) estimate the effect of grade repetition on subsequent school performance. Using a simultaneous equations model in which identification is ensured through IV techniques, they find that the effect of retention in grade 1 or 1 on test scores in grade 3 is positive, while the effect on test scores in grade 6 is negative. This result suggests that the initial positive effect is transitory. Using international data from multiple waves of the PISA international assessment test, Diris (2017) evaluates the effects of delayed school entry and grade retention in primary school on academic achievement. While the results of his IV models highlight a negative effect of primary school grade retention, he finds that delayed school entry can produce positive results, especially for students with the lowest performance and for girls. Mahjoub (2017) estimates the treatment effect of grade repetition in French upper secondary schools. He finds that repetition in grades 6-8 increases the probability of reaching grade 10 by 2.5 percentage points.

Other estimation strategies include difference-in-differences and propensity score matching approaches. The former can be used to compare retained students with comparable non-retained students, both before and after the decision to repeat a grade. By looking for differences in the trajectories of the two groups in terms of academic achievement, a potential effect of repetition can be identified. The latter is generally used to choose groups of retained and non-retained students who are comparable in terms of a selection of observable characteristics (such as gender, socioeconomic status, etc.). Differences in academic achievement can then be attributed to grade repetition rather than to differences related to these characteristics. For example, Chen et al. (2010) use both methods to examine the effect of grade retention on the educational performance of students in elementary schools in Shaanxi province (China). They find no significant positive effect of grade retention on school performance. Using propensity score matching methods on a panel of low- income minority students in Chicago, Ou and Reynolds (2010) instead find a negative effect of retention in primary school on participation in higher education. Goos et al. (2013) construct a two-level logistic regression model to examine the effects of first-grade retention on children's academic growth and future school career in Belgium. They find that students who repeat a year outperform their at-risk peers who

are promoted (in math and reading fluency during the retention year), but this benefit is very short-lived. Lamote et al. (2014) focus on the effect of grade retention in primary education on language achievement. The results of their propensity score matching analysis suggest that grade retention has no negative effect on achievement in the short term. However, the effects become negative in the long term. More recently, Nunes et al. (2018) found that in some situations, retention might have a positive, but small impact on future achievement for Portuguese grade 4 students. Finally, Contini and Salza (2024) use a highly original matching strategy that leverages the variation in exposure to grade retention within Italian high schools to estimate its impact. Their approach accounts for differences in school environments and student characteristics, providing a more precise evaluation of the effects of grade retention. They find that grade retention has a negative impact on educational outcomes by dramatically increasing dropout rates.

To the best of our knowledge, there are only a few studies investigating the effects of grade retention on labor market outcomes, particularly on wages. Eide and Showalter (2001), using instrumental variable techniques, find that grade retention can benefit students by increasing their labor market earnings. Similarly, using fixed effects regressions, Babcock and Bedard (2011) find that a 1 SD increase in grade retention through grade 2 is linked to a 0.7 percent rise in average hourly wages for males. This positive wage effect does not appear to be confined to the lower end of the wage distribution, but is observed throughout the entire distribution. In contrast, also using instrumental variable techniques, Brodaty et al. (2008) find a robust, significant, and negative effect of delay on wages: a one-year delay is associated with a decrease of approximately 9% in starting wages.

On the whole, research on the effects of grade retention on students' educational outcomes has produced mixed results. Some studies find no effect, while others observe only limited positive or negative effects, whether in the short term or the long term. These findings often depend on the grade level(s) considered, the estimation method, and differences in retention practices between schools and countries. There is also considerable heterogeneity based on factors such as gender and SES.

3 Institutional context

3.1 Grade repetition in European educational systems

Most European countries have specific regulations on grade repetition for students struggling to meet educational objectives (Borodankova and Coutinho, 2011). While the majority allow grade repetition, it is often subject to certain restrictions (*e.g.* it is limited to primary school, or there is a cap on the number of years a student can repeat). However, some countries, such as the United Kingdom lack specific regulations. Instead, they

consider that education should be tailored to a student's age, abilities, and skills. Even among countries with similar rules, however, the implementation of grade repetition varies widely. For example, Spain and Luxembourg have much higher repetition rates in primary education than Cyprus or Slovakia. Furthermore, some countries where repetition is permitted do not implement it extensively. In Greece, for instance, a complex process determines whether a student should repeat a grade, while in Italy, unanimous agreement among teachers is required. Both limit the implementation of grade retention in practice. In Norway and Iceland, automatic promotion is the norm, while in Bulgaria, grade retention is implemented primarily in primary education. In addition, countries that involve parents in the decision process, such as Denmark and Sweden, tend to have lower repetition rates, although the final decision is usually made by the school.

Thus, the practice of grade repetition is not universally widespread across educational systems (Gary-Bobo et al., 2008). In some countries, repetition is common and students only advance to the next grade if their exam results are deemed satisfactory. Students who do not succeed are either guided toward tailored pathways or required to repeat the grade. This model, observed notably in France and Germany, emphasizes mastery of knowledge before progression. Conversely, in countries with policies favoring mass education, such as the Scandinavian countries and the United Kingdom, automatic promotion predominates: in general, students advance regardless of their academic performance. These opposing approaches reflect the diversity of educational systems, from heavily repetition-based to automatic promotion models.

Factors beyond regulations, such as tradition, cultural elements, and societal beliefs about the benefits and effectiveness of grade repetition significantly influence its implementation. Belgium, Luxembourg, Spain, Portugal, the Netherlands, and France are all countries where repetition is widely practiced for these reasons (Goos et al., 2013). The use of repetition varies significantly between countries and evolves over time, reflecting the institutional cycles of educational systems within broader social systems. Grade repetition may emerge or re-emerge under specific social, economic, and political conditions (Gary-Bobo and Robin, 2014). In some contexts, it may ultimately serve as the primary form of academic support, even though its effectiveness and cost are often debated.

3.2 The French education system and the practice of grade repetition

French primary schooling is organized into three educational "cycles", each with distinct learning objectives. The first cycle, called the "early learning cycle", covers preschool years for children aged 3 to 6 and focuses on socialization, language development and motor skills. The second cycle, known as the "fundamental learning cycle", spans grades 1 to 3 and emphasizes the mastery of reading, writing, basic math, arts and physical

education. The third cycle, referred to as the "consolidation cycle", extends from grades 4 to 5 and aims to deepen students' knowledge of subjects such as history, geography, and the sciences, while fostering autonomy and preparing students for secondary school.

Lower secondary school ("*collège*" in French) serves (non-retained) students aged 11 to 15 and encompasses grades 6 through 9. The four years of lower secondary school are themselves divided into two distinct cycles. The first, called the "consolidation cycle," covers grades 6 and 7. This cycle is focused on reinforcing students' knowledge in core subjects such as French, mathematics, science, history-geography and modern languages. The second, known as the "in-depth cycle," comprises grades 8 and 9. In these years, students deepen their understanding of various subjects and prepare for the Diplôme National du Brevet (DNB), the lower secondary diploma validated through exams at the end of grade 9, which conditions access . In this final year, some students also have the option to pursue a pre-vocational pathway.

In France, grade repetition has long been used to help students consolidate their knowledge before progressing to the next level. Over recent decades, this practice has been increasingly debated, and its implementation reduced, due to concerns about its effectiveness and impact on student motivation. In 1960, over half of primary school students completed their education with at least one year of delay. By 2004, this figure had decreased to approximately 21%, while by 2022 it had dropped to below 2% (Mattenet and Sorbe, 2014; Dauphin et al., 2022). Today, grade repetition is most prevalent in grade 1. The large decline in the past two decades is partly attributable to the decree of November 18, 2014, which prohibited grade repetition in kindergarten and significantly limited its application in primary and lower secondary schools. Under this decree, grade repetition is now permitted only to "address a significant disruption in learning."

In primary school, teachers and educational teams must deliberate to determine whether repeating a year would genuinely benefit a student. Grade repetition may be proposed if a student has not acquired essential skills at the end of a cycle (grades 2] and 5 are key years). Alternative solutions to retention, such as individualized tutoring and enhanced support programs, are prioritized to help struggling students progress. In lower secondary school, grade repetition is still allowed, but has become increasingly rare. This decline is largely due to the 2014 reform, which introduced stricter conditions for its use. In 2022, the average repetition rate in the first three years of lower secondary school (grades 6 to 8) was about 1%. It was higher in grade 9, with 2.2% of students repeating (Dauphin et al., 2022). With the reform, grade repetition can only be proposed in exceptional situations, such as when students face significant and persistent difficulties despite existing support measures. The decision must involve the agreement of the student's parents, and cannot be imposed without consultation. At the end of grade 9, grade repetition is sometimes suggested if a student's performance does not allow them to pursue their desired path in high school. However, such decisions are often avoided in favor of tailored pathways in

vocational high schools or specialized second-year classes. In these cases, grade repetition plays a strategic role, serving as a mechanism to align students' academic progress with their long-term goals or to reconcile the gap between family ambitions and school recommendations.

During the period covered by the survey used in this study, grade repetition was widespread. For instance, in 1993, 46% of students in grade 9 were at least one year behind in school. By 2013, this figure had fallen to 24%. In the 2012 PISA survey, grade repetition rates in France remained the fifth-highest in the OECD, despite a significant decline (11 percentage points in grade 9 repetition rates between PISA 2003 and PISA 2012). Although grade repetition has become less common over the years, until 2013 it remained fairly widespread due to permissive regulations that left schools with considerable leeway to use it. The implementation of the 2014 decree drastically limited its use in primary and lower secondary schools.

In this context, there are considerable advantages to using the 1995-96 grade 6 cohort to study the effects of grade repetition. Nearly one in three members of this cohort repeated at least one grade by the time they completed or left lower secondary school. More recent cohorts would include fewer grade repeaters with more specific characteristics (more severe academic difficulties or disrupted learning). In these later cohorts, using matching methods (described below) with highly diverse samples in terms of observable characteristics would make comparisons more difficult and less informative. Conversely, the large number of grade repeaters in the 1990s creates more homogeneous groups (*i.e.* of repeaters vs. non-repeaters), facilitating the matching process by reducing the heterogeneity in observable characteristics.

4 Data and descriptive statistics

4.1 Data

The data for this study come from the Panel d'élèves du second degré, recrutement 1995[1,2] [Panel of lower secondary school students, 1995 recruitment], a panel survey conducted by the French Ministry of National Education. The students were subsequently followed in higher education and during their initial years in the labor market by the National Institute of Statistics and Economic Studies (INSEE) as part of the Enquête sur l'entrée dans la vie adulte[3] (Survey on entry into adult life, or EVA). This data initially tracks a sample of 17,830 students who entered grade 6 in the 1995-96 academic year, and provides

[1]Panel d'élèves du second degré, recrutement 1995 - 1995-2006 - (2006) [electronic file], DEPP [producer], Centre Maurice Halbwachs (CMH) [distributor

[2]For the remainder of this study, we will refer to these data as the "Panel 1995".

[3]Enquête sur l'entrée dans la vie adulte des élèves entrés en 6ème en 1995 (EVA) - 2005-2012 - (2012) [electronic file], INSEE [producer], ADISP [distributor].

a comprehensive observation of their educational pathways through secondary school and their transitions to higher education and the labor market. The dataset provides detailed information on individual characteristics, educational pathways, and early labor market experiences. It thus offers a particularly rich longitudinal perspective on a single cohort of students in France.

The available data in this panel include individual sociodemographic and educational characteristics as well as school characteristics. These are supplemented by post-secondary information, such as the chosen field of study, the degree obtained, employment status, and income level once the students have exited the education system. Data on higher education were collected semi-annually for the first nine years (for students who had not repeated a year) after students complete secondary school, while labor market information, including salary and employment status, was updated annually between 2005 and 2012. This structure enables individuals to be followed up to 17 years after their entry into grade 6, providing valuable insights into their long-term trajectories.

The richness of this longitudinal data set allows for a detailed analysis of the long-term trajectories of students who experienced grade retention, making it particularly valuable for assessing the impact of retention on educational and labor market outcomes over an extended period. However, partly as a result of this extensive follow-up, the dataset is subject to some potential biases related to attrition. With such long-term data collection, certain individuals may not respond to surveys due to moving abroad or becoming unreachable for some other reason. Additionally, some individuals did not report their income, resulting in a degree of heterogeneity in the coverage of financial variables within the sample.

The four primary outcome variables fall into two categories. The first are educational outcomes: the completion of a secondary school diploma (*baccalauréat*) and a higher education qualification. The second are labor market outcomes: employment status and monthly wage. Completion of the *baccalauréat* is defined as holding any of these three types of secondary school diplomas in France—general (academic), technical, or vocational—regardless of the year of attainment. The variable for a higher education qualification is a binary variable indicating whether the student's highest level of qualification in 2012 represents at least two years of post-secondary education. Employment status indicates whether the individual was employed at the time of follow-up in 2012, while the monthly wage reflects their reported income in 2012, providing an estimate of their early-career earnings.

Finally, the sample is restricted to students with measurable characteristics from grade 6, including entrance test scores for lower secondary school, which are used for estimating propensity scores in the matching analyses. We restrict the sample to students with recorded scores in French and mathematics in both grade 6 and grade 9, who completed their lower secondary education in the general academic track—thus ensuring a homoge-

neous sample in terms of initial educational pathways. To be included, individuals may have repeated a grade in lower secondary school, but no more than once. Our final sample includes 11,225 students, among whom one quarter (2716 students) repeated a grade once during lower secondary education.

4.2 Descriptive statistics

The results presented in the table 1 compare the individual sociodemographic and educational characteristics of students who did and did not repeat a grade, and those of their schools. The results show significant differences between the two groups. For instance, 45.4% of students who repeated a grade are from low socioeconomic status (SES) backgrounds, versus 30.4% of students who did not. Conversely, students from high SES backgrounds were less likely to repeat a grade: they made up only 10.2% of repeaters, versus 24.8% of non-repeaters. Parental education also plays a key role: 51% of students who repeated a grade had mothers who held no educational qualifications or whose level of education was not reported, versus only 36.8% of non-repeaters. Similarly, the results on father's level of education show that students with fathers holding no qualifications were much more likely to repeat a grade (they made up 51.5% of all repeaters, vs. only 37.2% of non-repeaters). Family structure is another factor. Repeaters were less likely than non-repeaters to come from two-parent households (78.1% vs. 84.6%) and more likely to come from single-parent households (16.8% vs. 11.8%). In terms of school characteristics, grade retention is more common among students attending schools in priority education zones, with 11.3% of repeaters attending these schools in grade 6 , compared to 9.4% of non-repeaters. In terms of students characteristics, nsurprisingly, grade retention is strongly linked to academic performance in the first year of lower secondary education: students who repeated scored lower on both math and French tests in grade 6, with average scores of 44.8 and 41.4 respectively, compared to 57.0 and 51.2 for those who did not repeat. These findings highlight the strong influence of socioeconomic status, family background, and early academic performance on the likelihood of grade retention.

Table 2 presents key educational and labor market outcomes by grade retention status. Students who repeated a grade were far less likely to obtain a secondary school diploma (24.8%, vs. 74.55% of non-repeaters). Similarly, grade repeaters were less likely to obtain a higher education qualification (21.9% vs. 62.4%). Interestingly, the difference in employment status between the two groups is not statistically significant, with 85.1% of grade repeaters and 87.2% of non-repeaters being employed in 2012. However, there was a substantial wage difference between the two group, with those who did not repeat earning a significantly higher monthly wage (€1736) than those who repeated (€1490), suggesting that grade retention is associated with long-term differences in earnings.

5 Empirical approach

5.1 Estimation strategy

In Table 3, we compare the results obtained by applying two different methods, instrumental variables (IV) estimation and propensity score matching (PSM), to two outcomes, test scores in mathematics and reading in grade 9. Brodaty et al. (2008) conduct a similar analysis using data from the 1995 panel, focusing mainly on comparing different IV specifications to test the stability of these methods. For this comparison of methods, we use data from the 2007 panel to estimate the effect of grade retention on academic outcomes, because it includes standardized evaluations in grade 6 and national exam scores for grade 9, which are not influenced by the school context and ensure comparability. In contrast, the data on academic performance in grade 9 from the 1995 panel represent a mixture of test scores and continuous assessment, which cannot be separated and which are affected by the school environment. They are thus not as suitable for comparison with the standardized grade 6 evaluations (Cayouette-Remblière and Moulin, 2019). Three specifications are tested for each method: (1) a regression estimating the effect of grade retention without controlling for initial scores, (2) a regression including initial mathematics or reading scores, and (3) a regression including both initial mathematics and reading scores.

The results show that PSM is more robust across the specifications in terms of coefficient stability, as reflected by the low standard errors. For instance, in mathematics, the estimated coefficients for the effect of grade retention with PSM vary slightly across specifications, from 0.350 to 0.054, with low standard errors (0.019-0.020). In comparison, the IV method produces coefficients with much higher standard errors (e.g., 1.070 and 5.973), especially in the adjusted specification. Moreover, PSM consistently yields significant coefficients across all specifications, with very low p-values. In contrast, the coefficients calculated using the IV method in the adjusted specification are not significant.

Thus, PSM appears to provide more stable results, with consistently significant coefficients and reduced standard errors, suggesting that this method offers a more robust way to evaluate the effect of grade retention on mathematics and reading outcomes.

5.2 Propensity score matching

The objective of this study is to estimate the causal impact of grade retention in lower secondary school on various long-term outcomes. To address the inherent selection bias in grade retention, we employ propensity score matching[4] (PSM), grounded in the potential

[4]The implementation of the PSM procedure used here is inspired by the presentation of the method in the extensive works of Marco Caliendo and his co-authors (see for example Caliendo and Künn, 2011; Caliendo et al., 2016; Caliendo and Tübbicke, 2022).

outcomes framework developed by Roy (1951) and Rubin (1974). In France, the decision to have students repeat a grade is primarily based on academic performance, but there is no specific threshold or standardized criterion. In the period covered by this study, discretion was left mainly to the student's teachers. Teachers, often with input from the school council, would assess a student's preparedness to advance to the next grade based on their academic results and overall behavior. This lack of an objective, fixed standard means that various factors, such as family support or socioeconomic conditions, can influence the decision to retain a student. In this context, PSM is particularly well suited to estimating the causal effect of grade retention (Contini and Salza, 2024), as it controls for selection bias by balancing observable characteristics between students who repeat a grade and those who do not. By equating these characteristics across groups, PSM can improve the estimation of the impact of grade retention, in a way that resembles a comparison with a control group.

Each student has two potential outcomes: $Y(1)$ if they repeated a grade in lower secondary school (treatment group, $D = 1$), and $Y(0)$ if they did not (control group, $D = 0$). The parameter of interest is the average treatment effect on the treated (ATT), calculated as:

$$\tau_{ATT} = E[Y(1)|D = 1] - E[Y(0)|D = 1], \tag{1}$$

where $E[Y(0)|D = 1]$ is the hypothetical outcome for students who repeated a grade, had they not repeated. Relying on data from students who did not repeat, without proper adjustment, would lead to selection bias, as assignment to the treatment and control groups is not random but influenced by differences in both observed and unobserved characteristics. PSM addresses this bias by matching students based on observed characteristics, creating a comparable control group. PSM relies on the conditional independence assumption (CIA) or assumption of unconfoundedness, which holds that any systematic differences in outcomes between treated and comparison groups, after controlling for covariates, can be attributed solely to the treatment. Formally, the CIA can be stated as follows:

$$Y(0) \perp D|P(X), \tag{2}$$

where $P(X) = \Pr(D = 1|X)$ represents the propensity score: in other words, the probability of receiving the treatment given observed covariates X. For the CIA to hold, this implies that all variables influencing both treatment assignment and potential outcomes must be observed and included in the analysis. While this is a strong assumption, and must be justified by the quality and richness of the available data, it is essential for attributing differences in outcomes to the treatment itself. In this study, we include sociodemographic, family, school, and academic performance variables, making the assump-

tion that they capture the key factors susceptible to influencing both retention decisions and outcomes.

PSM also relies on the common support condition, which requires that individuals have a positive probability of being both treated and untreated across the range of X values:

$$0 < \Pr(D = 1|X) < 1, \quad \forall X. \tag{3}$$

This condition prevents the treatment status from being perfectly predictable based on covariates, ensuring that both groups (retained and non-retained students) can be compared across the full range of observable characteristics.

Assuming that the conditional independence and common support assumptions are met, the ATT can be estimated as the average difference in outcomes between matched treated and control students. It is calculated as:

$$\tau_{ATT} = E[Y(1)|P(X), D = 1] - E[Y(0)|P(X), D = 0]. \tag{4}$$

We implement various matching algorithms—including nearest neighbor matching, kernel matching, radius matching, and inverse probability weighting (IPW)—to ensure the robustness of our estimates to the choice of matching method. Additionally, we conduct two further robustness checks. First, the bounding method (Rosenbaum, 2002) assesses how much unobserved bias would be needed to invalidate the observed effects, providing a measure of sensitivity to potential hidden biases. Second, simulation-based sensitivity analysis (Nannicini, 2007; Ichino et al., 2008) introduces a hypothetical unobserved confounder to evaluate the impact of unobserved factors on the estimated treatment effects, ensuring robustness to potential hidden heterogeneity.

5.3 Estimation procedure

The first step in the PSM procedure involves estimating the likelihood of grade retention using a logistic regression model. The results are presented in Table 4. They reveal that several sociodemographic, family, and school-related factors are significantly associated with the likelihood of grade retention. First, gender plays a key role, with boys more likely to repeat a grade than girls (coefficient: -0.588). Socioeconomic status (SES) is also an important factor: students from lower-SES backgrounds, particularly those in the medium-low and low SES categories, were more likely to repeat a grade than those from higher SES backgrounds. Parental education levels also significantly affected the likelihood of grade retention. Students whose parents have lower educational attainment, especially those with no qualifications or with vocational education, were more likely to repeat a grade. For example, having a mother or father with no educational qualifications

increased the likelihood of grade retention (0.319 and 0.403 respectively). Family structure also has an impact: students from single-mother households were more likely to repeat a grade (0.545) than those from two-parent households. Additionally, having parents who were born outside of France reduces the likelihood of grade retention (−0.344), meaning that students from immigrant backgrounds were less likely to repeat a grade. School-related characteristics also show significant associations. Students attending schools in rural areas or smaller towns were less likely to repeat a grade compared to those in larger urban areas. Furthermore, students attending grade 6 in a priority education zone were less likely to repeat a grade (−0.646), as were those who had previously repeated a grade in primary school (−0.724). Finally, and unsurprisingly, academic performance, as measured by grade 6 test scores in mathematics and French, has a strong inverse relationship with grade retention. Higher test scores in both subjects significantly reduced the likelihood of grade retention.[5]

5.4 Matching quality

Table 5 looks at the balance of characteristics between repeaters and non-repeaters before and after the matching process. Before matching, there are significant differences in means at the 1% level for 18 variables. After matching, the number drops to zero, demonstrating much-improved balance between the treatment and control groups. Standardized bias also decreases significantly, with the mean absolute bias dropping from 16.67% before matching to just 1.40% after, well within the 3–5% threshold established by Caliendo and Kopeinig (2008), confirming the robustness and quality of the matching process. This substantial reduction in bias means that the differences in outcomes can be more reliably attributed to the treatment, as confounding variables are now well-balanced between treatment and control groups. Pseudo-R^2 falls from 0.2551 to 0.0021, and the joint significance test confirms the quality of the match, with a p-value of 0.9998 after matching, indicating that no significant differences remain between the groups. These results confirm the quality of the matching process. The significant reduction in standardized bias, along with the near elimination of differences in means between the treatment and control groups, demonstrates that the matching successfully balanced the covariates.

The distributions of estimated propensity scores for both the treatment and control groups are shown in Figure 1. Before matching, there is a clear difference in the propensity

[5]Because we use the 2007 panel in 3, we replicate the analysis from Table 4 in Table A.2 using this panel. The results are relatively similar between the two tables. The key difference between the two tables lies in the effect of attending a private school: while being enrolled in a private school had no impact on the likelihood of grade retention for the 1995 panel, it had a significant negative effect for the 2007 panel, reflecting changes in retention policies over time. This shift likely reflects a change in the approach of private schools, where stricter academic standards or policies may have been implemented, leading to a higher likelihood of grade retention in 2007. Private schools may thus have placed more emphasis on maintaining academic performance through grade retention, diverging from public school policies during this period.

score distributions between the two groups, with the treatment group generally displaying higher scores. This pattern suggests that, prior to matching, individuals in the treatment group (those retained in a grade) had an inherently higher likelihood of retention, likely due to factors correlated with their initial propensity for retention, such as socioeconomic background or prior achievement. After matching, the distributions are closely aligned, indicating that the matching process has successfully balanced the propensity scores across the treatment and control groups. The fulfillment of the common support condition ensured that no observations had to be excluded, thus preserving the full sample for analysis.

6 Estimation results

6.1 Grade retention effects

Table 6 presents the effects of grade retention on educational and employment outcomes.[6] The results indicate that being held back a grade has a significant negative impact on the likelihood of obtaining both a secondary school diploma (-0.211) and a higher education qualification (-0.238). These negative coefficients suggest that grade retention substantially reduces the likelihood of reaching these key academic milestones. These findings are in accordance with those of Roderick and Nagaoka (2005), who studied a sample of students in Chicago followed longitudinally from grade 6, and Manacorda (2012), who focused on a sample of students in Uruguay followed longitudinally from lower secondary school. Both studies show that grade repetition leads to higher dropout rates and lower educational attainment. Similarly, Contini and Salza (2024) also demonstrate that grade retention in high school negatively impacts students' educational outcomes by significantly increasing dropout rates. However, previous research focusing on the French system has produced somewhat different results, finding some positive effects of grade repetition on academic achievement. For instance, d'Haultfoeuille (2010) provides evidence that the short-term effects of grade retention among fifth-grade students are more likely to be positive. Additionally, Mahjoub (2017) finds that grade repetition increases the probability of obtaining a secondary school diploma (*baccalauréat*). These contrasting conclusions can be explained by variations in study samples and differences in estimation methods, reflecting varying approaches to addressing potential biases.

[6] As robustness check, we modified the grade repetition variable in primary school, which is used as a covariate, by imputing missing values based on students' date of birth. This adjustment reduced the number of missing observations in the estimations. We focus on this variable because grade repetition in primary school strongly affects the likelihood of repeating a grade in secondary school (see Table 4). The results, presented in Table A.2, are consistent with the main findings reported in Table 6 Specifically, the negative and statistically significant effects observed for secondary school completion, obtaining a higher education qualification, and monthly wage remain unchanged, confirming the robustness of our analysis. This suggests that our conclusions are not sensitive to missing data in the grade repetition variable.

Interestingly, the results do not show a significant effect of grade retention on the likelihood of being employed. This suggests that while retention may harm academic outcomes, it may not immediately translate into lower employability, potentially because employment prospects depend on various other factors beyond educational attainment, such as personal networks or job market conditions. However, the significant negative impact of grade retention on monthly wages (-0.059) suggests that when students who were retained find employment, they tend to earn less than their peers who were not retained. This finding is in accordance with Brodaty et al. (2008), who find a negative effect of school delay on wages. This wage penalty might reflect slower career progression, lower job quality, or the long-term effect of delayed entry into the labor market due to extended schooling. Together, these findings underscore the lasting negative consequences of grade retention, not only in terms of academic achievement but also in terms of future earning potential.

While few, if any, studies on grade repetition in the economics of education focus on such long-term effects, some do examine the consequences of grade repetition several years later. This is the case for Jacob (2005) and Figlio and Özek (2020), who, in examining the long-run effects of retention (at least five years later), find that it does not significantly affect educational outcomes. We then confirm over a longer period that repetition seems to produce strong and significant negative effects. Over time, the potentially negative effect is significantly diminished.

6.2 Heterogeneity

Table 7 presents the gender-specific effects of grade retention on educational and employment outcomes. For both boys and girls, grade retention has a significant negative impact on the likelihood of secondary school graduation and obtaining a higher education qualification. The measured effects are stronger for girls, which may be partly due to the fact that they are less likely to be held back. As a result, the subset of girls who do experience grade retention might differ in unobserved ways from the broader population, potentially making them more vulnerable to its negative impacts. Boys are more likely to be held back, suggesting that those who are retained may have a wider range of academic profiles. Girls who do repeat a grade may thus experience a more concentrated effect on their educational trajectory. In addition, girls are known to be more likely to experience anxiety and self-confidence issues (Jacobs et al., 2002). Grade repetition might exacerbate these feelings and thereby harm their academic performance. Nevertheless, this extra negative effect for girls has not always been found. For example, Contini and Salza (2024) find that grade retention negatively affects girls more than boys in traditional lyceums and technical schools, but not in other types of secondary schools. And the results of Alet et al. (2013) and Diris (2017) run in the opposite direction, suggesting that repetition

harms boys more than girls.

In our analysis, grade retention does not have a statistically significant effect on the likelihood of employment for either boys or girls. The key difference between the genders lies in the effect on monthly wages. For boys, grade retention is associated with a significant reduction in wages (-0.074), but not for girls. This suggests that the long-term wage penalty of grade retention may be greater for boys.

Table 8 presents the effects of grade retention on educational and employment outcomes by socioeconomic status (SES). For all SES groups, grade retention has a significant negative effect on secondary school completion and obtaining a higher education qualification, with similar effect sizes across the different SES levels. Our results thus argue against the hypothesis that grade repetition penalizes students from less advantaged social backgrounds more severely. Retained children whose parents have a higher socioeconomic status do not seem to benefit disproportionately from their parents' social capital and network. This contrasts with the recent finding of Contini and Salza (2024) that negative effects of repetition are stronger for students with low-educated parents.

The effects on the likelihood of being employed are not statistically significant for any SES group. Grade retention does have a socioeconomically differentiated effect on monthly wages, however. For the lowest SES group, grade retention is associated with a marginally significant reduction in wages (-0.052), while for higher SES groups, no effect on wages is found. This suggests that grade retention may have a more pronounced long-term economic impact on individuals from lower socioeconomic backgrounds. This may be because students from the lowest SES groups lack the supportive environments needed to mitigate the negative effects of retention, which tends to undermine self-confidence and career aspirations (Bernardi, 2014). It may thereby increase their chances of settling for lower-paying jobs, even when they possess the qualifications and skills required. Students from more advantaged groups benefit from stronger social and family networks, which can help preserve their sense of self-worth and ambition.

6.3 Robustness checks

In Table 9, we present the results of the impact of grade retention on secondary school completion, higher education qualifications, employment, and monthly wages using various matching algorithms. Whether using nearest neighbor (with or without replacement), kernel matching (with various bandwidths), radius matching, or inverse probability weighting (IPW), the results remain consistent in terms of size and significance. This stability across algorithms confirms the robustness of the results, indicating that the estimated effects are not sensitive to the specific matching technique applied. The standard errors remain reasonable across all methods, further confirming the reliability of the findings. The consistent negative impact of grade retention on academic outcomes and wages, as well as

the consistent lack of an effect on the likelihood of employment, suggests that the results are robust to different matching algorithms.

In Table 10, we use a bounding technique to assess the sensitivity of our results to unobserved heterogeneity, building on the framework of Rosenbaum (2002). This approach introduces a hypothetical unobserved factor, or confounder, to evaluate the extent of hidden bias required to nullify our findings on the outcomes of interest: secondary school completion, obtaining a higher education qualification, employment status, and monthly wage. Specifically, for each tested value of a hypothetical hidden bias factor Γ, which represents the extent of unobserved bias needed to invalidate the effect, we assess whether the estimated treatment effect remains significant. The results for educational attainment and wage outcomes are particularly robust. For both secondary school completion and obtaining a higher education qualification, the treatment effects remain significant up to a Γ of approximately 2.80 to 3.05, depending on the significance level. This high threshold implies a strong resilience to hidden biases, and suggests that even substantial unobserved heterogeneity would not easily undermine the observed effects for these educational outcomes. For monthly wage, the estimated effect remains robust up to a Γ of approximately 1.20 to 1.40, depending on the significance level. Although this result is weaker than the one for educational attainment, it still indicates strong robustness with respect to plausible levels of hidden bias. The employment outcome is already non-significant, and the sensitivity analysis with critical Γ values close to 0.00 confirms that this result is not driven by unobserved biases.

In Table 11, we use a simulation-based sensitivity analysis, inspired by Nannicini (2007) and Ichino et al. (2008), to evaluate the robustness of our estimated treatment effects to unobserved heterogeneity. This approach introduces a simulated unobserved confounder with characteristics similar to an observable variable, allowing us to assess how unobserved factors might impact our findings. The selection effect represents the impact of the unobserved confounder on the likelihood of being retained, where a selection effect below (above) 1 indicates a negative (positive) selection bias, suggesting lower (higher) retention likelihood for these groups. The outcome effect represents the estimated impact of a given variable on the outcome of interest, such as obtaining a higher education qualification or employment status. These simulated confounders demonstrate consistent patterns across outcomes, suggesting that unobserved factors are unlikely to alter the magnitude or direction of the ATT effects. For example, even when simulating an unobserved factor that mimics characteristics with strong selective effects, such as primary school grade repetition or high SES, the ATT on all outcomes remains stable across most subgroups, supporting the robustness of our baseline estimates. This analysis reinforces our confidence that grade retention consistently impacts educational and employment outcomes, indicating that our findings are unlikely to be influenced by potential unobserved heterogeneity.

7 Conclusion

This study provides a comprehensive analysis of the causal impact of grade retention in French lower secondary schools on a set of long-term outcomes, focusing on students' later educational attainment and labor market performance. Leveraging a unique longitudinal dataset that tracks a cohort of French students who entered grade 6 in the 1995-1996 academic year over a period of 17 years, we offer robust evidence on the consequences of grade retention using propensity score matching methods. This approach enables us to account for selection bias and to balance observable characteristics between retained and non-retained students, providing credible causal estimates. It also addresses some of the limitations of instrumental variable methods, which have previously yielded less stable results on student performance. The originality of our contribution lies in the use of French data to directly analyze the effects of grade retention itself, in contrast to the broader educational delays examined in Gary-Bobo et al. (2008). To the best of our knowledge, this is the first study to do so with European data.

Our results indicate that grade retention has significant negative effects on educational outcomes. Specifically, students who repeated a grade in lower secondary school are less likely to complete secondary school or obtain a higher education qualification. These effects are consistent across socioeconomic groups and for both boys and girls, underscoring the broad and persistent nature of these adverse consequences. When examining labor market outcomes, we find that while grade retention does not appear to affect the likelihood of being employed, it significantly reduces monthly wages. This wage penalty is observed for boys and students from low socioeconomic backgrounds, but not for girls or for students from other socioeconomic groups, suggesting that grade retention may contribute to labor market inequalities affecting these specific populations.

The mechanisms underlying these findings are multifaceted, and reflect the complex interplay between educational trajectories, psychological factors, and socioeconomic contexts. Grade retention may prolong students' time in the education system, leading to delayed entry into the labor market and a corresponding loss of potential labor income. It can also undermine students' self-confidence, motivation, and social integration, which may further hinder their ability to succeed academically and develop a stable career path. Additionally, grade retention may act as a negative signal to employers, who might perceive it as an indicator of lower ability or productivity, further contributing to the wage penalties observed. While these potential mechanisms provide a plausible explanation for our findings, the precise identification of these pathways is beyond the scope of the present study. Further research is needed to determine the relative contributions of delayed educational attainment, psychological impacts, and signaling effects to the long-term consequences of grade retention.

These results have important policy implications. First, the significant and lasting

negative effects of grade retention call for a reevaluation of its use as a remedial measure within education systems. Policymakers should consider alternative interventions, such as individualized support programs, early academic assistance, and tutoring, which may provide more effective and less costly solutions to help struggling students. That said, any discussion or evaluation of grade retention policies must prioritize their impact on academic outcomes, as these are the central focus of educational policy objectives. Second, efforts to reduce educational inequalities and ensure smoother transitions into higher education and the labor market must focus on mitigating the stigma and long-term consequences associated with grade retention.

Overall, this study highlights the negative long-term impacts of grade retention on educational attainment and labor market outcomes, emphasizing the need for evidence-based policies that better support struggling students.

Declaration of generative AI in scientific writing

During the preparation of this work the authors used ChatGPT in order to improve the clarity of the writting. After using this tool, reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

Declaration of competing interest

None.

References

Alet, E., Bonnal, L. and Favard, P. (2013), 'Repetition: Medicine for a short-run remission', *Annals of Economics and Statistics* (111/112), 227–250.

Andrew, M. (2014), 'The scarring effects of primary-grade retention? a study of cumulative advantage in the educational career', *Social Forces* **93**(2), 653–685.

Autor, D., Figlio, D., Karbownik, K., Roth, J. and Wasserman, M. (2019), 'Family disadvantage and the gender gap in behavioral and educational outcomes', *American Economic Journal: Applied Economics* **11**(3), 338–81.

Babcock, P. and Bedard, K. (2011), 'The wages of failure: New evidence on school retention and long-run outcomes', *Education Finance and Policy* **6**(3), 293–322.

Bernardi, F. (2014), 'Compensatory advantage as a mechanism of educational inequality: A regression discontinuity based on month of birth', *Sociology of Education* **87**(2), 74–88.

Borodankova, O. and Coutinho, A. A. (2011), *Grade retention during compulsory education in Europe : regulations and statistics*, Publications Office.

Brodaty, T., Gary-Bobo, R. J. and Prieto, A. (2008), 'Does speed signal ability? the impact of grade repetitions on employment and wages', *CEPR Discussion Paper* (DP6832).

Buhs, E. S. and Ladd, G. W. (2001), 'Peer rejection as antecedent of young children's school adjustment: An examination of mediating processes', *Developmental Psychology* **37**(4), 550–560.

Caliendo, M. and Kopeinig, S. (2008), 'Some practical guidance for the implementation of propensity score matching', *Journal of Economic Surveys* **22**(1), 31–72.

Caliendo, M. and Künn, S. (2011), 'Start-up subsidies for the unemployed: Long-term evidence and effect heterogeneity', *Journal of Public Economics* **95**(3-4), 311–331.

Caliendo, M., Künn, S. and Weißenberger, M. (2016), 'Personality traits and the evaluation of start-up subsidies', *European Economic Review* **86**, 87–108.

Caliendo, M. and Tübbicke, S. (2022), 'Do start-up subsidies for the unemployed affect participants' well-being? a rigorous look at (un-)intended consequences of labor market policies', *Evaluation Review* **46**(5), 517–554.

Cayouette-Remblière, J. and Moulin, L. (2019), 'How inequalities in academic performance evolve in lower secondary school in france: A longitudinal follow-up of students', *Population* **74**(4), 507–540.

Chen, X., Liu, C., Zhang, L., Shi, Y. and Rozelle, S. (2010), 'Does taking one step back get you two steps forward? grade retention and school performance in poor areas in rural china', *International Journal of Educational Development* **30**(6), 544–559.

Contini, D. and Salza, G. (2024), 'Children left behind. new evidence on the (adverse) impact of grade retention on educational careers', *European Sociological Review* p. jcae014.

Dauphin, L., Dieusaert, P., Juzdzewski, L. and Miconnet, N. (2022), Les élèves du second degré àla rentrée 2022, Technical Report 22.39, DEPP.

Dehejia, R. H. and Wahba, S. (2002), 'Propensity score-matching methods for nonexperimental causal studies', *The Review of Economics and Statistics* **84**(1), 151–161.

Diris, R. (2017), 'Don't hold back? the effect of grade retention on student achievement', *Education Finance and Policy* **12**(3), 312–341.

d'Haultfoeuille, X. (2010), 'A new instrumental method for dealing with endogenous selection', *Journal of Econometrics* **154**(1), 1–15.

Eide, E. R. and Showalter, M. H. (2001), 'The effect of grade retention on educational and labor market outcomes', *Economics of Education Review* **20**(6), 563–576.

Eren, O., Depew, B. and Barnes, S. (2017), 'Test-based promotion policies, dropping out, and juvenile crime', *Journal of Public Economics* **153**, 9–31.

Figlio, D. and Özek, U. (2020), 'An extra year to learn English? Early grade retention and the human capital development of English learners', *Journal of Public Economics* **186**(C).

Fortin, N. M. (2015), 'Gender role attitudes and women's labor market participation: Opting-out, aids, and the persistent appeal of housewifery', *Annals of Economics and Statistics/Annales d'Économie et de Statistique* (117/118), 379–401.

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

Gary-Bobo, R. J., Brodaty, T. and Prieto, A. (2008), Does Speed Signal Ability? The Impact of Grade Repetitions on Employment and Wages, CEPR Discussion Papers 6832, C.E.P.R. Discussion Papers.

Gary-Bobo, R. and Robin, J.-M. (2014), 'La question des redoublements. Analyse économique et problèmes statistiques', *Revue économique* **65**(1), 5–45.

Gleason, K., Kwok, O.-M. and Hughes, J. (2007), 'The short-term effect of grade retention on peer relations and academic performance of at-risk first graders', *The Elementary school journal* **107**, 327–340.

Goldin, C. (2006), 'The quiet revolution that transformed women's employment, education, and family', *American Economic Review* **96**(2), 1–21.

Goos, M., Pipa, J. and Peixoto, F. (2021), 'Effectiveness of grade retention: A systematic review and meta-analysis', *Educational Research Review* **34**, 100401.

Goos, M., Van Damme, J., Onghena, P., Petry, K. and de Bilde, J. (2013), 'First-grade retention in the flemish educational context: Effects on children's academic growth, psychosocial growth, and school career throughout primary education', *Journal of School Psychology* **51**(3), 323–347.

Greene, J. P. and Winters, M. A. (2007), 'Revisiting grade retention: An evaluation of florida's test-based promotion policy', *Education Finance and Policy* **2**(4), 319–340.

Heiskala, L., Erola, J. and Kilpi-Jakonen, E. (2020), 'Compensatory and multiplicative advantages: Social origin, school performance, and stratified higher education enrolment in finland', *European Sociological Review* **37**(2), 171–185.

Ichino, A., Mealli, F. and Nannicini, T. (2008), 'From temporary help jobs to permanent employment: What can we learn from matching estimators and their sensitivity?', *Journal of applied econometrics* **23**(3), 305–327.

Jacob, B. A. (2005), 'Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools', *Journal of Public Economics* **89**(5-6), 761–796.

Jacob, B. A. and Lefgren, L. (2004), 'Remedial Education and Student Achievement: A Regression-Discontinuity Analysis', *The Review of Economics and Statistics* **86**(1), 226–244.

Jacob, B. A. and Lefgren, L. (2009), 'The effect of grade retention on high school completion', *American Economic Journal: Applied Economics* **1**(3), 33–58.

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S. and Wigfield, A. (2002), 'Changes in children's self-competence and values: Gender and domain differences across grades one twelve', *Child Development* **73**, 509–527.

Jimerson, S., Carlson, E., Rotert, M., Egeland, B. and Sroufe, L. (1997), 'A prospective, longitudinal study of the correlates and consequences of early grade retention', *Journal of School Psychology* **35**(1), 3–25.

Jimerson, S. R. (1999), 'On the failure of failure: Examining the association between early grade retention and education and employment outcomes during late adolescence', *Journal of School Psychology* **37**(3), 243–272.

Jimerson, S. R. (2001), 'Meta-analysis of grade retention research: Implications for practice in the 21st century', *School Psychology Review* **30**(3), 420–437.

Lamote, C., Pinxten, M., Noortgate, W. V. D. and Damme, J. V. (2014), 'Is the cure worse than the disease? a longitudinal study on the effect of grade retention in secondary education on achievement and academic self-concept', *Educational Studies* **40**(5), 496–514.

Mahjoub, M.-B. (2017), 'The treatment effect of grade repetitions', *Education Economics* **25**(4), 418–432.

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

Mattenet, J.-P. and Sorbe, X. (2014), Forte baisse du redoublement : un impact positif sur la réussite des élèves, Technical Report 36, DEPP.

Nannicini, T. (2007), 'Simulation-based sensitivity analysis for matching estimators', *The stata journal* **7**(3), 334–350.

Nunes, L. C., Reis, A. B. and Seabra, C. (2018), 'Is retention beneficial to low-achieving students? evidence from portugal', *Applied Economics* **50**(40), 4306–4317.

OECD (2015), *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence*, PISA, OECD Publishing.

Ou, S.-R. and Reynolds, A. J. (2010), 'Grade retention, postsecondary education, and public aid receipt', *Educational Evaluation and Policy Analysis* **32**(1), 118–139.

Piaget, J. and Inhelder, B. (1962), 'Le développement des images mentales chez l'enfant', *Journal de psychologie* **59**, 75–108.

Roderick, M. and Nagaoka, J. (2005), 'Retention under chicago's high-stakes testing program: Helpful, harmful, or harmless?', *Educational Evaluation and Policy Analysis* **27**(4), 309–340.

Rosenbaum, P. R. (2002), *Observational studies*, Springer.

Rosenbaum, P. R. and Rubin, D. B. (1983), 'The central role of the propensity score in observational studies for causal effects', *Biometrika* **70**(1), 41–55.

Roy, A. D. (1951), 'Some thoughts on the distribution of earnings', *Oxford economic papers* **3**(2), 135–146.

Rubin, D. B. (1974), 'Estimating causal effects of treatments in randomized and nonrandomized studies.', *Journal of educational Psychology* **66**(5), 688.

Ryan, R. M. and Deci, E. L. (2020), 'Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions', *Contemporary educational psychology* **61**, 101860.

Schwerdt, G., West, M. R. and Winters, M. A. (2017), 'The effects of test-based retention on student outcomes over time: Regression discontinuity evidence from florida', *Journal of Public Economics* **152**, 154–169.

Shepard, L. and Smith, M. L. (1989), *A review of research on kindergarten retention*, Falmer Press.

Stearns, E., Moller, S., Blau, J. and Potochnick, S. (2007), 'Staying back and dropping out: The relationship between grade retention and school dropout', *Sociology of Education* **80**(3), 210–240.

Tafreschi, D. and 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.

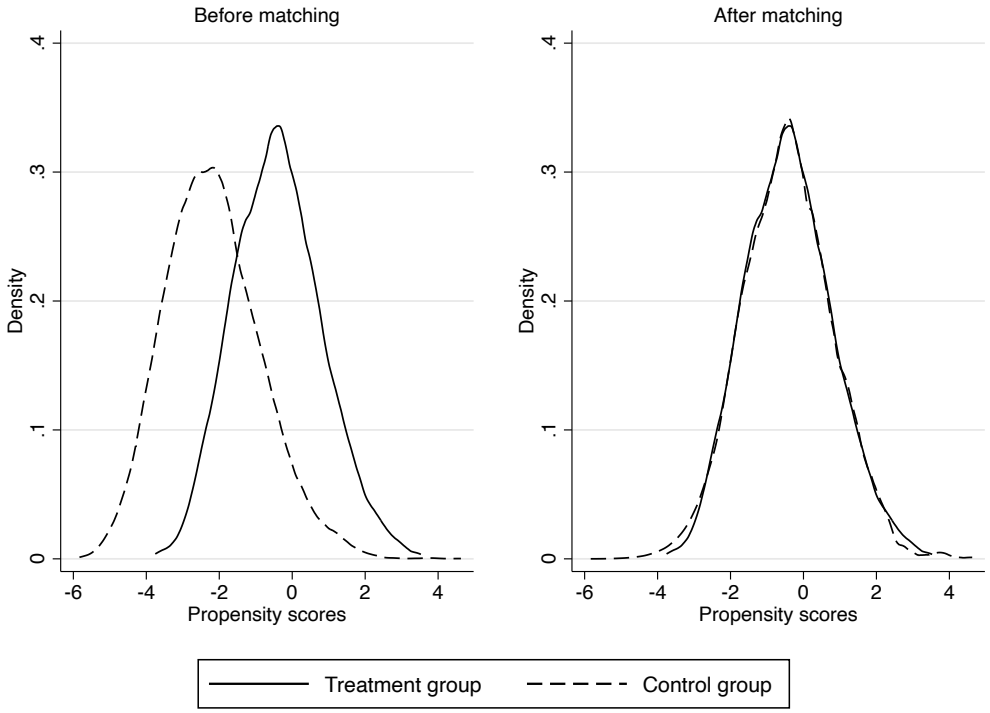
Valbuena, J., Mediavilla, M., Álvaro Choi and Gil, M. (2021), 'Effects Of Grade Retention Policies: A Literature Review Of Empirical Studies Applying Causal Inference', *Journal of Economic Surveys* **35**(2), 408–451.

Vygotsky, L. (1978), *The Role of Play in Development*, Harvard University Press, pp. 199–210.

Winters, M. A. and Greene, J. P. (2012), 'The medium-run effects of florida's test-based promotion policy', *Education Finance and Policy* **7**(3), 305–330.

Xia, N. and Kirby, S. N. (2009), 'Retaining students in grade: A literature review of the effects of retention on students' academic and nonacademic outcomes', *RANDTechnical Report* **678**.

Figure 1: Distribution of Estimated Propensity Scores



Note: Propensity score distributions for treatment and control groups, based on estimations in Table 4.

Table 1: Descriptive statistics on grade retention

	Grade retention	No grade retention	p-value
Sociodemographic characteristics			
Gender			
Boys	55.38%	44.81%	0.000
Girls	44.62%	55.19%	0.000
Socioeconomic status			
High	10.21%	24.84%	0.000
Medium-high	14.03%	18.43%	0.000
Medium-low	30.33%	26.31%	0.000
Low	45.43%	30.42%	0.000
Mother's educational attainment			
Tertiary education	8.76%	23.54%	0.000
General upper secondary education	6.14%	9.04%	0.000
Technical or vocational upper secondary education	5.04%	7.33%	0.000
Vocational education	29.04%	23.23%	0.000
No qualification or NR	51.01%	36.85%	0.000
Father's educational attainment			
Tertiary education	7.25%	21.74%	0.000
General upper secondary education	3.53%	5.02%	0.001
Technical or vocational upper secondary education	4.85%	7.03%	0.000
Vocational education	32.90%	29.02%	0.000
No qualification or NR	51.46%	37.18%	0.000
Family type			
Two parents	78.06%	84.57%	0.000
Single mother	15.16%	10.77%	0.000
Single father	1.62%	1.01%	0.042
Blended family	4.11%	3.12%	0.048
Other situations	1.05%	0.54%	0.035
Parents born in France	29.11%	23.85%	0.000
Number of siblings	1.812	1.596	0.000
School characteristics			
Size of the urban area			
Rural	8.77%	10.07%	0.056
$<$ 5,000 inh	9.94%	10.73%	0.269
[5,000 ; 10,000] inh	10.80%	9.52%	0.078
[10,000 ; 20,000] inh	8.84%	7.63%	0.066
[20,000 ; 50,000] inh	10.21%	10.39%	0.795
[50,000 ; 100,000] inh	8.29%	7.98%	0.628
[100,000 ; 200,000] inh	7.35%	8.12%	0.217
[200,000 ; 2,000,000] inh	20.25%	21.47%	0.203
Paris agglomeration	15.56%	14.10%	0.090
Private school (attended in grade 6)	16.52%	18.46%	0.029
Priority education zone (grade 6)	11.26%	9.36%	0.011
Student's educational characteristics			
Grade repetition in primary school	15.57%	9.34%	0.000
Enrolled at least once in a private primary school	15.06%	17.71%	0.002
Grade 6 test score in mathematics	44.839	56.95	0.000
Grade 6 test score in French	41.399	51.157	0.000
Number of observations	2716	8509	

Note: Reported values are proportions for categorical variables (e.g., gender, socioeconomic status) and means for continuous variables (e.g., test scores). All p-values are derived from t-tests for equality of means between students who experienced grade retention in lower secondary school and those who did not.

Source: 1995 Panel – EVA.

Table 2: Outcomes by Grade Retention Status

	Grade retention	No grade retention	p-value
Secondary school completion	24.76%	74.55%	0.000
Obtaining a higher education qualification	21.87%	62.41%	0.000
Being employed	85.12%	87.18%	0.163
Monthly wage	1490.33	1736.03	0.000

Note: Reported values are proportions for categorical variables (e.g., secondary school completion) and means for continuous variables (e.g., monthly wage). All p-values are derived from t-tests for equality of means between students who experienced grade retention in lower secondary school and those who did not.

Source: 1995 Panel – EVA.

Table 3: Comparison of IV and PSM methods

	IV		PSM	
	Math	Reading	Math	Reading
Unadjusted for initial scores	4.747***	3.610***	0.350***	0.281***
	(1.070)	(0.887)	(0.020)	(0.020)
Adjusted for initial test scores	7.844	3.725	0.076***	0.064***
	(5.973)	(3.219)	(0.019)	(0.021)

Note: Comparison of estimated coefficients for the effect of grade retention on mathematics and reading scores in grade 9, using instrumental variables (IV) and propensity score matching (PSM) methods. Three specifications are shown for each method: (1) unadjusted for initial (grade 6) test scores (VA) and (2) adjusted for the initial scores in mathematics and reading. The logit estimates of the propensity scores are provided in Table A.1. Average treatment effects on the treated (ATT) are calculated using the Epanechnikov kernel function with a bandwidth of 0.06. Standard errors are in parentheses. $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

Source: Panel 2007.

Table 4: Logistic Regression on Likelihood of Grade Retention

	Coefficient
Sociodemographic characteristics	
Girls	-0.588***
	(0.065)
Socioeconomic status (ref. High)	
Medium-high	0.174
	(0.118)
Medium-low	0.201*
	(0.115)
Low	0.228*
	(0.121)
Mother's educational attainment (ref. Tertiary education)	
General upper secondary education	0.194
	(0.143)
Technical or vocational upper secondary education	0.200
	(0.155)
Vocational education	0.482***
	(0.116)
No qualification	0.319***
	(0.116)
Father's educational attainment (ref. Tertiary education)	
General upper secondary education	0.502***
	(0.184)
Technical or vocational upper secondary education	0.311*
	(0.162)
Vocational education	0.442***
	(0.134)
No qualification	0.403***
	(0.134)
Family type (ref. Two parents)	
Single mother	0.545***
	(0.103)
Single father	0.353
	(0.265)
Blended family	0.278
	(0.180)
Other situations	0.640
	(0.409)
Parents born outside France	-0.344***
	(0.085)
Number of siblings	0.029
	(0.028)
School characteristics	
Size of the urban area (ref. Paris agglomeration)	
Rural	-0.329**
	(0.131)

Continued on next page

Table 4 (continued): Logistic Regression on Likelihood of Grade Retention

	Coefficient
< 5,000 inhabitants	-0.478***
	(0.135)
[5,000 ; 10,000] inhabitants	-0.085
	(0.127)
[10,000 ; 20,000] inhabitants	0.002
	(0.137)
[20,000 ; 50,000] inhabitants	0.016
	(0.129)
[50,000 ; 100,000] inhabitants	-0.162
	(0.135)
[100,000 ; 200,000] inhabitants	-0.129
	(0.147)
[200,000 ; 2,000,000] inhabitants	-0.109
	(0.108)
Private school in grade 6	0.048
	(0.104)
Priority education zone in grade 6	-0.646***
	(0.123)
Student's educational characteristics	
Grade repetition in primary school	-0.724***
	(0.112)
Enrolled at least once in a private primary school	-0.026
	(0.107)
Grade 6 test score in mathematics	-0.075***
	(0.004)
Grade 6 test score in French	-0.067***
	(0.005)
Constant	5.985***
	(0.287)
Observations	8,673

Note: Logistic regression estimates for the likelihood of grade retention in lower secondary school. Coefficients represent the log odds of grade retention associated with each variable, with reference categories specified for categorical variables. Standard errors are in parentheses. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.
Source: 1995 Panel – EVA.

Table 5: Matching Quality

	Before matching	After matching
Number of variables with significant differences in means		
At the 1% level	18	0
At the 5% level	3	1
At the 10% level	5	0
Number of variables with absolute standardized bias:		
<1%	2	13
1% - <3%	2	16
3% - <5%	6	2
5% - <10%	9	1
10% - <15%	5	0
>15%	8	0
Mean absolute standardized bias	16.67	1.40
Median absolute standardized bias	8.09	1.06
Pseudo-R^2 for propensity score estimation	0.2551	0.0021
p-value of joint significance test	0.0000	0.9998
Total number of variables	32	32
Off-support treated students		0

Note: Summary of matching quality, showing the number of variables with significant differences in means and levels of absolute standardized bias before and after matching. Mean and median absolute standardized biases are reported as percentages. Pseudo-R^2 and the p-value of the joint significance test characterize the overall balance of covariates.
Source: 1995 Panel – EVA.

Table 6: Main Results

	Secondary school completion	Obtaining a higher education qualification	Being employed	Monthly wage
Kernel	-0.211***	-0.238***	-0.006	-0.059***
(bw=0.06)	(0.015)	(0.022)	(0.015)	(0.019)
Number of observations	8,673	5,137	5,211	4,065

Note: Estimated effects of grade retention on educational and employment outcomes using kernel matching with a bandwidth of 0.06. Coefficients represent the average treatment effects on the treated (ATT) for each outcome. The logit estimates used to calculate propensity scores are provided in Table 4. Standard errors are in parentheses. Significance levels: $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.
Source: 1995 Panel – EVA.

Table 7: Heterogeneity by Gender

	Secondary school completion	Obtaining a higher education qualification	Being employed	Monthly wage
Boys	-0.185***	-0.191***	-0.007	-0.074***
	(0.022)	(0.030)	(0.020)	(0.024)
Girls	-0.235***	-0.300***	-0.009	-0.028
	(0.023)	(0.033)	(0.027)	(0.031)

Note: Estimated effects of grade retention on educational and employment outcomes, disaggregated by gender. Coefficients represent the average treatment effects on the treated (ATT) for each outcome. The variables used to calculate the propensity scores are listed in Table 4. Standard errors are in parentheses. Significance levels: $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.
Source: 1995 Panel – EVA.

Table 8: Heterogeneity by Socioeconomic Status (SES)

	Secondary school completion	Obtaining a higher education qualification	Being employed	Monthly wage
High SES	-0.207***	-0.192**	-0.019	-0.117
	(0.053)	(0.079)	(0.057)	(0.073)
Medium-high SES	-0.229***	-0.326***	-0.036	-0.015
	(0.042)	(0.060)	(0.039)	(0.052)
Medium-low SES	-0.221***	-0.260***	0.020	-0.066
	(0.028)	(0.043)	(0.030)	(0.045)
Low SES	-0.210***	-0.203***	-0.005	-0.052*
	(0.024)	(0.032)	(0.026)	(0.029)

Note: Estimated effects of grade retention on educational and employment outcomes, disaggregated by SES. Coefficients represent the average treatment effects on the treated (ATT) for each outcome. The variables used to calculate the propensity scores are listed in Table 4. Standard errors are in parentheses. Significance levels: $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.
Source: 1995 Panel – EVA..

Table 9: Sensitivity to Matching Approach for Selected Outcomes

	Secondary school completion	Obtaining a higher education qualification	Being employed	Monthly wage
Nearest neighbor (1:1) without replacement	-0.243***	-0.253***	-0.007	-0.060***
	(0.016)	(0.024)	(0.017)	(0.020)
Nearest neighbor (1:1) with replacement	-0.203***	-0.198***	-0.018	-0.024
	(0.022)	(0.033)	(0.024)	(0.029)
Nearest neighbor (1:3) with replacement	-0.199***	-0.236***	-0.017	-0.057**
	(0.019)	(0.028)	(0.020)	(0.025)
Nearest neighbor (1:5) with replacement	-0.198***	-0.233***	-0.013	-0.060***
	(0.019)	(0.027)	(0.019)	(0.023)
Kernel (bw=0.02)	-0.206***	-0.228***	-0.003	-0.054***
	(0.016)	(0.022)	(0.016)	(0.020)
Kernel (bw=0.06)	-0.211***	-0.238***	-0.006	-0.059***
	(0.016)	(0.021)	(0.015)	(0.019)
Kernel (bw=0.2)	-0.236***	-0.271***	-0.008	-0.072***
	(0.015)	(0.021)	(0.015)	(0.018)
Radius (caliper=0.1)	-0.220***	-0.252***	-0.007	-0.064***
	(0.015)	(0.020)	(0.016)	(0.018)
IPW	-0.151***	-0.194***	-0.012	-0.049*
	(0.035)	(0.029)	(0.018)	(0.025)

Source: Panel 1995 – EVA. *Note:* The table compares the estimated effects of grade retention on four outcomes—secondary school completion (obtaining a *baccalauréat*), obtaining a higher education qualification, employment status, and monthly wage—across different matching techniques. Matching approaches include nearest neighbor (with varying ratios and with or without replacement), kernel (with bandwidths of 0.02, 0.06, and 0.2), radius (with a caliper of 0.1), and inverse probability weighting (IPW). Coefficients represent the average treatment effects on the treated (ATT) for each outcome. The logit estimates used to calculate propensity scores are provided in Table 4. Standard errors are reported in parentheses. Significance levels: $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Table 10: Sensitivity Analysis – Bounding Approach

Γ	Secondary school completion	Obtaining a higher education qualification	Being employed	Monthly wage
1.00	0.000	0.000	1.000	0.000
1.50	0.000	0.000	1.000	0.389
2.00	0.000	0.000	1.000	0.998
2.50	0.000	0.000	1.000	1.000
3.00	0.083	0.106	1.000	1.000
Critical values				
1%	2.80-2.85	2.70-2.75	0.00-0.00	1.20-1.25
5%	2.95-3.00	2.85-2.90	0.00-0.00	1.30-1.35
10%	3.00-3.05	2.95-3.00	0.00-0.00	1.35-1.40

Note: Results obtained using Rosenbaum bounds Rosenbaum (2002) to evaluate sensitivity to unobserved heterogeneity. Critical Γ values indicate thresholds at which estimated effects lose significance, with higher values suggesting greater robustness against hidden bias.
Source: 1995 Panel – EVA.

Table 11: Sensitivity Analysis – Simulation Approach

	Secondary school completion			Obtaining a higher education qualification			Being employed			Monthly wage		
	Outcome effect	Selection effect	ATT	Outcome effect	Selection effect	ATT	Outcome effect	Selection effect	ATT	Outcome effect	Selection effect	ATT
Girls	1.340	0.699	-0.213	1.324	0.629	-0.243	0.775	0.610	-0.009	0.538	0.683	-0.065
High SES	3.636	0.388	-0.196	2.534	0.387	-0.226	1.146	0.359	-0.004	1.457	0.331	-0.054
Mother's educational attainment – Tertiary education	4.041	0.353	-0.196	2.532	0.380	-0.227	1.036	0.329	-0.006	1.309	0.293	-0.056
Father's educational attainment – Tertiary education	3.929	0.312	-0.197	2.680	0.340	-0.226	0.987	0.291	-0.007	1.484	0.264	-0.054
Family types – Two parents	2.058	0.744	-0.213	1.485	0.739	-0.247	1.795	0.693	-0.004	0.727	0.673	-0.061
Parents born outside France	0.644	1.248	-0.215	0.883	1.324	-0.248	0.581	1.340	-0.004	1.751	1.320	-0.062
Size of the urban area - Paris agglomeration	1.100	1.092	-0.218	1.730	1.201	-0.250	0.893	1.142	-0.007	2.062	1.084	-0.060
Private school in grade 6	1.302	0.925	-0.217	1.504	1.109	-0.250	1.275	1.040	-0.007	1.074	0.901	-0.059
Priority education zone in grade 6	0.564	1.147	-0.217	0.522	1.112	-0.248	0.790	1.230	-0.007	1.250	1.243	-0.059
Grade repetition in primary school	0.139	1.311	-0.209	0.198	1.691	-0.239	0.624	2.178	-0.004	1.269	1.843	-0.061
Enrolled at least once in a private primary school	1.234	0.852	-0.217	1.520	0.930	-0.248	1.088	0.878	-0.007	1.080	0.812	-0.059

Note: Outcome effects represent the estimated impact of the variable on the outcome; selection effects are the expected effects based on selection into the treatment; ATT is the average treatment effect on the treated after adjustment for observed characteristics. Values below (above) 1 for outcome and selection effects indicate a negative (positive) influence of the confounder. In the absence of unobserved heterogeneity, both effects are set to zero. *Source:* 1995 Panel – EVA.

Table A.1: Logistic Regression on Likelihood of Grade Retention

	Coefficient
Sociodemographic characteristics	
Girls	-0.509***
	(0.053)
Socioeconomic status (ref. High SES)	
Medium-high	0.143
	(0.098)
Medium-low	0.057
	(0.095)
Low	0.163*
	(0.097)
Mother's educational attainment (ref. Tertiary education)	
General upper secondary education	0.357***
	(0.121)
Technical or vocational upper secondary education	0.441***
	(0.098)
Vocational education	0.614***
	(0.083)
No qualification	0.509***
	(0.090)
Father's educational attainment (ref. Tertiary education)	
General upper secondary education	0.317**
	(0.158)
Technical or vocational upper secondary education	0.255**
	(0.121)
Vocational education	0.333***
	(0.099)
No qualification	0.362***
	(0.102)
Family type (ref. Two parents)	
Single mother	0.426***
	(0.073)
Single father	0.353
	(0.265)
Blended family	0.278
	(0.180)
Other situations	0.640
	(0.409)
Parents born outside France	-0.344***
	(0.085)
Number of siblings	0.031*
	(0.018)
School characteristics	
Size of the urban area (ref. Paris agglomeration)	
Rural	-0.273**
	(0.112)

Continued on next page

Table A.1 (continued): Logistic Regression on Likelihood of Grade Retention

	Coefficient
< 5,000 inhabitants	-0.069
	(0.104)
[5,000 ; 10,000] inhabitants	-0.089
	(0.105)
[10,000 ; 20,000] inhabitants	-0.176
	(0.111)
[20,000 ; 50,000] inhabitants	-0.111
	(0.110)
[50,000 ; 100,000] inhabitants	-0.020
	(0.107)
[100,000 ; 200,000] inhabitants	0.077
	(0.111)
[200,000 ; 2,000,000] inhabitants	0.065
	(0.085)
Private school in grade 6	0.565***
	(0.081)
Priority education zone in grade 6	-0.523***
	(0.074)
Grade repetition in primary school	-1.705***
	(0.097)
Private school in grade 6	0.036
	(0.085)
Grade 6 test score in mathematics	-0.043***
	(0.002)
Grade 6 test score in French	-0.027***
	(0.002)
Constant	1.308***
	(0.173)
Observations	21,821

Note: Logistic regression estimates for the likelihood of grade retention in lower secondary school. Coefficients represent the log odds of grade retention associated with each variable, with reference categories specified for categorical variables. Standard errors are in parentheses. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.
Source: 2007 Panel.

Table A.2: Main Results with Imputation

	Secondary school completion	Obtaining a higher education qualification	Being employed	Monthly wage
Kernel (bw=0.06)	-0.203***	-0.210***	-0.005	-0.064***
	(0.014)	(0.020)	(0.015)	(0.017)
Number of observations	10,024	5739	5825	4508

Note: Estimated effects of grade retention on educational and employment outcomes using kernel matching with a bandwidth of 0.06. Coefficients represent the average treatment effects on the treated (ATT) for each outcome. The logit estimates used to calculate propensity scores are provided in Table 4. Standard errors are in parentheses. Significance levels: $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.
Source: Panel 1995 – EVA.