Replication Materials

Kopp: Do recruiters penalize men who prefer low hours? Evidence from online labor market data (accepted in 2024)

Date: January 3th 2025

Data Availability

The programs require proprietary data from the Swiss State Secretariat for Economic Affairs (SECO) that are not provided with the replication files. Access to the data requires an application granted by the Swiss State Secretariat for Economic Affairs (SECO). Details on how to access the different datasets are below.

The following datasets have to be requested:

- Recruiter click data from Job-Room linked to the Swiss unemployment register, March-December 2017, provided by the Swiss State Secretariat for Economic Affairs (SECO) and processed by KOF Swiss Economic Institute (KOF). The data can be accessed by signing a data sharing agreement with the Swiss State Secretariat for Economic Affairs, Labour Directorate, Holzikofenweg 36, 3003 Bern, Switzerland and ETH Zurich, KOF Micro Data Centre, Leonhardsstrasse 21, 8092, Zurich, Switzerland. The name of the dataset is *jobroom_data_full_201703_201712.dta*.
- Job ads published on Job-Room, July 2018-June 2021, provided by the Swiss State Secretariat for Economic Affairs (SECO). The data can be accessed by signing a data sharing agreement with the Swiss State Secretariat for Economic Affairs, Labour Directorate, Holzikofenweg 36, 3003 Bern, Switzerland.
  Researchers have to request the following micro-level datasets, that contain information on job ads directly reported to Job-Room between July 2018 and June 2021. The data delivery contract between KOF and SECO has the contract number 210922. Please mention this number when applying for the data to ensure that your datasets will be identical to the one used in the paper.
  - *20220407_SECO2KOF_Tab1_OSTE.csv*
  - *20220303_SECO2KOF_Tab4_OSTE.csv*
  - *20220329_SECO2KOF_Tab1_OSTE_Beruf.csv*

  In addition, researchers have to request lists of all job ads published on Job-Room between January 1st 2020 and May 31 2021. Based on these lists, researchers can access the content of the ads via the Job-Room API (with the script *jobroom_api_info.R*).
  - kof-ids_20200101-20200630.csv
  - kof-ids_20200701-20201231.csv
  - kof-ids_20210101-20210531.csv
- Click data of registered jobseekers on job ads on Job-Room between June 2020 and May 2021, that can be linked to the Swiss unemployment register. The data is owned by the Swiss State Secretariat for Economic Affairs (SECO) and can be accessed by signing a data sharing agreement with the Swiss State Secretariat for Economic Affairs, Labour Directorate, Holzikofenweg 36, 3003 Bern, Switzerland. The data delivery contract between KOF and SECO has the contract number 210922. Please mention this number when applying for the data to ensure that your datasets will be identical.

- 20220329SECO2KOF_Tab2_STES.csv
- OSTE_JR.Rdata
- STES_JR_200522.Rdata

The subfolder "Misc_files" contains the empty subfolder "jobroom_ads" the following auxiliary files that are provided with the replication package:

- *key_gemeinde_kanton.dta* is a crosswalk between the municipality number according to the Federal Statistical Office and the canton.
- *plz_kanton.dta* is a crosswalk between the zip code of a municipality and the canton.
- *lse_isco_women.dta* shows female shares in total employment by ISCO 2-digit occupation in 2016 according to the Swiss Earnings Structure Survey 2016. The data is provided by the Swiss Federal Statistical Office (FSO).[1]
- *employment_by_isco4.dta* shows employment by ISCO 4-digit occupation according to the structural survey of the Swiss Federal Statistical Office (FSO).[2] To compute these numbers, we average employment counts in the waves 2016 and 2017 of the survey.
- *abstimmung_vaterschaft.dta* shows the percentage of no votes in the referendum on the extension of paternity leave for fathers that took place on 27 September 2020.
- *branchen.xlsx* is a crosswalk between NOGA (NACE) letters and 2-digit codes.
- *AVAM_ISCO.xlsx* is a crosswalk between the occupational classification of the Swiss Public Employment Service and the ISCO classification.
- *oecd_hours_worked_paid_unpaid_gender.csv* provides information on how women and men spend their time. The data distinguish between paid and unpaid work and refer to the year 2019 or the latest year available up to 2019. Data source is the OECD Time Use Database, except for Switzerland, where the data come from the Swiss Federal Statistical Office, as the OECD does not report such data for Switzerland.
- *working_time_v_flex_isco2.dta* shows the share of workers who report being flexible in determining the start and end of their work day and who report that they have no block hours by ISCO-2 occupations. The data come from the 2015-2019 waves of the Swiss Labor Force Survey.[3]
- *take_time_off_by_isco2.dta* shows the share of workers who report that they can easily take one or two hours off at short notice by ISCO-2 occupations. The data come from the 2015-2019 waves of the Swiss Labor Force Survey.
- *overtime_extrawork_by_isco2.dta* shows the share of workers who worked overtime the week before by ISCO-2 occupations. The data come from the 2015-2019 waves of the Swiss Labor Force Survey.
- *lse_isco_overtime.dta* shows the share of overtime pay in gross monthly pay by ISCO-2 occupations. The data come from the Swiss wage structure survey 2016.
- *schichtarbeit_s_by_isco2.dta* shows the share of workers reporting shift work by ISCO-2 occupations. The data come from the 2015-2019 waves of the Swiss Labor Force Survey.
- *time_pressure_by_isco2.dta* shows the share of workers who report that they often or very often experience time pressure at work by ISCO-2 occupations. The data come from the 2015-2019 waves of the Swiss Labor Force Survey.
- *work_autonomy_by_isco2.dta* shows the share of workers who report that they can influence the order and the content of taks by ISCO-2 occupations. The data come from the 2015-2019 waves of the Swiss Labor Force Survey.

[1] To get access to the raw data, researchers may use the contact form provided on the following FSO website: https://www.bfs.admin.ch/bfs/en/home/statistics/work-income/surveys/ess.html.
[2] To get access to the raw data, researchers may use the contact form provided on the following FSO website: https://www.bfs.admin.ch/bfs/en/home/statistics/population/surveys/se.html.
[3] To get access to the raw data, researchers may use the contact form provided on the following FSO website: https://www.bfs.admin.ch/bfs/en/home/statistics/work-income/surveys/slfs.html

- *cantons.csv* is a crosswalk between the canton key used by the Swiss Federal Statistical Office and the key used by the State Secretariat for Economic Affairs (SECO).

The subfolder "Help_files" contains the following scripts that are provided with the replication package:

- *label_isco08_1.do*
- *label_isco08_2.do*
- *label_isco08_2_99_no_inform.do*
- *label_isco08_3.do*
- *label_isco08_4.do*
- *sample_restrictions_parttime.do*
- *extract_gender_jobtitle.do*
- *labelfromstring.do*
- *progr_appendmodels.do*

Software requirements

The replication requires the following user-written Stata programs. To install them, run the corresponding code in 00_masterfile once:

- Coefplot
- estout
- reghdfe
- distinct
- ftools
- egenmore
- addplot
- appendfile
- gtools

The replication requires the following R-packages:

- data.table
- Matrix
- Foreach
- Glmnet
- glmnetUtils
- readr
- readstata13
- foreign
- tidyverse
- MASS
- Biglasso
- Bigmemory
- bigmemory.sri
- ncvreg
- lfe
- fastDummies
- stringr
- lubridate
- httr
- jsonlite

- readxl
- tictoc

Version and runtime information:

The code was last run with Stata MP Version 17.0 and R version 4.2.3.

The analyses were last run on a 64-core 2.25GHZ Windows server with 512 GB RAM. The replicator should expect the code to run for about 7 days.

How to replicate results

The provided code reproduces all tables and figures in the paper.

General note: All Stata scripts are called in the correct order in the master-do file. However, the following three R scripts must be run before the Stata scripts:

- 1_1_jobroom_api_info.R
- 1_2_persnr_date_to_stes_id.R
- 1_3_import_ad_clicks.R.

Also, after Stata script 5_1_preparation_lasso_selection_parttime.do, the master-do file stops and one has to manually run R script 5_2_lasso_selection_parttime.R, before calling the remaining Stata scripts from the master-do file. I indicate in the master-do file when the R scripts need to be run.

STEP 1: Preparation

- Download and unpack the zipfile into a preferred location
- The main folder (replication_files) should contain the following elements
    - Folders
        - Data_raw (empty)
        - Data_processed (empty)
        - Help_files (9 files)
        - Misc_files (16 files and one empty folder)
        - Save_estimates (empty)
        - Utils (empty)
        - Results_and_paper (empty)
    - Files
        - 19 Stata Do-Files
        - 4 R-Scripts
        - README.pdf
- Acquire the main datasets mentioned under “Data Availability” (10 datasets) and store it in the “data_raw” folder.
- Define the replication folder “replication_files” as the project folder in the Master do file on line 15.

STEP 2: Preparation of analysis datasets

- Scripts 1_* prepare the dataset with all ads published on Job-Room (job_ads_from_api) and the dataset with all jobseeker clicks (ad_clicks). The scripts must be executed one after the other in R.

- Scripts 2_* prepare the datasets of registered jobseekers (stes) and ads directly reported to Job-Room (oste). The scripts can be called from the master-do file.

Step 3: Analysis of the click and job ad data

- Scripts 4_* analyse the the jobseeker click data and the job ad data. They produce the following tables and figures that appear in the paper:
    - Table 1
    - Figure 4
    - Table D.4
    - Table D.7
    - Table D.8
    - Table D.9
    - Table D.10
    - Table D.11
    - Table D.12
    - Figure D.11

Step 4: Post-double selection with Lasso

- Scripts 5_* perform the post-double selection with Lasso. The first two scripts (5_0_Dataset_prep_gen_interact and 5_1_paraparation_lasso_selection_parttime) can be called from the master-do file. The last script 5_2_lasso_selection_parttime.R must be run in R.

Step 5: Descriptive analyses of the recruiter click data and OECD time use data

- Scripts 6_* perform descriptive analyses of the recruiter click data. They produce the following tables and figures that appear in the paper:
    - Table 2
    - Table D.2
    - Table D.3
    - Table D.5
    - Figure D.2
    - Figure D.3
- Script "7_1_Gen_share_unpaid_oecd" produces figure D.1.

Step 6: Regression analyses of the recruiter click data

- Script 8_1 performs the regression analysis on the recruiter click data.

Step 7: Generate figures and tables from regression analysis based on recruiter click data

- Scripts 9_* generate figures and tables from the regression analysis of the recruiter click data. They produce the following tables and figures that appear in the paper:
    - Figure 2
    - Figure 3

- Figure 5
- Figure B.1
- Figure D.4
- Figure D.5
- Figure D.6
- Figure D.7
- Figure D.8
- Figure D.9
- Figure D.11
- Table C.2
- Table D.6