Replication Notes for Foreign Direct Investment and Repression: An Analysis Across Industry Sectors

Nicole Janz *University of Nottingham*
School of Politics and International Relations
University Park
Nottingham, NG7 2RD
Email: nicole.janz@nottingham.ac.uk

September 14, 2016

List of Replication Materials

1. Dataset (replicationdata.Rdata): The data set contains all variables employed in the main models and the robustness checks.

2. Rscript (JHR_Replication.R): The R code contains all commands to load the data, transform the variables as necessary, and run the models to produce the results.

3. Codebook (Variable_Codebook.xlsx): The codebook contains the variables, their description, sources and exact variable names as in the data set.

Instructions for Replication

- First, load the data. This is a panel data set with country-years. Each table or figure is identified by its header as in the article or online appendix. If you are looking for a specific table, search the (admittedly long) Rscript for that particular heading.

- Run the code chunk by chunk. Many variables are created 'on the go' and re-used at a later point. Make sure to run the code in the same order, and run all the code even if you only want to replicate e.g. the last table.

- In order to preserve the largest possible sample size, I have created data sets for each model separately (based on the original table that you have loaded in R). This way, depending on the particular model, slightly different country-years are included. Make sure to create all these data sets (again, in order of the code) to run the models.

- To create nice tables, I used the R package 'stargazer' to create latex files. Each output object is entered into the stargazer function, and all results are displayed in the R output as well as saved as .tex file into your working directory.

- For the ordered logit models, I use the functions lrm() and robcov() from the R package "rms" version 4.3-0, which produces the same (or very similar) results as the corresponding STATA command ologit with the cluster() modification (STATA version 13.0). For OLS with PCSE, I use the functions plm() and vcovBK() from

the R package "plm" version 1.3.1, which produces the same (or very similar) results as the STATA command xtpcse with the pairwise specification.

- Sometimes, when replicating code and models that were conducted using previous versions of R packages, the results slightly change; sometimes models may not converge because the R package was updated. In order to get to the exact same results as in the study, I recommend installing the exact same R version and package versions. Calling the command line sessionInfo() I have displayed all versions and packages necessary at the beginning of the Rscript.

- All these steps refer to a replication, i.e. duplication of the exact same models with exactly the same data as provided in the replication files. You may find that re-collecting these data from the original sources yields slightly different results, probably due to more recent data revisions, the inclusion of more or different years and countries.

Citation

If you work with the data for your own study (replication or original work based on these data), please cite my article as well as the data set. It would be great if you could let me know about your results. A suitable citation of the data is provided by Dataverse where you downloaded the data.

Contact

Please contact me if you have any questions about the study or replication files at: nicole.janz@nottingham.ac.uk or nicolejanz@gmail.com. Any feedback on your replication attempts is more than welcome.

Acknowledgements

Gu Li and Sergio Cuesta have verified that the provided Rscript produces the tables and figures for the study. All errors remain my own.