EdWorkingPaper No. 23-810

COVID-19 Diagnoses and University Student Performance: Evidence from Linked Administrative Health and Education Data

Timothy F. Harris
Illinois State University

C. Lockwood Reynolds
Kent State University

We analyze the impact of COVID-19 diagnoses on student grades, retention, and on-time graduation at a large public university. Even though COVID-19 rarely causes major health complications for a typical university student, diagnosis and quarantine may cause non-trivial disruptions to learning. Using event study analysis, we find that a COVID-19 diagnosis decreased a student's term grade point average (GPA) modestly by 0.08 points in the semester of diagnosis without significant effects afterward. The results were the most pronounced for male students, individuals with face-to-face instruction, and those with higher GPAs before the pandemic. We do not find a significant increase in the incidence of failing or withdrawing from a course due to diagnosis. In addition, we find no general evidence that the diagnoses delayed graduation or significantly altered first-year retention. However, the University experienced significant grade inflation during the pandemic, which exceeded the estimated effects of any COVID-19 diagnoses.

VERSION: July 2023

COVID-19 Diagnoses and University Student Performance: Evidence from Linked Administrative Health and Education Data

Timothy F. Harris & C. Lockwood Reynolds*

May 22, 2023

Abstract: We analyze the impact of COVID-19 diagnoses on student grades, retention, and on-time graduation at a large public university. Even though COVID-19 rarely causes major health complications for university students, diagnosis and quarantine may cause non-trivial disruptions to learning. Using event study analysis, we find that a COVID-19 diagnosis decreased a student's term grade point average (GPA) modestly by 0.08 points in the semester of diagnosis without significant effects afterward. The results were more pronounced for male students, individuals with face-to-face instruction, and those with higher GPAs before the pandemic. We do not find a significant increase in the incidence of failing or withdrawing from a course due to diagnosis. In addition, we find no general evidence that the diagnoses delayed graduation or significantly altered first-year retention. However, the University experienced significant grade inflation during the pandemic, like other institutions, which exceeded the estimated effects of any COVID-19 diagnoses.

Keywords: COVID-19, Higher Education, Student Success, Student absences

JEL Codes: I10, I21, I23

* Harris: Department of Economics, Illinois State University, Campus Box 4200, Normal, IL, 61790, tfharr1@ilstu.edu; Reynolds: Department of Economics, Kent State University, PO Box 5190, Kent, OH, 44242 creynol9@kent.edu. We thank German Blanco, Jeremiah Harris, Todd Jones, and Aaron Yelowitz for helpful comments. Data availability statement: the authors use confidential, non-public data that is deidentified to protect university students' privacy and comply with the Family Educational Rights and Privacy Act (FERPA) and Health Insurance Portability and Accountability Act (HIPPA). Disclosure statement: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. As the data was previously collected and deidentified, IRB approval was not required.

I. Introduction

By the end of May 2021, there were at least 700,000 cumulative cases of COVID-19 among students and employees at colleges and universities in the United States (The New York Times, 2021). While a vast majority of these university students ultimately recovered, and many did not experience severe symptoms, the diagnosis itself had the potential to influence academic performance significantly. Possible disruptive channels include COVID-19 symptoms (e.g., cough, fatigue, fever, nausea)[1] (Ihm et al., 2021; Centers for Disease Control and Prevention, 2022*b*), increased stress and anxiety (Essadek and Rabeyron, 2020; Wathelet et al., 2020), negative psychological responses to mandatory quarantine (Brooks et al., 2020), missed lectures and assessments (especially for face-to-face instruction), and "brain fog" from long COVID, which affects memory and cognitive function (Ceban et al., 2022; Premraj et al., 2022).[2]

While the general effects of the COVID-19 *pandemic* on education are well-studied,[3] this paper represents, to the best of our knowledge, the first to analyze the influence of a COVID-19 *diagnosis* on university student academic performance. We use longitudinal administrative records of students from a large public university (henceforth "the University") combined with detailed data on reported absences to estimate the effect of a COVID-19 diagnosis on term grade-point average (GPA), student retention, and on-time graduation at the University.

Using event study analysis with individual fixed effects on students who began at the University in the fall cohorts from 2017 to 2020, we find that a COVID-19 absence through the Spring 2021 semester resulted in a 0.08 point decline in GPA in the semester of diagnosis but no evidence that the effects extended to the next semester. While not a seemingly debilitating effect, we demonstrate that it is roughly the same magnitude as absences for bereavement and family emergencies. Furthermore, we find that significantly larger effects (0.15 decrease in GPA) for students with face-to-face (F2F) courses or labs suggesting

[1]Incidence of symptoms is not well-documented for this exact population. Some early evidence suggested that about one out of five positive cases in a college setting were asymptomatic (Lewis and Bhavnani, 2020).

[2]Among students who reported that they had COVID-19 in the 2022 survey, four percent were diagnosed with long-COVID, and another 15 percent suspect they may have or have had long-COVID (American College Health Association, 2022). Landry et al. (2023) collect data on students, faculty and staff at Georgetown University who were diagnosed with COVID-19 finding a 36% prevalence of long COIVD. These long-term effects are increasingly being highlighted as potentially problematic for students (Magee, 2022).

[3]Brodeur et al. (2021) provides a nice summary of the early literature in economics studying the effects of the pandemic across a wide range of topics, noting that by November 2020, there were over 200 IZA working papers and nearly 250 NBER working papers related to COVID-19.

that the general transition to online instruction—the common modality for institutions during the time—mitigated the adverse effects. Outside of differences in course modality, we find substantial heterogeneity in treatment effects with larger estimates for male students, sophomores and juniors, and those with the highest pre-pandemic GPA.

In contrast to the negative effects of COVID-19 diagnosis, we show substantial grade inflation in general associated with the pandemic (0.21 point increase in term GPA). These general increases in grades may have offset some of these negative diagnosis effects as we find no evidence that COVID-19 absences lowered on-time graduation or first-year retention generally. The elevated grades continued even when classes returned to F2F instruction suggesting an acceleration of a trend in grade inflation that has positively affected graduation rates in higher education (Denning et al., 2022).

This study makes several contributions to the literature. First, while there is a large literature on the general effects of the pandemic on higher education (see Section II), we know of no research on the impacts of COVID-19 diagnosis on student academic success.[4] The safeguards and sensitive nature of student educational records linked to protected health information have impeded previous research on this crucial topic. To protect university students' privacy and comply with the Family Educational Rights and Privacy Act (FERPA) and Health Insurance Portability and Accountability Act (HIPPA), we, in consultation with University administrators and general counsel, took several precautionary steps to appropriately silo and deidentify the data. To the best of our knowledge, Essadek and Rabeyron (2020)—which use *survey* data early in the pandemic to document increased anxiety among French university students with a COVID-19 diagnosis—is the only other empirical study of university students with information on COVID-19 cases.[5]

Second, we add to the overall literature on the effects of the pandemic in higher education by placing diagnosis within the context of institutional policies. Institutions changed policies to respond to both the health threats of the pandemic (e.g. online courses, mitigation rules, vaccination policies) as well as academic effects for students.[6] Furthermore, faculty may

[4] There is some evidence about negativity effects on productivity in other contexts, for example among professional soccer players in Europe (Fisher, 2021).

[5] The focus of Essadek and Rabeyron (2020) was on the pandemic effects on mental health but they were able to also explore the effects of diagnoses reported in their survey. Jaeger et al. (2021) also ask about COVID-19 diagnoses in their surveys but their focus is on exploring incidence of COVID-19 as part of studying how the pandemic affected students.

[6] While our paper focuses on the academic effects, recent research has demonstrated that university COVID-19 vaccination policies had spillover effects lowering mortality in the local community Acton et al. (2022).

have adjusted academic standards, whether in response to formal policy changes or informal encouragement to be particularly sensitive to students during the pandemic. However, it is unclear whether those responses were effective or appropriately calibrated. Bird, Castleman and Lohner (2022) finds evidence that grading became more lenient during the pandemic, and faculty focus groups in McDaniel et al. (2020) reported struggling to balance flexibility and academic rigor. The same faculty further worried that the online learning environment led to more academic misconduct among students (i.e., cheating), which could increase GPAs. Additionally, some evidence suggests that flexible grading policies during the Spring 2020 semester led to substantial grade inflation (Rodríguez-Planas, 2022*a*).

Third, we add to the literature on student absences and shocks in higher education. The literature has noted that students may be too present-oriented and rely too much on routines (Lavecchia, Liu and Oreopoulos, 2016), and thus students may not respond well to disruptions. There is a suggestion that mandatory course attendance might be important for students to overcome some of their behavioral biases, as there is strong evidence that course attendance improves course outcomes (Dobkin, Gil and Marion, 2010). However, this literature typically does not have information on why a student is absent, and interventions are often designed to incentivize attendance, implicitly targeting students making decisions to not attend class as opposed to students who cannot attend class. Thus, it is unclear what this research tells us about the impact of shocks on the individual that cause absences.

Similarly, there is little longitudinal research on the effect of health shocks to students or households on student success in higher education. Eisenberg, Golberstein and Hunt (2009) investigates mental health and student success in higher education using longitudinal student records and individual fixed effects models, finding evidence that depression and anxiety are associated with lower term GPA. Similarly, research has shown that the propensity for poor mental health is associated with lower educational outcomes in longitudinal data in the United Kingdom (Cornaglia, Crivellaro and McNally, 2015). In terms of non-mental health issues, a small study in the UK finds evidence that mononucleosis (mono), a disease that causes extreme fatigue, was associated with less study time in general and an increased likelihood of discontinuing school for female students (Macsween et al., 2010).[7] Household, instead of individual, health shocks may also matter as Johnson and Reynolds

[7] Williams-Harmon, Jason and Katz (2016) note that up to 5% of students are infected with infectious mononucleosis during college.

(2013) finds evidence that week-long hospitalizations of household members negatively affect college attendance and attendance shifts to closer institutions or two-year colleges, particularly among the oldest male children. More research is needed about the effect of these shocks so that institutions can craft appropriate policies to assist students.

The paper is organized as follows. Section II gives an overview of the influence of the pandemic on higher education; Section III provides institutional details and a timeline, including the transition to online learning; Section IV describes the administrative student and absence data; Section V presents the empirical analysis; and Section VI concludes.

II. Effects of Pandemic on Students in Higher Education

While we are unaware of any existing research on the effects of COVID-19 diagnosis on university student outcomes, there is significant work on the general effects of the pandemic on students in higher education. The most extensive related literature has focused on the general pandemic effects on students' mental health (see Jehi et al. (2022) for a recent review). College students historically under-utilize mental health services (Browning et al., 2021), the pandemic strained existing services (Lederer et al., 2021), and there were fewer students helped in college counseling centers during the pandemic compared to before (Center for Collegiate Mental Health, 2021*b*). Data from counseling centers collected by the Center for Collegiate Mental Health (2021*a*) suggested that a third of students sought treatment due to the pandemic, while students seeking help for any reason reported that the pandemic affected mental health (72 percent), motivation/focus (68 percent), loneliness/isolation (67 percent), and academics (66 percent). However, students also reported concerns about finances (35 percent) and worries about their own health (26 percent), and the health of others (30 percent).

Similarly, academic studies based on student populations across the United States and other countries have documented that the spread of the pandemic was associated with an increase in stress, anxiety, and depression among college students (e.g., Aucejo et al., 2020; Browning et al., 2021; Essadek and Rabeyron, 2020; Fruehwirth, Biswas and Perreira, 2021; Jaeger et al., 2021; Logel, Oreopoulos and Petronijevic, 2021; McDaniel et al., 2020; Odriozola-González et al., 2020). Additional work shows significant heterogeneity in these effects, frequently finding larger impacts for female students (Jaeger et al., 2021; Logel, Oreopoulos and Petronijevic, 2021; Browning et al., 2021; Padrón et al., 2021; Aucejo et al., 2020), minorities (Jaeger et al., 2021), and lower-income and first-generation college students (Jaeger et al., 2021; Aucejo et al., 2020; Rudenstine et al., 2021).

Students were concerned about the disease for themselves, family, and friends (Jaeger et al., 2021) but were also concerned about the quality of their education, the likelihood of continuing their studies, and the lack of social connections. The rapid switch to online education was not seen as particularly effective (McDaniel et al., 2020), and students often experienced technological problems (Logel, Oreopoulos and Petronijevic, 2021; McDaniel et al., 2020; Jaeger et al., 2021). Even beyond the initial switch, the online coursework was seen as less effective, and some research demonstrated that students did worse in this environment (e.g., Bird, Castleman and Lohner, 2022; Kofoed et al., 2021). Social connectedness was an important mediating strategy for students (Logel, Oreopoulos and Petronijevic, 2021), and recent evidence suggests that online peer mentoring helped to mitigate some of the adverse effects of the pandemic (Hardt, Nagler and Rincke, 2022). Other research showed that loss of social interactions was an important mechanism for mental health decline (Browning et al., 2021; Logel, Oreopoulos and Petronijevic, 2021; Prowse et al., 2021). Jaeger et al. (2021) indicated the lack of in-person classes was a major reason students were considering not returning to school, and students expected the pandemic would decrease the likelihood of on-time graduation, particularly among lower-income and first-generation college students (Aucejo et al., 2020; Rodríguez-Planas, 2022*b*). Consequently, Aucejo, French and Zafar (2021) found a high willingness to pay for in-person classes and social activities.

Overall, this literature suggests that the pandemic could have non-trivial negative effects on educational outcomes of college students, although that may be counteracted by flexibility in grading practices (Rodríguez-Planas, 2022*a*). Notably, few of these papers provide specific information about COVID-19 diagnosis itself. Some surveys ask about diagnosis: the US subsample in Jaeger et al. (2021) shows slightly elevated rates of COVID-19 diagnosis, as well as elevated COVID-related deaths among family and acquaintances, among Black and Hispanic students while Essadek and Rabeyron (2020) finds evidence of higher depression among students diagnosed in France. However, the literature is clear that students were disadvantaged during the pandemic, which could lead them to be more susceptible to disruptions caused by COVID-19 diagnosis and quarantine.

III. Institutional Details and Timeline

Prior to describing the main analysis, we present a comparison of the University to other similarly sized public universities using data from the Integrated Postsecondary Education Data System (IPEDS) to gauge the extent to which the results would apply in other settings. Specifically, we compare the University to the 229 public universities with enrollment of at least 10,000 students, comprising the two largest institution size categories in the IPEDS data, and complete data for all variables utilized. Combined with the University, this sample enrolled approximately 4.5 million undergraduate students in 2018, equivalent to 59 percent of all undergraduates at public four-year institutions and 41 percent of undergraduates at all four-year institutions. Figure 1 reports the comparison and illustrates that the University is comparable across multiple dimensions to the median of similarly-sized institutions. Nonetheless, the University has a higher proportion of White (non-Hispanic) students, with the difference originating from a relatively lower share of Hispanics and international students. The University students are also more likely to receive aid, use loans to finance their education, and had marginally lower entrance exam scores. However, except for the percent of students receiving loans, the University characteristics are within the interquartile range of similarly-sized institutions. The similarities suggest that the results of this analysis likely have some external validity.

Related to any difference in student populations across universities, the varying responses by higher education administrators to the pandemic also potentially impact the influence of a COVID-19 diagnosis and are discussed here. Like many institutions, the University transitioned to fully online courses starting in March 2020. The University first extended spring break and later announced that online instruction would continue for the remainder of the semester. Due to the sudden change in instructional format, for one semester only (Spring 2020), the University allowed students to change the grading criteria to pass/fail throughout the semester.[8] Furthermore, the University encouraged "a spirit of flexibility that will help ensure positive student outcomes." A combination of this encouragement and empathy for the educational disruptions resulted, anecdotally, in more flexible deadlines,

[8]Technically, students needed to elect the pass/fail option by early May. Nonetheless, students could switch the grading criteria back to letter grades even after final exams were administered and graded. Consequently, strategic students changed all courses to pass/fail and merely switched classes back to letter grades that they did well in. We explore this further in section V.C.

additional extra credit opportunities, and less stringent grading criteria.

The entirety of the Fall 2020 semester and the Spring 2021 semester were also primarily conducted online. Nonetheless, 21.9 and 31.4 percent of students had at least one class categorized as "in person" respectively during the Fall 2020 and Spring 2021 semesters due largely to courses with a laboratory component, directed research, or professional practice requirements.[9] Even though the mode of instruction was primarily online for these two semesters, the number of students with active contracts to live in the dorms still accounted for roughly 56 percent of the Spring 2020 active contracts. Thus a significant number of students were still living on campus and would therefore be subject to substantial disruptions from isolation policies, including potentially having to temporarily switch housing during isolation (including going home or staying in an alternate dormitory). Unfortunately, we do not have campus residency data to isolate this effect, but this is one potential avenue through which a COVID-19 diagnosis could affect student outcomes.

At the tail end of the Spring 2021 semester (mid-April), the general population in the state became eligible to receive a COVID-19 vaccine. Consequently, some students might have been vaccinated at the end of the Spring 2021 semester. Nonetheless, the demand for vaccinations exceeded the supply initially leading to rationing and non-trivial waits. Thus, the number of vaccinated students during the Spring 2021 semester was likely small. Furthermore, individuals were only considered fully vaccinated two weeks after their last injection. Thus the vaccine had no influence on students during the Fall 2020 semester and likely did not significantly influence most students for the Spring 2021 semester. In contrast, the widespread availability of the COVID-19 vaccine during the summer of 2021 led to the University reopening for F2F instruction for the Fall 2021 semester. For the main analysis, we exclude the Fall 2021 semester due to increased concerns for endogeneity resulting from selection into vaccination status, which is correlated with COVID-19 diagnosis and potentially correlated with the primary outcome variables.[10]

[9] Only a small share of *lectures* continued to be held in a F2F environment during the Fall 2020 and Spring 2021 semesters. Respectively, 4.2 and 5.8 percent of lectures (weighted by number of students) continued to be held in the F2F environment.

[10] In addition to the concerns for endogeneity in the Fall 2021 semester, the absence data we possess only accounts for a portion of the semester.

IV. Data

We use administrative education and absence data from the University which were deidentified to comply with FERPA and HIPPA requirements and maintain the confidentiality of sensitive student records. For the primary analysis, we use a sample of first-time in college (FTIC) undergraduate students from the Fall 2017, 2018, and 2019 cohorts. This restriction creates a more homogeneous sample and allows for observation prior to the pandemic-induced educational changes. The estimation sample consists of 10,390 unique students, of whom 7,467 were enrolled during the Spring 2021 semester.

To analyze COVID-19 diagnoses, we use detailed information on excused absences at the University. The institution uses a centralized system wherein any student seeking an official absence contacts the "Dean of Students" office and provides documentation. The office then emails *all* of the student's instructors specifying the duration and a broad reason for the excused absence (e.g., bereavement, illness, family emergency). Students are incentivised to officially report legitimate absences as they will be able to take exams late and turn in assessments after the due dates without being penalized for the duration of the excused absence (e.g., for a COVID-19 diagnosis, students were generally granted 10 days worth of excused absences). Instructors commonly direct students to the Dean of Students office when approached about absences as it provides uniformity, less hassle for the instructor (e.g., the instructor does not need to verify the reason for the absence), and consistent communication to all of the student's instructors.

Overall, there are 546 reported COVID-19 diagnoses for the main estimation sample including 428 (5.5 percent) in the Fall 2020 semester and 118 (1.6 percent) in the Spring 2021 semester. These diagnoses are in-line with the percentage of reported statewide COVID-19 cases (5.2 percent during the fall 2020 semester and 2.3 percent during the spring 2021 semester).[11] Additionally, these rates a substantially higher than other types of absences in the data (see Appendix Figure A1). Thus, there is little evidence of underreporting.[12]

[11]The county COVID-19 rates are similar to the statewide COVID-19 rates (5.5 percent during the fall 2020 semester and 2.9 percent during the spring 2021 semester). Nonetheless, the statewide rates likely serve as a better comparison as most students have a permanent residence in-state, but were not physically on campus for the semesters of interest.

[12]We do not observe COVID-19 cases during school breaks or for the Spring 2020 semester. That means some control units (i.e., those coded as not having the diagnosis) were potentially diagnosed with COVID-19 outside the semester. This would attenuate our results but only if there are long-term effects of COVID-19 (such as long COVID), since we can measure treatment and control in the same semester. Given results later that show little to long-term effects, we believe the absence of out-of-semester COVID-19 case data does not significantly affect our estimates.

Nonetheless, in Section V.B we discuss and analyze potential bias to our estimates originating from possible mismeasurement of cases.

In addition to the absence data, we have detailed information on student demographics, college entrance information (e.g., high school GPA and ACT/SAT scores), term GPA, graduation date, and department of major.[13] Furthermore, we use the zip code from a student's permanent address to provide additional information about the student's background. Using 5-year American Community Survey (ACS) data from 2015-2019, we calculate the share of adults with a four-year college degree, median household income, and the percentage with broadband internet at the zip code level. In addition to those variables, we also have student-course level information, including instructional method, credit hours, department of course, and the grade earned, which we use in the analyses.[14]

A. Control and Treatment Group Comparison

To understand the differences between students diagnosed with COVID-19 and those who were not during the Fall 2020 and Spring 2021 semesters, we present a simple mean comparison in Figure 2. The most pronounced demographic difference between the two populations—albeit not statistically significant—is that those who contracted COVID-19 were more likely to be female than those who did not. Nonetheless, this gender difference goes away if nursing majors—predominately female and higher risk of getting COVID-19 due to patient interactions—are excluded. In addition, the mean comparison shows that students from more advantaged areas, as measured by household income, college graduates, and internet access, are marginally more likely to get COVID-19. However, while these locational differences are statistically significant, the magnitudes likely do not rise to the level of "economic" significance.

To further understand any selection into the "treatment," Figure 3 reports the coefficients from a simple cross-section regression of COVID-19 absence on our student characteristics. As shown, there is no evidence of a significant correlation based on academic performance

[13] The University accepts either the ACT or SAT scores for college entrance. For consistency, we convert SAT scores into ACT scores for the analysis using concordance tables. Only the department of the student's major was provided to maintain anonymity, as some majors have few students.

[14] In the estimation sample, 49 do not have an ACT score, and 45 do not have HS GPA listed in the data. In addition, we are missing the zip code for 40 students and missing zip code level information on median household income, percent with a college degree, and share with broadband internet access for 46, 40, and 40 students, respectively. Nonetheless, individual fixed effects are used in the main specifications, which allows these individuals to be included except in subsample analysis directly using these variables.

predating the pandemic (e.g., cumulative GPA in the Fall 2019 semester and ACT scores), demographics (e.g., gender, race/ethnicity), or likely household characteristics (e.g., income, education, and internet access). The student's department of major was also uncorrelated with COVID-19 diagnosis except in the case of nursing and technology majors (positively associated with diagnosis). Nursing students were involved with patient care, and their elevated infection seems reasonable. It is less apparent why technology students have relatively high COVID-19 diagnoses. One potential explanation is that their jobs might have required them to physically go to different locations to maintain technology and consequently had more opportunities for infection. Nonetheless, the main takeaway from this simple regression is that there does not appear to be significant selection—especially across academic measures—into who contracted COVID-19 for this sample of undergraduate students during this period. Furthermore, as described in the next section, our primary estimation strategy will include individual fixed effects, which account for time-invariant factors that might be associated with COVID-19 infection.

V. Empirical Analysis

A. *Influence of COVID-19 on Term GPA*

To analyze the impact of a COVID-19 diagnosis on term GPA, evaluate the validity of a common trends assumption, and gauge the duration of any COVID-19 effects, we estimate the following event study model:

$$Y_{it} = \sum_k \eta_k \mathbb{I}(t - E_i = k) + \theta_1 X_{it} + \alpha_i + \gamma_t + \delta_m + \varepsilon_{it} \tag{1}$$

where Y_{it} is student i's semester GPA for term t, E_i is the semester in which individual i contracted COVID-19, and $\mathbb{I}(t - E_i = k)$ is an indicator for being k semesters from the COVID-19 diagnosis. Failure to reject the hypothesis that $\eta_k = 0\ \forall k < 0$ supports the parallel trends assumption. X_{it} is a vector of time-varying covariates, including other excused absences (e.g., bereavement, other illness, family emergency), credit hours enrolled, and the share of credit hours for major courses.[15] Fixed effects for individuals, term-year, and major are given respectively by α_i, γ_t, and δ_m.[16]

[15] We also estimated alternative models excluding X_{it}, and the main conclusions remain largely unchanged.

[16] Note that we allow student majors to change over time so δ_m allows us to compare students within a given major

An alternative approach would be to estimate a simple difference-in-difference model by replacing the set of $\mathbb{I}(t - E_i = k)$ indicators with a single indicator for having COVID-19 in that semester or previous semesters. We will consider that approach later when exploring heterogeneous responses across student characteristics. A recent literature, however, has noted that both models, while being standard and common in the literature, require stronger assumptions than previously considered and involve a weighting of underlying average treatment effects that can bias estimates (e.g., Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021; Sun and Abraham, 2021). For example, bias will occur if treatments are staggered in timing and treatment effects are not homogeneous. This is not likely to be a significant concern in our case because we only have two time periods of treatment (Fall 2020 and Spring 2021), and we will show that treatment effects do not seem to persist across periods, reducing the concern about "forbidden" comparisons of treated units to previously treated units. We will also show that our main event study results are consistent with alternative approaches, specifically those proposed in Callaway and Sant'Anna (2021) and that our difference-in-difference model produces similar estimates to both sets of event study approaches.

The primary identifying assumption for our empirical models is that the outcomes of students diagnosed with COVID-19 would have trended in the same direction if they had not been infected as students without a COVID-19 diagnosis (i.e., parallel trends). Note that random assignment of COVID-19 diagnosis is not necessarily required for parallel trends to hold. Additionally, individual fixed effects will control for underlying differences in unchanging student characteristics.[17] For example, if students with lower academic ability were more likely to be diagnosed with COVID-19 (which does not appear to be the case based on the above correlation regression), then the individual fixed effects would account for the lower acumen, and the assumption would be satisfied inasmuch as lower-ability students' grades trended in the same fashion as high ability students. Furthermore, time-invariant differences in individuals that would affect choices to employ mitigation techniques such as masking (CDC, 2022c) would also be captured by the individual fixed effects.

Figure 4 presents the event study results. The first and second panels report coefficients

during each semester.

[17] See Bird, Castleman and Lohner (2022); Kofoed et al. (2021); Rodríguez-Planas (2022*a*) for examples of studies using individual fixed effects in related literature.

for a treatment of being diagnosed in the Fall 2020 and Spring 2021 semesters, respectively. To avoid contamination across the two time periods, we exclude observations diagnosed in both semesters (three students) and those diagnosed in the non-designated treatment semester.[18] In the recent literature about event study and difference-in-difference estimates with two-way fixed effects, our approach to the sample is equivalent to comparing the students with a COVID-19 diagnosis (treated) to the set of students who are never treated in our time period.[19] Note that the prior results show that student characteristics do not predict COVID-19 diagnosis, so the treated and "never treated" are balanced on covariates suggesting that the "never treated" may be a suitable control group. Furthermore, as shown in both regressions, there does not seem to be evidence of a differential trend prior to diagnosis, which provides evidence that the estimation identifies the influence of COVID-19.[20] For example, the lack of a pre-trend for those diagnosed in the Spring 2021 semester suggests that mitigation techniques did not both significantly increase GPA and decrease infection rates. Otherwise, there should be evidence of divergence in the period before diagnosis when mitigation techniques were applied. For example, if students that were diagnosed with COVID-19 in Spring 2021 got the virus due to less "social distancing" (e.g., partying) which had a negative impact on their grades, then we would expect their grades to be worse in the Fall 2020 semester than those that took more preventative measures. Given the lack of a pre-trend in the second panel of Figure 4, this is less likely to be the case.

The statistically significant point estimates indicate that COVID-19 was associated with a 0.08 and 0.13 point decrease in GPA, respectively, for those diagnosed in the Fall 2020 and Spring 2021 semesters. In the semester prior to the pandemic (Fall 2019) the mean semester GPA was 2.95 with a standard deviation of 0.88. Thus, the effect of a COVID-19 absence is approximately 9.3 to 14.9 percent of the standard deviation of GPA prior to the

[18]For example, we exclude students diagnosed with COVID-19 in the Spring 2021 semester for the specification that is estimating the effect of getting COVID-19 in the Fall 2020 semester.

[19]Removing the Fall 2020 COVID-19 cases from the Spring 2021 results removes the "previously treated" individuals from the analysis. This group is a particular concern in the recent literature on event studies and difference-in-difference models because including them among the controls causes the estimated average treatment effects of Spring 2021 cases to depend on the evolution of Fall 2020 treatment effects. Removing the Spring 2021 COVID-19 cases from the Fall 2020 analysis removes the "not-yet-treated" observations, which are not generally seen as problematic and could be valid controls. However, removing them keeps consistency of our approach across the two time periods.

[20]While commonly investigated, a lack of pre-trends does not guarantee that the common trends assumption necessary for identification actually holds. Furthermore, the estimated pre-trends can be contaminated in situations with multiple time periods and heterogeneous treatment effect (Sun and Abraham, 2021). However, as previously discussed, our limited number of time periods, lack of evidence of long-term effects, and the way we construct our data to use only the "never treated" as controls should reduce this problem.

pandemic. Additionally, the graphs can provide some evidence of whether the COVID-19 effect carries beyond the semester. The results do not suggest longer-term effects, as the Fall 2020 treatment specification does not provide evidence that the impacts of a COVID-19 diagnosis persisted for more than the semester of diagnosis.

To get a combined estimate, we follow Callaway and Sant'Anna (2021) to estimate the group-time average treatment effects and then aggregate the separate effects to arrive at our estimate of the overall average treatment effects. In our implementation, we use "never treated" observations instead of "not yet treated" as the control group in the specification because there are only two periods of treatment. This also means we use a similar treatment-control comparison as in Figure 4. Figure 5 shows that the results are largely consistent with the separate event studies, with the average treatment on the treated (ATT) indicating that a COVID-19 diagnosis decreased term GPA by 0.08 points or 9.5 percent of a standard deviation.

B. COVID-19 Case Measurement

Given the incentives for students to report cases and the fact that University rates of COVID-19 are similar to county and state data, there is little reason to suspect that underreporting is a serious concern in our analysis. Nonetheless, while it is easy to report an absence and students are encouraged to do so, it is certainly possible that not all absences are reported (including those for COVID-19). There are three possible types of COVID-19 cases among students: 1) a COVID-19 case known to the student and reported to the University for an absence, 2) a COVID-19 case known to the student but not reported to the University, and 3) a COVID-19 case unknown to the student (and therefore not reported). Our absence data is the universe of Type 1. We have no data on Type 3, but argue that these cases are not policy relevant because there should be no effect on student academic outcomes.[21] During this time, students with symptoms of illness would likely get tested for COVID-19, which would put them into categories 1 or 2. Consequently, Type 2 cases are more concerning because they have the potential to bias our estimates.

On the one hand, if unreported cases experience negative outcomes then underreporting in general would bias our results toward zero, as treated observations would be wrongly cat-

[21] These unknown cases were likely asymptomatic or had minimal symptoms implying minimal or no no effects on course outcomes.

egorized as control observations in our analysis (Bound, Brown and Mathiowetz, 2001).[22] On the other hand, if only the most severe cases were reported by students (i.e., unreported cases represent circumstances where the students were unaffected), then the observed outcome under treatment would be too large to be interpreted as the outcome of *any* COVID-19 infection. Our results would, however, still be a valid measure of outcomes among *reported* COVID-19 absences, which could have policy relevance but would affect the interpretation of our results.

Given that we cannot identify unreported cases, we run a simulation exercise to see the potential impacts of underreporting on our point estimates. In this exercise, we assume that only half of COVID-19 cases were reported for absences (an extreme assumption given the similarity of case rates for the University, state, and county). Specifically, we randomly assigned 546 students who did *not* report having COVID-19—along with the 546 students who actually reported a COVID-19 diagnosis—to our "treated" group and estimate a two-way fixed effect model of the influence of the treatment on term GPA. Figure 6, reports the point estimates (panel a) and p-values (panel b) from 1,000 simulations. As shown, some point estimates in the simulation are larger in magnitude than what we have estimated previously. This could be associated with non-reporting of actual COVID-19 cases that had a detrimental effect or could simply be where the random assignment went to students who had a bad semester. The vast majority of the point estimates, however, are smaller in magnitude reflecting either circumstances where non-reported cases had a less detrimental effect in comparison to reported cases or more likely, where the randomly assigned student did not actually have COVID-19, which diluted the treatment effect. Regardless, the main take away is that even if there were massive amounts of underreporting, the treatment effects are still negative with a vast majority being statistically significant.

Outside of severity of COVID-19, we might be concerned that reported absences are correlated with other factors. For example, perhaps lower-income households would be more likely to get COVID-19 and have fewer technical resources (e.g., broadband internet), which could cause these students to have reduced academic achievement during the online semesters. If this were the case, any negative treatment effects would be overstated.

[22]Similarly, if students were diagnosed with COVID-19 during the summer of 2020 (which we do not see in the data for a vast majority of students) and had lingering effects (i.e., long COVID), then the treatment effect would be biased toward zero. However, Figure 5 showed no evidence of lingering effects across semesters.

Nonetheless, the correlation regression results do not indicate an increased incidence of COVID-19 for university students from lower-income areas or areas without stable internet access. Similarly, if students with more challenging courses in a particular semester were more likely to report a COVID-19 diagnosis, the estimated effects could be overstated. Nonetheless, semester difficulty—as measured by the realized grades of other students in an individual's courses—is uncorrelated with COVID-19 diagnosis (correlation coefficient of -0.01), which lessens this concern.

In addition, we might be concerned that underlying behaviors correlated with the risk of contracting COVID-19 are also correlated with academic performance. For example, perhaps students who are less likely to mitigate COVID-19 risk by social distancing or avoiding groups/crowds are less likely to do well academically; essentially, the people with more social activities during the pandemic may be less likely to spend time studying. Conversely, students "staying home" may have more time to devote to studying. We do not believe this is likely to be a significant issue in our estimates for several reasons. First, the individual fixed effects will capture any time-invariant characteristics that affect the likelihood of prioritizing social activities over school, regardless of the pandemic. Second, it is not clear that students who are being more cautious about COVID-19 are converting spare time into academic performance, particularly given that students were feeling stressed academically and struggling with motivation (e.g., Center for Collegiate Mental Health, 2021*a*). Lastly, exposure to COVID-19 is not a simple function of individual choices and activities. The airborne nature of COVID-19, combined with occasional mild symptoms, means that students could contract COVID-19 because of exposure from people they live with. Thus, it is unclear whether behavioral choices around COVID-19 would form a tight link between COVID-19 diagnosis and academic outcomes in our estimation strategy.

Lastly, it is worth noting that any selection story would have to orthogonal to the covariates and the individual fixed effects. While we cannot prove such selection is not occurring, it seems implausible that it would be large enough to drive our estimates. Nor is it likely to produce magnitudes of effects that appear "reasonable" when compared to other types of absences as we will later show. Nor is it likely to vary in such a way to produce some of the heterogeneity effects we later present. For example, the story above would most likely suggest that more academically marginal students would be most affected by a COVID-19 diagnosis. In fact, we later present evidence that the highest-performing students are the most affected. Similarly, since we might expect less negative effects on student achievement in a completely virtual setting, negative selection would likely suggest large negative effects regardless of course delivery. Again, we will later show evidence that the effects are larger for students with a face-to-face component to their courses during the semester. Thus we argue that some form of selection bias is unlikely to be the primary driver of our estimates.

C. Contextualizing the Results

While our estimation strategy provides identification of COVID-19 absences on same term GPA at the University, it is important to consider how we should interpret the magnitudes of the estimates. To do that, we leverage additional information in our data to provide more context about the COVID-19 effects within the institutional setting. First, we benchmark the impact of a COVID-19 diagnosis on term GPA by comparing the magnitude to the effects on term GPA of other student absences. To do that, we replace our event study approach with the following two-way fixed-effect model

$$Y_{it} = \beta_0 + \beta_1 COVID_{it} + X'_{it}\Gamma + \alpha_i + \gamma_t + \delta_m + \varepsilon_{imt} \tag{2}$$

where the variables are defined the same as in the event study analysis.

Figure 7 reports that a COVID-19 diagnosis decreases term GPA by 0.08 points. Thus, our switch to a difference-in-difference model with fixed effects produces a similar estimate as both the simple event study models in Figure 4 and the aggregated approach of Callaway and Sant'Anna (2021) in Figure 5. Again, this suggests that our empirical approach is not being driven by recent concerns about two-way fixed effects models.

Nonetheless, the main reason for presenting this figure is to compare the point estimate for COVID-19 to other absences.[23] As shown, the disruptive influence of a COVID-19 diagnosis is comparable to the impact of losing a close family member, an event which we might expect to have a negative impact on student success. COVID-19 absences, with a median of 12 excused days, are similar in magnitude to illnesses with absences of less than seven calendar days but have lower effects than illnesses that result in long absences

[23]Note that previously we controlled for "other illnesses" but here we split this by duration to provide a clearer comparison with COVID absences.

(e.g., extended hospitalization). These results provide context for our estimated COVID-19 effects and, perhaps, provide additional reassurance that our results are not driven by measurement problems as the estimates appear "reasonable" compared to other absences.

In addition to understanding the magnitude of the effect of a COVID-19 diagnosis on student grades, it is also vital to understand the University's general "pandemic" response. Figure 8 presents results from a simple fixed effect regression with term GPA as the outcome with only term and individual fixed effects as the independent variables.[24] As illustrated, the pandemic itself led to significant levels of grade inflation generally. The effect is particularly pronounced in the Spring 2020 semester (0.43 point increase) due in part to the option for students to change their courses to pass/fail in practice even after completion of final exams. Nonetheless, as shown in the results that use an imputed measure of term GPA (i.e., using underlying course grades for students who used the pass/fail option), there is still non-trivial grade inflation.

This finding is consistent with the evidence in Rodríguez-Planas (2022*a*) that official grade policies explained approximately half of elevated grades during the Spring 2020 semester at another institution.[25] Our evidence, however, is that this grade inflation continued well beyond the beginning of the pandemic, and beyond when there were formal policies allowing for expanded pass/fail options. One potential explanation is that students learned the content better in an online environment, although that is not consistent with the research showing lower performance in online settings in economics coursework (Kofoed et al., 2021). Perhaps academic integrity decreased (as was a faculty concern in McDaniel et al., 2020). Nonetheless, the high grades continued even in the Fall 2021 semester, which was predominately F2F. Perhaps the most likely explanation is increased leniency and flexibility by faculty members, consistent with the guidance of University administrators previously mentioned. This is also consistent with evidence that grade inflation has increased graduation rates in higher education (Denning et al., 2022) but would suggest that the pandemic has accelerated these longer term trends.

[24] To mitigate concerns that sample attrition is causing any change in GPA, we restrict the sample to continuously-enrolled individuals once they initially start in Fall 2017, Fall 2018, or 2019. For the Fall 2017 cohort, we only require that they are continuously enrolled through Spring 2021 when many of them graduate. Regardless, the overall results remain essentially unchanged without the balanced panel restriction.

[25] Students might have decreased effort in courses where they took a class as pass/fail, which would cause the calculated GPA without pass/fail to be even lower. Nonetheless, GPA is still elevated relative to the Fall 2019 semester. Alternatively, students could have increased effort to pass, but this likely represents a smaller share of students.

Regardless of the root cause, the general grade elevation of approximately 0.21 points more than compensated for the -0.08 point influence of a COVID-19 diagnosis on semester GPA. Consequently, even though the students were at a relative disadvantage *on average* to their peers who did not get COVID-19, their GPAs were higher than if the pandemic did not transpire. This suggests that stated concerns about the pandemic on grades (but not necessarily learning) by students early in the pandemic (Aucejo et al., 2020; Rodríguez-Planas, 2022*b*) may not have been realized. If the cause of the grade inflation was the reaction of the University to the pandemic, it suggests that it overcompensated, at least in terms of student academic outcomes. Of course, the University's pandemic policies (e.g., online courses) also had public health goals for which it is possible the benefit of the reduced spread of COVID-19 and potentially reduced anxiety of students may have been worth the inflated grades.

D. Heterogeneity Analysis

The main analysis identifies a modest yet meaningful influence of COVID-19 on term GPA. In this section, we analyze the heterogeneous impact of COVID-19 across several dimensions. We estimate Equation 2, including various interaction terms of COVID-19 absence with other covariates, primarily student characteristics, to identify differential effects.

Figure 9 presents results across different models, where each model includes different interactions of a COVID-19 absence with another set of covariates. The first model tests for differential effects by sex by interacting COVID-19 with an indicator for being male. The estimated effect for the base category, female, is small in magnitude and not statistically significant, suggesting little impact of a COVID-19 diagnosis on semester GPA for female students. In contrast, the interaction term shows that male students drove the negative effect of COVID-19 diagnosis with a 0.14 point decrease relative to female students. This effect is almost twice the average effect previously estimated, and would therefore represent approximately a 4.7 percent decline relative to pre-pandemic average GPA, or a decline of 15.9 percent of a standard deviation in pre-pandemic GPA. This finding is consistent with male students having become more marginal in education over time (e.g., see Fortin, Oreopoulos and Phipps, 2015). Thus, a COVID-19 diagnosis may affect male students more because they are more at risk of not being successful in college. In contrast to these gender differences, the next set of estimates suggests there is no differential impact across racial and ethnic groups at the University. The base category, White non-Hispanic, shows a decrease in term GPA from a COVID-19 diagnosis, but the interaction terms are not statistically different from zero.

The following two sets of results use information derived from the student's permanent zip code and both suggest that the negative effects are concentrated among students from less privileged backgrounds (lower-income and lower-educated zip codes). For household income and percent of the adult population with a four-year college degree, students from zip codes in the lowest tercile show large and statistically negative effects of COVID-19 absences. In contrast, in both cases, the interaction terms on higher terciles are positive but are not statistically significant. These results are potentially suggestive (based on the insignificant but positive interaction terms) of more negative effects of a COVID-19 diagnosis for students from less-privileged backgrounds. This would be consistent with the larger negative effects of the pandemic found among lower-income and first-generation college students for mental health (Jaeger et al., 2021; Aucejo et al., 2020; Rudenstine et al., 2021) and expectations of student success (Aucejo et al., 2020; Rodríguez-Planas, 2022*b*).

Next, the figure reports analysis by class standing and shows that the negative COVID-19 effect is concentrated among sophomores and juniors, with the point estimates implying that GPA decreased by 0.19 points more than freshmen.[26] What might be driving this is unclear. The lack of effect for first-year students might be due to the relatively more straightforward courses they typically take. It could be the case that experienced seniors with larger networks were better able to handle disruptions or, perhaps more plausibly, the University's policies (formal and informal) served to protect students who were on the verge of graduation. Consistent with these effects, we later find no evidence that a COVID-19 diagnosis reduced the likelihood of on-time graduation.

In the next model, we test for differences across historical student performance by interacting the COVID-19 absences with indicators for terciles of the Fall 2019 GPA. The results illustrate that students whose cumulative GPA is in the highest tercile prior to the pandemic were negatively impacted by a COVID-19 diagnosis, whereas students with lower GPAs were not. Maintaining a higher GPA presumably requires students to devote signif-

[26] We define class standing by accumulated credits earned. Thus a "freshman" is a student who has 30 or less accumulated credits, a sophomore has between 30 and 60 credits, etc.

icant amounts of time and focus to their studies, and disruptions—such as a COVID-19 diagnosis—represent a binding constraint. Conversely, students in the lowest tercile might not devote as much time to their studies, and disruptions might not necessarily cut into their study time in the same fashion. Another conjecture is that some students have low GPAs due to other time commitments, such as employment. In this scenario, these students would not have to work due to their COVID-19 diagnosis and consequently may not be as negatively influenced academically as those without employment.

There is also the potential that students with a F2F requirement are more negatively impacted than those who can complete all required tasks online. As discussed previously, the primary course modality was fully online during this time, but some courses met in person and others had an in-person component. We create an indicator for having any course that is in person during a given semester and include both the variable itself and its interaction with the COVID-19 indicator in the specification. As shown, the negative effects of COVID-19 were the most pronounced for students with an in-person component (e.g., laboratory requirement) with an estimated 0.15 point reduction in term GPA. This magnitude is quite large, approximately twice the size of the average effect estimated across all students and would represent a 5 percent decline relative to average GPA prior to the pandemic, and a decline of approximately 17 percent of a standard deviation of GPA. This result implies that the University's decision to move the vast majority of courses online likely mitigated the negative impacts of COVID-19 diagnosis.[27]

The last results presented in Figure 9 report an analysis of how the timing of diagnosis *within* a semester influences term GPA. A diagnosis early in a given semester might give the student more time to catch up in their classes, mitigating the COVID-19 effect. Alternatively, getting COVID-19 early in the semester might be more detrimental if course content builds upon early concepts (e.g., mathematics). To investigate, we interacted the COVID-19 absence with an indicator for contracting the virus in the second half of the semester. The results do not show any difference in the effect based on diagnosis in the first or second half of the semester. This could be because the effects described above cancel on

[27]Given that nursing students, and to some extent Applied Technology students, are more likely to report a COVID absence we also explored heterogeneity across majors. The results in Appendix Figure A2 show that nursing students do experience large and negative effects of a COVID diagnosis. Some of the effect appears to be due to the role of face-to-face classes as nursing students have a large share of face-to-face classes (see Appendix Figure A3). Furthermore, while not statistically significant, the point estimates are negative for Applied Science & Technology students as well as Fine Arts students, both of who have large shares of face-to-face courses.

average. It is also possible that whether an early or late diagnosis is more harmful depends on course subject, course level, and course structure in such a way that there is no apparent difference with timing on average.

E. Additional Educational Outcomes

Thus far, we have focused on the influence of a COVID-19 diagnosis on term GPA, which is, of course, an average of underlying course grades. While there is not a large effect on that average, there is a potential for the diagnosis to impact more consequential educational outcomes. For example, it is more detrimental if a COVID-19 diagnosis caused a student to earn a D or F in a course, than if the student earns a B instead of an A. Earning a D or F might require the student to retake the course or change their major. Furthermore, students may be more likely to drop a course because a poor grade could affect financial assistance. However, a student withdrawing from a course could delay their graduation. Importantly, withdrawals would not be represented in the above GPA analysis as the courses would not count towards the term GPA.

Figure 10 presents two event study analyses using the approach of Callaway and Sant'Anna (2021) for dependent variables of earning a D or F in any course and withdrawing from a class altogether. In both cases, a COVID-19 diagnosis did not increase the likelihood of these negative outcomes. This is consistent with the evidence in 9 that effects are concentrated among those students with higher GPA who are unlikely to be on the margin for either of these outcomes. Nonetheless, Appendix Figures A4 and A5 show some heterogeneous influences with some evidence of differential below-C grades by males and more course withdrawals by Black, low-income, and lower-GPA students. Thus, there is some evidence that students from more at-risk higher education populations experienced consequential negative educational outcomes due to a COVID-19 absence.

Lastly, we can consider some additional outcomes by switching to a cross-sectional analysis of a specific entering cohort. For example, students were concerned early in the pandemic that the likelihood of on-time graduation would be reduced by the various effects of the pandemic (Aucejo et al., 2020). We can test whether a COVID-19 diagnosis lowered on-time graduation among students in the Fall 2017 cohort who were scheduled to graduate "on-time" at the end of the Spring 2021 semester. Because we are focusing on a single cohort, we cannot estimate models with individual fixed effects and instead must rely on a cross-sectional analysis. This approach relies on there being no selection on unobservable characteristics, which is unknowable. However, as discussed previously, we have seen little evidence that a COVID-19 diagnosis is correlated with our set of observable characteristics, and have discussed why we believe unobserved selection is not likely to be a significant problem in our context. So we think it is informative to investigate some additional outcomes using a cross-sectional analysis, with the caveat that identification may not be as strong as our prior estimates.

To analyze the influence of a COVID-19 diagnosis on on-time graduation, we estimate the following cross-section for the Fall 2017 cohort:

$$OnTime_i = \beta_0 + \beta_1 COVID_i + X_i'\Lambda + \delta_m + \gamma_z + \alpha_a + \varepsilon_i \tag{3}$$

where $OnTime_i$ is one if the student graduates in the Spring or Summer 2021 semesters (i.e., four years following initial enrollment). X_i is a vector of covariates including gender, race/ethnicity, major, high school GPA, ACT score, cumulative GPA as of Fall 2019, and separate indicators for ever switching majors, excused absence for bereavement, excused absence for a family emergency, and excused absence for "other illness." Major, zip code and age group fixed effects are given respectively by δ_m, γ_z and α_a. The sample includes undergraduates from the Fall 2017 cohort who were enrolled in the Spring 2020 semester. Therefore, the comparison will be those that graduate "on-time" compared to anyone that does not either due to course failure, drop-out (related or not related to the pandemic), major change, etc. Note that any selection on unobservables here would have to be orthogonal to the covariates, including both major and zip code fixed effects. While possible, we do not think such selection is likely to drive our estimates given our previous discussion.

To strengthen our regression approach, we also consider a reweighting scheme based on entropy weighting (Hainmueller, 2012). The goal is to estimate our regression with weights that help to balance the covariates across treatment and control groups. This approach is considered a "doubly-robust" estimator in that the results are valid if either the regression model is correctly specified or if the weights are correctly specified (Zhao and Percival, 2017). One could create weights for the regressions using propensity score methods (predicting the treatment using the covariates), but that relies on the researcher to check that balance of the covariates is achieved. While not hard in principle, it is not always done in practice and determining at what point "covariate balance" achieved is somewhat subjective (choice of statistical tests and level of statistical significance are up to the researcher and just because covariates are not statistically difference does not mean they are the same). Unlike weighting techniques that use the propensity score, entropy weighting constructs weights that *exactly* balance multiple moments of a covariate distribution. In our case, we match both the first and second moments of the covariates and then estimate Equation 3 with the weighted sample.

The results in the top of Figure 11 indicate that a COVID-19 diagnosis does not influence on-time graduation for the Fall 2017 cohort. The null result for COVID-19 is insensitive to use of the reweighting scheme that balances the first two moments of the student characteristics for those who did and did not have a COVID-19 diagnosis. Thus, despite concerns expressed by students that the pandemic would reduce the likelihood of on-time graduation, it does not appear that COVID-19 absences themselves delayed graduation. This very well could be because of the responses by the University leading to grade inflation, which may have cushioned the impact of COVID-19 absences on graduation.[28]

Higher education research also often focuses on first-year student success, as students are most at risk of dropping out prior to the second year of college. We can explore how COVID-19 diagnoses affect outcomes of first-year students using the Fall 2020 cohort in our data and estimating versions of Equation 3 using outcomes relevant to first-year students, namely first-year GPA and second-year enrollment (retention to a second year of college). The identifying assumptions are the same as for on-time graduation and we consider both the unweighted and weighted approaches used previously. These results are also presented in Figure 11. Once again, the results indicate that a COVID-19 diagnosis did not meaningfully influence either first-year GPA or retention. These results are also consistent with Figure A2, which found little evidence of term GPA effects among first-year students.

[28]While there is not a general influence on on-time graduation, Appendix Figure A6 illustrates that on-time graduation of Hispanic students was differentially influenced by COVID-19 indicating that on-time graduation decreased by 31.0 percentage points more than non-Hispanic Whites. Nonetheless, the point estimate was derived from a total of 19 COVID-19 diagnoses of Hispanic students.

VI. Conclusion

Early evidence suggested that higher education students were quite concerned about the pandemic's effect on their educational outcomes. Universities attempted to respond to both those academic concerns as well as the public health issues of the pandemic. Research has detailed that the pandemic negatively affected mental health and student success, despite evidence of grade inflation. Despite this large literature on the effects of the *pandemic*, there is no direct evidence of the effects of COVID-19 diagnoses on university student success. This paper begins to fill that gap by utilizing student longitudinal academic records along with data on COVID-19 absences in an event study approach to explore how a COVID-19 diagnosis affected student outcomes at a large, public university.

We find that a COVID-19 diagnosis reduced term GPA by approximately 0.08, or 9.5 percent of a standard deviation. This average effect is roughly equivalent to the effect of absences due to other short-term illnesses and bereavements. We find evidence of heterogeneity in the effects, with larger negative impacts for male students and students with higher pre-pandemic GPAs. There is some evidence of smaller effects on term GPA for first-year students and seniors, which is consistent with additional evidence that COVID-19 diagnoses did not affect on-time graduation or first-year retention. And while we do not have data to explore the influence of COVID-19 diagnosis in a predominately F2F environment, our results from students with some F2F component suggest that the negative effects would likely have been exacerbated. Finally, all of these results should be understood within the context of an institutional response that led to significant grade inflation, potentially mitigating the effects of a COVID-19 diagnosis.

REFERENCES

Acton, Riley K, Wenjia Cao, Emily E Cook, Scott A Imberman, and Michael F Lovenheim. 2022. "The Effect of Vaccine Mandates on Disease Spread: Evidence from College COVID-19 Mandates." Working Paper 30303.

American College Health Association. 2022. "National Survey of College Student COVID-19 Vaccination Uptake, Attitudes, Experiences, and Intentions." *https://www.acha.org/documents/CoVAC/National_Survey_of_College_Student_COVID-19_Vaccination_Uptake_Attitudes_Experiences_and_Intentions.pdf*.

Aucejo, Esteban M., Jacob F. French, and Basit Zafar. 2021. "Estimating Students' Valuation for College Experiences." *National Bureau of Economic Research Working Paper Series*, No. 28511.

Aucejo, Esteban M, Jacob French, Maria Paola Ugalde Araya, and Basit Zafar. 2020. "The Impact of COVID-19 on Student Experiences and Expectations: Evidence from a Survey." *Journal of Public Economics*, 191: 104271.

Bird, Kelli A., Benjamin L. Castleman, and Gabrielle Lohner. 2022. "Negative Impacts From the Shift to Online Learning During the COVID-19 Crisis: Evidence From a Statewide Community College System." *AERA Open*, 8: 23328584221081220.

Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. "Measurement Error in Survey Data." In *Handbook of Econometrics.* Vol. 5, 3705–3843. Elsevier.

Brodeur, Abel, David Gray, Anik Islam, and Suraiya Bhuiyan. 2021. "A literature review of the economics of COVID-19." *Journal of Economic Surveys*, 35(4): 1007–1044.

Brooks, Samantha K, Rebecca K Webster, Louise E Smith, Lisa Woodland, Simon Wessely, Neil Greenberg, and Gideon James Rubin. 2020. "The Psychological Impact of Quarantine and How to Reduce It: Rapid Review of the Evidence." *Lancet*, 395(10227): 912–920.

**Browning, Matthew H. E. M., Lincoln R. Larson, Iryna Sharaievska, Alessandro Rigolon, Olivia McAnirlin, Lauren Mullenbach, Scott Cloutier, Tue M. Vu, Jennifer Thomsen, Nathan Reigner, Elizabeth Covelli Metcalf, Ashley

D'Antonio, Marco Helbich, Gregory N. Bratman, and Hector Olvera Alvarez. 2021. "Psychological impacts from COVID-19 among university students: Risk factors across seven states in the United States." *PLOS ONE*, 16(1): 1–27.

Callaway, Brantly, and Pedro H.C. Sant'Anna. 2021. "Difference-in-Differences with multiple time periods." *Journal of Econometrics*, 225(2): 200–230. Themed Issue: Treatment Effect 1.

Ceban, Felicia, Susan Ling, Leanna M.W. Lui, Yena Lee, Hartej Gill, Kayla M. Teopiz, Nelson B. Rodrigues, Mehala Subramaniapillai, Joshua D. Di Vincenzo, Bing Cao, Kangguang Lin, Rodrigo B. Mansur, Roger C. Ho, Joshua D. Rosenblat, Kamilla W. Miskowiak, Maj Vinberg, Vladimir Maletic, and Roger S. McIntyre. 2022. "Fatigue and Cognitive Impairment in Post-COVID-19 Syndrome: A Systematic Review and Meta-analysis." *Brain, Behavior, and Immunity*, 101: 93–135.

Center for Collegiate Mental Health. 2021*a*. "Part 1 of 5: COVID-19 Impact on College Student Mental Health." *https://ccmh.psu.edu/index.php?option=com_dailyplanetblog&view=entry&year=2021&month=02&day=01&id=9:part-1-of-5-covid-19-impact-on-college-student-mental-health*, Accessed: 7-19-2022.

Center for Collegiate Mental Health. 2021*b*. "Part 4 of 5: Impact of COVID-19 on Students Served at College Counseling Centers." *https://ccmh.psu.edu/index.php?option=com_dailyplanetblog&view=entry&year=2021&month=02&day=01&id=9:part-1-of-5-covid-19-impact-on-college-student-mental-health*, Accessed: 7-19-2022.

Centers for Disease Control and Prevention. 2022*a*. "How to Protect Yourself & Others." *https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html*, Accessed: 07-21-2022.

Centers for Disease Control and Prevention. 2022*b*. "Symptoms of COVID-19." *https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html*, Accessed: 06-27-2022.

Cornaglia, Francesca, Elena Crivellaro, and Sandra McNally. 2015. "Mental Health and Education Decisions." *Labour Economics*, 33: 1–12.

Denning, Jeffrey T., Eric R. Eide, Kevin J. Mumford, Richard W. Patterson, and Merrill Warnick. 2022. "Why Have College Completion Rates Increased?" *American Economic Journal: Applied Economics*, 14(3): 1–29.

Dobkin, Carlos, Ricard Gil, and Justin Marion. 2010. "Skipping Class in College and Exam Performance: Evidence from a Regression Discontinuity Classroom Experiment." *Economics of Education Review*, 29(4): 566–575.

Eisenberg, Daniel, Ezra Golberstein, and Justin B Hunt. 2009. "Mental Health and Academic Success in College." *The B.E. Journal of Economic Analysis & Policy*, 9(1).

Essadek, Aziz, and Thomas Rabeyron. 2020. "Mental Health of French Students during the Covid-19 Pandemic." *Journal of Affective Disorders*, 277: 392–393.

Fisher, Kai; Reade, J. James; Schmal W. Benedikt. 2021. "The Long Shadow of an Infection: COVID-19 and Performance at Work." Düsseldorf Institute for Competition Economics (DICE) Discussion Paper 368.

Fortin, Nicole M., Philip Oreopoulos, and Shelley Phipps. 2015. "Leaving Boys Behind: Gender Disparities in High Academic Achievement." *Journal of Human Resources*, 50(3): 549–579.

Fruehwirth, Jane Cooley, Siddhartha Biswas, and Krista M. Perreira. 2021. "The Covid-19 Pandemic and Mental Health of First-year College Students: Examining the Effect of Covid-19 Stressors Using Longitudinal Data." *PLOS ONE*, 16(3): 1–15.

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with Variation in Treatment Timing." *Journal of Econometrics*, 225(2): 254–277. Themed Issue: Treatment Effect 1.

Hainmueller, Jens. 2012. "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies." *Political Analysis*, 20(1): 25–46.

Hardt, David, Markus Nagler, and Johannes Rincke. 2022. "Can Peer Mentoring Improve Online Teaching Effectiveness? an RCT during the COVID-19 Pandemic." *Labour Economics*, 78: 102220.

Ihm, Laura, Han Zhang, Alexandra van Vijfeijken, and Mark G. Waugh. 2021. "Impacts of the Covid-19 Pandemic on the Health of University Students." *International Journal of Health Planning and Management*, 36(3): 618–627.

Jaeger, David A., Jaime Arellano-Bover, Krzysztof Karbownik, Marta Martínez-Matute, John M. Nunley, Richard A Seals, Mackenzie Alston, Sascha O. Becker, Pilar Beneito, Rene Boheim, José Emilio Boscá Mares, Jessica Brown, Kang-Hung Chang, Deborah A. Cobb-Clark, Shooshan Danagoulian, Sandra Donnally, Marissa Eckrote-Nordland, Lidia Farre, Javier Ferri, Margherita Fort, Jane Fruewirth, Rebecca Gelding, Allen Goodman, Melanie Guldi, Melanie Guldi, Simone Häckl, Janet Hankin, Scott Andrew Imberman, Joanna Lahey, Joan Llull, Hani Mansour, Jaakko Meriläinen, Tove Mortlund, Martin Nybom, Stephen D. O'Connell, Rupert Sausgruber, Amy Schwartz, Jan Stuhler, Petra Thiemann, Roel van Veldhuizen, Marianne Wanamaker, and Maria Zhu. 2021. "The Global COVID-19 Student Survey: First Wave Results." *IZA Working Paper*, , (14419).

Jehi, Tony, Raihan Khan, Hildemar Dos Santos, and Nader Majzoub. 2022. "Effect of COVID-19 Outbreak on Anxiety among Students of Higher Education; a Review of Literature." *Current Psychology*.

Johnson, Eric, and C. Lockwood Reynolds. 2013. "The Effect of Household Hospitalizations on the Educational Attainment of Youth." *Economics of Education Review*, 37: 165–182.

Kofoed, Michael, Lucas Gebhart, Dallas Gilmore, and Ryan Moschitto. 2021. "Zooming to Class?: Experimental Evidence on College Students' Online Learning During Covid-19." *IZA Discussion Paper*, , (14356).

Landry, Megan, Sydney Bornstein, Nitasha Nagaraj, Gary A Sardon Jr, Amanda Castel, Amita Vyas, Karen McDonnell, Mira Agneshwar, Alyson Wilkinson, and Lynn Goldman. 2023. "Postacute Sequelae of SARS-CoV-2 in University Setting." *Emerging Infectious Diseases*, 29(3).

Lavecchia, A.M., H. Liu, and P. Oreopoulos. 2016. "Chapter 1 - Behavioral Economics of Education: Progress and Possibilities." In . Vol. 5 of *Handbook of the Economics* *of Education*, , ed. Eric A. Hanushek, Stephen Machin and Ludger Woessmann, 1–74. Elsevier.

Lederer, Alyssa M., Mary T. Hoban, Sarah K. Lipson, Sasha Zhou, and Daniel Eisenberg. 2021. "More Than Inconvenienced: The Unique Needs of U.S. College Students During the COVID-19 Pandemic." *Health Education & Behavior*, 48(1): 14–19. PMID: 33131325.

Lewis, Megan; Sanchez, Ruth; Auerbach Sarah; Nam Dolly; Lanier Brennan; Taylor Jeffrey; Jaso Cynthia; Nolan Kate; Jacobs Elizabeth; Parker F. Hudson, and Darlene Bhavnani. 2020. "COVID-19 Outbreak Among College Students After a Spring Break Trip to Mexico — Austin, Texas, March 26–April 5, 2020." *Morbidity Mortality Weekly Report*, 69(26): 830–835.

Logel, Christine, Philip Oreopoulos, and Uros Petronijevic. 2021. "Experiences and Coping Strategies of College Students During the COVID-19 Pandemic." *National Bureau of Economic Research.*

Macsween, Karen F., Craig D. Higgins, Karen A. McAulay, Hilary Williams, Nadine Harrison, Anthony J. Swerdlow, and Dorothy H. Crawford. 2010. "Infectious Mononucleosis in University Students in the United Kingdom: Evaluation of the Clinical Features and Consequences of the Disease." *Clinical Infectious Diseases*, 50(5): 699–706.

Magee, Skylar; Imad, Mays. 2022. "Long COVID Is Here to Stay." *https://www.insidehighered.com/views/2022/05/31/colleges-must-support-students-long-covid-opinion?v2*, Accessed: 06-27-2022.

McDaniel, Christy, Catherine Suffern, Jenna Joo, and Rayane Alamuddin. 2020. "Student and Faculty Experiences with Emergency Remote Learning in Spring 2020: Insights from a Small Exploratory Study." Ithaka S+R.

Odriozola-González, Paula, Álvaro Planchuelo-Gómez, María Jesús Irurtia, and Rodrigo de Luis-García. 2020. "Psychological Effects of the COVID-19 Outbreak and Lockdown among Students and Workers of a Spanish University." *Psychiatry Research*, 290: 113108.

Padrón, Isabel, Isabel Fraga, Lucía Vieitez, Carlos Montes, and Estrella Romero. 2021. "A Study on the Psychological Wound of COVID-19 in University Students." *Frontiers in Psychology*, 12.

Premraj, Lavienraj, Nivedha V. Kannapadi, Jack Briggs, Stella M. Seal, Denise Battaglini, Jonathon Fanning, Jacky Suen, Chiara Robba, John Fraser, and Sung-Min Cho. 2022. "Mid and Long-term Neurological and Neuropsychiatric Manifestations of Post-COVID-19 Syndrome: A Meta-analysis." *Journal of the Neurological Sciences*, 434: 120162.

Prowse, Rebecca, Frances Sherratt, Alfonso Abizaid, Robert L. Gabrys, Kim G. C. Hellemans, Zachary R. Patterson, and Robyn J. McQuaid. 2021. "Coping With the COVID-19 Pandemic: Examining Gender Differences in Stress and Mental Health Among University Students." *Frontiers in Psychiatry*, 12.

Rodríguez-Planas, Núria. 2022*a*. "COVID-19, College Academic Performance, and the Flexible Grading Policy: A Longitudinal Analysis." *Journal of Public Economics*, 104606.

Rodríguez-Planas, Núria. 2022*b*. "Hitting Where It Hurts Most: COVID-19 and Low-income Urban College Students." *Economics of Education Review*, 87: 102233.

Rudenstine, Sasha, Kat McNeal, Talia Schulder, Catherine K. Ettman, Michelle Hernandez, Kseniia Gvozdieva, and Sandro Galea. 2021. "Depression and Anxiety During the COVID-19 Pandemic in an Urban, Low-Income Public University Sample." *Journal of Traumatic Stress*, 34(1): 12–22.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics*, 225(2): 175–199. Themed Issue: Treatment Effect 1.

The New York Times. 2021. "Tracking Coronavirus Cases at U.S. Colleges and Universities." Accessed: 7-14-2022.

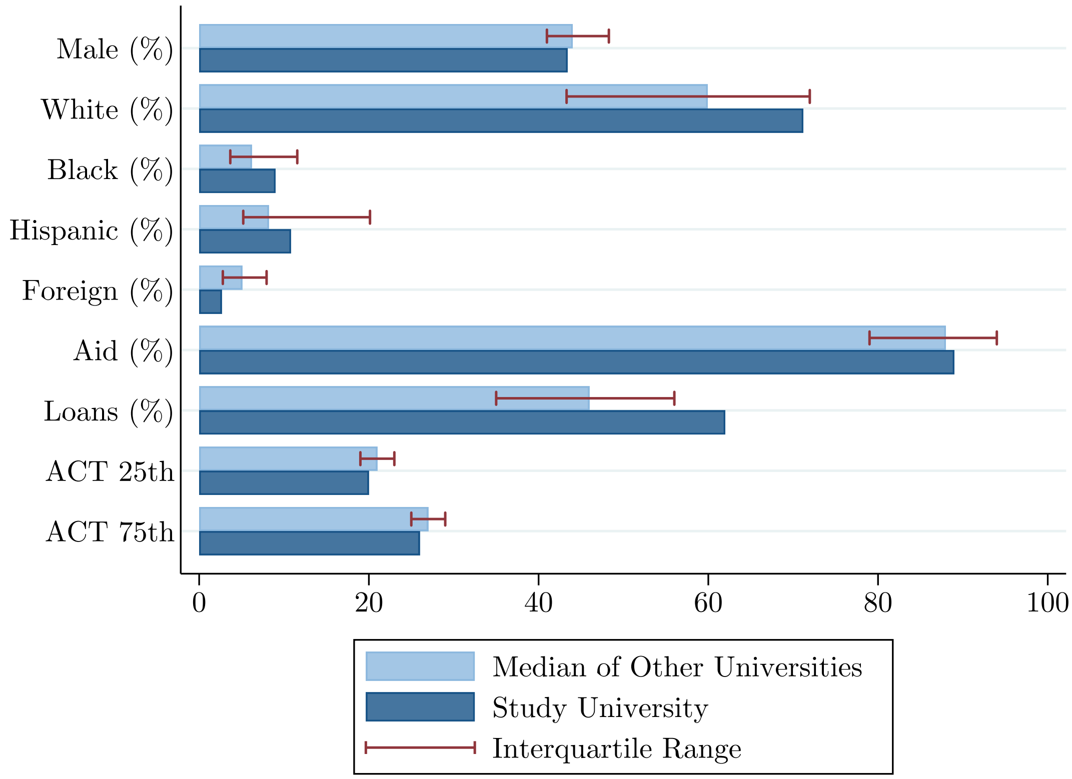
Wathelet, Marielle, Stéphane Duhem, Guillaume Vaiva, Thierry Baubet, Enguerrand Habran, Emilie Veerapa, Christophe Debien, Sylvie Molenda,

Mathilde Horn, Pierre Grandgenèvre, et al. 2020. "Factors Associated with Mental Health Disorders among University Students in France Confined during the Covid-19 Pandemic." *JAMA network open*, 3(10): e2025591–e2025591.

Williams-Harmon, Yolonda J., Leonard A. Jason, and Ben Z. Katz. 2016. "Incidence of Infectious Mononucleosis in Universities and U.S. Military Settings." *Journal of Diagnostic Techniques and Biomedical Analysis*, 5.

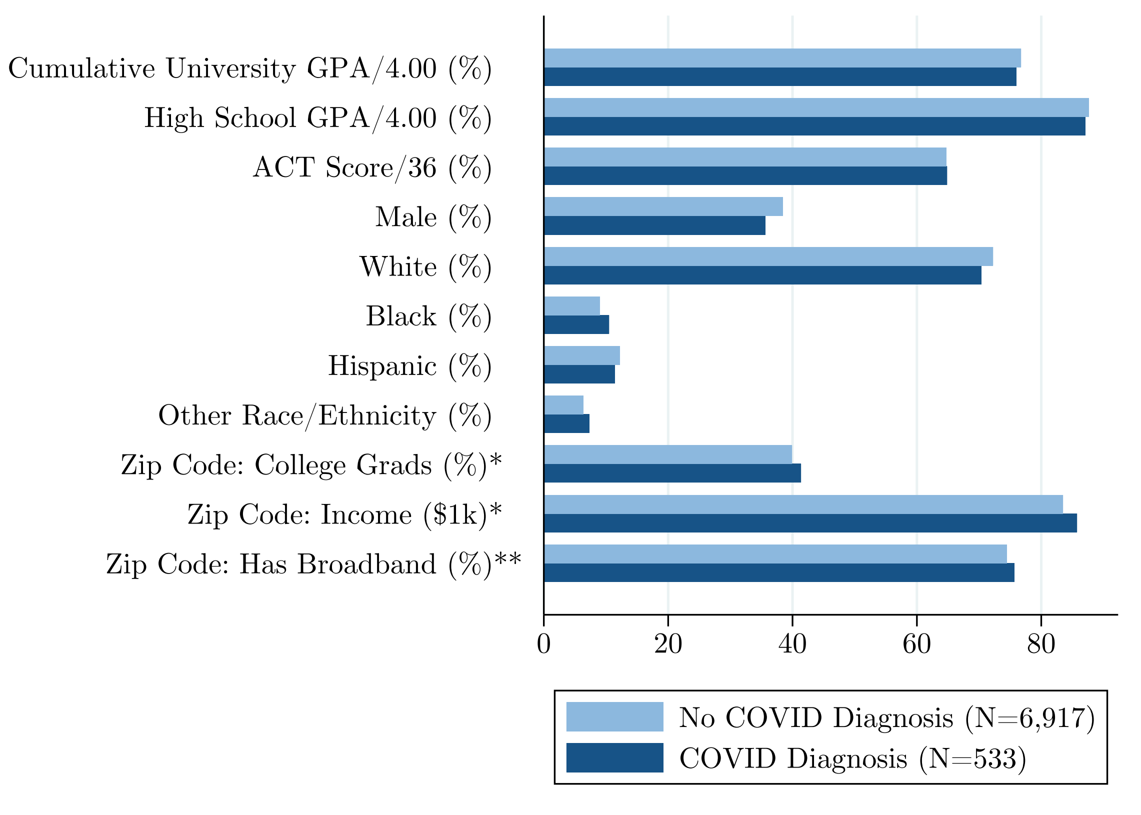
Zhao, Qingyuan, and Daniel Percival. 2017. "Entropy Balancing is Doubly Robust." *Journal of Causal Inference*, 5(1).

Figure 1. IPEDS University Comparison



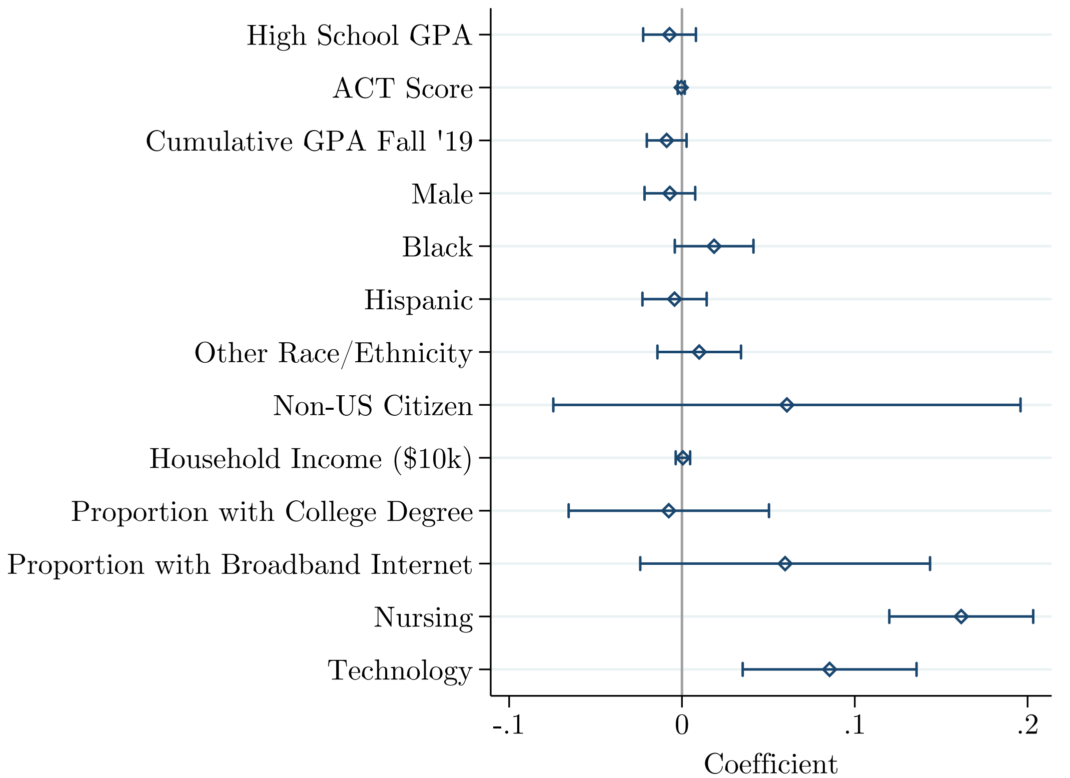
Note: The figure compares the University to 229 other public universities with an enrollment over 10,000 students in 2018.

Figure 2. Estimation Sample Summary Stats



Note: The figure compares undergraduate university students from the Fall 2017, 2018, and 2019 cohorts enrolled during the Spring 2021 semester. The statistic are presented separately based on the whether the student was diagnosed with COVID-19 (either during the Fall 2020 or Spring 2021 semester). Cumulative GPA is reported for the Fall 2019 semester, and the other metrics are time invariant.

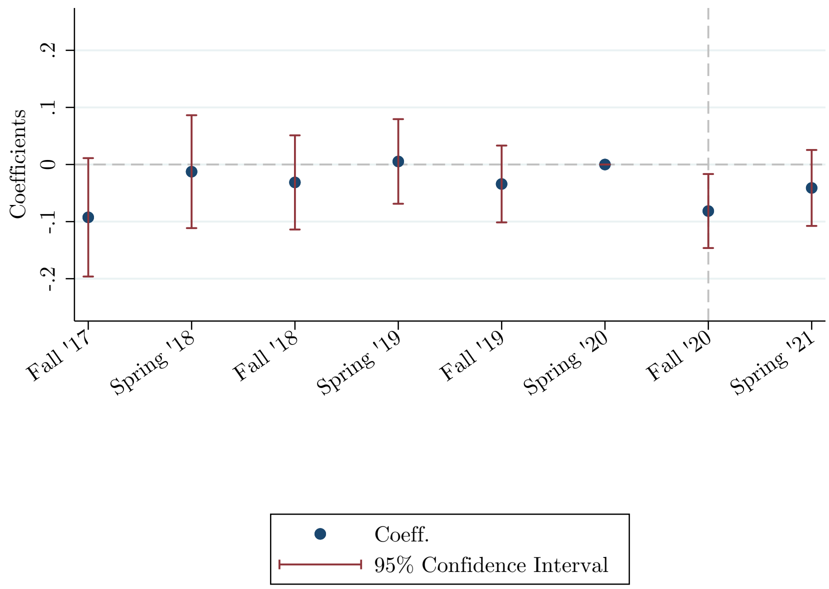
Figure 3. Correlates with COVID, Dependent Variable: COVID-19 in Fall '20 or Spring '21



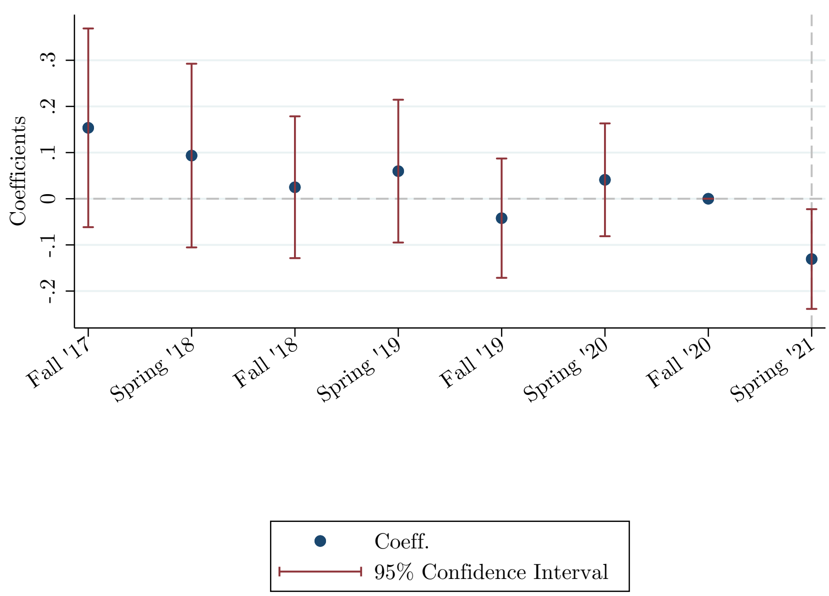
Note: The sample includes 7,450 undergraduate students from the Fall 2017, 2018, and 2019 cohorts enrolled during the Spring 2021 semester. Controls for department of major fixed effects were included but not reported with the exception of nursing and technology majors (omitted major is accounting). The bars represent 95% confidence intervals.

Figure 4. Event Study, Dependent Variable: Term GPA

(a) Diagnosed Fall 2020

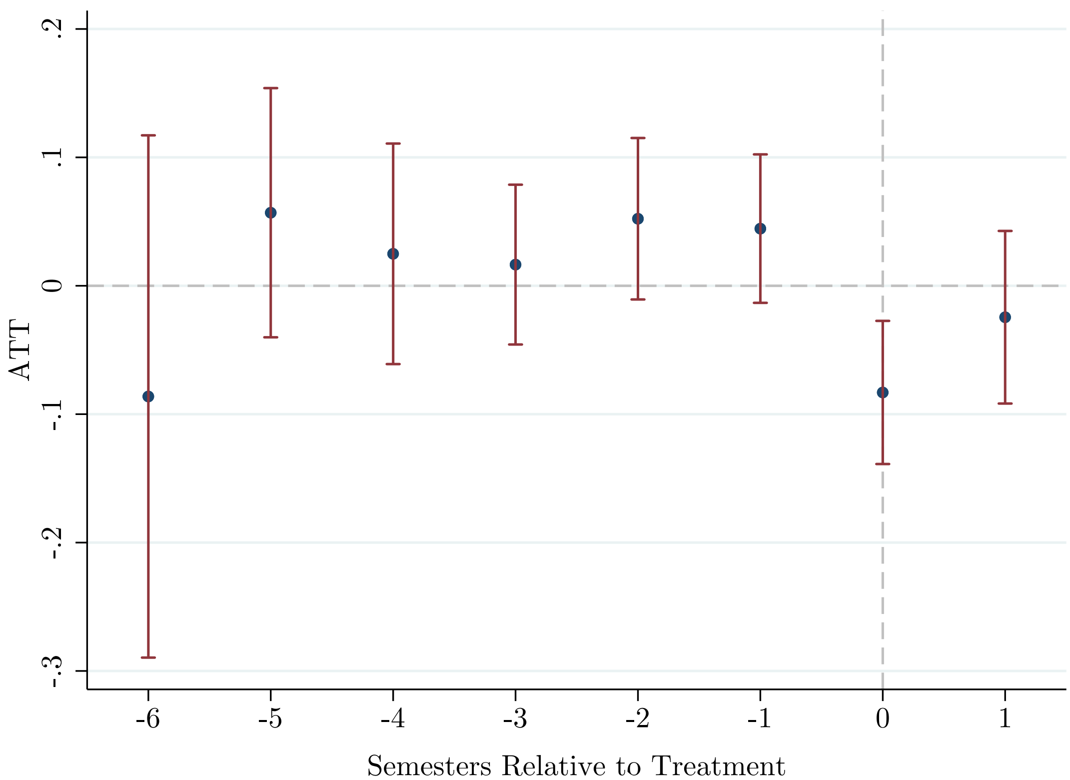


(b) Diagnosed Spring 2021



Note: The sample includes FTIC students from the Fall 2017, Fall 2018, and Fall 2019 cohorts who never had a COVID diagnosis or only had a COVID diagnosis is the semester considered. Individual, date, and department of major fixed effects, along with variables for bereavement, family emergency, other illness, major course share, and credit hours, were included but not reported here. The bars represent 95% confidence intervals, and the standard errors are clustered at the student level.

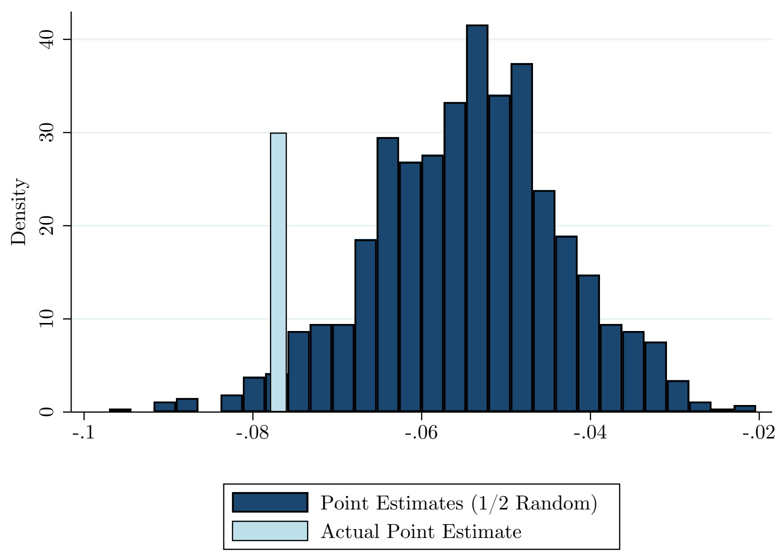
Figure 5. Staggered DID Estimation
Callaway and Sant'Anna (2021)



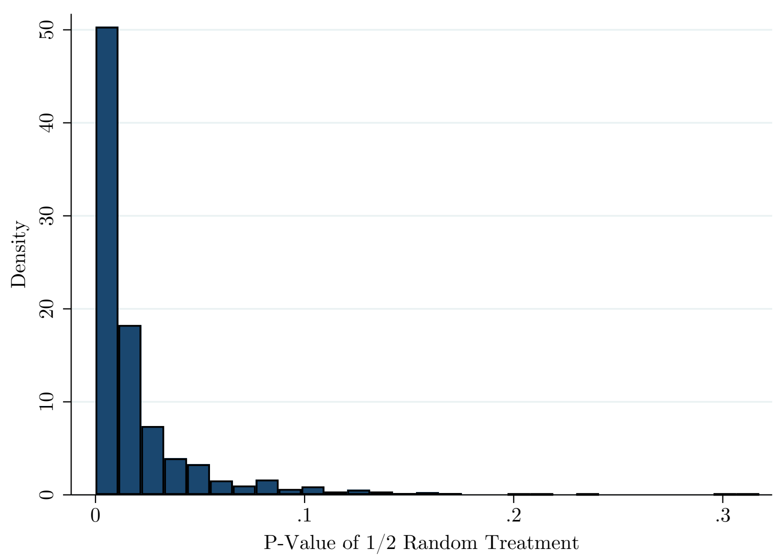
Note: The figure represents an event study approach following Callaway and Sant'Anna (2021). The sample includes 48,956 observations for 9,571 unique undergraduate students. Individual, date, and department of major fixed effects were included but not reported here. The bars represent 95% confidence intervals.

Figure 6. Simulations of the Influence of COVID-19 on Term GPA Assuming Substantial Underreporting

(a) Distribution of point estimates

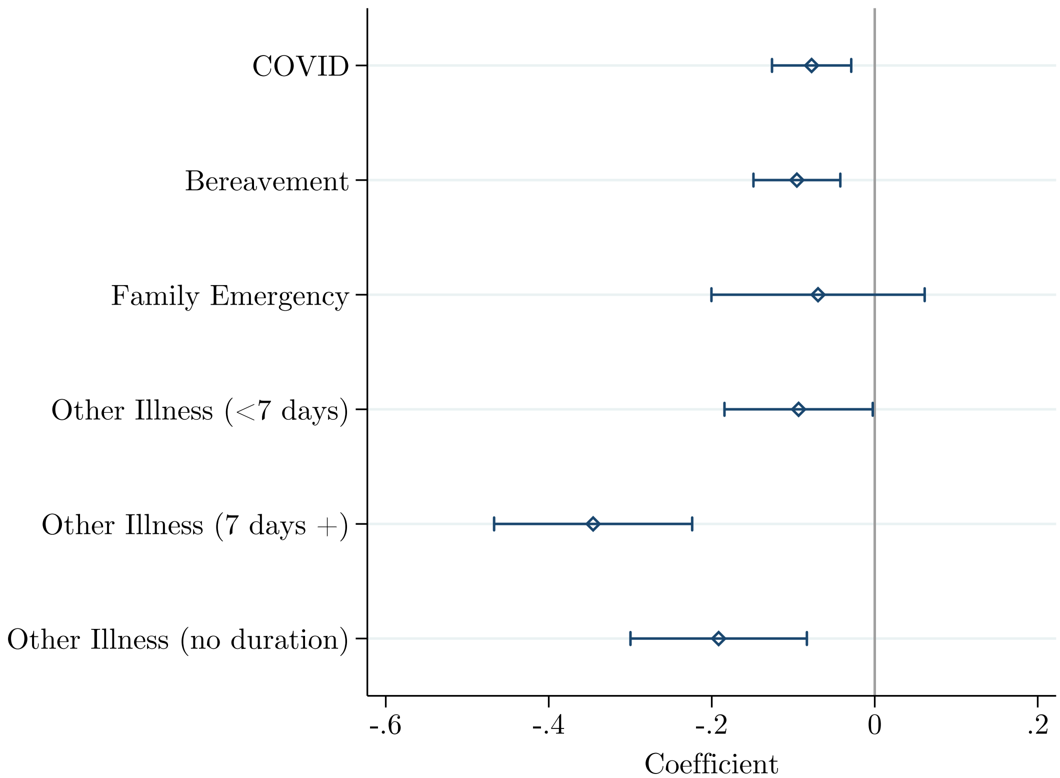


(b) Distribution of P-values



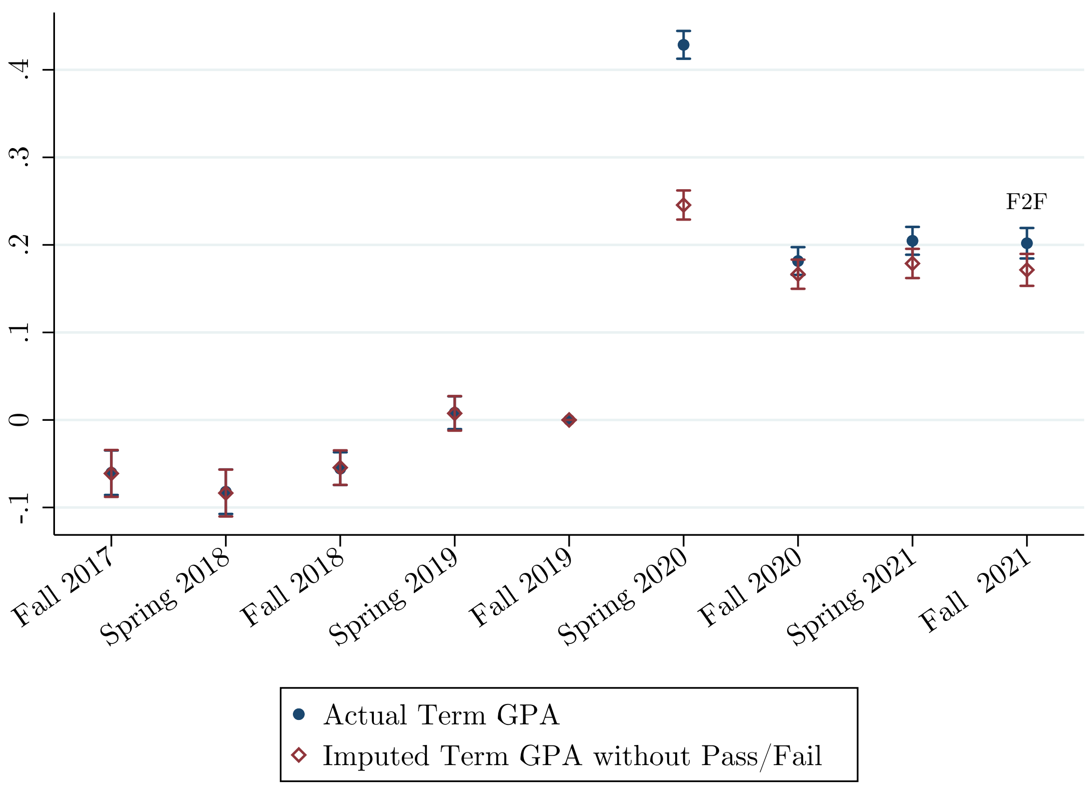
Note: For these simulations, we randomly assigned 546 students who did *not* report having COVID-19—along with the 546 students who actually reported a COVID-19 diagnosis—to our "treated" group and estimate a two-way fixed effect model given in Equation (2) of the influence of the treatment on term GPA. The distributions are based on 1,000 simulations.

Figure 7. Benchmarking the Influence of COVID-19 on Term GPA



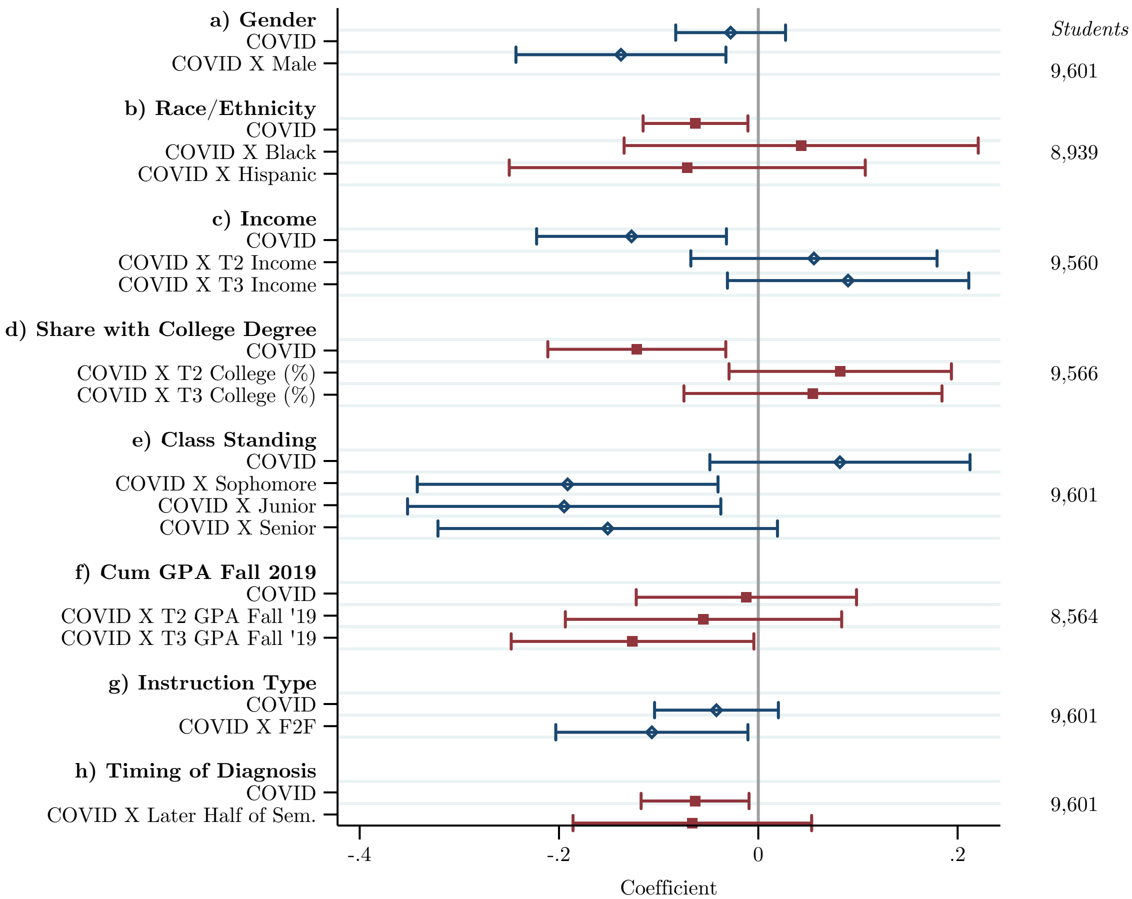
Note: The sample includes 49,169 observations for 9,601 unique undergraduate students. Individual, date, and department of major fixed effects, and controls for share of credit in a students major, and credit hours, were included but not reported here. The bars represent 95% confidence intervals, and the standard errors are clustered at the student level.

Figure 8. Semester Effects, Dependent Variable: Term GPA



Note: Student fixed effects were included but not reported here. The bars represent 95% confidence intervals, and the standard errors are clustered at the student level.

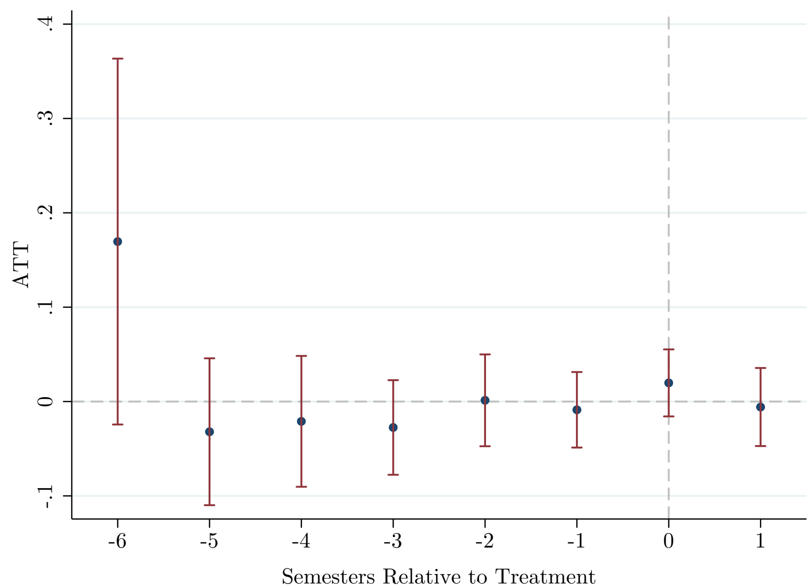
Figure 9. Heterogeneity Analysis, Influence of COVID-19 on Term GPA



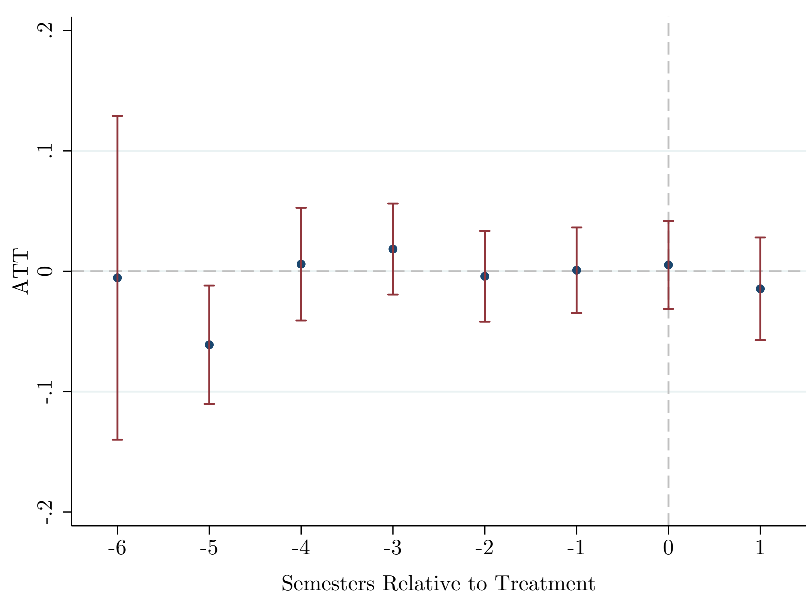
Note: Figure reports results from multiple models, iterating through different interactions of COVID-19 absence and other variables. Individual, date, and department of major fixed effects along with other absences (e.g., bereavement, other illness), share of major course credits, and credit hours were included but not reported here. For the class standing specification we additionally include indicators for class standing and for the F2F regression we similarly add a covariate for any F2F instruction. The bars represent 95% confidence intervals, and the standard errors are clustered at the student level.

Figure 10. Event Study, Alternative Educational Outcomes

(a) Dependent Variable: Below C grade in any class

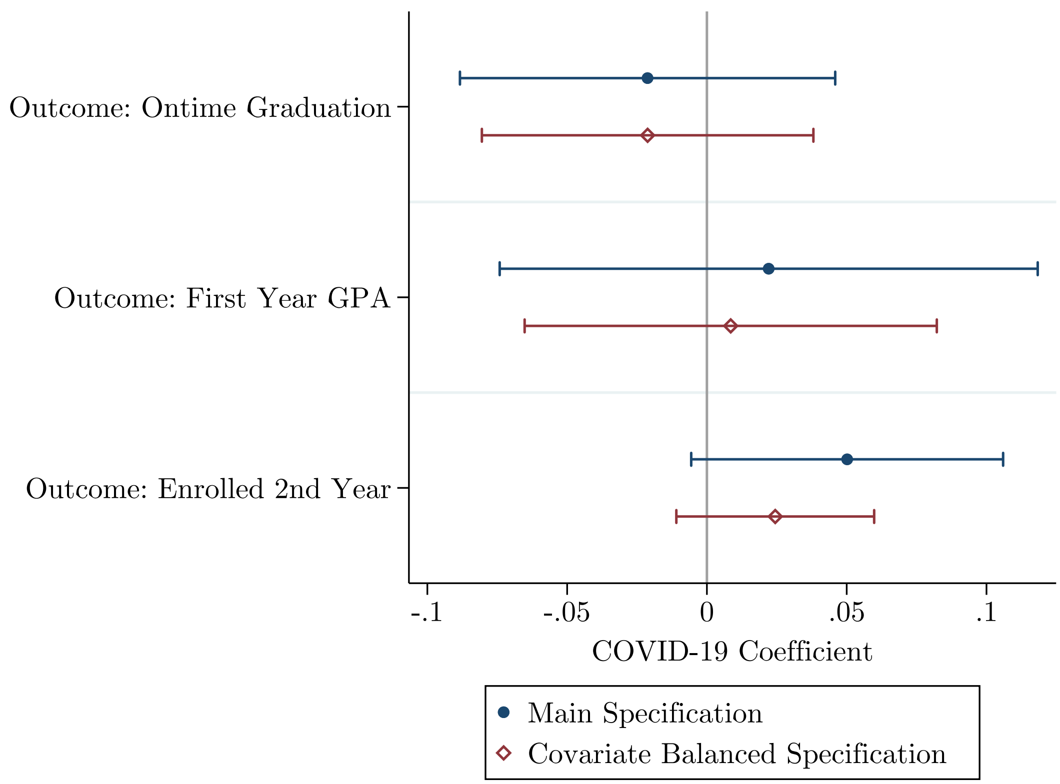


(b) Dependent Variable: Withdraw from any class



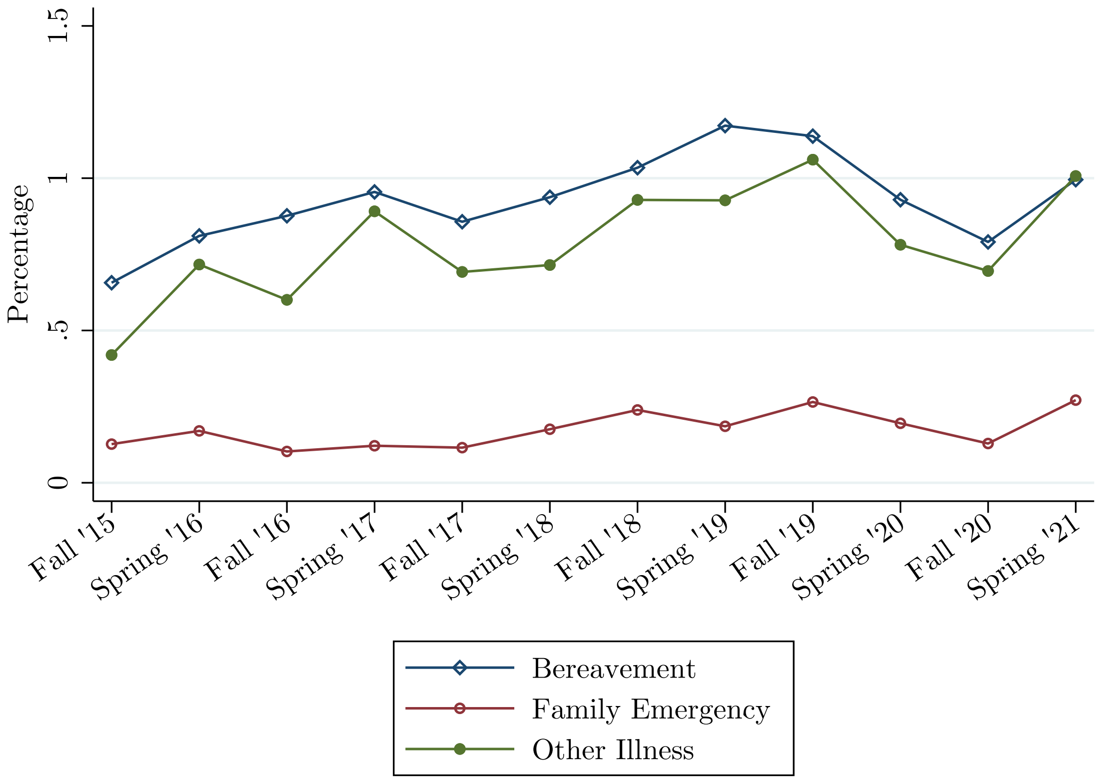
Note: Each figure represents an event study approach following Callaway and Sant'Anna (2021). The sample includes FTIC students from the Fall 2017, Fall 2018, and Fall 2019 cohorts and excludes students who contracted COVID-19 during both semesters. Individual, date, and department of major fixed effects along with variables for bereavement, family emergency, other illness, major course share, and credit hours, were included but not reported here. The bars represent 95% confidence intervals, and the standard errors are clustered at the student level.

Figure 11. Influence of COVID-19 on "On-Time" Graduation, First Year GPA, and Second Year Enrollment



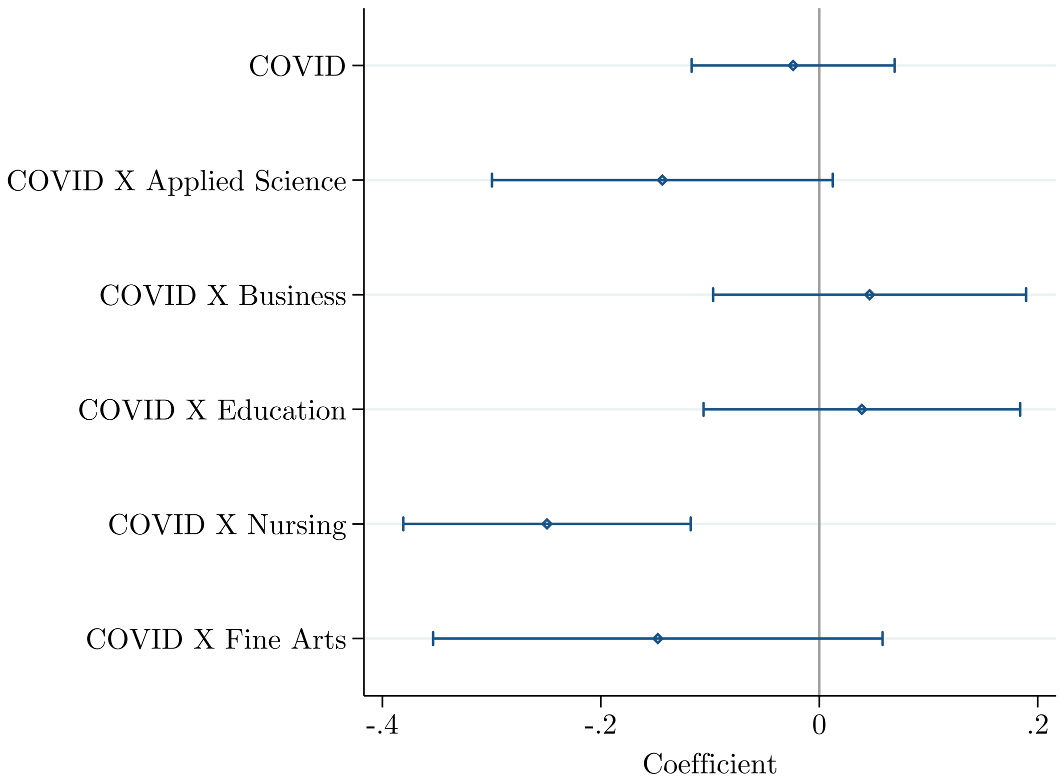
Note: Each specification was estimated using Equation 3. The regressions with on-time graduation as the dependent variable has a sample that includes 2,063 FTIC students from the Fall 2017 cohorts enrolled in the Spring 2020 semester. The samples for the next two specifications include FTIC students from the Fall 2020 cohorts. The bars represent 95% confidence intervals, and the standard errors are clustered at the student level. The covariate balanced specifications match the covariates using entropy weighting as a pre-processing step.

Figure A1. Absences over Time as a Percentage of Enrollment



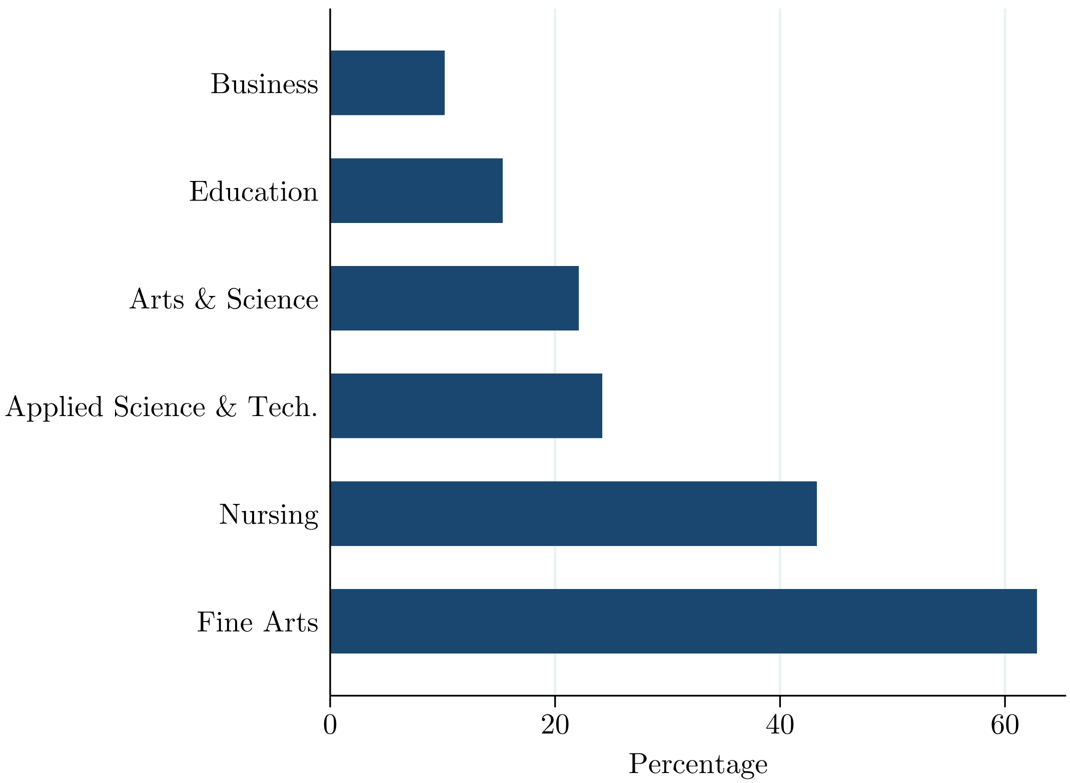
Note: The sample includes FTIC students from the Fall 2017, Fall 2018, and Fall 2019 cohorts.

Figure A2. Influence of COVID-19 Diagnosis on GPA: College Interactions



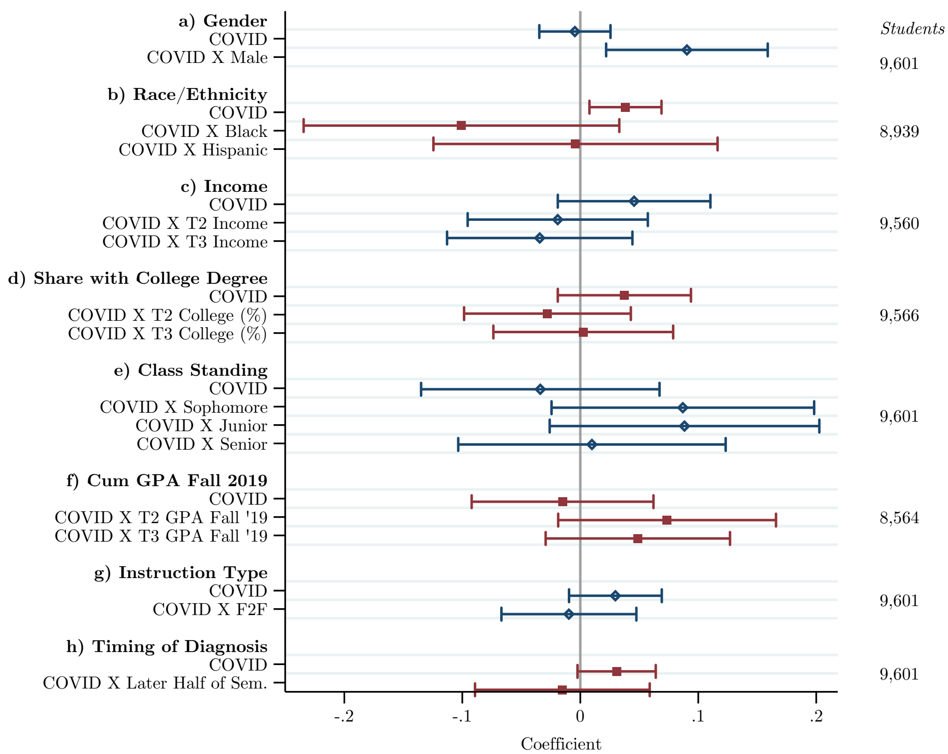
Note: There were 9,601 unique students and 49,169 total observations. Individual, date, and department of major fixed effects along with other absences (e.g., bereavement, other illness), share of major course credits, and credit hours were included but not reported here. The bars represent 95% confidence intervals, and the standard errors are clustered at the student level.

Figure A3. Share of Students with at least one Face-to-Face Course, Fall 2020



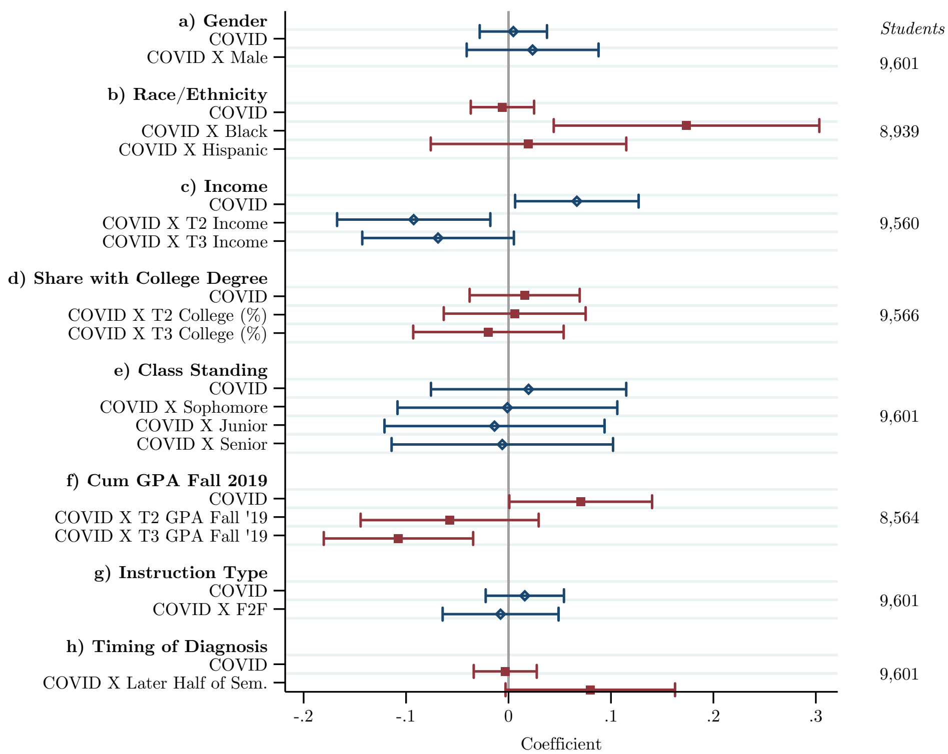
Note: The sample includes FTIC students from the Fall 2017, Fall 2018, and Fall 2019 cohorts.

Figure A4. Below C Heterogeneity



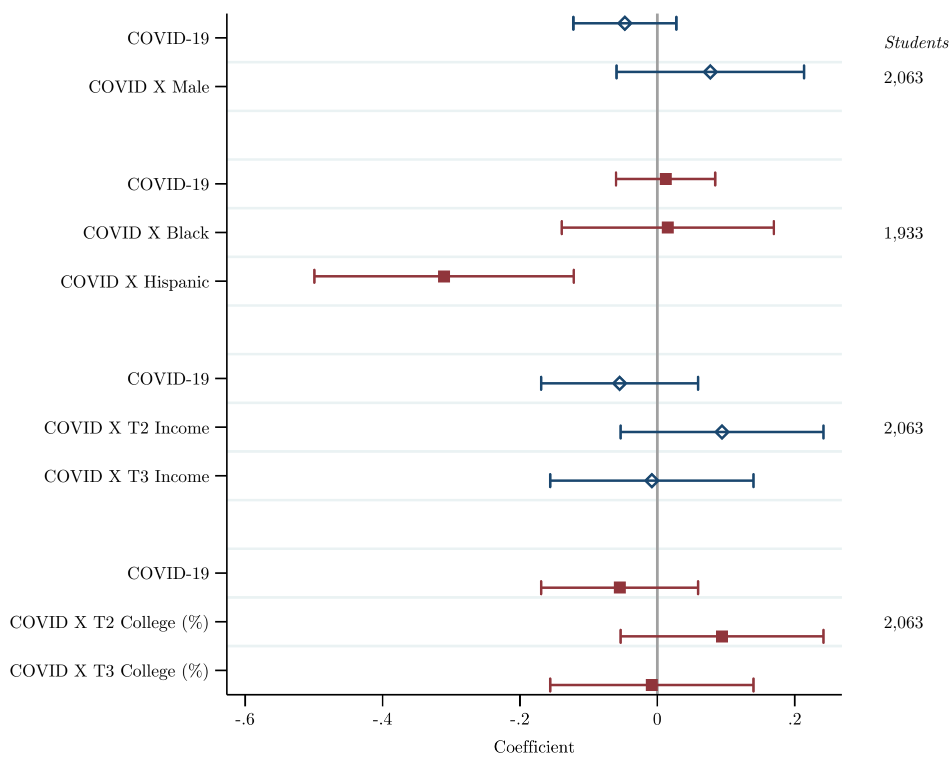
Note: Figure reports results from multiple models, iterating through different interactions of COVID-19 absence and other variables. Individual, date, and department of major fixed effects along with other absences (e.g., bereavement, other illness), share of major course credits, and credit hours were included but not reported here. The bars represent 95% confidence intervals, and the standard errors are clustered at the student level.

Figure A5. Any Withdrawal Heterogeneity



Note: Figure reports results from multiple models, iterating through different interactions of COVID-19 absence and other variables. Individual, date, and department of major fixed effects along with other absences (e.g., bereavement, other illness), share of major course credits, and credit hours were included but not reported here. The bars represent 95% confidence intervals, and the standard errors are clustered at the student level.

Figure A6. Heterogeneous Influence of COVID-19 on "On-Time" Graduation



Note: Figure reports results from multiple models, iterating through different interactions of COVID-19 absence and other variables. The sample includes FTIC students from the Fall 2017 cohorts enrolled in the Spring 2020 semester. The bars represent 95% confidence intervals, and the standard errors are clustered at the student level.