Restatement Matrices for Bilateral External Portfolios

Antonio Coppola* Matteo Maggiori† Brent Neiman‡ Jesse Schreger§

December 2022, Version 1.5

This README file provides documentation for the file "Restatement_Matrices.dta", containing matrices that can be used to restate residency-based data on a nationality or sales-weighted basis, as outlined in Coppola et al. (2021). We provide several sets of matrices, corresponding to the various methodologies introduced and discussed in Coppola et al. (2021), which are identified via the field "Methodology" in the files:

1. The "Fund Holdings" methodology yields the baseline reallocation matrices introduced in Coppola et al. (2021), which are estimated using data on the holdings of global mutual fund and ETF portfolios obtained from Morningstar, and use our algorithm to aggregate associated issuances across corporate subsidiaries. We provide these matrices for the following economies as sources of the investments: the U.S., the European Monetary Union (EMU), Great Britain, Canada, Switzerland, Australia, Sweden, Denmark, and Norway. For all methodologies, we only consider the EMU as a block when on the investor side since mutual funds are concentrated in Luxembourg and Ireland, but collect investments from the rest of the countries in the European Union. Individual countries are instead kept as separate entries when on the issuer side.

2. The "Issuance" methodology yields the issuance distribution matrices introduced in Coppola et al. (2021), which are estimated using commercially-available data on bond and equity issuance from Dealogic, Factset, and Refinitiv. These matrices are not investor-specific, but rather reflect the total investable set of relevant securities globally.

3. The "Enhanced Fund Holdings" methodology provides analogs of our baseline reallocation matrices (point (1) above) that are constructed by complementing the Morningstar mutual funds and ETF holdings data with data on the holdings of other institutional sectors, where

*Coppola: Harvard University, Department of Economics. Email: acoppola@g.harvard.edu.
†Maggiori: Harvard University, Department of Economics, NBER, and CEPR. Email: maggiori@fas.harvard.edu.
‡Neiman: University of Chicago, Booth School of Business, and NBER. Email: brent.neiman@chicagobooth.edu.
§Schreger: Columbia Business School and NBER. Email: jesse.schreger@columbia.edu.

available. We provide estimates using this methodology for two countries: the United States, for which we complement the Morningstar data with data on the holdings of U.S. insurance companies, and Norway, for which we add data on the holdings of the Norwegian sovereign wealth fund. Both of these two additional holdings datasets are described in detail in Coppola, Maggiori, Neiman and Schreger (2021).

4. The "Sales-Based, Fund Holdings" methodology yields the baseline sales-based matrices introduced in Coppola et al. (2021), which link investment in a firm with the geographic distribution of its sales. These matrices use the Morningstar holdings (same as point (1) above) in order to establish the firm-level positions of various investing countries. We provide these matrices for the same set of countries as the "Fund Holdings" ones.

5. The "Sales-Based, Issuance" methodology provides a single set of sales-based matrices that are not investor-specific, but rather reflect the total investable set of relevant securities globally. They are estimated using data on total outstanding bond and equity issuance from Dealogic, Factset, and Refinitiv.

6. The "Sales-Based, Enhanced Fund Holdings" methodology is analogous to the "Sales-Based, Fund Holdings" one, but are constructed by complementing the Morningstar mutual funds and ETF holdings data with data on the positions of U.S. insurers and of the Norwegian sovereign wealth fund. As a consequence, they are only available for the United States and Norway.

We request that users of these data acknowledge their source and recommend inclusion of the following sentence:

> "These data are based on the work in Coppola, Maggiori, Neiman and Schreger (2021) and were obtained from: www.globalcapitalallocation.com".

We include matrices estimated separately for holdings in the following asset classes: equities; all bonds; corporate bonds; sovereign, agency, and local government (muni) bonds; asset-backed securities.[1] This release of the data includes matrices estimated for the years 2007 to 2020. The matrices are always estimated using data as of the last reporting date of the year (December). In order to reduce the file size, the matrices are in "sparse" format, meaning that only non-zero entries are included in the data.

[1]For the sales-based matrices, we only include equities, all bonds, and corporate bonds.

1 File Structure

The file has the following fields:

1. Methodology – The methodology underlying the estimates, one of the six defined above.
2. Year – Positions are measured at the end of December of this year.
3. Investor – 3-letter ISO code for the domicile country corresponding to Investor_Name; for the methodologies "Issuance" and "Sales-Based, Issuance", this field is populated with "World".
4. Asset_Class – Asset class of the investments.
5. Asset_Class_Code – A code corresponding to the "Asset_Class" field. The codes are the following: (i) "E" for common equities; (ii) "BC" for corporate bonds; (iii) "BG" for sovereign, agency, and local government (muni) bonds; (iv) "BSF" for asset-backed securities; (v) "B" for all bonds.
6. Destination – 3-letter ISO code for country (residency) of issuer of bond or stock.
7. Destination_Restated – 3-letter ISO code for country to which investments are reallocated.
8. Value – The share of the investments of "Investor" that should be reallocated from "Destination" to "Destination_Restated" according to "Methodology"; for each triplet of "Investor", "Destination", and "Methodology", the values in this field sum to 1.

References

Coppola, Antonio, Matteo Maggiori, Brent Neiman, and Jesse Schreger, "Redrawing the Map of Global Capital Flows: The Role of Cross-border Financing and Tax Havens," *The Quarterly Journal of Economics*, 2021, *136* (3), 1499–1556.