Do Deadlines Affect Project Completion? Experimental Evidence from Israeli Vocational Colleges *

NAOMI GERSHONI† MIRI STRYJAN‡

April, 2022

Abstract

We study a large-scale intervention aimed at increasing graduation rates in Israeli vocational colleges. In this context, the main reason for low graduation rates is students' failure to complete the required final project. This may be the result of procrastination which is prevalent among students in many settings. To address procrastination, we introduced a deadline for final project defense in randomly selected departments while control group departments maintained the practice of scheduling defense dates on a rolling basis. We compare student performance over time in treated and control departments in a difference-in-differences framework and find no effect of deadlines on project defense or on graduation rates. A potential explanation for these findings is other constraints faced by students, such as academic difficulties or a low perceived value of the diploma, which are not alleviated by the deadline. Using administrative and survey data, we find that deadlines have no effect even when the alternative constraints are not binding.

Keywords: Deadlines; Higher education; Vocational Education; Field experiment; Procrastination; Educational attainment.
JEL classification: I23, C93

*We would like to thank the Israeli National Institute for Technical Training, the National Program for Practical Engineers, the Joint Distribution Committee (JDC), and the Myers-JDC-Brookdale Institute. We are grateful in particular to Tair Ifergan, Rotem Amram-Nativ, Avivit Alfasi, Ilana Lankri, Tirza Wilner, Anna Schwartz, Revital Assayg Shlush, Natali Nahir and Yitzchak Nahmani for their collaboration and practical help in the planning and implementation of the intervention and data collection. We highly appreciate the engagement and support of Ariel Tikotsky from Bar Ilan University and the Israeli government Nudgeathon team. Excellent research assistance was provided by Hadas Stiassnie, Tomer Gofman, and Tom Parizat. We also thank Amjad Mahajna and Mies Abu-Kandiill for their help with the preparation of the survey. This paper and project has benefited from feedback from Analia Schlosser, Itay Saporta-Eckstein, Shirlee Lichtman-Sadot, Amnon Maltz, Anna Sandberg and Topi Miettinen. This work was supported by the Washington University Social Sciences Lab. The experiment reported in this paper is registered in the AEA RCT registry with number AEARCTR-0007146. Declaration of interest: none.

†Department of Economics, Ben-Gurion University of the Negev, naomige@bgu.ac.il.
‡Department of Economics, Aalto University School of Business, miri.stryjan@aalto.fi.

1 Introduction

Improving outcomes and graduation rates in post-secondary education is an important policy goal. This is especially relevant in vocational education (see e.g., OECD, 2016). A growing literature in economics and psychology studies the potential of interventions that address behavioral barriers to educational success (Lavecchia et al., 2016; Clark et al., 2020). One such barrier is present-bias, which is linked to the tendency to procrastinate (O'Donoghue and Rabin, 1999, 2001) and has been well documented among students (Solomon and Rothblum, 1984; Steel and Klingsieck, 2016).[1] A *deadline* can serve as a commitment device that helps the agent overcome present-bias, thereby improving task performance (Herweg and Müller, 2011; Frederick et al., 2002). However, despite theoretically compelling arguments in favor of deadlines, few well-identified empirical studies of their effects exist to date, particularly for "real life" tasks with long-term consequences beyond the experiment. Furthermore, those studies present mixed findings regarding the effect of deadlines on task completion, which may be attributed to their relatively small and selective samples (Ariely and Wertenbroch, 2002; Bisin and Hyndman, 2020; Burger et al., 2011).

In this paper, we present results from a field experiment where deadlines were introduced to randomly selected departments from 32 vocational colleges in Israel. These vocational colleges struggle with low graduation rates and the final project, which is a key component in the vocational degree, has been identified as the main barrier to graduation (Israeli Ministry of Economy and Industry, 2017). Specifically, departments in our treatment group introduced deadlines for the *final project* defense, while control group departments continued to schedule defense dates on a rolling basis. To address possible inattention problems, the deadline treatment was accompanied by reminder messages to students in the treatment group (Ericson, 2017). Failing to meet the deadline imposed a cost by causing a delay and introducing uncertainty due to limited re-scheduling options, yet it did not completely prohibit students from graduating. Beyond this cost, the deadline could potentially affect students' behavior by exogenously framing a target date for project completion at the early stages of their work process, rather than allowing them to work at their own pace without a clear goal in mind. Other "goal setting" interventions were shown to have substantial effects on study outcomes, even in the absence of external sanctions

[1] Descriptive work from psychology documents a correlation between the tendency to procrastinate and failures in self-regulation (Steel, 2007; Steel and Klingsieck, 2016), and links the tendency to procrastinate to indecision and low self-esteem (Beswick et al., 1988).

for failing to meet the goals (see e.g., Clark et al., 2020).

Using individual level administrative data, we estimated the effect of these deadlines on students' likelihood to successfully defend their final project and to obtain a diploma, as well as on other academic achievements such as mandatory exams and project grades. We worked with a total of 89 departments within the colleges. A paired design was implemented in which departments from the same field of study with similar graduation rates and student composition prior to the intervention were matched and one department in each pair was randomly assigned to treatment. Balanced treatment and control groups would allow us to draw causal conclusions about the effect of deadlines and reminders on project completion and diploma rates. After implementation, we received more granular data on previous cohorts of students in treatment and control group departments. Based on this additional data, it became clear that despite the randomization procedure, pre-experiment final project completion as well as graduation took place earlier in the departments that were later assigned to treatment. We therefore employed a *difference-in-differences* (DID) approach building on the exogenous treatment assignment, where we compare treatment and control departments, pre- and post-intervention. We found no effect of the deadline treatment on either the likelihood or the timing of project completion (defense) or on the likelihood of receiving a diploma. We implemented several robustness tests and conducted an event study analysis to confirm that these findings can be interpreted as a null-effect.

To support our premise that students procrastinate with respect to the final project, we combine survey and administrative data, and show that there is a large gap between students' expectations and actual graduation rates. Therefore, our results could either indicate that deadlines are not a useful way to deal with procrastination or that there are additional constraints preventing students from completing their final project and graduating successfully. We consider and discuss three main alternative constraints: low expected returns to project completion and graduation, liquidity constraints, and academic constraints, and show that even when they are not binding, the deadline treatment is not effective. More specifically, we show that: (i) students believe that the final project is useful and there are significant expected labor market returns to a diploma, both among students who later on complete the project and among those who do not; (ii) the estimated treatment effect is small and insignificant even for students from areas that rank higher in socioeconomic terms; and (iii) the deadlines appear to have no effect on students with better exam performance, or in colleges that have higher teaching quality scores.

This work contributes to the growing body of work in the economics of education that tests and evaluates policies designed to overcome behavioral barriers in education (for surveys of this literature see Lavecchia et al., 2016; Damgaard and Nielsen, 2018; Koch et al., 2015). This literature is concerned with identifying policies that are both effective and scalable. We specifically contribute to recent work that evaluates policies designed to address present biased behavior among students. In a field experiment with German university students, Himmler et al. (2019) find that inducing students to set soft (non-binding) commitments to stay on track for graduation led to better attendance and study results on exams, especially for students with a tendency to procrastinate. Clark et al. (2020) show that requiring college students to set goals about a specific task to complete (online practice exams) led to positive effects on both task completion and grades. Other related work includes studies in social psychology and economics that evaluate more general goal setting programs (e.g. Morisano et al., 2010; Schippers et al., 2015; Dobronyi et al., 2019) and coaching programs for students (Oreopoulos et al., 2020).

Our study also complements experimental work in behavioral economics which focuses on the effect of deadlines. It is well documented that people impose costly deadlines to deal with their own procrastination tendencies (Ariely and Wertenbroch, 2002; Augenblick et al., 2015; Bisin and Hyndman, 2020; Burger et al., 2011). In an influential study, Ariely and Wertenbroch (2002) found that exogenously set deadlines increased the timely submission of class projects among two classes of MBA students, compared to deadlines set by the students themselves. Meanwhile, the results of two more recent framed field experiments with students were less positive. Neither exogenously given nor self-imposed deadlines led to higher task completion as compared to a setup with no deadlines (Bisin and Hyndman, 2020), whereas a negative effect was reported for evenly spaced interim deadlines (Burger et al., 2011).

Relative to this work, our main contributions are twofold. First, while other interventions that were recently tested in "real life" educational settings deploy new infrastructure and training aimed at changing students' mindset or internal commitment strategies, our intervention merely introduces a change in schedule within the existing curriculum. Such an exogenous change in rules is relatively easy to implement and scale up. As it relies less on students' cooperation and active take-up than previously studied interventions, it is also likely to be more successful in a diverse student population. Second, compared to existing work on deadlines, this study allows us to draw more general conclusions because it is implemented at a national scale, for a task that is part of the standard curriculum requirements for graduation and is thus associated with

high stakes.[2]

By focusing on vocational post-secondary education, our work also contributes to the small but growing literature on vocational education. Rapid technological advancement and automation is expected to increase the need for workers with general skills who can adapt to changing conditions and job requirements (Autor, 2015; Malamud and Pop-Eleches, 2010; Golsteyn and Stenberg, 2017; Hall, 2016; Hanushek et al., 2017; Hampf and Woessmann, 2017). Formal vocational and technical training has the potential to increase the prevalence of such skills and its quality has thus become a pressing policy matter (Nedelkoska and Quintini, 2018; Brunello and Rocco, 2017).

Our results suggest that deadlines are not a straightforward policy tool to increase completion rates among vocational education students. Given the limited empirical literature on the causal effect of deadlines in education, these findings contribute usefully toward our understanding of the applicability of deadlines. More generally, our work adds to recent studies that report null effects and aim to identify when behavioral interventions are less effective (see e.g., Kristal and Whillans, 2020; Oreopoulos and Petronijevic, 2019).

The rest of the paper is organized as follows: In section 2, we describe the context of our study, and the intervention. In section 3, we describe the different sources of data and our empirical approach, and present summary statistics. Thereafter, we present our results in section 4, and section 5 concludes the paper.

2 Context and intervention

2.1 Context

We focus on post-secondary technological education tracks in Israel's system of technical colleges which, in addition to specific vocational skills, similarly to academic engineering studies, provide broader knowledge and expertise. These tracks award official practical-engineering diplomas and, in some majors, also licensure. Colleges offer both full-time studies that typically take two years to complete (morning track) and evening studies that target the working

[2]Previous experiments on the effects of deadlines are concerned with measuring the demand for and response to deadlines in limited student samples, often with tasks that are imposed as part of the research design and with incentives in the form of relatively modest monetary rewards (Burger et al., 2011; Bisin and Hyndman, 2020) or a better course grade Ariely and Wertenbroch (2002).

population and typically take three years (evening track). We will therefore label cohorts by the year they are supposed to *complete* their mandatory coursework and exams. Figure 1 presents the typical timeline for the final semester of studies and the following months when students take their certification exams and defend their final project.[3]

The Israeli National Institute for Technical Training (NITT), which operates within the Ministry of Labor, Social Affairs, and Social Services, sets the criteria and regulates these tracks, covering over 70 colleges and more than 30 fields of study. Each year, approximately 10,000 students begin their first year of practical engineering training. The government regulates tuition and subsidizes the cost of this training, as is also the case for academic education in Israel. We focus our intervention on the four largest majors: architecture, electrical engineering, civil engineering, and software programming which together constitute more than half of a typical cohort of students. Licensure, which is subject to graduation, is required for the former three occupations, and therefore we expected these students to have higher incentives to graduate with an official diploma. Our unit of treatment is college by major units, which we refer to as "departments." The average rate of timely graduation in these departments for the pre-intervention cohort was around 56%, considerably lower than that for Israeli university students where the timely graduation rate was 72.4% (Feniger and Ayalon, 2016).

To graduate with a diploma, students are required to (i) complete at least 2,176 academic hours (with some variation across study fields); (ii) pass the standardized national certification exams in their field of study; and (iii) submit a final project which demonstrates their acquired knowledge and requires the application of practical skills. NITT administers and grades both the national exams and the final project defense. While exam dates are set unilaterally by the NITT, defense dates are coordinated on a rolling basis between the NITT and the departments. While the timing is flexible, defense dates are set for groups of students (usually ten students) since the required committee of reviewers cannot be summoned for each individual student. Therefore, departments will wait until enough students are ready to submit their project and then ask the NITT to schedule a session of consecutive defenses. It should also be noted that while students are expected to submit and defend their project within a few months after completing their course work, legally, they are permitted to do this up to seven years after their coursework is completed.

[3]The academic year in Israel is based on the Jewish calendar and holidays, and the exact Gregorian dates of the semesters may vary substantially by year. The academic year typically starts in October.

Our intervention focuses on the final project which is a major barrier to graduation. Table 2 shows the relationship between completing mandatory exams and defending the final project for the 2018 cohort that was unaffected by our intervention. As seen, the share that completed the project but not the exams is only 4.7% while the share of students who passed their exams but failed to complete the project is substantially higher at 37%.

2.2 Intervention

Our intervention was implemented in close collaboration with NITT, and its main component was the introduction of deadlines. For a randomly selected set of departments, NITT and the research team instructed the department heads to set strict deadlines for the submission and defense of the final project and to ensure that students are informed about these set dates by their departments well in advance. To enforce deadlines, treated units are notified that students who fail to meet the deadline would have to wait *at least* two months to reschedule a defense. Moreover, as defenses are not scheduled on an individual basis, a missed deadline causes the student considerable uncertainty about the next chance to defend the project. While we wanted to impose a cost for not meeting the deadline, completely preventing students from submitting after the deadline could discourage students and work against the aim of increasing graduation rates. We also intended for the deadline to serve as a behavioral commitment device by framing a specific goal in terms of the timing of project completion. Such goals can have an impact regardless of the actual cost of failing to meet the goal (see e.g., Himmler et al., 2019). In addition, to make the intervention more salient, students in treated departments receive personal reminders and encouragement messages via SMS.[4] Control units maintain NITT's common practice of scheduling defense dates on a rolling basis as projects are submitted.

The intervention was implemented for the chosen majors in the 2019 cohort (i.e. the cohort scheduled to complete exams in fall 2019; we will also refer to this cohort as the experimental cohort). After randomization, in January 2019, the heads of the treatment group departments were invited to a meeting in which the research team and the NITT administrator of final projects presented the goals of our joint work and described the potential of behavioral interventions, and in particular, deadlines, to improve educational outcomes. Graduation rates are an

[4] Ericson (2017) shows the relevance of combining deadlines with (unanticipated) reminders in order to address inattention, and Bettinger et al. (2019) show that for US high school students personalized text message reminders can decrease drop-out rates.

important factor in the funding formula of the colleges, and department heads are aware of this. Therefore, the aim of the meeting was to convince department heads that deadlines could be an efficient way to improve these rates and to motivate them to communicate the deadlines to the students effectively. Several months later, during July-August 2019, the same department heads were contacted by the NITT team and were asked to coordinate defense dates and relay them promptly to their students. Pre-determined deadlines were scheduled within the time interval between October 2019 and February 2020. SMS reminders were sent to students in the treated departments once a month from August 2019 until their scheduled defense date.

A second round of the intervention was implemented for the same majors in the 2020 cohort (with the same departments assigned to treatment and control) in which the deadline component was maintained, but reminders were not sent. Due to the severity of the Covid-19 crisis during the period of the planned deadlines for the 2020 cohort, it was difficult to verify whether the deadlines were kept or even communicated to the students according to plan. Our analysis will therefore focus on the 2019 cohort, but we will show that our main findings are robust also to including the 2020 cohort.

3 Data and empirical approach

3.1 Data

We combine data from several sources to estimate the impact of the intervention and to analyze potential mechanisms. The main part of the analysis, as well as the randomization, are based on college and individual level administrative data.

Our individual level data covers five cohorts of students all of whom completed their last year of course work between 2017-2020 in one of the four majors. The individual level records provide basic demographic characteristics (gender, year of birth and locality of residence) as well as detailed information on their studies track, registration date, test grades and dates, final project grade and defense date, and graduation (diploma) date. Missing values for grades and test or defense dates indicate that the student did not take the specific test or did not yet defend the final project. Diplomas are recorded with a two months delay and therefore when we address diploma outcomes, we observe a shorter time frame than for defense outcomes. We match these data with a socioeconomic Z-score calculated by the Israeli Central Bureau of Statistics at the

locality of residence level.

At the college level, the records include the different majors and tracks (morning/evening) offered by each college, an indicator for whether the college specifically focuses on the Arab population (teaching mostly in Arabic, located in an Arab town, etc.), an indicator for the college being part of an academic institution, and a 7 year average graduation rate for each specific studies track within each college.[5] This information is supplemented by survey-based rankings of the colleges and of each specific department within the college which are published by the NITT and measure student satisfaction along several dimensions. For details on this measure, see Section A.2.

In addition, in September 2020, we conducted an online survey of students from the 2020 cohort in order to gain some insight on students' beliefs regarding the importance of the final project and the diploma. The survey was administered in collaboration with the NITT and timed right after the final exams period. At the time of the survey, the vast majority of the students were expected to be working on their final projects (only 11.4% of the students had already defended their project when the survey was launched). The predetermined deadlines were coordinated and communicated to students after the survey was completed. Students were contacted through personalized SMS messages and answered the survey online via their phones. The response rate (after two reminders) was 37.8% out of 4,094 students. For additional details on the survey methodology, see section A.2.

The survey included questions about demographic characteristics, living arrangements, and past work experience in general and in their field of study in particular. In addition, students were asked to estimate expected own earnings with or without an official diploma, one and five years after completing their studies. They were not asked about their earnings immediately after graduation in order to diminish the effects of the Covid-19 situation on their labor market expectations.[6] To capture a more objective view on the returns to a diploma, respondents were also asked to estimate the expected earnings of a *typical* student in these different scenarios. We further asked about their attitude towards the final project to understand if they see any benefits in this task beyond the instrumental value of meeting the diploma requirements.[7] Lastly, we

[5]NITT funds colleges according to a formula that is partially based on these rates.

[6]When the survey was launched it seemed reasonable to expect the economy to recover and labor market conditions to return to a pre-pandemic situation over the next 12 months. The survey was launched after a period of steady caseload but infections began to rapidly increase by the time we sent the last round of reminders.

[7]The specific statements were (1) The final project is an unnecessary hurdle for obtaining a diploma. (2) The final project provides an opportunity to apply the knowledge and skills that I have acquired during my studies. (3)

asked the students to choose the statement that best describes their chance of graduating with a diploma, from "very likely" to "no chance".

3.2 Randomization

The main experiment was designed and implemented on the 2019 cohort. Treatment was assigned using pair-wise randomization (Bruhn and McKenzie, 2009) in which departments from the same major were matched based on previous rates of graduation with a diploma (7-year averages calculated by the NITT), college target population (whether it specifically targets the Arab population), and number of students in the department.[8] After this procedure, the matched pairs were reviewed and approved by the NITT team to verify their similarity on additional dimensions which cannot be measured directly (e.g. management quality). Overall, 46 departments ($college \times majors$ units) were assigned to treatment, while 43 serve as control.

As explained above, we also planned and implemented a second round on the 2020 cohort which was interrupted by the Covid-19 pandemic. To avoid contamination, we used the same allocation of departments to treatment and control as in the first round, and did not reveal our preliminary results to any of the departments (we note that departments can voluntarily implement deadlines even if they are defined as control).

3.3 Empirical strategy

3.3.1 Sample, summary statistics and balance check

At the time of randomization, we only had access to seven-year averages of one of our two main outcome variables, namely diploma rates, and we used that information to create comparable pairs. This measure was used in the matching procedure and was balanced across treatment and control units. In 2021, after both rounds of the experiment were carried out, we received more granular individual level administrative data for the three cohorts that preceded the experimental cohort. Based on these data, as shown in Figure 2, we discovered that the *timing* of project defense and of graduation were not balanced between treatment and control units pre-intervention.

The final project provides an opportunity to connect with potential employers. (4) The final project is an academic task that has no relevance to my future career. (5) The final project provides an opportunity to signal my ability to potential employers.

[8]For three of the majors there was an uneven number of departments and thus we formed a triplet. In each of the three resulting triplets, two departments were assigned to treatment and one to the control condition.

Therefore, to avoid bias, we will estimate the treatment effect with a DID approach, utilizing the data on pre-treatment cohorts.

To efficiently implement this approach, the main sample that will be used includes a sub-set of the 89 departments that participated in our experiment which had students for three consecutive cohorts (the last being the experimental cohort). Moreover, we exclude departments whose paired unit does not meet this requirement, leaving us with 62 department and 31 matched pairs. This balanced panel of pairs includes data on 11,130 students from three cohorts (2017-2019).

Table 1 presents summary statistics and balance tests for the 2019 cohort (the experimental cohort) in this sample. The table reports the control-group mean in column (1) and in column (2), the difference between treatment and control with the corresponding standard errors in parentheses. All differences are quite small and none are statistically significant indicating that the sample is indeed balanced on these characteristics, including graduation rates. Panel A shows that students in technical colleges are relatively old (the average age is 27), mostly male (76.9%), and come from below average socioeconomic backgrounds. It is also important to note that the age variation is high and ranges between 19 and 69 (44 is the 95th percentile). Panel B presents department-level characteristics. The departments in our sample have on average 61 students in the 2019 cohort; their average 7-year graduation rate is slightly below 50%; a third of these departments are located in a technical college within an academic institution; almost 80% offer evening tracks intended to serve working students; and 3.3% focus on Arab students.

To further understand the average characteristics of the studied population, Appendix Table A1 reports additional student characteristics based on the 2020 cohort survey. In this sample of students, more than 30% are either first or second generation immigrants (have at least one parent who is not native born), and more than 70% have fathers with high school level education or less.[9] Approximately 40% were married or cohabiting and almost 30% had children, yet close to 50% were still living with their parents (or with other older relatives). These facts point out the diversity of the studied population (as also demonstrated by the age distribution) — some are independent and provide for themselves and their families, while others still depend on their parents. Nonetheless, most of the students (93%) had previous work experience, and 33% worked in their study field before their studies. As expected, since most colleges provide

[9]The share is similar for mothers. For reference, according to OECD statistics, more than 50% of the working age population in Israel have tertiary education. While the parents of students in our sample are probably older than this group, tertiary education rates in Israel were also considerably higher two decades ago (42% in 2002) than for the average parent of our respondents.

evening tracks, a large share of students also reported working *during* their studies (86% overall and 30% specifically in their study field).

To test the validity of the design and the robustness of our results, we also use a secondary sample which includes more cohorts and departments but does not maintain the paired structure. In this sample, we include 84 departments that appear in at least one of the three pre-treatment cohorts 2016-2018 and also in the experimental cohort (post-treatment). Descriptive statistics and balance tests for the secondary sample are presented in Appendix Table A2 demonstrating that the treatment and control students are balanced for this sample as well.

3.3.2 Estimation

The red lines in Figure 2 present the cumulative rates of defense (panel (a)) and graduation with a diploma (panel (b)) over time for the 2017 and the 2018 cohorts which preceded the experimental cohort. These rates are shown separately for the departments which were randomly selected to our treatment group and for their control group counterparts. Despite the fact that these two cohorts of students were not subject to any treatment or experiment, the figures clearly show that both the rate of defense and the rate of graduation were substantially higher for the treatment group departments over the first 12 months since the beginning of the students' last semester. After this period, the rates in the two groups converge and remain very similar (although the treatment group still has slightly higher graduation rates after 27 months).

As a result, we cannot reliably estimate treatment effects based on a simple comparison of the treatment and control group outcomes as we initially intended. Instead, we build on the random assignment of treatment and compare treatment and control department using a DID approach. This approach does not require treatment and control groups to be identical or similar in their characteristics and pre-treatment outcomes. Instead, the main identifying assumption is that treated departments would follow the same (or a very similar) trend as control departments in the absence of treatment. A common way to assess the plausibility of this assumption is to compare pre-trends in the outcome variable, assuming that if trends were parallel pre-treatment they are expected to remain parallel in the post-treatment period if it were not for treatment. This latter assumption could be violated if, for example, the composition of the treatment and control groups would differentially change over time.

To implement this empirical approach we use the main sample of 62 paired departments

which were active (had students) of three consecutive cohorts: 2017-2018 are the pre-treatment cohorts and 2019 is the post-treatment cohort. This balanced panel of departments allows us to estimate the following DID specification:

$$Y_{idpc} = \beta_0 Treated_d + \beta_1 Treated_d \times Post_c + \gamma' X_{id} + \delta_p + \eta_c + \varepsilon_{idpc} \tag{1}$$

where y is the outcome of student i in department ($college \times majors$) d of pair p and cohort c. Outcomes are always measured for a specific number of months after the first month of each cohort's last semester (February of the relevant cohort year), and we are able to follow students for up to 29 months (our data goes up to June 2021). $Treated_d$ is an indicator for treatment (determined at the department level) and $Post_c$ indicates that the student belongs to the experimental cohort (2019 in the main analysis and 2019-2020 in some robustness checks). X_{id} is a vector of individual and department level controls (presented in Table 1) which are added in some of our estimations. This specification also includes cohort fixed effects (which are collinear with the omitted post indicator) and pair fixed effects due to the nature of the randomization procedure. β_1 is the coefficient of interest in our estimation since it measures the effect of the intervention on outcomes within matched pairs. Standard errors are clustered at the department ($college \times major$) level, which is our unit of randomization.

In column (3) of Table 1 we report estimates of β_1 from a set of regressions where the outcome variable is one of the cohort-varying control variables. The magnitude of the point estimates and their insignificance confirm that treatment and control departments do not change differentially over cohorts, supporting the plausibility of the parallel-trends assumption. Corresponding estimates are reported for the secondary sample in Appendix Table A2 where only one of the characteristics, namely the number of students in the department, shows a differential, positive trend over time. We will use this sample and estimate the same DID specification to test the robustness of the main results. Although the sample is larger and covers more cohorts than our main sample, it has a substantial disadvantage in that we cannot include pair fixed effects when pairs do not consistently appear across cohorts. To increase the precision of our estimates and to control for unobserved, constant department characteristics, we replace the pair fixed effects with department fixed effects.

To further establish that pre-trends are parallel and to test the robustness of our results, we will also estimate the following event study specification with this extended sample, which will

allow us to look at three pre-treatment periods:

$$(2) \quad Y_{idc} = \alpha_1 Treated_i \times D_{2016} + \alpha_2 Treated_i \times D_{2017} + \alpha_3 Treated_i \times D_{2019} + \zeta_c + \omega_d + \varepsilon_{idt}$$

where ζ_c and ω_d are cohort and department fixed effects (respectively), D_{year} are cohort indicators which are interacted with the treatment indicator, and 2018 is the omitted cohort (the cohort that precedes the experimental cohort). Therefore, α_1, α_2 and α_3 are interpreted as the 2016, 2017, and 2019 differences between treatment and control relative to the baseline difference in the 2018 cohort.

Additional robustness tests incorporate the 2020 cohort which received the deadline treatment without reminders and during the acute stages of the Covid-19 pandemic into the analysis. In this analysis, the pooled 2019 and 2020 cohorts constitute the post-intervention sample, yet student outcomes can only be followed for up to 19 months from the beginning of the final semester.

4 Results

4.1 Main results

Figure 2 presents the raw data, cumulative defense and graduation rates pre- and post-treatment. The outcomes for the pre-period cohorts are indicated by the red lines, demonstrating that students in treated departments tended to defend their final projects and graduate earlier than students in the control departments even before our intervention. The dark blue lines present the same outcomes for the experimental cohort. The lines for the pre- and post-periods practically converge up to the thirteenth month (marked by the vertical line) for both treatment and control. After that, both lines flatten out completely for the 2019 cohort during the stringent lock down imposed by the government due to the Covid-19 outbreak. During this period all project defenses were suspended. When the lock down was eased, defenses were rescheduled and the 2019 defense rates continued to increase but remained lower than previous cohorts until the 29th month. Regardless of the overall decrease in project defense and graduation rates which was most likely caused by the Covid-19 shock, the difference between treatment and control is very similar between the pre- and post treatment cohorts. The fact that the differences are

very similar across cohorts indicates that the intervention did not have an impact on student outcomes.

These differences are formally compared through the estimation of equation 1, which also controls for pair and cohort fixed effects. Figure 3 presents by month, the results of this estimation — point estimates and 95% confidence intervals for the coefficient $\beta 1$. Both panels clearly show that we cannot reject the null-hypothesis that the intervention did not change defense or graduation rates at any point in time, namely that neither the rate nor the timing of project defense and graduation were affected. Moreover, the point estimates are very close to zero and remain quite constant over time which we interpret as strong evidence of a null effect of the intervention.

The results remain practically identical when we add control variables to the specification (Figure A3) or when we replace the pair fixed effects with department ($college \times major$) fixed effects (Figure A2). We also estimate the latter specification using our secondary sample which includes more departments and cohorts but where the pairs are not necessarily maintained. Figure A5 presents these results, and confirms the robustness of the main findings.

We then use this extended sample to estimate the event study specification (equation 2) because it includes more cohorts pre-treatment, thus allowing us to test whether trends in defense or graduation rates were parallel between treatment and control in the pre-period. In this estimation we look at defense or graduation for the last available month in our data (29th and 27th, respectively). The results are presented in Figure A1 and clearly confirm that the difference between treatment and control departments did not change over cohorts either pre- or post-treatment.

As an additional robustness test, in Figure A4, we present raw data figures and regression results for a third sample where we pool the experimental 2019 cohort with the 2020 cohort, for which the deadline intervention was implemented during the acute stage of the Covid-19 pandemic. This sample includes the same 62 departments as our main sample but we are only able to follow students' defense status for 19 months and graduation for 17 months. Nonetheless, this analysis further confirms that the gaps between treatment and control department remain similar over cohorts regardless of the imposed deadlines.

After establishing that the rate of project defense and its timing was not affected by the deadline intervention, we verify that the intervention did not have any unintended effects on other outcomes. Specifically, it is possible that while the deadlines were not efficient in increasing

project defense rates, they affected the quality of the projects or students' exam performance. We already partially address this question by showing that graduation rates were not affected (since graduation requires both a successful project defense and passing the mandatory exams). In Table 3, we show more directly that students' project grades (conditional on defending the project) and their achievements in the mandatory exams did not change differentially post-treatment. We estimate equation 1 for three outcome variables: project grade, exam grade point average (GPA)[10] and fraction of mandatory exams passed.

Overall our analysis presents consistent and robust evidence that the deadline intervention had a null-effect on student achievements and graduation rates.

4.2 Descriptive evidence on potential barriers to graduation

The deadline treatment was designed to address *procrastination*. Evidence that students do procrastinate when it comes to the final project and graduation can be observed in our survey of the 2020 cohort which was conducted when most students were at the early stages of work on their final project. We find a substantial gap between students' expectations at that point in time, and their actual graduation rates as observed in our administrative data. While 74% of the respondents stated that there is a high or very high likelihood that they will graduate and complete the final project, only 41.78% of them actually graduated on time (within six months from the survey). The reported expectations are also high relative to graduation rates in previous cohorts, which can be followed for a longer time period (below 60%).

Although this evidence is suggestive of a tendency to procrastinate, the deadlines that we introduced did not affect student outcomes. One interpretation of these findings is that this kind of deadlines cannot solve students' procrastination problems. An alternative explanation would be that students in our setting face other binding constraints that are not addressed by the deadlines intervention. If this is the case, a null effect may be found even if deadlines decrease procrastination. In this subsection, we present evidence on the main alternative constraints, namely low expected returns to graduation, liquidity constraints, and academic constraints, and conclude that they are not likely to explain the null results.

[10]The grade point average (GPA) is computed over the mandatory exams in each student's study field. Students receive a grade on a scale of 1 to 100 points. The threshold for passing an exam is 55 points. There are students that did not take some or any of the mandatory exams. In light of the fact that this is equivalent to failing, when we calculate their GPA, these students are assigned with a grade equal to the median grade among students who took the exam and failed. Thereby we avoid overestimating their exam performance by conditioning the sample on actually taking the exam.

4.2.1 Beliefs about returns

Students will not complete the final project if they believe that the returns will be low. "Returns to the project" can be thought of as whether the student believes that they will acquire useful and/or valued skills for the labor market. In our survey of the 2020 cohort, we asked students to choose one out of five statements, three positive and two negative about the final project that was closest to their own opinion. 76.56% of the respondents stated either that the project provides an opportunity to apply the knowledge they acquired during their studies (64%) or that it is a way to signal their abilities to or connect with potential employers (12%).[11] When we limit the sample to students who have not yet defended their final project, the share of students with a positive attitude is very similar (74.54%).[12]

Students may also choose not to complete the project because they believe the labor market returns to a diploma to be low, in terms of expected earnings and their progression over time. To examine this hypothesis, we presented the students with hypothetical scenarios where students were asked to estimate what would be their salary level after one and five years, with and without a diploma. The results in Figures 4 (a) and (b) present kernel density plots and show that the distribution of expected wages *with* a diploma is shifted to the right compared to expected wages without diploma, both one and five years after study completion. Furthermore, Kolmogorov-Smirnoff tests confirm that the two distributions are different. In other words, students in our population believe that there are positive returns to a diploma, both in the short and medium run. The expected returns to a diploma are also very similar when restricting our data to students who did not complete the final project or when using estimates of the expected returns for a typical student. Taken together it would seem that the lack of treatment effect is not due to the fact that students do not value the diploma.

4.2.2 Liquidity constraints

Regardless of their beliefs about the returns to graduation, students may choose to enter the labor market before they graduate due to liquidity constraints. This is likely to hold relevance

[11] The two more negative statements refer to the project being an unnecessary hurdle or to it being an academic exercise with no practical implications.

[12] The survey was conducted in September 2020, right after the students took their final exams so only 11% of the respondents already completed their projects, out of which 92.27% report positive attitudes. This could be explained by selection but may also indicate that the students gain a positive experience when they actually work on the project which leads them to change their initial expectations.

in our context, where many students come from economically disadvantaged backgrounds, are relatively old, and have financial responsibilities. If students start working full-time, they will not have time to work on their final projects in which case a deadline cannot assist. Therefore, we might expect to see differential treatment effects for students from different socio-economic standings. In columns 1-2 of Table 5 we present results separately for students that live in a locality ranked below the median and for those whose locality of residence is at or above the median socioeconomic score of the Israeli Central Bureau of Statistics. Our results do not indicate any effect of the deadlines for either subgroup, and in particular the estimated coefficient on $Treated \times Post$ is negative and small for students from richer areas, where it is much less likely that students are financially constrained. This implies that liquidity constraints cannot explain our main results.

4.2.3 Academic constraints

An additional constraint to project completion and graduation is that some students may simply lack the academic knowledge needed to complete the project. However, this is partially ruled out in Table 2 which shows that among students who did not defend their project, 60% did pass all the mandatory exams, indicating that they were able to complete other high stake academic requirements. In line with this, when we divide students by whether they had passed the mandatory exams by March the year after the first round of exams (see columns 3-4 of Table 5), the estimated treatment effect is negative and small. Finally, we exploit college level data from an online survey administered by the NITT in 2018 where students anonymously ranked the quality of teaching in their department. Results are presented in columns 5-6 of Table 5 and show that there is no differential treatment effect for students in departments with high teaching scores.[13] Overall, we find no indication of a positive treatment effect even among the subset of students who are not academically constrained.

5 Conclusion

In this paper, we estimate the effect of a deadline policy in vocational education on task completion and graduation rates. Specifically, we designed a field experiment where a deadline was

[13] We also did not find differential effects based on other measures of students' satisfaction, such as satisfaction with the college facilities, or students' overall level of satisfaction with their college.

imposed for a lengthy and high stakes task: the "final project" in randomly selected departments of Israeli vocational colleges. To the best of our knowledge, this is the first study to evaluate the effect of a large scale deadline policy in the educational sector. Employing a DID strategy in which we compare treatment and control group departments pre- and post-intervention, we find no effect of the deadline policy on project defense rates, timing, or diploma rates. We present several robustness checks that support the interpretation that this finding is a null effect of the diploma policy.

The deadline treatment is designed to address one possible constraint to project completion, highlighted in the behavioral economics and psychology literature: *procrastination*. Since we do not find any effect of our treatment on task completion, one interpretation would be that procrastination is not a concern in the student population we study. A more nuanced interpretation is that there are other and more binding barriers to project completion, in particular that students may have low expected returns to the project, or face binding liquidity or academic constraints. We present suggestive evidence that such explanations are not the main cause of weak response to the deadline treatment.

Our results are somewhat surprising given the theoretical literature on time inconsistency (see e.g. Frederick et al., 2002) combined with the widespread notion of procrastination in academia. Deadlines are also widely believed to be effective, as exemplified by reports from the grant giving organizations ESRC and the NSF that removing deadlines had led them to receive fewer submitted grant applications.[14]

While our findings are robust and based on a large national sample they should not be taken to suggest that deadlines never work. Rather, they contribute to our understanding of how and when deadlines in educational settings can be applied in an effective manner. Two aspects that are likely to be important for the efficiency of deadlines are their target population and the specific setup and consequences of the deadline. The vocational college population we study differs from the samples of competetive, high performers for whom deadlines were reported to be effective (e.g. the Executive MBA students studied by Ariely and Wertenbroch (2002) or the academic research community that submit grant applications to the ESRC and the NSF). If this population is less likely to internalise the reframing of targets associated with the deadline,

[14]The example from the NSF can be found at https://www.science.org/news/2016/04/no-pressure-nsf-test-finds-eliminating-deadlines-halves-number-grant-proposals (accessed on January 17, 2022). The example from the ESRC is discussed in, and based on informal communication with the article authors of Frederick et al. (2002).

higher penalties may be required in order for the deadline to have an effect. The optimal level of such sanctions remains a subject for further research, since they can also have a discouraging effect on task completion.

References

Ariely, D. and K. Wertenbroch (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological science 13*(3), 219–224.

Augenblick, N., M. Niederle, and C. Sprenger (2015). Working over time: Dynamic inconsistency in real effort tasks. *The Quarterly Journal of Economics 130*(3), 1067–1115.

Autor, D. H. (2015). Why are there still so many jobs? the history and future of workplace automation. *The Journal of Economic Perspectives 29*(3), 3–30.

Beswick, G., E. D. Rothblum, and L. Mann (1988). Psychological antecedents of student procrastination. *Australian psychologist 23*(2), 207–217.

Bettinger, E. P., B. L. Castleman, and Z. Mabel (2019). Finishing the last lap: Experimental evidence on strategies to increase college completion for students at risk of late departure.

Bisin, A. and K. Hyndman (2020). Present-bias, procrastination and deadlines in a field experiment. *Games and economic behavior 119*, 339–357.

Bruhn, M. and D. McKenzie (2009). In pursuit of balance: Randomization in practice in development field experiments. *American economic journal: applied economics 1*(4), 200–232.

Brunello, G. and L. Rocco (2017). The effects of vocational education on adult skills, employment and wages: What can we learn from piaac? *SERIEs 8*(4), 315–343.

Burger, N., G. Charness, and J. Lynham (2011). Field and online experiments on self-control. *Journal of Economic Behavior & Organization 77*(3), 393–404.

Clark, D., D. Gill, V. Prowse, and M. Rush (2020). Using goals to motivate college students: Theory and evidence from field experiments. *Review of Economics and Statistics 102*(4), 648–663.

Damgaard, M. T. and H. S. Nielsen (2018). Nudging in education. *Economics of Education Review 64*, 313–342.

Dobronyi, C. R., P. Oreopoulos, and U. Petronijevic (2019). Goal setting, academic reminders, and college success: A large-scale field experiment. *Journal of Research on Educational Effectiveness 12*(1), 38–66.

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*(3), 692–719.

Feniger, Y. and H. Ayalon (2016). Employment and satisfaction of practical engineers from mahat, 2-3 years after the graduation (2013-2014 graduates). In Hebrew.

Frederick, S., G. Loewenstein, and T. O'donoghue (2002). Time discounting and time preference: A critical review. *Journal of economic literature 40*(2), 351–401.

Golsteyn, B. H. and A. Stenberg (2017). Earnings over the life course: General versus vocational education. *Journal of Human Capital 11*(2), 167–212.

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.

Hampf, F. and L. Woessmann (2017). Vocational vs. general education and employment over the life cycle: New evidence from piaac. *CESifo Economic Studies 63*(3), 255–269.

Hanushek, E. A., G. Schwerdt, L. Woessmann, and L. Zhang (2017). General education, vocational education, and labor-market outcomes over the lifecycle. *Journal of human resources 52*(1), 48–87.

Herweg, F. and D. Müller (2011). Performance of procrastinators: on the value of deadlines. *Theory and Decision 70*(3), 329–366.

Himmler, O., R. Jäckle, and P. Weinschenk (2019). Soft commitments, reminders, and academic performance. *American Economic Journal: Applied Economics 11*(2), 114–42.

Israeli Ministry of Economy and Industry (2017). Employment and satisfaction of practical engineers from mahat, 2-3 years after the graduation (2013-2014 graduates). In Hebrew.

Koch, A., J. Nafziger, and H. S. Nielsen (2015). Behavioral economics of education. *Journal of Economic Behavior & Organization 115*, 3–17.

Kristal, A. S. and A. V. Whillans (2020). What we can learn from five naturalistic field experiments that failed to shift commuter behaviour. *Nature Human Behaviour 4*(2), 169–176.

Lavecchia, A. M., H. Liu, and P. Oreopoulos (2016). Behavioral economics of education: Progress and possibilities. In *Handbook of the Economics of Education*, Volume 5, pp. 1–74. Elsevier.

Malamud, O. and C. Pop-Eleches (2010). General education versus vocational training: Evidence from an economy in transition. *The review of economics and statistics 92*(1), 43–60.

Morisano, D., J. B. Hirsh, J. B. Peterson, R. O. Pihl, and B. M. Shore (2010). Setting, elaborating, and reflecting on personal goals improves academic performance. *Journal of applied psychology 95*(2), 255.

Nedelkoska, L. and G. Quintini (2018). Automation, skills use and training. OECD Social, Employment and Migration Working Papers.

O'Donoghue, T. and M. Rabin (1999). Doing it now or later. *American economic review 89*(1), 103–124.

O'Donoghue, T. and M. Rabin (2001). Choice and procrastination. *The Quarterly Journal of Economics 116*(1), 121–160.

OECD (2016). Education at a glance 2016: Oecd indicators.

Oreopoulos, P. and U. Petronijevic (2019). The remarkable unresponsiveness of college students to nudging and what we can learn from it. Technical report, National Bureau of Economic Research.

Oreopoulos, P., U. Petronijevic, C. Logel, and G. Beattie (2020). Improving non-academic student outcomes using online and text-message coaching. *Journal of Economic Behavior & Organization 171*, 342–360.

Schippers, M. C., A. W. Scheepers, and J. B. Peterson (2015). A scalable goal-setting intervention closes both the gender and ethnic minority achievement gap. *Palgrave Communications 1*(1), 1–12.

Solomon, L. J. and E. D. Rothblum (1984). Academic procrastination: frequency and cognitive-behavioral correlates. *Journal of counseling psychology 31*(4), 503.

Steel, P. (2007). The nature of procrastination: a meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological bulletin 133*(1), 65.

Steel, P. and K. B. Klingsieck (2016). Academic procrastination: Psychological antecedents revisited. *Australian Psychologist 51*(1), 36–46.

6 Tables and Figures

Table 1: Summary Statistics and Balance Tests

	Treatment cohort			All cohorts	
Variable	Control group mean	Diff. T-C	Obs.	DID coeff.	Obs.
Panel A - Individual level					
Age (years)	26.96	0.493	4,064	0.129	11,130
		(0.389)		(0.333)	
Female	0.231	0.005	4,064	0.006	11,130
		(0.009)		(0.011)	
Socioeconomic Z-score	-0.045	0.092	4,046	-0.027	11,090
		(0.086)		(0.039)	
Panel B - Department level					
# of students in department	60.93	6.293	62	0.932	186
		(7.655)		(4.101)	
Graduation rate (%)	49.35	2.138	62		
		(1.826)			
Within Academic institution	0.333	0.087	62		
		(0.133)			
Focus on Arab population	0.033	0.022	62		
		(0.030)			
Evening track	0.767	0.000	62		
		(0.103)			
# of majors offered by college	3.2	-0.163	62		
		(0.178)			

Notes: Columns 1-3 of the table report summary statistics for key variables in the administrative data at the student level (Panel A) and department level (Panel B), divided by treatment status of the department in the experimental cohort. Column 2 reports the results of a DID estimation (equation 1) for the subset of background variables for which there is variation across years. The variables which are constant over cohorts are: Graduation rate average taken over the seven years prior to our intervention, a dummy for whether the department belongs to a college that is part of an academic institution, a dummy for whether the department belongs to a college that focuses on the Arab population, and a dummy for the department having an evening track. * p<0.1, ** p<0.05, *** p<0.01.

Table 2: Graduation, Exams and Defense Status

	Graduated	Did not graduate		
	Exams Defense	Exams No Defense	No Exams No Defense	No Exams Defense
Number	1,652	1,774	1,187	228
Share	0.341	0.367	0.245	0.047

Notes: The table reports numbers and shares of the 2018 cohort students by their graduation status, mandatory exams completion (passing) and project defense by March 1, 2019.

Table 3: Treatment Effect on Additional Outcomes

	Project grade		Exams avg. grade		Exams passed (frac.)	
	(1)	(2)	(3)	(4)	(5)	(6)
TreatedXPost	0.420	-0.457	0.343	0.343	0.015	0.002
	(0.758)	(0.669)	(0.990)	(1.102)	(0.021)	(0.023)
Control Mean	80.93	80.93	64.15	64.15	0.74	0.74
Obs	7,493	7,493	11,130	11,130	11,130	11,130
Pair FE	YES	NO	YES	NO	YES	NO
CollegeXMajor FE	NO	YES	NO	YES	NO	YES

Notes: The table reports treatment effect estimates from a *difference-in-difference* specification (equation 1). In column (1)—(2) the outcome variable is final project grade and the sample is conditioned on project defense. In columns (3)—(4) the outcome variable is the average grade in all mandatory exams. Missing grades (indicating that the student did not take the exam) are replaced with the average grade of students who failed in the same exam subject. In columns (5)—(6) the outcome is the fraction of mandatory exams that the student passed. The sample includes all $college \times majors$ pairs which had students enrolled in each of the three cohorts (2017-19). Cohort year refers to the year of the last semester of studies. In columns (2), (4) and (6) the pair fixed effects are replaced with $college \times majors$ fixed effects. * p<0.1, ** p<0.05, *** p<0.01.

Table 4: Heterogeneous Effects by Major

	Major			
	Civil Eng. (1)	Architecture (2)	Electrical Eng. (3)	Software Eng. (4)
Panel A - Defense				
$Treated \times Post$	-0.013	-0.008	-0.021	-0.030
	(0.071)	(0.069)	(0.051)	(0.073)
Control mean	0.487	0.601	0.720	0.509
Panel B - Diploma				
$Treated \times Post$	-0.028	-0.011	-0.065	-0.015
	(0.063)	(0.076)	(0.056)	(0.067)
Control mean	0.414	0.531	0.643	0.412
Obs	5,626	2,084	2,321	1,099

Notes: The table reports treatment effect estimates from a difference-in-difference specification (equation 1) for each study field (major) separately. The sample includes all $college \times major$ pairs which had students enrolled in each of the three cohorts (2017-19). Cohort year refers to the year of the last semester of studies. * p<0.1, ** p<0.05, *** p<0.01.

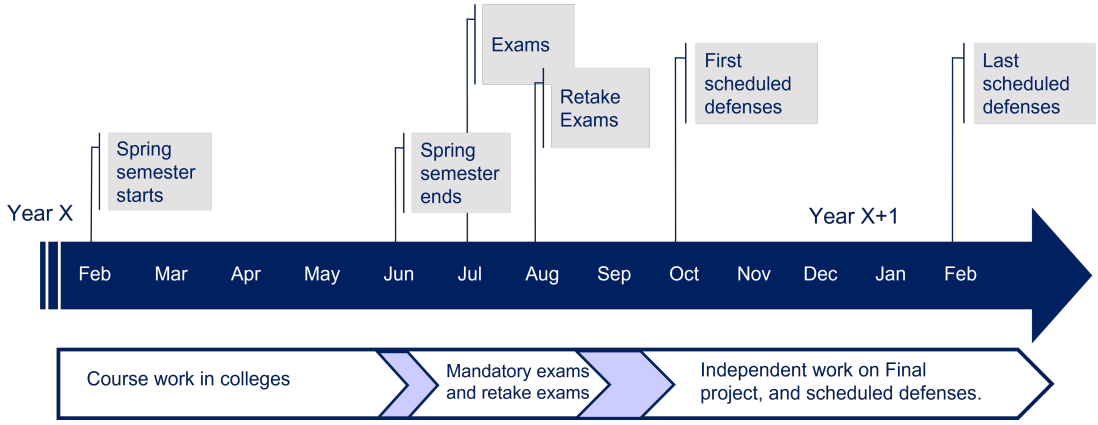


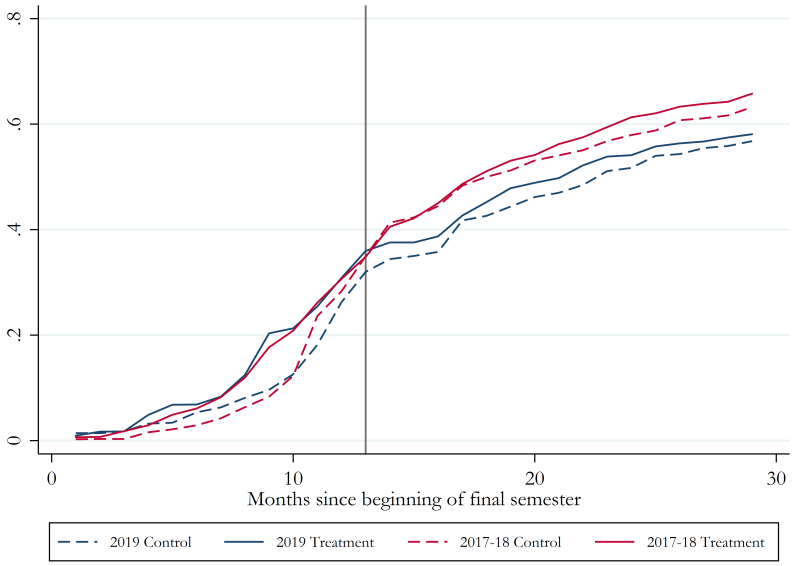
Figure 1: Typical Timeline for Study Completion, Students in Cohort X

Notes: Outline of typical final study semester and subsequent schedule for exams and final project, for "Cohort X". The boxes above the timeline indicate the scheduled activities that take place during the period, while the arrow below the timeline outlines what study related activities the students are supposed to devote their time to in each sub-period.

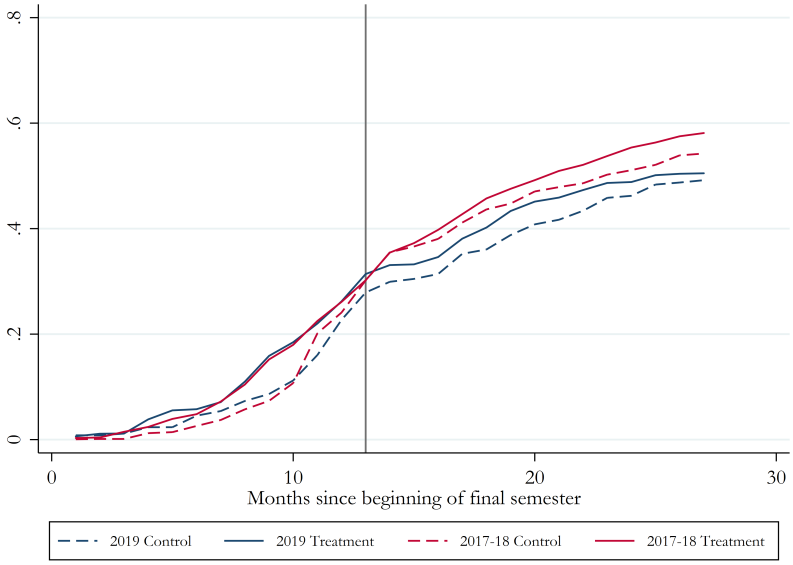
Table 5: Heterogeneous Effects by Student and Department Characteristics

	Socioeconomic Z-Score		Passed Exams		Teaching Satisfaction	
	Below Median (1)	Median or Above (2)	No (3)	Yes (4)	Below Median (5)	Median or Above (6)
Panel A - Defense						
Treated × Post	-0.016	-0.023	0.013	-0.015	0.006	-0.040
	(0.044)	(0.060)	(0.038)	(0.043)	(0.074)	(0.049)
Control mean	0.53	0.60	0.18	0.75	0.60	0.53
Panel B - Diploma						
Treated × Post	-0.030	-0.035	-0.014	-0.024	-0.023	-0.042
	(0.039)	(0.058)	(0.024)	(0.043)	(0.063)	(0.049)
Control mean	0.43	0.54	0.08	0.69	0.54	0.43
Obs	4,962	6,168	2,952	8,178	5,719	5,411

Notes: The table reports treatment effect estimates from a difference-in-difference specification (equation 1) for for different subgroups. Columns (1)—(2) split the sample by SES score based on locality of residence (below or above the median student in our sample). Columns (3)—(4) separates students who did or did not pass all the mandatory exams by March of the year after their final semester (e.g., for the 2019 cohort — exams passed by March 2020). Columns (5)—(6) divide the sample by the average level of students' satisfaction from the quality of teaching, lectures and exercises (below and above the median department in our sample). The sample includes all *college × majors* pairs which had students enrolled in each of the three cohorts (2017-19). Cohort year refers to the year of the last semester of studies. Regressions include *college × major* fixed effects (instead of pair fixed effects as in the main estimation). * p<0.1, ** p<0.05, *** p<0.01.



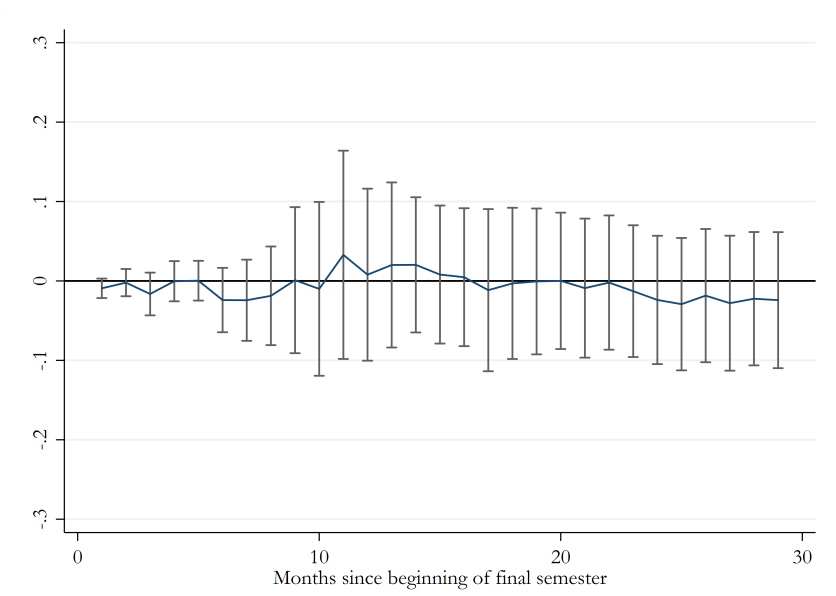
(a) Defense



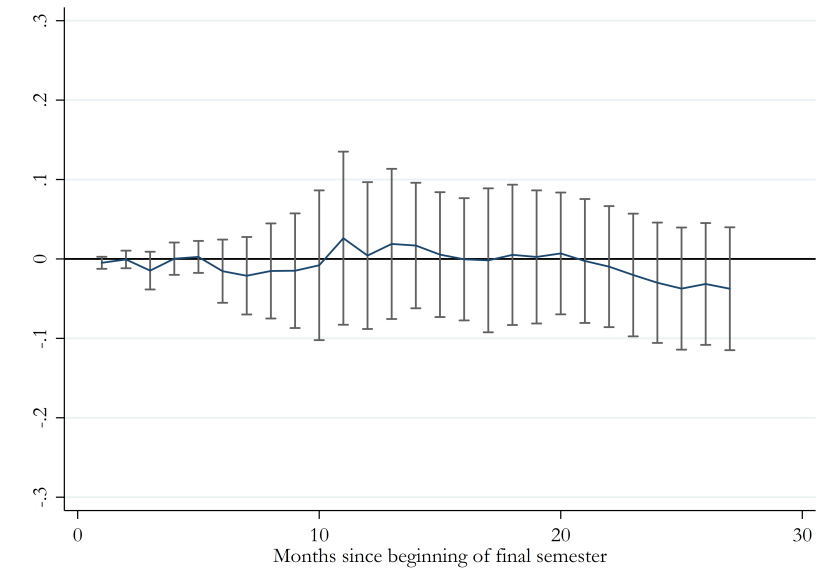
(b) Graduation (Diploma)

Figure 2: Cumulative Defense and Graduation Rates by Treatment Status, 2019 vs 2017-18 cohorts

Notes: Figure (a) presents cumulative rates of successful defense by month since the beginning of the last semester of studies (February of the relevant cohort year). This is shown separately for treatment and control units. Figure (b) shows the same for graduation (diploma) rates. The sample includes all $college \times major$ pairs which had students enrolled in each of the three cohorts (2017-19). Cohort year refers to the year of the last semester of studies. Month 0 denotes January of the last study year for each cohort (e.g., January 2019 for the 2019 cohort). The 2019 cohort is the treated cohort. The vertical line marks February 2020 for the 2019 cohort, which is the month before the first Covid-19 lock-down was imposed in Israel (on March 14th 2020).



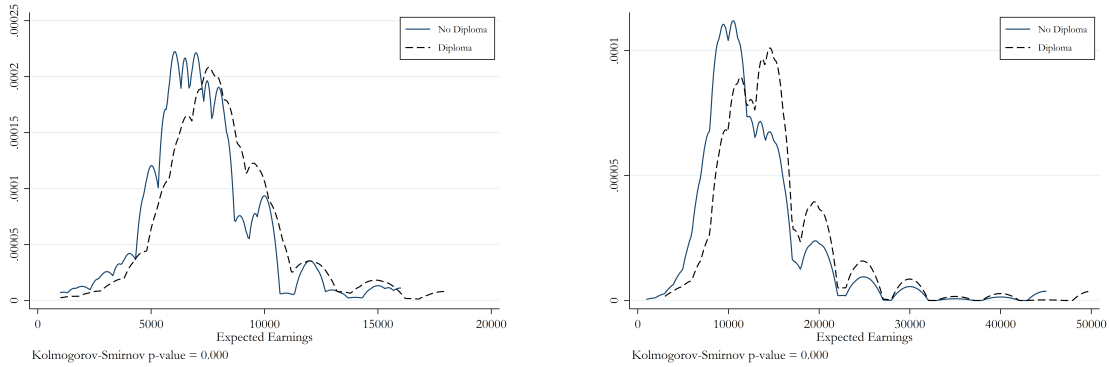
(a) Defense



(b) Graduation (Diploma)

Figure 3: DID Treatment Effect Estimates by Month (since final semester)

Notes: The figures presents DID coefficient estimates and confidence intervals comparing treatment to control units between the post- (2019) and the pre-treatment (2017-18) cohorts. The regression specification is given by equation 1. Figure (a) shows the estimated effects on defenses and figure (b) shows the estimated effects on graduation with a diploma. The sample includes all $college \times major$ pairs which had students enrolled in each of the three cohorts (2017-19). Cohort year refers to the year of the last semester of studies. Month zero is January of the last studies year for each cohort (e.g., January 2019 for the 2019 cohort).



(a) 1 year post graduation

(b) 5 years post graduation

Figure 4: Distribution of expected earnings after studies completion with and without a diploma

Notes: Both figures present kernel density plots of students' expected earnings after completing their studies with or without an official diploma, as reported in our survey. The distributions are presented for the expectations after 1 year (a) and after 5 years (b). We apply optimal bandwidth selection. P-values from a Kolmogorov-Smirnov test for equality of distributions are listed below each figure.

1 Appendix

Table A1: Descriptive Statistics — 2020 Cohort Survey

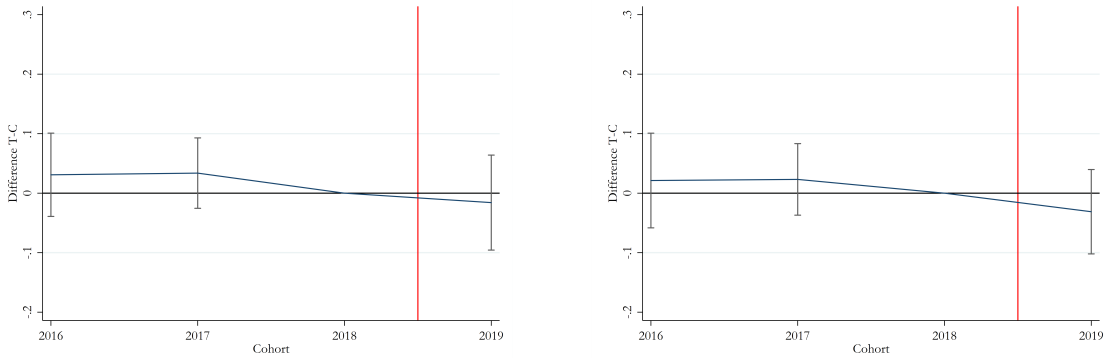
Variable	Mean (1)	SD (2)
Married	0.382	0.486
Has Children	0.296	0.456
Living with Parents or Other Older Relatives	0.486	0.500
Mother with Tertiary Education	0.309	0.462
Father with Tertiary Education	0.287	0.452
1st Generation Immigrant	0.089	0.285
2nd Generation Immigrant	0.231	0.422
Worked Before Studies	0.930	0.255
Worked in Study Field Before Studies	0.331	0.471
Worked During Studies	0.860	0.347
Worked in Study Field During Studies	0.301	0.459
Observations	1,587	

Notes: The table reports summary statistics from the 2020 cohort survey. The number of observations for parents' education is slightly smaller (1,441 for mothers and 1,424 for fathers) because we allowed an "irrelevant or unknown" category which is coded as missing values.

Table A2: Summary Statistics and Balance Tests - Extended Sample 2016-2019

	Treatment cohort			All cohorts	
Variable	Control group mean	Diff. T-C	Obs.	DID coeff.	Obs.
Panel A - Individual level					
Age (years)	26.91	0.348	4,786	0.159	16,461
		(0.658)		(0.222)	
Female	0.241	-0.072	4,786	0.005	16,461
		(0.066)		(0.007)	
Socioeconomic z-score	-0.105	-0.001	4,767	-0.095	16,410
		(0.146)		(0.023)	
Panel B - Department level					
# of students in department	52.66	8.434	84	9.843***	310
		(10.008)		(1.977)	
Graduation rate (%)	47.23	2.634	84		
		(3.287)			
Within Academic institution	0.317	0.032	84		
		(0.104)			
Focus on Arab population	0.073	0.020	84		
		(0.061)			
Evening track	0.732	0.036	84		
		(0.096)			
# of majors offered by college	3.1	-0.053	84		
		(0.169)			

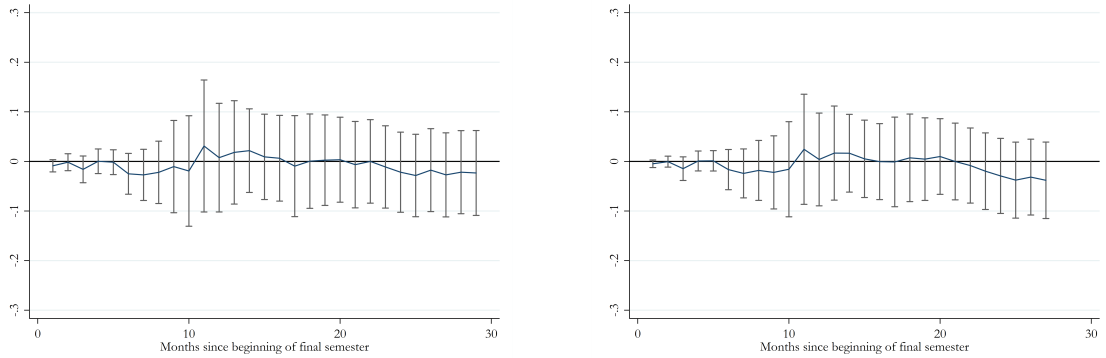
Notes: Columns 1-3 of the table report summary statistics for key variables in the administrative data at the student level (Panel A) and department level (Panel B), divided by treatment status of the department in the experimental cohort. Column 2 reports the results of a DID estimation (equation 1 with $college \times major$ fixed effects instead of pair fixed effects) for the subset of background variables for which there is variation across years. The variables which are constant over cohorts are: Graduation rate average taken over the seven years prior to our intervention, a dummy for whether the department belongs to a college that is part of an academic institution, a dummy for whether the department belongs to a college that focuses on the Arab population, and a dummy for the department having an evening track. * p<0.1, ** p<0.05, *** p<0.01.



(a) Defense **(b)** Graduation (Diploma)

Figure A1: Event Study Analysis

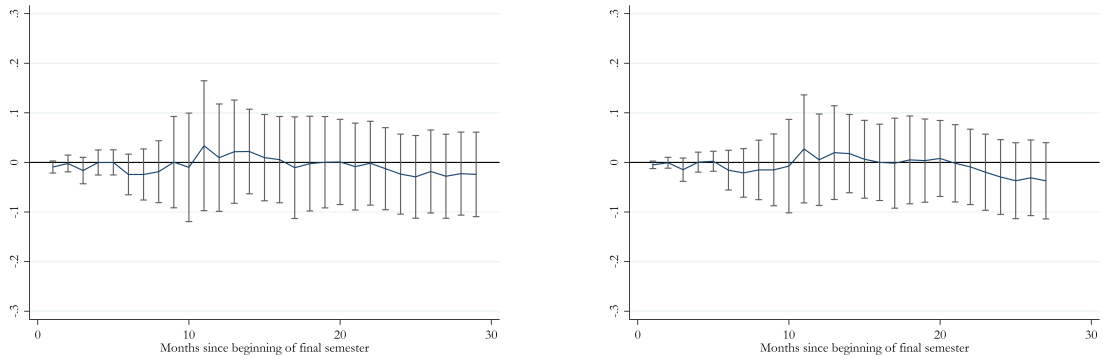
Notes: The figures present point estimates and confidence intervals from an event study specification (equation 2) which measure the difference between treatment and control units for each cohort separately (relative to the 2018 cohort difference which is normalized to zero). In figure (a), the outcome variable is final-project defense up to 29 months after the beginning of the final semester. In figure (b), the outcome is graduation (with a diploma) up to 27 months after the beginning of the final semester. The sample includes all $college \times major$ pairs which had students enrolled in the treated cohort (2019) and in *any* of the three pre-treatment cohorts (2016-2018). Cohort year refers to the year of the last semester of studies.



(a) Defense **(b)** Graduation (Diploma)

Figure A2: DID Treatment Effect Estimates by Month (since final semester)
Specification without pair fixed effects

Notes: The figures presents DID coefficient estimates and confidence intervals comparing treatment to control units between the post- (2019) and the pre-treatment (2017-18) cohorts. The regression specification is given by equation 1 but with $college \times major$ instead of pair fixed effects. The sample includes all $college \times major$ pairs which had students enrolled in each of the three cohorts (2017-19). Cohort year refers to the year of the last semester of studies.

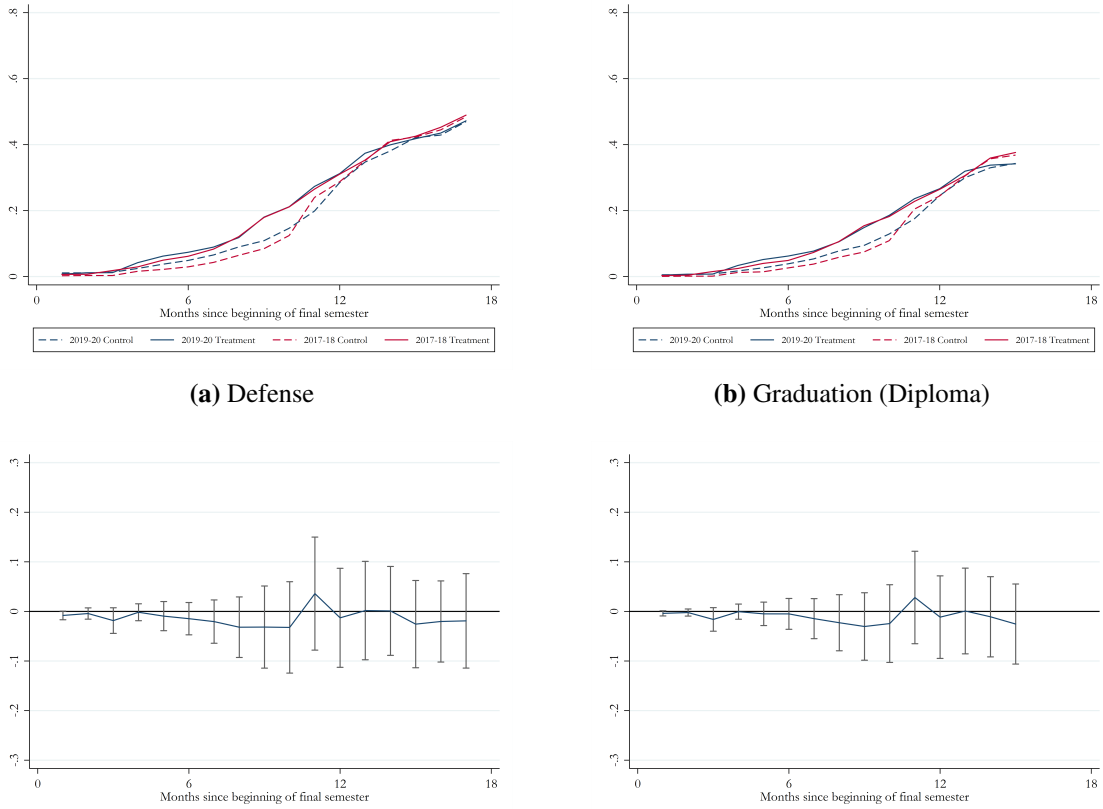


(a) Defense

(b) Graduation (Diploma)

Figure A3: DID Treatment Effect Estimates by Month (since final semester) (Specification with additional controls)

Notes: The figures presents DID coefficient estimates and confidence intervals comparing treatment to control units between the post- (2019) and the pre-treatment (2017-18) cohorts. The regression specification is given by equation 1 with additional controls for gender, age (quadratic), SES by locality of residence and the whether the college specifically targets the Arab population. The sample includes all $college \times major$ pairs which had students enrolled in each of the three cohorts (2017-19). Cohort year refers to the year of the last semester of studies.



(a) Defense

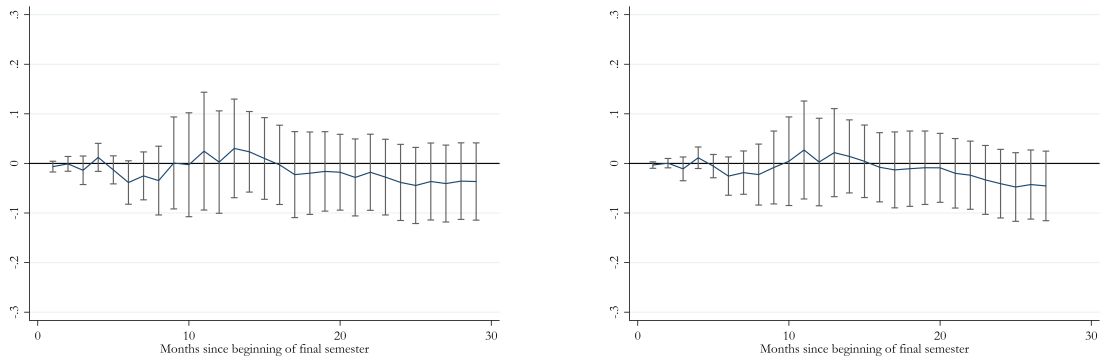
(b) Graduation (Diploma)

(c) Defense

(d) Graduation (Diploma)

Figure A4: DID Treatment Effect Estimates by Month (since final semester)
Extended sample 2017-2020

Notes: Figures (a) and (b) presents cumulative rates of successful defense and graduation (with a diploma) respectively, by month since the beginning of the last semester of studies (February of the relevant cohort year). This is shown separately for treatment and control units. Figures (c) and (d) present DID coefficient estimates and confidence intervals comparing cumulative rates of successful defense and graduation (with a diploma) respectively for treatment versus control units between the post- (2019-2020) and the pre-treatment (2017-18) cohorts. The regression specification is given by equation 1. The sample includes all $college \times majors$ which had students enrolled in each of the three cohorts (2017-19), as in our main estimation.



(a) Defense

(b) Graduation (Diploma)

Figure A5: DID Treatment Effect Estimates by Month (since final semester)
Extended sample 2016-2019

Notes: The figures presents DID coefficient estimates and confidence intervals comparing treatment to control units between the post- (2019) and the pre-treatment (2017-18) cohorts. The regression specification is given by equation 1. The sample includes all $college \times majors$ which had students enrolled in 2019 and in *any* of the pre-treatment cohorts 2016-2018, regardless of the paired unit's presence.

A.2 Additional details on survey data

A.2.1 Online survey on student satisfaction

We use data from an online, anonymous survey on college rankings collected by the NITT in 2018. These rankings are on a five point scale and refer to six different statements which focus on different aspects of students' satisfaction with their college and department. We construct three department-level metrics based on these six statements. First, teaching satisfaction is based on the average score of two statements referring to teachers' skills and expertise, and to the quality of lectures and exercises. The second metric adds the scoring of a third statement on practical training facilities (labs and materials). The third metric measures overall satisfaction and averages the scores of all six statements, including three statements on maintenance levels, job-finding assistance and general satisfaction (whether one would recommend the college to friends).

A.2.2 Details on survey data collection

We administered a survey to students in the 2020 cohort in September 2020: just after completion of final exams, and before deadlines were announced for the final project. The survey was collected electronically. A link to the survey was sent via SMS messages to 4,094 students, 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 can not 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 in the survey questions.