Journal for Labour Market Research

ORIGINAL ARTICLE Open Access

The COVID-19 pandemic: a threat to higher education? Evidence from a large university in Northern Italy

Marina Bonaccolto-Töpfer[2*] and Carolina Castagnetti[1]

Abstract

Transition to online teaching during the first wave of the COVID-19 pandemic has led to various concerns about educational quality. The study investigates the consequences of this transition on student performance for a prominent university in Northern Italy. Comprehensive administrative data allows us to monitor students' performance and to have detailed socioeconomic information about them. Using a difference-in-differences design with lecturer fixed effects, we exploit the unique timing of the lockdown that fell together with the start of the summer term in 2020. We find no marked effects on students' average grades in higher education in the immediate aftermath of the lockdown.

Keywords Online teaching, COVID-19 pandemic, Difference-in-differences, Higher education

JEL Classification G18, I18, I23, I24

1 Introduction

At the beginning of the year 2020, the whole world was unexpectedly and harshly hit by the COVID-19 pandemic, what prompted governments to implement measures aimed at curbing the spread of the virus. Italy, and especially the Northern Lombardy region, became a global COVID-19 hotspot in February 2020 as the number of reported cases grew rapidly. As a consequence, urgent measures to contain and manage the COVID-19 emergency were implemented. Lombardy as well as several other regions in Northern Italy began to close schools, kindergartens, and universities at the end of February,[1] leading to an accelerated transition to online teaching.[2] Educational institutions responded to the closures and transition to online teaching in different ways. Primary and secondary schools generally experienced delays in implementing distance learning resulting in reduced teaching hours. Conversely, universities quickly adapted to the change and implemented online teaching without significant disruptions in the hours of instruction or exam evaluations. Consequently, the measures implemented may have had distinct effects on schools and universities.

[1] To be precise on February 23, 2020.

[2] In the rest of Italy, a decree was issued by the President of the Council of Ministers on March 8, 2020 that mandated the suspension of classes in all Italian educational institutions, with the provision for distance learning activities such as online teaching.

*Correspondence:
Marina Bonaccolto-Töpfer
marina.toepfer@unikore.it
[1] Department of Economics and Management, University of Pavia, Pavia, Italy
[2] Department of Economics and Law, Kore University of Enna, Enna, Italy

The treatment, i.e. the transition to online teaching due to the pandemic, is identified during the summer term when only the treated cohort experiences the change in the mode of instruction shifting from in-person to online teaching.

We use administrative data from the University of Pavia for the academic years from 2016/2017 to 2019/2020. The University of Pavia is one of the largest universities in Lombardy dating back to its establishment in 1361.[3] The university has up to 25,000 active students (MIUR 2023)[4] from all over Italy as well as from many different countries worldwide.

The case of the University of Pavia is interesting for several reasons. First, it allows us to base our analysis on rich administrative data. Our administrative dataset includes detailed information on students' academic career (such as grades, credits earned, department of study) and socio-demographic characteristics (gender, age, economic background). Second, the case of the University of Pavia is particularly well-suited for identifying the effect of the transition to online teaching on student performance as, in Lombardy, the decision to suspend in-person classes coincided with the beginning of the summer term. As a consequence, students were exposed to online teaching during the complete summer term. In contrast, in other Italian regions, particularly those in Central and Southern Italy, there was an initial period of in-person learning during the summer term of the academic year 2019–2020, followed by the transition to online instruction.

Most of the papers that look at the effects of the pandemic on education focus on primary and secondary education with rather clear-cut results suggesting mainly negative effects. For example, Agostinelli et al. (2022) found that school closures had a large and presumably persistent effect on educational outcomes of high school students that is highly unequally distributed. High school students from poor neighborhoods suffered from a learning loss, while those from rich neighborhoods remained unaffected. Similarly, Chetty et al. (2020) found that student progress of an online program in Maths decreased in poorer areas.

In addition to the impact on students, it is crucial to acknowledge that the quality of teaching may have been influenced as well. The transition to online teaching has brought about various challenges, including technical difficulties and the adaptation struggles of educators to this new educational landscape. Consequently, these factors may have contributed to a deterioration in the quality of (online) teaching. Additionally, personal issues faced by lecturers could have further hampered their teaching performance, resulting in a decline in student learning outcomes and reduced educational quality. However, the shift to online teaching has also fostered the use of innovative teaching formats. For instance, the massive utilization of online whiteboards as well as of the flipped classroom model during the pandemic have emerged as promising approaches. These novel methodologies have the potential to positively impact educational quality, fostering enhanced engagement and interaction among students and teachers.

The existing literature on higher education suggests that the effects of the pandemic on academic outcomes are theoretically ambiguous and that the empirical evidence is mixed. Aucejo et al. (2020) conducted a study based on survey data from a US university, which suggested that the economic factors, such as a lack of financial resources to complete studies, as well as the health-related effects, such as the fear of becoming sick, resulting from the pandemic may have contributed to the exacerbation of inequality within higher education. Rodríguez-Planas (2022a) carried out research using an online student survey from a US college, which revealed several negative effects on higher education. For instance, graduation plans and freshmen's retention rates were adversely affected by the pandemic. In the analysis, difficulties in attending online classes and an augmented probability of dropping out of the study program became evident. Further, these aspects were found to be significantly affected by the socioeconomic level of students.

Jaeger et al. (2021) analyzed data from the Global COVID-19 Student Survey, which involved 28 universities worldwide, including one Italian university. Their research aimed to examine students' reactions to various aspects of the pandemic, particularly their current and future academic situations. The results indicated negative effects on education, challenges related to computer and library access as well as disruptions in the educational process.

While the studies mentioned above use survey data to investigate the impact of the pandemic on academic expectations, other approaches use administrative data to assess how the transition to online teaching affected students' performance in terms of grades. For instance, the study of Rodríguez-Planas (2022b) is based on administrative records from a US college and reveals that, despite students reporting challenges with online learning, student performance was comparatively higher, and there was no significant impact on credits earned during the spring of 2020. Bulman and Fairlie (2022) analyzing

[3] It is recognized as one of the oldest academic institutions around the globe and is ranked among the top 561–570 universities worldwide in the QS World University Rankings 2022.

[4] For comparison consider, for instance, that the majority of US colleges and universities has student numbers within the range of 15,000 and 25,000 students (StudyPoint 2021).

administrative college data from California, found that GPAs were higher in spring 2020 compared to spring 2019. Thus, there may be factors having led to better performance of students during the first lockdown. For example, a different assessment process with potentially easier exams and more lenient grading may have led to higher (or unchanged) grades. Further, the absence or reduction of social events may have allowed students to focus more on learning and thus to improve their grading. However, the positive effect found by Bulman and Fairlie (2022) diminished in subsequent semesters.

We contribute to this literature as follows. Our paper provides new insights for Italy, a country where the higher education system before the pandemic was mostly traditional. Following the COVID-19 outbreak, Italian universities swiftly adapted to online teaching methods. Moreover, focusing on a University in the Northern Lombardy region, where the transition to online learning coincided with the start of the summer term 2020, allows us to identify the treatment effect in the DID set-up outlined above. Additionally, we employ a rich administrative dataset. The use of administrative data offers clear advantages, including increased reliability and comprehensiveness, as well as the absence of self-reporting bias compared to survey data. Furthermore, administrative data eliminates sample limitations allowing to analyze the entire population. Our dataset covers all students who have taken at least one exam at the University of Pavia during the considered time period.

Compared to the US studies mentioned earlier, our approach differs in the following aspect. During the pandemic, certain US colleges and universities introduced a flexible grading policy.[5] This policy temporarily modified the assessment of exams and the composition of the GPA and may be responsible for a positive impact on student performance. In Italy, and in particular at the University of Pavia, there have been no similar official changes in the grading policy.

Our findings reveal only small effects of the transition to online teaching on student performance (at most + 1%). We conduct various tests, including separate estimations based on gender, performance levels (top- and bottom-performing students), socioeconomic backgrounds (students from poorer or richer families), as well as restricting the sample to freshmen and exams of mandatory courses. Yet, as we cannot rule out that the pandemic may have affected lecturers' evaluation of student performance being more lenient and comprehensive, we also examine the presence of grade inflation. To do so, we focus on the performance of students who, after attending in-person courses in the winter term have chosen to postpone their exams. Prior to spring 2020, students were able to take the exams as usual (either written or oral, both in person), while in spring 2020 they were unexpectedly required to switch to remote examination. Looking at changes in the performance before and after spring 2020 confirms the results that, during the pandemic university students' performance did not change markedly.

In our study, we focus on the first wave of the pandemic as during this period, teaching in the summer term for the treatment group was completely and exclusively online. This resulted in an immediate and non-anticipated transition to online teaching. In subsequent waves, there were alternating periods between in-person and online learning. Often, these hybrid periods allowed students to attend classes flexibly either in person or via live-streaming from home. The results shown in this paper may be significant beyond the University of Pavia as they analyze the impact within a university that, due to its research characteristics, teaching offers, and organization can be considered representative of an average well-ranked university.

2 Experimental setting and identification strategy

Lombardy, being the hardest-hit region in Italy during the initial wave of the pandemic in spring 2020, was among the first regions to implement closures of schools and universities. Notably, the timing of the university shutdown in Lombardy was particularly significant for our empirical approach as it coincided with the start of the summer term in 2020. School and university closures were enacted in Lombardy, as well as in other regions such as Piedmont, Veneto, and Emilia-Romagna, on February 25, 2020, and remained in force until September 2020. At the University of Pavia, situated in Lombardy, the beginning of lectures was originally scheduled for February 24, 2020, and subsequently started (entirely) online only one week later.

Figure 1 shows the timing of events and presents how our DID approach works. We begin our analysis by focusing on the academic years 2018/2019 and 2019/2020.[6] Academic years are typically divided into two semesters or terms: the winter term and the summer term. The winter term typically spans from early October to late February, while the summer term begins at the end of February and concludes in late September. The treatment group consists of students that were affected by the transition

[5] The flexible grading policy allows students a 20-day window, starting from when the instructor submits the final grade, to decide whether they wish to exclude lower-than-expected grades (B+ to D for undergraduate students, B+ to C- for graduate students) from affecting their GPA while still earning credits for those courses, or to convert an F grade into a no-credit grade.

[6] In further analyses, we will consider a longer time horizon ranging from academic year 2015/2016 or 2016/2017 to 2019/2020.

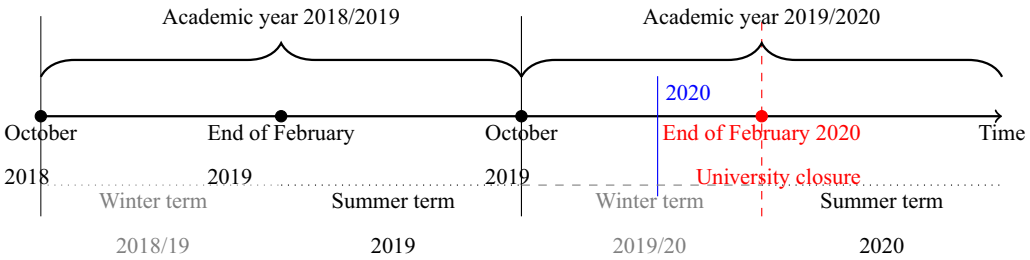


Fig. 1 Timeline of events. The control group consists of students from academic year 2018/2019, while the treatment group includes students from academic year 2019/2020. The treatment occurs during the summer term. The difference-in-differences effect is calculated as: (average grades in summer term—average grades in winter term)$^{\text{academic year 2019/2020}}$ - (average grades in summer term - average grades in winter term)$^{\text{academic year 2018/2019}}$. Average grades are obtained for each term at the the student level

to online teaching during the summer term of 2020. To establish a comparison, we define the control group as students from the academic year 2018/2019, while the treatment group consists of students from the academic year 2019/2020. Therefore the DID effect is derived from a comparison between the average grade differences during the summer and winter semesters for the academic years 2019/2020 and 2018/2019, respectively.

Analyzing students' grades, we want to address several key questions. First, we seek to determine whether the COVID-19 pandemic has had an impact on students' academic performance. As stated above, the literature focusing on the first waves of the pandemic in higher education is ambiguous in this regard and found negative (Jaeger et al. 2021; Rodríguez-Planas 2022a) as well as positive (Bulman and Fairlie 2022) and/or insignificant effects (Rodríguez-Planas 2022b). Second, we look at several socio-economic dimensions of students. For instance, we explore whether family wealth influences student performance during this challenging period. We also examine whether there are differential effects between top-performing and bottom-performing students.

3 University career data

The data set used in this study includes the entire population of students enrolled at the University of Pavia who have taken at least one exam during either the academic year 2018/2019 or the academic year 2019/2020. For each student, we have access to specific grades for each exam they have taken, as well as information on gender, date and place of birth, municipality of residence, details about the degree course in which the student is enrolled, and the year of enrollment. Additionally, we have data indicating whether a student is in his or her freshman year and whether a particular course is mandatory or not. Another important variable we observe is the student's ISEE (*Indicatore della Situazione Economica Equivalente*), which is a standardized indicator of economic situation. The ISEE declaration is an official document issued by a public institution, typically the municipality of residence. The ISEE takes into account the family's annual income, non-labor income (such as assets), and factors related to the family composition, such as single parent and number and age of siblings.

At the University of Pavia, tuition fees are determined based on the ISEE declaration submitted by students. The university defines 60 different income brackets based on the ISEE to determine the corresponding tuition fees for students. Tuition fees vary from €0 to €4845 per year. Students who do not submit an ISEE declaration are automatically assigned to the highest income bracket. Given that using the 60 income brackets defined by the university may lead to limited observations in each bracket, this study divides the ISEE into four broader categories: bottom (< 25th percentile), medium-bottom (≥ 25th and < 50th percentile), medium-top (≥ 50th and < 75th percentile), top (≥ 75th percentile). We focus the analysis on bachelor and master courses, five-year degree in Law and six-year degree in Medicine.

To identify top and bottom students, we consider their average grades before the summer term 2020. Top (bottom) students are defined as those whose average grades are at or above the 75th (below the 25th) percentile of the grade distribution. We also examine the impact of family background by using the ISEE. Students with a low ISEE have an ISEE below the 25th percentile of the income brackets defined by the university, indicating a relatively lower family income. Conversely, students with a high ISEE have an ISEE equal to or above the 75th percentile of the income brackets, indicating a relatively higher family income.

Student performance is computed as a weighted average of grades, considering the credit units (CFU)[7] assigned to each course. The credits for a course are determined by the number of instructional hours, and achieving a predefined number of credits is a requirement for graduation. We

[7] The number of credits corresponds to the CFU (*Crediti Formativi Universitari*) associated to each exam.

compute for each student the average grade per term and year weighted by the number of credits assigned to each exam. Further, we only consider exams taken in the same semester as the corresponding course. Courses are organized into semesters, and exams are scheduled at the end of each semester, with the option for additional extraordinary exam sessions available at various times throughout the year. Students have the flexibility to decide when to take exams, regardless of the semester in which the course was conducted. To ensure comparability between the control and treatment groups, we exclusively consider exams taken in the same semester as the course was taught. This approach aims to assess the performance of students who promptly undertook exams after attending and actively participating in the classes. In fact, typically students tend to postpone exams of courses that were not attended or adequately followed. Note that the number of exams is stable throughout the academic years, i.e. not more students postpone their exam due to online teaching.[8]

The analysis focuses on regular students, defined as those enrolled and accepted for enrollment at the University and studying within the expected duration of their degree program. Notably, in the Italian university system, there is no time limit for completing a degree. By excluding non-regular students from the sample, we eliminate those who likely do not attend classes and are only marginally involved in academic activities. After applying these restrictions, the final sample consists of 28,052 students. The study uses a cohort-based approach to compare student performance between different academic years. A cohort is defined as a class of students educated during the same period of time. By examining cohorts from consecutive academic years, the study can compare the performance of students in one cohort with the performance of students in the previous cohort.

In Italy, the grading system ranges from 18 to 31, with 18 being the minimum passing grade and 31 representing the highest grade achievable, often referred to as *30 e lode* or *cum laude*. In Italian public universities, there is no limit to the number of times a student can retake or refuse an exam grade.[9] Table 1 reports differences in means (together with standard deviations) of students' university career data by treatment status. As can be seen in Table 1, the treatment (academic year 2019/2020) and control groups (academic year 2018/2019) are well balanced across students' individual characteristics.[10]

Table 1 Descriptive statistics student performance by treatment status, selected controls

	(1)	(2)	(3)	(4)	(5)
Academic year	**2019/2020**		**2018/2019**		
Variable	**Mean**	**Std.dev.**	**Mean**	**Std. dev.**	**Difference**
Average grade	27.08	2.66	27.12	2.58	– 0.040
Female (dummy)	0.611	0.48	0.602	0.49	0.009
Age (in years)	22.29	3.27	22.23	3.14	0.060*
Freshmen (dummy)	0.243	0.43	0.239	0.43	0.004
ISEE bottom	0.221	0.42	0.221	0.42	0.000
ISEE medium bottom	0.260	0.44	0.262	0.44	– 0.002
ISEE medium top	0.220	0.41	0.216	0.41	0.004
ISEE top	0.300	0.46	0.302	0.46	– 0.002
Number of exams per term	3.234	1.07	3.232	1.05	0.002
Bachelor (dummy)	0.465	0.49	0.465	0.49	0.000
Master (dummy)	0.285	0.45	0.281	0.45	0.004
Observations	14,441		13,611		28,052

The table shows pre- and post-treatment characteristics of the treatment group constituted of students in academic year 2019/2020 as well as pre- and post-treatment characteristics of the control group constituted of students in academic year 2018/2019

* denotes significance at the 10%-level. Standard errors clustered at the individual level are used

4 Estimation approach

Our empirical framework employs a DID approach using a basic two-group, two-period model. Under the common trend assumption, we presume that relevant unmeasured variables are either time-invariant attributes of each group or time-varying factors that affect both groups uniformly. This assumption is necessary for the validity of our DID model: it posits that in the absence of the treatment, the outcome trajectories for both groups would have remained parallel over time. In applying this assumption, we infer that the performance of the control group (students before the academic year 2019/2020) serves as a valid counterfactual for the observed trend in the treated group (students during the academic year 2019/2020). This is not an empirical verification but rather a theoretical stipulation based on the structural consistency and stability observed in the pre-intervention periods across both groups. This assumption allows us to attribute differences in outcomes to the intervention rather than to other external changes. Further, we assume that the trajectory of average grades obtained by students in courses during the winter and summer terms follows a parallel trend across the various academic years.

Fig. 2 shows the mean grades for the summer and winter term in different academic years. This graphical representation supports the assumption of a

[8] We have verified that the percentage of semester-delayed exams relative to the total number of exams has not changed in the semesters considered for the estimation.

[9] Exceptions to this rule exist in some private universities, where students cannot refuse grades.

[10] Table A.1 in the Appendix provides descriptive statistics by academic year and term. The table documents that all characteristics were balanced across treatment and control group as well as across term.

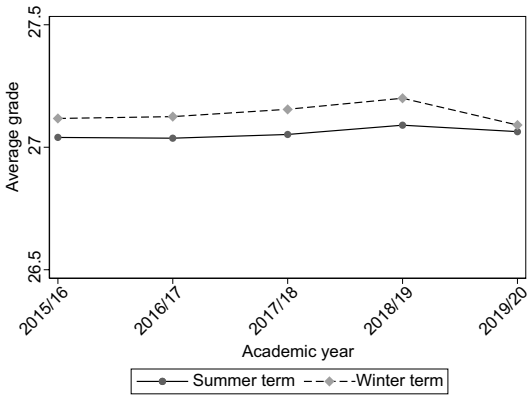


Fig. 2 Average student performance in each semester in the academic years from 2015/2016 to 2019/2020. 79,682 observations over academic years 2015/2016 to 2019/2020

parallel trend across different academic years in the trajectory of average grades during the first and second semester.

In order to evaluate the impact of the transition to online teaching on students' performance, we compare average grades in the academic years 2019/2020 and 2018/2019 over terms (winter and summer). In our design, the control group consists of the cohort of students from academic year 2018/2019, while the treatment group refers to the cohort of students from academic year 2019/2020. We estimate the following baseline equation:

$$v_{ist} = \alpha + \gamma coh_{it} + \delta term_s + \beta coh_{it} * term_s + \varepsilon_{ist} \tag{1}$$

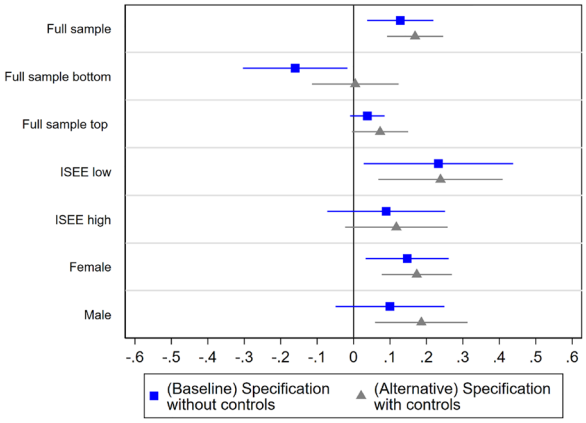
where v_{ist} is the average grade obtained by student i in the exams of year t and term s, coh_{it} is the cohort student i is part of in academic year t with $coh = 1(t = 2019/2020)$ and $term_s$ identifies the summer term, ε_{ist} is the corresponding error term. The interaction term $coh_{it} * term_s$ defines the treatment. As anticipated, we run the above regression for several subsamples of the data such as men, women, top- and bottom-performing students, students from wealthier families, from poorer families as well as for Bachelor and Master students.

In an alternative specification, we augment our baseline model specification by incorporating a vector of control variables, $\mathbf{x}_{ist}$, which encompasses various individual characteristics. These controls include age, gender, student status (freshmen or not), ISEE indicator, number of exams taken, department of enrollment, and study program (e.g., Bachelor or Master program). In this alternative specification, we modify our dependent variable by considering the average grade after accounting for lecturer fixed effects. To achieve this correction or depuration, we employ a two-step process.[11] First, we estimate lecturer fixed effects in the following regression: $v_{clst} = \alpha_0 + c_c * \iota_c + \theta_l + \phi_s + \omega_t + \tilde{\varepsilon}_{lcst}$, where θ_l, ϕ_s and ω_t represent lecturer, term and academic year fixed effects, respectively. c are course or exam dummies, α_0 is an intercept and $\tilde{\varepsilon}$ is the corresponding error term. Second, the corrected grades net of the lecturer time-constant heterogeneity or the lecturer fixed effects are defined as: $\tilde{v}_{iclst} = v_{iclst} - \hat{\theta}_l$. Observe that student i writes exam c of lecturer l in term s and academic year t. The corrected grades are then averaged over exams c of student i in term s and academic year t: $\bar{\tilde{v}}_{ist} = \tilde{v}_{ist}$. The latter is the dependent variable in the alternative specification.

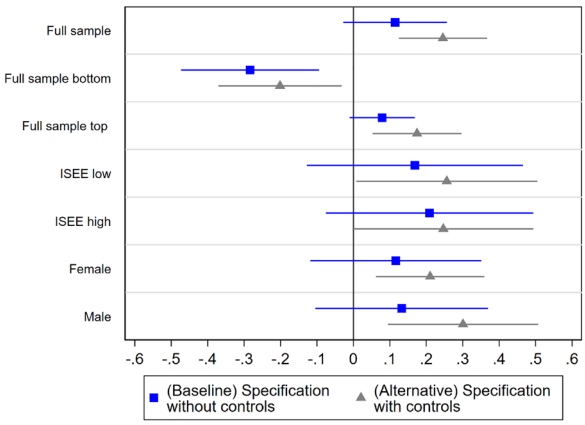
To account for potential changes in course scheduling and instructors over time, we limit our sample to courses that remained constant across these dimensions. By excluding courses that were relocated to different terms or experienced changes in the assigned lecturers, we create a more stable and comparable subset of data for our analysis. As the potential self-selection bias caused by students choosing specific exams during the pandemic, could distort our results, we conduct additional analyses by focusing on two subgroups: (a) freshmen students and (b) mandatory courses (see Appendix C). The additional test (a) represents a group of freshmen that are more likely to stick to the original study schedule compared to higher-semester students, while (b) are compulsory courses at a specific stage of the study program that may not be postponed. Restricting the analysis to freshmen and mandatory courses reduces potential biases from students changing courses due to the new learning format. Descriptive statistics in Table A.2 in the Appendix show that students in academic years 2018/2019 and 2019/2020 did not take substantially less courses in the winter and summer term, respectively.

To address the concern of more lenient grading during the spring semester 2020, we conduct a comparative analysis between two groups of students: (a) those who attended courses in person and took exams using traditional methods in the winter semester 2019/2020, and (b) those who attended courses in person during the same semester but took exams online in the spring semester 2020. Both groups experienced in-person teaching during the winter 2019/2020 semester, which serves as the baseline for comparison, while only group (b) students did the exam online. We examine the variation in grades between these two groups and compare it to the corresponding groups of students in academic year 2018/2019, where no online exams or online teaching took place.

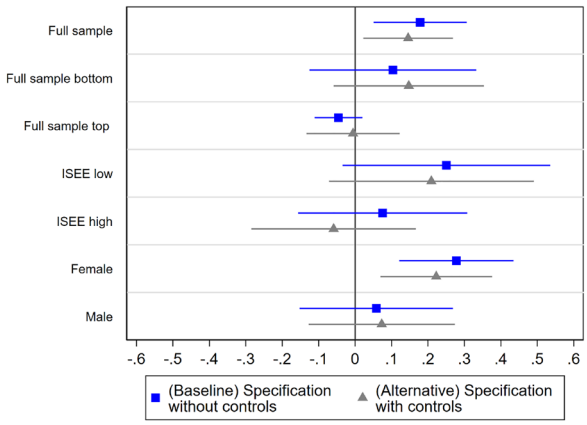
[11] Observe that the correction is similar to the approach of Canay (2011) for panel data (see e.g. Bargain et al. 2018; Castagnetti and Giorgetti 2019; Bonaccolto-Töpfer et al. 2022, for empirical applications).



(a) Full sample



(b) Bachelor



(c) Master

Fig. 3 Effects on students' average grades of transition to online teaching. The figure shows estimates of the effects on students' average grades from a difference-in-differences specification. In the (alternative) specification with controls, the dependent variable is the student-level average of corrected grades $\tilde{v}_{ist}$. Control variables used are gender, age, dummies for being a freshmen student, ISEE, master, 6-year degree or 5-year degree, 13 department dummies (department of medicine is the base category, see Appendix B for details). Point estimates with 95% confidence intervals, standard errors clustered at the individual level. Table A.2 shows the corresponding number of observations

5 Empirical results

In this section, we present the results of our estimation. In addition to analyzing the full sample of bachelor and master students, we also examine specific subgroups to gain further insights. These subgroups include top- and bottom-performing students, as well as the role of family background in student performance. Moreover, we investigate potential gender effects by separately analyzing male and female student performance.

Figure 3 presents the overall impact of the first wave of the COVID-19 pandemic on student performance examining both the full sample and selected subsamples (panel (a)). The estimation results from the specification without control variables indicate a slightly positive and statistically significant effect for the full sample, with an increase in grades of 0.128 grade points. This corresponds to a 0.47% improvement in grades based on the average grade of 27.10. Bottom-performing students are slightly negatively affected, while top-performing students show no significant impact. These findings suggest that the pandemic has had a negative effect on the academic performance of lower-performing students, whereas higher-performing students seem to be unaffected.

Examining the effects based on family background, we find slightly positive effects for students from lower-income families, but no significant effects for students from higher-income families. However, the estimated effects are not statistically different from each other, indicating that family wealth may not have played a significant role, and the pandemic may not have exacerbated

educational inequality along this socioeconomic dimension. When analyzing the effects separately for men and women, no statistically significant differences are observed. The effects are slightly positive for women, while no significant effects are found for men. These patterns generally persist when examining bachelor and master students separately (panel (b) and (c) respectively). Specifically, the negative effect on performance for bottom students is primarily driven by bachelor students, while the slightly more positive effect for women is driven by master students.

The results from the alternative specification with control variables show that the effect size for the full sample is slightly higher compared to the specification without control variables, amounting to 0.18 grade points. However, like before, this effect remains quantitatively small, representing an increase of 0.68% in average grades (based on an average grade of 27.1) due to the pandemic. The results indicate that there are no longer statistically significant effects for bottom students, top students, or students from higher-income families. However, we now find slightly positive and statistically significant effects for both female and male students. The point estimates for the different subgroups do not differ significantly, consistent with the results from the specification without control variables. Thus, in line with previous findings, we find no significant heterogeneous effects across various socioeconomic dimensions.

By looking at the effects for bachelor (panel (b)) and master students (panel (c)), we see again that bottom bachelor students are negatively affected. Panel (c) of Fig. 3 shows that the positive effect for female students is driven by master students. The effects do again not differ statistically significantly from each other. The findings for the full sample do not change substantially for bachelor and master students. The only exception is that bachelor-bottom students experience a statistically significant, though economically small negative effect (−0.20 grade points or 0.88%).[12]

Overall, our analysis indicates that the first wave of the COVID-19 pandemic had no substantial effect on average grades among university students. Overall, there seems to be a slight tendency for average grades to increase following the transition to online teaching. However, when examining the effects within different subsamples, the direction of the effects becomes more ambiguous. For example, we observe a negative effect on average grades for bottom-performing bachelor students, no significant effect for top-performing master students, and a positive effect for female master students. It is important to note that even within these subsamples, the effects are small, generally amounting to less than 1% of the average grade.

Moreover, our findings suggest that the pandemic did not lead to significant disparities in student performance based on factors such as gender, academic standing (top or bottom students), or family background. These results contrast with the findings from studies examining the effects of the pandemic on schools, which have reported substantial differences in outcomes for different groups of students. For example, existing literature has identified variations in the effects of the pandemic on academic performance for men and women (Engzell et al. 2021), top and bottom students (Aucejo et al. 2020), and students from different socioeconomic backgrounds (Agostinelli et al. 2022; Bacher-Hicks et al. 2021; Chetty et al. 2020).

6 Accounting for potential grade inflation

In this section, we focus specifically on students who attended courses during the winter term and took the corresponding exams either in the winter or the summer. This approach allows us to isolate the effects of the examination procedure, as the courses (held in the winter) were in both academic years 2018/2019 and 2019/2020 not impacted by the anti-COVID restrictions. It is important to note that the examination procedure did differ between the winter and summer terms in the academic year 2019/2020 due to the implementation of anti-COVID measures. However, there were no changes in the examination procedure throughout the academic year 2018/2019, which serves as our baseline comparison with no anti-COVID measures in place.

In the DID specification, the key distinction lies in the choice of the exam term, which serves as the treatment variable. Students who opted to postpone the exam of a winter course to the summer experienced an unexpected change in the examination procedure during the academic year 2019/2020. Therefore, the academic year 2019/2020 is considered as treatment group in this analysis. Note that the change in the examination mode occurred for both standard oral and written exams, with the transition from in-person exams to online exams. This change in the mode of examination was an unforeseen event for students who attended the course during the winter term. By comparing the outcomes of students who took their exams in different terms within the same academic year, we can assess the impact of this unexpected change on student performance.

Even though this estimation strategy allows us to compare two relatively homogeneous groups of students, we cannot rule out that apart from the mode of examination we may also capture other time-variant changes. Further, there may be negative selection in postponing

[12] Given an average grade of 22.7 in the sample of bottom-bachelor students.

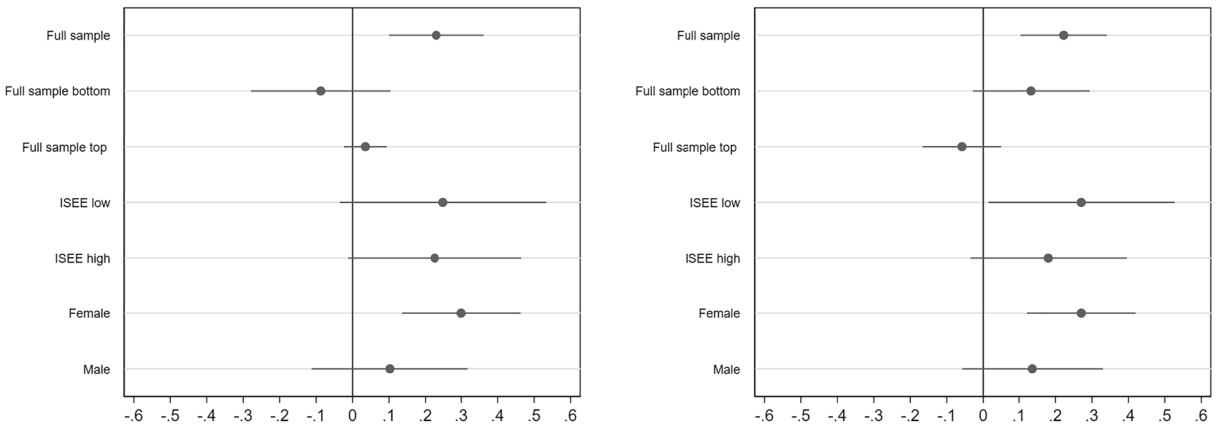


Fig. 4 Accounting for potential grade inflation—Effects on students' average grades of transition to online teaching (specification with controls). Estimates of effects on average grades for students attending classes in the winter term and doing exams in the winter or summer term in academic years 2018/2019 and 2019/2020. The figure shows estimates of the effects on students' average grade from a difference-in-differences specification with controls. Control variables used are gender, age, dummies for being a freshmen student, ISEE, master, 6-year degree or 5-year degree, 13 department dummies (department of medicine is the base category, see Appendix B for details). In panel (**a**), the dependent variable is the student-level average of grades. In panel (**b**), the dependent variable is the student-level average of corrected grades $\tilde{v}_{ist}$. Standard errors clustered at the individual level are used. Point estimates with 95% confidence intervals. Table A.2 shows the corresponding number of observations

exams. Indeed, we expect to find a tendency of bottom-performing students to postpone their exams. As a result, we may observe a decline in average exam grades between the group of students who took exams during the winter term and those who postponed their exams to the summer term. However, it is important to note that this selection process is not specific to the pandemic but occurs in regular academic years as well. If there is no significant difference in the change of exam grades over the two years, it may be considered as evidence against grade inflation.

Figure 4 displays the results of the estimation. The point estimates, either statistically insignificant or below 0.27 grade points, indicate that there is no significant impact on the outcome. The highest estimate we obtain from our analysis is 0.27 grade points, which is observed in the subsample of women when using average grades without considering lecturer fixed effects as the dependent variable. This outcome implies a marginal increase of 0.99% in grades, based on the average grade of 27.24 for female students. Similarly, the figure shows that—when accounting for grade inflation—the results for students from different social backgrounds, i.e. across different ISEE categories, are robust to those of the main analysis (see Fig. 3). We thus conclude that grade inflation does not affect the accuracy of our estimation results.[13]

During the onset of the COVID-19 pandemic, there were concerns that grades may not effectively measure student learning in certain educational systems, such as the USA, where grading is often based on a relative scale or curve. This means that grades are determined based on students' performance relative to their peers, and any changes in student learning may not be accurately reflected if the relative distribution of performance remains unchanged. However, it is important to note that at the University of Pavia and generally in Italy, grading practices do not systematically assign the highest grade to the best-performing student or follow a strict relative scale. This suggests that grades at universities without grade adjustments, like the University of Pavia, may provide a closer monitoring of actual student learning and may be more sensitive to changes in performance.

In the USA, to help students cope with the challenges posed by the pandemic, higher education institutions implemented flexible grading policies that included alternative grading options like pass-fail or credit-no credit (Mostafa et al. 2023). Such policies were not applicable in Italian universities or specifically at the University of Pavia. Mostafa et al. (2023) examine the impact of flexible grading policies in the USA during the pandemic and find that they had a positive effect in mitigating the disruptive effects on students' academic paths, particularly for freshmen. Rodríguez-Planas (2022b) also suggest

[13] Figure C.4 in the Appendix shows that grade inflation did not affect Bachelor and Master students differently.

that low-performing students from lower-income families outperformed their wealthier counterparts during the pandemic, but this advantage was fully explained by their utilization of the flexible grading policy. However, it is important to note that these findings are specific to the USA context and do not directly apply to the University of Pavia or Italy.

In our case, while the aforementioned grading practices are not common in Pavia or Italy, it is possible that professors were more lenient with students during the initial phase of the pandemic, considering the challenging circumstances students faced. To address these concerns, we conducted a regression analysis that examined the relationship between grades, course, and exam dates. We have estimated the mean and variance of the grades per course with respect to the exam session to see if the instructors use a grading curve to evaluate the students. We included dummy variables for multiple pre-treatment terms, excluding the first term, to account for any variations in grading practices over time. The regression includes term dummies for two winter and two summer terms over the academic years 2015/2016–2018/2019, as well as one term (winter) for the academic year 2019/2020.

In Fig. 5, we present evidence that grades exhibit statistically significant variations over time, contradicting the notion that a student's grade solely reflects their position within the distribution of students in a specific course during a particular term. The figure illustrates that grading practices have changed over time and indicates that grades are not solely determined by the relative performance of students within a given term. The figure also suggests that there are no statistically significant patterns during the pre-COVID period (i.e., before spring 2020). This finding is further supported by a test of joint significance of the coefficient estimates in the pre-COVID period, which rejects the null hypothesis of 0 (p-value = 0.3505). Despite this reassuring evidence, we cannot completely exclude a positive bias in student performance. If grade inflation was present during the pandemic, it may only become apparent in the long run as it would imply lower long-term human capital accumulation of students affected by the pandemic.

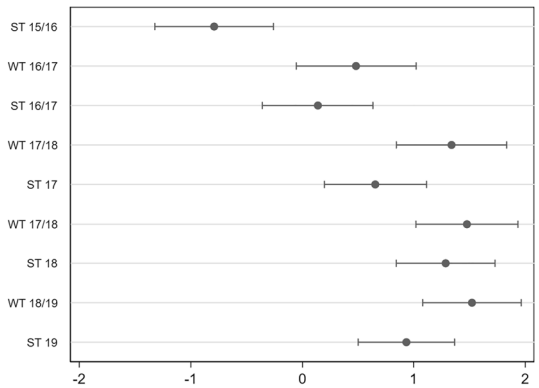


Fig. 5 Regression of the variance of grades by course and exam date in academic years 2015/2016–2019/2020. The figure shows coefficient estimates of dummies from the various multiple pre-treatment terms (both winter and summer term in academic years 2015/2067–2019/2020, respectively). The corresponding dependent variable in the regression is the grade variance per course and exam date. Point estimates with 95% confidence intervals calculated based on standard errors clustered at the course level

7 Discussion

Our analysis suggests at most tiny consequences of the shift to online teaching during the initial wave of the COVID-19 pandemic at the University of Pavia. Overall, students' average grades are slightly positively affected. However, the estimated effects varied by course characteristics and students' socioeconomic characteristics and academic profiles. The effects are small throughout (at most 1%).[14] Considering this unexpected finding of no significant variation in student performance attributed to the COVID-19 pandemic, it is crucial to take the potential factors that may have influenced the outcomes for students during this period into account.

One factor that might have been expected to impact student performance is the deterioration of health and economic conditions, both for students themselves and their families. The unprecedented challenging circumstances during the pandemic may have had implications for students' abilities to concentrate on their studies or to access necessary resources. In order to better understand these trajectories, we specifically analyze the performance of top- and bottom-performing students as well as students from low- and high-income families. However, our results indicate that the COVID-19 pandemic did not lead to an exacerbation of performance disparities based on family background or individual ability.[15]

To address concerns about self-selection bias, we examine the possibility that students may have taken fewer exams in the summer of 2020, opting for subjects where they believed they had a higher chance of success. This selective behavior could introduce bias if certain subjects were perceived as easier or more appealing during the pandemic. However, our analysis indicates that there is no observable disparity in the average number of exams taken during this period. Nonetheless, we

[14] This applies to the regressions with both dependent variables: average grades and average grades free of lecturer fixed effects.

[15] These findings are in contrast to findings from the literature of the COVID-19 pandemic on schools (among others Aucejo et al. 2020; Bacher-Hicks et al. 2021; Chetty et al. 2020; Agostinelli et al. 2022).

cannot rule out that students took different courses with different contents possibly influenced by perceptions of difficulty or personal interests. For instance, due to the pandemic, students may have chosen to delay a theoretical mandatory course instead of one involving lab work. Nonetheless, we exclude elective courses, where students have greater flexibility in their choices, from the analysis. Focusing only on mandatory courses reduces the likelihood of self-selection bias in subject choice. Also when considering only mandatory exams, we find no significant impact of the pandemic on student performance or the assessment of educational quality, as detailed in the results provided in Appendix C.

Further, lecturers may have lowered the level of difficulty of exams or graded student outcomes more generously in order to compensate students for the special situation. As mentioned above, at the University of Pavia grades are not normalized but lecturers stick to the same grading scale. That is, lecturers do not systematically attribute to the best-performing student in the course the highest grade. To investigate the possibility of lecturer heterogeneity in exams and grading practices, we employ several strategies in our analysis. First, we apply lecturer fixed effects to account for any differences in grading tendencies across instructors. By including these fixed effects in our models, we can isolate the variation in grades that is specific to each lecturer. This allows us to examine whether there are systematic differences in grading across instructors that could potentially influence the results. We also calculate average grades that are free of lecturer fixed effects to provide a more comprehensive view of student performance. Second, we restrict the sample to students who experienced in-person teaching before the pandemic but differ in terms of the type of examination they underwent. By comparing students who were assessed using different methods, such as oral or written exams, we can examine if there are any differential effects on student grades based on the way of interrogation. This approach helps us to explore whether the transition to online exams during the pandemic had a significant impact on student outcomes. Finally, we analyze the association between student grades and their performance in the pre-treatment terms. By examining the relationship between grades before the pandemic and grades during the pandemic, we can assess whether there is any consistent pattern or significant deviation in student performance that can be attributed to the pandemic-related changes in teaching and assessment methods. Our findings suggest that the effects observed in student grades are primarily attributed to the transition to online teaching rather than other factors like economic uncertainty or health concerns arising from the pandemic.

Nevertheless, the results concerning student grades immediately following the lockdown could still be positively biased if we assume that lecturers have become more accommodating and lenient in their evaluations to grade students. Both lecturers and students had to quickly adapt to online teaching within a short period. Unfortunately, with the available data, we are unable to isolate and account for these effects. However, if this grade "inflation" has indeed led to reduced accumulation of human capital by students during online teaching, it could have significant long-term consequences on the labor market, potentially disrupting standard labor market signals. As studies on educational institutions indicate (e.g. Engzell et al. 2021), the full impact of these consequences can only be verified in the long run. Mostafa et al. (2023), for instance, underline that the flexible grading policy in the USA may have negative effects for students in subsequent courses.

Our findings stand in contrast with large part of the existing literature, which primarily focuses on schools and often relies on survey or macro data. The use of administrative data in our study provides a different perspective and adds to the literature on this topic. Additionally, it is important to recognize that the effects of the transition to online teaching may vary between school and university students. The incentive mechanisms and voluntary nature of university education differ from compulsory schooling, which may contribute to different outcomes and experiences for students.

The efforts invested in online teaching during the pandemic may have been particularly significant for university students due to their understanding of the importance of human capital accumulation for their future careers and lives. As young adults, university students are generally adept at adapting to changes and dealing with technical challenges. Furthermore, university students were likely better prepared for online learning as many of them already owned computers or other necessary technical devices prior to the pandemic. This advantage may not have been as prevalent among primary or secondary school students, especially considering the increased demand for such devices by parents due to remote work arrangements (Adams-Prassl et al. 2020). Additionally, it is worth noting that legal measures implemented for universities and schools differed, especially during the first wave of the pandemic. While university students did not experience a reduction in workload, primary and secondary schools in Italy typically reduced the number of instructional hours per week from 30 to 15. Taken together, these factors suggest that university students may have been more prepared for and more capable of adapting to online teaching compared to

students at lower education levels. This could have influenced their ability to navigate the transition effectively and potentially mitigate the negative effects on their academic performance.

Finally, as highlighted in the existing literature (Wagner et al. 2016; Boring 2017; Aucejo et al. 2022), it is important to consider gender issues in education. To account for these issues, we include lecturer fixed effects in our regression analysis and examine the effects separately for male and female students. The results indicate that there are no substantial gender differences observed in student outcomes during the transition period.

8 Conclusion

This paper suggests that the first wave of the COVID-19 pandemic did not represent a threat to higher education in the direct aftermath of the pandemic at the University of Pavia. We estimate the effect of the transition to online teaching due to the first wave of the COVID-19 pandemic on student performance (average grades). This study considers the effects of the pandemic on higher education using administrative university data. The rich data set allows us to control for various socioeconomic dimensions such as family income or gender. Our results may be of interest for academic institutions all around the world as the coronavirus raged worldwide. Nonetheless, the case of Pavia represents a worst-case scenario characterized by an early and prolonged period of online teaching (spanning the entire semester).

During the spring of 2020, the transition to online teaching took place in Lombardy and other areas of Northern Italy contemporaneously with the start of the summer term. This temporal alignment allows us to identify and examine the specific effects of the COVID-19 pandemic on higher education. The estimation approach consists of a difference-in-differences setup, where the corresponding effect is derived from a comparison between the average grade differences during the summer and winter semesters for the academic years 2019/2020 and 2018/2019, respectively.

The results suggest no substantial effect of the COVID-19 pandemic on students' grades. The general tendency of the estimated effects on student performance is slightly positive. Our results are robust across various subsamples (freshmen, mandatory courses) and along different dimensions (gender, rich/poor family background, top/bottom performance students). Our findings differ from those of the COVID-19 lature on schools that identified significant negative effects along various dimensions (Agostinelli et al. 2022; Engzell et al. 2021; Bacher-Hicks et al. 2021). Similarly, our results differ from those of Aucejo et al. (2020) and Jaeger et al. (2021) that found pronounced negative effects of the pandemic for university students' outcome and expectations based on survey data. In contrast to the literature of the COVID-19 pandemic on education (Aucejo et al. 2020; Bacher-Hicks et al. 2021; Chetty et al. 2020; Agostinelli et al. 2022), we find no marked adverse effects for top or bottom students, for students from poorer or richer families or for men and women.

Several tests support our conclusions providing no evidence for grade inflation. Moreover, we consider multiple pre-treatment periods when analyzing the pattern of grades before the COVID-19 pandemic hit higher education in Lombardy. The data reveals no time-consistent grading pattern and hence does not support concerns of grading on a relative scale or curve where students' grades represent only their position in the distribution of students in a given term and course. Nonetheless, a potential positive bias in student performance cannot be entirely ruled out. If grade inflation occurred during the pandemic, it may have had negative implications for students' long-term human capital accumulation, which may only become apparent over a longer time horizon. This paper focuses solely on the short-term effects of the COVID-19 pandemic and does not address its impact in the medium and long run. Further research is needed to investigate the potential medium and long-term consequences of the pandemic on student outcomes and educational quality.

A caveat of the study is that it does not look at higher education dropout. Indeed, socioeconomic disparities may contribute to different dropout rates among diverse student groups. Factors such as economic constraints, lack of family support, limited access to educational resources, external pressures, and academic challenges could all play a role. This aspect merits further investigation and should be on the research agenda of future studies.

In summary, our analysis suggests that the institution we examine was only slightly affected by the COVID-19 pandemic in terms of average grades indicating no marked impact in the short term. Thus, in the specific context of the University of Pavia, our study finds no clear evidence of a negative impact on student performance challenging concerns of increased socioeconomic gaps due to the pandemic. Our study motivates further research for other Italian and international universities. Further, it would be interesting to investigate whether the transition to online teaching has any negative effects on labor market outcomes for the cohorts of students affected by the pandemic. Only time will show whether students' outcomes and the overall quality of education have remained stable or have experienced a decline.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12651-024-00371-0.

Supplementary Material 1.

Acknowledgements
We thank Paolo Bertoletti, Andrea Geraci, Giulio Pedrini and Claus Schnabel for helpful comments and suggestions and in particular the University of Pavia for data provision.

Author contributions
All authors contributed equally to all sections and steps of the study.

Funding
The authors have no relevant financial or non-financial interests to disclose.

Availability of data and materials
Access to the research data on student performance at the University of Pavia is confidential and owned by the University of Pavia. Therefore, we cannot deposit the files in an openly accessible data repository.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors have no Competing interests to declare that are relevant to the content of this article.

Received: 16 October 2023 Accepted: 22 June 2024
Published online: 17 August 2024

References

Adams-Prassl, A., Boneva, T., Golin, M., Rauh, C.: Inequality in the impact of the coronavirus shock: evidence from real time surveys. J. Public Econ. **189**, 104245 (2020)
Agostinelli, F., Doepke, M., Sorrenti, G., Zilibotti, F.: When the great equalizer shuts down: schools, peers, and parents in pandemic times. J. Public Econ. **206**, 104574 (2022)
Aucejo, E.M., French, J., Ugalde Araya, M.P., Zafar, B.: The impact of COVID-19 on student experiences and expectations: evidence from a survey. J. Public Econ. **191**, 104271 (2020)
Aucejo, E.M., Fruehwirth, J.C., Kelly, S., Mozenter, Z.: Teachers and the gender gap in reading achievement. J. Human Capital **16**(3), 372–403 (2022)
Bacher-Hicks, A., Goodman, J., Mulhern, C.: Inequality in household adaptation to schooling shocks: covid-induced online learning engagement in real time. J. Public Econ. **193**, 104345 (2021)
Bargain, O., Etienne, A., Melly, B.: Public sector wage gaps over the long-run: evidence from panel administrative data. IZA Discussion Paper No., **11924**, (2018)
Bonaccolto-Töpfer, M., Castagnetti, C., Prümer, S.: Understanding the public-private sector wage gap in Germany: new evidence from a fixed effects quantile Approach. Econ. Modell. **116**, 106037 (2022)
Boring, A.: Gender biases in student evaluations of teachers. J. Public Econ. **145**, 27–41 (2017)
Bulman, G., Fairlie, R.: The impact of COVID-19 on community college enrollment and student success: evidence from California administrative data. Educ Financ. Polic. **17**(4), 745–764 (2022)
Canay, I.A.: A simple approach to quantile regression for panel data. Econom. J. **14**(3), 368–386 (2011)
Castagnetti, C., Giorgetti, M.L.: Understanding the gender wage-gap differential between the public and private sectors in Italy: a quantile approach. Econo. Modell. **78**, 240–261 (2019)
Chetty, R., Friedman, J.. N.., Hendren, N., Stepner, M., Team, T.. O..: I: how did COVID-19 and stabilization policies affect spending and employment? A new real-time economic tracker based on private sector data, vol. 91. National Bureau of Economic Research, Cambridge (2020)
Engzell, P., Frey, A., Verhagen, M.: Learning loss due to school closures during the covid-19 pandemic. Proc. Natl. Acad. Sci. **118**(17), e2022376118 (2021)
Jaeger, D.A., Arellano-Bover, J., Karbownik, K., Martínez-Matute, M., Nunley, J.M., Seals, A., Alston, M., Becker, S.O., Beneito, P., Böheim, R., et al.: The global COVID-19 student survey: first wave results, p. 14419. IZA, Bonn (2021)
MIUR. Portale dei dati dell'istruzione superiore. http://ustat.miur.it/dati/didattica/italia/atenei-statali/pavia (2023) Accessed 16 Jun 2023.
Mostafa, S. A., Ferguson, R., Tang, G., Ashqer, M.: An analysis of the covid-19-induced flexible grading policy at a public university. Higher Education Policy 1–34 (2023)
Rodríguez-Planas, N.: COVID-19, college academic performance, and the flexible grading policy: a longitudinal analysis. J. Public Econ. **207**, 1–11 (2022a)
Rodríguez-Planas, N.: Hitting where it hurts most: COVID-19 and low-income urban college students. Econ. Educ. Rev. **207**, 1–11 (2022b)
StudyPoint (2021). 15000-25000 students – college admissions profiles. https://www.studypoint.com/admissions/category/size/15000-25000-students/ (2021). Accessed 8 Dec 2021
Wagner, W., Göllner, R., Werth, S., Voss, T., Schmitz, B., Trautwein, U.: Student and teacher ratings of instructional quality: consistency of ratings over time, agreement, and predictive power. J. Educ. Psychol. **108**(5), 705–721 (2016)

Publisher's Note