# README and Guidance

## Overview

The code in this replication package constructs the analyses files from a number of data sources using Stata. The analyses are partly carried out using proprietary data made available for research, under specific and strict regulations, by Statistics Norway. Other parts of the analyses are based on survey data collected by the researchers and will be freely available to others. Below we list the different data sources.

The replicator should expect the code to run for less than 30 minutes on a laptop computer. The authors used Stata 17 and a laptop computer with Windows 10.

## Data Availability and Provenance Statements

### Statement about Rights

- X I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

### Summary of Availability

- X Some data **cannot be made** publicly available.

As mentioned in our initial cover letter, we will be able to upload data to replicate almost all of the analyses. Unfortunately, however, we cannot upload data to replicate the analyses of the selection into the experiment as these are based on Norwegian (encrypted) administrative data that have been leased from Statistics Norway. It is unlikely that permission will be given to access the identifiers used to match the register data to the survey data as this was based on informed consent by the participants to be matched with our specific data. All code is, however made available and replicators at the journal can be given access for a limited time to the servers at the Frisch center to run all analyses.

## Dataset list

The data files used in the analyses are the following:

| Filename | Description |
| --- | --- |
| `survey_data.dta:` | Dataset including the survey responses. The dataset also includes data from ILO (2014), World Development Indicators Online Database (2020), World Economic Forum (2016), World Value Survey (2015) and Statistics Faroe Islands (2005). |
| `studies_coded.dta:` | Dataset including results from other papers. From Table 3 in Dariel et al. (2017). |
| `replication_competitiveness.dta:` | Coded dataset from our survey (also generated as part of the package) |
| `survey_data_with_registerdata:` | Coded dataset including register data. Not shared to the wider public but replicators can gain access at the internal servers of the Frisch Centre. The register data are from Statistics Norway (2020). |

All datasets are available as .dta, .csv, and .txt.

Data references
Dariel, A., Kephart, C., Nikiforakis, N. and Zenker, C. (2017). `Emirati women do not shy away from competition: Evidence from a patriarchal society in transition', Journal of the Economic Science Association, vol. 3(2), pp. 121–136.

ILO (2014). `ILOSTAT Database', Geneva.

Statistics Faroe Islands (2005). `Population registers', Statistics Faroe Islands, https://statbank.hagstova.fo/pxweb/en/H2/H2 AM/ (last accessed February 2020).

Statistics Norway (2020). `Population registers', Statistics Norway, https://www.ssb.no/en/data-tilforskning (last accessed February 2020).

World Development Indicators Online Database (2020). `Female labour force participation', The World Bank, https://databank.worldbank.org/data/home (last accessed February 2020).

World Economic Forum (2016). `The global gender gap report 2016.', World Economic Forum, https://www3.weforum.org/docs/GGGR16/WEF Global Gender Gap Report 2016.pdf.

World Value Survey (2015). `World value survey version v2015.04.18'.

## Computational requirements

### Software Requirements

- Stata (code was last run with version 17)

Replication requires the following libraries to be installed and use.

- Binscatter
- Coefplot

### Memory and Runtime Requirements

#### *Summary*

Approximate time needed to reproduce the analyses on a standard (2022) desktop machine:

- ☐<10 minutes
- X 10-60 minutes
- ☐1-8 hours
- ☐8-24 hours
- ☐1-3 days
- ☐3-14 days
- ☐> 14 days
- ☐Not feasible to run on a desktop machine, as described below.

#### *Details*

Estimations were partly done on a windows 10 virtual machine, with 4 virtual cores at 2.10Ghz, and 16GB RAM. All estimations can, however, be run on a laptop computer with windows 10.

## Description of programs/code

### ANALYSIS

All tables and graphs are constructed using a set of STATA programs.
The results are reproduced by running the master dofile "S00_master_dofile_surveydata.do" for the surveydata analysis and the dofile "R01_register_data_analysis.do" for the data including the administrative data. The dofile "R01_register_data_analysis.do" can only be run after temporary access to the Frisch Centre server has been granted.

These programs produce the different graphs and table inputs and puts everything in the folders "Results".

**Here is an overview of tables and figures, <mark>results in yellow are based on restricted</mark>**

**data where temporary access will be granted.**

<u>Main paper</u>
Table 1 Table_1, Just text no replication code.
Table 2: Consists of two tables:
Table_2a (unrestricted)
Table_2b (restricted)
Table 3 Table_3
Table 4 Table_4
Consists of three tables:
Table_4a (unrestricted) , creates column 1
Table_4b (unrestricted), creates most of column 2 but does not include voting
Table_4c (restricted), creates voting, i.e. the rest of column 2.
Table 5 Table_5 (restricted)
Table 6 Table_6
Table 7 Table_7
Table 8 Table_8
Table 9 Table_9
Table 10 Table_10 (Restricted)
Table 11 Consists of two tables:
Table_11a (Restricted)
Table_11b (Restricted)
Figure 1 Figure_1
Figure 2 Figure_2
Figure 3 Figure_3
Figure 4 Figure_4

<u>Appendix</u>
Table A1, created by hand from the log file: Appendix_A1.log
Table A2 Table_A2
Table A3 Table_A3
Table A4 Table_A4
Table A5 Table_A5
Table A6 Table_A6
Table A7 Table_A7
Table A8 Table_A8
Table A9 Consists of three tables:
 Table_A9a (Restricted)
Table_A9b (Restricted)
Table_A9c (Restricted)
Table A10 Table_A10
Table A11 Table_A11
Table A12 Table_A12
Table A13 Table_A13
Table A14 Consists of two tables:
Table_14a (Restricted)
Table_14b (Restricted)
Table A15 Table_A15  (Restricted)
Table A16 Table_A16  (Restricted)
Table A17 Table_A17
Table A18 Table_A18
Table A19 Table_A19
Table A20 Table_A20
Table A21 Table_A21
Table A22 Table_A22
Table A23 Table_A23

Figure A1, Just a picture no replication code
Figure A2 Figure_A2
Figure A3 Figure_A3
Figure A4 Figure_A4
Figure A5 Figure_A5
Figure A6, consists of two figures: Figure_A6a and Figure_A6b
Figure A7, consists of two figures: Figure_A7a and Figure_A7b

## Instructions

The replication package contains the following folders:

    Results

    Data

    Dofiles

**Step 1:** Copy the filepath to where the folders are stored.

**Step 2**: Open the dofiles and change the globals for the results and data folders so that it matches your folder paths. In particular change the globals on the following lines:

S00: Lines 3-5
S01: Lines 10 and 11
R01: Lines 6 and 7
S02: Lines 10 and 11

**Step 3:** Run the master dofile "S00_master_dofile_surveydata.do" and the dofile "R01_register_data_analysis.do". When everything is finished, you will find the graphs and inputs for tables in the folder "Results".