Overview

The code in this part of the replication package constructs the analysis files from confidential U.S. Bureau of Economic Analysis (BEA) data that are processed using 9 sequential programs in SAS and Stata. The programs and the order in which they should be run are described here. The replicator should expect the code to take about 90 minutes to process.

Data Availability and Provenance

The confidential BEA data used in the paper were collected on the Annual Survey of U.S. Direct Investment Abroad (BE-10 or BE-11) and the Annual Survey of Foreign Direct Investment in the United States (BE-15). The microdata research files and the SAS and Stata programs are housed at BEA on secure servers.

*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.

☐ I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package. Appropriate permission are documented in the [LICENSE.txt](https://social-science-data-editors.github.io/template_README/LICENSE.txt) file.

*Summary of Availability*

☐ All data **are** publicly available.

☐ Some data **cannot be made** publicly available.

X **No data can be made** publicly available.

*Details on each Data Source*

The formulary adjustments in the paper use the confidential microdata from BEA. To gain access to the BEA microdata, follow the directions here on how to write a proposal for access to the data via a Federal Statistical Research Data Center: <https://www.bea.gov/research/special-sworn-researcher-program>. You must request the following data sets in your proposal: BE-10, BE-11, and BE-15 microdata research files for 1982-2016.

Instructions to Replicators

The following SAS and Stata programs should be run in sequence. These programs were last run on November 4, 2021.

Run first: “(1) USDIA Financial and Operating Data 1982-1992 AER Replication.sas” – This SAS program reformats BEA research files to generate consistent variable names and field characteristics. The resulting dataset is a panel of financial and operating data on U.S. reporters and their foreign affiliates for 1982-1992: usdiafo82to92.sas7bdat. The output dataset is used in “(3) RD Stock 1982-2016 AER Replication.sas” and “(5) Working Data Sets 1982-2016 AER Replication.sas”.

Run second: “(2) USDIA Financial and Operating Data 1993-2016 AER Replication.sas” – This SAS program reformats BEA research data files to generate consistent variable names and field characteristics. The resulting dataset is a panel of financial and operating data on U.S. reporters and their foreign affiliates for 1993-2016: usdiafo93to16.sas7bdat. The output dataset is used in “(3) RD Stock 1982-2016 AER Replication.sas” and “(5) Working Data Sets 1982-2016 AER Replication.sas”.

Run third: “(3) RD Stock 1982-2016 AER Replication.sas” – This SAS program generates R&D stocks built from R&D expenditures for U.S. parents and foreign affiliates. The resulting dataset is a panel of R&D stocks for 1982-2016: rdstock.sas7bdat. The output dataset is used in “(5) Working Data Sets 1982-2016 AER Replication.sas”.

Run fourth: “(4) Reporter PPE Interpolation.do” – This Stata program generates a dataset with net PPE interpolated between benchmark years for U.S. reporters: interpolate\_rep\_netppe.dta. The output dataset must be converted to SAS, which is used in “(5) Working Data Sets 1982-2016 AER Replication.sas”.

Run fifth: “(5) Working Data Sets 1982-2016 AER Replication.sas” – This program generates two datasets. The first dataset includes financial and operating data on U.S. reporters and foreign affiliates plus a direct investment earnings proxy for U.S. reporters and foreign affiliates for 1982-2016: usdiafoir82to16.sas7bdat. The output dataset must be converted to Stata, which is used in “(8) Standard Errors Panels AER Replication.do”. The second dataset includes an adjustment to the earnings proxy for U.S. reporters and foreign affiliates from the first dataset based on formulary apportionment for 1982-2016: usdiaapportion82to16.sas7bdat. The output dataset is used in “(6) Formulary Apportionment 1982-2016 AER Replication.sas” and “(7) Hines and Rice 1982 AER Replication.sas”.

Run sixth: “(6) Formulary Apportionment 1982-2016 AER Replication.sas” – This program generates aggregates of formulary adjustments to direct investment earnings proxies between U.S. reporters and foreign affiliates for 1982-2016.

Run seventh: “(7) Hines and Rice 1982 AER Replication.sas” – This program includes tax rates from Hines and Rice (1994) and an adjusted earnings proxy calculated using the formulary adjustment to generate elasticities for 1982: elasticities.sas7bdat.

Run eighth: “(8) Standard Errors Panels AER Replication.sas” – This program randomly generates 100 CSV panels with an equal number of observations to our original panel. The CSV panels are used in “(9) Standard Errors AER Replication.sas”.

Run ninth: “(9) Standard Errors AER Replication.sas” – This program calculates formulary adjustments and standard errors for randomly generated panels with an equal number of observations to our original panel.