This README file provides a description of the data and programs used to generate the results in the paper "Semester or Quarters? The Effect of the Academic Calendar of Postsecondary Student Outcomes", published in AEJ: Policy. Please contact Stefanie Fischer, stefaniejfischer@gmail.com or Valerie Bostwick, vkbostwick@gmail.com with any questions.

Replication data and do files can be found at ICPSR repository project ID: **openicpsr-124861**. To run any of the programs, one will first have to change the file path at the top of the .do files.

Description of the data

Publicly available data

- The main analysis uses the Integrated Postsecondary Education Data System (IPEDS) data. These data are administered by the US Department of Education and were downloaded from the National Center for Education Statistics in 2016: https://nces.ed.gov/ipeds/. They have been deposited in the openicpsr-124861 repository in the "Data" folder.

  These data are used for the institution-level analysis outlined in Section 3 of the manuscript and are used to produce the results in Tables 1-3, Appendix Tables A1-A2, Figures 1-2 and Appendix Figures A1-A3.

Restricted-access data

- The mechanism analysis uses the Ohio Longitudinal Data Archive (OLDA) data. These are restricted-access data files but may be obtained with Data Use Agreements with the Center for Human Resources Research at Ohio State University. Researchers interested in access to the data may contact Dr. Lisa Neilson at ohioanalytics@oerc.osu.edu also see http://chrr.osu.edu/projects/ohio-longitudinal-data-archive. It can take some months to negotiate data use agreements and gain access to the data. The author will assist with any reasonable replication attempts for two years following publication.

  These data are used for the individual-level analysis portion of this project outlined in Section 4 in the manuscript. They are used to produce the results in Tables 4-10, Appendix Tables A3-A8, Figures 3-7 and Appendix Figure A4.

- The employment analysis uses the Ohio Department of Job and Family Services (ODJFS) data. These are restricted-access data files but may be obtained with Data Use Agreements with the Center for Human Resources Research at Ohio State University. Researchers interested in access to the data may contact Dr. Lisa Neilson at ohioanalytics@oerc.osu.edu also see http://chrr.osu.edu/projects/ohio-longitudinal-data-archive. It can take some months to negotiate data use agreements and gain access to the data. The author will assist with any reasonable replication attempts for two years following publication.

  These data are used for the individual-level employment analysis portion of this project outlined in Section 4.6 in the manuscript. They are used to produce the results in Tables 8-10, Appendix Tables A7-A8 and Figures 6-7.

Supplementary Publicly Available Data

- The American Community Survey data used to support the findings of this study have been deposited in openicpsr-124861 repository in the "Data" folder. They can be found at this site: https://doi.org/10.18128/D010.V10.0. The following site contains the zip code files that we use in our analysis: https://www.psc.isr.umich.edu/dis/census/Features/tract2zip.

  These data are used to perform a heterogeneity analysis presented in Table 5, columns 6 and 7.

- National Center for Education Statistics (NCES) Classification of Instructional Programs (CIP) Crosswalk data are used to identify a student's major choice. These data come from the following sites: https://nces.ed.gov/ipeds/cipcod and https://www.ohiohighered.org/node/2104. These data have been deposited in the openicpsr-124861 repository in the "Data" folder.

  These data are used in the major switching analysis found in Tables 5, 6, A4, A5, A6, Figure 5 and Figure A4.

- Bureau of Labor Statistics (BLS) Consumer Price Index (CPI) data are used to convert earnings data from nominal into real dollars. These data come from the following site: https://www.bls.gov/cpi/data.htm. These data have been deposited in the openicpsr-124861 repository in the "Data" folder.

  These data are used for the individual-level employment analysis portion of this project outlined in Section 4.6 in the manuscript.

Software requirements

Stata: The intuitional-level analysis was last run with Stata 14.2 and the individual-level analysis was last run with Stata 16.1. All programs were last run on 10/19/2020.

User-written Stata commands to download:

- outreg2 (as of 10/19/2020)
- sutex (as of 10/19/2020)
- coefplot (as of 10/19/2020)
- reghdfe (as of 10/19/2020)
- boottest (as of 10/19/2020)

Hardware requirements

The analysis was last run with an Intel® Core™ i5-8500 CPU @ 3.00GHz 3.00 GHz with 32.0GB installed RAM with the Windows 10 operating system.

Description of the programs

The following programs clean the data for the main analysis (the institution-level analysis).

Master_Create_IPEDS.do: This program creates the cleaned IPEDS data for the main analysis. The resulting dataset is: ipeds_cleaned_final.dta. The program: (1) cleans and merges the six IPEDS files on

enrollment, graduation rates, tuition, faculty, finances and calendar system, and (2) generates relative time indicators (pre and post semester adoption) for the event studies.

Master_Create_IPEDS.do calls in the following sub-programs:

- **directorycreate1_2018.do:**
  - This file cleans years 1987-2016 of directory files by calling in the sub-programs directory1987.do-directory2016.do and merges all the years together to create directorymaster1_2018.dta.
- **facultycreate1_2018.do:** 1990-2016
  - This file cleans years 1990-2016 of faculty files by calling in the sub-programs faculty1990.do- faculty2016.do and merges all the years together to create facultymaster1_2018.
- **gradcreate1_2018.do:** 1991-2010
  - This file cleans years 1991-2010 of graduation files by calling in the sub-programs grad1991.do-grad2010.do and merges all the years together to create gradmaster1_2018.dta.
- **institutioncreate1_2018.do:** 1986-2016
  - This file cleans years 1986-2016 of institution files by calling in the sub-programs institution1986.do-institution2016.do and merges all the years together to create institutionmaster1_2018.dta.
- **revenuecreate1_2018.do:**
  - This file cleans years 1991-2016 of the cost and revenue files by calling in the sub-programs publicrevenues1997.do- publicrevenues2016.do, privaterevenues1997.do-privaterevenues2016.do, costs1991.do-costs1996.do, revenues1991.do-revenues1996.do. and merges all the years together to create revenuesmaster1_2018.dta.
- **tuitioncreate1_2018.do:** 1990-2016
  - This file cleans years 1990-2016 of tuition files by calling in the sub-programs tuition1990.do- tuition2016.do and merges all the years together to create tuitionmaster1_2018.dta.

The following programs produce the results for the main analysis (institution-level analysis).

- **IPEDS_Analysis_Tables.do:** This generates:
  - Table 1, Output files: table_1.text
  - Table 2, Output files: table_2.text; table_2_means.text
  - Table 3, Output files: table_3a.text, table_3b.text
  - Table A2, Output files: table_3a_appa2.text, table_3b_appa2.text
- **IPEDS_Analysis _Table_A1:** This generates:
  - Table A1, Output files: table_A1a.text, table_A1b.text
- **IPEDS_Analysis_Figures.do**: This generates:
  - Figure 1, Output files: figure_1.pdf
  - Figure 2, Output files: figure_2a.pdf, figure_2b.pdf
  - Figure A1, Output files: figure_A1.pdf, figure_A1_DnD.text
  - Figure A2, Output files: figure_A2a.pdf, figure_A2a_DnD;
    figure_A2b.pdf, figure_A2b_DnD;
    figure_A2c.pdf, figure_A2c_DnD;
    figure_A2d.pdf, figure_A2d_DnD;
    figure_A2e.pdf, figure_A2e_DnD;

figure_A2f.pdf, figure_A2f_DnD

- Figure A3, Output files: figure_A3.pdf

The following programs clean the data for the mechanism analysis (the institution-level analysis). Note: the mechanism analysis uses restricted-access data and, as such, they are not available in the repository.

01_00_Create_Data_forAnalysis.do: This program cleans and combines raw data on enrollment, demographics, credits, GPA, high school zip code, major choice and degrees into 1 wide file with 1 observation per student. The resulting dataset is: Final_Enrollment_Data_SM99-SP17.dta.

01_00_Create_Data_forAnalysis.do calls in the following sub-programs:

- 01_01_01_Create_Enrollment_allOH_SM99-SP17.do
- 01_01_02_Create_Demographics.do
- 01_01_03_Calc_Credit_Hours.do
- 01_01_04_Clean_GPA.do
- 01_02_01_Merge_Files.do
  - 01_02_01a_Match_Majors_toCIP.do
- 01_02_02_Collapse_to_Individual.do
- 01_03_Merge_Degrees_FTFresh_allOH.do
- 01_04_Identify_Transfers_Out.do
- 01_05_01_Clean_HS_Data.do
- 01_05_02_Merge_HS_Data.do
- Create_CIPdtafiles.do

02_00_Create_Variables_forAnalysis.do: This program cleans data for the analysis and defines primary outcome variables and treatment variables. The two resulting datasets are: Full_Data_Sample.dta which include all cohorts (1999-2016), and Main_Data_Sample.dta which is the subset of cohorts observed for 4+ years (1999-2013).

The following programs produce the results for the mechanism analysis (individual-level analysis).

- **03_00_Run_Main_Analyses.do:** This generates Tables 4-7 and Figures 3-5.
  - The following sub-programs are called: 03_01_01_Heterogeneity_Analysis.do and 03_01_02_Event_Study_Analysis.do

The following programs clean the data for the employment analysis (the individual-level analysis). Note: the employment analysis uses restricted-access data and, as such, they are not available in the repository. Once access to the employment data has been approved, use the following procedure to extract the relevant sample from the Data Investigator:

1. Create a master list of key_ids from the HEI data you want to match employment data to. Save it as a .txt file as a single column of ids, with missings removed, no column header.
2. In Investigator choose as the substudy “ODJFS- UI Wages All Quarters (research extracts)”

3. In the Variable Search tab select the relevant data elements by clicking the boxes in the left column
4. When you have all of your variables selected, click the Advanced Download tab
5. Observation Selection: Use Id List - choose filepath from Step #1
6. Create download of data – select Stata dictionary file and codebook
7. Assign a filename at the bottom of the window and download.

04_00_Prep_Employment_Sample.do: This program cleans raw employment data and merges the cleaned employment data into the estimation sample. It also creates employment outcome variables. The resulting dataset is: FullSample_withEmp.dta

04_00_Prep_Employment_Sample.do calls in the following sub-programs:

- 04_01_Clean_UI_Data.do
- 04_02_Merge_UI_to_FullSample.do
    - 04_02a_Run_UI_Merge.do

The following programs produce the results for the employment analysis (individual-level analysis).

- **05_00_Run_Employment_Analyses.do:** This generates Tables 8-10 and Figures 6-7.
    - The following sub-program is called: 05_01_Employment_Event_Study.do

The following programs produce the results for the appendix (individual-level analysis).

- **06_00_Run_Appendix_Analyses.do:** This generates Tables A3-A8 and Figure A4.
    - The following sub-program is called: 03_01_02_Event_Study_Analysis.do