

# Discussion Paper Series

IZA DP No. 18445

March 2026

## Disconnected: The Unequal Impact of Online Learning on Minority Students

**Naomi Gershoni**

Ben-Gurion University of the Negev and IZA@LISER

**Miri Stryjan**

Aalto University School of Business

The IZA Discussion Paper Series (ISSN: 2365-9793) ("Series") is the primary platform for disseminating research produced within the framework of the IZA@LISER Network, an unincorporated international network of labour economists coordinated by the Luxembourg Institute of Socio-Economic Research (LISER). The Series is operated by LISER, a Luxembourg public establishment (établissement public) registered with the Luxembourg Business Registers under number J57, with its registered office at 11, Porte des Sciences, 4366 Esch-sur-Alzette, Grand Duchy of Luxembourg.

Any opinions expressed in this Series are solely those of the author(s). LISER accepts no responsibility or liability for the content of the contributions published herein. LISER adheres to the European Code of Conduct for Research Integrity. Contributions published in this Series present preliminary work intended to foster academic debate. They may be revised, are not definitive, and should be cited accordingly. Copyright remains with the author(s) unless otherwise indicated.



# Disconnected: The Unequal Impact of Online Learning on Minority Students\*

## Abstract

Online instruction holds the promise of expanding access to education for disadvantaged groups, yet it often deepens existing performance gaps. This study examines its impact on highstakes exam outcomes, focusing on mechanisms driving differential effects for minority students. We leverage rich administrative data on students from five consecutive cohorts in 31 Israeli vocational colleges and the abrupt transition to online instruction during the COVID-19 pandemic. A key advantage of our setting is that exams were held in person and graded centrally, ensuring comparability to pre-pandemic performance. Using a difference-in-differences design that compares outcomes within students and across cohorts, we find significant declines in both exam attendance and demonstrated knowledge following the switch to online instruction. These effects are not explained by local infection rates or childcare responsibilities, and are especially pronounced among Arabic-speaking minority students, regardless of their socioeconomic status. Drawing on variation in internet infrastructure, residential crowdedness, language of instruction, and prior academic performance, we show that poor internet access is a central mechanism underlying the adverse effects among minority students. For majority students, negative effects are concentrated among weaker students, suggesting widening disparities within the majority group. For minority students, the difference by prior achievement is not statistically significant, and if anything, point estimates are larger for higher achievers.

## JEL classification

I21, I23, I24

## Keywords

online instruction, education and inequality, minorities, vocational education, higher education, COVID-19

## Corresponding author

Miri Stryhjan

[miri.stryhjan@aalto.fi](mailto:miri.stryhjan@aalto.fi)

---

\* This research did not receive any specific grant from funding agencies in the public, commercial, or not-forprofit sectors.

---

# 1 Introduction

The COVID-19 pandemic, along with the resulting lockdowns and social distancing policies, has accelerated a shift toward online interactions both in education and in other domains such as public service delivery. As uncertainty grows due to conflicts and natural disasters exacerbated by climate change, this trend is likely to persist. However, replacing face-to-face interactions with online alternatives risks amplifying existing inequalities stemming from gaps in digital literacy and infrastructure access (Bacher-Hicks et al., 2021; OECD, 2021, 2024). In education, evaluations have shown that online setups often result in greater grade dispersion (Barrow et al., 2024), with disadvantaged and minority students disproportionately falling behind (see, e.g., Cacault et al., 2021). At the same time, online learning environments hold the promise of lowering costs and improving accessibility for disadvantaged groups (Bettinger et al., 2017; Novella et al., 2024). The pandemic forced students who typically would not opt for online instruction into this mode of learning, thereby offering a “natural experiment” for evaluating its impact.<sup>1</sup> However, in most settings, concurrent shifts in exam formats and instructors’ grading practices made it difficult to separate the specific effects of online instruction.

This paper exploits a unique opportunity to pinpoint the effect of online instruction on learning outcomes in a previously understudied educational domain — post-secondary vocational education — where a high share of students come from disadvantaged backgrounds and minority groups. In 2020, due to the pandemic, students in 31 technical colleges across Israel were compelled to switch to online learning at the start of the semester. However, by the end of the semester, they attended their mandatory state exams in-person. These exams were graded centrally and anonymously, minimizing any concerns about cheating or lenient grading, which have frequently been discussed in related literature (see, e.g., Bilen and Matros, 2021; Bird et al., 2022). This setting, combined with detailed administrative data on more than 15,000 students’ backgrounds, exam performance, and residential characteristics, including the local spread of the pandemic, the quality of internet infrastructure, and locality-level household crowdedness

---

<sup>1</sup>Bacher-Hicks and Goodman (2021) discuss the clear limitations of this “natural experiment” given the simultaneous changes to labor markets and health condition caused by the pandemic. Some of these confounders are less relevant in our setting while others are addressed in our analysis.

(average residents per room), enables us to isolate the effect of online instruction and to understand the determinants of its effectiveness across different groups of students.

The empirical approach leverages the fact that Israeli technical college students are required to pass several state-administered exams in core subjects at different stages of their studies. This allows us to compare each student's performance, in "late" versus "early" exams across pre- and post-pandemic cohorts using a *difference-in-differences* (DID) framework with individual-level fixed effects. For the affected cohort, the final semester coincided with the first lockdown and the shift to online learning, with their "late" exams taken after this semester, when social distancing policies had been relaxed. In contrast, the comparison cohorts completed their courses and took all of their exams before 2020, without experiencing any remote learning. Using supplementary data on local COVID-19 infection rates, socio-economic status (SES), and students' surveys, we rule out potential impacts of other pandemic-related factors besides online learning, including health-related factors and increased childcare responsibilities.

Our main finding is that exposure to online learning negatively affected exam attendance: the likelihood of *not* taking the exam increased by 22% for affected students compared to unaffected cohorts who took the same exams before 2020. Absence from these mandatory state exams carries severe academic and labor-market consequences, as it delays graduation and, in some fields, prevents students from obtaining the required license to practice their profession.<sup>2</sup>

We also assess exam performance and find that the likelihood of attaining at least 55% of the total score (the standard passing threshold) *decreased* on average by 12% relative to unaffected cohorts. A limitation of the latter finding is that the decrease in performance was not reflected in official grades since the original score was amended to compensate students for the hardships caused by the pandemic. While this policy may have somewhat altered students' incentives to exert effort during the exam, this concern is less relevant for those near the passing threshold.

Building on previous findings that low-SES students and ethnic minorities are more susceptible to the adverse effects of online learning (see, e.g., [Figlio et al., 2013](#); [Cacault et al.,](#)

---

<sup>2</sup>Focusing on exam attendance is also consistent with the broader literature, which commonly uses course completion as a measure of performance in studies of online learning ([Bettinger et al., 2017](#); [Xu and Jaggars, 2013](#)) and during the COVID-19 pandemic ([Bird et al., 2022](#); [Johnson et al., 2023](#)). More broadly, course completion and graduation are commonly used as key indicators of educational efficiency (see, e.g., [Banerjee and Duflo, 2014](#); [OECD, 2019a](#)).

2021), we focus on these aspects and their interaction. In our setting, the most relevant distinction is whether a student belongs to the Jewish majority or the Arabic-speaking minority, which comprises approximately 20% of the Israeli population and 31.4% of our study sample.<sup>3</sup> Examining the effects separately for Jewish and Arab students from high versus low socio-economic backgrounds reveals that the negative effects are consistently more pronounced among minority students, regardless of their SES. The effect on exam attendance for Arab students is almost twice as large in percentage points compared to Jewish students (though the difference is not statistically significant), and the differential effect on exam performance is even larger (and statistically significant at the 5% level).

To better understand what shapes these differential impacts, we utilize additional variation in elements that may be central to rendering online instruction effective. Examining heterogeneity by language of instruction (Hebrew or Arabic), internet infrastructure quality, and household crowdedness, we find that poor internet infrastructure is the primary driver of adverse effects for Arab students, whereas these factors are largely irrelevant for Jewish students. With regard to internet access, our data clearly show that it is highly correlated with minority status. While we do not detect larger effects on Jewish students with limited internet access, the sample size is too small to conclusively refute such effects.

When we assess heterogeneity by prior academic performance, we find that adverse effects among Jewish students are driven almost exclusively by low-achievers, suggesting that online learning further increases disparities even within the majority population. This pattern is consistent with previous findings (Bettinger et al., 2017). In contrast, prior grades do not play a distinct role among minority students, and, if anything, high-achievers within this population display larger losses. Overall, the significant initial achievement gap between Arab and Jewish students widened following the transition to online instruction.

Lastly, institutions and instructors may consider combining online and in-person teaching to exploit the advantages of both methods (see e.g., Bowen et al., 2014; Alpert et al., 2016). Based on a survey of department heads, we found that some departments offered in-person classes

---

<sup>3</sup>We use the term “Arab” to refer to this group of minority students who are citizens of Israel and may identify primarily as Palestinian, Arab, Druze, or Bedouin (Hawker, 2019).

towards the end of the semester, when the pandemic restrictions were alleviated. Exploiting this variation, we compare departments by the share of instruction delivered in person during the final weeks of the semester and find no evidence of differential impacts on students' outcomes for either minority or majority students. These results suggest that a partial return to in-person instruction did not offset the adverse effects of the transition to online learning.

Our findings contribute to the growing literature on the effects of online instruction in higher education. Experimental work on university students has found that learning outcomes are worse in online courses in economics and finance compared to face-to-face formats (Alpert et al., 2016; Kofoed et al., 2024), with larger performance decreases reported for low-ability and minority students (Figlio et al., 2013; Cacault et al., 2021; Kofoed et al., 2024). While these studies offer clean causal evidence, they focus on variation in online teaching intensity for just one course or for specific elements of a course. At the same time, observational studies aim to draw more general conclusions based on the growing prevalence of online instruction in higher education. However, their findings are limited to students who select into online courses. Using student, instructor, and course fixed effects models, Altindag et al. (2024) show that students consistently perform better in face-to-face courses than in online courses. Similarly, Bettinger et al. (2017) and Xu and Jaggars (2013) found that opting for online instruction, as determined by variation in distance to campus or course availability across semesters, led to lower grades and retention rates. An additional body of work evaluates the impact of the abrupt shift to online learning along with broader changes induced by the COVID-19 pandemic, including modified exam formats and grading, and challenging economic and health conditions. Several of these studies emphasize that outcomes of disadvantaged students, including ethnic minorities, were more negatively affected during this period (see, e.g., Bird et al., 2022; Bulman and Fairlie, 2022; Rodríguez-Planas, 2022a,b).

Our framework and data possess several advantages. First, the unique timing of the spring semester in Israel provided an almost ideal setting for this study — all instruction was moved online for an entire semester, whereas exams were taken in-person during a period of very low infection rates and mild restrictions. This timeline enables us to analyze the impact of full online instruction on students who would not otherwise select this mode of instruction while

avoiding the added effects of changes in exam setup. Moreover, exam content and grading are standardized and, thus, do not depend on college directives or on the specific student-instructor relationship. Second, we focus on students in their last semester, taking high-stakes exams that are mandatory to obtain their diploma or license to practice the profession. Compared to other settings where students could repeat courses without any delay to graduation, failing to take these mandatory exams on time necessarily delays the start of their professional careers, leading to income losses.<sup>4</sup> Third, the scale and richness of our administrative data allows us to rule out other effects of the pandemic and to explore different mechanisms, specifically focusing on minority students.

Finally, this paper provides evidence on the unique setting of technical colleges, thereby contributing to the active discussion on improving educational content and outcomes in vocational post-secondary education — an important policy goal (see, e.g., [Malamud and Pop-Eleches, 2010](#); [OECD, 2016](#); [Hall, 2016](#); [Hanushek et al., 2017](#); [Nedelkoska and Quintini, 2018](#); [OECD, 2019b](#)). As studies point out, remote learning may be particularly challenging in vocational studies due to their practical nature ([Xu and Jaggars, 2014](#); [Asgari et al., 2021](#)). Furthermore, graduation rates tend to be lower in these studies compared to academic education, highlighting the importance of focusing on student retention as an outcome ([Xu and Jaggars, 2013](#)).<sup>5</sup>

The remainder of the paper is organized as follows. Section 2 describes the institutional background of Israel's technical colleges and the events of the pandemic. Section 3 discusses potential mechanisms through which online instruction might affect learning outcomes. In Section 4 we describe the data sources and the sample used in the analysis, and Section 5 details the empirical strategy. The results are presented and discussed in Section 6, and Section 7 concludes.

---

<sup>4</sup>Labor market conditions for the fields of study that we focus on were largely unaffected by the pandemic either because the relevant industry was excluded from the restrictions (construction) or since the job could be performed remotely (software programming).

<sup>5</sup>In our setting the likelihood not to graduate because of one or more failed state exams is approximately 30% ([Gershoni and Stryjan, 2023](#)).

## 2 Background and Setting

### 2.1 Vocational Technical Colleges in Israel

We study post-secondary technological education tracks in Israel’s system of technical colleges which provide not only specific vocational skills but also broader knowledge, similarly to academic engineering programs. These tracks are regulated by the National Institute for Technical Training (NITT), and their costs are subsidized by the government. We focus on three of the largest fields of study: electrical engineering, civil engineering, and software programming, which together account for about one-third of a typical student cohort (approximately 3,000 students).<sup>6</sup> In addition to awarding official practical engineering diplomas, the first two fields grant licensure that enables graduates to certify construction and electrical plans for small residential buildings.

To graduate, students are required to complete about 2,180 academic hours of coursework, submit a final project, and pass the mandatory state exams in their field of study. Colleges offer both full-time programs, which take two years to complete (morning track), and part-time evening programs, which take three years (evening track). We therefore label cohorts by the final year in which students are expected to complete their coursework and exams.

The NITT administers and grades the state exams, which take place at the end of each semester. Fall semester exams are typically held from February to early March, while spring semester exams occur between July and mid-September.<sup>7</sup> Depending on their major, each student takes either three or four state exams over the course of their studies. The exact timing of each exam is predetermined by the NITT for each major and study track. Although students may retake or postpone exams, opportunities to do so are limited, typically requiring them to wait until the following semester. As a result, absence or failure directly delays graduation.

---

<sup>6</sup>The second-largest field — architecture — was excluded from our analysis because its exam schedule differs substantially from that of other fields, limiting our ability to assess the impact of online learning on exam outcomes.

<sup>7</sup>The academic year in Israel is based on the Jewish calendar, and the exact Gregorian dates of the semesters vary by year. The academic year typically begins in October, and the spring semester in early March.

## 2.2 The Israeli Technical Colleges System During Spring 2020

On March 14, 2020, Israel imposed its first nationwide lockdown, which included strict restrictions on gatherings and the closure of businesses and public spaces. Non-essential workers were required to work from home or were furloughed (Bodas et al., 2022). The lockdown ended in May 2020, though certain restrictions remained in place throughout the summer. On September 6, the government introduced the “traffic-light” plan, which classified localities daily by COVID-19 infection rates into red, orange, yellow, or green categories. This plan enabled differentiated restrictions based on local caseloads. However, as case numbers continued to rise, a second nationwide lockdown was implemented on September 18, 2020.

In response to the first lockdown, kindergartens, schools, and higher education institutions — including vocational colleges — fully transitioned to remote learning, and daycares were closed. After restrictions were lifted in May, in-person classes at higher education institutions were gradually and partially resumed, depending on the specific decisions of each institution. Nevertheless, most courses remained online through the end of the spring semester, and many of the spring semester exams were also administered online.<sup>8</sup>

However, the NITT maintained its standard exam protocol, including *in-person* exams, anonymized and centralized grading, and an unchanged passing threshold. Figure 1 presents mobility trends in Israel during 2020, demonstrating that mobility levels during the exam period (highlighted in gray) were comparable to pre-pandemic levels. This suggests that students, similar to other residents, likely felt safe participating in in-person activities.

The only formal change to the exam policy was that students were allowed to answer additional questions, with the possibility of earning up to 120 points instead of the usual 100. The NITT expanded the exam to maintain students’ usual level of choice among exam questions.<sup>9</sup> While the reported grade remained capped at 100,<sup>10</sup> this change effectively raised students’ grades.

Because our interest lies in the impact of online instruction on learning, rather than on

---

<sup>8</sup>Schools and daycares also continued to operate under substantial restrictions throughout this period.

<sup>9</sup>Each exam question is typically worth 20 points and consists of several sections.

<sup>10</sup>For example, a student scoring 90 out of 120 would receive a grade of 90, while one scoring 110 would be capped at 100.

officially reported grades, we use the percentage of points earned out of the total available — referred to as the *adjusted grade* — as our measure. This serves as a proxy for students’ actual level of learning, though it is admittedly imperfect. When asked to answer more questions, performance may be affected by fatigue, tighter time constraints, or strategic behavior, such as answering only enough questions to meet a target grade. In practice, these concerns appear minimal in our context, as stakes are high, time limits are not particularly strict, and, even with the expanded format, the number of questions remained manageable.

To further mitigate these concerns, our primary outcome measure is exam absence rather than grades. In addition, we use an indicator for whether a student attained the typical passing threshold — 55% of total exam points. We focus on this threshold, which corresponds to relatively low raw scores (55 and 66 under the standard and modified policies, respectively), since fatigue and strategic behavior are more likely to affect students aiming for higher scores.

### **3 Potential Mechanisms Linking Online Instruction and Learning Outcomes**

A growing body of research suggests that online instruction has heterogeneous effects on student learning. In this section, we discuss several mechanisms through which online instruction may influence learning outcomes, with particular attention to how these mechanisms are likely to affect minority and low-SES students, including Arab students in our context.

A fundamental reason for low educational outcomes is the suboptimal exertion of effort (Clark et al., 2020; Oreopoulos et al., 2022). This may result from flawed time perception and poor time management (Kahneman and Tversky, 1977; Francis-Smythe and Robertson, 1999), procrastination and present bias (O’Donoghue and Rabin, 1999), or inattention and distractibility (Ericson, 2017). These constraints are exacerbated by remote learning, which has been shown to negatively affect students’ attention, time management, and tendencies toward procrastination (De Paola et al., 2023; Kofoed et al., 2024). Without the structure of a physical classroom, students must exercise greater self-discipline and motivation to stay engaged (Xu

and Jaggars, 2014). This challenge is compounded by household distractions and the absence of routine, which can disproportionately affect students with less supportive or more crowded home environments. Furthermore, the practical focus of vocational studies presents additional challenges for online instruction compared to academic education (Asgari et al., 2021).

Students belonging to ethnic minorities are often more adversely affected by online learning (see, e.g., Figlio et al., 2013; Cacault et al., 2021; Kofoed et al., 2024), and several factors may account for this disparity. First, language barriers may be exacerbated in online settings. In our setting, students who are less proficient in Hebrew may find it more difficult to follow lectures or engage with instructors when communication is mediated through digital platforms. While this issue may affect some Jewish students — specifically those from immigrant backgrounds — it is especially relevant for Arab students, whose first language is Arabic and who, in most cases, attended Arabic-speaking high schools (Tehawkho et al., 2020; The Knesset Research and Information Center, 2024).

Second, digital literacy is significantly lower among the Arab population in Israel (Tehawkho et al., 2021). This affects students' ability to use online learning platforms, access course materials, and communicate digitally with instructors and peers. Third, disparities in internet access create direct barriers to participation. Reliable high-speed internet is not universally available, and in our setting, this issue is especially pronounced in peripheral areas where minority students are concentrated (The Knesset Research and Information Center, 2020).<sup>11</sup>

In addition to the direct effects on students' ability to participate in online classes, internet access may further reduce opportunities for interaction with peers and instructors outside class hours. Qualitative evidence highlights that students in online courses frequently experience a sense of isolation, reduced engagement with peers, and weaker connections to instructors (see e.g., Jaggars, 2014; Kofoed et al., 2024). This diminished interaction can be especially detrimental for those who depend on collaborative learning environments or lack academic support at home.

Lastly, the effectiveness of online learning may depend on the availability of a quiet space at

---

<sup>11</sup>It is plausible that limited internet access is correlated with lower levels of digital literacy. Students residing in areas with poor internet infrastructure may therefore face not only technical barriers to participating in virtual classes, but also challenges related to navigating and collaborating effectively in an online environment.

home and on household crowdedness. On average, minority students are more likely to live in crowded residences, which may limit their ability to attend online classes effectively, regardless of digital literacy or internet access. More broadly, minority status is often correlated with low SES and lower educational attainment which may also amplify the adverse effects of online instruction (Xu and Jaggars, 2014).

While most mechanisms suggest negative effects of online learning, there could also be potential benefits (see e.g., Jaggars, 2014). Notably, online instruction eliminates commuting time, which can be significant for students living far from campus. Other possible advantages include greater scheduling flexibility and reduced transportation costs, though these benefits are likely to be realized primarily by students with the digital resources and home environments conducive to effective remote learning.

Finally, pandemic-related stressors such as childcare responsibilities, illness, and economic hardship may have also affected students during this period, regardless of the switch to online platforms. Heightened stress due to financial uncertainty or fear of infection may limit students' cognitive bandwidth and ability to focus on academic tasks (Shah et al., 2015). More broadly, the pandemic has been associated with increased mental health challenges, including anxiety and depression, which have been linked to declines in cognitive function and academic performance among students (see e.g., Barbosa-Camacho et al., 2022). Evidence from Israel points to similar increases in mental health challenges, including among Arab students (e.g., Gilbar et al., 2022; Alfayumi-Zeadna et al., 2022). Educational constraints are also amplified for older students with commitments in addition to their studies (Bettinger et al., 2022), and students with young children are affected by daycare and school closures.

Leveraging our rich dataset, we conduct extensive heterogeneity analysis to shed light on the specific mechanisms underlying the observed outcomes for majority and minority students, and to distinguish the effects of online instruction from broader pandemic-related shocks.

## 4 Data and Sample

### 4.1 Administrative Data

The administrative database we analyze covers all students in the three study fields described above, spanning five consecutive cohorts from 2016 to 2020, defined by their final year of study.<sup>12</sup> The 2020 cohort experienced online instruction during their final semester, whereas the four preceding cohorts completed their coursework entirely in-person.

The dataset includes students' gender, age, ethnicity, and locality of residence, along with their study track, college ID, and the timing and grades of all state exams they took. To pass the exam, students must score at least 55 points out of 100, namely, answer correctly on 55% of the questions. As noted earlier, during the 2020 spring semester, the maximum possible score was raised to 120, while the passing threshold remained unchanged. As described below in Section 5, we convert the reported grades to percentages out of the maximum score (either 100 or 120) to ensure comparability with pre-pandemic semesters.

Our primary outcome is exam absence. To assess learning outcomes, we also construct an indicator for underperformance, defined as failing to reach the standard passing threshold of 55%. Because being absent from a mandatory exam is, in practice, equivalent to failing it, we classify absentees as underperforming as well. To explore heterogeneous treatment effects, we also use grades from “early” exams to classify students as high- or low-achievers based on the median grade.

Another key source of heterogeneity in our study is ethnicity. Ethnicity data come from the population registry but are missing for 12% of students.<sup>13</sup> For students with missing ethnicity data, we classify them as Arab or Jewish using information on the predominant ethnicity in their locality of residence and the prevalence of their first names within each ethnic group.<sup>14</sup> Using

---

<sup>12</sup>We exclude fewer than 5% of these students because they follow slightly different curricula or began their studies in the spring semester.

<sup>13</sup>The registry records religious affiliation, which in Israel largely corresponds to the two main population groups – Jewish and Arab (including Muslims, Druze, and Christians). The “Other” category includes smaller religious groups or individuals without a religious affiliation (less than 1.5% of our sample).

<sup>14</sup>A comparison with registry data for students with known ethnicity shows that our classification method is nearly 100% accurate. This high accuracy is largely due to the pronounced residential segregation between Jewish and Arab populations in Israel, as well as the distinctiveness of first names by ethnic group.

college ID, we also determine the language of instruction for each student, either Hebrew (90% of colleges) or Arabic (10%).

## 4.2 Locality-Level Characteristics

Leveraging students' locality of residence, we augment the data with four additional indices. First, we incorporate a socioeconomic Z-score, calculated by the Israeli Central Bureau of Statistics, based on various locality-level characteristics such as education levels, family size, labor force participation, and income.

Second, we construct a locality-level indicator of poor internet infrastructure using data on service coverage by Israel's second-largest internet provider, Hot Telecommunication Systems Ltd. In Israel's duopolistic internet market, the largest provider, Bezeq, is legally required to serve all households, but service quality is lower in localities where Bezeq faces no competition ([The Knesset Research and Information Center, 2020](#)). Clearly, the presence of competitive internet providers is not random but rather determined by profitability factors such as population size, density, and geographic distance from other localities. [Figure A3](#) shows the geographical variation in internet quality across Israeli localities. Although poor infrastructure is concentrated primarily in peripheral areas, there is substantial variation even among neighboring localities of comparable size.

Third, we construct a locality-level measure of pandemic severity using data on COVID-19 related hospitalizations between March and September 2020, our study period. We compute the per-capita monthly average of this variable for each locality. While the actual number of infections in each locality is unobservable, hospitalizations provide a reliable proxy for local disease burden and the associated risk environment in students' localities of residence. [Figure A4](#) shows the geographical dispersion of this measure across Israeli localities.<sup>15</sup>

Fourth, we use restricted data from the 2020-2021 Israeli Social Survey to construct a

---

<sup>15</sup>Hospitalization data is available for most localities in our sample, excluding very small ones. We also employ a second measure of infection rates based on the "traffic-light" classification, where a "red" designation indicates the highest caseload. This classification is available at a daily frequency starting in September 2020 and therefore does not overlap with the semester period, only with part of the exam period. Despite this limitation, the measure captures a broader set of pandemic-related conditions, as it also reflects variation in lockdown stringency. We construct this measure as the share of "red" days in each locality between September and December 2020.

locality-level measure of residential crowdedness. Respondents report the number of household residents per room, allowing us to classify localities based on the prevalence of overcrowded housing. We classify a locality as high in household crowdedness if at least 40% of households report more than one resident per room and at least 10% report two or more residents per room.<sup>16</sup>

### 4.3 Survey Data

We supplement the administrative records with data from an online survey conducted in collaboration with NITT during September 2020, immediately after the exam period. Students from the 2020 cohort were contacted via personalized SMS messages and completed the survey on their mobile phones.<sup>17</sup> The response rate, after two reminders, was 36%. The survey collected information on demographics, living arrangements, employment history before and during the pandemic, and interactions with peers and academic supervisors. We mainly use these data to describe changes in students' behavior and experiences during the 2020 spring semester and to gain some insights on potential barriers that limit the effectiveness of online learning.

Additionally, we use data from a survey distributed to the heads of the 67 departments in our study, of which 37 responded (55%). Administered in June 2020 (the last month of the semester), this survey covered a period when limited in-person teaching was permitted, allowing us to classify departments by the share of classes conducted in person during that time. The department heads were also asked about the challenges their students encountered with remote learning, as well as the broader impacts of COVID-19.

### 4.4 Descriptive Statistics

Table 1 presents summary statistics for the students in our sample based on the administrative data (Panel A) and for the subsample of students who answered our survey in the 2020 cohort

---

<sup>16</sup>These thresholds are motivated by international definitions of overcrowding, which emphasize the adequacy of living space relative to household size (OECD, 2025). While these definitions are more nuanced and account for differences in space needs by household composition, age, and gender, our data do not allow us to fully implement them. Nevertheless, localities with a high share of households exceeding one resident per room are likely to face similar constraints on private space, negatively affecting the home learning environment.

<sup>17</sup>Appendix section A.2 provides more details on the administration of this survey.

(Panel B). To account for potential selection into the survey, we also present the estimated population averages using probability weights and the corresponding standard errors, in Columns (3)-(4).<sup>18</sup> Reassuringly, the estimated population means are very similar to the survey sample averages.

The majority (93.2%) of students in our sample are male, they are on average older than a typical university student, and come from relatively low socioeconomic backgrounds, as indicated by the negative average socioeconomic Z-score. Accordingly, 46% study in evening tracks designated for working students. In addition, the table demonstrates that minority students are over-represented in these colleges, as Arab students make up 31.4% of our sample, compared to approximately 20% of the Israeli population. 13.5% of students attend colleges that specifically cater to Arab students and teach in Arabic.

From the survey we further learn that almost all the students have previous work experience, with approximately 40% having worked in their study field, and that 87% work during their studies. Despite this high share of working students and although many have spouses (37%) and children (28%), 51% are living with their parents (or other older relatives). During the pandemic, close to 15% of the students were furloughed or lost their job,<sup>19</sup> while we see smaller changes in living arrangements — about 5% of students moved in with parents or relatives and around 4% moved to a more independent setup. Overall, 21% of students reside in localities with high residential crowdedness.

## 5 Empirical Strategy

To identify the effect of the transition to online instruction on exam outcomes, we compare performance on exams scheduled before the final semester of studies (“early exams”) to those scheduled afterward (“late exams”). For students in the 2020 cohort, early exam results could not have been influenced by online instruction, whereas late ones, taken after the COVID-19 lockdown, were affected. To account for potential differences between early and late exams

---

<sup>18</sup>The probability of responding to the survey is predicted based on the students’ age, gender, SES, and specific department, from the administrative data.

<sup>19</sup>We group together both options since it was not clear at the time of the survey whether furloughed workers would return to their jobs.

unrelated to the switch to online learning, we include “pre-pandemic” cohorts from 2016-2019 in our analysis and implement a DID approach.<sup>20</sup> This approach allows us to control for both exam-subject and individual fixed effects, capturing unobserved, time-invariant characteristics of the exam and the student, respectively. Causal effects are identified under the assumptions that exam outcomes would have followed similar trends in the absence of the switch to online learning (common/parallel trends), and that students could not respond to this switch before it was implemented (no anticipation). Given the circumstances of the pandemic, it is highly unlikely that students or colleges could predict either the transition to online instruction or its precise timing in advance. We present several validity tests to assess the plausibility of these assumptions.

Table 2 reports the total number of mandatory exams for each major, along with the number of exams classified as early or late by major and track. Departments are expected to follow a standard exam schedule, but in practice, there is some variation in exam timing across departments, and students do not always adhere to the official schedule. Defining treatment based on the realized individual timing of exam taking would introduce selection, as students’ decisions to advance, postpone, or skip an exam may be correlated with unobserved individual, exam or period characteristics. This is particularly concerning in our setting, where the primary outcome, exam absence, is itself determined by the same individual decision. In addition, students who never take the exam would mechanically be classified as “late,” although this reflects dropout rather than delayed timing.

For this reason, we base our classification on the intended timing of exams. We classify an exam as “early” if more than 50% of students in the same track who took it for the first time did so before their final semester, and as “late” otherwise. Appendix Table A1 provides the full list of mandatory exams, the share of students who take each exam early (by major and track), and their classification based on these shares. Among exams classified as early, the share of students taking them early ranges from 54% to 99%, while for exams classified as late, the corresponding range is 1% to 32%.<sup>21</sup>

---

<sup>20</sup>Some students in these cohorts took exams after March 2020 because they were behind schedule.

<sup>21</sup>As evident in Table A1, for Exam 8 in the morning track and Exam 9 in the evening track, the share of students taking the exam late is only slightly above 50%, making these cases borderline under our classification rule. We

The exam outcomes we consider include an indicator for exam absence and an indicator for underperformance, defined as failing to reach the standard passing threshold, either by obtaining less than 55% of the maximum score or by being absent from the exam. As explained in section 2.2, we use percentages of the maximum score (adjusted grades) rather than reported grades, as they provide a more consistent proxy for students' actual learning levels, given the temporary increase in the maximum score during the summer of 2020.<sup>22</sup> As discussed earlier, although this increase could raise concerns about fatigue, time pressure, or strategic answering, these issues are likely minimal in our setting given the structure and timing of the exams and the nature of the outcomes analyzed. Appendix Figure A1 presents the distributions of reported and adjusted grades in late and early exams.<sup>23</sup>

Figures 2 (a)-(b) compare the distribution of adjusted grades in the 2020 cohort to that of previous cohorts, separating late and early exams. This raw-data representation of the DID framework indicates that the distribution of grades in “untreated” early exams remained stable across cohorts, lending support to the common trends assumption. In contrast, clear and substantial differences emerge in the outcomes of late (“treated”) exams, where the distribution of adjusted grades for the 2020 cohort is shifted to the left, indicating a decline in students' acquired knowledge.<sup>24</sup>

---

treat these exams as late throughout the analysis, but our results are robust to excluding them. Our results are also robust to using the actual timing of exam taking at the individual-level. However, this definition may introduce selection bias, as individual-level exam timing is endogenous in our setting.

<sup>22</sup>Results based on unadjusted, reported grades are presented in the appendix.

<sup>23</sup>Appendix Figure A1(a) compares the distributions of reported and adjusted grades in late exams. Mechanically, the adjustment reduces the variation in scores. In addition, since the reported grades are capped at 100, the 100 score bunches together all the scores between 100 and 120. In the adjusted grades distribution, these students are bunched at 83.33% which is the maximal adjusted grade (100 divided by 120). However, the inability to distinguish among raw scores above 100 does not affect our results as the grade outcome is defined relative to the 55% threshold. Due to imperfect compliance with our classification of early and late exams, some students took late exams before their final semester, and their grades remain unchanged in the adjustment procedure. Consequently, some adjusted grades exceed 83.33. Some lack of compliance is also evident in the corresponding distributions for early exams, shown in Appendix Figure A1(b). Still, as expected, the differences between unadjusted and adjusted grades in early exams are minimal.

<sup>24</sup>Appendix Figure A2(a)-(b) presents the corresponding figures for reported grades. In this case, the distribution for the late exams shifted to the right for the 2020 cohort, implying that students in this cohort improved their official state exam achievements relative to earlier cohorts. However, given the observed decline in their level of knowledge, this improvement is most likely a result of the temporary grading policy.

To formally test these initial findings, we next estimate the following DID specification:

$$(1) \quad y_{tic(i)d(i)} = \alpha_1 \text{late}_{ti} + \alpha_2 \text{late}_{ti} \times d2020_{c(i)} + \theta_{c(i)} + \omega_{d(i)} + \vartheta_t + \lambda' X_i + \psi_{tic(i)d(i)}$$

where  $y_{tic(i)d(i)}$  is the outcome of exam  $t$  for student  $i$  in cohort  $c$  and department  $d$ .  $\text{late}_{ti}$  indicates whether the exam is typically taken after the final semester based on the exam subject  $t$  and student  $i$ 's department and study track (morning or evening), and  $D2020_{c(i)}$  indicates that the student belongs to the 2020 cohort. Therefore,  $\alpha_2$  is our main coefficient of interest, interpreted as the average impact of online learning on exam outcomes. We first estimate this model with cohort, department and exam fixed effects, as well as individual-level controls including a quadratic in age, gender, ethnicity (Arab/Jewish), SES, and study track. In our main specification, we add individual fixed effects that further control for unobserved factors and identify within-student effects.<sup>25</sup> In this specification, the main coefficient of interest,  $\alpha_2$ , measures the average treatment effect *within* students.

To allow for potential correlation in the error term within students and across students within the same department, in our main results, we report two sets of clustered standard errors — at the department level and at the individual level. Department-level clustering accounts for a broader within-cluster correlation structure and therefore constitutes a more conservative approach.

To address potential concerns about differential pre-trends, we also estimate an event-study specification and conduct several placebo tests using pre-pandemic cohorts.

## 6 Results

As discussed above in Section 3, online instruction may have both negative and positive effects on students' learning and exam performance. These effects are likely to vary based on student characteristics and features of the learning environment. We first estimate average treatment effects, and then conduct a comprehensive heterogeneity analysis to shed light on key determinants and to rule out alternative explanations, such as pandemic-related factors.

---

<sup>25</sup>All time-invariant individual-level covariates, as well as cohort and department fixed effects, are absorbed by the individual fixed effects.

## 6.1 Average Effects

In Table 3, we estimate the effect of the switch to online instruction on our main outcome, exam absence, and on underperformance. Department-level clustered standard errors are reported in parentheses, and individual-level clustered standard errors in square brackets. The results in Column (1) show a substantial point estimate for increase in exam absence which is marginally significant when clustering errors at the department level (p-value of 0.101) and significant at the 1% level with student level clustering.

In our preferred specification that also adds individual fixed effects, this effect is more pronounced and statistically significant at conventional levels (Column (2)). The estimated 2.1 percentage point increase corresponds to a 22% higher likelihood of *not* taking the exam, relative to the mean absence rate in late exams for the pre-treatment cohorts. This effect is considerable, particularly given its direct implications for timely graduation.<sup>26</sup>

To verify that these effects are not driven by differential pre-trends for late exams, Figure 3 presents an event-study analysis. We estimate the same DID specifications as in Columns (1) and (2) but interact the indicator for a late exam with cohort dummies rather than pooling all pre-pandemic cohorts. The estimated differences in the pre-treatment period are close to zero and statistically insignificant for both specifications, indicating that trends in exam absence were parallel prior to the switch to online instruction.<sup>27</sup> Table A5 further supports this finding by presenting six placebo estimates based on pre-treatment cohorts. Each placebo estimate is obtained from a regression analogous to equation (1), in which the late-exam indicator is interacted with a cohort dummy from the pre-period. We conduct each placebo exercise including and excluding the actual treated cohort (2020). Across all specifications, the placebo estimates are substantially smaller than the estimated treatment effect and are not statistically significant. These results are consistent with the interpretation that the estimated effects reflect a distinct break associated with the transition to online instruction, rather than reflecting pre-existing trends or random variation.

---

<sup>26</sup>Appendix Table A4 reports the same specifications, displaying the coefficients on all control variables.

<sup>27</sup>In our setting, there is only a single post-treatment period, and therefore we cannot assess treatment dynamics in the post-event period, which is typically a central objective of event-study analyses.

Turning to exam performance, Columns (3)-(4) of Table 3 report a significant increase in the likelihood of scoring below the standard passing threshold. This outcome combines students who did not attend the exam and those who attended but underperformed, implying that their actual level of knowledge did not meet the one required pre-2020. Overall, underperformance increased by approximately 9 percentage points, or nearly 40% relative to the control mean. Subtracting the estimated 2 percentage point increase in absence suggests a 7 percentage point increase in the share of students who failed to demonstrate adequate knowledge when taking the exam — a 50% increase relative to the mean.

Columns (5)-(6) further examine performance at the intensive margin by estimating the effect on underperformance conditional on exam attendance. Restricting the sample based on an endogenous variable may introduce selection bias, and these estimates should therefore be interpreted with caution. This concern is partially mitigated when individual fixed effects are included, to the extent that selection into attendance reflects time-invariant individual characteristics, such as ability and motivation. Nonetheless, the estimated effects in this subsample are very similar to those in Columns (3)-(4), suggesting that students who did not attend the exam are, if anything, positively selected.<sup>28</sup> We return to this issue in Section 6.2, where we examine heterogeneous effects on absence by prior achievement.

Notably, these adverse effects were not reflected in students' reported grades or official failure rates due to the modification in the raw exam scoring. Appendix table A3 presents the corresponding estimates for officially failing the exam (scoring below 55 raw points, based on unadjusted reported grades), and find very small, statistically insignificant effects.

Taken together, our findings demonstrate clear adverse effects on students' learning and labor market skills. In the following sections, we continue to examine these effects and their

---

<sup>28</sup>As a complementary approach, we also assess the robustness of our findings to endogenous selection into late-semester exams using the changes-in-changes correction for attrition proposed by Ghanem et al. (2024). To implement this method, we collapse the exam-level data to the student level, treating early-semester exam outcomes as baseline outcomes and late-semester outcomes as follow-up outcomes. We then compare the 2020 cohort to earlier cohorts, interpreting absence in late exams as attrition. Under the identifying assumptions that, absent the switch to online learning, the relationship between early- and late-semester performance would have remained stable across cohorts and that students' relative standing is preserved, this procedure identifies a selection-adjusted average treatment effect. It also provides estimated average treatment effects on the treated (ATT) for non-attriters. Reassuringly, the resulting estimate intervals, although wide, are bounded away from zero. For example, the ATT is estimated between 0.05 and 0.28.

underlying mechanisms. All subsequent analyses report the more conservative standard errors, clustered at the department level. As shown above, department-level clustering yields larger standard errors across specifications, and inference remains unchanged when clustering at the student level.

## **6.2 Heterogeneous Effects and Evidence on Mechanisms**

### **6.2.1 Potential Effects of the COVID-19 Pandemic**

Our identification strategy relies on the abrupt shift to online learning caused by the COVID-19 pandemic, which brought about many concurrent changes. Three main factors that may have affected students' ability to exert effort and succeed in their studies include increased childcare responsibilities, adverse health effects, and financial distress. According to the survey-based descriptive statistics in Table 1, most students (72%) had no children, and 75% had no children under age 6, reducing the likelihood that our results are driven by school and daycare closures.

Approximately 4.8% of students reported moving in with parents or relatives during the semester which is suggestive of financial hardship. Yet, a similar share of students (3.7%) moved to live independently during the same period, resulting in a minor change to overall living arrangements. In addition, while 15% of students lost their job following the outburst of the pandemic, it is unlikely that they suffered substantial income declines due to the generous safety net provided by the Israeli government to workers who were laid off or furloughed at the time.

From the perspective of department heads, online learning was identified as the primary difficulty students faced during the semester. In the survey, 32 out of 37 respondents reported that students experienced learning difficulties. Among those who elaborated in an open-ended follow-up question, approximately 86% (18 out of 21) described challenges related to online learning, whereas only a small minority (3) mentioned financial hardship, health issues, or family obligations.<sup>29</sup>

---

<sup>29</sup>For example: *"The preparation takes place over Zoom, and it is difficult to assess the students' true level of understanding."* *"They feel less confident. Zoom classes are less effective. It's harder for them to concentrate."* *"It is harder to follow [students] on a personal level...and therefore we expect the exam results to be worse (students may think they are familiar with the content but will discover in the exam that this is really not the case)."*

To further assess whether our results are driven by other pandemic-related factors, we estimate heterogeneous effects along three dimensions: parenthood, local infection rates, and socioeconomic status. If the effects were due to broader pandemic disruptions rather than on-line instruction itself, we would expect more negative outcomes among students with children, in high-infection areas, or with lower financial resources. Although we cannot completely rule out the possibility that a general increase in mental health challenges contributed to the observed decline in performance, such challenges are likely correlated with the specific stressors observed in our data and examined below.

Since information on parenthood is only available for survey respondents, we use age as a proxy. We split the sample at the median age (25), above which 49.8% of respondents reported having children, compared to just 5.5% among younger students. Local infection rates are measured using COVID-19 hospitalization per capita. We classify students as living in high-infection areas if their locality's rate is above the median.<sup>30</sup> Finally, we use locality-level SES as a proxy for financial resources and estimate effects separately for students in the bottom SES tercile versus those in the upper two terciles.<sup>31</sup>

Table 4 presents the results of this analysis, focusing on exam absence. Columns (1)-(2) report nearly identical point estimates for students in high- and low-infection areas, respectively. The p-value for the difference between these estimates is 0.85 (reported at the bottom of the table). If anything, since baseline absence was lower among students in low-infection areas, the relative effect is larger for this group, contrary to what would be expected if reduced attendance was primarily driven by illness. The estimated effect for older students (Column (3)), who are more likely to have children, is not statistically different from that for younger students (Column (4)). Moreover, it is small and insignificant. Similar patterns are observed in the analysis of underperformance, reported in Appendix Table A6, suggesting that these two pandemic-related factors cannot explain our main findings.

In Columns (5)-(6) of Table 4 and Appendix Table A6, we split the sample by SES. Sub-

---

<sup>30</sup>We obtain similar results when we instead classify localities based on the median share of “red” days using the “traffic-light” score, which captures broader pandemic restrictions (see Section 4 above for details).

<sup>31</sup>We split at the tercile rather than the median because, in the minority student analysis that follows, the above-median sample becomes too small.

stantial negative effects are observed for both subsamples and their magnitudes cannot be distinguished statistically although the point estimates are slightly higher for the low SES group. In relative terms, these point estimates range between 17.4% for low SES students and 28.7% for high SES students.

The interpretation of these results is complicated by the fact that low SES communities were also more likely to be affected in terms of COVID-19 infection, as can be seen in Appendix figure A5, plotting local hospitalization rates against locality SES scores. Nevertheless, as neither low SES nor high infection rates seem to explain the negative effect on learning outcomes, this potential confounding factor is not a primary concern. Overall, the reported heterogeneity patterns indicate that the main effects that we estimate on learning outcomes are most likely attributed to online instruction.

### **6.2.2 Heterogeneous Effects by Minority Status**

Students' SES may also be correlated with other characteristics, particularly minority status. Prior studies have found both factors to be important in the context of online learning. Contrary to these findings, we do not detect substantial heterogeneity by SES in our data. In Table 5, we further examine exam absence for Arab and Jewish students separately, both in the full sample (Columns (1)-(2)) and within each SES category (Columns (3)-(6)). The point estimates for Arab students are larger, ranging between 2.1 and 2.6 percentage points, compared to 1.4 to 1.6 percentage points for Jewish students. Moreover, none of the estimated effects for Jewish students are statistically significant, whereas the effects for Arab students are statistically insignificant only for the high-SES group (Column (5)), where the sample size is very small. However, the difference between the estimated effects for Arab and Jewish students cannot be statistically distinguished from zero.

When we estimate the effects on underperformance (Appendix Table A7), the differences between Arab and Jewish students are larger and statistically significant (or marginally significant within SES categories). The likelihood of underperforming increased by approximately 13 percentage points for both low- and high-SES Arab students, representing relative increases of 33% and 45.7% from baseline rates of 39% and 30%, respectively. The corresponding in-

creases for Jewish students range from an insignificant 3-percentage-point estimate for low-SES students to a significant 6-percentage-point increase for high-SES students. These estimates are also smaller than those for Arab students in relative terms, despite Jewish students' lower baseline rates. Based on these results, we conclude that the adverse effects of online learning were consistently larger among minority students, regardless of their socioeconomic background.

### 6.2.3 What Determines the Effectiveness of Online Instruction?

To better understand the sources of the observed ethnic disparities, We leverage student-level variation in resources and academic background, alongside department-level variation in instructional practices. We focus on four dimensions of heterogeneity: the quality of internet infrastructure, the crowdedness in housing, the language of study (Hebrew or Arabic), and students' prior academic achievement. In addition, using the subsample of departments whose heads responded to our survey, we examine whether the effects vary with the extent of in-person instruction provided toward the end of the semester. We estimate heterogeneous effects along these dimensions separately for Arab and Jewish students, both because these factors are highly correlated with ethnicity and because the underlying mechanisms may differ across groups.

Interestingly, in our survey of department heads, several responses highlighted variation in students' ability to study effectively online. This heterogeneity was primarily attributed to differences in baseline academic ability. Internet access was also mentioned as a contributing factor, and some mentioned aspects of the student-instructor relationship, such as reduced monitoring and lower class attendance.<sup>32</sup>

**Internet Infrastructure** We begin by examining internet access, which is clearly a key constraint on students' ability to engage with online instruction. Beyond technical difficulties during classes, limited internet access may hinder students' ability to stay connected with peers and faculty, further exacerbating the challenges of remote learning.

Appendix Table A2 presents the share of students that report various hardships in the stu-

---

<sup>32</sup>Quotes include: "Not all students have the possibility to connect to Zoom, and their circumstances also differ from one another." "Taking classes via zoom does not suit all students, especially weaker students." and "All the studying over zoom significantly hurt the weak students and some of the 'average' students."

dent survey by the quality of their internet infrastructure, either standard (Column (1)) or poor (Column (2)). The differences and the corresponding standard errors are reported in column (3). Those with poor internet infrastructure were 85% more likely to report having no recent contact with fellow students and 52% more likely to report no contact with academic supervisors, relative to students with standard internet quality. They also reported more academic challenges, difficulty in completing tasks and less optimism about their own graduation prospects. In contrast, internet access is not correlated with students' self-reported ability to concentrate and work independently, nor with external constraints such as family obligations or COVID-related disruptions.

Columns (1)-(2) of Table 6 present the estimated effects of online learning on exam absence by internet quality. Panel A focuses on Arab students. Among this group, those residing in areas with poor internet access are 4.8 percentage points (24%) more likely to miss the exam — a sizable and statistically significant effect. In contrast, we find no effect for students with standard internet access. The difference between the two estimates is meaningful and statistically significant ( $p$ -value = 0.08). In Panel B, we turn our attention to Jewish students. For this group, the estimated effects are insignificant, and, if anything, the point estimate is larger for the standard internet subsample. However, given the strong correlation between minority status and poor infrastructure, the number of Jewish students with poor internet access is very small. As a result, our ability to draw meaningful conclusions for this group is limited. Columns (1)-(2) of Table A8 report the corresponding estimates for underperformance. While Arab students with poor internet access show larger increases in underperformance than those with standard access, the difference is less pronounced than for exam absence and is not statistically significant. These results are consistent with the interpretation that limited internet access weakens students' engagement to the extent that they do not attempt the exam at all, suggesting that it primarily influences the extensive margin of exam performance. As the vast majority of students with poor internet access are Arab, it may also be the main reason that online instruction was more harmful to Arab students.

**Household Crowdedness** Remote learning may be less effective in crowded living conditions, where limited private space can hinder students' ability to attend virtual classes and concentrate on academic work. In Columns (3) and (4) of Table 6, and the corresponding columns in Table A8, we split the sample by residential crowdedness in students' localities of residence.

We find no statistically significant differences in the estimated effects on absence or underperformance between students residing in high- and low-crowdedness localities, for neither Jewish nor Arab students. If anything, the point estimates are larger for students in low-crowdedness localities. For absence, the effect is statistically significant only for Arab students residing in low-crowdedness areas. Notably, high residential crowdedness is far more prevalent among Arab students (45.9%) than among Jewish students (9.6%). Although this disparity could potentially contribute to the larger adverse effects observed for Arab students, the absence of differential effects across crowdedness categories suggests that crowded living conditions do not drive the observed ethnic gap.

**Language Barriers** Language barriers may pose additional unique challenges to minority students. In our setting, the main language for Arab students is Arabic, yet 56% study in colleges where instruction is in Hebrew. Learning in one's native language likely facilitates comprehension and reduces cognitive load, so we might expect online learning to be more effective for Arab students studying in Arabic. We test this hypothesis for Arab students in Columns (5)-(6) of Table 6 (Panel A), and find that they are negatively affected regardless of the language of instruction. Although the estimated effect is larger for those studying in Hebrew (4.57 vs. 2.45 percentage points), the difference is not statistically significant. Moreover, when we examine underperformance in Appendix Table A8, we detect a statistically significant difference in the opposite direction of our hypothesis. The increase in underperformance is more than twice as large for Arab students in Arabic-instruction colleges compared to those in Hebrew-instruction ones (22.6 vs. 9.16 percentage points), and the relative gap is similarly large (51.7% vs. 28.6%). This suggests that language barriers alone do not explain the observed gaps between Arab and Jewish students.

**Prior Achievements** Finally, the heterogeneity analysis by prior academic performance reveals a striking pattern. To characterize students' previous achievement levels, we calculate their average grades in early exams, and split the sample by the median average grade in Columns (7)-(8).<sup>33</sup> Among Arab students, only those with *above*-median early-exam grades experience a significant decline in exam participation (3.7 percentage points), more than doubling their baseline absence rate. The effect among lower-performing Arab students is not statistically significant, but the difference in the effect between these two groups is also not significant. In contrast, among Jewish students the decline is concentrated among those with below-median prior grades (3.2 percentage points, or 34%), while the point estimate for higher-achieving students is close to zero. The difference between these two groups is marginally significant ( $p$ -value = 0.09).

**Combining Online and in-Person Classes** The last dimension we examine in Table 7 and Appendix Table A9 is the potential complementarity between in-person and online instruction, which may serve as an alternative to fully remote teaching. We exploit variation across departments in the share of in-person classes offered during the final weeks of the Spring 2020 semester, when COVID-19 restrictions were partially lifted. This measure is available only for departments whose heads responded to our survey, and thus, the sample size is smaller. We divide departments at the median share of in-person classes in June 2020. Low-intensity departments are those in which more than 40% of June classes remained online.

Columns (1)-(2) show that for Arab students, exam absence increased in both low- and high-intensity departments, although the effect is statistically significant only in the latter (4.1 percentage points, or 21%). The difference between the two groups is not statistically significant. These findings suggest that a partial return to in-person instruction did not offset the negative effects of a semester conducted largely online. This interpretation is reinforced by the absence of differential effects on underperformance (Columns (1)-(2), Appendix Table A9). Among

---

<sup>33</sup>Students who did not take any of the early exams are considered to have an average grade of zero and thus, categorized as "low". The average grade is calculated conditional on taking the exam, meaning that students with some missing exams have their average calculated over fewer grades which may lead to bias. Reassuringly, our results are robust to replacing missing grades with either zero or 55, namely the lower and upper bound of a "failing" grade.

Jewish students (Columns (3)-(4)), no effect is detected on absence under either intensity of in-person instruction, and the difference in the estimated effects on underperformance is likewise statistically insignificant.

However, the provision of in-person classes may itself be an endogenous departmental response. Instructors may have expanded in-person instruction precisely in departments where students were perceived to struggle more, so differences in estimated effects across intensity groups may partly reflect selection rather than the causal impact of combining in-person and online instruction.

Taken together, our findings point to the critical role of internet infrastructure in mediating the effects of online instruction, particularly for Arab students. Neither crowded living conditions nor language of instruction appear to drive the observed results, and a partial return to in-person teaching did not offset the learning losses incurred during the online semester. A notable finding concerns heterogeneity by prior academic achievement across minority and majority students. The concentration of negative effects among previously high-achieving Arab students is especially concerning, as it implies that online learning may have derailed students who had already overcome structural disadvantages. In contrast, for Jewish students, the effects of online learning are more modest and primarily concentrated among those who were already struggling academically.

## **7 Conclusion**

This paper examines the effect of online instruction on educational outcomes in post-secondary vocational education. The COVID-19-induced shift to online learning — coinciding with the final semester — provides a unique opportunity to study its impact in this overlooked domain. Using high-stakes, centrally graded, in-person exams and a difference-in-differences design comparing early (unaffected) and late (affected) exams within students and across cohorts, we find significant increases in exam absences and declines in learning outcomes. These effects are not explained by infection rates or childcare responsibilities and are evident across socioeconomic groups, but are especially pronounced among Arabic-speaking minority students. Drawing on

a comprehensive heterogeneity analysis, we attribute these declines largely to poor internet infrastructure. Unlike prior studies of online instruction, we find that high-performing minority students suffer the most, suggesting that online learning disrupted the progress of students who had previously overcome structural disadvantages.

While our study provides valuable insights into the challenges faced by minority students in online learning setups, certain limitations are important to acknowledge. Although our data include reasonable measures of pandemic-related pathways, such as infection rates, financial stress and childcare burden, allowing us to rule out these factors as primary drivers of the estimated effects, we do not observe any information that relates to students' mental health and digital literacy. Accordingly, we cannot exclude the possibility that these factors affect student's ability to adapt to online instruction and explain some of the heterogeneity observed by internet access and ethnicity.

Another potential caveat is that vocational programs may require greater adaptation to online formats than academic programs, which the abrupt shift in 2020 did not allow time for. However, our findings highlight that internet quality, rather than course design, was the primary constraint for minority students. Where internet access was adequate, as among most majority students, academically weaker students were most affected, likely due to challenges related to focus and attention.

Our results underscore the importance of examining online learning effects separately for minority groups, particularly in settings with significant segregation. In contexts with unequal infrastructure access, digital learning risks intensifying existing inequalities — an important consideration as online education continues to expand.

## **Declaration of generative AI use in the writing process.**

In the process of preparing this draft for submission, we utilized ChatGPT to refine the language and improve the readability of the text. After using this tool, we reviewed and edited the content as necessary, and we take full responsibility for the final submitted text.

## References

- Alfayumi-Zeadna, S., Gnaim-Abu Touma, L., Weinreich, M., OâRourke, N., 2022. Covid-19 and mental health of minority arab higher-education students in israel: Social, economic, and academic factors. *International Journal of Environmental Research and Public Health* 19, 13466.
- Alpert, W.T., Couch, K.A., Harmon, O.R., 2016. A randomized assessment of online learning. *American Economic Review* 106, 378–382.
- Altindag, D.T., S. Filiz, E., Tekin, E., 2024. Is online education working? *Educational Evaluation and Policy Analysis* , 01623737241274802.
- Asgari, S., Trajkovic, J., Rahmani, M., Zhang, W., Lo, R.C., Sciortino, A., 2021. An observational study of engineering online education during the COVID-19 pandemic. *Plos one* 16, e0250041.
- Bacher-Hicks, A., Goodman, J., 2021. The covid-19 pandemic is a lousy natural experiment for studying the effects of online learning. *EducationNext* .
- Bacher-Hicks, A., Goodman, J., Mulhern, C., 2021. Inequality in household adaptation to schooling shocks: Covid-induced online learning engagement in real time. *Journal of Public Economics* 193, 104345.
- Banerjee, A.V., Duflo, E., 2014. (dis) organization and success in an economics mooc. *American economic review* 104, 514–518.
- Barbosa-Camacho, F.J., Romero-Limón, O.M., Ibarrola-Peña, J.C., Almanza-Mena, Y.L., Pintor-Belmontes, K.J., Sánchez-López, V.A., Chejfec-Ciociano, J.M., Guzmán-Ramírez, B.G., Sapién-Fernández, J.H., Guzmán-Ruvalcaba, M.J., et al., 2022. Depression, anxiety, and academic performance in covid-19: a cross-sectional study. *Bmc Psychiatry* 22, 443.
- Barrow, L., Morris, W.T., Sartain, L., 2024. The expanding landscape of online education: Who engages and how they fare. *Journal of Labor Economics* 42, S417–S443.

- Bettinger, E.P., Castleman, B.L., Choe, A., Mabel, Z., 2022. Finishing the last lap: Experimental evidence on strategies to increase attainment for students near college completion. *Journal of Policy Analysis and Management* 41, 1040–1059.
- Bettinger, E.P., Fox, L., Loeb, S., Taylor, E.S., 2017. Virtual classrooms: How online college courses affect student success. *American Economic Review* 107, 2855–2875.
- Bilen, E., Matros, A., 2021. Online cheating amid covid-19. *Journal of Economic Behavior and Organization* 182, 196–211.
- Bird, K.A., Castleman, B.L., Lohner, G., 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.
- Bodas, M., Adini, B., Jaffe, E., Kaim, A., Peleg, K., 2022. Lockdown efficacy in controlling the spread of COVID-19 may be waning due to decline in public compliance, especially among unvaccinated individuals: A cross-sectional study in Israel. *International Journal of Environmental Research and Public Health* 19, 4943.
- Bowen, W.G., Chingos, M.M., Lack, K.A., Nygren, T.I., 2014. Interactive learning online at public universities: Evidence from a six-campus randomized trial. *Journal of Policy Analysis and Management* 33, 94–111.
- Bulman, G., Fairlie, R., 2022. The impact of COVID-19 on community college enrollment and student success: Evidence from California administrative data. *Education Finance and Policy* 17, 745–764.
- Cacault, M.P., Hildebrand, C., Laurent-Lucchetti, J., Pellizzari, M., 2021. Distance learning in higher education: evidence from a randomized experiment. *Journal of the European Economic Association* 19, 2322–2372.
- Clark, D., Gill, D., Prowse, V., Rush, M., 2020. Using goals to motivate college students: Theory and evidence from field experiments. *Review of Economics and Statistics* 102, 648–663.

- De Paola, M., Gioia, F., Scoppa, V., 2023. Online teaching, procrastination and student achievement. *Economics of Education Review* 94, 102378.
- Ericson, K.M., 2017. On the interaction of memory and procrastination: Implications for reminders, deadlines, and empirical estimation. *Journal of the European Economic Association* 15, 692–719.
- Figlio, D., Rush, M., Yin, L., 2013. Is it live or is it internet? experimental estimates of the effects of online instruction on student learning. *Journal of Labor Economics* 31, 763–784.
- Francis-Smythe, J.A., Robertson, I.T., 1999. On the relationship between time management and time estimation. *British Journal of Psychology* 90, 333–347.
- Gershoni, N., Stryjan, M., 2023. Do deadlines affect project completion? experimental evidence from israeli vocational colleges. *Journal of Economic Behavior & Organization* 205, 359–375.
- Ghanem, D., Hirshleifer, S., Kédagni, D., Ortiz-Becerra, K., 2024. Correcting attrition bias using changes-in-changes. *Journal of Econometrics* 241, 105737.
- Gilbar, O., Gelkopf, M., Berger, R., Greene, T., 2022. Risk factors for depression and anxiety during covid-19 in israel: A two-wave study before and during the pandemic. *Stress and Health* 38, 736–745.
- Hall, C., 2016. Does more general education reduce the risk of future unemployment? evidence from an expansion of vocational upper secondary education. *Economics of Education Review* 52, 251–271.
- Hanushek, E.A., Schwerdt, G., Woessmann, L., Zhang, L., 2017. General education, vocational education, and labor-market outcomes over the lifecycle. *Journal of Human Resources* 52, 48–87.
- Hawker, N., 2019. *The politics of Palestinian multilingualism: Speaking for citizenship*. Routledge.

- Jaggars, S.S., 2014. Choosing between online and face-to-face courses: Community college student voices. *American Journal of Distance Education* 28, 27–38.
- Johnson, M., Engelhardt, B., Taengnoi, S., 2023. Business school grades, assessment scores, and dropout rates in the covid-19 pandemic. *Journal of Education for Business* .
- Kahneman, D., Tversky, A., 1977. Intuitive prediction: Biases and corrective procedures. Technical Report. Decisions and Designs Inc Mclean Va.
- Kofoed, M., Gebhart, L., Gilmore, D., Moschitto, R., 2024. Zooming to class?: Experimental evidence on college students' online learning during COVID-19. *American Economic Review: Insights*, Forthcoming.
- Malamud, O., Pop-Eleches, C., 2010. General education versus vocational training: Evidence from an economy in transition. *The Review of Economics and Statistics* 92, 43–60.
- Nedelkoska, L., Quintini, G., 2018. Automation, skills use and training. OECD Social, Employment and Migration Working Papers.
- Novella, R., Rosas-Shady, D., Freund, R., 2024. Is online job training for all? experimental evidence on the effects of a coursera program in costa rica. *Journal of Development Economics* 169, 103285.
- O'Donoghue, T., Rabin, M., 1999. Doing it now or later. *American Economic Review* 89, 103–124.
- OECD, 2016. Education at a glance 2016: OECD indicators. URL: <http://dx.doi.org/10.187/eag-2016-en>.
- OECD, 2019a. Benchmarking Higher Education System Performance. Technical Report. OECD Publishing. URL: <https://doi.org/10.1787/be5514d7>.
- OECD, 2019b. Working better with age, ageing and employment policies. Technical Report. OECD Publishing. URL: <https://doi.org/10.1787/c4d4f66a>.

- OECD, 2021. Bridging connectivity divides. OECD Digital Economy Papers 315. OECD Publishing. Paris. URL: <https://doi.org/10.1787/e38f5db7>.
- OECD, 2024. OECD Digital Economy Outlook 2024 (Volume 1): Embracing the Technology Frontier. Technical Report. OECD. URL: <https://doi.org/10.1787/a1689dc5>.
- OECD, 2025. OECD affordable housing database: HC2.1 Housing space. Last updated 02/10/2025; Accessed 02/02/2026.
- Oreopoulos, P., Patterson, R.W., Petronijevic, U., Pope, N.G., 2022. Low-touch attempts to improve time management among traditional and online college students. *Journal of Human Resources* 57, 1–43.
- Rodríguez-Planas, N., 2022a. COVID-19, college academic performance, and the flexible grading policy: A longitudinal analysis. *Journal of Public Economics* 207, 104606.
- Rodríguez-Planas, N., 2022b. Hitting where it hurts most: COVID-19 and low-income urban college students. *Economics of Education Review* 87, 102233.
- Shah, A.K., Shafir, E., Mullainathan, S., 2015. Scarcity frames value. *Psychological Science* 26, 402–412.
- Tehawkho, M., Axelrad, H., Matar, H., 2021. The Digital Challenge in Arab Society. Technical Report 2021.05. Aaron Institute for Economic Policy, Reichman University. URL: [https://www.runi.ac.il/media/vqlfpo4i/the\\_digital\\_challenge\\_in\\_arab\\_society.pdf](https://www.runi.ac.il/media/vqlfpo4i/the_digital_challenge_in_arab_society.pdf). (Policy Paper in Hebrew).
- Tehawkho, M., Kalisher, I., Moskalev, K., 2020. Hebrew Language Proficiency in Arab Society: Barriers to Acquisition and Policy Recommendations. Technical Report 2020.06. Aaron Institute for Economic Policy, Reichman University. URL: <https://www.runi.ac.il/media/ytadpbvt/policy-paper.pdf>. (Policy Paper in Hebrew).
- The Knesset Research and Information Center, 2020. Fixed internet infrastructure and broadband speeds in peripheral localities URL: <https://fs.knesset>.

[gov.il/globaldocs/MMM/54b02e53-d9b6-ea11-8116-00155d0af32a/2\\_54b02e53-d9b6-ea11-8116-00155d0af32a\\_11\\_16397.pdf](https://fs.knesset.gov.il/globaldocs/MMM/54b02e53-d9b6-ea11-8116-00155d0af32a/2_54b02e53-d9b6-ea11-8116-00155d0af32a_11_16397.pdf). (Policy Paper in Hebrew).

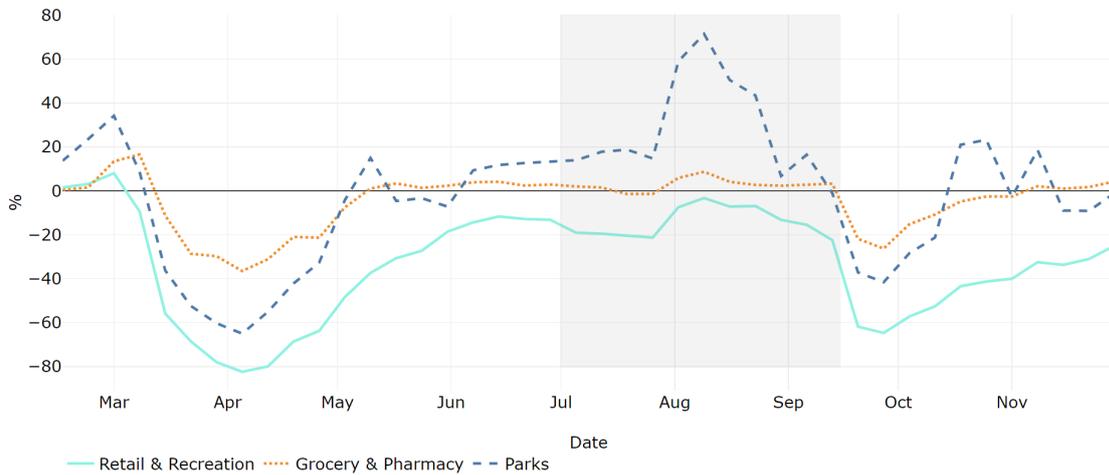
The Knesset Research and Information Center, 2024. Hebrew language instruction and proficiency among arab youth and young adults URL: [https://fs.knesset.gov.il/globaldocs/MMM/b827739a-28f7-ee11-815f-005056aac6c3/2\\_b827739a-28f7-ee11-815f-005056aac6c3\\_11\\_20638.pdf](https://fs.knesset.gov.il/globaldocs/MMM/b827739a-28f7-ee11-815f-005056aac6c3/2_b827739a-28f7-ee11-815f-005056aac6c3_11_20638.pdf). (Policy Paper in Hebrew).

Xu, D., Jaggars, S.S., 2013. The impact of online learning on students' course outcomes: Evidence from a large community and technical college system. *Economics of Education Review* 37, 46–57.

Xu, D., Jaggars, S.S., 2014. Performance gaps between online and face-to-face courses: Differences across types of students and academic subject areas. *The Journal of Higher Education* 85, 633–659.

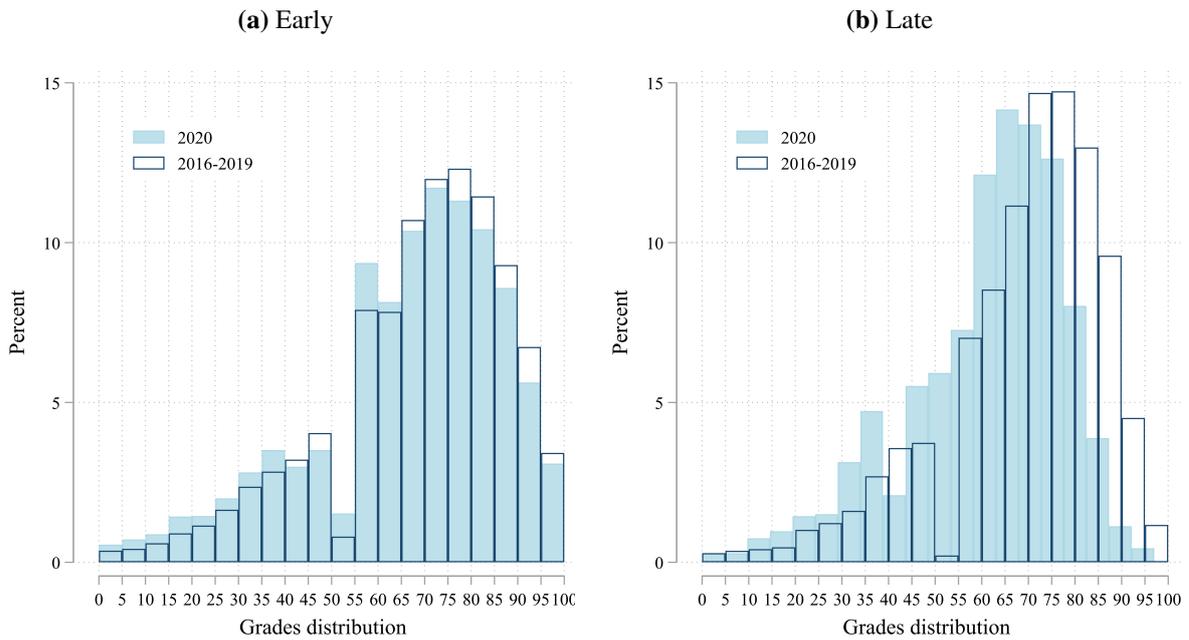
## 8 Figures and Tables

**Figure 1:** Weekly Mobility Trends in Israel during 2020



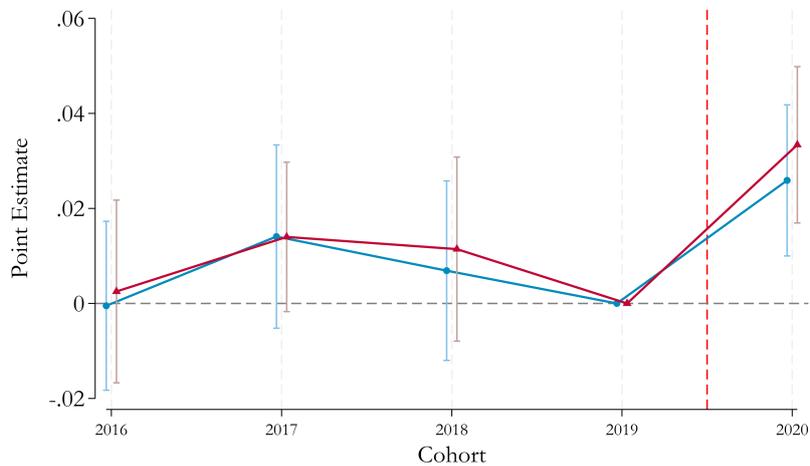
*Notes:* The figure presents mobility trends in Israel during 2020 at the weekly level, using data provided by Google’s Community Mobility Reports. These trends are shown separately for three categories of places — Retail and recreation (restaurants, shopping centers, theme parks, museums, etc.), Grocery and pharmacy, and Parks (including public beaches, marinas, dog parks, plazas, and public gardens). For each week, the percentage change in movement was calculated relative to the median value from from a “baseline period” selected by Google to be January 3 - February 6, 2020 (the weeks just before the pandemic). The shaded area marks the exam period in our study.

**Figure 2:** The Distribution of Early and Late Exam Grades (Adjusted), 2020 vs. Previous Cohorts



*Notes:* The figures plot grade distributions for early and late exams, comparing the 2020 cohort to 2016-2019. For exams administered after Spring 2020, reported grades are divided by 120 (the maximum exam score for that period).

**Figure 3:** Event Studies — Exam Absence



*Notes:* The figure plots event study estimates of treatment effects on absence and the corresponding confidence intervals. The blue line with circles presents estimates from a specification similar to equation (1) but with separate indicators for each cohort and their interaction with an indicator for late exam. This specification includes department and exam fixed effects, as well as individual level controls for age (quadratic), sex, ethnicity, SES, and study track (morning/evening). The red line with triangles presents estimates of the same specification but with individual fixed effects.

**Table 1: Summary statistics**

Variable	Mean	Std. Dev.	Mean (Estimated)	SE	N
<i>Panel A - All cohorts, Administrative data</i>					
Age (years)	27.42	6.88			15,352
Female	0.068	0.252			15,352
Arab	0.314	0.464			15,352
Morning track	0.540	0.498			15,352
Civil Engineering	0.608	0.488			15,352
Electrical Engineering	0.279	0.448			15,352
Software Programming	0.113	0.316			15,352
Teaching Language Arabic	0.135	0.342			15,352
<i>Locality-level</i>					
Socioeconomic Z-score	-0.062	0.855			15,295
High Household Crowdedness	0.211	0.408			13,245
<i>Panel B - 2020 cohort, Survey data</i>					
	Survey Sample		Population Inference		
Married	0.37	0.483	0.364	0.016	1,059
Has children	0.286	0.452	0.280	0.015	1,059
Has children under 6	0.25	0.433	0.244	0.014	1,059
Lives with parents or other older relatives	0.509	0.5	0.515	0.017	1,058
Moved to parents/relatives during pandemic	0.051	0.22	0.048	0.007	1,058
Moved to more independent setup during pandemic	0.034	0.181	0.037	0.007	1,058
Mother with Tertiary Education	0.28	0.449	0.256	0.014	960
Father with Tertiary Education	0.271	0.445	0.259	0.015	953
1st generation immigrant	0.276	0.447	0.310	0.016	1,059
1st or 2nd gen immigrant	0.5	0.5	0.518	0.017	1,059
Worked before studies	0.941	0.235	0.946	0.007	1,059
Worked in study field before studies	0.375	0.484	0.415	0.017	1,059
Worked during studies	0.868	0.339	0.870	0.011	1,059
Lost job during pandemic	0.145	0.353	0.153	0.012	1,059

*Notes:* Panel A reports summary statistics for our sample of students based on administrative data from the NITT. “Arab” students are students from the Arabic speaking minority within Israel who may identify primarily as Palestinian, Arab, Druze, or Bedouin. The socioeconomic score is based on the Z-score calculated by the Israeli Central Bureau of Statistics for each student’s locality of residence in 2017. The locality of residence is missing for less than 0.4% of our sample. Household crowdedness is calculated at the locality level based on restricted data from the Israeli Social Survey, and it is missing for very small localities. A “High-crowdedness” locality is one where at least 40% of households report more than one resident per room and at least 10% report two or more residents per room. Panel B reports summary statistics from the 2020 cohort survey. The number of observations for parents’ education is smaller than for the other variables because we allowed for an “irrelevant or unknown” category which is coded as missing values. 2nd generation immigrant means that at least one of the parents is an immigrant. “Moved to parents/relatives during pandemic” is an indicator for having lived alone or with a partner or roommates before the pandemic and reporting moving in with parents or relatives during the pandemic, while “Moved to more independent setup” is an indicator for moving out from parents or relatives. Data on current living setup is missing for one respondent. Columns 1-2 show mean values and standard deviations for the survey sample, while columns 3-4 present estimated population means and corresponding standard errors, using inverse probability weights to account for selection into the survey sample. Weights are generated by regressing an indicator for being in the survey sample on students’ age, gender, SES and fixed effects for *major*  $\times$  *college*. The population N in the 2020 cohort = 2,893.

**Table 2:** Number of Mandatory Exams and Their Timing by Major and Track

Major	Number of mandatory exams	Track	Number of early exams	Number of late exams
Software Programming	3	Morning	2	1
		Evening	2	1
Electrical Engineering	3	Morning	1	2
		Evening	2	1
Civil Engineering	4	Morning	2	2
		Evening	4	0

*Notes:* The table presents the total number of mandatory exams in each study field and their classification into late and early exam timing by track (morning or evening). An exam is defined as late (early) if less (more) than 50% of the students take the exam before their last semester, conditional on taking the exam. For the full classification, see Table A1.

**Table 3:** The Effects of Online Instruction on Exam Outcomes

	Absence		Underperformance		Underper. Attendance	
	(1)	(2)	(3)	(4)	(5)	(6)
Late Exam × D2020	0.0153 (0.0092) [0.00701]***	0.0210 (0.0082)** [0.0062]***	0.0935 (0.0157)*** [0.0099]***	0.0888 (0.0175)*** [0.0097]***	0.0927 (0.0192)*** [0.0097]***	0.0834 (0.0206)*** [0.010]***
Outcome Mean	0.0941	0.0949	0.235	0.236	0.155	0.154
Effect in %	16.3	22.1	39.9	37.6	59.8	54.1
Observations	55,179	55,395	55,179	55,395	50,255	49,927
Major×College FE	✓		✓		✓	
Track FE	✓		✓		✓	
Individual FE		✓		✓		✓

*Notes:* The table presents results for our main outcomes. *Absence* is a dummy variable which equals one if the student did not take the exam. *Underperformance* indicates that the student scored less than 55% of the exam points. Columns 5 and 6 show estimates for Underperformance with the sample restricted to students attending the exam. Columns 1, 3, and 5 include department and track fixed effects and controls for age (quadratic), sex, ethnicity, enrollment in an Arabic-speaking college, and socioeconomic rank based on the locality of residence, while columns 2, 4, and 6 include individual fixed effects. All columns include exam-subject fixed effects. The outcome mean is calculated using only pre-treatment cohorts' outcomes in late exams. Standard errors clustered by department (*Major × College*) in parentheses; standard errors clustered at student level in square brackets. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 4:** Heterogeneous Effects on Exam Absence by Pandemic Related Factors

	Infection Rate		Age		SES	
	High	Low	> 25	≤ 25	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)
Late Exam × D2020	0.0205* (0.0107)	0.0235* (0.0119)	0.0171 (0.0111)	0.0228** (0.0111)	0.0283** (0.0116)	0.0159* (0.00926)
Outcome Mean	0.118	0.0779	0.0607	0.118	0.162	0.0552
Effect in %	17.3	30.1	28.1	19.2	17.4	28.7
P-value (Diff.)	0.85		0.71		0.33	
Observations	24,782	25,173	28,670	26,725	18,346	36,833
Individual FE	✓	✓	✓	✓	✓	✓

*Notes:* The table presents results splitting the sample by potential changes related to the pandemic other than online learning. Infection rates and socioeconomic status (SES) are calculated at the locality level based on students' locality of residence. A locality is considered to have high infection rates if its average monthly rate of COVID-19 related hospitalizations in March-September 2020 is above the median. Age is split by the sample median. SES is defined based on the socioeconomic Z-score calculated by the Israeli central Bureau of Statistics, and we classify the bottom SES tercile as "Low SES" and the two upper terciles as "High SES". For more details on the variables, see section 4. The outcome mean is calculated using only pre-treatment cohorts' outcomes in late exams. All columns include exam-subject fixed effects. Standard errors are clustered by department (*Major × College*). \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 5:** Heterogeneous Effects on Exam Absence by Ethnicity and Socioeconomic Status

	All		Low SES		High SES	
	Arab	Jewish	Arab	Jewish	Arab	Jewish
	(1)	(2)	(3)	(4)	(5)	(6)
Late Exam × D2020	0.0258* (0.0136)	0.0144 (0.00941)	0.0281** (0.0134)	0.0158 (0.0232)	0.0211 (0.0224)	0.0144 (0.00925)
Outcome Mean	0.187	0.0473	0.197	0.0425	0.128	0.0479
Effect in %	13.8	30.4	14.2	37.1	16.4	30.1
P-value (Diff.)	0.45		0.65		0.76	
Observations	17,715	37,680	14,181	4,165	3,357	33,476
Individual FE	✓	✓	✓	✓	✓	✓

*Notes:* The table presents results for exam absence by SES group and minority status (Arab or Jewish). SES is defined based on a socioeconomic Z-score calculated by the Israeli central Bureau of Statistics, and we classify the bottom SES tercile as "Low SES" and the two upper terciles as "High SES". Students are classified as either "Jewish" or "Arab" based on administrative student level records. Where ethnicity records are missing, they are imputed based on names and localities of residence, that are highly segregated in Israel. For more details on the variables, see section 4. The outcome mean is calculated using only pre-treatment cohorts' outcomes in late exams. All columns include exam-subject fixed effects. Standard errors are clustered by department (*Major × College*). \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 6:** Heterogeneous Effects on Exam Absence by Attributes Related to Learning Effectiveness

	Internet Infrastructure		Household Crowdedness		Language of Instruction		Previous Grades	
	Poor	Standard	High	Low	Hebrew	Arabic	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A — Arab</i>								
Late × D2020	0.0481*** (0.0178)	0.00749 (0.0165)	0.0155 (0.0201)	0.0303** (0.0125)	0.0457** (0.0182)	0.0245*** (0.00413)	0.0148 (0.0252)	0.0369*** (0.0132)
Outcome Mean	0.198	0.178	0.184	0.176	0.0938	0.279	0.285	0.0351
Effect in %	24.3	4.21	8.44	17.3	48.7	8.76	5.21	104.9
P-value (Diff.)	0.08		0.50		0.26		0.54	
Observations	7,586	10,129	7,239	8,536	9,856	7,859	10,896	6,819
Individual FE	✓	✓	✓	✓	✓	✓	✓	✓
<i>Panel B — Jewish</i>								
Late × D2020	-0.0504 (0.0318)	0.0157 (0.00954)	0.00399 (0.0287)	0.0133 (0.0107)			0.0324* (0.0177)	0.00157 (0.00703)
Outcome Mean	0.0710	0.0468	0.0462	0.0496			0.0949	0.0155
Effect in %	-71.0	33.5	8.63	2.68			34.1	10.1
P-value (Diff.)	0.04		0.75				0.09	
Observations	739	36,941	3,141	28,738			15,736	21,944
Individual FE	✓	✓	✓	✓			✓	✓

*Notes:* The table presents results for exam absence by potential mechanisms related to or interacting with online learning. A student's internet infrastructure is considered to be poor if there is only one network provider serving their locality of residence, and standard otherwise. A "High Household Crowdedness" locality is one where at least 40% of households report more than one resident per room and at least 10% report two or more residents per room. The language of instruction is defined at the college level and only varies within the Arab population, as all Jewish students in our data attend Hebrew language colleges. Previous grades are split by the median average grade in our sample, so that "Low" refers to below median grades. The outcome mean is calculated using only pre-treatment cohorts' outcomes in late exams. The crowdedness measure is missing for relatively small localities, affecting 7,741 observations (from 2,107 students). All columns include exam-subject fixed effects. Standard errors are clustered by department (*Major × College*). \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

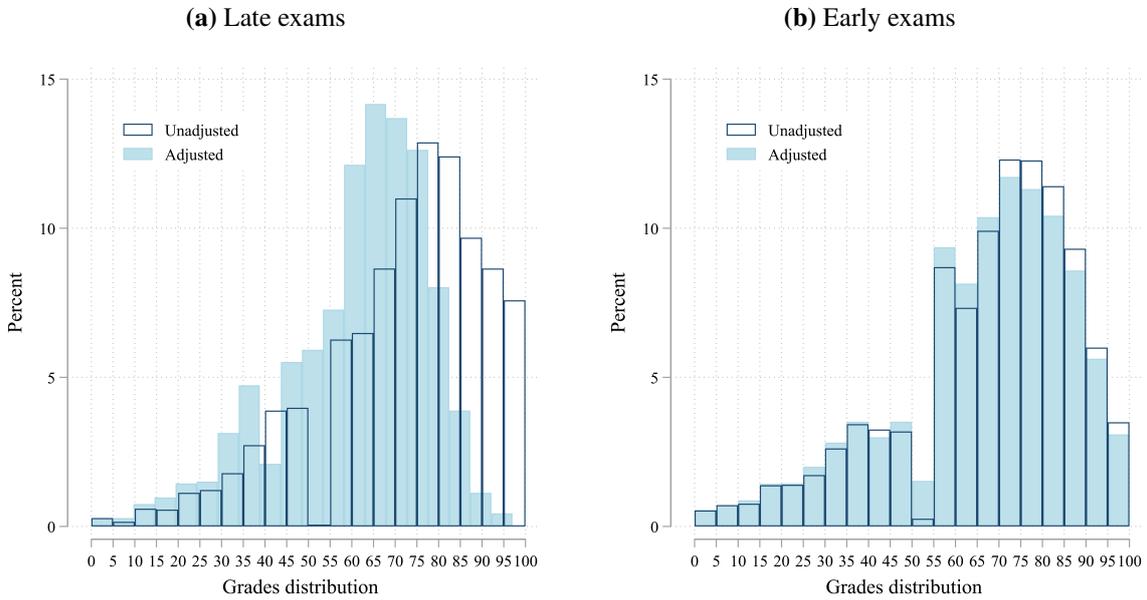
**Table 7: Heterogeneous Effects on Exam Absence by Extent of in-Person Classes**

	Arab		Jewish	
	Low	High	Low	High
	(1)	(2)	(3)	(4)
Late Exam $\times$ D2020	0.0212 (0.0165)	0.0411** (0.0159)	0.0146 (0.0190)	0.0144 (0.0223)
Outcome Mean	0.204	0.196	0.0496	0.0420
Effect in %	10.4	20.9	29.5	34.2
P-value (Diff.)	0.45		0.99	
Observations	3,724	8,640	11,762	9,179
Individual FE	✓	✓	✓	✓

*Notes:* The table presents results for exam absence by the share of in-person classes held during June 2020, separately estimated for Arab students and for Jewish students. This share is defined based on the response in our department head survey during the last weeks of the spring semester, when some in-person teaching took place. Departments with above median share of in-person classes are categorized as “High”. The outcome mean is calculated using only pre-treatment cohorts’ outcomes in late exams. All columns include exam-subject fixed effects. Standard errors are clustered by department (*Major  $\times$  College*). \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

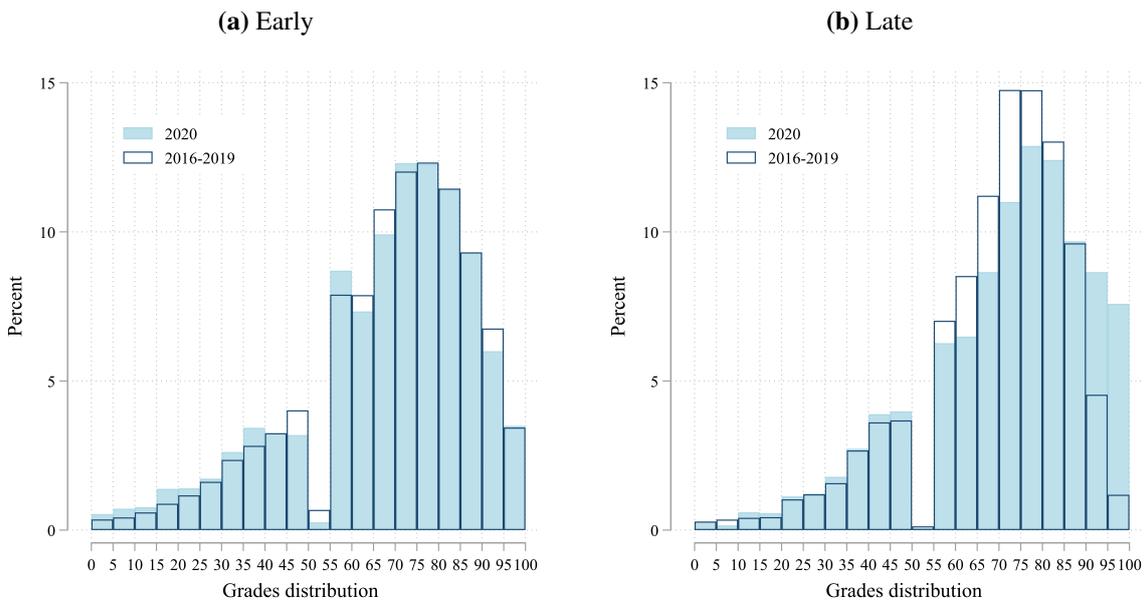
# A.1 Appendix Figures and Tables

**Figure A1:** The Distribution of Adjusted and Unadjusted Grades in the 2020 Cohort



*Notes:* The figures plot adjusted and unadjusted grade distributions for the 2020 cohort, split into early and late exams. Unadjusted = reported grades. Adjusted = reported grades through Spring 2020; after that, reported grades divided by 120 (the maximum exam score for that period).

**Figure A2:** The Distribution of Early and Late Exam Reported Grades, 2020 vs. Previous Cohorts



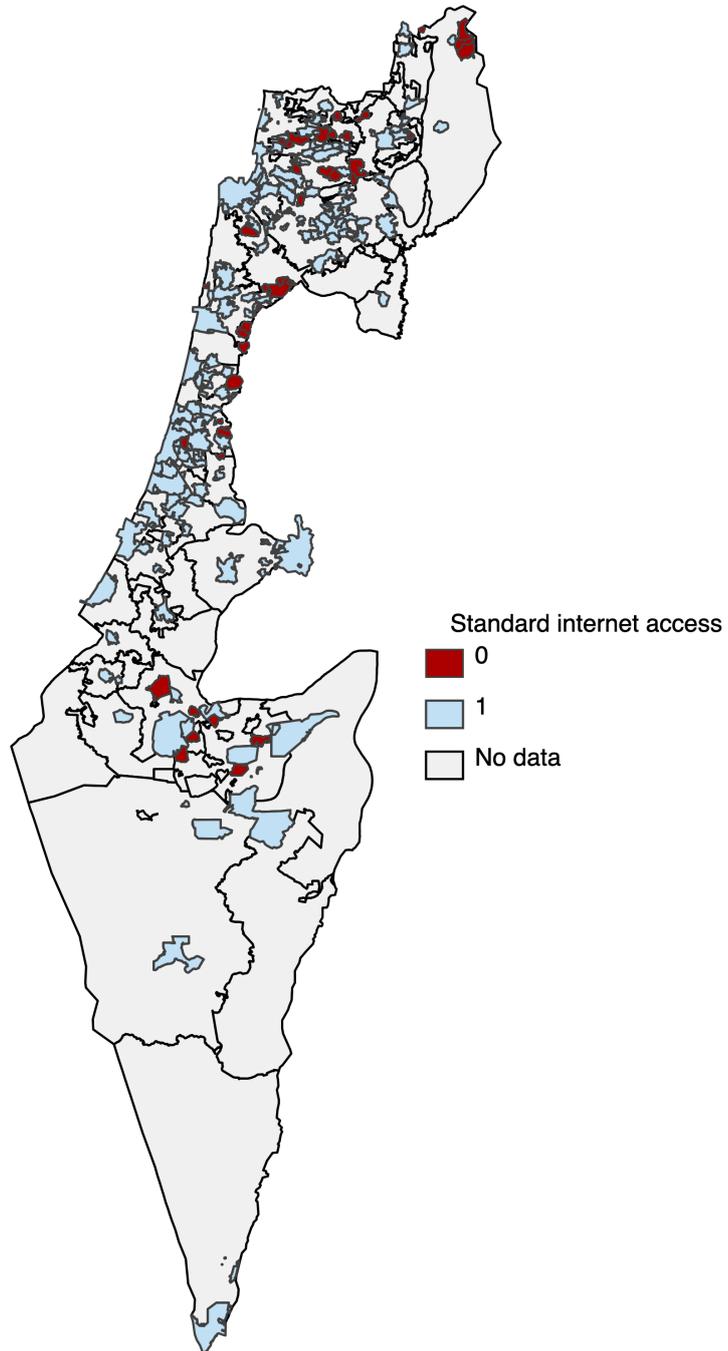
*Notes:* The figures present the distribution of reported grades (unadjusted) in either late or early exams, for students in the 2020 cohort compared to the 2016-2019 cohorts.

**Table A1:** Exam Classification into Late and Early

Major	Exam id	Track	Share of students taking the exam before their final semester		Late Exam
			All students	Conditional on attending	
			(1)	(2)	
Software Programming	1	Evening	96%	99%	0
		Morning	88%	96%	0
	2	Evening	14%	14%	1
		Morning	1%	1%	1
	3	Evening	85%	94%	0
		Morning	59%	70%	0
Electrical Engineering	4	Evening	86%	98%	0
		Morning	85%	95%	0
	5	Evening	86%	91%	0
		Morning	9%	9%	1
	6	Evening	25%	27%	1
		Morning	3%	4%	1
Civil Engineering	7	Evening	92%	97%	0
		Morning	11%	13%	1
	8	Evening	80%	88%	0
		Morning	49%	55%	0
	9	Evening	49%	54%	0
		Morning	28%	32%	1
10	Evening	92%	97%	0	
	Morning	89%	95%	0	

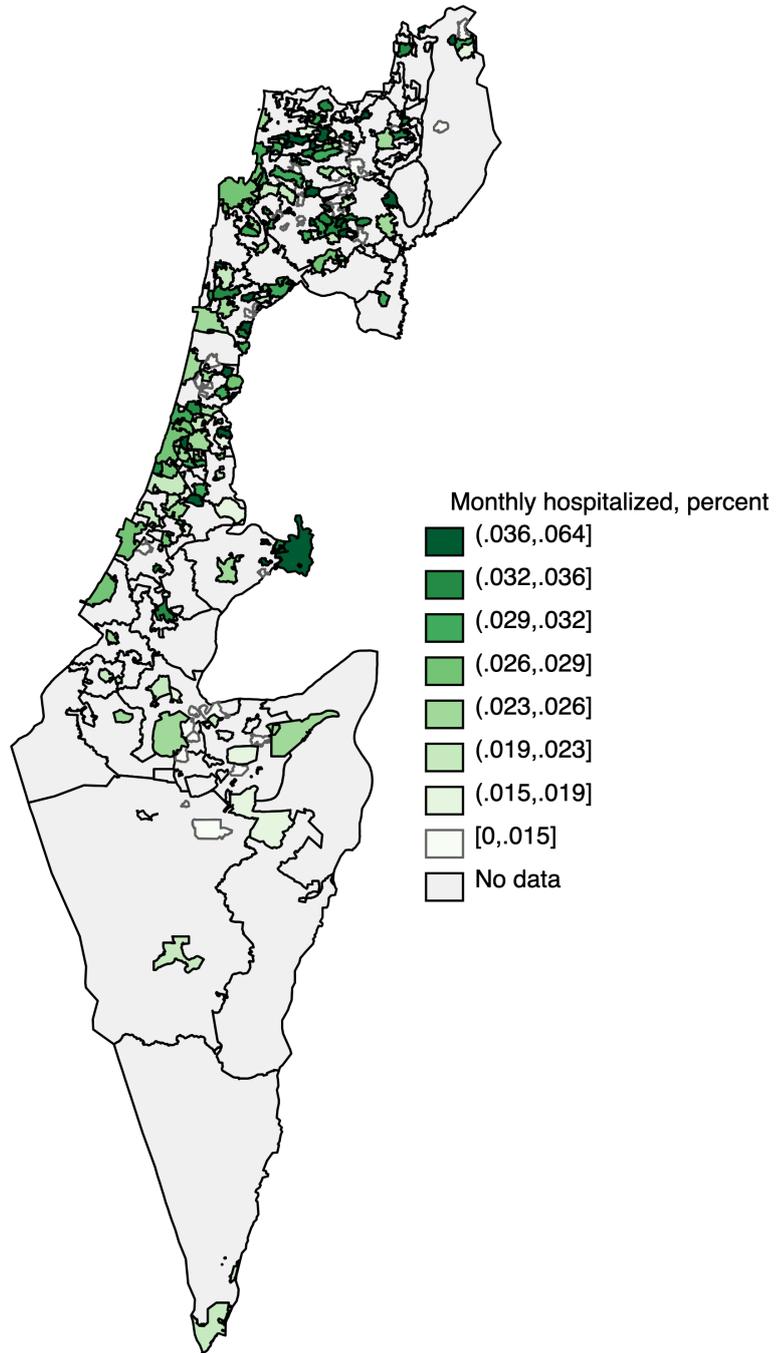
*Notes:* The table reports the share of students that took the exam before their final semester for each exam subject by study field. Column (1) presents these shares out of all students in the track, whereas column (2) shows these shares only for students who actually took the exam. Column (3) indicates how the exam was categorized in each track for the purpose of our analysis.

**Figure A3: Geographical Variation in Internet Quality**



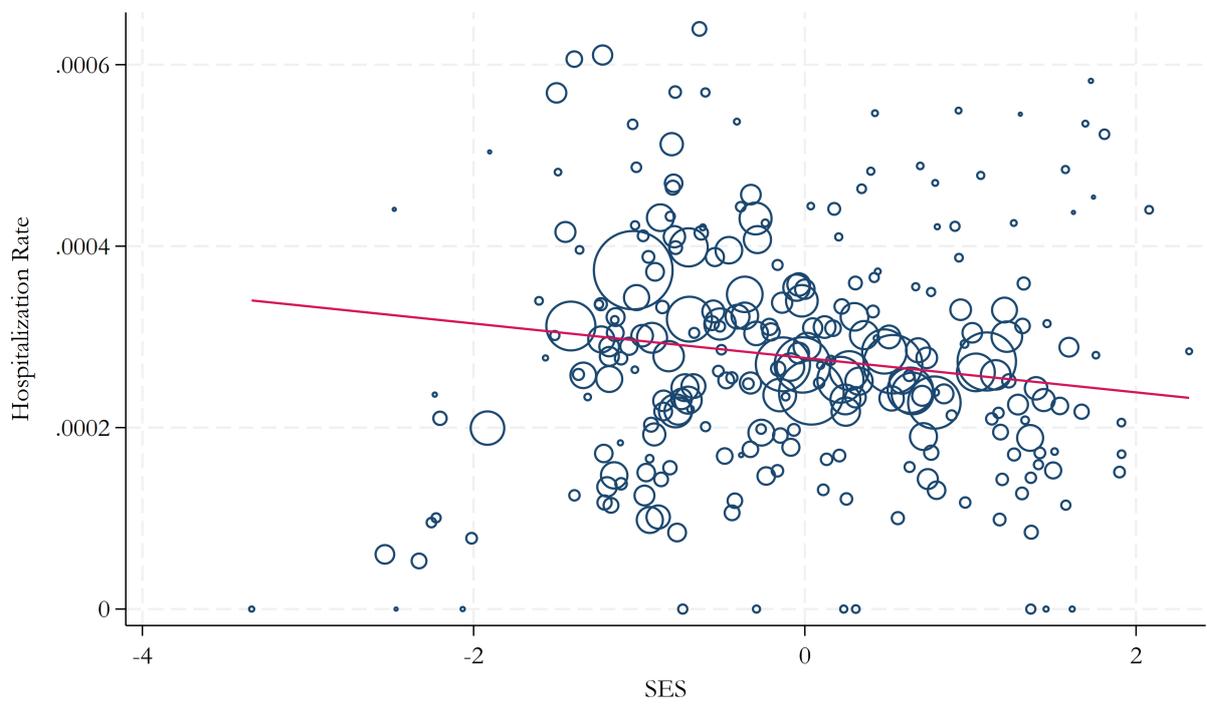
*Notes:* The figure presents the geographical variation in internet quality across localities in Israel according to the measure explained in section 4. Light blue color denotes standard internet quality (due to more than one provider) and dark red color denotes poor internet quality.

**Figure A4: Geographical Variation in Infection Rates**



*Notes:* The figure presents the geographical variation in average monthly COVID-19 related hospitalization rates (as a percentage of locality population) during the period March to September 2020 across localities in Israel. Darker colors denote higher hospitalization rates.

**Figure A5: Socioeconomic Status and Infection Rates**



*Notes:* The figure plots local COVID-19 related hospitalization rates (number of hospitalized individuals divided by locality population size) against the localities' socioeconomic Z-score provided by the Israeli Central Bureau of Statistics. Circle sizes represent the number of students in our sample from each specific locality (number of individual-level observations).

**Table A2: Student-Reported Challenges and Expectations by Internet Quality**

	Standard internet mean (s.d.) (1)	Poor internet mean (s.d.) (2)	Difference [s.e.] (3)	Observations (4)
No contact with fellow students	0.052 (0.222)	0.096 (0.296)	0.044* [0.025]	1,059
No contact with supervisor	0.392 (0.488)	0.596 (0.492)	0.204*** [0.043]	1,059
I always complete tasks	0.875 (0.331)	0.776 (0.419)	-0.099*** [0.035]	1,059
I am focused when working independently	0.828 (0.377)	0.782 (0.414)	-0.046 [0.035]	1,059
Self reported graduation probability	5.039 (1.606)	3.506 (2.157)	-1.532*** [0.180]	1,059
Academic challenges	0.730 (0.444)	0.808 (0.395)	0.078** [0.035]	1,059
Family obligation challenges	0.319 (0.466)	0.263 (0.442)	-0.056 [0.039]	1,059
COVID-related challenges	0.623 (0.485)	0.570 (0.497)	-0.053 [0.043]	1,059

*Notes:* All the data in the table is based on the student survey. Columns 1 and 2 show the mean and standard deviation of each variable for the sample of students with standard internet and poor internet, respectively. Column 3 shows the difference (coefficients and standard errors) between the two groups obtained from a regression of each variable on the indicator for poor internet. Robust standard errors are shown in square brackets. *No contact with fellow students/supervisor* are dummies indicating that the respondent had had no contact with this group in the past 30 days. *I always complete tasks* and *I am focused when working independently* are dummy variables indicating that the student replied strongly agree or agree on a 5-point scale to the question “I finish every task that I start” and “When I work independently, I am focused,” respectively. *Self reported graduation probability* captures the student’s reported probability to graduate on a scale from 1 (zero probability) to 6 (very high). The *X challenges*-questions are dummy variables indicating that the student replied Agree or Somewhat agree on a statement about whether they were facing difficulty *X* during the months before the exams. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table A3:** The Effects of Online Instruction on Official Exam Failure

	Failure		Failure Attendance	
	(1)	(2)	(3)	(4)
Late Exam $\times$ D2020	0.0139 (0.0151)	0.00428 (0.0137)	0.00113 (0.0155)	-0.0131 (0.0167)
Outcome Mean	0.233	0.234	0.153	0.152
Effect in %	5.97	1.83	0.74	-8.6
Observations	55,179	55,395	50,255	49,927
Major $\times$ College FE	✓		✓	
Track FE	✓		✓	
Individual FE		✓		✓

*Notes:* The table presents results for the outcome *Failure*, a dummy for scoring less than 55 points on the exam according to the reported grade. The reported grades are based on the raw exam score, and were truncated at 100 after the spring semester when the maximum attainable exam-score was changed to 120. Columns 3 and 4 show estimates for Underperformance with the sample restricted to students attending the exam. Columns 1 and 3 include department and track fixed effects and controls for age (quadratic), sex, ethnicity, enrollment in an Arabic-teaching college, and socioeconomic rank based on the locality of residence, while columns 2 and 4 include individual fixed effects. All columns include exam-subject fixed effects. The outcome mean is calculated using only pre-treatment cohorts' outcomes in late exams. Standard errors are clustered by department (*Major  $\times$  College*). \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

**Table A4:** The Effects of Online Instruction on Exam Outcomes – All Controls Displayed

	Absence		Underperformance		Underper. Attendance	
	(1)	(2)	(3)	(4)	(5)	(6)
Late Exam × D2020	0.0153 (0.0092)	0.0210** (0.0082)	0.0935*** (0.0157)	0.0888*** (0.0175)	0.0927*** (0.0192)	0.0834*** (0.0206)
Late Exam	0.0226** (0.0103)	0.0216** (0.0098)	0.0694*** (0.0178)	0.0674*** (0.0171)	0.0541*** (0.0141)	0.0621*** (0.0161)
Age	-0.00801** (0.0037)		-0.0235*** (0.0058)		-0.0196*** (0.0041)	
Age <sup>2</sup>	0.0001** (0.0000)		0.0003*** (0.0000)		0.0002*** (0.0000)	
Arab=1	0.0399*** (0.0116)		0.125*** (0.0206)		0.103*** (0.0163)	
Female=1	-0.0068 (0.0081)		0.0100 (0.0094)		0.0163* (0.0087)	
2017 SES Rank	-0.0090** (0.0043)		-0.0134** (0.0064)		-0.0076* (0.0043)	
Constant	0.197*** (0.0608)	0.0810*** (0.0032)	0.551*** (0.0940)	0.230*** (0.0062)	0.432*** (0.0665)	0.155*** (0.0059)
Outcome mean	0.0941	0.0949	0.235	0.236	0.155	0.154
Effect in %	16.3	22.1	39.9	37.6	59.8	54.1
Observations	55,179	55,395	55,179	55,395	50,255	49,927
Major×College FE	✓		✓		✓	
Track FE	✓		✓		✓	
Cohort FE	✓		✓		✓	
Controls	✓		✓		✓	
Exam Subject FE	✓	✓	✓	✓	✓	✓
Individual FE		✓		✓		✓
R <sup>2</sup>	0.090	0.665	0.126	0.589	0.101	0.506

Notes: The table presents results for our main outcomes from the same estimations as shown in Table 3, and also displays the coefficients for all control variables that are omitted from the output in Table 3. The outcome mean is calculated using only pre-treatment cohorts' outcomes in late exams. Standard errors clustered by department (*Major × College*) in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table A5: Placebo Tests**

	Including 2020			Excluding 2020		
	2017	2018	2019	2017	2018	2019
	(1)	(2)	(3)	(4)	(5)	(6)
Late $\times$ Placebo Cohort	-0.00288 (0.00518)	-0.00634 (0.00799)	0.00614 (0.00693)	0.00227 (0.00490)	-0.00135 (0.00780)	0.0121 (0.00787)
Observations	55,395	55,395	55,395	44,905	44,905	44,905
Individual FE	✓	✓	✓	✓	✓	✓

*Notes:* The table presents results from placebo tests for our main outcome variable absence. We have estimated a regression analogous to the one for Table 3, column 2, in which the late-exam indicator is interacted with cohort dummies for the pre-period years 2017, 2018 and 2019, respectively. Columns 1-3 display results from an estimation using the same data as in the main analysis, while columns 4-6 display results excluding the actual treated cohort (2020). All columns include exam-subject fixed effects. Standard errors are clustered by department (*Major  $\times$  College*). \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

**Table A6: Heterogeneous Effects on Exam Underperformance by Pandemic Related Factors**

	Infection Rate		Age		SES	
	High	Low	> 25	$\leq$ 25	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)
Late Exam $\times$ D2020	0.103*** (0.0237)	0.0873*** (0.0211)	0.0813*** (0.0240)	0.0910*** (0.0234)	0.118*** (0.0311)	0.0711*** (0.0181)
Outcome Mean	0.263	0.223	0.198	0.263	0.329	0.181
Effect in %	39.1	39.2	41.2	34.7	35.8	39.3
P-value (Diff.)	0.55		0.77		0.47	
Observations	24,782	25,173	28,670	26,725	18,346	36,833
Individual FE	✓	✓	✓	✓	✓	✓

*Notes:* The table presents results for the outcome *Underperformance*, an indicator for scoring less than 55% of the exam points, splitting the sample by potential changes related to the pandemic other than online learning. Infection rates and socioeconomic status (SES) are calculated at the locality level based on students' locality of residence. A locality is considered to have high infection rates if its average monthly rate of COVID-related hospitalizations in March-September 2020 is above the median. Age is split by the sample median. SES is defined based on the socioeconomic Z-score calculated by the Israeli central Bureau of Statistics, and we classify the bottom SES tercile as "Low SES" and the two upper terciles as "High SES". For more details on the variables, see section 4. The outcome mean is calculated using only pre-treatment cohorts' outcomes in late exams. All columns include exam-subject fixed effects. Standard errors are clustered by department (*Major  $\times$  College*). \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

**Table A7: Heterogeneous Effects on Exam Underperformance by Ethnicity and SES**

	All		Low SES		High SES	
	Arab	Jewish	Arab	Jewish	Arab	Jewish
	(1)	(2)	(3)	(4)	(5)	(6)
lateXpost	0.129*** (0.0328)	0.0575*** (0.0178)	0.129*** (0.0355)	0.0331 (0.0494)	0.137*** (0.0441)	0.0604*** (0.0185)
Outcome Mean	0.379	0.162	0.392	0.110	0.300	0.169
Effect in %	34.1	35.5	33.0	30.0	45.7	35.7
P-value (Diff.)	0.05		0.11		0.08	
Observations	17,715	37,680	14,181	4,165	3,357	33,476
Individual FE	✓	✓	✓	✓	✓	✓

*Notes:* The table presents results for exam underperformance, an indicator for scoring less than 55% of the exam points, by SES group and minority status (Arab or Jewish). SES is defined based on a socioeconomic Z-score calculated by the Israeli central Bureau of Statistics, and we classify the bottom SES tercile as “Low SES” and the two upper terciles as “High SES”. Students are classified as either “Jewish” or “Arab” based on administrative student level records. Where ethnicity records are missing, they are imputed based on names and localities of residence, that are highly segregated in Israel. For more details on the variables, see section 4. The outcome mean is calculated using only pre-treatment cohorts’ outcomes in late exams. All columns include exam-subject fixed effects. Standard errors are clustered by department (*Major* × *College*). \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table A8:** Heterogeneous Effects on Exam Underperformance by Attributes Related to Learning Effectiveness

	Internet Infrastructure		Household Crowdedness		Language of Instruction		Previous Grades	
	Poor (1)	Standard (2)	High (3)	Low (4)	Hebrew (5)	Arabic (6)	Low (7)	High (8)
<i>Panel A — Arab</i>								
Late × D2020	0.139*** (0.0367)	0.121** (0.0553)	0.101** (0.0439)	0.166*** (0.0442)	0.0916*** (0.0326)	0.226*** (0.0301)	0.125*** (0.0351)	0.146*** (0.0306)
Outcome Mean	0.401	0.362	0.387	0.361	0.321	0.437	0.545	0.123
Effect in %	34.5	33.6	26.1	46.1	28.6	51.7	22.9	118.4
P-value (Diff.)	0.80		0.24		0.003		0.49	
Observations	7,586	10,129	7,239	8,536	9,856	7,859	10,896	6,819
Individual FE	✓	✓	✓	✓	✓	✓	✓	✓
<i>Panel B — Jewish</i>								
Late × D2020	0.0516 (0.0798)	0.0575*** (0.0174)	0.0446 (0.0573)	0.0629*** (0.0191)			0.0812** (0.0338)	0.0411*** (0.0131)
Outcome Mean	0.160	0.162	0.104	0.177			0.315	0.0600
Effect in %	32.3	35.5	429	0.355			25.8	68.5
P-value (Diff.)	0.94		0.76				0.24	
Observations	739	36,941	3,141	28,738			15,736	21,944
Individual FE	✓	✓	✓	✓			✓	✓

*Notes:* The table presents results for exam underperformance, an indicator for scoring less than 55% of the exam points, by potential mechanisms related to or interacting with online learning. A student’s internet infrastructure is considered to be poor if there is only one network provider serving their locality of residence, and standard otherwise. A “High Household Crowdedness” locality is one where at least 40% of households report more than one resident per room and at least 10% report two or more residents per room. The language of instruction is defined at the college level and only varies within the Arab population, as all Jewish students in our data attend Hebrew language colleges. Previous grades are split by the median average grade in our sample, so that “Low” refers to below median grades. The outcome mean is calculated using only pre-treatment cohorts’ outcomes in late exams. The crowdedness measure is missing for relatively small localities, affecting 7,741 observations (from 2,107 students). All columns include exam-subject fixed effects. Standard errors are clustered by department (*Major × College*). \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

**Table A9:** Heterogeneous Effects on Exam Underperformance by Extent of in-Person Classes

	Arab		Jewish	
	Low	High	Low	High
	(1)	(2)	(3)	(4)
Late Exam $\times$ D2020	0.157* (0.0856)	0.152*** (0.0496)	0.0602** (0.0243)	0.0380 (0.0292)
Outcome Mean	0.371	0.373	0.118	0.151
Effect in %	42.5	40.7	50.8	25.2
P-value (Diff.)	0.96		0.61	
Observations	3,724	8,640	11,762	9,179
Individual FE	✓	✓	✓	✓

*Notes:* The table presents results for exam underperformance, an indicator for scoring less than 55% of the exam points, by the share of in-person classes held during June 2020, separately estimated for Arab students and for Jewish students. This share is defined based on the response in our department head survey during the last weeks of the spring semester, when some in-person teaching took place. Departments with above median share of in-person classes are categorized as “High”. The outcome mean is calculated using only pre-treatment cohorts’ outcomes in late exams. All columns include exam-subject fixed effects. Standard errors are clustered by department (*Major  $\times$  College*). \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

## **A.2 Additional Details on Survey Data Collection**

We administered a survey to the 2,893 students in the 2020 cohort in September 2020 just after completion of final exams. The survey was collected electronically. A link to the survey was sent via SMS, where the sender's name was NITT (MAHAT in Hebrew). To increase response rates, the messages were personalized, and addressed students by their first name. The message further asked them to answer a short survey in order to help the NITT improve their study programs. To encourage survey response and completion rates, it also announced that survey respondents could participate in a lottery with the chance to win attractive prizes. The lottery announcement was translated to Arabic in order to increase Arab students' attention and response rates. Once opening the link to the survey, students could choose their preferred language for the survey, either Arabic or Hebrew. To make the tone more personal, the survey addressed each student according to their gender (as registered in the administrative data from NITT). This can matter, since both Arabic and Hebrew use different pronouns for males and females also in their plural forms. Therefore, many questions cannot be phrased in a gender neutral way, and we wanted to avoid the widespread practice of using the male forms to address all respondents. Wherever possible, we did however refrain from using gendered language.