

# Replication package for “Automation and New Tasks: How Technology Changes Labor Demand”

Daron Acemoglu and Pascual Restrepo

The replication package consists of three components: **Data**, **Programs** and **Outputs**.

## Data

The Data component of the paper is made up by the following folders:

### 1. **raw\_data**

#### I) aggregates includes:

- values of consumer price index (*consumer\_price\_pce.dta*), employment (*employment.dta*), population (*population.dta*) from BEA;
- manufacturing output (*quantities\_manuf.dta*) and total factor productivity *tfp.dta* between 1947 and 2017 from FRED, St. Louis FED;
- estimates on cost of capital (*rognlie\_r.dta*) from Rognlie (2015),
- tax rates (*taxes\_KN.dta*) from Karabarbounis and Neimann (2016).

#### II) bea\_klems includes:

two files reporting the data underlying the BEA/BLS integrated Industry-level Production Account for the United States respectively for the years:

- 1987-1998 (*BEA-BLS-industry-level-production-account-1987-1998.xlsx*)
- 1998-2016 (*BEA-BLS-industry-level-production-account-1998-2016.xlsx*)

#### III) industry includes:

- Measure of offshoring based on Feenstra and Hanson (1999) and extended by Wright (2014) (*FH\_offshoring\_naics.dta*)

- Data on Value Added, GDP and Worker Compensation from BEA Industry Accounts following 2007 NAICS classification for the years 1947 - 2014 (*GDPbyInd\_VA\_1947-2017.xlsx*) and following SIC 1972 classification for the years 1947 - 1987 (*GDPbyInd\_VA\_SIC.xls*)
- Data on hours worked from BLS Multifactor Productivity Measures (*hrs.xlsx*).
- Data on price and quantity of capital as well as on price, quantity and income of labor respectively for the:
  - \* manufacturing (*klemsmfbymeasure.xlsx*)
  - \* non-manufacturing (*klemsmfpxgbymeasure.xlsx*)
 sectors from the BLS Multifactor Productivity data.
- NBER-CES Manufacturing Industry Database at the NAICS (*naics5811.dta*) and SIC versions (*sic5811.dta*) for the years 1958 - 2011 . Cross-section of the same dataset for the year 1997 reporting only employment and total cost of materials (*NBER-ces-naics-emp.dta*)
- Data on:
  - \* stock of private equipment (Tables 3.1E, 3.2E and 3.4E),
  - \* investment in equipment (Tables 3.7E and 3.8E),
  - \* stock of all fixed assets (Tables 3.1ESI, 3.4ESI, 3.2ESI)
  - \* investment in fixed assets (Tables 3.7ESI, 3.8ESI)
 from NIPA, Fixed Asset Tables (*Section3All.xls.xlsx*).
- Data on adoption of automation technologies for the years
  - \* 1988 (*smt88.dta*)
  - \* 1993 (*smt93.dta*)
 from the Survey of Manufacturing Technology.
- Data on value added and labor share for a detailed set of industries reported in the BEA Input-Output Accounts (1987, 1992, 1997, 2002, 2007) harmonized into a consistent set of of four-digit industries classified according to the 1987 SIC from Christina Patterson.  
*Crosswalks used to create these data can be requested to their author: Christina Patterson at cpatt@mit.edu*

IV) ipums (raw data not shared due to its size. Please contact authors if you need it)  
 Census data on industry characteristics and on employment by occupation for the 1990 Census and the 2012-2016 ACS (*ipum\_census\_jep.dta*).

V) occs:

- Do files by Acemoglu and Restrepo (2018a) to extend the consistent classification of occupation titles developed by Autor and Dorn (2013)

to the most recent ACS vintages (*recode\_acs.do* and *create\_occ1990dd\_acs.do*).

The labels of the respective job titles are listed in *occ1990dd\_acs\_names.dta*

- Share of new job titles by occupation for: 1980 (*new1980-wk.dta*), 1991 (*new91-wk.dta*), 2000 (*new2000-wk.dta*) from Lin (2011).
- Crosswalks between U.S. Census occupation codes to balanced panel of occupations consistent across Census waves (“occ1990dd” classification) developed by David Dorn (Autor and Dorn, 2013) (*occ1980\_occ1990dd.dta*, *occ1990\_occ1990dd.dta*, *occ2000dd\_occ1990.dta*, *occACS\_occ1990.dta*).
- David Dorn’s do file (Autor and Dorn, 2013) to create a balanced panel of industries for the Census 5% samples 1980-2000, and the 2005 ACS (5% sample) (*subfile\_ind1990dd.do*).
- Measure of emerging tasks by occupation for the years 2008-2018 from ONET, data are classified by occupation in a consistent way between 1980 and the most recent ACS vintage (*task\_changes\_occ1990dd\_acs\_allyrs.dta*).

## 2. temp\_data:

This folder contains the files created in the process of data-cleaning from the programs included in the “cleaners” folder.

## 3. clean\_data:

- Measure on adjusted penetration of robots in 19 industries as developed by Acemoglu and Restrepo (2018b): *apr\_measure.dta*, the underlying data can be purchased contacting the *IFR* ;
- Data on Chinese import and offshoring by BEA industry codes (*china-sag\_bea.dta*) and by sic87dd (David Dorn’s consistent classification for SIC87, *china-sag\_sic87dd.dta*), the underlying data are available on David Dorn’s *website* ;
- Data on value added, labor share, compensation, employment, population, price index for 1850 - 1910 (authors’ elaboration on Budd (1960) ’s estimates): *panel\_budd.dta*
- *beaio\_smt.dta* is the output of the cleaning of the data from the Survey of Manufacturing Technology (corresponding do file in the cleaners folder: *clean\_smt.do*)
- *census\_characteristics\_BEA.dta*, *census\_education1990\_BEA.dta*, *census\_education2014\_BEA.dta*, *censusoccdiversity\_BEA.dta* are the outputs of the cleaning of the census data

(corresponding do file in the cleaners folder: *clean\_census\_industry.xs.do*)

- *FH\_offshoring\_bea.dta*, *FH\_offshoring\_sic87dd.dta* are the outputs of the cleaning of the Feenstra and Hansons (1999) data , extended by Wright (2014).  
The corresponding do file in the cleaner folder is: *clean\_offshoring.do*
- The panel of BEA industries and their respective measures of quantities and prices of inputs and aggregates for the years 1948-1987 (*panel\_bea72.dta*) and and for the years 1987 - 2017 (*panel\_beaNAICS.dta*).  
The respective do files in the cleaners folder are: *clean\_bea\_broad\_1947-1987.do* and *clean\_bea\_broad.do*
- The folder includes also the file *industry\_contribution\_sigma\_0.8.dta* which records each industry's contribution to the task change of production in terms of reinstatement and displacement effect.  
This file is produced in the *decomposition 1987-2017.do* present in the dofiles folder.
- The datasets: *price\_bea.dta*, *qty\_bea.dta*, *tfp\_bea.dta* contain cleaned BEA data respectively on prices, quantities, TFP they are the outputs of the corresponding cleaning file: *clean\_bea\_productivity.do*.

#### 4. xwalks

- *master\_xwalk.dta* crosswalk to map consistently BLS industry codes, BEA industry codes, the 19 codes of industries for which we have a measure of robot penetration (*industry\_ifr1990*), the Census' industry Codes for 1990 (*ind1990*).
- *consolidate\_nipa.dta* This crosswalk maps NIPA categories to the BEA - SIC 72 industry codes (used in the 1947-1987 analysis) and to the BEA - NAICS 2007 industry codes (used in the 1987 - 2017 analysis).
- *cw\_n97-s87.dta* This crosswalk maps SIC 87 - 4 digit industry codes into NAICS 97 - 6 digit industry codes. The share of employment in each NAICS 1997 industry is used as weight. Data are from David Dorn.
- *naics4\_xwalk.dta*: maps NAICS 97 - 4 digit classification into BEA industry classification.
- *sic72-sic87.dta*: maps SIC 72 - 4 digit industry codes into SIC 82 - 4 digit classification.
- *sic72\_xwalk.dta*: maps SIC 72 industry descriptions into consistent BEA industry descriptions.
- *sic87-sic87dd.dta*: output of the do file *subfile\_sic87dd.do* created by David Dorn that develops a consistent 4-digit SIC classification that can be matched to NAICS industry codes and to the NBER-CES Manufacturing Industry database.

## Programs

The programs used in the papers are included in the folders:

### 1. dofiles

- *executer\_final.do* is the main do file that recalls the programs required to produce the analysis reported in the paper. **Before running the file set the path to the working directory at line 16.**
- *sectoral trends.do* analyzes labor share and sectoral evolutions for three periods: 1850-1910, 1947-1987 and 1987-2017.
- *decomposition 1987-2017.do*, *decomposition 1947-1987.do* implement decomposition of the wage bill for the whole economy respectively for 1987-2017 and for 1947-2017. The files *decomposition 1987-2017*, *manufacturing.do*, *decomposition 1947-1987*, *manufacturing.do* replicate the analysis on the manufacturing sector. *decomposition 1987-2017*, *BLS* replicates the decomposition for the whole economy using BLS data in place of BEA data.
- *analyze\_var\_jep.do*, *correlates automation final.do*, *correlates newtasks final.do*, *correlates prices and quantities.do* analyze the relationship between change in task content of production and: proxies of automation, proxies of new tasks, quantities produced, TFP and skill intensity of industries.

### 2. cleaners:

This folder contains the do files used to clean the raw-data. The outputs are stored in the folders *temp\_data* and *clean\_data*.

Notice that the do file *clean\_bea\_broad.do* must be run last as it requires inputs created from other do files.

## Outputs

The outputs of the analysis carried out in the paper are stored in the two folders:

### 1. tables:

this is a repository for the results of the models developed in the *analyze\_var\_jep.do*, *correlates automation final.do*, *correlates newtasks final.do*, *correlates prices and quantities.do*. It contains the results reported in Table 1 and Table A1

### 2. figs:

this is a repository for the figures produces by the analysis. Notice that more figures are produced than the ones reported on the paper.

## References

**Acemoglu, Daron and Pascual Restrepo (2018a)** “The Race Between Machine and Man: Implications of Technology for Growth, Factor Shares and Employment” *American Economic Review*, 108(6): 1488–1542.

**Acemoglu, Daron and Pascual Restrepo (2018b)** “Robots and Jobs: Evidence from US Labor Markets” NBER WP No. 23285.

**Autor, David, David Dorn, and Gordon Hanson (2013)** “The China Syndrome: Local Labor Market Effects of Import Competition in the United States.” *American Economic Review* 103(6): 2121–68

**Budd, Edward C. (1960)** “Factor Shares, 1850-1910,” in *Trends in the American Economy in the Nineteenth Century*, Princeton University Press.

**Feenstra, Robert, and Gordon Hanson (1999)** “The Impact of Outsourcing and High-Technology Capital on Wages: Estimates for the United States, 1979-1990.” *The Quarterly Journal of Economics*, 114(3): 907–940.

**Karabarbounis, Loukas and Brent Neiman (2018)** “Accounting for Factorless Income,” NBER WP 24404.

**Lin, Jeffrey (2011)** “Technological Adaptation, Cities, and New Work” *Review of Economics and Statistics* 93(2): 554–574.

**Rognlie, Matthew (2015)** “Deciphering the Fall and Rise in the Net Capital Share: Accumulation or Scarcity?,” *Brooking papers on Economic Activity*.

**Wright, Greg (2014)** “Revisiting the Employment Impact of Offshoring,” *European Economic Review* 66:63–83.