READ ME

This directory contains the data and code to replicate the results in

"Human Capital Accumulation and Development: The Role of On-the-Job Training" (Xiao Ma, Alejandro Nakab and Daniela Vidart), 14 October 2023

It should be noted that some of the estimation tasks are computationally intensive, and certain files may take several hours to execute even on a fast computer. Additionally, minor rounding errors may be present in the simulation. Please feel free to contact the authors with any questions or concerns regarding the replications.

1) SUMMARY

This table summarizes the files that replicate figures and tables in the paper:

Table 1	List of common parameters, which are described in Appendix H.1
Table 2	Run "quantitative\US calibration\main_simulation_run", extract targeted moments in model from lines 343-358
Table 3	Run "quantitative\US calibration\main_simulation_run", extract parameter values from lines 360-375
Table 4	Run "quantitative\Tables and Figures\Table 4\code_Table4.do"
Figure 1	Run "Figure1\code_Figure1"
Figure 2	Run "empirics\Clean Files\ESandCVT_results.do" (lines 167-174)
Figure 3	Run "empirics\ Clean Files\ESandCVT_results.do" (Fig 3.a in lines 177-181, Fig 3.b in lines 185-190)
Figure 4	Run "empirics\Clean Files\CVT_privatedata_results.do" (Fig 4.a lines 160-167, fig 4.b lines 150-157)
Figure 5	Run "empirics\Clean Files\AES_CVT_EU_results.do" (Fig 5.a in lines 297-303, Fig 5.b in lines 306-313, Fig 5.c in lines 316-322)
Figure 6	Run "quantitative\Tables and Figures\Figure 6\code_Figure6.do"
Figure 7	Run "quantitative\Tables and Figures\Figure 7\code_Figure7.do"
Figure 8	Run "quantitative\Tables and Figures\Figure 8 9 10\code_Figure_8_9.do"
Figure 9	Run "quantitative\Tables and Figures\Figure 8 9 10\code_Figure_8_9.do"
Figure 10	Run "quantitative\Tables and Figures\Figure 8 9 10\code_Figure_10.do"
Figure 11	Run "quantitative\Tables and Figures\Figure 11\code_Figure11.do"
Figure 12	Run "quantitative\Tables and Figures\Figure 12\code_Figure12.do"

This table summarizes the files that replicate figures and tables in the online appendix:

Table A.1	Review of literature
Table C.1	Examples and definitions following ISCED 2011
Figure D.1	Run "empirics\Clean Files\ESandCVT_results.do (lines 138-144)

Figure D.2	Run “empirics\ Clean Files\AES_CVT_EU_results.do (Fig D.2.a in lines 331-335, Fig D.2.b in lines 338-341, Fig D.2.c in lines 344-349, Fig D.2.d in lines 352-356)
Figure D.3	Run “empirics\ Clean Files\AES_CVT_EU_results.do (Fig D.3.a in lines 361-365, Fig D.3.b in lines 368-3722, Fig D.3.c in lines 376-380, Fig D.3.d in lines 384-389, Fig D.3.e in lines 392-397, Fig D.3.f in lines 400-405)
Figure D.4	Run “empirics\ Clean Files\ESandCVT_results.do (Fig D.4.a in lines 146-151, Fig D.4.b in lines 153-158)
Figure D.5	Run “empirics\ Clean Files\AES_CVT_EU_results.do (lines 409-414)
Table D.1	Run “empirics\Clean Files\CVT_privatedata_results.do” (lines 184-224)
Figure E.1	Run “empirics\Clean Files\AES_privatedata_composition.do (lines 303-322)
Figure E.2	Run “empirics\Clean Files\AES_privatedata_composition.do (lines 518-524)
Figure E.3	Run “empirics\Clean Files\AES_privatedata_composition.do (Industry in lines 371-379, occupation in lines 408-415, education in lines 445-452, firm size in lines 481-488)
Table E.1	Run “empirics\Clean Files\AES_privatedata_composition.do (Industry in lines 366-370, occupation in lines 401-405, education in lines 437-441, firm size in lines 473-477, all in lines 507-513)
Table F.1	Run “empirics\Clean Files\ESandCVT_results.do (lines 268-301)
Figure H.1	Run “quantitative\Tables and Figures\Figure H1\code_FigureH1.do”
Figure I.1	Run “quantitative\cross_country\main_file”, “quantitative\robustness checks\endogenous layoff\cross_country\main_file”, “quantitative\robustness checks\firm_productivity_dist\main_file”, “quantitative\robustness checks\learning_by_doing\cross_country\main_file”, “quantitative\robustness checks\no_contract_breaking\cross_country\main_file” to obtain the quantitative results from result_report variable for the cases of baseline, considering different productivity distributions, considering learning-by-doing, considering no breaking-contract costs and considering endogenous layoffs. Finally, run “quantitative \robustness checks\country_specific calibration\summarize_code.do” for the results for the case of country-specific calibration.
Table I.1	Run “quantitative\Tables and Figures\Figure H1\code_FigureI1.do”
Table I.2	Run “quantitative\Tables and Figures\Table I2\code_TableI2.do”

2) DIRECTORY STRUCTURE

Files should be run in the order that they are listed below.

empirics	Contains the code to replicate the empirical results in the paper
quantitative	Contains the code to replicate the quantitative model in the paper

3) Original Data

To generate the input database required to produce the results, it is necessary to use the ".do" files located in the "Clean Files" folder, which clean the raw data. Once cleaned, the databases are saved in the same folder to be utilized by the following primary ".do" files that generate the desired output:

ESandCVT_results.do
AES_CVT_ results.do
CVT_ privatedata _ results.do
AES_privatedata_composition.do

To access the publicly available raw data and clean them for obtaining all inputs required for the aforementioned ".do" files, please refer to the following sequence of folders where the data is stored.

Raw Data	Folder raw data	Do file that prepare data ("Clean Files" folder)
Data on separation rates from Donovan, Lu and Shoellman (2023)	Donovan, Lu and Shoellman 2019 Raw	Cleaning_DLS_seprates.do
EUROSTAT public data (AES and CVT surveys) downloaded from https://ec.europa.eu/eurostat /web/main/data/database	EUROSTAT public data\Public Raw Data: • aes_101 • aes_121 • aes_147 • aes_189 • aes_202 • cvt_01s • cvt_12s • cvt_13s • cvt_19s • cvt_22s • cvt_25s • cvt_29s • cvt_34s	Cleaning_AES_publicdata.do Cleaning_CVT_publicdata.do
Penn world table version 9.1	Penn World Table 9.1 Raw	Cleaning_pwt91.do
Penn world table version 10.1	Penn World Table 10 Raw	Cleaning_pwt101.do
World Bank Enterprise Survey Wave 1 (2000-2005) and wave 2 used for main analysis (2005-2017).	World Bank Enterprise Survey Raw	Cleaning_WBES.do
Worldwide Governance Indicators downloaded from https://databank.worldbank.org /source /worldwide-governance-indicators in 2019	Worldwide Governance Indicators Raw	Cleaning_WGI.do (this do file just move locations of .dta)
Worldbank self employment data downloaded from WB data in 2019	Self Emp WB Raw	Cleaning_SelfEmpWB.do (this do file just move locations of .dta)
Worldbank GDP PPP data downloaded from WB data in 2019	Clean Files: gdp_percapita_ppp.dta	
EU AES survey private data (can get access applying directly through		AES_privatedata_composition.do

EUROSTAT: https://ec.europa.eu/eurostat/web/microdata)		
EU CVT survey private data (can get access applying directly through EUROSTAT: https://ec.europa.eu/eurostat/web/microdata)		CVT_privatedata_results.do

To replicate all empirical results follow the order:

Cleaning_pwt91.do
Cleaning_pwt101.do
Cleaning_DLS_seprates.do
Cleaning_WGI.do
Cleaning_SelfEmpWB.do
Cleaning_AES_publicdata.do
Cleaning_CVT_publicdata.do
Cleaning_WBES.do

Then get results in following order (to get results from ESandCVT_results need to first run AES_CVT_results):

AES_CVT_ results.do
ESandCVT_results.do

Then get access to AES and CVT through EUROSTAT Microdata Unit (how to access data is detailed in folder "replication file\empirics\AES and CVT private data access information") and run:

CVT_ privatedata _ results.do
AES_privatedata_composition.do

4) Software Requirements (Windows 10 Version)

Stata 16.0 released on June 2019
Matlab R2019b released on Sep 2019