Replication files for: "Occupational exposure to capital-embodied technical change"

The replication package contains two main blocks, empirics and quantitative.

(a) Empirics includes Stata do-files that process and analyze the relevant series for this project, including the results in Sections 2 and 3 of the paper. To ease the usage of the dataset we provide capital and price series by occupation that can be directly downloaded by the user in folder "empirics/finaldata". This data is also shared through our dedicated web at https://capitalbyoccupation.weebly.com/downloads.html. Empirical results can be replicated by running "[0]Master-run.do".

(b) Quantitative includes Matlab m-files which run the accounting exercises presented in Section 5 and 6 of the paper. Quantitative results can be replicated by running "A-main.m".

There are a total of 28 do files, 33 m files.

To extract tool requirements from historical DOT data, we refer the user to a supplementary CodeOcean capsule containing additional 5 python routines. https://doi.org/10.24433/CO.5242776.v1

1 Empirics

Instructions. After downloading the replication folder, researchers should open the do file
.../empirics/dofiles/[0]Master-run.do file.
Researchers must specify the home directory in which the replication folder has been saved. This is the main program for the replication. The program installs the necessary packages and then calls the do files that build the intermediate and final datasets by processing the raw data. This do-file is organized in three blocks, one for the data construction, one for tables, and one for figures. Each block can be run independently. Results shown in the paper and in the appendix are stored under */empirics/results/*

If researchers would like to explore the NPL routines used to extract tool requirements from historical DOT data, they can access those routines at https://doi.org/10.24433/CO.5242776.v1.

Packages. All stata packages needed are installed by [0]Master-run.do file (lines 34-35): weibullfit, ivreg2 and ranktest.

Hardware and duration: We use a iMac Apple M1 chip with 8-core CPU 16GB memory and the code takes less than 6 minutes to run.

Finally, we provide a capsule that would allow users to run the python routines that were used to extract the historical tool requirements of different occupations from the text of the DOT. "https://doi.org/10.24433/CO.5242776.v1"

Hardware and duration: When this code runs in an iMac Apple M1 chip with 8-core CPU 16GB memory, it takes approximately 9 hours to run.

1.1 Data sources

All data used in the analysis is public and can be found in the rawdata folder, as described in Table 1. Additional details on these sources and methodology can be found in the online appendix of the paper.

Table 1: Data sources and location

Data Name and url	Data File	Location replication file	Downloaded
NIPA Table, U.S. Bureau of Economic Analysis [2021a]	Table2.1_BEA_current_cost_stock.xls	/empirics/rawdata/BEA_NIPA	January-21
	Table2.7_BEA_current_cost_investment.xls		
	Table5.5.4*		
	Table5.6.4*		
	Table6.1D_BEA_valueadded_ind.xls		
	Table115_BEA_GDP.xls		
	TableHistorical_BEA_valueadded_ind.xls		
ONET, U.S. Department of Labor [2016]	Tools and Technology_23_1.xlsx	/empirics/rawdata/Tool_data_ONET	January-19
Crosswalks, US Census Bureau [2016]	occ1950_occ1990dd.dta	/empirics/rawdata/Crosswalks	January-19
	occ1960_occ1990dd.dta		
	occ1970_occ1990dd.dta		
	occ1980_occ1990dd.dta		
	occ1990_occ1990dd.dta		
	occ2000_occ1990dd.dta		
	occ2005_occ1990dd.dta		
	occ2010_occ1990dd.dta		
IPUMS-CPS, Flood et al. [2019]	cps_00003.dat	/empirics/rawdata/IPUMS_CPS	January-19
DOT, U.S. Department of Labor [1991] ICPSR membership required	'DS0013/06100-0013-Data.txt', 'DS0011/06100-0011-Data.txt', DS0017/06100-0017-Data.txt'	https://doi.org/10.24433/CO.5242776.v1	January-21
FED-PCEPI, U.S. Bureau of Economic Analysis [2021b]	Raw_Data_FED.xls	/empirics/rawdata/FED_PCEPI	June-21
FED-working-age-population Organization for Economic Co-operation and Development [2021]	working_age_pop.xls	/empirics/rawdata/FED_working_age_population	June-21
Quality-adj. price of capital, DiCecio [2009]	Capital_embodiment_measure.xls	/empirics/rawdata/PriceKquality	Unpublished Data *
WDI, World Bank National Accounts [2021]	Data_trade_US.dta	/empirics/rawdata/WDI	June-21

Table 5.5.4* full reference: Table5.5.4_Price_Indexes_for_Private_Fixed_Investment_in_Equipment_by_Type.xls
Table5.6.4* full reference: Table5.6.4_Price_Indexes_for_Intellectual_property_product.xls
*Data shared by R. Dicecio.

2 Quantitative

Instructions. The files in the folder replication_quantitative produce the main outcomes of the quantitative analysis in Sections 4.2, 5, and 6 – that is, model calibration, counterfactuals, alternative exercises, and predictions, and produce corresponding tables (in *tables_paper.xslx*, sheet *Tables_paper*) and figures (numbered as in the paper), saved in the folder *quantitative/results*. All files are coded for Matlab and saved in the folder *quantitative/codes*. The file the runs all the above mentioned exercises is *A_main.m*. This file calls sub-routines which can be run separately and are described below.

Data preparation. The main file that loads and prepares the data generated in the *Empirics* is *AA_data.m*; this program calls the following programs:

step 1: *estimates_load.m* – loads the estimates of the elasticity of substitution between capital and labor, σ_o, the tail of the Frechet distribution, θ, and the elasticity of substitution across capital goods, ϕ.

step 2: *data_load.m* – loads the occupational data on capital, employment, and their rental rates at both annual and decennial frequency and saves these data in the files $data\_info\_yby.mat$, $data\_main\_yby.mat$ (annual), $data\_info.mat$, $data\_main.mat$ (decennial).

step 3: *data_derive.m* – derives statistics on the data loaded in step 2.

step 4: *figures_appendix.m*– generates Figures BI and BII.

Estimation or the elasticity of substitution across occupational output. The main file for the estimation of the elasticity of substitution across occupational goods, ρ, is *AA_estimation_rho.m*; this program calls the following programs:

step 1: *sel_elast.m* and *data_for_model.m* – computes the statistics needed in the exercise (annual frequency).

step 2: *model_solve.m* – runs the calibration of the model. Note 1: The notation used in the code separates the matrix T_{oh} in two components, T_h and z_{oh}, and calls the occupational price per efficiency unit of labor $\tilde{\varkappa}_o$ instead of λ_o^n, so that, for example, wages can be written as $w_h = (T_h \sum_o (z_{oh}\tilde{\varkappa}_u)^\theta)^{1/\theta}$ instead of $w_h = (\sum_o (T_{oh}\tilde{\varkappa}_u)^\theta)^{1/\theta}$. Note 2: Despite not of relevance for the specifics of the calibration in the paper, the code allows for the case where the occupational production technology is Cobb-Douglas in capital and labor ($\sigma = 1$).

step 3: *est_rho.m* – runs the estimation of ρ and saves the IV estimates in the $rho.mat$.

Baseline exercise. The main file that runs the main counterfactual and prediction exercises is *AA_baseline_model.m*; this program calls the following programs:

step 1: Model fit:

step 1.1: *sel_elast.m* and *data_for_model.m* – computes the statistics needed in the exercise (decennial frequency).

step 1.2: *model_solve.m* – runs the calibration of the model.

step 1.3: *model_fit.m* – runs Table B.IV.

step 2: Counterfactuals:

step 2.1: *z_decomposition.m* – runs the decomposition of the T_{oht} matrix for the counterfactuals.

step 2.3: *main_counterfactuals.m* – runs the main counterfactual exercise outlined in Section 5.1.

step 2.3: *counterfactual_stats.m* and *counterfactual_figures.m* – runs column *CETC/baseline* of Tables 1, 2, columns *CETC/baseline, demand, demographics, composition, CA* of Table B.IV, Table B.V, and Figures 5 and 6.

step 3: Prediction:

step 3.1: *sel_elast.m* and *data_for_model.m* – computes the statistics needed in the exercise (annual frequency).

step 3.2: *model_solve.m* – runs the calibration of the model at yearly frequency.

step 3.2: *prediction.m* – runs the prediction exercise and Figure 7.

Alternative exercises. The main file that runs the alternative exercises is *AA_alternative_exercise.m*; this program calls a subset of the programs in the baseline exercise and outputs columns *CETC/identical elasticities* and *CETC/identical CETC* of Tables 1, 2, and B.VI.

Multiple capital goods. The main file that runs the main counterfactual in the model with multiple capital goods is *AA_model_multiple_equipment.m*; this program calls a subset of the programs in the baseline exercise and outputs columns *CETC/computers*, *CETC/communication equipment* and *CETC/software* of Table 3.

The remaining Matlab files in the folder are sub-routines.

Packages. Users must have installed Matlab's Optimization Toolbox. The code was run on Matlab R2022b.

Hardware and duration: MacBook Pro M1 (2020, 8CPU) 816.04 seconds. Time info is saved in *A-main* file.

A complete reference of tables and figures can be found in Table 2.

Table 2: Tables and Figures, reference

Figure/Table Number	Program	Output file
Figure 1	empirics/figures/Figure 1.do	empirics/Results/[Figure1_panelA]Perworker_lev_stock-byoccupation_DOTONET_base85_tmech
		empirics/Results/[Figure1_panelB]usercostbytime_implied_052021_base85_tmech.pdf
Figure 2	empirics/figures/Figure2.do	empirics/Results/[Figure2_panelA]employment_cetc_intuition_base85_tmech.pdf
		empirics/Results/[Figure2_panelB]inputratio_cetc_intuition_base85_tmech.pdf
Figure 3	empirics/figures/Figure3.do	empirics/Results/[Figure3]elasticities_1d.pdf
Figure 4	empirics/figures/Figure4.do	empirics/Results/[Figure4]Occupational_exposure_CETC.pdf
Figure 5	quantitative/codes/counterfactual_figures.m	quantitative/Results/Figure_5_panel1.pdf
		quantitative/Results/Figure_5_panel2.pdf
Figure 6	quantitative/codes/counterfactual_figures.m	quantitative/Results/Figure_6_panel1.pdf
		quantitative/Results/Figure_6_panel2.pdf
Figure 7	quantitative/codes/prediction.m	quantitative/Results/Figure_7_panel1.pdf
		quantitative/Results/Figure_7_panel2.pdf
		quantitative/Results/Figure_7_panel3.pdf
		quantitative/Results/Figure_7_panel4.pdf
Figure B.I	quantitative/codes/tool_shares.m	quantitative/Results/Figure_BI_panel1.pdf
		quantitative/Results/Figure_BI_panel2.pdf
Figure B.II	quantitative/codes/tool_shares.m	quantitative/Results/Figure_BII_panel1.pdf
		quantitative/Results/Figure_BII_panel2.pdf
Figure B.III	empirics/figures/Figure B.III.do	empirics/Results/[FigureB.III]inputratio_cetc_intuition_base85_tmech_demog.pdf
		empirics/Results/[FigureB.III]inputratio_cetc_intuition_base85_tmech_demogsk.pdf
Table 1	quantitative/codes/AA_model_baseline.m	quantitative/Results/Tables_paper.xls sheet "tables_paper"
	quantitative/codes/AA_alternative_exercise.m	
Table 2	quantitative/codes/AA_model_baseline.m	quantitative/Results/Tables_paper.xls sheet "tables_paper"
	quantitative/codes/AA_alternative_exercise.m	
Table 3	quantitative/codes/AA_model_baseline.m	quantitative/Results/Tables_paper.xls sheet "tables_paper"
	quantitative/codes/AA model multiple equipment.m	
Table B.I	empirics/tables/TableBI.do	empirics/Results/[TableB.I]CETC and changes in the labor market 1984-2015.xls
Table B.II	empirics/tables/TableBII.do	empirics/Results/[TableB.II]P_values_elasticities.xls
Table B.III	empirics/tables/TableBIII.do	empirics/Results/[TableB.III]Elasticity of substitution between capital and labor.xls
Table B.IV	quantitative/codes/AA_model_baseline.m	quantitative/Results/Tables_paper.xls sheet "tables_paper"
Table B.V	quantitative/codes/AA_model_baseline.m	quantitative/Results/Tables_paper.xls sheet "tables_paper"
Table B.VI	quantitative/codes/AA_model_baseline.m	quantitative/Results/Tables_paper.xls sheet "tables_paper"
	quantitative/codes/AA_alternative_exercise.m	

References

Riccardo DiCecio. Sticky wages and sectoral labor comovement. *Journal of Economic Dynamics and Control*, 33(3):538 – 553, 2009. ISSN 0165-1889. doi: https://doi.org/10.1016/j.jedc.2008.08.003. URL http://www.sciencedirect.com/science/article/pii/S016518890HICK8001449.

Sarah Flood, Miriam Kind, Renae Rodgers, Steven Ruggles, and Robert Warren. Integrated Public Use Microdata Series, Current Population Survey: Version 9.0 [dataset]. Technical report, 2019. https://doi.org/10.18128/D030.V9.0.

Organization for Economic Co-operation and Development. Working age population: Aged 15-64: All persons for the united states [lfwa64ttusa647n], June 2021. https://fred.stlouisfed.org/series/LFWA64TTUSA647N.

U.S. Bureau of Economic Analysis. Nipa tables, 2021a. https://apps.bea.gov/iTable/?reqid=19&step=2&isuri=1&categories=survey.

U.S. Bureau of Economic Analysis. Personal consumption expenditures: Chain-type price index [pcepi], June 2021b. https://fred.stlouisfed.org/series/PCEPI.

US Census Bureau. Industry and Occupation Code Lists Crosswalks: [2010 Census Occupation Codes with Crosswalk], 2016. URL https://www.census.gov/topics/employment/industry-occupation/guidance/code-lists.html.

U.S. Department of Labor. Dictionary of occupational titles: Revised fourth edition [dataset], 1991. https://doi.org/10.3886/ICPSR06100.v1.

U.S. Department of Labor. Onet version 23.1 [dataset]. Used under the cc by 4.0 license, Employment and Training Administration, 2016. https://www.onetcenter.org/db_releases.html.

World Bank National Accounts. Exports of goods and services (% of gdp) - united states, 1960-2020 [ne.exp.gnfs.zs]. Used under the cc by 4.0 license, World Bank, 2021. https://data.worldbank.org/indicator/NE.EXP.GNFS.ZS?name_desc=true&locations=US.