Global Fund Holdings Summary Statistics

The Global Capital Allocation Project

April 2020, Version 1.1

This README file provides documentation for the series of downloadable .CSV and .DTA files with summary statistics on global mutual fund and ETF portfolios. These tables make use of the fund holdings data obtained from Morningstar and described and analyzed in Maggiori, Neiman, and Schreger (2020), as well as the algorithm to aggregate associated issuances across corporate subsidiaries developed in Coppola, Maggiori, Neiman, and Schreger (2020). We are providing these statistics in support of those papers.[1] We request that users of these data acknowledge their source and recommend inclusion of the following sentence:

> "We obtained these data, which are based on the work in Maggiori et al. (2020) and Coppola et al. (2020), from: www.globalcapitalallocation.com."

The downloadable files include:

- Aggregate_Positions.csv
- Bilateral_Asset_Shares_X.csv
- Bilateral_Country_Shares_Y_X.csv
- Bilateral_Currency_Shares_Y_X.csv
- Bilateral_Industry_Shares_Y_X.csv,

where "X" indicates whether the results are based on an issuer or borrower's "residency", tax-haven-only version of nationality "nationality_th", or Full-nationality "nationality_full".[2] "Y" indicates the applicable type of asset and may be "eqty" for equities, "corp" for corporate bonds, or "govt" for government bonds. Though the underlying fund holding data includes assets such as cash or real estate, we only include equity and bond positions in these summary statistics.

[1] The statistics in these files differ somewhat from related statistics reported in our papers. For example, we have continued to update our methodology since publishing Maggiori et al. (2020), so some values will differ. And restating aggregate portfolios on a nationality basis, the focus in Coppola et al. (2020), requires a different weighting scheme across countries and asset classes compared to what is implicit in these summary statistics on mutual fund portfolios.

[2] See Coppola et al. (2020) for the distinction between the two versions of our nationality-based estimates.

As discussed in our work, the European mutual fund industry has assets concentrated in Ireland and Luxembourg. While we report our statistics for individual European countries, we emphasize that in most cases, they likely capture only a small share of the mutual fund and ETF investments made by their households.[3] For this reason, the baseline files we provide (as described above) report results only for the European Monetary Union (EMU) as an investor domicile, and do not report results separately for funds domiciled in EMU countries like France or Germany. We believe that for nearly all purposes, these baseline files are best for researchers to use.

Our results for the EMU as an investor simply pool together all the countries belonging to the EMU (for a given year) that are available in the data. These results therefore disproportionately reflect those EMU countries with better coverage.[4] For completeness, therefore, we also include a set of files equivalent to the five listed above but with "_emu_details" appended to the filenames. These files are equivalent to the baseline files but report for individual EMU countries.

We report individual EMU countries as issuers of securities, even in our baseline files, since the same concern does not exist when trying to determine the recipient of investments as exists when trying to think about the investor. For example, we report positions of U.S. investors in both France and in the EMU as a whole. Due to this double-counting, in some files, relevant shares may sum to a total that exceeds 100 percent.

Finally, the splits reported here might change over time as we improve our procedures and keep performing data checks. Starting from the initial release, Version 1.1, we will archive older releases of the splits.

1 Aggregate_Positions.csv

This file gives a summary of the overall positions and breakdown by asset class held in each country and year. It has the following fields:

1. Year – Positions are measured at the end of December of this year
2. Investor_Name – Domicile of the mutual fund or ETF
3. Investor – 3-digit ISO code for the domicile country corresponding to Investor_Name
4. Share_Corp_Bonds – Share of total assets held as corporate bonds

[3] For example, a significant share of mutual fund investments by Italians, say, might be with funds domiciled in Luxembourg.

[4] We also note that our construction of the EMU is not a constant panel.

5. Share_Govt_Bonds – Share of total assets held as government bonds

6. Share_Other_Bonds – Share of total aseets held as other bonds

7. Share_Equities_Bonds – Share of total assets held as equities

8. MNS_Total – Total value of bonds and stocks held by funds in Investor/Year

The shares in fields 4-7 sum to 100%. In all files, all positions in levels are measured in U.S. dollar and counted in millions.

2 Bilateral_Asset_Shares_X.csv

These files list the share of investment held in each asset class for bilateral relationships. There are three files with this naming convention: Bilateral_Asset_Shares_nationality_full.csv, Bilateral_Asset_Shares_nationality_th.csv, and Bilateral_Asset_Shares_residency.csv. These files list not only an Investor but also an Issuer (or borrower), so the files differ in terms of which notion (residency or each type of nationality) of issuing country is reported.[5]

For any given non-EMU Investor, the file only includes Issuers that account for at least 1 percent of total foreign investment in these data. For EMU Investors, all non-zero positions in EMU Issuers are included but the file only includes non-EMU Issuers that account for at least 1 percent of total investment outside of the EMU. Domestic investment, for which the Investor and Issuer are the same, is included for all Investor countries.

Each of the three files has the following fields:

1. Year – Positions are measured at the end of December of this year

2. Investor_Name – Domicile of the mutual fund or ETF

3. Investor – 3-digit ISO code for the domicile country corresponding to Investor_Name

4. Issuer_Name_X – Country (residency or nationality) of issuer of bond or stock

5. Issuer_X – 3-digit ISO code for country (residency or nationality) of issuer of bond or stock

6. Share_Corp_Bonds – Share of portfolio by Investor in Issuer held in corporate bonds

[5]Users should be aware that the procedure used to determine issuance by nationality is based on data as of December 2017. We apply the 2017-based aggregation of Coppola et al. (2020) to all years 2005-2017. As explained in Coppola et al. (2020) the procedure remains to a large extent accurate since corporate M&A activity normally results in updated securities identifiers, which reduces the need for a dynamic nationality assignment. Future releases of these splits might contain further improvements to the historical accuracy of our procedure.

7. Share_Govt_Bonds – Share of portfolio by Investor in Issuer held in government bonds
8. Share_Other_Bonds – Share of portfolio by Investor in Issuer held in other bonds
9. Share_Equities_Bonds – Share of portfolio by Investor in Issuer held in equities

The residency of an Issuer is the country of residency (i.e. incorporation of the issuing entity) of the immediate issuer of a security. The nationality of an Issuer is the main country of operation of the ultimate parent entity of the issuer of a security. See Coppola et al. (2020) for precise definitions of these fields. The shares in fields 6-9 sum to 100 percent.

3 Bilateral_Country_Shares_Y_X.csv

These files list the share across countries of investment in a given asset class for a given Investor. There are nine files with this naming convention: Bilateral_Country_Shares_govt_nationality_full.csv, Bilateral_Country_Shares_corp_nationality_full.csv, Bilateral_Country_Shares_eqty_nationality_full.csv, Bilateral_Country_Shares_govt_nationality_th.csv, Bilateral_Country_Shares_corp_nationality_th.csv, Bilateral_Country_Shares_eqty_nationality_th.csv, Bilateral_Country_Shares_govt_residency.csv, Bilateral_Country_Shares_corp_residency.csv, and Bilateral_Country_Shares_eqty_residency.csv. These files list not only an Investor but also an Issuer (or borrower), so the files differ in terms of which notion (residency or nationality) of issuing country is reported. The files also separately cover corporate bond investment (corp), government bond investment (govt), and equities (eqty).

For any given non-EMU Investor, the file only includes Issuers that account for at least 1 percent of total foreign investment in that asset class in these data. For EMU Investors, all non-zero positions in EMU Issuers are included but the file only includes non-EMU Issuers that account for at least 1 percent of total investment outside of the EMU in that asset class. Domestic investment, for which the Investor and Issuer are the same, is included for all Investor countries.

Each of the nine files has the following fields:

1. Year – Positions are measured at the end of December of this year
2. Investor_Name – Domicile of the mutual fund or ETF
3. Investor – 3-digit ISO code for the domicile country corresponding to Investor_Name
4. Issuer_Name_X – Country (residency or nationality) of issuer of bond or stock
5. Issuer_X – 3-digit ISO code for country (residency or nationality) of issuer of bond or stock
6. Portfolio_Share – Share of Investor's portfolio in that asset class issued by Issuer

4 Bilateral_Currency_Shares_Y_X.csv

These files list the share of corporate and government bond investment for any given Investor and Issuer bilateral pair that is denominated in any given currency. There are six files with this naming convention: Bilateral_Currency_Shares_govt_nationality_full.csv, Bilateral_Currency_Shares_corp_nationality_full.csv, Bilateral_Currency_Shares_govt_nationality_th.csv, Bilateral_Currency_Shares_corp_nationality_th.csv, Bilateral_Currency_Shares_govt_residency.csv, and Bilateral_Currency_Shares_corp_residency.csv. These files list not only an Investor but also an Issuer (or borrower), so the files differ in terms of which notion (residency or nationality) of issuing country is reported. The files also separately cover corporate bond investment (corp) and government bond investment (govt). We do not offer these files for other bonds or for equities, where the notion of currency of denomination is less well-defined.

For any given non-EMU Investor, the file only includes Issuers that account for at least 1 percent of total foreign investment in that asset class in these data. For EMU Investors, all non-zero positions in EMU Issuers are included but the file only includes non-EMU Issuers that account for at least 1 percent of total investment outside of the EMU in that asset class. Domestic investment, for which the Investor and Issuer are the same, is included for all Investor countries. Within any included Investor-Issuer pair, we only include currencies accounting for at least 1 percent of the total value.

Each of the six files has the following fields:

1. Year – Positions are measured at the end of December of this year
2. Investor_Name – Domicile of the mutual fund or ETF
3. Investor – 3-digit ISO code for the domicile country corresponding to Investor_Name
4. Issuer_Name_X – Country (residency or nationality) of issuer of bond or stock
5. Issuer_X – 3-digit ISO code for country (residency or nationality) of issuer of bond
6. Currency – 3-digit ISO code for currency of denomination of the bond
7. Currency_Name – Name of currency of denomination of the bond
8. Share – Share of Investor's portfolio of bonds from Issuer denominated in Currency

The shares in point 8 sum to 100 percent within an Investor/Issuer pair by asset class.

5 Bilateral_Industry_Shares_Y_X.csv

These files list the share of corporate bond and equity investment between any given Investor and Issuer bilateral pair that is accounted by any given industry. There are six files with this naming convention: Bilateral_Industry_Shares_corp_nationality_full.csv, Bilateral_Industry_Sh ares_corp_nationality_th.csv, Bilateral_Industry_Shares_corp_residency.csv, Bilateral_Industry_ Shares_eqty_nationality_full.csv, Bilateral_Industry_Shares_eqty_nationality_th.csv, and Bilater al_Industry_Shares_eqty_residency.csv. These files list not only an Investor but also an Issuer (or borrower), so the files differ in terms of which notion (residency or nationality) of issuing country is reported. The files also separately cover corporate bond investment (corp) and equity investment (eqty). We do not offer these files for government or other bonds, where the notion of industry is more abstract.

For any given non-EMU Investor, the file only includes Issuers that account for at least 1 percent of total foreign investment in that asset class in these data. For EMU Investors, all non-zero positions in EMU Issuers are included but the file only includes non-EMU Issuers that account for at least 1 percent of total investment outside of the EMU in that asset class. Domestic investment, for which the Investor and Issuer are the same, is included for all Investor countries. Within any included Investor-Issuer pair, we only incude industries accounting for at least 1 percent of the total value.

The industry classification is based on 2-digit NAICS codes. We aggregate these codes to the following industries at 1-digit level:

- "Agriculture, Mining, Utilities, and Construction" (NAICS ≤ 29)
- "Manufacturing" (NAICS $\in [30, 39]$)
- "Wholesale and Retail Trade" (NAICS $\in [42, 45]$)
- "Transportation and Warehousing" (NAICS $\in [48, 49]$)
- "Information" (NAICS $= 51$)
- "Finance, Insurance, Real Estate" (NAICS $\in [52, 53]$)
- "Business Services" (NAICS $\in [54, 56]$)
- "Education, Health, Public Administration, and Other Services" (NAICS $\in \{61, 62, 81, 92\}$)
- "Arts, Entertainment, Hotels, and Dining" (NAICS $\in [71, 72]$)

- "Missing"

For each issuing entity we use the primary industry code if more than one code is available for that entity. In the residency view we use the industry code of the immediate issuer. In the nationality view we use the industry code of the ultimate parent entity of the immediate issuer.

Each of the six files has the following fields:

1. Year – Positions are measured at the end of December of this year
2. Investor_Name – Domicile of the mutual fund or ETF
3. Investor – 3-digit ISO code for the domicile country corresponding to Investor_Name
4. Issuer_Name_X – Country (residency or nationality) of issuer of bond or stock
5. Issuer_X – 3-digit ISO code for country (residency or nationality) of issuer of bond or stock
6. Sector – 1-digit Industry characterizing the activity of the bond or equity issuer
7. Share – Share of bilateral position in each Sector

The shares in point 7 sum to 100 percent within an Investor/Issuer pair by asset class.

References

Coppola, A., M. Maggiori, B. Neiman, and J. Schreger (2020). Redrawing the map of global capital flows: The role of cross-border financing and tax havens. *NBER Working Paper No. 26855*.

Maggiori, M., B. Neiman, and J. Schreger (2020). International currencies and capital allocation. *Forthcoming in Journal of Political Economy*.