

README: Replication files for “The Lasting Effects of Working while in School: A Long-Term Follow-Up.”
Mery Ferrando, Noemi Katzkowicz, Thomas Le Barbanchon and Diego Ubfal.

July 29, 2025

Overview

The code in this replication package constructs all tables and figures in the text and Online Appendix using Stata. There is one master do file to clean the data, which calls 9 do files, and one master do file for the analysis, which generates all tables and figures and calls 8 do files.

The replicator should expect the master cleaning do file to run for 55 minutes and the master analysis do file to run for 32 minutes.

Data Availability and Provenance Statements

This paper uses confidential data from the Uruguayan Ministry of Labor and Social Security (MTSS) and Social Security Administration (BPS), and Universidad de la República (UDELAR). These data were accessed through a confidential agreement. Researchers interested in accessing the data will need to send applications and research proposals to MTSS, BPS, and UDELAR. The authors will assist with any reasonable replication attempts.

Statement about Rights

- I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

Summary of availability:

- Data cannot be made publicly available.

Data Sources

The project uses 7 datasets, which are described in the following table. The first four datasets were used to replicate and obtain the same sample from Le Barbanchon, Thomas, Diego Ubfal, and Federico Araya. “The Effects of Working While in School: Evidence from Employment Lotteries.” *American Economic Journal: Applied Economics*, 15, no. 1 (January 2023): 383–410. <https://doi.org/10.1257/app.20210041>. They were all obtained through confidential agreements, and a special authorization was requested from the Uruguayan Ministry of Labor and Security (MTSS) to re-use again.

The 5th dataset contained new data from Social Security Bank (BPS) and obtained through a confidential agreement between MTSS and the Instituto de Economía, Universidad de la República (UDELAR). The 6th dataset contains data from University Graduations and was obtained through a confidential agreement with UDELAR.

Finally, the 7th dataset is from the 2022 Household Survey in Uruguay: Instituto Nacional de Estadística Uruguay (2022): “Encuesta Continua de Hogares 2022,” The original dataset was retrieved at <https://www4.ine.gub.uy/Anda5/index.php/catalog/730/get-microdata>. It was transformed into Stata

and renamed “p22”; it includes some additional variables created by previous teams, but for our current analysis we only use the original variables in the survey.

Dataset	File name:	Location:	Date Accessed	Description:
1	data_newrequestmay3.dta	`project'/data/in/AEJapplied/	May 2023	Confidential
2	data_uniqueapp.dta	`project'/data/in/AEJapplied/	July 2021	Confidential
3	data_uniqueapp_alled_newreq2023.dta	`project'/data/in/AEJapplied/	April 2023	Confidential
4	earnings_panel_alled.dta	`project'/data/in/AEJapplied/	April 2021	Confidential
5	datos_dec2023.dta	`project'/data/in/MTSSiecon/	Dec 2023	Confidential
6	egresos_2006_2023.csv	`project'/data/in/UDELAR/	Oct 2024	Confidential
7	p22.dta	`project'/data/in/ECH/	April 2024	Original dataset is Public and Downloadable

Computational Requirements

This replication package requires Stata. The analysis was conducted in Stata 18.0 but the code was set to run with Stata version 16. Codes were run on a standard Dell Latitude 7440 laptop running Windows 11 with 32GB of installed RAM. Run time was 55 minutes for the master cleaning do file and 42 minutes for the master analysis do file.

The Stata do-file will require the following packages to be installed:

Stata packages: unique, distinct, estout, outreg, outreg2, coefplot (these dependencies are included in the package in the folder do/ado)

Description of code

All figures and tables in the paper can be generated by running first the master cleaning do file “O_cleaning_master.do” and then the analysis master do file “O_analysis_master.do”

- O_cleaning_master.do: runs 9 different do files
 - `project'/do/O_paths: sets file paths
 - `project'/do/cleaning/1_cleaning_emprecords.do: cleans employment records
 - `project'/do/cleaning/1a_ipc.do: used to convert nominal values in real terms
 - `project'/do/cleaning/1b_define_yrlyvariables_earnhours.do: used to define the yearly values of earnings

- `project'/do/cleaning/1c_define_yrlyvariables_sect.do: used to define sectoral earnings
- `project'/do/cleaning/1d_define_yrlyvariables_vf_contract.do: used to define type of contract
- `project'/do/cleaning/2_cleaning_ECH.do: used to clean data from Uruguay Household Survey.
- `project'/do/analysis/ 1b_irpf.do: used to generate income tax data (both for the MVPF analysis and to convert net to nominal incomes in ECH)
- `project'/do/cleaning/3_cleaning_university: used to clean data from university records

Note: The cleaning part of the code was **not shared** with the replication team because it contains multiple pieces of personal and identifiable information. For that reason, this portion of the code is **not included** in the replication package. Interested replicators can contact the author at dubfal@worldbank.org for further clarifications.

- 0_analysis_master.do: runs 8 different do files: in the `project'/do/analysis folder
 - 1a_ipc.do: used to convert nominal values in real terms
 - 1b_irpf.do: used to generate income tax data (both for the MVPF analysis and to convert net to nominal incomes in ECH)
 - 2a_cost_benefit.do: used to generate data for MVPF analysis
 - 2b_ipc_bpc.do: used to convert nominal values and generate income tax brackets for MVPF analysis
 - 2c_comparison other programs_21Vs8 years.do: used for Figure A3 to compare MVPF with other papers
 - 2d_Job Corps MVPF 8 years.do: used to calculate the 8-year MVPF for Job Corps for Figure A3.
 - 2e_get_tax_rate.do: used to calculate tax rates for the 8-year MVPF for Job Corps.
 - 2f_deflate_to.do: used to convert nominal values into real terms for the 8-year MVPF for Job Corps.

Instructions for replicators

- Sign a confidentiality agreement to obtain databases listed above as “Confidential,” and contact the researchers to obtain the cleaning do files, which contain confidential information and cannot be shared publicly.
- Add path file to file “0_paths.do”
- Install required packages: `ssc install distinct`, `ssc install unique`, `ssc install estout`, `ssc install outreg`, `ssc install coefplot`.
- Run `0_cleaning_master.do`, and then `0_analysis_master.do` to generate all tables and results.
- Create all relevant folders:
 - analysis, including subfolders: `results>cost-benefit`, `results>fig`, and `results>tab`
 - data, including subfolders: `in>AEJApplied`, `in>ECH`, `in>MTSSiecon`, `in>UDELAR`, `out>ECH`, `out>MTSiecon`
 - do: including subfolders analysis and cleaning

Code Locations for Figures and Tables

All Tables and Figures are produced by the file `O_analysis_master.do`. Tables and Figures are clearly indicated, and appear at the following lines, and stored in the listed output files:

Figure/Table #	Line Number	Output file
Figure 1	173	fig_qrtlyearn_d5.pdf
Table 1	327	tab_mainout.tex
Table 2	555	tab_aggsector.tex
Figure 2 Panel A	621	figure2a.pdf
Figure 2 Panel B	671	figure2b.pdf
Figure 2 Panel C	691	figure2c.pdf
Figure 2 Panel D	714	figure2d.pdf
Figure 2 Panel E	730	figure2e.pdf
Table A1	759	tab_prog.tex
Table A2	809	balance.tex
Table A3	897	tab_firststage.tex
Table A4	927	tab_returns_regmonth_cont.tex
Table A5	1001	tab_graduation.tex
Table A6	1056	tab_hetgender.tex
Table A7	1126	tab_hetvuln.tex
Figure A1	1198	fig_MVPFvertime.pdf
Figure A2	1241	fig_MVPFcomparison_21VS8.pdf
Table A8	1251	tabmain_nocontrols.tex
Table A9	1297	tabmain_clusloc.tex
Table A10	1341	tabmain_nowins.tex
Table A11	1418	tab_ITT.tex
Table A12	1463	tab_altdeft.tex
Table A13	1530	tab_bounds.tex

References

Instituto Nacional de Estadística Uruguay (2022). “Encuesta Continua de Hogares 2022.” Database retrieved at <https://www4.ine.gub.uy/Anda5/index.php/catalog/730/get-microdata>.

Le Barbanchon, Thomas, Diego Ubfal, and Federico Araya. “The Effects of Working While in School: Evidence from Employment Lotteries.” *American Economic Journal: Applied Economics* 15, no. 1 (January 2023): 383–410. <https://doi.org/10.1257/app.20210041>.