README File for

„Caught between Cultures: Unintended Consequences of Improving Opportunity for Immigrant Girls"

By Gordon Dahl, Christina Felfe, Paul Frijters and Helmut Rainer

This readme file contains instructions on how to replicate the results of the paper "Caught between Cultures: Unintended Consequences of Improving Opportunity for Immigrant Girls". It contains a data availability statement and replication instructions.

Data Availability

The authors of the manuscript conducted their own data collection and have legitimate access to and permission to use the data used in this manuscript. The original questionnaire is in German and is provided as part of this replication package. The authors further draw upon confidential data from the German Microcensus, and this readme file explains how replicators can gain access to this data.

Replication Instructions

The Replication folder is organized according to the following structure:

- The subdirectory Data\ contains all the data. The Data/Raw subdirectory contains the raw data files and the Data\Clean is an empty subdirectory to save the final datasets used for the analysis.
- The subdirectory Codes\ includes all codes necessary to replicate the results. The subdirectory Codes\Ado\ contains all ado files needed to make the codes run.
- The subdirectory Output\ is a subdirectory in which all figures (Output\Figures\) and tables (Output\Tables\) are saved.
- The subdirectory Survey\ contains our original survey instrument.

Before running the codes it is necessary to update the global path to the Replication folder.

In order to generate the results of the paper, replicators need first to run the "Prepare_Data.do" file, available in the subdirectory Codes. This file draws upon the raw data "data_analysis.dta" contained in the subdirectory Data/Raw/. Note that the raw data already contain basic regional data (local unemployment rates) and gender norms in the mother's country of birth that was merged in using data from the Federal Employment Agency (2015) and the World Economic Forum (2006), respectively. The raw data is restricted to all data needed for replication. Specifically, it provides data on all native and immigrant children born between July 1, 1999 and June 30, 2000 that have successfully participated in the survey. It also provides all variables needed for replication.

The "Generate_Results.do" file, available in the subdirectory Codes, contains all codes to generate Tables 1-10 and Appendix Tables A1-A7 as well as Figure 2.

The "Graph_Citizenship.do" file, available in the subdirectory Codes, contains the code to generate Figure 1. The figure has two panels. Panel A is produced using confidential data from

the German Microcensus (Wave 2001). Specifically, it uses the Microcensus Scientific Use File 2001 (Federal Statistical Office and Statistical Offices of the Länder, 2001). A description of this data file and information on how to gain access to this data is available at https://doi.org/10.21242/12211.2001.00.00.3.1.0. Once replicators have gained access to this data file, they have to run the "Microcensus_Citizenship.do" file to prepare the data for this subfigure. For the convenience of researchers, we have included the numbers used to create Panel A in the file "citizenship_microcensus_collapsed.dta", saved in the subdirectory Data/Clean, so that replicators can see how we created the graph. To produce Panel B, replicators have to first run the data preparation code "Survey_Citizenship.do" available in the subdirectory Codes. This file draws upon the data file "data_citizenship_survey.dta" from the subdirectory Data/Raw. This dataset is restricted to all variables needed for the replication of Panel B.

Finally, the "Master.do" file, available in the subdirectory Codes, generates all results in this paper by calling the aforementioned do-files in correct order.

Software and computational requirements

The software and computational requirements for running the relevant analysis code, as well as the relevant packages, that must be installed, are as follows:

Stata version 15.1 having installed the following ado files (which can be found in the subdirectory codes/ado):

- estout (as of 2021-09-14)
- estpost (as of 2021-09-14)
- eststo (as of 2021-09-14)
- esttab (as of 2021-09-14)
- grc1leg (as of 2021-09-14)
- outreg2 (as of 2021-09-14)

The code was last run on an Intel ® Core ™ i7-8550U.CPU with 8 GB of RAM on Windows 10.

Data citations

Federal Statistical Office and Statistical Offices of the Länder, 2001. Microzensus 2001, Scientific Use File (SUF). Access information available at: https://doi.org/10.21242/12211.2001.00.00.3.1.0.

Federal Employment Agency, 2015. Arbeitsmarktstatistik der Bundesagentur für Arbeit 2015. Available at: www.regionalstatistik.de.

World Economic Forum, 2006. The Global Gender Gap Report 2006. Available at: https://www.weforum.org/reports/global-gender-gap-report-2006.