

Overview

The code in this replication package produces the tables and figures in the manuscript using Stata version 18. There are four Stata do files that rely on four datasets (available in both .dta and .csv format) to produce the results. The code does not take a long time to run; users should expect approximately one minute (60 seconds) to replicate all tables and figures.

Data Availability and Provenance Statements

The data come from two sources. First, the authors ran experiments with students from the University of Toronto from 2014 to 2020, during which time they collected data via direct surveys of students. Second, the authors then matched the survey data to student administrative records provided directly by the University of Toronto. Data from both sources are confidential and the authors are prohibited from sharing the raw data in their original form. Under our ethics agreement, however, we can share anonymized versions of the analysis files, which contain limited variables needed for most of the analysis and anonymized student identifiers. The data files we share in this replication package do not have many student background characteristics (control variables in robustness checks) or campus location information. Four such files are included in this replication package, and while anonymous, the student identifiers are consistent across the files, allowing one to merge across datasets if desired (although this is not required for the replication).

An interactive copy of all surveys and experimental materials that generated the author-collected data used in the paper can be found at <https://www.studentachievementlab.org/>. Users can complete each survey or experimental module in the format in which it was presented to students. The authors are happy to provide assistance with doing so and to provide clarity on the variables provided in the data available in this replication package.

The administrative data from the University of Toronto is available through a request to the University Registrar's office.

Dataset list

- **treatment_effects_analysis_file_anonymized1.dta** (also available in .csv format)
Brief description: this is the analysis data file that is used to produce all the estimates of treatment effects on grades and study habits across all cohorts.
- **year5_April2022FlwUp_anlaysis_file_anonymized1.dta** (also available in .csv format)
Brief description: this is the analysis data file that is used to produce all the descriptive analysis of the year 5 (2018-19) cohort along with the analysis of the April 2022 study time follow-up survey for this cohort.

- **year6_anlaysis_file_anonymized1.dta** (also available in .csv format)
Brief description: this is the analysis data file that is used to produce all the descriptive analysis of the year 6 (2019-20) cohort.
- **year6_weeklysurveys_anlaysis_file_anonymized1.dta** (also available in .csv format)
Brief description: this is the analysis data file that is used to produce all the descriptive analysis of the weekly dynamics in grade expectations and study time for the year 6 (2019-20) cohort.

Software Requirements

Stata (version 18) was used to carry out all the analysis. The following additional programs were installed (via ssc install):

- binscatter
- outreg2
- cv_regress

Memory and Runtime Requirements

Running all the analysis code should take only about one minute (60 seconds). The code was last run on a 4-core Intel-based laptop with Windows 11.

Description of programs/code

There are four analysis programs that produce all the tables and figures in the manuscript. There is a “master script” program that runs each of the four analysis programs. We also include two programs that create the estimates of the student-specific intercept and slope (expected ability and return to study time parameters) that are discussed in Section 4.1 of the manuscript. Note the user does not need to run these last two programs, as these parameters (expected ability and return to study time) have already been appended to the main analyses data files; we only include these programs for users who wish to replicate the parameter estimates.

Running the “master script” program (and, in turn, the four analysis programs) will create the tables and figures in the manuscript and store them, respectively, the following folders:

- \3 replication package\output\figures
- \3 replication package\output\tables

There are three exceptions: the output for Table 1, Table 5, and Table A2 is not exported to a standalone file; instead, users can retrieve the output for these tables directly from the Stata output window by running the appropriate analysis program (described below). All other figures and tables are produced as their own separate .png or .xml files and logically named.

The following is a detailed description of the analysis programs, which all appear in \3 replication package\code

- Master Analysis Script.do runs the four analysis programs.
- Estimating Effects on Grades & Study Time Across All Cohorts.do produces treatment effect estimates for academic outcomes and study time across all cohorts.
- Analysis of Year 5 (2018-19) Cohort & the April 2022 Study Time Followup.do produces all the descriptive analysis of the year 5 (2018-19) cohort along with the analysis of the April 2022 study time follow-up survey for this cohort.
- Analysis of Year 6 (2019-20) Cohort.do produces all the descriptive analysis of the year 6 (2019-20) cohort.
- Analysis of Weekly Dynamics in Year 6 (2019-20) Cohort.do produces all the descriptive analysis of the weekly dynamics in grade expectations and study time for the year 6 (2019-20) cohort.
- **[Optional].** Estimating Student-Specific Expected Returns for Year 5 (2018-19) Cohort.do produces the estimates of the student-specific intercept and slope (expected ability and return to study time parameters) for the year 5 (2018-19) cohort that are discussed in Section 4.1 of the manuscript.
- **[Optional].** Estimating Student-Specific Expected Returns for Year 6 (2019-20) Cohort.do produces the estimates of the student-specific intercept and slope (expected ability and return to study time parameters) for the year 6 (2019-20) cohort that are discussed in Section 4.1 of the manuscript.

Instructions to Replicators

The following are step-by-step instructions for replicating the figures and tables in the manuscript.

- Each of the do files (analysis programs) described above begins by defining directories that contain the analysis data and that will contain the output produced by the code. Users should change the portion of directory path that appears prior to “\3 replication package” at the start of each file to the appropriate directories on their local machine. This should be done where it says, “YOUR_PATH_HERE”.
- Make sure the do files (analysis programs) are stored in “YOUR_PATH_HERE”\3 replication package\code
- Make sure the data files are stored in “YOUR_PATH_HERE”\3 replication package\data

- Run the analysis program Master Analysis Script.do
- Find the resulting figures and tables in “YOUR_PATH_HERE”\3 replication package\output\figures and “YOUR_PATH_HERE”\3 replication package\output\tables

List of tables and programs

Note: while the output for each table appears in the locations and files indicated below, the output does need to be reorganized manually in Excel to achieve the presentation that appears in the manuscript.

Figure/Table	Program	Output file	Note
Table 1	Estimating Effects on Grades & Study Time Across All Cohorts.do Analysis of Year 6 (2019-20) Cohort.do	Stata Output Window	Execute the code from lines 62-69 and see the output. (Columns 2 to 9) Execute line 131 and see the output. (Column 10)
Table 2	Estimating Effects on Grades & Study Time Across All Cohorts.do	Table_2.xml	
Table 3	Estimating Effects on Grades & Study Time Across All Cohorts.do	Table_3.xml	In publicly available data, can only replicate panel B without campus fixed effects.
Table 4	Analysis of Year 5 (2018-19) Cohort & the April 2022 Study Time Followup.do	Table_4.xml	In publicly available data, can only replicate columns 1, 3, 5 (without controls) and we do so without campus fixed effects (table in paper has campus fixed

			effects).
Table 5	Analysis of Year 6 (2019-20) Cohort.do Analysis of Year 5 (2018-19) Cohort & the April 2022 Study Time Followup.do	Stata Output Window	Columns 4 & 5 (line 151) Columns 2 & 3 (line 386)
Table 6	Analysis of Year 5 (2018-19) Cohort & the April 2022 Study Time Followup.do	Table_6.xml	In publicly available data, can only replicate columns 1, 3, 5 (without controls).
Table 7	Analysis of Year 6 (2019-20) Cohort.do	Table_7.xml	In publicly available data, can only replicate columns 1, 3, 5, 7 (without controls).
Table 8	Analysis of Year 5 (2018-19) Cohort & the April 2022 Study Time Followup.do	Table_8.xml	In publicly available data, can only replicate without campus fixed effects (table in paper has campus fixed effects).
Table A1	Estimating Effects on Grades & Study Time Across All Cohorts.do	Table_A1.xml	Cannot replicate with publicly available data
Table A2	Estimating Effects on Grades & Study Time Across All Cohorts.do	Stata Output Window	Execute the code from lines 204-280 and see the output
Table B1	Estimating Effects on Grades & Study Time Across All	Table_B1.xml	Cannot replicate with publicly available data

	Cohorts.do		
Table B2	Estimating Effects on Grades & Study Time Across All Cohorts.do	Table_B2.xml	
Table B3	Estimating Effects on Grades & Study Time Across All Cohorts.do	Table_B3.xml	Cannot replicate with publicly available data
Table B4	Estimating Effects on Grades & Study Time Across All Cohorts.do	Table_B4.xml	Cannot replicate with publicly available data
Table B5	Analysis of Year 5 (2018-19) Cohort & the April 2022 Study Time Followup.do	Table_B5_diff.xml Table_B5_baseline.xml	Cannot replicate with publicly available data
Table B6	Analysis of Year 5 (2018-19) Cohort & the April 2022 Study Time Followup.do Analysis of Year 6 (2019-20) Cohort.do	Table_B6b_col2&3.xml Table_B6b_col4&5.xml	Cannot replicate with publicly available data
Figure 1a	Analysis of Year 5 (2018-19) Cohort & the April 2022 Study Time Followup.do		
Figure 1b	Analysis of Year 6 (2019-20) Cohort.do	Figure_1b.png	
Figure 2	Analysis of Weekly Dynamics in Year 6 (2019-20) Cohort.	Figure_2.png	

Figure 3a	Analysis of Year 5 (2018-19) Cohort & the April 2022 Study Time Followup.do	Figure_3a.png	
Figure 3b	Analysis of Year 5 (2018-19) Cohort & the April 2022 Study Time Followup.do	Figure_3b.png	
Figure 3c	Analysis of Year 6 (2019-20) Cohort.do	Figure_3c.png	
Figure 3d	Analysis of Year 6 (2019-20) Cohort.do	Figure_3d.png	
Figure 4a	Analysis of Year 5 (2018-19) Cohort & the April 2022 Study Time Followup.do	Figure_4a.png	
Figure 4b	Analysis of Year 5 (2018-19) Cohort & the April 2022 Study Time Followup.do	Figure_4b.png	
Figure 4c	Analysis of Year 6 (2019-20) Cohort.do	Figure_4c.png	
Figure 4d	Analysis of Year 6 (2019-20) Cohort.do	Figure_4d.png	
Figure 5a	Analysis of Weekly Dynamics in Year 6 (2019-20) Cohort.	Figure_5a.png	
Figure 5b	Analysis of Weekly Dynamics in Year 6 (2019-20) Cohort.	Figure_5b.png	

Figure 5c	Analysis of Weekly Dynamics in Year 6 (2019-20) Cohort.	Figure_5c.png	
Figure 5d	Analysis of Weekly Dynamics in Year 6 (2019-20) Cohort.	Figure_5d.png	
Figure 6	Analysis of Year 5 (2018-19) Cohort & the April 2022 Study Time Followup.do	Figure_6.png	
Figure A1	Estimating Effects on Grades & Study Time Across All Cohorts.do	Figure_A1.png	
Figure A2	Estimating Effects on Grades & Study Time Across All Cohorts.do	Figure_A2.png	
Figure A8	Code for Figure A8.do	Figure_A8.png	Cannot replicate with publicly available data
Figure A9	Estimating Effects on Grades & Study Time Across All Cohorts.do	Figure_A9a.png to Figure_A9f.png	
Figure B1a	Analysis of Year 5 (2018-19) Cohort & the April 2022 Study Time Followup.do	FigureB1_a.png	
Figure B1b	Analysis of Year 5 (2018-19) Cohort & the April 2022 Study Time Followup.do	FigureB1_b.png	

Figure B1c	Analysis of Year 6 (2019-20) Cohort.do	FigureB1_c.png	
Figure B1d	Analysis of Year 6 (2019-20) Cohort.do	FigureB1_d.png	