

# Replication files for: “Occupational exposure to capital-embodied technical change”

The replication package contains two main blocks. (a) Empirics, which 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 also provide outputs from the data construction exercise 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>. (b) Quantitative, which includes Matlab m-files which run the accounting exercises presented in Section 5 and 6 of the paper.

## Empirics

The raw data subfolder includes public data from the following sources. Additional details on these sources and methodology can be found in the online appendix of the paper.

1. BEANIPA: Nipa Tables from U.S. Bureau of Economic Analysis (Retrieved on June 14, 2021).
2. Crosswalks: Occupation crosswalks from Deming, D. J. (2017). The growing importance of social skills in the labor market.
3. DOT-1977: 4th edition of the Dictionary of Occupational Titles, <https://www.icpsr.umich.edu/web/ICPSR/studies/6100>
4. FED-PCEPI: Federal Reserve Bank of St. Louis (Retrieved on June 14, 2021).
5. FED-working-age-population: FRED, Federal Reserve Bank of St. Louis (Retrieved on January 2019).
6. IPUMS-CPS: Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren and Michael Westberry. Integrated Public Use Microdata Series, Current Population Survey: Version 9.0 [dataset]. Minneapolis, MN: IPUMS, 2021. <https://doi.org/10.18128/D030.V9.0> (Retrieved on January 2019).
7. Quality-adjusted price of capital: DiCecio (2009) extending the methodology in Cummins and Violante (2002). Most updated data was kindly provided by the author on January 2019.
8. Tool-data-ONET: ONET Data from: [https://www.onetcenter.org/db\\_releases.html](https://www.onetcenter.org/db_releases.html) (see version 23.1 of the tools and technology module). (Retrieved on January 2019). The data includes information from the O\*NET 23.1 Database by the U.S. Department of Labor, Employment and Training Administration (USDOL/ETA). Used under the CC BY 4.0 license. O\*NET® is a trademark of USDOL/ETA.
9. WDI: World Bank. “Exports of goods and services (% of GDP)” World Development Indicators. The World Bank Group. (retrieved on June 2021).
10. Depreciation rates estimates from BEA by equipment type [https://apps.bea.gov/national/pdf/BEA\\_depreciation\\_rates.pdf](https://apps.bea.gov/national/pdf/BEA_depreciation_rates.pdf), (retrieved on January 2019).

All the data used in the analysis is public and it is contained in the rawdata folder.

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/*

NOTE: We provide a dedicated folder “*empirics/toolsDOT*” that contains the python routines used to extract the historical tool requirements of different occupations.

## Quantitative

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.xlsx*, 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 that 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,  $\sigma_o$ , the tail of the Frechet distribution,  $\theta$ , and the elasticity of substitution across capital goods,  $\phi$ .
- 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 of the elasticity of substitution across occupational output.** The main file for the estimation of the elasticity of substitution across occupational goods,  $\rho$ , 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  $\tilde{\lambda}_o$  instead of  $\lambda_o^n$ , so that, for example, wages can be written as  $w_h = (T_h \sum_o (z_{oh} \tilde{\lambda}_u)^\theta)^{1/\theta}$  instead of  $w_h = (\sum_o (T_{oh} \tilde{\lambda}_u)^\theta)^{1/\theta}$ .

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  $\rho$  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  $T_{oh}$  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.

## References

- CUMMINS, J. G. and VIOLANTE, G. L. (2002). Investment-Specific Technical Change in the US (1947-2000): Measurement and Macroeconomic Consequences. *Review of Economic Dynamics*, **5** (2), 243–284.
- DICECIO, R. (2009). Sticky wages and sectoral labor comovement. *Journal of Economic Dynamics and Control*, **33** (3), 538 – 553.