Bilateral International Investments: The Big Sur?
by Fernando Broner, Tatiana Didier, Sergio L. Schmukler, and Goetz von Peter
Journal of International Economics

Readme File for the Replication Package

This readme file describes how to replicate the tables and figures in the paper, starting from the raw data downloaded from the original sources. It lists the databases used in the paper and explains how the data were cleaned. It also describes the Stata code (do-files) used to generate the tables and figures.

1. Data

1.1. Investment databases

The databases used in the paper are publicly available from their original sources, except for bank loans and deposits.

Bank loans and deposits: For bank loans and deposits, we use data from the Locational Banking Statistics (LBS) from the Bank for International Settlements (BIS), which compiles assets and liabilities of internationally active banks on a residency basis, in line with balance of payments statistics.

The loans and deposits series used in our analysis are based on the full LBS data, which cannot be shared publicly due to data confidentiality restrictions. The replication package we share includes aggregated series to preserve confidentiality, as described in Section 4 below. The series in the replication package are based on the full LBS data and reflect various adjustments we made (such as the use of mirror data, reconciliation of interbank positions, and reallocation of regional residuals). The BIS makes a more aggregated version of the LBS data publicly available, which can be found here (or here in a single csv).

Portfolio investment: For portfolio investment, we use data from the IMF's Coordinated Portfolio Investment Survey (CPIS), which contains bilateral data on holdings of portfolio investment securities.

The data are available here, within the IMF data repository. We select the following options when downloading the data: Format: Panel. Selected variables: Assets total investments BPM6 US dollars, Liabilities total investments BPM6 US dollars, Assets equity investments BPM6 US dollars, Assets debt securities investments BPM6 US dollars. All countries. All counterparty countries. Sector: Total Holdings. Counterparty Sector: Total Holdings.

The raw data used in the paper are in the file *CPIS_raw.csv* in the folder …*1. Raw Databases\2.CPIS*.

Foreign direct investment (FDI): For Foreign Direct Investment, we use data from the IMF's Coordinate Direct Investment Survey (CDIS) databases and the UNCTAD Bilateral FDI Statistics: the CDIS data cover 2009-2018, while the UNCTAD data span 2001-2012. Section 2

below explains how to merge these databases and adjust the discrepancies in baseline values between the two databases to replicate the analysis in the paper.

The CDIS data are available here, within the IMF data repository. We select the following options when downloading the data: Format: Panel. Selected variables: Inward Direct Investments Positions, Outward Direct Investments Positions. All countries. All Counterparty Countries.

The UNCTAD data are available here. The raw data are available separately for each country, as a source to any other country (outstocks) and as a destination from any other country (instocks). Target countries are identified by name and not by ISO3 code. The UNCTAD table available here maps the country names to ISO3 codes.

The CDIS raw data used in the paper are in the file *CDIS_raw.csv* in the folder …*1. Raw Databases\\3.FDI\\FDI_IMF*. The UNCTAD raw data used in the paper are in the folders …*1. Raw Databases\\3.FDI\\FDI_UNCTAD\\Instocks\\Excel* and …*1. Raw Databases\\3.FDI\\FDI_UNCTAD\\Outstocks\\Excel.*

International reserves (IR): For International Reserves, we use data from the IMF's International Financial Statistics (IFS) on each country's international reserves.

The data are available here, within the IMF data repository. We select the following options when downloading the data: Format: Panel. Variable: International Liquidity, Total Reserves excluding Gold, US Dollars (RAXG_USD). All countries.

Public data on international reserves lack a counterparty breakdown. Thus, we construct a proxy for country-to-region investments using data from the IMF's Currency Composition of Official Foreign Exchange Reserves (COFER) (see Appendix A.4). The data are available here, within the IMF data repository.

The IFS raw data used in the paper are in the file *IR_raw.csv* in the folder ... *1. Raw Databases\\4.IR*.The COFER data used in the paper are in the file *CurrencyCompositionReserves.xls* in the folder ...*1. Raw Databases\\4.IR* which is used in the do-file *06_01_Clean IR.do.* Section 2 below explains what the do-file does.

1.2. Complementary databases

GDP: For GDP, we use data from the World Bank's World Development Indicators and the International Monetary Fund's World Economic Outlook, in current USD. We use IMF data only when World Bank data are not available.

The raw data used in the paper are in the files *GDP_IMF.xls* and *GDP_WB.xls* in the folder ...*1. Raw Databases\\1.GDP*.

LMF: For LMF, we use the Lane and Milesi-Ferretti's global international investment positions aggregates (summing across all countries). The data are available here.

The raw data used in the paper are in the file *LMF data_1970_2020.xlsx* in the folder ...*1.Raw Databases\\5.LMF data*.

Price deflator: For price deflator, we use the Consumer Price Index (CPI) from the Federal Reserve Bank of Minneapolis. The data are available here.

The raw data used in the paper are in the file *Price deflator_raw.csv* in the folder...*1. Raw Databases\\6.Price deflator*.

2. Data cleaning

The do-file *1.Data Cleaning.do* is a master do-file that calls different do-files from …*3.Dofiles\\1.Data Cleaning* to generate the data files for the analysis. These do-files are described below.

Part 1 – GDP
1. *01_Construct GDP Data.do* creates the GDP database used in the four databases (LBS, CPIS, FDI, and IR). Cleaned data are in millions of current USD. This do-file generates the *GPD.dta* database.

Part 2 – CPIS
1. *02_01_Construct CPIS Data_2018.do* imports and cleans the raw IMF data and merges them with GDP. This do-file generates the equity and debt databases *_CPIS_equity.dta* and *_CPIS_debt.dta*.
2. *02_02_Problem with zeros in 2004 and 2005.do* corrects a problem of reporting zeros and missing values between 2004 and 2005. As explained in the paper, we assume that if a country reports missing values for 2003 and 2006, the values for 2004 and 2005 are also missing. Thus, we replace the zero-valued observations in 2004 and 2005 with missing values for the corresponding country pairs. This is corrected for total assets. This do-file generates the portfolio investment database *_CPIS.dta*.

Part 3 – FDI
1. *03_01_Clean FDI_IMF.do* imports and cleans the raw IMF data and merges the FDI data with GDP. We use the maximum between assets (of the reporter) and liabilities (reported by the counterparty) when both are available; if one is missing, we use the non-missing value (see Appendix A.3). This do-file generates an auxiliary database *FDI_IMF_clean.dta*.
2. *03_02_01_Create Database_Outstock.do* imports and cleans Excel files of UNCTAD FDI outstock for each country (sender data). It generates a unique file with the data and merged GDP. This do-file generates an auxiliary database *UNCTAD_Outstock.dta*.
3. *03_02_02_Create Database_Instock.do* imports and cleans Excel files of UNCTAD FDI instock for each country (receiver data). It generates a unique file with the data and merged with GDP. This do-file generates an auxiliary database *UNCTAD_Instocks.dta*.
4. *03_02_03_Merge Outstock and Instock.do* merges the instock and outstock databases into one file. We use the maximum between assets and liabilities when both are available; if one is missing, we use the non-missing value. This do-file generates an auxiliary database *FDI_UNCTAD_clean.dta*.

5. *03_03_Merge UNCTAD and IMF.do* combines the IMF and UNCTAD datasets to obtain FDI data spanning 2001–2018. As explained in Appendix A.3, we adjust the UNCTAD data to match the IMF data. This do-file generates the FDI database *_FDI.dta.*

Part 4 – IR
1. *04_Clean IR.do* imports and cleans the raw IMF data. To construct country-region data of international reserves, we use data on the currency composition of reserves of advanced and emerging regions. These shares are in the Excel file *CurrencyCompositionReserves.xls* and are already included in the do-file. Appendix A.4 provides additional information. The do-file generates the international reserves database *_IR.dta.*

Part 5 – LMF
1. *05_Clean LMF data.do* imports and cleans the LMF data. This do-file generates the LMF database *_LMF_clean.dta.*

Part 6 – Price deflator
1. *06_Clean raw data.do* imports the price data and generates the price deflator database *_PRICE_DEFLATOR.dta.*

Part 7 – GDP by classification
1. *07_01_Regions GDP.do* aggregates the GDP of the North and South and generates the database *regions_gdp.dta.*
2. *07_02_Regions GDP regions.do* aggregates the GDP of the South regions and generates the database *region_gdp_regions.dta.*
3. *07_03_Regions GDP OFC.do* aggregates the GDP of the North and South including OFCs and generates the database *regions_gdp_OFC.dta.*

3. Tables and figures

The do-file *2.Tables and Figures.do* is a master do-file that calls different do-files from …*3.Dofiles\2.Tables and Figures* to generate the output reported in the tables and figures in the paper. These do-files are described below.

To obtain the tables and figures in the exact format displayed in the paper, users would need to copy the output generated by the do-file and paste it into the corresponding sheet (in the red rectangle) in the Excel file *"Bilateral Investments-T&Fs.xlsx".* Figures 3, 5, 6, and 7 and Online Appendix Figure 4 (OAF4) need additional steps, as explained below.

2.Tables and Figures.do runs auxiliary do-files: *North-South classification.do* classifies countries into blocks. *North-South classification_regions.do* separates the South block into regions. *North-South classification_OFC.do* allocates OFCs into the North and South.

Tables:
1. *T1_T2.do* generates the output reported in Table 1: International Investment Positions/World GDP and Table 2: Share of International Investment Positions.
2. *T3.do* generates the output reported in Table 3: Growth in the Value of International Investment Positions.

3. *T4.do* generates the output reported in Table 4: Role of Offshore Financial Centers.
4. *T5.do* generates the output reported in Table 5: International Investment Position/Regional GDPs.
5. *T6.do* generates the output reported in Table 6: Growth in the Value of International Investment Positions over Regional GDPs.
6. *T7_T8.do* generates the output reported in Tables 7: Share of Active Links in the Global Total, and Table 8: Growth of Active Links.
7. *T9.do* generates the output reported in Table 9: International Investment Positions/World GDP: Old versus New Links.
8. *AT1_OAT1.do* generates the output reported in Appendix Table 1: List of Countries and Online Appendix Table 1: List of Countries and Data Availability.
9. *OAT2.do* generates the output reported in Online Appendix Table 2: Portfolio Investment International Positions: Equity vs Debt.
10. *OAT3.do* generates the output reported in Online Appendix Table 3: FDI: Growth in the Value of International Investment Positions: Robustness.
11. *OAT4.do* generates the output reported in Online Appendix Table 4: FDI: Growth of Active Links: Robustness.

Figures:
12. *F1.do* generates the output reported in Figure 1: International Investment Positions: Bilateral vs. Country-Level Data.
13. *F2.do* generates the output reported in Figure 2: Size of International Investments Positions.
14. *F3_F4_F5_OAT5.do* generates the output reported in Figure 3: Growth of International Investment Positions, by South Region, Figure 4: Growth of North and South, by South Region, Figure 5: International Investment Positions in 2018, by South Region, and Online Appendix Table 5: International Investment Positions/World GDP, by South Region.
15. *F6.do* generates the output reported in Figure 6: Growth in the Value of International Investment Positions over Regional GDPs.
16. *F7.do* generates the output reported in Figure 7: The Proliferation of South-South Links.
17. *OAF1.do* generates the output reported in Online Appendix Figure 1: Number of Source and Destination Countries by Year.
18. *OAF2.do* generates the output reported in Online Appendix Figure 2: Average Share of Counterparty Countries, by Source Country per Year and Block.
19. *OAF3_OAF4.do* generates the output reported in Online Appendix Figure 3: Growth of North and South, by South Regions: China and Hong Kong as a Region and Online Appendix Figure 4: International Investment Positions in 2018, by South Regions: China and Hong Kong as a Region.

Figures that require additional steps
- Figure 3. Growth of International Investment Positions, by South Region
    a. To create this figure, generate first the Online Appendix Table 5 (OAT5). Copy and paste its output into the red square of the F3 sheet. The sheet orders the value copied from OAT5 to produce the figure.
- Figure 5 and Online Appendix Figure 4. International Investment Positions in 2018, by South Region

a. To create these figures, copy the international investment positions in 2018 from *F3_F4_F5_OAT5.do* and *OAF3_OAF4.do* and paste it into the red square of the F5 and OAF4 sheets. These Excel sheets change the scales of Figure 5 and Online Appendix Figure 4 to represent the North-to-North values with discontinuous y-axis for better legibility.

- Figure 6. Growth in the Value of International Investment Positions over Regional GDPs
    a. To create this figure, copy the output from *F6.do* and paste it into the red square of the F6 sheet. This Excel sheet orders the output into tables by international investment positions and region-to-region groups and graphs it after normalizing the series to start at zero in the initial year.
- Figure 7. Proliferation of South-South Links
    a. Run the *F7.do* file. This do-file generates tables with the links of South-to-South country-pairs for each year and type of investment.
    b. Run the *F7.R* file to generate each map. This R script shows a connection between a source and a destination country if an active link exists. This script generates the output reported in Figure 7 as PNG files.
    c. To produce Figure 7, import each PNG file into the F7 sheet.

4. Replication for bank and loans deposits

Due to the confidential nature of banking data owned by BIS reporting countries, the BIS cannot release the granular Country-to-Country (C-C) dataset. To preserve confidentiality, the LBS data shared in this package has been aggregated to the Country-to-Region (C-R), Region-to-Country (R-C), and Region-to-Region (R-R) levels. In addition, some observations have been concealed to comply with BIS data sharing rules; missing observations in the dataset are either unreported or concealed positions.

The sharing restrictions prevent replication of the C-C and C-R results of the paper. However, the aggregate R-R series are based on the full dataset, allowing for exact replication.

The R-C, C-R and R-R LBS data are available in the folder ... *2.Cleaned Databases\LBS Public data.*

Do-file *3.Tables and Figures Banking R-R.do* is a master do-file that uses different do-files from … *3.Dofiles\3.Tables and Figures Banking R-R* to generate the R-R output reported in the paper. These do-files are described below.

1. *T1_T2.do* generates the output reported in Table 1 (Panel A): International Investment Positions/World GDP and Table 2 (Panel A): Share of International Investment Position.
2. *T4.do* generates the output reported in the first two columns of Table 4 (Panel A): Role of Offshore Financial Centers.
3. *T5.do* generates the output reported in Table 5 (Panel A): International Investment Position/Regional GDPs.
4. *F2.do* generates the output reported in Figure 2 (Panel A): Size of International Investments Positions.

5. Folder structure

The replication files are available in the following folders:
- ...
- ... *\1. Raw Databases*
- ... *\1. Raw Databases\1.GDP*
- ... *\1. Raw Databases\1.GDP\Reporting*
- ... *\1. Raw Databases\2.CPIS*
- ... *\1. Raw Databases\3.FDI*
- ... *\1. Raw Databases\3.FDI\FDI_IMF*
- ... *\1. Raw Databases\3.FDI\FDI_UNCTAD\Instocks\DTA*
- ... *\1. Raw Databases\3.FDI\FDI_UNCTAD\Instocks\Excel*
- ... *\1. Raw Databases\3.FDI\FDI_UNCTAD\ Outstocks\DTA*
- ... *\1. Raw Databases\3.FDI\FDI_UNCTAD\ Outstocks\Excel*
- ... *\1. Raw Databases\3.FDI\Merge IMF and UNCTAD*
- ... *\1. Raw Databases\4.IR*
- ... *\1. Raw Databases\5.LMF data*
- ... *\1. Raw Databases\6.Price Deflator*
- ... *\1. Raw Databases\Country-Pairs*
- ... *\2.Cleaned Databases*
- ... *\2.Cleaned Databases\LBS Public data*
- ... *\3.Dofiles*
- ... *\3.Dofiles\1.Data Cleaning*
- ... *\3.Dofiles\2.Tables and Figures*
- ... *\3.Dofiles\3.Tables and Figures Banking* R-R
- ... *\4.Tables and Figures*
- ... *\4.Tables and Figures \IBS public*