This README file attempts to replicate figure 1-2 and tables 1-3 in “Semester or Quarters? The effect of the academic calendar of postsecondary student outcomes”, published on AEJ: Policy. This file also includes an extension of the data to year 2016.

**Description of data:**

* This uses publicly available data from IPEDS, which can be accessed from here: <https://nces.ed.gov/ipeds/datacenter/DataFiles.aspx?year=-1&surveyNumber=-1&sid=f1df9aa1-9f9a-4aad-993f-77c865a91556&rtid=1>
* These data are used for the institution-level analysis outlined in Section 3 of the manuscript and are used to produce the results in Table 1-3, Appendix Tables A1-A2, Figures 1-2 and appendix Figures A1-A3.
* To download the correct datasets, you need to open the do-files provided in the replication package and check what dataset is used from the “label data” command in the do-files.

**Software requirements**

Stata: The institutional-level analysis was last run with Stata 18.0. All programs were last run on 2/22/2024.

User-written Stata commands to download:

* outreg2 (as of 2/22/2024)
* sutex (as of 2/22/2024)
* coefplot (as of 2/22/2024)
* reghdfe (as of 2/22/2024)
* boottest (as of 2/22/2024)

**Hardware requirements:**

The programs were last run with Apple M2 Pro with 16GB RAM with MacOS Ventura operating system.

**Description of the folders on GitHub:**

The directory, graduation rate, revenues, faculty, institution, and tuition folders hold do-files for cleaning the datasets in each above-mentioned category in varying years. Do-files before and including 2016 are accessed from the replication package. Do-files from 2016 onwards are downloaded from the IPEDS website under the Programs column in the same row of the corresponding dataset. Below is a description of what each of the code folders on GitHub hold.

* **Cost code**:
  + This folder contains programs costs1991.do – costs1996.do which cleans the cost datasets and gives costs1991.dta – costs1996.dta
* **Directory code**: 1987-2022
  + This folder contains programs directory1987.do – directory2022.do which cleans the directory datasets and gives directory1987.dta – directory2022.dta.
* **Faculty data cleaning code**:
  + This folder contains programs faculty1990.do – faculty2022.do which cleans the faculty datasets and creates faculty1990.dta – faculty2022.dta.
* **Graduation rate code**:
  + This folder contains programs grad1991.do – grad 2016.do which cleans the graduation rate datasets and creates grad1991.dta – grad2016.dta
* **Institution code**:
  + This folder contains programs institution1986.do – institution 2016.do which cleans the institution characteristics datasets and creates institution1986.dta – institution2016.dta
* **Public revenue code**:
  + This folder contains programs publicrevenues1997.do – publicrevenues2022.do which cleans the revenue datasets of public institutions and creates publicrevenues1997.dta – publicrevenues2022.dta
* **Private inst. Revenue code**:
  + This folder contains programs privaterevenues1997.do – privaterevenues2022.do which cleans the revenue datasets of private institutions and creates privaterevenues1997.dta – privaterevenues2022.dta
* **Revenue code**:
  + This folder contains programs revenues1991.do – revenues1996.do which cleans the revenues datasets (public and private institutions) and creates revenues1991.dta – revenues19962016.dta
* **Tuition cleaning code**:
  + This folder contains programs tuition1990.do – tuition2022.do which cleans the tuition datasets and creates tuition1990.dta – tuition2022.dta
* **Tables latex code**:
  + This folder contains the latex codes needed to output tables 1-3 in the desired format.

**Master list do-files**: This folder contains other do-files from each of the directory, graduation rate, revenue, tuition, faculty, and institution to create merged category lists.

**Master\_Create\_IPEDS.do** cleans and merges all six datasets and creates ipeds\_cleaned\_final.dta. This dataset is used for institution-level analysis and generates relative time indictors (pre and post semester adoption) for event studies. To merge all 6 datasets, it calls in the following sub-programs:

* **Directorycreate\_new\_2022.do**: 1987-2022
  + This file cleans years 1987-2022 of directory data by calling in sub-programs directory1987.do – directory2022.do. It merges all years together to create directorynew\_2022.dta.
* **Facultycreate\_new\_2022.do:** 1990-2022
  + This file cleans years 1990-2022 of directory data by calling in sub-programs faculty1990.do – faculty2022.do. It merges all years together to create facultynew\_2022.dta.
* **Gradcreate\_new\_2022.do**: 1991-2016:
  + This file cleans years 1991-2016 of graduation rate files by calling in sub-programs grad1991.do-grad2016.do and merges all years together to create gradmaster\_new\_2022.dta.
* **Institutioncreate\_new\_2022.do:** 1986-2022
  + This file cleans years 1986-2022 of institution files by calling in sub-programs institution1986.do-institution2022.do and merges all years together to create institutionmaster\_new\_2022.dta.
* **Revenuecreate\_new\_2022.do: 1991-2022**
  + This file cleans years1991-2022 of the cost and revenue files (revenue, private revenue, public revenue) by calling in sub-programs publicerevenues1997.do-2022.do, privaterevenues1997.do-privaterevenues2022.do, costs1991.do-costs1996.do, revenues1991.do-revenues1996.do. and merges all the years together to create revenuesmaster\_new\_2018.dta.
* **tuitioncreate1\_2018.do:** 1990-2022
  + This file cleans the tuition data from 1990 to 2022 by calling in sub-programs tuition1990.do- tuition2022.do and merges all the years together to create tuitionmaster\_new\_2022.dta.

**IPEDS\_Analysis\_Figures\_new.do** generates**:**

* Figure 1, Output files: figure\_1.pdf
* Figure 2 (panel A), output files: figure\_2a.pdf
* Figure 2 (panel B), output files: figure\_2b.pdf

**IPEDS\_Analysis\_figures\_new.do** generates:

* Table 1, Output files: table\_1.text
  + The do-file outputs a text log of Latex codes
* Table 2, Output files: table\_2.text, table2\_means.text
  + To output into a word document, change the format to table\_2.doc.
* Table 3A, Output files: table\_3a.text
* Table 3B, Output files: table\_3b.text