IMF Working Paper


by Robin Koepke and Simon Paetzold

IMF Working Papers describe research in progress by the author(s) and are published to elicit comments and to encourage debate. The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.
IMF Working Paper
Asia & Pacific Department

Prepared by Robin Koepke and Simon Paetzold

Authorized for distribution by Todd Schneider
August 2020

Abstract
This paper provides an analytical overview of the most widely used capital flow datasets. The paper is written as a guide for academics who embark on empirical research projects and for policymakers who need timely information on capital flow developments to inform their decisions. We address common misconceptions about capital flow data and discuss differences between high-frequency proxies for portfolio flows. In a nowcasting “horse race” we show that high-frequency proxies have significant predictive content for portfolio flows from the balance of payments (BoP). We also construct a new dataset for academic use, consisting of monthly portfolio flows broadly consistent with BoP data.

JEL Classification Numbers: F21, F32, F37, G15

Keywords: portfolio flows, fund flows, balance of payments, coincident indicators

Author’s E-Mail Addresses: rkoepke@imf.org, paetzold.simon@gmail.com

1 We are grateful for helpful suggestions by Gustavo Adler, Swarnali Ahmed, Serkan Arslanalp, Roland Beck, Mahir Binici, Robin Brooks, Eugenio Cerutti, Nathan Converse, Dimitris Drakopoulos, Thomas Elkjaer, Eric Fischer, Jonathan Fortun, Joji Ishikawa, Sebnem Kalemli- Özcan, Yevgeniya Korniyenko, Swapan-Kumar Pradhan, Gian Maria Milesi-Ferretti, Gurmain Pasricha, Daniel Rodriguez, Russell Rollow, Todd Schneider, Beatrice Scheubel, and Ling Zhu. We would also like to acknowledge the valuable comments from participants in several IMF seminars on monitoring capital flows, including during presentations given to the IMF’s Capital Flows Group, the Asia & Pacific Department, the Western Hemisphere Department, and two ICD lectures on “Tracking, Interpreting and Forecasting Capital Flows Developments.”
## Contents

**I. Introduction** .......................................................................................................................... 4  
**II. Overview of Capital Flow Data** ............................................................................................... 6  
   A. Key Capital Flow Data Sources ......................................................................................... 7  
   B. Balance of Payments Framework: Misconceptions and Limitations .............................. 14  
**III. Capital Flow Data in the Literature** .................................................................................... 18  
   A. Meta-Study of Data Sources Used in the Empirical Literature ........................................ 18  
   B. Constructing a Monthly Portfolio Flow Dataset Consistent with BoP Principles ........... 20  
**IV. Real-Time Tracking of Portfolio Flows** ............................................................................. 25  
**V. Conclusion** ............................................................................................................................ 28  

**Box**  

**Figures**  
1. Start-date of Common Portfolio Flow Data Sources ............................................................ 4  
2. IIF Portfolio Flows vs. EPFR Fund Flows ........................................................................... 4  
3. Equity and Debt, RMSFE in Percent of GDP ................................................................... 6  
4. Data Sources in Academic Dataset ................................................................................... 7  
5. High Frequency Capital Flow Data in Policy Dataset, 2010-20 ....................................... 7  
6. Why Capital Flow Data are Difficult to Navigate ................................................................ 15  
7. Advanced Economy Holdings of EMDE Equity ................................................................. 17  
8. Data Frequency in Meta-Dataset ....................................................................................... 18  
9. Capital Flow Type in Meta-Dataset ................................................................................... 18  
10. BoP-Consistent Flows by Frequency ............................................................................... 20  
11. Number of Papers by Flow Component and Frequency: Pull vs Push Factors .............. 20  
12. Total Monthly Equity and Debt Flows from KP Dataset ................................................ 21  
13. Comparison Balance of Payments and EM Portfolio Flow Proxies ............................... 23  
14. Comparison KP Dataset and Balance of Payments Portfolio Flows ............................. 24  
15. Comparison EPFR and Balance of Payments Portfolio Flows ........................................ 24  
16. Equity Portfolio Flows to EMs ....................................................................................... 27  
17. Debt Portfolio Flows to EMs ............................................................................................ 27
Tables
1. Overview of Capital Flow Data Sources ................................................................. 9
2. Sign Conventions under BPM5 vs. BPM6.............................................................. 16
3. Description of Data in the KP Monthly Portfolio Flow Dataset .............................. 22

Annexes
I. Bilateral Capital Flows Data (“From-Whom-to-Whom” Data)................................. 35
II. Daily IIF Portfolio Flows Data.............................................................................. 37
III. List of Papers Included in the Academic Dataset ............................................ 38
IV. Portfolio Flow Nowcasting Results by Country................................................ 45

References.............................................................................................................. 30
I. INTRODUCTION

In early 2020, the COVID-19 pandemic triggered one of the sharpest reversals in portfolio flows to emerging markets (EMs) on record. Timely capital flow proxies served as early warning indicators, alerting policymakers to the severity of the shock. While traditional quarterly balance of payments (BoP) data would only become available several months after 2020:Q1, proxy data on monthly, weekly and even daily flows from Emerging Portfolio Fund Research (EPFR) and the Institute of International Finance (IIF) showed as early as March 2020 that the reversal was extremely severe in speed and magnitude (IIF 2020; IMF 2020).

Data availability on international capital flows has improved dramatically over the past two decades, particularly high-frequency proxies for portfolio flows (Figure 1). New data sources provide valuable opportunities for advancing academic research and informing policy decisions in real time. However, these data sources also pose new challenges to researchers and policymakers. There are numerous conceptual and measurement issues surrounding these datasets, which are compounded by long-standing misconceptions about capital flow data in general.

Many of these conceptual and measurement issues are reflected in the differences between EPFR and IIF proxies for portfolio flows to EMs. While both datasets generally signal turning points correctly, they often provide widely different estimates of EM portfolio flows (Figure 2). For example, from 2010:Q1-2019:Q2 average quarterly fund flows to EMs as a group were about $11 billion according to EPFR fund flow data, compared to $68 billion for the IIF Portfolio Flows Tracker and $71 billion for IMF BoP portfolio flow data.\(^1\)

---

\(^1\) The numbers referenced for IIF and IMF data relate to the concept of “gross” portfolio inflows, defined as non-resident purchases of EM stocks and bonds. Section 2 provides a detailed discussion of differences in scope and data coverage of BoP, EPFR and IIF data.
These issues are exacerbated by the fact that private sector data providers often make available only limited documentation of the methodological principles underlying the construction of their datasets. In particular, data providers (and the empirical literature) devote little attention to discussing how these principles differ from standard balance of payments accounting conventions. Moreover, only very limited work has been done comparing and contrasting different datasets in order to assess what data are most appropriate to answer which research questions.

The objective of this paper is to provide clarity on these issues. The paper is written as a guide for academics who embark on empirical research projects and for policymakers and market participants who need timely information on capital flow developments to inform their decisions. From a policy perspective, the data discussed in this paper are an important ingredient for understanding the policy challenges created by international capital flows, documented in an extensive literature (some surveys include Dooley 1996; BIS 2009; Milne 2014).

This paper makes four main contributions. First, we provide an overview of the most widely used datasets available today, along with a discussion of the strengths and weaknesses of various data sources. Special attention is given to high frequency measures of financial transactions like portfolio flow and fund flow data. We also address common misconceptions about capital flow data in the context of the BoP accounting framework and discuss recent research highlighting the limitations of that framework (e.g., Lane and Milesi-Ferretti 2018; Coppola et al. 2020).

Second, we conduct a meta-study on the use of data sources in the empirical literature, exploring how the choice of data may have shaped the empirical results that were obtained in these studies. We argue that some empirical findings may be distorted by data availability. For example, the widespread use of fund flow data may have over-emphasized the importance of external ("push") factors in empirical results, given that fund flows by construction are subject to “benchmark effects” that are common to many emerging markets (Raddatz et al. 2017; Arslanalp et al. 2020).

Third, we address a gap in the availability of portfolio flow data for academic research. Our meta-study shows that academic research has made little use of BoP-consistent portfolio flow data at high data frequencies. This is in part due to subscription requirements and issues with the construction of country samples. We make a contribution towards closing the gap in data

---

2 An important benefit of using capital flow data on a BoP basis is that they are consistent with other macroeconomic accounts, such as current account data, government finance statistics, and the international investment position. Capital flow data are often analyzed through the prism of external financing (of a current account deficit) and financing of a government’s budget deficit. Indeed, a key motivation of the voluminous literature on the role of external (“push”) and domestic (“pull”) drivers of capital flows has been to advance our understanding of the degree to which foreign capital inflows are likely to be a temporary rather than permanent source of external finance (e.g., Calvo et al. 1993, Chuhun et al. 1993; Fernandez-Arias 1996; surveyed in Koepeke 2019). This motivation argues for the use of BoP-consistent data that can be analyzed and interpreted in conjunction with other standard macroeconomic data.
availability by providing a free online dataset with monthly emerging market portfolio flow data, specifically geared towards academic use.

Fourth, we provide a quantitative assessment of how well widely used portfolio flow proxies track portfolio flow developments in real-time. Results of a nowcasting “horse race” suggest that IIF and EPFR data have significant predictive content for BoP-based portfolio flows, reducing forecast errors by 80-90 percent relative to an autoregressive model (Figure 3). Portfolio flow proxies at the daily and weekly data frequencies outperform monthly data in the first half of the current quarter, and IIF data generally outperform EPFR data.

The rest of this paper is structured as follows. Section 2 provides an overview of capital flow data sources and addresses common misconceptions. Section 3 presents a meta-study on the usage of capital flow data in the empirical literature and introduces a new monthly portfolio flow dataset. Section 4 assesses the predictive content of several portfolio flow proxies. Section 5 concludes the paper.

II. OVERVIEW OF CAPITAL FLOW DATA

This section provides a concise overview of the most commonly used capital flow data sources. We construct two sets of metadata to assess what data sources academics and policymakers have used the most. The first dataset represents the use of capital flow data in the empirical literature (“academic dataset”), while the second dataset represents the use of timely portfolio flow proxies by policymakers (“policy dataset”). Both datasets are posted on the IMF website, along with this paper.

In the 88 studies that constitute the academic dataset, IMF balance of payments data are the most common data source (used in 39 percent of all studies), followed by BIS data (20 percent) and EPFR data (14 percent; Figure 4). Less commonly used data sources include

---

3 For the academic dataset, we build on the work of Koepke (2019), which surveys 34 empirical studies to analyze the empirical drivers of capital flows to emerging markets. We extend this sample of studies by including an additional 54 studies in this same strand of literature, focusing on papers with a large number of citations and/or published in leading economic journals. Overall, we evaluate a total of 88 studies published over the course of the last 27 years and listed in Annex III. For the policy dataset, we review 220 reports published since 2010, including financial stability reports by G-20 emerging market central banks and reports of policy-oriented international financial institutions, notably the IMF’s Global Financial Stability Reports, the BIS’s Annual Reports, and the World Bank’s Global Development Prospects report.
the World Bank’s World Development Indicators and the U.S. Treasury International Capital (TIC) data.

In the 111 studies that constitute the policy dataset, high-frequency portfolio flow proxies are used in about 50 percent of studies since 2010 (and about 60 percent of those published in 2017-19). EPFR accounts for about two thirds of those studies (and about 50 percent in 2017-19), followed by IIF data (21 percent in the full sample, 34 percent in 2017-19; Figure 5). The remaining 12 percent of studies make use of other high-frequency data sources, mostly national or regional stock exchange data.

There are a number of additional useful capital flow data sources that deserve attention even though they have seen little or no use in the studies included in our academic and policy datasets. Some of these data sources are relatively new and likely to receive future attention in the literature, notably “big data” initiatives such as SWIFT’s data on transactions facilitated by banks. Another area where data coverage has expanded significantly in recent years is bilateral capital flow data (“from-whom-to-whom” data), which are discussed in Annex I. Additional notable datasets include BIS debt securities statistics, the ECB’s Securities Holdings Statistics, Morningstar’s data on investment funds, UNCTAD data on FDI, and Bureau van Dijk’s Zephyr on M&A and equity investments.

A. Key Capital Flow Data Sources

Table 1 presents detailed information on data characteristics as well as advantages and caveats relating to most commonly used data sources among scholars and practitioners.4

4 The use of capital flow data sources in the academic literature is discussed in detail in section 3. Data sources commonly used by practitioners for tracking capital flows in real time are discussed in section 4.
1. IMF Balance of Payments Statistics

Capital flow data as reported in the IMF’s Balance of Payment Statistics (BOPS) provide the most comprehensive country coverage and refined methodology. The data are reported to the IMF by country authorities and cover the various capital flow components (presented as subcomponents of the financial account) as well as all other major international transactions, including current and capital transfers. Capital flow data are generally available both on a gross and net basis for each major component (such as “net FDI”). For the majority of countries, data are available on a quarterly and annual basis and are typically released with a lag of two to four months.

The IMF’s balance of payments accounting principles serve as an important anchor for the analysis of capital flow data. Now in its sixth edition, the IMF’s Balance of Payments and International Investment Position Manual (BPM6) and its companion Compilation Guide provide detailed guidance on methodological standards. These internationally recognized standards ensure that capital flow data are comparable across countries and time, which is an important precondition for empirical research and policy analysis.

Some of the key accounting principles include:

- **Residency**: Capital flows arise from the acquisition and disposal of financial assets and liabilities between residents of different countries.\(^6\)\(^,\)\(^7\)

- **Quadruple entry bookkeeping**: Each transaction is recorded twice in each of the two countries’ BoP, reflecting the source of funds and the use of funds in each country.\(^8\)

- **Transactions at market value**: To the extent possible, capital flows are recorded using the market value at the time of the transaction. Changes in valuations (“valuation effects”) do not affect BoP data.\(^9\)

---


\(^6\) The BPM6 notes that “residence of each institutional unit is the economic territory with which it has the strongest connection, expressed as its center of predominant economic interest” (¶4.113).

\(^7\) It is worth noting that the term “flow” has a different meaning in the capital flow literature than in the terminology of balance of payments statistics. In the capital flow literature, a “flow” relates to transactions. In macroeconomic statistics, a “flow” refers to all changes in the stock of a variable from one period to the next, including transactions, valuation changes, and other flows (see BPM6 ¶2.2).

\(^8\) The BPM6 specifies that “a single transaction between two counterparties thus gives rise to four entries. […] As a liability of one unit is mirrored in a financial asset of another unit, for instance, they should be identically valued, allocated in time, and classified to avoid inconsistencies in aggregating balance sheets of units into regional or global totals” (¶3.29).

\(^9\) The BPM6 notes that “market prices are the basis for valuation in the international accounts” (¶3.67). However, some financial instruments (e.g., loans) are recorded at nominal values.
## Table 1. Overview of Capital Flow Data Sources

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Data Frequency</th>
<th>Typical Release Lag</th>
<th>Country Coverage</th>
<th>Scope</th>
<th>Advantages</th>
<th>Caveats</th>
<th>Public (P) or Subscription-based (S)</th>
<th>Start date</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>International Monetary Fund (IMF): Balance of Payments Statistics</strong></td>
<td>quarterly</td>
<td>2-4 months</td>
<td>Almost all EMs</td>
<td>All major capital flow components</td>
<td>Comprehensive coverage of cross-border transactions; well-defined and established methodology</td>
<td>Statistical break in the 2005-2008 period due to shift from BPM5 to BPM6 (the IMF BOPS adjusts and publishes historical data in the BPM6 format)</td>
<td>P</td>
<td>Jan-2002 (equity); Jul-2002 (debt) 1/</td>
</tr>
<tr>
<td></td>
<td>annual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S</td>
<td>1977 (equity and debt) 1/</td>
</tr>
<tr>
<td><strong>Bank of International Settlements (BIS); Locational Banking Statistics (LBS)</strong></td>
<td>quarterly</td>
<td>4 months</td>
<td>Based on counterparty reporting for up to 200 countries/jurisdictions, around 95% of cross-border banking activity captured</td>
<td>Cross-border banking activity: Loans and deposits, debt securities, other instruments</td>
<td>Comprehensive coverage of cross-border banking activity; currency composition and counterpart sector available; data on both immediate and ultimate counterparty basis</td>
<td>No direct mapping between BIS data and standard BoP components of capital flows; flows are constructed based on stock data, adjusted for FX valuation effects</td>
<td>P</td>
<td>Oct-1977</td>
</tr>
<tr>
<td><strong>BIS Consolidated Banking Statistics (CBS)</strong></td>
<td>quarterly (since 1999Q4); bi-annual (since 1983Q4)</td>
<td></td>
<td>Based on counterparty reporting for up to 200 countries/jurisdictions</td>
<td></td>
<td>Additional data on guarantor (since early 2005) and consolidated basis</td>
<td>No data on constructed flows</td>
<td></td>
<td>Jan-1984</td>
</tr>
<tr>
<td></td>
<td>daily/weekly</td>
<td>1-3 days</td>
<td>15 countries for equity flows, 9 for debt flows, 2 for equity and debt combined</td>
<td>Portfolio equity and debt flows (only non-resident flows to EMs)</td>
<td>Good proxy for quarterly BoP data, especially for equity flows; high data frequencies and short release lag</td>
<td>For some countries, debt flows cover only local currency and/or sovereign bonds; does not capture bond purchases in the primary market and maturing bonds</td>
<td></td>
<td>Jan-2005 (equity and debt)</td>
</tr>
<tr>
<td><strong>Institute of International Finance (IIF): Portfolio Flows Trackers and Capital Flows Databases</strong></td>
<td>monthly</td>
<td>1-5 days</td>
<td>35 countries, mostly EMs and some frontier countries; some have only equity/debt flows</td>
<td>Portfolio equity and debt flows (only non-resident flows to EMs)</td>
<td>Good proxy for quarterly BoP data; short release lag</td>
<td>Data for the most recent 2-3 months get revised; for some countries, debt flows cover only local currency and/or sovereign bond</td>
<td>S</td>
<td>Jan-2005 (equity and debt)</td>
</tr>
<tr>
<td></td>
<td>quarterly</td>
<td>2-4 months</td>
<td>25 EMs</td>
<td>FDI, portfolio equity/debt flows, other investment (only non-resident flows to EMs)</td>
<td>Includes forecasts for the current year and year ahead</td>
<td>Limited country sample relative to IMF BOPS</td>
<td></td>
<td>Jan-2005 (equity and debt)</td>
</tr>
<tr>
<td></td>
<td>annual</td>
<td>2-4 months</td>
<td>25 EMs</td>
<td>FDI, portfolio equity/debt flows, other investment</td>
<td>Comprehensive cross-country data; includes forecasts for current year</td>
<td>Low data frequency; limited country sample relative to IMF BOPS</td>
<td></td>
<td>1978 (equity and debt)</td>
</tr>
<tr>
<td><strong>Emerging Portfolio Fund Research (EPFR): Fund Flows Data</strong></td>
<td>daily</td>
<td>1 day</td>
<td>Almost all EMs, some frontier markets</td>
<td>Inflows into EM-dedicated investment funds (equity, debt)</td>
<td>High data frequencies and short release lags; provides insights into the behavior of ultimate investors of investment funds; granular data available by currency denomination, fund domicile, and investor type, among others</td>
<td>Conceptually different from BoP data; based on large sample of reporting funds; institutional investors are underrepresented; country-level flows are estimated top-down</td>
<td>S</td>
<td>May-2007 (equity and debt)</td>
</tr>
<tr>
<td></td>
<td>weekly</td>
<td>1 day</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S</td>
<td>Oct-2000 (equity); Jul-2004 (debt)</td>
</tr>
<tr>
<td></td>
<td>monthly</td>
<td>15 days</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S</td>
<td>Feb-1996 (equity); Nov-2003 (debt)</td>
</tr>
</tbody>
</table>

Sources: BIS, BOPS, EPFR, IIF, and IMF.
1/The start date is determined by the first time more than 50% of current EMDEs report their data in the IMF BOPS.
2. **BIS Locational and Consolidated Banking Statistics**

The cross-border banking data from the BIS provide comprehensive coverage of international banking flows and positions, including currency composition, instrument type, and sector and residency of counterparty (BIS 2019). The data are reported to the BIS by national central banks and compiled following methods that are broadly consistent with BoP accounting principles. The Locational Banking Statistics (LBS) data are available on a quarterly basis since December 1977, with instrument breakdown available from December 1995. The LBS data are most closely related to the “other investment” component in the BoP, which includes cross-border bank loans and deposits, but the LBS data also capture banks’ holdings of debt securities, which in the BoP are classified as portfolio investment. The reported currency breakdown and break-in-series allow the BIS to derive “FX and break-adjusted changes” and unlike BoP flows, the adjusted changes are technically not “flows” in a narrow sense.

In the Consolidated Banking Statistics (CBS), intra-banking group claims are excluded, in contrast to the LBS. The CBS data on an immediate counterparty basis are available at the semi-annual frequency from December 1983 for 16 years and on a quarterly basis from December 1999 onwards. In addition, the CBS provide data both on an immediate counterparty basis and on a guarantor basis, where the guarantor is the entity assuming contractual responsibilities if the immediate counterparty defaults (BIS 2019). The quarterly CBS data on a guarantor basis are publicly available since early 2005. Guarantor basis data record a cross-border banking flow between country A and the country of residence of the counterparty’s guarantor, rather than the country of residence of the counterparty itself.

3. **IIF Capital Flow Data and Portfolio Flow Trackers**

The Institute of International Finance (IIF) has provided independent, private sector estimates of capital flows to and from emerging market economies since the 1990s (IIF 2007). Up until 2013, IIF data was exclusively at the annual frequency, covering all major capital flow components, and estimated using a proprietary methodology to ensure consistency with stock data on external debt. In 2014 and 2015, the IIF began making available datasets on monthly and daily portfolio flow data that are now widely used (Koepke and Mohammed 2014; Koepke and Kunii 2015). These datasets serve as timely proxies for BoP-based portfolio flow data and are constructed using country-level data from individual national sources. These national sources are often used as input data for authorities’ official balance of payments data on portfolio flows (typically published at the quarterly or annual frequency). Separately, the

---

11 Pradhan and Silva (2019) show that the data contained in the LBS align relatively closely with the IMF’s International Investment Position Statistics but there are several sources of discrepancies, including differences in reporting populations, mix of data sources, treatment of instruments, and country-specific factors.

12 The LBS provide bilateral data for both claims and liabilities whereas the CBS provide data mainly for claims and very limited data on liabilities (overall total and local currency liabilities in host countries).

13 More specifically, the data serve as a proxy for non-residents’ net acquisition of EM assets, i.e. inward portfolio flows to EMs. For some countries, the sum of daily or monthly observations is identical to the official quarterly BoP data on inward portfolio flows for the same period.
IIF also provides monthly estimates of net capital flows (i.e., the financial account balance) for a group of 23 emerging market economies (Fortun 2020).

The coverage of IIF monthly and especially daily portfolio flows data varies at the country level, depending on the information provided by the underlying source. For equity flows, coverage is quite comprehensive, reflected in a close match with BoP portfolio equity flows. For bond flows, some country data do not include certain types of securities such as hard currency debt or corporate debt (Farnham and Koepke 2015; see also Tables A2.1 and A2.2 in Annex II for detailed country information).

**Monthly flows:** The IIF’s monthly tracker for overall emerging market flows uses an econometric model in which underlying portfolio flow data are supplemented with financial variables and bond issuance data to minimize the statistical deviation from quarterly BoP portfolio flow data. The calibration of the model has evolved over the years and saw its fourth update in 2018 (Farnham and Tiftik 2018).

**Daily flows:** In contrast to the monthly dataset, the aggregate daily data are simply the sum of all reported country data. Some countries are excluded from the aggregate daily tracker because the release lag is longer than for the other countries (e.g., for Mexico, the release lag is typically 10 days). In terms of the underlying data, it is worth noting that for some countries, daily flows are estimated based on stock data, which are sometimes published in local currency terms and thus are being converted into U.S. dollar terms. Purchases of newly issued bonds and redemptions of maturing bonds are generally not included.

4. **Fund Flow Data**

Another popular high-frequency proxy for portfolio flows is fund flow data. Several private sector data providers make available subscription-based data on flows to investment funds (i.e., mutual funds and exchange-traded funds). The most widely used data provider is EPFR, which makes available monthly, weekly and daily estimates of fund flows for EMs as a group and at the country level, based on a large sample of reporting funds. For monthly data, more than 18,000 reporting equity funds and more than 9,000 reporting debt funds cover around 96 percent of assets under management (AUM) of the global investment fund industry, using the Investment Company Institute’s global AUM measure (as of September 2019, Informa Financial Intelligence 2020a). Weekly and daily data comprise more than 15,000 reporting equity funds and 7,000 reporting debt funds with about 65 percent of the AUM of monthly EPFR data (Informa Financial Intelligence 2020b).

EPFR also publishes estimates of country-level fund flows, which are constructed by applying the monthly average of a fund group’s country specific portfolio allocation share to the flows reported by that fund group. These estimates rely on several simplifying assumptions. For example, valuation changes affecting the change in country allocations from one period to the next are assumed to be zero. Moreover, not all funds make available the country-level portfolio allocations needed to estimate country flows at the fund level, so EPFR applies the average country allocation of one fund group to all funds in this group. Therefore, EPFR’s data on country-level flows are considerably less robust than its data on
flows to emerging markets as a group and typically differ greatly from country-level portfolio flows data (see also Figure 15).

A major benefit of EPFR data is that fund flow data are available by fund types, such as ETFs vs. mutual funds, funds oriented towards retail vs. institutional investors, and active vs. passive funds. Moreover, the data can be disaggregated by the fund domicile, investment benchmark, thematic fund category, and by local vs. hard currency denomination. Additionally, the data can be further disaggregated by specific categories for both equity and bond funds (e.g., corporate/sovereign, duration, capitalization, sector, etc.).

An important caveat is that EPFR data do not cover all types of emerging market investors, only those investing via mutual funds and exchange traded funds. For example, large institutional investors like sovereign wealth funds, pension funds, hedge funds, and banks’ proprietary trading desks typically purchase EM securities directly and are generally not reflected in EPFR data.

Additional providers of fund flow data include Lipper Fund Flows and Trounceflow, State Street and BNY Mellon offer data products based on their custody holdings.

**Box 1. The Rise of Portfolio Flow Proxies: What Accounts for the Differences between EPFR, IIF and BoP Data?**

**In recent years, portfolio flows have received substantial attention from academics, policymakers and market participants.** One reason is that portfolio flows have grown rapidly after the global financial crisis (notably debt flows) and have been the most volatile component of capital flows. Moreover, portfolio debt and equity flows are most closely tied to asset price and exchange rate fluctuations (Bergant and Schmitz 2018), making them highly relevant for central bank policy decisions. In addition, data availability on portfolio flows is far better than for any other component of capital flows, since there are no comparable monthly, weekly or daily data on banking flows, FDI, or “other investment” flows.

**The growing use of high-frequency portfolio flow proxies is evident in both the academic and policy meta-datasets compiled for this paper (Figures B.1 and B.2).** EPFR data in particular have been used frequently in recent academic studies (e.g., Fratzscher 2012; Anachotikul and Zhang 2014; Puy 2016; Bonizzi 2017; Converse et al. 2020).

**Figure B.1. High Frequency Capital Flow Data in the Academic Dataset, 2010-18 1/**

<table>
<thead>
<tr>
<th>Year</th>
<th>EPFR</th>
<th>TIC</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010-12</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013-15</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2016-18</td>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: IMF staff calculations.

1/ The chart shows 14 papers published since 2010.

**Figure B.2. High Frequency Capital Flow Data in Financial Stability Reports, 2010-19 1/**

<table>
<thead>
<tr>
<th>Year</th>
<th>EPFR Data</th>
<th>IIF Data</th>
<th>Other Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>10</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2019</td>
<td>18</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: IMF staff calculations.

1/ The chart shows 111 reports by G20 EMs, the IMF, World Bank, and BIS.
Box 1. The Rise of Portfolio Flow Proxies: What Accounts for the Differences between EPFR, IIF and BoP Data? (continued)

Available datasets on high-frequency portfolio flows differ widely in scope. There are three main sources of discrepancies between EPFR data, IIF data, and quarterly balance of payments data. They relate to conceptual differences, data coverage, and EM country samples (Koepke and Mohammed 2014; Farnham and Tiftik 2018).

Conceptual differences: The IIF's monthly and daily data aim to be consistent with balance of payments accounting principles, in which an inflow is recorded if there is a transaction between a non-resident and a resident. By contrast, flows into investment funds do not necessarily result in cross-border transactions. For example, an EM-dedicated mutual fund may see an inflow from ultimate investors and increase its cash buffer rather than buying EM securities. Similarly, if the mutual fund is domiciled in the U.S. and buys foreign securities from another U.S. investor, this transaction would not be recorded as a capital flow in the balance of payments. This issue also applies to residents of emerging markets who purchase shares of funds that invest domestically (which again would not be recorded in the balance of payments). In EPFR’s country flow data, this issue is particularly relevant for Thailand, and India, which have sizeable investment funds that are domiciled and invested locally.

Data coverage: In the IIF monthly and especially daily flows, some of the country level data do not include specific types of portfolio transactions if these are not reported by the original data provider (such as a national central bank). For example, some countries’ data do not include hard currency debt or corporate debt.

EPFR data do not cover all types of emerging market investors, only mutual funds and exchange traded funds. For example, large institutional investors like sovereign wealth funds, pension funds, and hedge funds typically purchase EM securities directly and generally do not report to EPFR.

Country sample: There is no universally accepted definition of which countries are emerging markets. For EPFR data, the emerging market universe is guided by which countries are included in key benchmark indices and covers data for a total of 98 countries for equity and 113 countries for debt flows. For IIF data, the emerging market universe is additionally guided by which countries make available timely portfolio flow data at the monthly (35 countries) and daily (21 countries) frequencies.

Bottom line: When to use EPFR vs IIF data? Both datasets are useful for monitoring directional shifts in investor interest in emerging market assets. Overall, EPFR data seem best suited for analyzing questions relating to (fund) investor behavior. In academic research, such questions are more likely to arise in the finance literature than in the international economics literature (e.g., Hu et al. 2014; Morris, Shim and Shin 2017; Shek, Shim and Shin 2018). Additionally, fund flow data may be useful to investment professionals for informing asset allocation decisions. By contrast, IIF data seem most appropriate for analyzing portfolio flows in a macroeconomic and external financing context, which are more common in the international economics literature than in the finance literature (e.g., Milesi-Ferretti and Tille 2011; Forbes and Warnock 2012; Koepke 2018, Cerdeiro and Komaromi 2019). IIF data may also be preferable for answering most policy-related questions (unless they relate to (fund) investor behavior). Finally, for country level analysis, IIF data are often preferable to EPFR data, given the simplifying assumptions used to construct the latter (see above, “Fund Flow Data”).
B. Balance of Payments Framework: Common Misconceptions and Measurement Limitations

Navigating capital flow data is more difficult than most other economic data, resulting in several common misconceptions and frequent miscommunication among economists (see, for example, the discussion in Avdjiev et al. 2020). This is in part due to the two-way nature of international capital flows: Inward investment reflects changes in foreign holdings of domestic assets, while outward investment reflects changes in domestic holdings of foreign assets. Both inward and outward investment flows can be positive or negative. And the two can be netted against each other to obtain a “net” measure of capital flows.

To provide clarity on common misconceptions, it is useful to illustrate how capital flows are captured in the financial account of the balance of payments, shown in Figure 6. In this illustration, changes in financial assets are separated from changes in liabilities. For each of the two, capital flows can be broken down further into the main components (FDI, portfolio equity, portfolio debt, and other investment). \(^{14}\)

\(^{14}\) The financial account also records financial derivatives, but these are excluded in this discussion of capital flow data.
There appear to be three main sources of misunderstandings, corresponding to the different levels of Figure 6:

- **First, there are several relevant levels of netting, making the commonly used terms of “net” and “gross” capital flows ambiguous.** For example, an important measure of capital flows is the net change in liabilities in the financial account, which captures the net acquisition of a country’s assets by non-residents. This is a “net” measure of capital flows in the sense that different liability components (inward FDI, inward portfolio flows, etc.) are netted against each other. But it is a “gross” measure of capital flows in the sense that changes in liabilities are not netted against changes in assets. Indeed, much of the literature refer to this concept as “gross capital flows” (e.g., Chuhan, Claessens and Mamingi 1998; Broner et al. 2013; Forbes and Warnock 2012; see also the discussion in Avdjiev et al. 2020).

- **Second, long-standing sign conventions changed with the introduction of the sixth edition of the IMF’s Balance of Payments and International Investment Position Manual (BPM6) in 2009.** The most notable changes included that an increase in foreign assets (including reserves) is shown with a positive sign (and the same holds for a decrease in liabilities; see Table 2 below).
Table 2. Sign Conventions under BPM5 vs. BPM6

<table>
<thead>
<tr>
<th></th>
<th>BPM5</th>
<th>BPM6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Current and capital account</strong></td>
<td>Credits with <em>positive</em> sign. Debits with <em>negative</em> sign.</td>
<td>Credits and debits with <em>positive</em> sign.</td>
</tr>
<tr>
<td><strong>Financial account</strong></td>
<td><em>Increase</em> in assets and <em>decrease</em> in liabilities with <em>negative</em> sign.</td>
<td><em>Increase</em> in assets or liabilities with <em>positive</em> sign.</td>
</tr>
<tr>
<td></td>
<td><em>Decreases</em> in assets and <em>increase</em> in liabilities with <em>positive</em> sign.</td>
<td><em>Decrease</em> in assets or liabilities with <em>negative</em> sign.</td>
</tr>
<tr>
<td><strong>Financial account balance</strong></td>
<td>Change in assets <em>plus</em> change in liabilities (<em>negative</em> sign increases IIP)</td>
<td>Change in assets <em>minus</em> change in liabilities (<em>positive</em> sign increases IIP)</td>
</tr>
</tbody>
</table>

Source: Authors' illustration adapted from IMF (2007) “FAQs on Conversion from BPM5 to BPM6.”

- **Third, there are several different types (or “components”) of capital flows.** The main components are foreign direct investment (FDI), portfolio equity and debt flows, and other investment (which mainly captures banks’ deposit and lending transactions). These different types of capital flows differ greatly from each other in their empirical determinants and dynamic behavior (Koepke 2019). Yet, the umbrella term “capital flows” is commonly used to refer to a particular type of capital flows, notably the more volatile portfolio and bank-related flows. Moreover, the term “capital flows” is sometimes used loosely to refer to net purchases of investment fund shares (more appropriately referred to as “fund flows”).

**Measurement Limitations: Residency vs. Nationality Based Recording**

The BoP records transactions based on legal residence, which has limitations in today’s integrated global economy. Many firms and investors operate across national boundaries, not just in their country of legal residence. Recent research finds that two important drivers of the growing disconnect between legal residence and economic exposure are capital market access and tax considerations (Bertaut et al. 2019; Lane and Milesi-Ferretti 2018). Common examples of issues that are not appropriately captured in the residency-based BoP framework include:

---

15 For example, the paper entitled “Capital Flows are Fickle: Anytime, Anywhere” (Bluedorn et al. 2013) mainly focuses on volatile components of capital flows. Similarly, a recent paper on “Capital Flows at Risk” (Gelos et al. 2019) focuses on portfolio equity and debt flows. The use of the term “capital flows” is thus similar to the use of the word “America” for the United States of America (a stylistic device referred to as *totum pro parte*, the whole for a part).
• **Corporate offshore borrowing:** Multinational corporations may borrow outside their main domicile and recycle funds through various channels (Avdjiev et al. 2014; Lane and Milesi-Ferretti 2018).

• **Corporate tax optimization:** Foreign direct investment to tax havens is driven by tax optimization strategies of multinational corporations (Damgaard et al. 2019; Bertaut et al. 2019; Coppola et al. 2020).

• **Portfolio allocation via financial centers:** Investment firms in financial centers manage portfolio assets on behalf of ultimate investors in third countries (Lane and Milesi-Ferretti 2018).

• **Portfolio allocation via investment funds:** The assets of investment funds are attributed to the country where the fund is registered (Bertaut et al. 2019; IMF 2014). Moreover, under standard BoP accounting principles the holdings of investment fund shares are classified as equity investments (even for bond funds, for example).

Lane and Milesi-Ferretti (2018) highlight that intermediation through international financial centers obscures underlying capital movements in residency-based databases. Reporting on a nationality basis can complement residency-based reporting to capture underlying capital flow dynamics. Coppola et al. (2020) quantify this effect, finding that advanced economy (AE) holdings of emerging market and developing economy (EMDE) equities are around 50 percent higher in 2017 on a nationality basis than in the residency-based IMF Coordinated Portfolio Investment Survey (CPIS) and the Treasury International Capital System (TIC, see Figure 7). Bertaut et al. (2019) estimate specifically for the United States that nationality-based reporting for 2017 would reclassify around USD 3.5 trillion or 30 percent of total US cross-border portfolio investments.

For foreign direct investment statistics, Damgaard et al. (2019) estimate that FDI in special purpose entities, predominantly located in financial centers, account for around 40 percent of global FDI. Reassigning the remaining “real” direct investments on a nationality basis, outward foreign direct investment stocks from AEs to EMDEs decrease by around USD 1.3 trillion, or around 7 percent, while outward investment stocks roughly double from

---

16 The study includes the following countries with high-quality fund holding data: Australia, Canada, Denmark, European Monetary Union, Norway, Sweden, Switzerland, the United States, and the United Kingdom.
USD 1.5 trillion to around USD 3.1 trillion for EMDEs compared to the immediate counterparty reporting in the IMF's Coordinated Direct Investment Survey and the authors' estimates.

### III. CAPITAL FLOW DATA IN THE LITERATURE

#### A. Meta-Study of Data Sources Used in the Empirical Literature

In this section, we conduct a meta-analysis of the use of different capital flow datasets in the empirical literature. We build on the work of Koepke (2019), which surveys 34 empirical studies to analyze the empirical drivers of capital flows to emerging markets. We extend this sample by including an additional 54 studies, focusing on papers with a large number of citations and/or published in leading economic journals. We evaluate a total of 88 studies published over the course of the last 27 years and listed in Annex III.17

Sixty-three studies in our sample (or about 70 percent) were published after 2007, likely reflecting increased scholarly interest in capital flow issues as a result of the global financial crisis (GFC) and the ensuing policy challenges. A breakdown by data frequency shows that data used up until the GFC was predominantly at the annual frequency (around 50 percent of studies published between 1993 to 2007), followed by a shift towards much greater use of higher data frequencies in studies published since 2008. Studies using annual data only account for around 20 percent of papers in our sample in the 2008 to 2019 period, with quarterly data around 50 percent and monthly or higher data frequencies around 25 percent.18

![Figure 8. Data Frequency in Meta-Dataset 1/](image)

![Figure 9. Capital Flow Type in Meta-Dataset 1/](image)

Source: IMF staff calculations.
1/ Shaded bars show extrapolation for future years by scaling the number of papers written from 2018 until 2019H1 to the entire five-year period. Totals per time period are not equal in both charts as information for capital component type or data frequency are unavailable or ambiguous for a limited number of papers.

---

17 Our sample begins in 1993 with the seminal study by Calvo, Leiderman and Reinhart. In our selection of papers, we give preference to studies with 20+ citations and those published in the top two quartiles of economic journals per the ranking by SCImago Institutions Rankings. For recently authored studies, we err on the side of including additional studies because citations will tend to increase with time.

18 Information on data frequency or data source was unavailable or ambiguous in some papers such that the percentages do not sum to 100 percent.
The growing focus on high-frequency data is mirrored in greater use of portfolio flow data, given that for many countries FDI or bank flow data are not available at frequencies greater than quarterly data (Figure 8. and Figure 9. above). One benefit of high-frequency data for empirical research is to facilitate event studies (e.g., Fratzscher 2012; Fratzscher et al. 2012). Among the papers with monthly, weekly, or daily data, the majority use fund flow data, notably EPFR. There are only a few studies that make use of BoP-consistent portfolio flows at a high data frequency (e.g., Converse 2018; Koepke 2018; Cerdeiro and Komaromi 2019; Figure 10. below19).

The dominance of fund flow data in recent academic research raises an important concern about the validity of academic findings for “true” capital flow data. For example, recent research has emphasized the importance of external (“push”) factors for portfolio flows (e.g., Feroli et al. 2014; Friedrich and Guerin 2019). However, fund flow data by construction are more affected by push factors, given the role of investment benchmarks in the portfolio allocation process (see, for example, Raddatz et al. 2017; Arslanalp et al. 2020). Therefore, using fund flows to examine the role of push versus pull factors is likely to bias empirical findings towards greater importance of push factors.

The use of different data frequencies itself may also affect the estimated statistical significance and economic importance of push versus pull factors. For example, Ananchotikul and Zhang (2014) find that at short time horizons, push factors play a dominant role in driving fund flows but at longer time horizons pull factors become more important as weekly fluctuations in push factors cancel out.20

We investigate this issue further by classifying papers in our sample according to whether they study push factors, pull factors, or both (Figure 11). We find that studies using high-frequency data generally tend to focus more on external factors while studies using lower frequency data tend to focus more on domestic factors. This pattern is consistent with the idea that the relative importance of push vs. pull factors may be partly determined by the data frequency.21

---

19 By BoP consistency we refer to data that broadly follows balance of payments accounting principles.

20 One implication of this effect could be that other capital flow components may also be more subject to common external forces at short time horizons. In the absence of high frequency data on bank flows and FDI, it is not clear whether this hypothesis can be tested empirically.

21 In contrast to other capital flow components, bank flows have rarely been studied at the annual frequency, given that the BIS provides quarterly data, which provides the benefit of a greater number of observations for quantitative analysis.
B. Constructing a Monthly Portfolio Flow Dataset Consistent with BoP Principles

To help address the potential biases stemming from the use of BoP-inconsistent portfolio flow proxies, we construct a dataset on monthly emerging market portfolio flows that is broadly consistent with balance of payments accounting principles and specifically geared towards academic use. The dataset is similar to the monthly data on country-level portfolio flows compiled by the IIF, but has the advantage that it is available online for free (whereas IIF data are not accessible to most academics because they are restricted to members, which are mostly financial institutions). The dataset is posted online along with this paper and will be updated periodically (Link).

In the rest of this paper, we refer to this dataset as the KP dataset (short for Koepke & Paetzold). The KP dataset is constructed using data from national sources in a set of 18 EMs, with data for total portfolio flows as well as debt and equity portfolio flows. Data availability begins in 2010 for most countries. Figure 12 shows total monthly equity and debt flows, respectively.

---

22 In particular, the data are based on transactions between residents and non-residents of emerging market countries. The data are not subject to valuation effects and reflect transactions at market prices.

23 The countries included are the Brazil, Bulgaria, Chile, Czech Republic, Hungary, India, Lebanon, Mexico, Pakistan, Philippines, Poland, Romania, South Africa, Korea, Sri Lanka, Thailand, Turkey, Ukraine. Of these countries, the Czech Republic and Korea are not part of the IMF’s classification of emerging markets, but are included in private sector classifications of EMs such as leading investment benchmark indices. Data for Bulgaria, Hungary, India, Romania, and South Africa were converted to USD using monthly average exchange rates.
Figure 12. Total Monthly Equity and Debt Flows from KP Dataset
(in USD billion, Jan. 2010 to Aug. 2019)

Source: KP dataset.

Of the 18 countries, 13 report monthly data that correspond to the quarterly balance of payments published by the central bank or statistical authority. For five countries, we collect monthly proxies for one or more portfolio flow type. Table 3 provides additional details on the data used in the KP dataset.

---

24 Among the proxy data, we use government debt flows for Mexico, non-government debt flows for Thailand, transactions on national stock exchanges by non-residents for South Africa and Sri Lanka and portfolio inflows by non-residents for India.
Table 3. Description of Data in the KP Monthly Portfolio Flow Dataset

<table>
<thead>
<tr>
<th>Country</th>
<th>Components</th>
<th>Data source</th>
<th>Release lag in months (approx.)</th>
<th>Currency</th>
<th>Proxy data</th>
<th>Correlation with total BoP flows</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td>Debt/Equity</td>
<td>Central Bank of Brazil</td>
<td>1-2</td>
<td>USD</td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>Debt/Equity</td>
<td>Eurostat</td>
<td>2-3</td>
<td>EUR</td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Chile</td>
<td>Debt/Equity</td>
<td>Central Bank of Chile</td>
<td>2-3</td>
<td>USD</td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Czech Republic*</td>
<td>Debt/Equity</td>
<td>Czech National Bank</td>
<td>2-3</td>
<td>EUR</td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Hungary</td>
<td>Debt/Equity</td>
<td>Eurostat</td>
<td>2-3</td>
<td>EUR</td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>India**</td>
<td>Debt/Equity</td>
<td>Securities &amp; Exchange Board of India</td>
<td>0-1</td>
<td>INR</td>
<td>Y</td>
<td>0.92</td>
</tr>
<tr>
<td>Korea*</td>
<td>Debt/Equity</td>
<td>Bank of Korea</td>
<td>1-2</td>
<td>USD</td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Lebanon</td>
<td>Debt/Equity</td>
<td>Bank of Lebanon</td>
<td>10-11</td>
<td>USD</td>
<td>Y</td>
<td>0.86</td>
</tr>
<tr>
<td>Mexico</td>
<td>Debt/Equity</td>
<td>Bank of Mexico</td>
<td>3-4</td>
<td>USD</td>
<td></td>
<td>0.98</td>
</tr>
<tr>
<td>Pakistan</td>
<td>Debt/Equity</td>
<td>State Bank of Pakistan</td>
<td>1-2</td>
<td>USD</td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Philippines</td>
<td>Debt/Equity</td>
<td>Central Bank of the Philippines</td>
<td>4-5</td>
<td>USD</td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Poland</td>
<td>Debt/Equity</td>
<td>National Bank of Poland</td>
<td>2-3</td>
<td>EUR</td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Romania</td>
<td>Debt/Equity</td>
<td>Eurostat</td>
<td>2-3</td>
<td>EUR</td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>South Africa</td>
<td>Debt/Equity</td>
<td>Johannesburg Stock Exchange</td>
<td>2-3</td>
<td>ZAF</td>
<td>Y</td>
<td>0.64</td>
</tr>
<tr>
<td>Sri Lanka</td>
<td>Debt/Equity</td>
<td>Colombo Stock Exchange</td>
<td>2-3</td>
<td>USD</td>
<td>Y</td>
<td>0.30</td>
</tr>
<tr>
<td>Thailand</td>
<td>Debt/Equity</td>
<td>Bank of Thailand</td>
<td>3-4</td>
<td>USD</td>
<td>Y</td>
<td>0.92</td>
</tr>
<tr>
<td>Turkey</td>
<td>Debt/Equity</td>
<td>Central Bank of Turkey</td>
<td>2-3</td>
<td>USD</td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Ukraine</td>
<td>Debt/Equity</td>
<td>National Bank of Ukraine</td>
<td>2-3</td>
<td>USD</td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

Sources: Authorities’ data, BOPS, and IMF.
* The Czech Republic and Korea are not part of the IMF’s classification of emerging market economies, but are included in private sector classifications of EMs such as leading investment benchmark indices.
** India’s portfolio flow data are recorded on a reporting day basis rather than on a trading day basis, which may contribute to a lower correlation between data used in the KP dataset and Balance of Payments data.

Figure 13 show total EM equity and debt flows, respectively, for the KP data and three relevant comparators (IIF monthly tracker, IIF daily, EPFR monthly). The sum of total flows to EMs in the KP dataset from 2010:Q1 to 2019:Q2 covers around 46 percent of total flows in the BOPS, with somewhat higher coverage of equity flows (53 percent) and lower coverage for debt flows (44 percent). Equity data from the KP dataset and the IIF’s monthly portfolio flow tracker are highly correlated with quarterly BoP-based equity flows (see Table 4). For the debt component, the KP data has a high correlation with BoP flows at the country level, but somewhat lower at an EM aggregate level.
Figure 13. Comparison Balance of Payments and EM Portfolio Flow Proxies (USD billion)


<table>
<thead>
<tr>
<th></th>
<th>Equity</th>
<th>Debt</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All EMs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IIF Monthly</td>
<td>0.86</td>
<td>IIF Monthly</td>
</tr>
<tr>
<td>EPFR Monthly</td>
<td>0.67</td>
<td>EPFR Monthly</td>
</tr>
<tr>
<td>KP</td>
<td>0.85</td>
<td>KP</td>
</tr>
<tr>
<td>IIF Daily</td>
<td>0.81</td>
<td>IIF Daily</td>
</tr>
<tr>
<td><strong>India</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IIF Monthly</td>
<td>0.95</td>
<td>IIF Monthly</td>
</tr>
<tr>
<td>EPFR Monthly</td>
<td>0.18</td>
<td>EPFR Monthly</td>
</tr>
<tr>
<td>KP</td>
<td>0.94</td>
<td>KP</td>
</tr>
<tr>
<td>IIF Daily</td>
<td>0.95</td>
<td>IIF Daily</td>
</tr>
<tr>
<td><strong>South Africa</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IIF Monthly</td>
<td>0.75</td>
<td>IIF Monthly</td>
</tr>
<tr>
<td>EPFR Monthly</td>
<td>0.25</td>
<td>EPFR Monthly</td>
</tr>
<tr>
<td>KP</td>
<td>0.74</td>
<td>KP</td>
</tr>
<tr>
<td>IIF Daily</td>
<td>0.75</td>
<td>IIF Daily</td>
</tr>
<tr>
<td><strong>Thailand</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IIF Monthly</td>
<td>0.80</td>
<td>IIF Monthly</td>
</tr>
<tr>
<td>EPFR Monthly</td>
<td>0.11</td>
<td>EPFR Monthly</td>
</tr>
<tr>
<td>KP</td>
<td>1.00</td>
<td>KP</td>
</tr>
<tr>
<td>IIF Daily</td>
<td>0.80</td>
<td>IIF Daily</td>
</tr>
<tr>
<td><strong>Turkey</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IIF Monthly</td>
<td>1.00</td>
<td>IIF Monthly</td>
</tr>
<tr>
<td>EPFR Monthly</td>
<td>0.54</td>
<td>EPFR Monthly</td>
</tr>
<tr>
<td>KP</td>
<td>1.00</td>
<td>KP</td>
</tr>
<tr>
<td>IIF Daily</td>
<td>0.99</td>
<td>IIF Daily</td>
</tr>
</tbody>
</table>

Source: IMF staff calculations.

Figures 14 and 15 illustrate how the KP data compare to balance of payments data. Average quarterly flows in 18 country cases, with the exception of South Africa, are equal or close to averages from the balance of payments. Co-movement for the majority of countries in our dataset with available data shows a standard deviation roughly in line with the standard deviation of balance of payments data, with the exception of Mexico, Sri Lanka, South
Africa, and Thailand. In comparison to investment fund data such as EPFR, the dataset provided in this paper more closely aligns with standard definitions in the BoP. Using this dataset may help researchers obtain empirical findings that are valid for “true” capital flow data, consistent with established BoP accounting principles.

**Figure 14. Comparison KP Dataset and Balance of Payments Portfolio Flows**

![Graph showing comparison between KP dataset and balance of payments portfolio flows.]

**Figure 15. Comparison EPFR and Balance of Payments Portfolio Flows**

![Graph showing comparison between EPFR and balance of payments portfolio flows.]

Source: IMF staff calculations.

1/ The large square shows the area where both variance and average of flows in the dataset and EPFR data are approximately equal to variance and average of flows in the balance of payments. Small squares with a cross inside show data for the EM aggregate.

Source: IMF BOPS, EPFR.

---

25 Potential reasons for these discrepancies could be data revisions as well as the use of proxy data and incomplete data for capital flows (i.e., stock exchange data in the case of equity flows for Sri Lanka and South Africa, the exclusion of government debt flows for Thailand, and the exclusion of private sector debt flows for Mexico and Sri Lanka).
IV. Real-Time Tracking of Portfolio Flows

In many emerging market countries, capital flows and exchange rate pressures have long been a key factor driving central banks’ policy decisions (Calvo and Reinhart 2002; Vegh and Vuletin 2012). Policymakers need timely information on capital flows to inform decisions on monetary policy and foreign exchange intervention. Official BoP data are of little help, given their long release lag. Instead, policymakers often rely on a combination of national data on foreign exchange transactions involving their own currencies and private sector data providers on portfolio flows, notably EPFR and IIF data.

In order to assess which portfolio flow proxies are best suited for real-time tracking, we conduct a quantitative assessment of how well EPFR and IIF data predict official BoP data. We run a nowcasting “horse race,” in which quarterly BoP portfolio flow data are predicted using available portfolio flow proxies as of a given point in time. To our knowledge, this is the first systematic assessment of the predictive content of commonly used portfolio flow proxies.

More broadly speaking, there is very limited literature discussing methods to track capital flows in real time. One important study is Miao and Pant (2012), which highlights the serious limitations of traditional capital flow data for surveillance and policymaking purposes. The authors propose two methods to address this data gap. One focuses on net capital flows (i.e., the financial account balance), which is estimated using timely data on trade in goods and changes in reserves, augmented with several global and EM financial variables (such as the VIX, US bond prices, and EM stock prices). In the other method, an error correction model is applied to EPFR data and several financial variables to obtain an estimate of gross portfolio flows to EM regions. However, both methods are still subject to considerable time lags, and the study does not assess predictive performance relative to other proxies.

Nowcasting Horse Race

We use several alternative portfolio flow proxies in a nowcasting setting with quarterly portfolio flows from the BoP as the dependent variable, where the accuracy of nowcasts is measured as the root mean squared forecast error (RMSFE). The portfolio flow proxies include:

- Weekly EPFR data, published with a lag of about 7 days.
- Monthly EPFR data, published with a lag of about 15 days.
- Daily IIF data, published with a lag of 1-3 days.
- Monthly IIF data, published with a lag of about 1-5 days.

---

26 In this section, the terms “nowcast,” “forecast,” and “prediction” are used interchangeably.
27 As noted in Miao and Pant (2012), variants of this approach have been used by several studies to obtain timely proxies of capital flow movements, including Calvo et al. (2004 and 2008). One limitation of this approach is that trade in services, primary income, and secondary income play major roles in some EMs, causing errors in the estimation. The IIF provides monthly estimates of net capital flows for 23 EMs, using the same basic approach.
We also include the monthly data from the KP dataset, described in the prior section. The underlying data are typically available after about 2-3 months.

The sample period for the nowcasting exercise runs from 2010:Q1 until 2019:Q2. In our baseline model, we use aggregate flows for those emerging market countries where all the portfolio flow proxies are available (separately for equity and debt flows).\(^{28}\) Aggregate flows to these countries are scaled by their combined GDP. We estimate a pseudo-out-of-sample forecast error, using only data that would have been available as of the relevant quarter. Therefore, the number of observations used in the regression is quite small at the beginning of the sample period and increases over time.

In the first step, observations for each portfolio flow proxy are summed up incrementally for each quarter. For example, daily capital flows are summed for days 1-2, days 1-3, days 1-4, and so forth, up until the last observation of the quarter is included. Similarly, weekly/monthly data are summed over an increasing number of weeks/months in each quarter.

In the second step, for each of these sums the following model is estimated:

\[
\text{BoP}_{\text{quad}} = c + \beta_{ij} \cdot \text{capflow\_sum}_{ij}
\]

Where \(\text{BoP}_{\text{quad}}\) is quarterly portfolio flow data from the balance of payments, \(c\) is a constant term, and \(\text{capflow\_sum}_{ij}\) is the sum of the high-frequency proxies for portfolio flows from the first observation to the \(i\)-th observation of the quarter, and \(j\) represents the portfolio flow proxy used.

This model is used to predict quarterly BoP portfolio flows using available data from 2010:Q1 onwards. For example, to predict the quarterly BoP flows in 2012:Q1, the model is estimated using the observations for 2010:Q1-2011:Q4, while for the prediction for 2012:Q2, the model is estimated using the observations for 2010:Q1-2012:Q1. For each quarter, \(i\) nowcasts are produced by multiplying the estimated coefficient from the regression and \(\text{capflow\_sum}_{ij}\) for that quarter (and adding the constant term).

We evaluate the predictive content of each portfolio flow proxy by computing the root-mean squared forecast error for each \(i\), for the period from 2012:Q1 until 2019:Q2:

\[
\text{RMSFE}_{ij} = \sqrt{\frac{1}{T} \sum_{t}(y_{tij} - \hat{y}_{tij})^2}
\]

Where RMSFE is the root-mean squared forecast error for sum \(i\) and proxy \(j\), \(T\) is the number of quarters for which a forecast (or “nowcast”) is produced, \(t\) is the time in quarters, \(y\) is actual portfolio flows for that quarter, and \(\hat{y}\) is the forecast for that quarter.

\(^{28}\) For equity flows, the sample contains Brazil, India, Pakistan, Philippines, South Africa, Korea, Sri Lanka, Thailand, and Turkey. For debt flows, the sample contains Hungary, India, Poland, South Africa, Thailand and Turkey.
The results of this exercise are shown in Figure 16 for equity flows and in Figure 17 for debt flows. As expected, forecast errors for all proxies generally decline over the course of the quarter as more data become available. Daily IIF data and weekly EPFR data are available as early as 2 and 7 days into the quarter, much earlier than either of the monthly proxies. RMSFEs based on the first few observations of daily and weekly proxies are relatively high, but the forecast performance of the daily IIF data improves notably once 20-30 days’ worth of data are included.

The monthly proxies first become available about 33 and 45 days into the quarter for IIF and EPFR data, respectively. The IIF’s monthly tracker consistently performs better than either of the EPFR datasets (weekly or monthly), but only outperforms the IIF daily flows data after the second data release in the quarter (on average). Similarly, EPFR’s monthly data mostly underperforms the weekly data, with the exception of debt flows towards the end of the quarter. We also compare the RMSFEs of the various proxies to that of a benchmark autoregressive regression with one lag and a constant term. All proxies outperform this benchmark by a wide margin, with all RMSFEs at least 80 percent lower.

The improvements in forecast performance of quarterly capital flows using IIF compared to EPFR data are economically important. Expressed as a share of absolute average quarterly flows, RMSFEs of debt and equity flows based on daily and monthly IIF data outperform weekly and monthly EPFR data, respectively, by around 20 percent early in the quarter and 50 percent at the end of the quarter. Compared to the benchmark regression, RMSFEs of IIF
flows as a share of quarterly flows are consistently lower while EPFR data only achieve marginally lower RMSFEs at the end of the quarter.

The results also show that the KP dataset has a similar performance to the IIF’s monthly tracker, although the monthly data from countries contained in the KP dataset is subject to a greater release lag within the current quarter. One caveat regarding the IIF’s monthly EM tracker and the KP dataset is that observations for the latest 1-2 months of data are revised with each monthly data release for the EM aggregate estimate and for more than half of the countries included in the country level flows. This may bias our results towards a smaller forecast error for these datasets. The other data sources generally do not get revised.

The main findings from the nowcasting horse race are that (1) all portfolio flow proxies considered have significant predictive content for BoP portfolio flows, (2) IIF proxies generally outperform EPFR proxies, and (3) higher frequency proxies generally outperform lower-frequency proxies, especially in the first half of the “current” quarter.

We test the robustness of these findings through a range of alternative specifications. For example, we conduct the same exercise using a longer sample period that begins in 2005:Q1, which did not substantially change average forecast errors. In an extension inspired by the coincident indicators proposed by Miao and Pant (2012), we include averages from the first to the i-th quarterly observation of global financial variables used in the literature on capital flow drivers (VIX, US Treasury yields) as additional predictors. The inclusion of these variables does not consistently reduce the RMSFEs and the relative magnitude of forecast errors does not change substantively.

We also run the nowcasting exercise using aggregate EM flows from each source (for equity flows, debt flows, and the sum of equity and debt flows), as well as for individual EM countries where data are available from all three data sources. The results are presented in Annex IV. For the EM aggregate exercise, the monthly proxy data often perform better than in our baseline results. Notably, the IIF’s monthly tracker generally outperforms other indicators for both total flows and equity flows, while EPFR tends to outperform for debt flows.

V. Conclusion

In this paper, we provide an overview of the main capital flow data sources used in the empirical literature and among policymakers. IMF balance of payments statistics remain a key data source and serve as an important benchmark for assessing the properties of other datasets. Data availability has improved significantly in recent years, especially with respect to high frequency proxies for portfolio flows at the daily, weekly, and monthly frequencies. We find that fund flow data have been used by a sizeable share of empirical studies (around 14 percent) over the past decade, allowing researchers to gain deeper insight into the drivers of rapid shifts in investor behavior.

This paper also sheds light on the conceptual and empirical differences between fund flow data and portfolio flows as measured in the IMF’s balance of payments statistics. These differences highlight that findings in the empirical literature are likely shaped by unique
properties of fund flow data. A greater focus on high frequency data that are consistent with balance of payments accounting principles seems warranted but may have been hampered by data subscription requirements and limitations of existing datasets.

We help address this gap by providing a monthly portfolio flow dataset designed specifically for academic use. The dataset covers 18 of the largest emerging market economies and tracks the quarterly balance of payments data of these countries closely. The dataset may be particularly useful for analyzing rapid shifts in portfolio flows, such as those witnessed during the COVID-19 pandemic. The dataset may also help facilitate future research investigating whether the role of external drivers (relative to domestic factors) is as important as suggested by papers using fund flow data (which are by construction subject to common factors).

Finally, we assess the predictive content of various portfolio flow proxies in a nowcasting “horse race.” The results are encouraging in that all portfolio flow proxies have significant predictive content for balance of payments portfolio flows. Among the various predictors, IIF portfolio flow trackers generally outperform EPFR fund flow data, and higher frequency proxies generally outperform lower-frequency data, especially in the first half of the current quarter.

Overall, this paper highlights the importance for researchers and policymakers to understand in detail the capital flow data and proxies they rely upon. This will help economists utilize the best-suited dataset depending on the question that is asked, and will help inform how the answers are framed. On this basis, the growing number of capital flow datasets provide valuable opportunities for advancing economic research and informing policy decisions in real time.
References


https://www.iif.com/Portals/0/Files/portfolio_flows_tracker_4.0_methodology_note.pdf.


Annex I. Bilateral Capital Flows Data ("From-Whom-to-Whom" Data)

The global financial crisis exposed the need for detailed statistics of cross-border financial linkages (FSB and IMF 2009). For example, research showed that global banking flows contributed to the spread of cross-border funding shocks in the global financial crisis (Cetorelli and Goldberg 2011). Specific recommendations by the IMF and FSB called on the G20 members to strengthen participation in the Coordinated Portfolio Investment Survey and Coordinated Direct Investment Survey (CPIS, CDIS) carried out by the IMF and the International Banking Statistics maintained by the Bank for International Settlements.

Data availability on bilateral portfolio and direct investment has improved in recent years. The CPIS contains annual data from 2001 and semi-annual data from 2013. The data are published with a 9-month lag. In 2018, all G20 members reported semi-annual data on bilateral portfolio investment stocks through the CPIS with overall coverage at 73 reporting countries and territories. Portfolio investment liability positions for participating and non-participating countries and territories are derived from portfolio investment asset positions reported by counterpart countries and territories. The CDIS begins with annual data in 2009. Data are released with a lag of approximately 12 months and 110 countries and territories reported data at the end of 2018. Even for countries that do not participate, the CDIS derives data using all participating countries’ outward direct investment positions in country A as a proxy for country A’s inward direct investment position.¹

The BIS’s International Banking Statistics provide the most comprehensive bilateral stock and flow data of cross-border banking flows. The quarterly data are released with a relatively short lag of 4 months or less. With around 48 central banks or monetary authorities reporting data, the data capture around 95 percent of global banking flows (BIS 2019).

Some national statistics providers and central banks publish bilateral capital flow and stock data at higher frequencies (see Table A1.1). While countries like the United States publish monthly portfolio transactions data, and Japan, Germany, and Canada publish quarterly data, most G7 or G20 economies do not publicly provide data on bilateral portfolio flows. While data availability for bilateral direct investment stocks or flows is more common at quarterly frequencies, more than half of G20 members do not publicly share this data.

¹ Country A’s outward direct investment position is also derived using all participating countries’ inward direct investment positions from country A as a proxy for country A’s outward direct investment position.
Table A1.1. Overview of Bilateral Portfolio and Direct Investment Data

<table>
<thead>
<tr>
<th>Country</th>
<th>Portfolio Investment</th>
<th>Direct Investment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>G7</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>Monthly flows data</td>
<td>Quarterly flows data; Annual stock data on ultimate beneficial owner basis</td>
</tr>
<tr>
<td>France</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Germany</td>
<td>Quarterly flows data</td>
<td>Quarterly flows data</td>
</tr>
<tr>
<td>Italy</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Japan</td>
<td>Quarterly flows data</td>
<td>Quarterly flows data</td>
</tr>
<tr>
<td>United States</td>
<td>Monthly flows and stock data</td>
<td>Quarterly flows data and annual stock and flows data on ultimate beneficial owner basis</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>G20</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Argentina</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Australia</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Brazil</td>
<td>N/A</td>
<td>Annual stock and flows data on ultimate beneficial owner basis</td>
</tr>
<tr>
<td>China</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>India</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Indonesia</td>
<td>N/A</td>
<td>Quarterly flows data</td>
</tr>
<tr>
<td>Mexico</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Russia</td>
<td>Quarterly stock data</td>
<td>Quarterly stock and flows data</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Korea</td>
<td>Annual flows data</td>
<td>Annual flows data</td>
</tr>
<tr>
<td>South Africa</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Turkey</td>
<td>Quarterly stock data (assets only)</td>
<td>Monthly stock and flows data</td>
</tr>
<tr>
<td>EU</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Other countries</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>Quarterly flows data including exchange rate/price adjustments</td>
<td>Annual stock and flows data on ultimate beneficial owner basis</td>
</tr>
<tr>
<td>Denmark</td>
<td>N/A</td>
<td>Quarterly flows data</td>
</tr>
<tr>
<td>Netherlands</td>
<td>N/A</td>
<td>Quarterly flows data</td>
</tr>
<tr>
<td>Spain</td>
<td>Quarterly stock data</td>
<td>Annual flows data</td>
</tr>
</tbody>
</table>

Source: Central banks or statistical agencies.
Annex II. Daily IIF Portfolio Flows Data

Table A2.1. Data Availability for Daily IIF Flows Data as of December 2, 2015 (Update to country coverage and data documentation ongoing)

<table>
<thead>
<tr>
<th>High Frequency Portfolio Flows: Data Sources</th>
<th>Frequency (Log in # of business days)</th>
<th>Currency Reported in</th>
<th>Reported as Stock or Flow Data (See FAQ #6)</th>
<th>% of EM Equity or Debt Portfolio Flows (2010-2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Equity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td>Daily (2)</td>
<td>BRL</td>
<td>Flow</td>
<td>12.0</td>
</tr>
<tr>
<td>India</td>
<td>Daily (1)</td>
<td>USD</td>
<td>Flow</td>
<td>14.7</td>
</tr>
<tr>
<td>Indonesia</td>
<td>Daily (0)</td>
<td>USD</td>
<td>Flow</td>
<td>2.8</td>
</tr>
<tr>
<td>Korea</td>
<td>Daily (0)</td>
<td>USD</td>
<td>Flow</td>
<td>15.0</td>
</tr>
<tr>
<td>South Africa</td>
<td>Daily (1)</td>
<td>USD</td>
<td>Flow</td>
<td>3.0</td>
</tr>
<tr>
<td>Thailand</td>
<td>Daily (0)</td>
<td>USD</td>
<td>Flow</td>
<td>3.0</td>
</tr>
<tr>
<td>Turkey</td>
<td>Weekly (5)</td>
<td>USD</td>
<td>Flow</td>
<td>2.9</td>
</tr>
<tr>
<td><strong>Debt</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungary</td>
<td>Daily (2)</td>
<td>HUF</td>
<td>Stock</td>
<td>2.7</td>
</tr>
<tr>
<td>India</td>
<td>Daily (1)</td>
<td>USD</td>
<td>Flow</td>
<td>5.4</td>
</tr>
<tr>
<td>Indonesia</td>
<td>Daily (1)</td>
<td>IDR</td>
<td>Stock</td>
<td>6.2</td>
</tr>
<tr>
<td>Mexico</td>
<td>Daily (7)</td>
<td>MXN</td>
<td>Stock</td>
<td>19.9</td>
</tr>
<tr>
<td>South Africa</td>
<td>Daily (1)</td>
<td>USD</td>
<td>Flow</td>
<td>4.4</td>
</tr>
<tr>
<td>Thailand</td>
<td>Daily (0)</td>
<td>USD</td>
<td>Flow</td>
<td>3.0</td>
</tr>
<tr>
<td>Turkey</td>
<td>Weekly (5)</td>
<td>USD</td>
<td>Flow</td>
<td>9.2</td>
</tr>
</tbody>
</table>


Table A2.2. Coverage of Daily Debt flows

<table>
<thead>
<tr>
<th>Coverage of Daily Debt flows</th>
<th>Currency denomination of bonds (Local=LC; Hard=HC)</th>
<th>Sectoral Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hungary</td>
<td>LC</td>
<td>Government</td>
</tr>
<tr>
<td>India</td>
<td>LC + HC</td>
<td>Government + Corporate</td>
</tr>
<tr>
<td>Indonesia</td>
<td>LC + HC</td>
<td>Government</td>
</tr>
<tr>
<td>Mexico</td>
<td>LC</td>
<td>Government</td>
</tr>
<tr>
<td>South Africa</td>
<td>LC</td>
<td>Government + Corporate*</td>
</tr>
<tr>
<td>Thailand</td>
<td>LC + HC</td>
<td>Government + Corporate</td>
</tr>
<tr>
<td>Turkey</td>
<td>LC</td>
<td>Government</td>
</tr>
</tbody>
</table>

Annex III. List of Papers Included in the Academic Dataset


https://doi.org/10.1093/wber/10.1.27.

https://doi.org/10.1016/0304-3878(95)00041-0.


https://doi.org/10.5089/9781451978827.001.


Annex IV. Portfolio Flow Nowcasting Results by Country

Figure A4.1. Root mean squared forecasting error (RMSFE) of total flows and broken down in equity and debt flows (in % of GDP)

Sources: BOPS, EPFR, IIF, IMF.
Figure A4.2. Average root mean squared forecasting error (RMSFE) of equity and debt flows by country (in % of GDP)

**Turkey 1/**

**South Africa 1/**

**India 2/**

**Thailand 2/**

Sources: BOPS, EPFR, IIF, IMF.
1/ Sample period starting in 2005. 2/ Sample period starting in 2010.