# Pre-Analysis Plan:
# Tutoring in (Online) Higher Education

David Hardt, Markus Nagler[†], Johannes Rincke[‡]

May 18, 2021

## Contents



david.hardt@fau.de

[†]markus.nagler@fau.de

[‡]johannes.rincke@fau.de

# 1 Introduction

## 1.1 Abstract

We study a cohort of students at the School of Business and Economics at a German university in their second semester which is entirely held online due to COVID-19. We design and implement a program that provides students with a tutor in economics subjects from a more advanced semester and induces peer-to-peer interaction. Tutors support randomly formed groups of 2 to 3 students. The tutors and the groups meet online and discuss problem sets in microeconomics and macroeconomics. To determine effects of the intervention, we measure the students' performance in both subjects as well as their self-reported mental health and motivation.

## 1.2 Motivation

During the past year, the COVID-19 pandemic has forced virtually all teaching to be conducted online around the world. For higher education, some observers argue that this has accelerated a trend where online education has risen in importance for years. Since online higher education can scale up higher education at low cost and reach students who would otherwise not attend tertiary education, politicians and academics alike are taken with the prospect of leveraging this technology to improve human capital [e.g., Goodman et al., 2019]. In contrast to in-person teaching, online teaching seems to be somewhat inferior to classical classroom-based teaching, however [Brown and Liedholm, 2002; Figlio et al., 2013; Joyce et al., 2015; Alpert et al., 2016; Bettinger et al., 2017; Kofoed et al., 2021].

In online teaching, many elements that make in-person teaching successful, such as meeting other peers or personalized education, are absent, potentially leading to worse learning gains. Students' mental health may also be affected by a lack of interactions, as evidenced by worse student mental health in the past year [e.g., Lai et al., 2020; Son et al., 2020; Browning et al., 2021; Logel et al., 2021a]. Among the primary correlates of worse mental health of students during the pandemic are loneliness or studying in isolation [e.g., Elmer et al., 2020; Logel et al., 2021a]. These features of online education may thus dampen the attractiveness of this format.

Recent research has found personalized tutoring to be a potentially successful way how to increase the effectiveness of teaching. In-person tutoring interventions have been shown to be effective across differing settings and for a wide variety of students [Fryer, 2017; de Ree et al., 2021]. The literature has so far primarily focused on preK-12 tutoring experiments, where tutoring increases learning outcomes by around 0.37SD on average, a large effect in comparison to other education interventions [Nickow et al., 2020]. For middle school kids, personalized tutoring in an online environment has been shown to be very effective across a number of outcomes including cognitive test scores and mental health during this pandemic [Carlana and La Ferrara, 2021]. This remarkable success of tutoring interventions is in contrast to mentoring interventions that have at best shown small improvements in average student performance or improvements only for subgroups of students [e.g., Angrist et al., 2009; Oreopoulos and Petronijevic, 2019; Hardt et al., 2020]. To date, little is known about whether tutoring is effective in (online) higher education settings. The few results in the literature do not provide a clear picture of the effectiveness of such interventions [see, e.g., Parkinson, 2009; Munley et al., 2010; Paloyo et al., 2016; Pugatch and Wilson, 2018, 2020]. However, tutoring is one element of highly successful student support programs such as the City University of New York's Accelerated Study in Associate Programs [ASAP, see Scrivener et al., 2015; Sommo et al., 2018].

This trial is designed to test whether a form of tutoring that aims at inducing peer-to-peer interactions affects the effectiveness of online higher education and the mental health of university students. The context of the trial is the School of Business, Economics, and Society at a German university during the summer term 2021 that is taking place online due to the COVID-19 pandemic. In each fall semester, about 890 students enroll in the three-year (six-semester) bachelor's program *Economics and Business Studies*. This program is broad and can lead to specializations in business administration, economics, information systems, and business and economics education. The program requires students to collect 180 credits to graduate. The study plan therefore assigns courses worth 30 credits to each of the six semesters. In each of the first two semesters, students are supposed to pass exams in six compulsory courses, each of them worth five credits. The specialization only starts after the first year in which students take compulsory modules.[1]

[1]Administrative data from the academic year 2018/19 shows that even in in-person teaching, many students underperform relative to the suggested curriculum in the first study year: After the first semester, only 59 percent of students still enrolled at this point in time have com-

Our study is meant to provide causal evidence on the effectiveness of a tutoring program and induced peer-to-peer interaction in the second semester. For this purpose, we designed a tutoring program that randomly groups students in their second semester into groups of two to three and assigns each group a tutor currently enrolled in the fourth or sixth semester of the same program and who performed well in the subjects that the tutoring program focuses on. The tutorials focus on microeconomics and macroeconomics, both of which are compulsory courses in the second term at the program. The following sections present further details on our experimental design and the planned analysis of the data.

### 1.3 Research Questions

- Does tutoring and peer-to-peer interaction improve the students' academic achievement in a context where all teaching is done online?
- Does tutoring and peer-to-peer interaction improve students' mental health?
- Do the effects of such an intervention on achievement and mental health differ by student gender and prior performance?

## 2 Experimental Design

### 2.1 Intervention

The study program *Economics and Business Studies* at the university where the trial is going to be implemented requires students to collect 180 credits to graduate. Students are expected to graduate after three years (six semesters). The study plan assigns courses worth 30 credits to each semester. Administrative data show that a large share of students do not complete 30 credits per semester, delaying their graduation. At the same time, survey data collected from an earlier cohort of students that were taught in-person suggests that most students do not work full-time even if one aggregates the hours studied and the hours worked

pleted courses worth at least 30 credits. The curriculum for the second semester comprises some courses involving more rigorous methods relative to the first semester. As a result, students typically further decrease their performance in the second semester: only about 25 percent have completed 60 credits at the end of the second semester.

to earn income.[2] The salient study plan and target of achieving 30 credits per term, the fact that most students do register for exams worth these credits, and the fact that students do not seem to work enough to pass these exams suggests that many students have problems in self-organizing and/or studying efficiently. This is where our program is supposed to intervene.

Due to the COVID-19 pandemic, in the summer term 2021 all courses of the School of Business, Economics, and Society will be conducted in online format. To this end, the university has acquired licenses of *Zoom* (already before the summer term 2020), an online video conference tool used widely in academic settings during this pandemic to digitize classes and seminars and to provide distance education. While the exact implementation of online teaching differs by subject and instructor, this should make the setting similar to the setting of other academic institutions around the globe during this pandemic.

The trial focuses on the second semester consisting of six compulsory courses. We recruited 15 tutors who are themselves students in the *Economics and Business Studies* program at the School of Business, Economics, and Society. We hired students as tutors who successfully completed the courses under consideration and during the current semester are enrolled in the fourth or sixth semester of the program.

In the first week of the semester, students were informed via e-mail about the launch of a new small-group tutoring program designed specifically for students in the second semester of the study program. They were invited to register for the program through a webpage. The page asked for the students' consent to use their personal information for research purposes in anonymized form and for their consent to pass along their name and email address to their tutors. We sent reminder emails to students who did not register for the program within two days. We subsequently randomly invited as many students as we have slots in the tutoring program based on our design to participate in the program. Students who were interested in the program but were not offered a slot in the randomization serve as our primary control group. A secondary control group may consist of students who did not indicate their willingness to participate in the program. Depending on the observed selection patterns into registering for the program, this secondary control group may or may not be used in our main analysis.

The tutoring program focuses on advancing students' knowledge of microe-

[2] On average in the first two semesters, survey participants spend about 13.3 hours per week attending courses, about 9.8 hours self-studying, and 7.5 hours to earn income.

conomics and macroeconomics, two compulsory courses in the second term of their study program, and on inducing peer-to-peer interaction. Students are supposed to work on the problem sets (available to all students) in advance of each tutoring session in randomly formed groups of three (their tutoring groups) every two weeks. In every other week (i.e., when the tutoring groups do not work on the problem sets themselves), tutors meet with the groups to discuss any issues that the tutoring group had while solving the problem sets. During the session, the tutor then explains the problems, asks for the issues that students had while solving the problem set, and also offers general advice on how to study effectively or on anything else that is related to students' second term. Each bi-weekly tutoring session is supposed to last for 90 minutes.

The idea of the program is to (i) induce students to take up tutoring services, (ii) induce peer-to-peer interaction between students in an online environment where this sort of interaction is missing and feelings of loneliness are pervasive and (iii) provide a commitment device to ensure that students study regularly during the term in an (online) environment where external structure (e.g., resulting from a fixed time schedule) is missing. Because of the personalized nature of the tutoring and the peer-to-peer interaction that is induced through our small groups, we hypothesize that students' mental health is positively affected by their program participation.

The tutors are asked to take brief notes about the content of the discussions and some background information during each meeting. Tutors are also instructed to prepare thoroughly for every individual meeting by recapturing their notes from the previous meeting. To limit the risk of spillovers, we ask all tutors to make sure that the information and tutoring is only provided to the students in their group and not to other students.

In the control group, there is no tutoring. However, the School of Business, Economics, and Society provides general practice sessions for students in both subjects that are less personalized and where peer-to-peer interaction is not directly induced. In terms of content, it is identical to what tutors and student groups are supposed to discuss in our intervention. In microeconomics, there are also additional practice tests that all students can take online that do not count towards students' grade.

After the end of the exam period (preliminarily scheduled for July and August 2021), we will collect individual data on exam performance. We may also collect additional performance data for a further research paper or research note at a later point in time to assess long-run benefits of the program.

We do not expect that the School of Business, Economics, and Society will switch from online to in-person teaching during the semester and therefore plan for a full teaching period with online courses being the only (or at least dominant) way of teaching. However, if the overall situation changes significantly during the experimental period, we may allow tutors and students to meet in person for the meetings.

## 2.2 Hiring and Training of Tutors

For administrative reasons, we had to initiate the hiring of the tutors about 4 weeks before the start of the program. In total, we hired 15 tutors. Work contracts are specified such that each tutor can handle a maximum of four groups of two to three students. We plan to have an about equal number of 2-person and 3-person groups. With 60 groups in total, the tutoring program's maximum capacity is therefore about 150 students. All tutors are students who successfully completed the courses that the program focuses on and during the current semester are enrolled in the sixth semester of the study program.

Shortly before the start of the tutoring program, all tutors took part in a kick-off meeting. In the kick-off meeting, the research team explained the purpose and the general structure of the program and laid out the planned sequence and contents of the tutoring sessions to be held with each student group. The tutors could also ask questions. The tutors were informed about the fact that the program's capacity is limited and that a random subset of all students in the second term was allowed to participate. After this initial meeting, the members of the research team send regular e-mails to the tutors and answer questions in response to individual queries by the tutors.

## 2.3 Data Collection

### 2.3.1 Administrative Data

We collect administrative data from the university to measure all outcomes related to exam participation and academic achievement. In addition, the university has provided us with background information on individual students. The individual characteristics include information on enrollment, gender, age, type of high school completed, and information on high-school GPA. The university's data protection officer authorized this data collection.

### 2.3.2 Survey Data

After the end of the intervention and shortly before the exams period, we will invite all students sampled at baseline to an online survey. The survey will be conducted using an existing platform at the department that is frequently used to survey students. Students who complete the survey will receive a payoff of €8. The survey will elicit the students' own study effort and students' self-perceived mental health. For details, see Subsections 3.2.1 and 3.2.2.

## 2.4 Sampling

### 2.4.1 Invitations and Randomized Treatment Assignment

About 890 students enrolled for the study program Business Studies for the fall semester of 2020. We excluded from the experiment students who dropped out after the first semester, who are not formally in their second semester, for example because of having been enrolled at another university before and having already completed courses from the first or second semester of the study program without having taken these exams at the university, and students who completed less than a full course (5 credits) in the first term.[3] This leaves us with 714 students entering the second term. These students were invited to participate in the tutoring program in the first week of the term. 226 students responded to this invitation and registered for the program. Students registered for the program and in the treatment group can drop out at any time with no penalty. If dropouts lead to study groups with only one student left, we reassign the remaining students to other groups.

In total, we have 150 slots in our program. We randomly assigned a corresponding number of students to the treatment group, and the other interested students to the control group. The randomization was done in office by a computer. We used a stratified randomization scheme with gender and number of credits completed in the first semester (three bins) as strata variables. We will drop students from the sample who are credited for courses in the second semester and earned the credits in an earlier term (either at the same university, or elsewhere). Such credits often show up with some delay in the administrative data. It is therefore possible that despite dropping students with such credits in the first semester, we have sampled some students who have already earned

[3] In Germany, some students enroll at a university because as students they have access to heavily subsidized health insurance.

credits for the second term.

#### 2.4.2 Assignment of Registered Students to Groups and Tutors

We randomly assign students registered for the program to their study group and their tutor.

### 2.5 Minimum Detectable Effects

Our primary academic outcome are the total credits students earned in the courses microeconomics and macroeconomics. As we do not have baseline data on this outcome for the sample of students interested in the tutoring program, we discuss minimum detectable effects for the secondary outcome most closely related to the main outcome, which is the number of overall credits earned in the second term.[4] We provide minimum detectable effects for a significance level of 0.05 and a statistical power of 0.8. We assume 230 participants in total, with about 150 allocated to the treatment group and about 80 to the control group.

From the baseline data (performance in the winter term of 2020) for the experimental cohort, we expect the mean of credits earned in the control group to be about 23.6 (SD 8.1). The minimum detectable effect (assuming independence within study groups) would then be 3.2 credits, or 40 percent of a standard deviation. Assuming perfect dependence within study groups, the minimum detectable effect would be 3.9 credits. We note, however, that the courses in the summer term 2021 differ from the courses in the winter term, possibly affecting the distribution of credits earned. The true minimum detectable effect size might therefore differ significantly from the value provided above.

## 3 Empirical Analysis

### 3.1 Balancing Checks

We will check balance between treatment and control by t-tests (mean-comparison tests) on individual characteristics and by standardized differences. The characteristics included in the balancing checks will comprise gender, age

[4]We prefer not to use baseline data on performance in the second term for two reasons. First, we could not model selection into program participation. Second, the curriculum of the study program was significantly changed for the experimental cohort. Using the second-term baseline data could thus lead to misleading estimates of minimum detectable effects.

(in years), high-school GPA, a dummy for the most common type of high school certificate ("Gymnasium"), a dummy for students who obtained their high school certificate abroad, credits earned in the first (winter) term, a dummy for students who are in their first year at university, and a dummy for full-time students.[5]

We will run the same balancing checks on the sample of survey respondents. We will also study the selectivity in survey participation by means of mean-comparison tests between survey participants and non-participants.

## 3.2 Treatment Effects

### 3.2.1 Primary Outcomes

Our primary academic outcome are the total credits students earned in the courses microeconomics and macroeconomics. We also focus on students' average grade in both subjects. We note that GPA is, in principle, affected by the student's decisions how many credits to attempt and which exams to take. If we find that the effects on credits earned in micro and macro are both insignificant, the effect on the GPA can however reveal a possible effect of the intervention on academic achievement.

Our primary mental health outcome is a mental health index. The index will standardize each reply to a mental health question to have mean zero and standard deviation one in the control group and then build the unweighted sum of the standardized variables [Kling et al., 2007].

We measure students' (mental) health outcomes drawing from Logel et al. [2021b,a] and Carlana and La Ferrara [2021] on 5-point-Likert scales:

- Students' overall happiness during the term
- Students' feelings of stress during the term
- Students' feelings of nervousness or anxiousness during the term
- Students' feelings of depression or hopelessness during the term
- Students' feelings of disconnectedness from peers during the term
- Students' sense of belonging during the term
- Students' overall assessment of mental health

[5] About 10% of students are enrolled as part-time students in regular times because their university education is integrated into a vocational training program.

- Students' overall assessment of physical health

Following Angrist et al. [2009], we will not exclude students who withdrew from the sample. Students who withdrew before earning any credits in the second term will be coded as having zero attempted and earned credits.

#### 3.2.2 Secondary Outcomes

**Program Take-up and Service Use** A first set of secondary outcomes is meant to capture the decision to actively take part in the tutoring program. We will use the following variables:

- indicator for students using some tutoring service (participation in at least one virtual session with the group and the tutor)
- number of tutoring sessions completed
- indicator for students taking part in all scheduled sessions with their tutor

**Separate Impacts** Secondary outcomes are the likelihood of passing microeconomics and of passing macroeconomics, respectively. We will also consider as secondary outcomes the respective grades.

**Credits Earned** A further secondary outcome is overall credits earned in the second term. This variable measures most directly the overall students' academic achievement during the term in which the intervention takes place. We will especially focus on credits earned outside the tutoring topics microeconomics and macroeconomics.

**Credits Registered For** Further secondary outcomes are the total of credits registered for in the second term and the total of credits registered for in the courses microeconomics and macroeconomics. The variables measure the students' effort during the term in which the intervention takes place.

**Likelihood of Reaching First Year Goal** A further secondary outcome is the likelihood of having completed the 60-credit goal at the end of the first year.

**GPA** We will also consider the second-term overall GPA and GPA among other subjects than microeconomics and macroeconomics as a secondary outcome. These other subjects are of interest since there may be spillovers from the treatment. We again note that GPA is, in principle, affected by the student's decisions how many credits to attempt and which exams to take. If we find that the effects on credits attempted and earned are both insignificant, the effect on the GPA can however reveal a possible effect of the intervention on academic achievement.

**Survey Outcomes** From the student survey to be conducted after the end of the intervention, we will construct several additional outcomes (all derived using a 5-point Likert scale) beyond the mental health outcomes described above. In particular, we will also elicit students' motivation and study behavior, mostly drawing from Hardt et al. [2020]:

- Students' response to whether they had contact with peers in their program and how much
- Rating of own continuous study effort during the teaching term
- Assessment of own motivation during the term
- Assessment of whether students feel they prepared for the exam timely
- Assessment of whether students feel they provided enough effort to reach their goals

We will again report an index-value [Kling et al., 2007] for all but the first question in this block.

**Participation and Performance in Low-Stakes Tests Before the Exams** As part of a different research project [Adler et al., 2021a], we plan to implement low-stakes voluntary tests in both microeconomics and macroeconomics about one week before the respective exams. All students enrolled for the respective classes will be invited to participate in the tests. We may consider as secondary outcomes the decision to take the tests as well as performance in the test (percent of correctly solved problems, separately for both tests). However, we do not commit to report those outcomes in the main paper.

**Medium and Long-term Outcomes** In addition to the outcomes for the second semester, we may later on also collect data on dropping out from the study program and graduation. As the study program is a three-year program and many students do not graduate in time, we will collect this data about four years after the beginning of the intervention. We may also collect data for medium-term outcomes (like credits earned after the second study year). However, we do not commit to report those outcomes in the main paper. Hence, we may publish a paper on the short-term outcomes, and a separate paper (or a note) on the medium- or long-term outcomes.

## 3.3 Estimation

To evaluate the treatment effects, we will run linear regressions. Each of the outcomes will be regressed on the treatment indicator and the vector of strata variables. We will report robust standard errors that allow for clustering within the tutoring groups. However, not all students in the treatment group will take up the offer to actually use the tutoring services. Thus, in addition to intent-to-treat estimation regressions, we may run instrumental variable regressions using the randomized treatment assignment as an instrument for actual take-up. The first variable describing program take-up will be participation in the first session. We will also estimate model variants where we use treatment assignment to instrument for continuous service use. Given our design of an oversubscribed lottery, it may however be the case that IV an ITT estimates are very close, in which case we will refrain from reporting IV estimates.

For several reasons, we consider it likely that the treatment will have heterogeneous effects. A first observation is that prior evidence on online education shows that its negative effects are more pronounced among weaker students [e.g., Figlio et al., 2013; Bettinger et al., 2017], but that treatment effects of other (mentoring) programs in our context seem to be larger for somewhat better students [Hardt et al., 2020]. We thus expect treatment effects to differ by prior student performance. This can be first measured through mentees' performance in the first term. Second, in the baseline, there is a positive correlation between the high school GPA and the probability to meet the 30 credit-points target in any term. This suggests to also use the high school GPA as a dimension to study the treatment effect heterogeneity by prior performance.

A second observation is that the literature has commonly found male students to suffer more from online relative to in-person education [e.g., Figlio et al., 2013; Xu and Jaggars, 2014]. In our context, male students seem to benefit more from similar (mentoring) interventions, if anything [Hardt et al., 2020], while take-up rates in such programs seem to be higher for female students [e.g., Angrist et al., 2009]. Thus, we expect the effects of mentoring on outcomes among randomly chosen students to be larger for male than for female students. We plan to study the effects by gender to inform on these questions.

We plan to study the treatment effect heterogeneity by running regressions including an interaction term between the variable capturing the dimension of heterogeneity and the treatment indicator, together with the variable capturing the dimension itself. The strata variables will be included as controls. We also plan to study treatment effect heterogeneity by splitting the sample along the dimension. For the effects by prior performance, we will also split the sample into terciles of prior performance and estimate baseline regressions in these sub-samples.

The dimensions of a possible treatment effect heterogeneity described above will be reported in the paper. Other exploratory dimensions will be reported in the paper only if we find some heterogeneity.

As an example, we also plan to study whether the effects of tutoring are larger when being tutored by female than by male tutors. Prior literature has found that interactions between student and instructor gender can matter for teaching effectiveness [e.g., Dee, 2005, 2007; Hoffmann and Oreopoulos, 2009]. As described above, we make sure to have an around equal number of female and male tutors. Given the limited number of tutors, this analysis will however likely run into power issues.

As another example, we also plan to study whether the effects of tutoring are larger when being tutored by more senior (from the 6th term) or by less senior (from the 4th term) tutors. Nickow et al. [2020] report that typically, more distant tutors are more effective than peer tutors. Given the limited number of tutors, this analysis will however likely run into power issues.

We also plan to investigate whether the effects of tutoring are larger in two-person or three-person study groups. Given the limited number of groups, this analysis will however likely run into power issues.

## 3.4 Other Variables

We may include data from a related project that elicited the behavioral traits of students who are now in the second term [Adler et al., 2021b]. This may help us identify additional heterogeneity that is important to understand the effects of tutoring on academic performance and mental health in this setting, e.g. by socio-economic status. These analyses will be exploratory.

# References

ADLER, K., D. HARDT, M. NAGLER, Y. OFEK-SHANNY, AND J. RINCKE (2021a): "Personality Traits and Low-Stakes Tests," *Mimeo*.

ADLER, K., F. KOSSE, M. NAGLER, AND J. RINCKE (2021b): "Student Sorting on Personality Traits," *Mimeo*.

ALPERT, W. T., K. A. COUCH, AND O. R. HARMON (2016): "A Randomized Assessment of Online Learning," *American Economic Review: Papers & Proceedings*, 106, 378–382.

ANGRIST, J., D. LANG, AND P. OREOPOULOS (2009): "Incentives and Services for College Achievement: Evidence from a Randomized Trial," *American Economic Journal: Applied Economics*, 1, 136–163.

BETTINGER, E. P., L. FOX, S. LOEB, AND E. S. TAYLOR (2017): "Virtual Classrooms: How Online College Courses Affect Student Success," *American Economic Review*, 107, 2855–2875.

BROWN, B. W. AND C. F. LIEDHOLM (2002): "Can Web Courses Replace the Classroom in Principles of Microeconomics?" *American Economic Review: Papers and Proceedings*, 92, 444–448.

BROWNING, M., L. R. LARSON, I. SHARAIEVSKA, A. RIGOLON, O. MCANIRLIN, L. MULLENBACH, S. CLOUTIER, T. M. VU, J. THOMSEN, N. REIGNER, E. C. METCALF, A. D'ANTONIO, M. HELBICH, G. N. BRATMAN, AND H. O. ALVAREZ (2021): "Psychological impacts from COVID-19 among university students: Risk factors across seven states in the United States," *PLOS One*.

CARLANA, M. AND E. LA FERRARA (2021): "Apart but Connected: Online Tutoring and Student Outcomes during the COVID-19 Pandemic," *Working Paper*.

DE REE, J., M. A. MAGGIONI, B. PAULLE, D. ROSSIGNOLI, AND D. WALENTEK (2021): "High dosage tutoring in pre-vocational secondary education: Experimental evidence from Amsterdam," .

DEE, T. S. (2005): "A Teacher Like Me: Does Race, Ethnicity, or Gender Matter?" *American Economic Review Papers and Proceedings*, 95, 158–165.

——— (2007): "Teachers and the Gender Gaps in Student Achievement," *Journal of Human Resources*, 42, 528–554.

ELMER, T., K. MEPHAM, AND C. STADTFELD (2020): "Students under lockdown: Comparisons of students' social networks and mental health before and during the COVID-19 crisis in Switzerland," *PLOS One*.

FIGLIO, D., M. RUSH, AND L. YIN (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.

FRYER, R. G. (2017): "The Production of Human Capital in Developed Countries: Evidence From 196 Randomized Field Experiments," in *Handbook of Economic Field Experiments*, ed. by A. V. Banerjee and E. Duflo, Elsevier, vol. 2.

GOODMAN, J., J. MELKERS, AND A. PALLAIS (2019): "Can Online Delivery Increase Access to Education?" *Journal of Labor Economics*, 37, 1–34.

HARDT, D., M. NAGLER, AND J. RINCKE (2020): "Can Peer Mentoring Improve Online Teaching Effectiveness? An RCT during the Covid-19 Pandemic," *CESifo Working Paper No. 8671*.

HOFFMANN, F. AND P. OREOPOULOS (2009): "A Professor Like Me: The Influence of Instructor Gender on College Achievement," *Journal of Human Resources*, 44, 479–494.

JOYCE, T., S. CROCKETT, D. A. JAEGER, O. ALTINDAG, AND S. D. O'CONNELL (2015): "Does Classroom Time Matter?" *Economics of Education Review*, 46, 64–77.

KLING, J. R., J. B. LIEBMAN, AND L. F. KATZ (2007): "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75, 83–119.

KOFOED, M. S., L. GEBHART, D. GILMORE, AND R. MOSCHITTO (2021): "Zooming to Class?: Experimental Evidence on College Students' Online Learning during COVID-19," *IZA Discussion Paper No. 14356*.

LAI, A., L. LEE, M. PING WANG, Y. FENG, T. T. KWAN LAI, L. MING HO, V. S. FUN LAM, M. S. MAN IP, AND T. HING LAM (2020): "Mental Health Impacts of the COVID-19 Pandemic on International University Students, Related Stressors, and Coping Strategies," *Frontiers in Psychiatry*, 11.

LOGEL, C., P. OREOPOULOS, AND U. PETRONIJEVIC (2021a): "Experiences and Coping Strategies of College Students During the COVID-19 Pandemic," *NBER Working Paper No. 28803*.

——— (2021b): “"Using Online and Text Message Coaching to Improve Student Outcomes During the COVID-19 Pandemic,” *AEA RCT Registry No. AEARCTR-0007000*, https://doi.org/10.1257/rct.7000-1.0.

MUNLEY, V. G., E. GARVEY, AND M. J. MCCONNELL (2010): “The Effectiveness of Peer Tutoring on Student Achievement at the University Level,” *American Economic Review: Papers & Proceedings*, 100, 277–282.

NICKOW, A., P. OREOPOULOS, AND V. QUAN (2020): “The Impressive Effects of Tutoring on PreK-12 Learning: A Systematic Review and Meta-Analysis of the Experimental Evidence,” *NBER Working Paper No. 27476*.

OREOPOULOS, P. AND U. PETRONIJEVIC (2019): “The Remarkable Unresponsiveness of College Students to Nudging And What We Can Learn from It,” *NBER Working Paper No. 26059*.

PALOYO, A. R., S. ROGANA, AND P. SIMINSKI (2016): “The effect of supplemental instruction on academic performance: An encouragement design experiment,” *Economics of Education Review*, 55, 57–69.

PARKINSON, M. (2009): “The effect of peer assisted learning support (PALS) on performance in mathematics and chemistry,” *Innovations in Education and Teaching International*, 46, 381–392.

PUGATCH, T. AND N. WILSON (2018): “Nudging study habits: A field experiment on peer tutoring in higher education,” *Economics of Education Review*, 62, 1.

——— (2020): “Nudging Demand for Academic Support Services: Experimental and Structural Evidence from Higher Education,” *IZA Discussion Paper No. 13732*.

SCRIVENER, S., M. J. WEISS, A. RATLEDGE, T. RUDD, C. SOMMO, AND H. FRESQUES (2015): “Doubling Graduation Rates: Three-Year Effects of CUNY’s Accelerated Study in Associate Programs (ASAP) for Developmental Education Students,” *MDRC*.

SOMMO, C., D. CULLINAN, M. MANNO, S. BLAKE, AND E. ALONZO (2018): “Doubling Graduation Rates in a New State Two-Year Findings from the ASAP Ohio Demonstration,” *MDRC Policy Brief 12/2018*.

Son, C., S. Hegde, A. Smith, X. Wang, and F. Sasangohar (2020): "Effects of COVID-19 on College Students' Mental Health in the United States: Interview Survey Study," *Journal of Medical Internet Research*, 22, e21279.

Xu, D. and S. S. Jaggars (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.