**README for “What Happens to Workers at Firms that Automate?”**

*James Bessen, Maarten Goos, Anna Salomons, Wiljan van den Berge*

*Last updated: 31 October 2022*

This replication folder contains the following subfolders:

1. **documentation**
2. **import**
3. **do**
   1. **prepare**
   2. **analysis**

The folder “**documentation**” contains the most recent official data dictionaries (in pdf) for the source files needed for this paper (in alphabetical order):

* ABR
* AOTOTJAARBEDRAGTAB
* BAANKENMERKENBUS
* BAANSOMMENTAB
* BETAB
* BIJSTANDJAARBEDRAGTAB
* GBAPERSOONTAB
* HOOGSTEOPLTAB
* ICTBEDRIJVEN
* IHG
* INVESTERINGEN
* POLISBUS
* PS\_BOUWNIJVERHEID
* PS\_COMMERCIELEDIENSTEN
* PS\_DETAILHANDEL
* PS\_GROOTHANDEL
* PS\_INDUSTRIE
* PS\_TRANSPORT
* SECMBUS
* SPOLISBUS
* VTVPERSOONTAB
* WWJAARBEDRAGTAB

We also include examples of questionnaires from the PS (Productie Statistieken, our main data source for automation costs), Investeringen and ICTBedrijven and some auxiliary documentation on the definition CBS (Statistics Netherlands) uses for a “firm” and any changes therein over time. The documentation files are only available in Dutch.

We cannot provide the source datasets as these are confidential administrative employer-employee data owned by Statistics Netherlands (CBS). At the time of writing, researchers can gain access to these data by first requesting access for their institution, if the institution is not already authorized.[[1]](#footnote-1) Second, researchers write a research proposal with an indication of the data files they will need and submit this to CBS. CBS has templates to use for requesting data access.[[2]](#footnote-2) Currently data access is only allowed for research institutes in countries that have data protections in place that are similar or equal to the European GDPR. All analyses must be run on the CBS remote server, with data export protocols in place to safeguard privacy. Data access fees are payable to Statistics Netherlands. Note that all access procedures are subject to change, and we recommend contacting Statistics Netherlands ([microdata@cbs.nl](mailto:microdata@cbs.nl)) directly to inquire about the most up to the date data access options and fee structure.

The folder “**do**” contains all programs used in the paper. This folder should be placed in “H:/automation/”. All other folders will be created in the first do-file. Programs are organized into two subfolders, “**prepare**”, containing data cleaning and sample creation; and “**analysis**”, containing data analysis. Programs in “prepare” are organized into two master dofiles which should be run in the following order: a00\_MASTER\_import\_raw\_data.do; b00\_MASTER\_prepare\_data.do. Programs in “analysis” are organized into one master dofile, MASTER\_analysis.do. These master dofiles will execute all other dofiles in their respective folders; and also contain a description of every dofile (further descriptions are also available at the top of every separate dofile).

The document **Tab-Fig-code-overview.xlsx** provides an overview of which Table, Figure, Appendix Table, and Appendix Figure is produced by which analysis do-file.

The folder “**import**” contains some auxiliary data files (sector prices, sector crosswalks) that can be imported to the Statistics Netherlands server. These should be placed in “H:/automaton/import” for all dofiles to run. The file **cpi9619.dta** should be placed in H:/.

We use Stata version 16 for all analyses. All analyses were run on the remote server of Statistics Netherlands which at the time of writing uses Windows (recent version). All required Stata packages are already installed on the servers of Statistics Netherlands.

Note that CBS might change the name and/or location of some original data files. Furthermore, if newer versions of data files become available (e.g. adding an additional year), they sometimes remove the old files. This can lead to errors when trying to run the dofiles. At the time of writing, all do-files are confirmed to run.

1. A list of already authorized institutions is available at: <https://www.cbs.nl/en-gb/our-services/customised-services-microdata/microdata-conducting-your-own-research/institutions-and-projects> (accessed on 31 October 2022). [↑](#footnote-ref-1)
2. See <https://www.cbs.nl/en-gb/our-services/customised-services-microdata/microdata-conducting-your-own-research/which-steps/application> for more detail on the required steps and links to the templates (accessed on 31 October 2022). [↑](#footnote-ref-2)