

READ ME

Replication Package for „Robots at Work? Pitfalls of Industry Level Data”

Karim Bekhtiar, Benjamin Bittschi & Richard Sellner

April 29th, 2024

Folder Structure:

After extracting the replication package into your local directory **[your-path]** you will find the following folder structure:

- **[your-path\data]**: includes several datasets we used to (1) augment the original Graetz & Michaels (2018)¹ dataset by additional control variables and (2) extend the analysis to the period 2010-2015. Notice that the robotics data, as well as all data sources needed to reconstruct the original G&M data are not contained in this folder! These data inputs can be found under **[your-path\g&m-replication-package\build\input]** (details see below).
- **[your-path\do]**: Includes all do-files written by us to (1) execute the entire replication package by G&M to generate their estimation data files, (2) augment these data files by additional variables, (3) generate a corresponding data set for the period 2010-2015 and (4) carry out our replication analysis. To execute the entire code (including G&M's original code) the user only has to change row 37 of **[your-path\do\000_master-file.do]** to her local directory **[your-path]** and execute the do-file. All do-files provided in this folder are extensively commented, to make each step of the computations as tractable as possible.
- **[your-path\g&m-replication-package]**: Contains the replication package from G&M's original paper (also available under [this link](#)). It contains all data sources and do-files needed to replicate the original G&M results. We have kept adjustments to these do-files at an absolute minimum to preserve the structure of the original G&M replication material. The very minor adjustments we had to make (to prevent the deletion of Stata macros) are documented in our do-file **[your-path\do\000_master-file.do]**.

Important Note: All data inputs needed to generate the G&M data set for the period 1993-2007 are included in this folder under **[your-path\g&m-replication-package\build\input]**. This includes the IFR-data!

- **[your-path\R]**: Includes two R-scripts which are used to generate the control variables for industry level trade exposure. These R-scripts need not be executed, as their outputs are provided in **[your-path\data]** and are merged to the remaining data sources in the do-file **[your-path\do\000_master-file.do]**. These scripts are thus mainly included for documentation purposes.
- **[your-path\tex]**: Contains all tables and figures included in our paper. After executing our analysis code via **[your-path\do\000_master-file.do]** all of these are re-estimated. Compile the tex-file in this folder to generate a pdf-file to view all results.

¹ Henceforth referred to as G&M.

IFR-Data:

We use the 2016 Version of the “World Robotics – Industrial Robotics” Data from the Industrial Federation of Robotics (IFR). This data is proprietary and is therefore deleted from this replication package. It can however be purchased online at the official IFR-Homepage under <https://ifr.org/worldrobotics/>. When purchasing the data be sure to select a package which allows online queries and excel-export of the data.

This data comes in two versions: robot deliveries and robot stocks. After this data has been purchased by the user it has to be put in the following folders:

- **[your-path\g&m-replication-package\build\input\IFR\robot deliveries]**
- **[your-path\g&m-replication-package\build\input\IFR\robot stocks]**

Each of these folders must contain industry level xlsx-files that are named in the following way:

s10_12.xlsx, s13_15.xlsx, s16.xlsx, s17_18.xlsx, s19.xlsx, s19_22.xlsx, s20_21.xlsx, s22.xlsx, s229.xlsx, s23.xlsx, s24.xlsx, s24_28.xlsx, s25.xlsx, s260.xlsx, s261.xlsx, s262.xlsx, s263.xlsx, s265.xlsx, s26_27.xlsx, s271.xlsx, s275.xlsx, s279.xlsx, s28.xlsx, s289.xlsx, s29.xlsx, s291.xlsx, s293.xlsx, s2931.xlsx, s2932.xlsx, s2933.xlsx, s2934.xlsx, s2939.xlsx, s299.xlsx, s2999.xlsx, s30.xlsx, s90.xlsx, s91.xlsx, s99.xlsx, sA.xlsx, sB.xlsx, sC.xlsx, sDE.xlsx, sF.xlsx, sP.xlsx

All other files in **[your-path\g&m-replication-package\build\input\IFR]** are either dummy data for robot prices (which are not available to us) or files indicating industry structure. Leave those files unchanged as they are required for the code to run smoothly.

Instructions for Replication:

Step 1: Purchase IFR-data and save it under **[your-path\g&m-replication-package\build\input\IFR]**. Make sure that all xlsx-files are correctly named.

Step 2: Change row 37 in the do-file **[your-path\do\000_master-file.do]** to your local directory.

Step 3: Execute **[your-path\do\000_master-file.do]**. All computations that are called by this do-file are extensively commented in the file.

Step 4: Compile the tex-file in **[your-path\tex]** to generate a pdf of all results.

Generated Estimation Data Files:

The data generating part of **[your-path\do\000_master-file.do]** generates the following final estimation data sets. Since some variables contain missings, not all observations in these data sets are used in the final estimations. The number of usable observations is indicated after the file name:

- **[your-path\data\cross-section_1993-2007_g&m.dta]** (224 Obs. with complete data)
- **[your-path\data\cross-section_2010-2015.dta]** (156 Obs. with complete data)

List of main variables used in analysis:

In the estimation data sets listed above, the following variables are used during the analysis. Other variables, which are not used by us, but where part of G&M's original paper/data sets are kept in the full estimation data sets for completeness (but are not described here, as they are not needed to replicate our results).

Variable Name	Content	Source
country	Country Name	IFR
num_c	Country Name converted to numeric	IFR
code_robots	IFR-Industry Code	IFR
ind_robots	IFR-Industry Name	IFR
code_euklems	EU-KLEMS-Industry Name	EU-KLEMS
ch_rob_pctile	Percentiles of change in robotization (all industries)	IFR
ch_rob_pctile_red	Percentiles of change in robotization (manufacturing)	IFR
ch_rob_pctile_nonmanuf	Percentiles of change in robotization (non-manufacturing)	IFR
ch_rob	Raw change in robots per million hours worked	IFR
ch_lrob	Change in Log(1 + robots per million hours worked)	IFR
hours_replace	Replaceable Hours Instrument	G&M
robots_dot91_phs	Reaching & Handling Instrument	G&M
ch_prod	Change in Log(Labor Productivity)	EU-KLEMS
ch_tfp	Change in Log(Total Factor Productivity)	EU-KLEMS
ch_p	Change in Log(Price Level)	EU-KLEMS
ch_lw	Change in Log(Avg. Hourly Wages)	EU-KLEMS
ch_kl	Capital/Labor Ratio (Change)	EU-KLEMS
kl0	Capital/Labor Ratio (Initial Value)	EU-KLEMS
lw0	Log(Avg. Hourly Wage Rate) (Initial Value)	EU-KLEMS
ch_kitsh	ICT-Capital Share (Change)	EU-KLEMS
kitsh0	ICT-Capital Share (Initial Value)	EU-KLEMS
ch_h	Total Hours Worked (Change)	EU-KLEMS
h0	Hours Worked (Initial Value)	EU-KLEMS
ch_lw_LS	Log(Avg. Hourly Wages) – Low Skill (Change)	EU-KLEMS
ch_h_sh_LS	Share in Hours Worked – Low Skill (Change)	EU-KLEMS
h_sh_LS0	Share in Hours Worked – Low Skill (Initial Value)	EU-KLEMS
ch_lw_MS	Log(Avg. Hourly Wages) – Medium Skill (Change)	EU-KLEMS
ch_h_sh_MS	Share in Hours Worked – Medium Skill (Change)	EU-KLEMS
h_sh_MS0	Share in Hours Worked – Medium Skill (Initial Value)	EU-KLEMS
ch_lw_HS	Log(Avg. Hourly Wages) – High Skill (Change)	EU-KLEMS
ch_h_sh_HS	Share in Hours Worked – High Skill (Change)	EU-KLEMS
h_sh_HS0	Share in Hours Worked – High Skill (Initial Value)	EU-KLEMS
h_sh_490	Share in Hours Worked – Age 30-49 (Initial Value)	EU-KLEMS
h_sh_50PLUS0	Share in Hours Worked – Age 50 and Older (Initial Value)	EU-KLEMS
ch_h_sh_F	Share in Hours Worked – Women (Change)	EU-KLEMS
netimports0	Net-Import Exposure (Initial Value)	Comtrade
share0_rob	Within country hours share of industries (serves as weight)	EU-KLEMS

References:

Graetz, Georg and Guy Michaels, "Robots at Work", The Review of Economics and Statistics, 2018, 100 (5), 753-768.