DNB Working Paper
No 757 / December 2022

A literature review of securities holdings statistics research and a practitioner's guide

Martijn Boermans

DeNederlandscheBank
EUROSYSTEEM

A literature review of securities holdings statistics research and a practitioner's guide

Martijn Boermans *

* Views expressed are those of the author and do not necessarily reflect official positions of De Nederlandsche Bank.

Working Paper No. 757

December 2022

De Nederlandsche Bank NV
P.O. Box 98
1000 AB AMSTERDAM
The Netherlands

A literature review of securities holdings statistics research and a practitioner's guide *

Martijn A. Boermans†

Updated version: January 2025

First version: December 2022

Abstract

Granular holdings data containing security-by-security portfolio investments features prominently in economics and finance research. One novel source is the granular Securities Holdings Statistics (SHS), managed by the European Central Bank. SHS covers different euro area investors with over 2 billion observations, representing +50 trillion euros in portfolio investments. This comprehensive SHS research review of 102 studies (31 published and 71 working papers) highlights the strong growth in SHS-based publications, especially since 2022. The review details how granular SHS data has catalyzed research across five fields: banking and finance, international investment, monetary policy, financial markets, and sustainable finance. Addressing the replication crisis in economics, this study emphasizes the need for rigorous data filtering, cleaning and documentation of aggregation choices when using SHS data. A practitioner's guide is provided to ease future research, enhancing replicability and to minimize non-standard errors, showcasing best practices through multidimensional time-series panel analyses of the holdings of different investors and their preferences for euro-denominated bonds, green bonds, and sustainability-linked bonds. The review concludes by suggesting potential avenues for future research.

Keywords — Securities holdings statistics, portfolio investment, literature review, Eurosystem data, ECB, home currency bias, green bonds, investor heterogeneity, replication crisis.

JEL codes — E52, E58, F14, F3, G11, G2, G51, Q56.

*I thank Maurice Bun, Johannes Breckenfelder, Daniel Carvalho, Daniel Fricke, Iman van Lelyveld, Matthias Sydow, Robert Vermeulen and participants of the DNB SHS workshop (2021), DNB research seminar (2022) and the World Statistics Congress (2022) for input. Views expressed are those of the author and do not necessarily reflect official positions of De Nederlandsche Bank or the Eurosystem. Data have been cleared by the Eurosystem for non-disclosure of confidential data. A database with detailed documentation on the literature and cleaning code (in Python and Stata) can be found on the DNB website: https://www.dnb.nl/en/research/personal-pages/martijn-boermans/shs-research-database.

†De Nederlandsche Bank. Email: m.a.boermans@dnb.nl

I. Introduction

Granular holdings data consisting of security-by-security portfolio investment details play a pivotal role in economics and finance research. Notably, focusing on US investments, many researchers have used line-by-line information on shareholdings from large investors through 13-F filings to the Securities and Exchange Commission (SEC) since 1999. These filings cover mainly US mutual funds' portfolios quarterly since 2004 (see e.g. Edison and Warnock, 2004; Agarwal et al., 2013; Aragon et al., 2013; Parida and Teo, 2018). Additionally, for US bonds, the Trade Reporting and Compliance Engine (TRACE) database records transactions at the bond level by various investors since 2002 (see Dick-Nielsen, 2014, for a review). Internationally, research has relied on portfolio holdings data mostly from individual mutual funds (see Cremers et al., 2019, for an overview) through commercial sources.[1]

This article covers a novel granular holdings source, where granular data is defined as security-by-security information on investment portfolios, where the disaggregated nature of the data implies one observes very precisely the behavior of investors and allows for more detailed analysis. Specifically, over the past decade, Securities Holdings Statistics (SHS) has catalyzed research by utilizing granular portfolio holdings primarily from euro area investors ECB (2015). SHS, managed by the European Central Bank, provide high-quality security-by-security information on portfolios, offering extensive coverage across different investor countries and sectors. SHS rely on harmonized reporting in accordance with Regulation (EU) No. 1011/2012. The database encompasses over 2 billion observations, representing over 50 trillion euros of euro area holdings at market value as of 2024-Q1, spanning stocks (15 trillion euros), funds (15 trillion euros), bonds (22 trillion euros), and money market paper (2 trillion euros).

This review contributes to the broad literature using portfolio holdings data. First, we outline a rich and growing set of papers using the Securities Holdings Statistics. We retrieve and review 102 journal articles and working papers that have been using SHS data and group these studies across five different fields. SHS research has filled the data gaps in (i) banking and finance, mostly related to interconnectedness, contagion and risky behavior of financial institutions (Anand et al., 2018; Boermans and Van Wijnbergen, 2018; Hüser et al., 2018; Bubeck et al., 2020; Tzamourani, 2021; Aldasoro et al., 2022; Jourde, 2022; Maddaloni and Scardozzi, 2022; Barbiero et al., 2024; Boermans and van der Kroft, 2024; Marques-Ibanez et al., 2024; Sydow

et al., 2024b), (ii) the international investment literature focusing on portfolio reallocations of different investors (Boermans et al., 2016; Boermans and Vermeulen, 2020; de Haan and Vermeulen, 2021; Carvalho, 2022; Boermans and Burger, 2023; Jansen, 2024), (iii) monetary policy research, analyzing its effectiveness and transmission channels, in particular focusing on the asset purchases program in the euro area (Arrata et al., 2020; Albertazzi et al., 2021; Koijen et al., 2021; Eser et al., 2023; Kabaca et al., 2023), (iv) financial market research, where security level data has provided novel testing grounds for theory (Boermans et al., 2016; Darmouni and Papoutsi, 2022; Kliatskova et al., 2023; Fricke et al., 2024), and (v) sustainable finance by examining the impact of climate change on portfolio allocations (Alessi and Battiston, 2022; Boermans, 2023; Alessi et al., 2024; Boermans and Galema, forthcoming). Several of these SHS papers have been published in top field journals, including *Journal of International Economics*, *Journal of Finance*, *Journal of Financial Economics* and *Review of Financial Studies*.

Second, our review suggests that SHS research was in a nascent stage between 2018 and 2021, entering a stage of growth in 2022 as witnessed by an exponential upward trend in the number of articles using granular SHS data: since 2022, 18 journal publications and 52 working papers have been published (up to the end of 2024), highlighting the strong growth in granular SHS research and the greater need for researchers to better appreciate the SHS data, which this review aim to do.

Third, against the backdrop of a replication crisis in economics and finance, reviewing the literature using granular SHS data highlights that best practices are still being developed when handling this raw data source. We demonstrate that most of the journal articles do not apply any detailed filtering to the raw dataset or neglect to report important information regarding data selection and aggregations. In a practitioner's guide in Section V we aim to guide users of SHS data when *conceptualizing* data filters, aggregations and setting up baseline (regression) analyses in a time-series panel with mostly unbalanced and high data skewness. We contribute by showing that without motivating and documenting essential data cleaning steps, estimates may strongly diverge. In an analytical exercise on investor sector heterogeneity, currency preferences and green bond and sustainability-linked bond holdings we show large differences between estimates with the raw data and the cleaned data.

Finally, we contribute by distilling a roadmap for future research using granular SHS data in

Section VI. Merging portfolio holdings with other datasets has initiated a kaleidoscopic view in different research agendas.

This literature study is structured as follows. Section II presents background information on SHS data. Section III explains how articles were selected for this literature review. Section IV provides an overview of studies with SHS data. Section V presents an overview of data preparation, filtering and cleaning steps done in the literature while arguing for a more coherent approach based on an illustration of a holdings regression with over 20 million observations on investor heterogeneity, home currency preferences and sustainability. Section VI distills a roadmap for future research. Section VII concludes.

II. An introduction to the ECB Securities Holdings Statistics (SHS)

Before the SHS, policymakers and researchers primarily used macroeconomic data like the IMF Coordinated Portfolio Investment Surveys (CPIS) to analyze portfolio holdings of investors across countries. The 2008-2009 financial crisis highlighted the need for detailed portfolio data to assess financial institutions' exposures. In response, the European Central Bank (ECB) established a legal framework in October 2012 to harmonize securities holdings statistics collection at the security level in the euro area. SHS covers money market paper, bonds, listed stocks, and investment funds, available since 2013-Q4 and updated quarterly. SHS data is available in two modules: SHS-Sectoral (SHS-S) and SHS-Group. SHS-S provides security-by-security holdings and transactions aggregated by investor sector and country, while SHS-Group covers over 120 individual large banking groups under the European Single Supervisory Mechanism (SSM), representing a total of 4 trillion euro of European bank investment positions by 2024-Q1.

The SHS-Sectoral is most widely used. This SHS-S modules includes granular holdings of securities by investor sector and investor country resident in the euro area, e.g. Dutch households or French insurance corporations. By 2024-Q1 total holdings by euro area investors were more than 50 trillion euro in the SHS-S data, covering over 1,000 million observations over the period 2013Q4-2024Q1. High coverage of total euro area portfolios across different investor sectors are guaranteed by the SHS legal framework. In addition, the SHS-S module includes partial and incomplete coverage for non-resident investors who deposited their holdings with a euro area custodian, e.g. Brazilian households living Portugal who deposit their portfolio at a Portuguese

custodian. Moreover, most non-euro area EU countries also started providing SHS-S data, including Bulgaria, the Czech Republic, Denmark, Hungary, Poland and Romania.

The granular SHS data allows for combining portfolio flows with other sources, offering flexibility in analyzing portfolio investments. The vast majority of securities held have an International Securities Identification Number (ISIN), which facilitates comparability across reporting agents and linking with other reference data (ECB, 2015). The ECB merged SHS with the Centralised Securities Database (CSDB), providing detailed security data for about 100 variables - including price data, yields, asset size indicators, issuer country and sector - facilitating common practices among researchers. Researchers within the Eurosystem gained access to SHS-S in 2015, leading to insights on interconnectedness, international investment, and unconventional monetary policy. Studies using the bank-level have mostly focused on financial intermediation and risk, linking SHS to bank characteristics.

Access to granular data is restricted to the European System of Central Banks, but detailed series at the sectoral level are published on the ECB website, which includes domestic investments. Public SHS-S data, available from 2021-Q1, is more detailed and broader in scope than CPIS, which focuses on foreign investment and has a broader number of investor countries in scope.[2] There are two ways of accessing the public SHS series. First, data is available in a bulk download (a single file is over 1 GB). Second, specific series can be retrieved using filters in an interface from the ECB website. These public SHS-S data have been used researchers to study disaggregated holdings or (gross) transactions at the investor country or investor sector level with various counterparty information, such as geographical exposures, instrument type or currency breakdown (see Du et al., 2024; Faia et al., 2024; Fang et al., 2022, for use cases of these disaggregated SHS series).

III. Data

A. Sample selection

We obtain our sample of journal publications and working papers through an online search. First, we use Google Scholar with broad search key words "SHS" and "securities holdings statistics" and select articles among the first 1,000 search results that use the Eurosystem SHS data. Next, we do the same for the first 250 searches using the Science Direct database from Elsevier. Only

studies benefiting from using granular SHS data are included. Only journal publications with an AJG/ABS or impact factor from Clarivate Analytics are included in the survey and working papers are listed in the Appendix. This methodology should suffice to reach global coverage as external SHS publications have to be cleared for non-disclosure of confidential data by the ESCB and authors should explicitly acknowledge the source. The search was completed on 31 December 2024.

B. Descriptive statistics

Our analysis yields 31 published journal articles for the survey and 71 recent working papers. Table 1 presents an overview of these published articles by ABS journal rankings by year.

Our survey highlights strong growth in the number of SHS articles in recent years as shown in Figure 1. In 2016 the first two published journal articles came out (Boermans et al., 2016; Boermans and Vermeulen, 2016). Since 2022, there has been a notable increase in studies. In the last three years, 19 out of 31 journal papers have been published and of the working papers, 51 out of 71 have emerged since 2022.

SHS-Sectoral data on holdings are most popular. One of the reasons for this popularity is that it allows a full euro area economic view of investment positions. In addition, the main advantage of the SHS-G data is that it allows research on specific individual bank characteristics, while in SHS-S one can still benefit from analyzing security-by-security positions on the banking sector by investor country. Portfolio holdings information on bonds and stocks is the most widely analyzed. Almost all data rely on the portfolio holdings and only a few papers utilize the transactions data.

[Insert Figure 1 here]

[Insert Table 1 here]

IV. Literature overview

In this section we discuss key insights of SHS research, focusing on journal publications. Figure 2 depicts the main topic of SHS research in a word cloud. To reflect on the academic contributions of empirical research using granular SHS data, we have grouped all papers across five fields:

banking and finance (Section IV.*A*), international investment (Section IV.*B*), monetary policy (Section IV.*C*), financial markets (Section IV.*D*) and sustainable finance (Section IV.*E*).

[Insert Figure 2 here]

A. Banking and finance

Our review covers 32 studies on banking and finance, of which 11 published in journals.

Granular SHS data provides an ideal testing ground for theories on *interconnectedness* in finance. In the SHS data one observes the investor, either at country-sector level in SHS-Sectoral or the individual banking group as investor in SHS-Group. The security level holdings further detail the exact exposure of the investor, allowing for various ways on calculating the interconnectedness of networks. These linkages drive contagion between individual banks (see Anand et al., 2018). Furthermore, SHS research highlights the role of mutual funds as intermediaries affecting liquidity and solvency risk. The presence of funds in the network may exaggerate losses through fire sales and reducing banks' capital (see Sydow et al., 2024b). Stress-tests highlight a prominent role of overlapping portfolios and associated negative price externalities due to fire sales (Anand et al., 2018; Aldasoro et al., 2022; Sydow et al., 2024b). These papers merge SHS data with bank balance sheet, own funds and capital requirements indicators, but related research also studies insurance corporations, suggesting they have become more interconnected (Jourde, 2022). While insurers are more vulnerable to shocks stemming from non-financial sectors, banks have stronger links with the rest of the financial industry.

In response to the financial crisis of 2008-2009 regulators required large banking groups to have sufficient capital to prevent future *bail-out*. In particular, contingent convertible capital (CoCo) allows banks to trigger write downs once capital buffers get depleted, bail-in debt. These products may introduce a new contagion channel in the banking sector from cross-holdings because these convertible bonds have equity conversion or write-down mechanics when being bailed in. Granular SHS data allows to directly observe these bond holdings. SHS research suggests there is no direct contagion in terms of creditor banks failing as a result of another bank being bailed in (Hüser et al., 2018; Boermans and Van Wijnbergen, 2018) because of low levels of securities cross-holdings in the interbank network related to CoCo bonds. Most European CoCos are held by foreign investors outside the euro area and investment funds located in

Ireland and Luxembourg (Boermans and Van Wijnbergen, 2018). With the introduction of the European bail-in regime, households reduced their exposures to these risky bonds, while there was a relative increase in the holdings of these instruments by European banks (Marques-Ibanez et al., 2024).

SHS research has further expanded our understanding of the *risk taking* channel of banks and financial institutions more generally in the bond markets. Bubeck et al. (2020) use a difference-in-differences setup to study how holdings of large European banks change in response to the introduction of negative policy rates, differentiating between high (retail) deposit ratio and low deposit ratio banks. They show how negative interest rates induce risk-taking by European banks as they hold more securities with higher yields, especially among more vulnerable banks. Boermans and van der Kroft (2024) analyze the bond holdings across investors and utilize a Solvency II shock to show how banks and insurance corporations tilt their portfolios towards bonds with high systematic yield. They argue that credit rating-based capital regulation in Europe induces risk taking which may hamper financial stability. Barbiero et al. (2024) show that the liquidation value of plegded collateral varies with the counterparty. Using SHS data they show that borrowers that hold both domestic and foreign securities in the same proportions are less likely to pledge domestic securities, suggesting they try to avoid the correlation premium associated with the default risk of the borrower and the risk of the collateral. Finally, Tzamourani (2021) uses SHS to derive residual maturity profiles of households to estimate how they are at risk of shifts in interest rates.

Work in progress: The following 21 working papers related to the field of banking and finance (see Table A1) include research on interconnectedness (Wang et al., 2018; Sydow et al., 2024a; Di Iasio et al., 2022; Craig et al., 2023), bail-in (Ringe and Patel, 2019; Attinà and Bologna, 2021; Altavilla et al., 2023), and risk taking (Bekaert and Breckenfelder, 2019; Fache Rousová and Giuzio, 2019; Mink et al., 2020; Montagna et al., 2020; Fukker et al., 2022; Del Vecchio et al., 2022; Henricot and Piquard, 2022; Allaire et al., 2023; Fay and Ghiselli, 2023; Altavilla et al., 2024; Hartung, 2024; Jochem and Lecomte, 2024; Kaufmann et al., 2024; Nicoletti et al., 2024).

B. International investments

Our review covers 13 studies on international investments, of which seven published in journals.

The international investments literature has grown with the availability of granular SHS data, providing new insights into investment patterns. Most notably, studies have shown how investor heterogeneity and security characteristics shape international investment patterns (Boermans and Vermeulen, 2020), including credit ratings (de Haan and Vermeulen, 2021) and domestic currency preferences (Boermans and Vermeulen, 2016; Boermans, 2023).

SHS data also has enabled look-through approaches for mutual fund holdings, revealing that foreign bond exposures are larger than aggregated data suggests (Carvalho, 2022) and how these indirect mutual fund positions shaped portfolio relocations during the Covid-19 pandemic (Carvalho and Schmitz, 2023). Finally, Jansen (2024) studies insurance corporations and pension fund sensitivities to long-term interest rates, while conversely showing how banks' greater price elasticity absorbs demand shocks in international bond markets.

Work in progress: The following 6 working papers related to the field of international investments (see Table A1) focus on the role of offshore investment funds (Beck et al., 2024), the effects of Brexit (Carvalho and Schmitz, 2025), emerging markets (Bergant et al., 2023), currency preferences (Jansen et al., 2023; Kubitza et al., 2024) and look-through approaches (Lambert et al., 2024).

C. Monetary policy

Our review covers 15 studies on monetary policy, of which five published in journals.

When the SHS data became available to researchers, the Eurosystem in early 2015 initiated quantitative easing (QE) to lower interest rates by purchasing long-term government bonds (see Benigno et al., 2023, for a review). Given the coincidental timing, several important studies on unconventional monetary policy effects have been published using granular SHS data, with more projects ongoing as working papers.

The majority of journal articles focus on the direct impact of quantitative easing. First, Koijen et al. (2021) study the effectiveness of QE in the euro area using SHS-S data. They find that government bond yields decreased by 65 basis points on average, and this estimate varies from 38 to 83 basis points across countries. Interestingly, the ECB purchases mostly came from abroad as foreign investors displayed more elastic demand than domestic investors when yields move. Koijen et al. (2021) further show that foreign investors do not reinvest in the euro area

after selling government bonds. This suggests that QE had little portfolio rebalancing effects when foreign investors substitute outside the euro area. Second, Albertazzi et al. (2021) focus on within euro area portfolio rebalancing but fail to find clear patterns. They show that in vulnerable countries investors moved towards more risky assets, while banks in non-vulnerable countries sold bonds to the ECB and replaced these proceeds mainly with loans. Relatedly, in a DSGE-model Kabaca et al. (2023) explore optional allocations of sovereign bond purchases, showing that given frictions, relatively more purchases in the periphery instead of the core are optimal to induce greater yield impact. Third, Eser et al. (2023) adopt a novel identification to analyze the effect of QE along the yield curve. They show that the ECB purchases drove long-term yields down by about 95 bps. Finally, Arrata et al. (2020) analyze the effect of euro area QE on the repo market. They find that QE aligns repo rates both by raising the scarcity of the bonds purchased and through more aggregated effects by boosting the amount of excess liquidity.

Work in progress: The following 10 working papers on monetary policy (see Table A1) focus on QE and preferred habitat investors (Elsayed et al., 2023; Boermans et al., 2024b), QE and portfolio rebalancing (Bergant et al., 2020; Hudepohl, 2022; Breckenfelder and De Falco, 2023), asset scarcity, portfolio concentration and market liquidity (Ferdinandusse et al., 2020; Boermans and Keshkov, 2018; Nguyen et al., 2023), and, central bank role of lender of last resort, repos and securities lending facilities (Jasova et al., 2021; Greppmair and Jank, 2022).

D. Financial markets

Our review covers 15 studies on financial markets, of which four published in journals.

SHS research on financial markets has set various crisis periods like the Taper Tantrum, Bund Tantrum and Covid shock early 2020 to the fore (Darmouni and Papoutsi, 2022). Much work is this work complementary to the banking and finance literature on how interconnectedness and risk taking by financial institutions affect financial markets (see Section *A*). Work in this area also provides a bridge with the monetary policy literature (see Section *C*), studying policy impact on asset scarcity, ownership concentration and liquidity (Boermans et al., 2016). SHS research also shows how markets with better insolvency regimes encourage portfolio investments (Kliatskova et al., 2023) or test the effects of financial market regulatory changes, specifically for

money-market funds in the US (Fricke et al., 2024).

Work in progress: The following 11 working papers on financial markets (see Table A1) cover a wide aspect of topics, including money markets (Breckenfelder and Schepens, 2022; Fricke et al., 2022b; Fache Rousová et al., 2023), mutual funds (Dötz and Weth, 2019; Bagattini et al., 2023; Gil-Bazo and Santioni, 2024), market liquidity and market regulatory changes (Breckenfelder and Ivashina, 2021), market flows during the Covid-19 pandemic (Breckenfelder and Schepens, 2022; Fache Rousová et al., 2023), and, safe assets Faia et al. (2024), while others are bridging the gaps between monetary policy and financial market outcomes in relationship to repo markets (Brand et al., 2019), corporate bond markets (Faia et al., 2022; Ahmed et al., 2023) and money market paper (Fourel and Schwenninger, 2024).

E. Sustainable finance

Our review covers 22 studies that fall within the research stream of sustainable finance, of which three published in journals.

SHS data allows to estimate investor portfolios in great detail, providing new insights on carbon footprints and portfolio energy intensity, which can be applied further to climate stress testing. Of the journal articles, first, Alessi and Battiston (2022) categorizes the European portfolio holdings of non-financial firms' bonds and equity based on NACE industry classifications to estimate firms' 'greenness' and climate transition risk. Second, Boermans (2023) studies preferred habitat in green bond markets, showing that mutual funds and pension funds overinvest in green bonds, regardless of price movements. Third, Alessi et al. (2024) find that investors reduced their holdings of stocks with high carbon after the Paris Agreement end-2015.

Work in progress: The following 19 working papers on sustainable finance highlight a recent surge in SHS research focusing on climate risk and the environment (see Table A1). A first category focuses on energy-intensive stock holdings (Mésonnier and Nguyen, 2020; Aghion et al., 2022) and bond holdings (Papoutsi et al., 2021). Boermans (2023) introduce the concept of carbon home bias, showing that investors not only overweight domestic stocks, but especially those with higher carbon intensity. Related work focuses on portfolio carbon risk (Jourde and Koné, 2023) and investor preferences (Emiris et al., 2024a,b; Liberati and Marinelli, 2024). SHS research on sustainable finance further analyses corporate bond pricing and green

innovation (Boermans et al., 2024a) while others focus on green bond markets, specifically on the greenium (Pietsch and Salakhova, 2022; Fricke et al., 2022a), clientele Levels et al. (2023) and decarbonization (Boermans and Jacobs, 2024). Finally, there are papers conducting climate stress tests (Alogoskoufis et al., 2021; Dubiel-Teleszynski et al., 2022), analyzing carbon price effects (Belloni et al., 2022) and climate policies for investors (Ehrenbergerová et al., 2023).

F. Other fields

There are a few studies linked to other fields, specifically our review covers five other working papers (see Table A1). These cover topics include household finance (Lamas and Martínez-Miera, 2021; Boermans et al., 2022; Della Corte and Santioni, 2023), supervision (Abidi et al., 2021) and investment patterns in the defense industry (Boermans et al., 2024c).

V. Practitioner's guide

A. SHS research guide outline

In this section we build a versatile regression framework using bond demand functions to analyze the general drivers of holdings, illustrating for bond markets the importance of currency denomination (see also Boermans and Vermeulen, 2016; Boermans and Burger, 2023) and investor heterogeneity more general (Boermans and Vermeulen, 2020). For this practitioner's guide we take the available SHS time-series for the period 2013-Q4 to 2024-Q1. To set out best practices from the literature based on the 27 journal publications (see Table 1), we provide a detailed overview of data selection, investor sector groupings, scope, filtering techniques and data cleaning applied in each of those studies. Researchers working with SHS-S data face important decisions with regard to these steps. However, often such choices are either not reported or not made, yet this may lead to non-standard errors in research, hence our best practices are easy to implement for future research.

B. Data preparation and selection

There are two types of SHS data, one covering all sectors (SHS-S) and one on individual banking groups (SHS-G). In this practitioner's guide we focus on SHS-S as the majority of work uses

this module of the data (only 6 out of 31 papers use only SHS-G). In addition, many of the data preparation and cleaning steps are also relevant for SHS-G.

Note that for SHS-G one also has to make importance choices on on the level of consolidation and treatment of intragroup holdings. Only two out of eight studies with SHS-G data explain their considerations, noting that both use the fully consolidated data, meaning aggregation over the individual bank entities, and, also opt for the inclusion of intragroup holdings.

SHS data is organized by quarter, covering 45 time periods from 2013q4 to 2024q1 (with regular quarterly updates) across four type of instruments. A raw data file for a single period (clean copy) of bond holdings includes 23.8 mln observations, 7.6 mln for stocks, 9,8 mln for investment funds and 8.3 mln for money market funds (based on the 2024-Q1 files provided by the ECB), hence nearly 50 millions observations of quarterly holdings data. Merging all these files would equate to about 2.3 billion observations.[3] With current standards in computational performance, such dataset would be unmanageable, especially given that it contains 88 variables including many string variables that use a lot of memory. Hence, data preparation steps are necessary when working the SHS data.

Several common data preparation steps are taken. Table 3 presents an overview.

[Insert Table 3]

First, most work relies on positions, so filtering on these holdings is accomplished using "amount_type" with value 'LE' (for transactions, one selects 'T'). This filter typically reduces the file size by half. With one exception, all papers report on the "amount_type" filtering applied for their research, meaning a choice between positions or net transactions. Only two of the 31 published papers use the transactions data ('T'), the rest focuses on positions ('LE').

Second, 14 of the 31 papers use market values. Using the filter for "valuation" with the value 'M' instead of nominal values 'N' this again reduces the file size by approximately half on average. Five papers use nominal valuation (one studies applies both nominal and market values).

Surprisingly, our analysis also shows that 12 published papers do not report about the valuation principle, which in most cases is necessary and required for replication purposes. The valuation choice can have large ramifications for the correct empirical specification (and thus may affect the results). When analyzing holdings, if these are at market value, one has to consider

that price effects determine the value of the holdings. Not accounting for price effects thus leads to biased results. Similarly, if one uses nominal values, this means that for the investor's portfolio rebalancing effects related to asymmetric price movements across securities are not accounted for. We argue that in most cases using market values is preferably because it will take into account how differential price shifts across securities may induce portfolio rebalancing effects. If one looks only at nominal valuations, such market equilibrium effects will be missed. Crucially, one then needs a security-by-security level analysis to control for individual bond price effects. Hence, for our analysis we focus on market valuations. Depending on the research topic one can take the positions at market value or nominal value.

Third, 13 published articles focus on euro area investors as they explicitly report this in the article. SHS-G is only available for banks from the euro area, thus another 6 studies implicitly focus on the euro area, although for these papers it is unclear how they deal with foreign branches and subsidiaries of banking groups located outside of the euro area. Surprisingly, 11 published papers do not report on any filtering for euro area investors, meaning they may included unintentionally incomplete data from non-euro area investors in their sample. In this practitioner's guide we propose to explicitly focus on the euro area for most research purposes because the coverage is generally above well 90 percent for each country and thus representative, while for non-euro area countries coverage is not guaranteed by law and is generally much lower and varies more, especially for non-European countries were reported relies exclusively on third-party holdings. Hence, to obtain a consistent dataset in general we propose to filter on the 20 euro area investor countries only based on the attribute "holder_area".

Fourth, these is the issue of potential double counting in the data. Surprisingly, as shown in Table 3, only three papers explicitly state these third-party holdings are excluded (Boermans and Vermeulen, 2020; Fricke et al., 2024; Boermans and Galema, forthcoming). National central banks as reporting country ("ref_area") require domestic custodians to report on foreign holdings worldwide. These so-called third-party holdings also include euro area positions. When a domestic investor has its administration abroad through a foreign custodian, say an Austrian bank with a German custodian, then double counting arises when the German central bank has that one custodian reporting, which should include the positions of the Austrian bank, and, at the same time, the Austrian bank directly reports to the Austrian central bank. To

avoid such potential double counting, these third-party holdings can be discarded from the sample. One exception is to include third-party holdings for households (S.14) and non-profit organizations (S.15) because these investors cannot suffer from double reporting given the custodian chains. There are no households and non-profit organizations directly reporting to central banks, hence the only way to cover these is via their direct custodian, regardless of its country of residence. Thus, one best practice is to filter on third-party holdings via matching "ref_area" and "holder_sector" given the "holder_area".

In summary, we apply the following generic data preparation steps focusing on bonds markets (F.32) by (i) selecting positions (LE), (ii) taking market values (M), (iii) focusing on euro area investors only and (iv) avoiding double counting with custodian data through third-party holdings. After this selection procedure, the total euro area bond holdings in 2024-Q1 are 15.5 trillion euros at market value, with 2,472,985 observations. As shown in Table 3 not all papers report on such choices. We strongly recommend as a "best practice" to clearly report the choices made for the data preparation and selection when using SHS data.

C. Data cleaning

In this section we set out general cleaning procedures for SHS research. These steps are important to obtain unbiased results. In addition, given the lack of common cleaning procedures, having a practitioner's guide should enhance comparability across papers. Arguably, lack of or variation in data cleaning could introduce non-standard errors in the findings of the outlined papers (see e.g. Menkveld et al., 2024).

*C.*1 General cleaning procedures for SHS

Our review of the additional cleaning steps in general highlights that most researchers rely on *raw* SHS data for their analysis. We find that only 6 of the 31 journal articles perform additional data cleaning steps on the raw SHS data, but the overlap in the procedures is very minimal.

Note that of the 25 papers that do not report any further cleaning steps, 8 of those mention they limit the scope of their data to certain issuer countries or sectors. In order to limit non-standard error in SHS research, we advocate a standard set of additional cleaning steps. We will benchmark the impact of our proposed standard cleaning steps against the *raw* SHS data of

euro area bond holdings at market values (15.5 trillion euros in 2024-Q1; n = 2,472,985).

First, focus on portfolio holdings. In practice, only two papers explicitly filter on portfolio holdings (Boermans and Vermeulen, 2020; Boermans and van der Kroft, 2024), while others do not mention filtering on portfolio investments or justify (implicitly) including foreign direct investments (FDI). Most studies analyze readily tradable securities, unlike loans or FDI. SHS data also cover direct investment positions, such as ownership in stock-listed subsidiaries, which are rarely traded but represent significant portions of security value (at least 10%) and these positions tend to be underreported in SHS-S. In this practitioner's guide, we propose focusing solely on portfolio holdings for SHS research unless otherwise motivated.

Second, exclude short positions. In our survey of published articles, only three studies exclude (aggregate) short positions (Boermans and Vermeulen, 2016, 2020; Boermans and Burger, 2023; Boermans and Galema, forthcoming). Although some researchers may want to focus on short positions, shorts in SHS-S are rare and can cause analytical issues. For instance, taking logs of positions excludes negatives, and not transforming holdings results in highly non-linear data with extreme outliers.

Third, exclude non-active securities (e.g. those in default status or already redeemed). The SHS data included 'old' securities, for example, debt positions of issuers that have long defaulted or misreported securities that are already redeemed. When these could still be part of the portfolio of the investors, such positions are not readily traded and are thus by definition no longer securities. Only one paper explicitly excludes these non-active securities (Boermans and Vermeulen, 2020).

Fourth, exclude investments in tax havens. These small countries often serve as pass-through rather than final recipients of capital flows, as SHS data are based on residency, not group-based principles (see Lane and Milesi-Ferretti, 2008; Coppola et al., 2021). While issuer destination may not always be crucial, it typically influences portfolio choice. Ignoring this can introduce bias. Excluding tax haven destinations does not significantly affect the overall sample coverage and this step could easily be dropped with proper motivations.

After these four cleaning steps, for bonds the coverage is 97.4% in terms of holdings while dropping 16.5% of the observations.

Fifth, a more debatable action is to select securities based on their holding amount or market

size. When researchers adopt s-b-s analysis, not doing so would create many observations of low relevance, similar to a small country focus when running unweighted world country analysis, where in economic terms US is underrepresented or for populations countries like India and China are underrepresented, compared to say country information from Malta or the Maldives . Because weighting introduces its own issues, and unless a researcher would be interested in say "small bonds" or "tiny holdings", we propose the exclude small values, although again the threshold level is hard to establish and somewhat arbitrary. If one puts it too high, the selection bias increases and relevant positions are discarded, whereas a too low threshold would render the whole exercise redundant. Another issue with data cleaning on size is that on the issuer side it requires to have data coverage for that variable ("amount_out_eur" for money market and bonds, and, "market_cap_eur" for stocks and investment funds). However, if such essential information is missing, other security information is often also less reliable or missing. Thus, using this information also trims the data and will in relative terms increase the coverage of other (merged) variables for the holdings data. We can benchmark the impact of different approaches and propose to only limit by market capitialisation or amount outstanding as a security level size indicator.

As a size threshold of the security, following Boermans and van der Kroft (2024) we suggest a threshold of 10 million euros at the security level. In terms of observations for bonds, such threshold implies moving from 2,065,462 observations after the aforementioned four cleaning steps to 1,207,091 observations, a reduction of 41.5%. Meanwhile, in terms of coverage, the total holdings go from 15,1 euros trillion to 14,2 trillion, representing a much smaller reduction of 6%. Clearly, as the majority of papers do not take into account the vast amount of observations at the lower end of the holdings scale, which could impact the results. In this practitioner's guide we recommend on setting a similar threshold level which does lead to large drops in the number of observations while at the same time maintaining a high coverage of the cleaned SHS data.[4]

In addition, cleaning can also be done for very small holdings positions for the data, which has similar advantages as filtering on security size (as proposed in cleaning step 5). Bubeck et al. (2020) use SHS-G data and filter at the bank level on securities positions of at least 0.5 million euros, while Boermans and van der Kroft (2024) use SHS-S data and exclude observations at the investor-sector-time level below 10 million euros in holdings. Additionally, they require that

at least 5% of the bond amount outstanding is held by euro area investors where at investor sector-time level at least one investor group holds at least 1 million euros of the bond. The latter approach ensures that only relevant securities in the intensive margin are included in the sample. If researchers are interested in the extensive margin, such filtering approaches should be avoided. No other published paper reports on cleaning based on holdings size, but we still suggest to filter small holdings for the holder sector by holder country level dataset at 10.000 euros to ensure small holdings do not dominate the sample and for bonds small incremental investment positions may hint at data quality issues as most bonds require a minimum of 10.000 euros.[5]

Finally, other data cleaning steps not performed in the published articles may be relevant. Working papers with SHS data suggest additional cleaning steps.

For bonds, researchers want to clean their data by excluding certain debt types. The attribute "debt_type" covers a very wide range of debt instruments, including certificates ('D.18'). These certificates are very large in number but small in volume (and hence holdings and transactions). Concretely, they cover 459,174 of the 1,207,091 observations, but represent only 1% in bond holdings coverage. The large majority of these certificates are issued by German entities and often include trackers and so-called Schuldschein debt, which are basically private loans (Franklin, 2020). Therefore we exclude these from the sample of bond holdings. Note that these instrument also have a notation basis in kind ('CCY') instead of a quotation in percentages ('PCL' or 'PDT' referring to clean and dirty prices), which is non-standard for bonds.

For quoted shares, in the finance literature penny stocks are often discarded from the analysis, especially when evaluating returns (Boermans and Galema, forthcoming). The U.S. Securities and Exchange Commission (SEC) takes a threshold of below USD 5.00 for the definition of 'penny stock' for firms with a low market capitalization, because these stocks are more volatile as investors often buy large quantities of shares with low investment value. However, the 5.00 threshold is relatively high, and would discard 15% of the observations on a cleaned file representing 4% of the total euro area holdings. Instead, we propose to exclude only stocks with a "price value" below 1.00, comprising about 3% of the observations and less than 0.5% of the holdings.

*C.*2 Other filters

This sector describes other filters applied in the literature which will not be executed in this practitioner's guide.

Data filtering on destinations is often very dependent on the research topic. Some papers focus only on the euro area issuers (Darmouni and Papoutsi, 2022; Alessi and Battiston, 2022; Arrata et al., 2020; Albertazzi et al., 2021) – or even a narrower subset (Boermans et al., 2016) – or the European Union (Alessi et al., 2024). Others focus on emerging markets (Boermans and Burger, 2023) or a selected list of countries (Kliatskova et al., 2023, who include 33 issuer countries).

There are several studies that also focus on certain issuer sectors only. 3 papers focus only on non-financial corporates by filtering on "issuer_sector" contains 'S.11' (Boermans and van der Kroft, 2024; Darmouni and Papoutsi, 2022; Kliatskova et al., 2023). 2 papers include only corporate issuers (Alessi and Battiston, 2022; Boermans, 2023) and de Haan and Vermeulen (2021) focus on sovereign issuers with "issuer_sector" equal to 'S.1311'. Finally, Boermans et al. (2016) take a sample with only banks ('S.122') and sovereigns.

In addition, 3 of the 27 published papers focus only on EUR-denominated securities (de Haan and Vermeulen, 2021; Arrata et al., 2020; Alessi and Battiston, 2022). Finally, one study focuses only on newly issued securities (Albertazzi et al., 2021).

D. Data aggregation of investors for SHS-S

SHS allows for a wide range of possibilities regarding data aggregration, consolidation and groupings over investors. In this Section we explain these choices based on the literature and advocate to work with investor sector and investor country details, using eight different investor sectors across 20 euro area holder countries.

*D.*1 Data aggregation of investor sectors for SHS-S

There are 24 different holder sectors available in the SHS-S data based on the ESA 2010 framework with subsectors and undefined groups. While greater granularity allows for various research directions for each given sector, e.g. by studying the highly particular holdings of local governments (S.1313) or non-profit institutions serving households (S.15), typically researchers

aggregate sectoral holdings to only a few and relatively large holder sectors.

[Insert Table 4]

As presented in Table 4, 5 of the 31 published articles do not report any information on their treatment of investor sectors. Most scholars (26 of the 31) perform some grouping over investor sectors, however, there is no consensus on these aggregations while such choices can affect the results. We propose to seek a harmonized approach to limit non-standard errors in SHS-S research by focusing on eight distinct investor sectors which have clear economic interpretation and sizable holdings as a group as shown in Table 4. Clearly, many studies have applied regroupings and aggregations of investor sectors that are close to ours, but these often go further in consolidating, e.g. by taking insurance corporations and pension funds as a single sector (Arrata et al., 2020; Boermans and Vermeulen, 2020; Koijen et al., 2021; Darmouni and Papoutsi, 2022; Eser et al., 2023; Jansen, 2024). Of course, such deliberations are fine, but they still may hamper comparability and, depending on the research topic, may neglect important investor sector heterogeneity, e.g. between preferences of insurance corporations and pension funds.

Figure 3 shows the bond holdings across different investor sectors from the euro area, highlighting that within global bond markets, mainly banks, investment funds and insurance corporations are dominant investor groups. Of course, for other asset classes in SHS-S including money market paper, quoted shares and investment funds, other investor groups become more dominant.

[Insert Figure 3]

*D.*2 Data aggregation over investor countries

Besides consolidating over the investor sectors, data can be grouped by investor country. The dimension of the investor country may not be relevant and the aggregation on the "holder _area is useful to obtain a perspective of an investor in the euro area. This also significantly reduces the size of the dataset (by about half, without affecting the total holdings/ data coverage). Alternatively, one may start with an euro area perspective analysis. Then one can explore the investor country-level dimension to test for holder country-specific effects that may drive the results (see e.g. Hudepohl, 2022).

Researchers may be interested in the investment patterns of foreign investors, i.e. those from outside the euro area. Conceptually, various papers rely on a rough estimate of the foreign sector holdings, often referred to as "Rest of the World" (ROW). The construction is completed by residual approach as follows:

$$HOLD^{ROW}_{i,s,j,t} \approx SIZE_{i,t} - \sum_{i=1}^{n} HOLD_{i,s,j,t} \tag{1}$$

There are two important caveats to this approach. First, the coverage of the SHS-S data is at least 95 percent for the main investor sectors and typically a somewhat lower for the other investor sectors. This means the residual representing $HOLD^{ROW}$ is always an *overestimation* because the granular holdings data lack a grossing up method at the security level. Second, the holdings of the central bank (including the ECB) are missing, which is especially problematic for bonds, given the various central bank bond purchase programmes in the euro area, in addition to central bank own holdings. This further exaggerates the overestimation of the residual. Combined this leads to misleading estimates of the foreign holdings ($HOLD^{ROW}$) as these investments will be *too high and stable* over time. There is one paper that aims to estimate the central bank positions with a simple algorithm that translates the macroeconomic data to the security level (Boermans and Keshkov, 2018), whereas more recent papers rely on ECB transactions to derive the ECB holdings to fill such data gaps (Boermans et al., 2024b; Elsayed et al., 2023).

*D.*3 Final cleaning after data aggregations

As explained in Section *C.*1, we propose to filter on bond size. However, after setting the final aggregation level one can identify holding positions that are larger than the amount outstanding or market capitalization, suggesting some minor data quality issues either with the reported holdings or the reference data on bond size.

Theoretically, the reporting holdings cannot exceed the size indicator unless there are exceptional shorts. In practice, data quality of the size indicator may lead to some distorted observations. Hence we propose to set *Size* equal to the total observed holdings in cases where $obs\_value_{i,s,t} > amount\_out\_eur_{i,s,t}$ for debt or $obs\_value_{i,s,t} > market\_cap\_eur_{i,s,t}$ for equity. This data quality adjustment is very rare, i.e. it affects 1,177 observations for bonds.

Finally, we can drop very small holding positions, where we suggest a threshold of 10,000 euros as a minimum about held by euro area investors from a given investor sector-country pair in a single security at the relevant time period i, s, j, t.

Our final sample for the practitioner's guide based on the 2024-Q1 bond holdings contains 664,247 observations, with total investments of 14.0 trillion euros at market value. Compared to the raw data, our aggregation and cleaning gives a coverage rate higher than 90% (and doing so for money markets, quoted stocks or investment funds would yield similar coverage rates.)

E. Securities holdings regression models

*E.*1 General model

A wide range of papers aims to explain the holdings as dependent variable in (panel) regressions. A typical empirical specification is the following:

$$HOLD_{i,s,j,t} = \alpha + \beta' X_{i,s,j,t} + \tau_t + \varepsilon_{i,s,j,t} \tag{2}$$

where $HOLD$ represents the positions (at market value), preferably *in logs* as $lnHOLD$ to obtain a distribution closer to a normal distribution by smoothing outliers (see Bubeck et al., 2020), i is the individual security, often the subset of identifiers with ISIN code (International Security Identification Number, an ISO-standard for unique coding of individual stocks, bonds and funds), s is the investor sector, defined by ESA-2010 subsector classes in 24 groupings or the proposed 8 investor sectors from the practitioner's guide (see Table 4, j the investor country - in principle all countries worldwide but preferable only euro area area investors (see Table 3) - and t the time period, quarterly, with time fixed effects τ.

Researchers must be aware that the holdings are at market value. Without any transactions these positions can still move along with currency and price changes, as well as changes in the data structure, e.g. when a new reporting agent enters the underlying sample. Proceeding in a time-series, because this is a multidimensional-panel the researcher has to choose the set of fixed effects, depending on the interest in X, the explanatory variables.

When X are general macroeconomic variables that change on a quarterly basis, time fixed effects are not feasible due to perfect multicollinearity with the macroeconomic time data. However, in most cases time-fixed effects are useful to ensure there are no specific time trends

in the series. Similarly, investor country fixed effects are often included to control for country-specific effects, unless holder country information is included in the vector X or when researchers analyze differences across holder countries. Likewise, investor sector fixed effects are included in most specifications to account for sector-specific effects, except when different investment patterns by investor sector are analyzed.

*E.*2 Security-by-security model for Euro Area investment

In most cases researchers can analyze 'aggregated' euro area investment positions using this aggregated framework. Concretely, aggregation means that all positions are summed over the investor country j and investor sector s dimension, but not over the individual security dimension i, yielding the following:

$$HOLD_{i,t} = \alpha + \beta' X_{i,t} + \tau_t + \varepsilon_{i,t} \tag{3}$$

In this way one treats all euro area investors as a single homogeneous group. This approach cancels out many potential outliers which need to be dealt with in investor country and investor sector granularity specifications that may suffer from "noise" in the holdings, e.g. think of small countries of tiny investor sectors which Section *B* outlines. Concretely, summed positions over investor country and investor sector reduce the noise of having a lot of very tiny positions in individual assets. In principle, applying weights to the OLS regressions mitigates this but to facilitate the ease of interpretation of the estimated coefficients we prefer an aggregate model.

Another important advantage of aggregation while keeping security-level data is that the processing time is very significantly reduced. For example, a raw dataset of with bond-level time-series contains about 100 million observations, making it difficult computationally or very time-consuming to estimate various models. Benefiting from the full security-by-security data for bonds reduces the number of observations by a factor of about 10 when aggregating the investor country j dimension to only euro area investors, and similarly another factor of about 10 for the investor sector s dimension, leaving about 8 million observations for the period 2013-Q4 to 2024-Q1.

*E.*3 Dealing with investor heterogeneity in regressions

To take differences in investor preferences into account we propose to group investor sector s across eight categories when using SHS-S data. Methodologically, true investor heterogeneity implies that the coefficients of each explanatory variable is allow to vary for each investor sector s. This can be achieved with interaction terms:

$$HOLD_{i,s,j,t} = \alpha + \beta' X_{i,s,j,t} * HolderSector_{i,j,t} + \theta' X_{i,t} + \gamma_s + \eta_j + \tau_t + \varepsilon_{i,s,j,t} \tag{4}$$

*E.*4 Fixed effects alternatives

Researcher using security-by-security regressions carefully select proper fixed effects. When investment patterns towards a set of destination countries are analyzed without further modeling, issuer country fixed effects are in order to absorb fixed destination country preferences of the specific euro area investor. Such destination country controls may not suffice because the issuers are also from different sectors. Including issuer sector fixed effects further capture unobserved issuer sector characteristics, which are often referred to as multilateral resistance terms in gravity models (see Anderson and Van Wincoop, 2003) when combined with holder sector-country and issuer sector-country dummies (Boermans and Vermeulen, 2020; Bergant et al., 2020).

The selection of fixed effects also depends on the level of aggregation. If investment behavior differs across investor sectors, then by re-introducing the sector level to Equation (2) one explicitly models investor sector preferences as follows with "granular" multilateral resistance terms that also pick up financial frictions (as long as you have no priors as to why the effects will differ across either investor countries or destination sector-countries), in a multidimensional time-series panel setting:

$$HOLD_{i,s,j,t} = \beta_s' X_{i,s,j,t} + FE_{s,j} + FE_{h,c} + \varepsilon_{s,j,t} \tag{5}$$

where FE_s, c are dummies for the investor sector and country (s, j), including sector*country, and destination sector and country (h, c), including interactions.[6] Instead of destination sector classifications one may use industry (NACE) classifications instead. We drop the constant α term here due to multiple fixed effects. Note that the fixed effect country and sector dummies

are not able to capture the different responses of certain investor countries or investor sectors for a given explanatory variable, but only control for the difference in holdings on average, against other countries and sectors.

Investor heterogeneity implies one estimates multiple coefficients across investor sectors. By estimating the coefficients for all investor sectors in a single regression with $\beta_s X_{i,s,j,t}$ one can infer statistical differences between coefficients across investor sectors, against a reference category or benchmark investor sector (see Boermans and Vermeulen, 2020). Here standard errors are clustered at the investor sector-country-level to correct for potential serial correlation of the error terms. Such "conservative" approach allowing for investor heterogeneity with multitude fixed effects and corrected standard errors is often lacking in studies analyzing international investment positions creating potential biased standard errors.

The investment positions ($HOLD$) depend on the size of the security. For money market paper and bonds the amount outstanding is thus an essential control variable to include in the vector X while for stocks and investment funds the market capitalization is required.

F. Application: Bond demand estimates with investor heterogeneity

We illustrate the modeling choices regarding the importance of investor heterogeneity and the level of aggregation.

Our regression model shows cross-sectional estimates in Section F.1 and multidimensional time-series panel specifications in Section F.2 related to investor sector differences on how covariates affect bond holdings, focusing on currency preferences.

F.1 Cross-sectional analysis on home currency bias with investor heterogeneity

In this section we present a cross-sectional analysis on how investor demand differs on bond, issuer and country characteristics in addition to investor sector preferences that are heterogeneous among different investors. Following Equation 3, neglecting time subscripts, for the covariates in vector X we include a simple EUR-denomination dummy (EUR) to test if euro area investors have preferences for their "home" currency, highlighting how this varies across investors. Because bond (market) value co-determines holdings, we further control for the market value of the amount outstanding ($lnSIZE$) as additional covariate in X. Doing so also allows us to show

if certain investors prefer to hold larger or smaller bonds, which also functions as a proxy for market liquidity.

We extend the baseline model of the euro area specification by disaggregating the data to the holder country level and holder sector level and then combine these two to analyze investor heterogeneity at the country*sector level in line with Equation 4.

We use the cleaned dataset that includes bond portfolio holdings ($lnHOLD$) of eight investor sectors from 20 countries across 68,373 unique bonds, yielding 635,453 observations for reference period 2024-Q1. The regression specification in most extensive form is:

$$\begin{aligned} \ln HOLD_{i,s,j} = \alpha &+ \beta' EUR_i \cdot HS_s + \gamma' \ln SIZE_i \cdot HS_s \\ &+ \zeta EUR_i + \eta \ln SIZE_i + HS_s + HC_j \\ &+ HS \cdot HC_{s,j} + \varepsilon_{i,s,j} \end{aligned} \tag{6}$$

We extend each specification based on the following (dis)aggregations: (1) at the euro area level, (2) at the investor-country level, (3), at the investor-sector level, (4) at the investor-country-sector level, and, (5) introducing investor sector-specific interaction terms, which test for the importance of investor heterogeneity (see Boermans and Vermeulen, 2020) in a single regression. We argue that for a cross-sectional analysis our regression model (5) as presented in Table 5 Column (5).

[Insert Table 5]

Table 5 highlights the importance of home currency for investors. Table 5 Column (1) shows that at the security-level both currency and bond size are important drivers of the portfolio allocations in bond markets. Generally, euro area investor hold 251 percent more in home currency bonds. In Columns (2) and (3) these effects are robust at the investor country-level and sector-level including relevant fixed effects. Column (4) shows that at the disaggregated holder-country, holder-sector level the results are robust, yet the size of the estimated coefficient is much smaller. Ceteris paribus, if a bond is denominated in euros, investors from the euro area will increase their holdings of this security by 121%. These results are very similar to Boermans and Vermeulen (2016) who analyze a more detailed gravity model to test for common currency preference in the euro area in a cross-section for the period 2014-Q4 they obtain an estimated coefficient of 87%.

Table 5 Columns (5) however shows that the common currency preference for euro-denominated debt varies strongly across different investors. In this preferred specification we explicitly allow the home-currency preference for EUR-denominated debt to vary across investor sectors. It includes the main effects of EUR and $lnSIZE$ for the benchmark investors and in the other 'Columns' the deviation of specific investor sectors from the main effect. For example, keeping all else constant, benchmark investors from the euro area hold 113 percent more of a home currency bond, comparable to Column (4). After accounting for investor heterogeneity our results show that euro area banks have a stronger preference for home currency bonds. Banks hold even 69% more than the reference euro area investor, translating in a portfolio tilt towards of 182% higher investment in home currency bonds. Similarly, insurance corporations also more strongly tilt towards home currency bonds. The more aggregated specifications hide the underlying variance at investor sector level, therefore we suggest to model investor heterogeneity with interaction terms in bond demand regressions.[7]

F.2 Time-series panel analysis on home currency bias with investor heterogeneity

In a multidimensional time-series panel setting home currency preference are also persistent. Now we move to the full-panel model following Equation 4 where we retain all fixed effects and include the interaction terms for investor-sector specific effects for the EUR-denomination (EUR) and bond size ($SIZE$) similar to Equation 6 with time fixed effects for 2013Q4-2024-Q1. The regressions are based on a cleaned dataset for eight investor sectors from 20 countries, now across 151,259 unique bonds with 20,112,593 observations.

Our panel settings are at i, s, j for a given quarter t with standard errors clustered at the investor. Table 6 Column (1) shows that euro area investors have a tendency to invest 110 percent more in home currency bonds, very consistent with the cross-sectional analysis for 2024-Q1 in Table 5 Column (5) where the reference investor group displayed a tendency to hold 113 percent more in home currency bonds.

Table 6 Column (2) demonstrates the importance of euro-denominated bonds is not equal across different investors, showing the importance of investor heterogeneity (similar to Boermans and Vermeulen, 2020). In this preferred specification especially banks and insurance corporations have a strong preference for home currency bonds, similar to the cross-sectional results in 5

Column (5).[8] The results highlight that for this case study the cross-sectional results are very close to those from the time-series, so doing time-series analysis for bond demand estimates may not always be necessary unless one is interested in the state-dependent nature of the demand.

[Insert Table 6]

To summarize, our application has set out how investor heterogeneity is relevant and can be incorporates with dummy interaction terms between the investor sector and the other explanatory variables. These bond demand function estimates appear persistent over the time period 2013-Q4 to 2024-Q1. Our case study highlights that a cross-sectional analysis may suffice for many studies and that controlling for home currency in addition to other factors is important.

G. Analyzing sustainability preferences of European investors

In this section we analyze the holdings of green bonds and sustainability-linked bonds (see also Pietsch and Salakhova, 2022; Fricke et al., 2022a; Boermans, 2023; Levels et al., 2023; Boermans and Jacobs, 2024, on green bond holdings using SHS-S data). Table 7 shows the results using the raw SHS-S data and the cleaned data, where in Column (1) we find no preferences for green bonds or sustainability-linked bonds among any of the European investor sectors in the cross-section, whereas in Column (2) European investors tend to invest 15 percent more in a bond with a green bond label and 16 percent more in a sustainability-linked bond while controlling for other factors. This overallocation depends on the investor sector. Banks, insurance corporations and households tend to invest less in bonds with a green label and the same applies to sustainability-linked bonds except for households. In Column (3) we find these preferences are very consistent over time as the estimated coefficients are of similar size in the cross-section and the time-series. This analysis highlights how data cleaning may affect the results, confirming the comparison of the main cross-sectional results from Table 5 and Table A2.

[Insert Table 7]

VI. SHS research: An agenda for the future

We identify four key areas that will shape the future agenda of SHS research. Firstly, studies will increasingly rely on merging datasets. The granular nature of SHS data allows for the

seamless integration of portfolio holdings with other datasets, providing a comprehensive view that could contribute to various research agendas. Despite this potential, most studies have primarily focused on SHS data alone. Ullersma and van Lelyveld (2021) highlight how granular data, particularly portfolio holdings, can be linked with reference data using identifiers like LEI and ISIN, along with supervisory data and other granular sources such as AnaCredit, SFTR, and EMIR. The SHS-Sectoral data is already premerged with several attributes from the CSDB, and this could be extended with more security characteristics and issuer/firm/sovereign details, including credit ratings, CDS spreads, industry classifications, and CO2 emission data to enhance comparability and ease research. Similarly, SHS-G data could be premerged with supervisory and commercial sources. Future research should focus on integrating SHS data with other databases in a transparent way such that others benefit and learning from existing projects, for example on how to group issuer information to the firm level (Beck et al., 2024).

Secondly, future SHS research will deepen our understanding of financial crises, their causes, and responses. Researchers can use SHS data to study investor responses and propagation channels during various financial crises periods, such as the Covid-19 pandemic (Breckenfelder and Ivashina, 2021; Carvalho and Schmitz, 2023; Breckenfelder and Schepens, 2022; Di Iasio et al., 2022), interest rate hikes Boermans et al. (2016), Brexit Carvalho and Schmitz (2025), and emerging market crises (Boermans, 2023; Jansen et al., 2023). Understanding how crises unfold or are prevented, future studies will also benefit from work on unconventional monetary operations, such as quantitative easing. However, the exclusion of Eurosystem holdings from SHS data limits research opportunities (for exceptions, see e.g. Boermans and Van Wijnbergen (2018); Elsayed et al. (2023)).

Thirdly, future research will expand the focus beyond holdings data to transactions data and portfolio performance data from SHS. Currently, most SHS research focuses on investment positions. Some researchers have emphasized the importance of net transactions SHS data for testing portfolio rebalancing (Bergant et al., 2020), short-term responses to global shocks Boermans (2023), and monetary policies (Arrata et al., 2020; Hudepohl, 2022). Future studies may increasingly analyze these transactions data. Additionally, the SHS-S module on the financial performance of portfolio investments is still in its early stages, leading to limited studies on price revaluations and exchange rate effects to analyze portfolio performance (Boermans

et al., 2022; Boermans and Burger, 2023), leaving ample room for more research.

Finally, the full research potential of securities data remains untapped. There are numerous unexplored areas and topics for investigation, such as various debt types, including asset-backed securities and covered bonds, with detailed analysis of subcategories like residential mortgage-backed securities and car loan securitizations. Other areas of interest include holdings of distressed bonds, Islamic financing bonds, and bonds with negative yields. Researchers could also study specific issuer industries, such as energy-intensive stocks in sustainable finance or government holdings. There are countless opportunities for analyzing individual industries, such as the aircraft industry, technology giants, automobile manufacturers, healthcare providers, bank sector ownership, and positions in supranational organizations like the newly issued EU bonds under the NGEU and SURE programs. SHS data could also be a valuable testing ground for topics ranging from AI to investor embeddings, complementing the standard set of firm characteristics used in finance and economics (Gabaix et al., 2024).

VII. Conclusion

This literature review has studied 102 studies using Securities Holdings Statistics at granular portfolio levels between 2016 and 2024. We document a recent surge in SHS research output using security-by-security data on European investments and expect this trend to continue. This paper has discussed important contributions from the literature, focusing on 31 journal articles grouped in five different research themes: banking and finance, international investment, monetary policy, financial markets and sustainable finance.

This survey of the literature using SHS data stresses that there is no consensus on the cleaning and organizing of the SHS data and we propose steps for proper handling of the raw SHS data related to filtering and aggregations. Another concern is the lack of consensus how to account for different investor preferences when using the SHS-S (Sectoral) data. Here we set out a method to take investor heterogeneity into account when analyzing portfolio choice. These best practices aim to foster future replication of papers as well as enhance comparability of results. Setting out our practitioner's guide we aim to streamline best practices for SHS future research. This is illustrated in this study by highlighting the importance of investor sector heterogeneity and home currency preferences for portfolio choice models. Finally, we analyze which European

investors seek green bonds and sustainability-linked bonds.

We anticipate that the accumulated knowledge on SHS data provides new researchers with an opportunity to quickly adapt and apply this large and growing dataset to their advantage using insights from previous research and best practices.

References

Abidi, N., L. Amato, I. Miquel-Flores, and Q. Vandeweyer (2021): "The bright side of transparency: Evidence from supervisory capital requirements," *Mimeo*.

Agarwal, V., W. Jiang, Y. Tang, and B. Yang (2013): "Uncovering hedge fund skill from the portfolio holdings they hide," *Journal of Finance*, 68, 739–783.

Aghion, P., L. Boneva, J. Breckenfelder, L. Laeven, C. Olovsson, A. A. Popov, and E. Rancoita (2022): "Financial markets and green innovation," *Mimeo*, 2686.

Ahmed, A., B. Hofmann, and M. Schmitz (2023): "Foreign institutional investors, monetary policy, and reaching for yield," *BIS Working Papers No*, 1153.

Albertazzi, U., B. Becker, and M. Boucinha (2021): "Portfolio rebalancing and the transmission of large-scale asset purchase programs: Evidence from the Euro area," *Journal of Financial Intermediation*, 48, 100896.

Aldasoro, I., A.-C. Hüser, and C. Kok (2022): "Contagion accounting in stress-testing," *Journal of Economic Dynamics and Control*, 137, 104354.

Alessi, L. and S. Battiston (2022): "Two sides of the same coin: Green Taxonomy alignment versus transition risk in financial portfolios," *International Review of Financial Analysis*, 84, 102319.

Alessi, L., S. Battiston, and V. Kvedaras (2024): "Over with carbon? Investors' reaction to the Paris Agreement and the US withdrawal," *Journal of Financial Stability*, 101232.

Allaire, N., J. Breckenfelder, and M. Hoerova (2023): "Fund fragility: The role of investor base," *ECB Working Paper No*, 2874.

Alogoskoufis, S., N. Dunz, T. Emambakhsh, T. Hennig, M. Kaijser, C. Kouratzoglou, M. A. Muñoz, L. Parisi, and C. Salleo (2021): "ECB economy-wide climate stress test: Methodology and results," *ECB Occasional Paper No*, 281.

Altavilla, C., J. Begenau, L. Burlon, and F. Hünnekes (2024): "Determinants of bank performance: Evidence from replicating portfolios," *ECB Working Paper No*, 2937.

Altavilla, C., C. M. Fernandes, S. Ongena, and A. Scopelliti (2023): "Bank bond holdings and bail-in regulatory changes: Evidence from euro area security registers," *ECB Working Paper No*, 2758.

Anand, K., I. Van Lelyveld, Á. Banai, S. Friedrich, R. Garratt, G. Hałaj, J. Fique, I. Hansen, S. M. Jaramillo, H. Lee, et al. (2018): "The missing links: A global study on uncovering financial network structures from partial data," *Journal of Financial Stability*, 35, 107–119.

Anderson, J. E. and E. Van Wincoop (2003): "Gravity with gravitas: A solution to the border puzzle," *American Economic Review*, 93, 170–192.

Aragon, G. O., M. Hertzel, and Z. Shi (2013): "Why do hedge funds avoid disclosure? Evidence from confidential 13F filings," *Journal of Financial and Quantitative Analysis*, 48, 1499–1518.

Arrata, W., B. Nguyen, I. Rahmouni-Rousseau, and M. Vari (2020): "The scarcity effect of QE on repo rates: Evidence from the euro area," *Journal of Financial Economics*, 137, 837–856.

Attinà, C. A. and P. Bologna (2021): "TLAC-eligible debt: Who holds it? A view from the euro area," *Bank of Italy Occasional Paper No*, 604.

Bagattini, G., F. Fecht, and A. Maddaloni (2023): "Liquidity support and distress resilience in bank-affiliated mutual funds," *SAFE Working Paper No*, 385.

Balazsi, L., L. Matyas, and T. Wansbeek (2018): "The estimation of multidimensional fixed effects panel data models," *Econometric Reviews*, 37, 212–227.

Barbiero, F., G. Schepens, and J.-d. Sigaux (2024): "Liquidation value and loan pricing," *Journal of Finance*, 79, 95–128.

Beck, R., A. Coppola, A. J. Lewis, M. Maggiori, M. Schmitz, and J. Schreger (2024): "The geography of capital allocation in the euro area," *NBER Working Paper No*, 32275.

Bekaert, G. and J. Breckenfelder (2019): "The (re) allocation of bank risk," *Mimeo*.

Belloni, M., F. Kuik, and L. Mingarelli (2022): "Euro area banks' sensitivity to changes in carbon price," *ECB Working Paper No*, 2654.

Ben-David, I., F. Franzoni, R. Moussawi, and J. Sedunov (2021): "The granular nature of large institutional investors," *Management Science*, 67, 6629–6659.

Benigno, P., P. Canofari, G. Di Bartolomeo, and M. Messori (2023): "The ECB's asset purchase programme: Theory, effects, and risks," *Journal of Economic Surveys*, 37, 890–914.

Bergant, K., M. Fidora, and M. Schmitz (2020): "International capital flows at the security level–evidence from the ECB's asset purchase programme," *IMF Working Paper No*, 46.

Bergant, K., G. M. Milesi-Ferretti, and M. Schmitz (2023): "Cross-border investment in emerging market bonds: stylized facts and security-level evidence from Europe," *CPER Discussion Paper No*, 17946.

Berninger, M., F. Kiesel, and J. Schnitzler (2024): "Commercial data in financial research," *Review of Corporate Finance*, 4, 293–335.

Boermans, M. A. (2023): "Preferred habitat investors in the green bond market," *Journal of Cleaner Production*, 421, 138365.

Boermans, M. A., M. J. Bun, and Y. van der Straten (2024a): "Funding the fittest?: Pricing of climate transition risk in the corporate bond market," *DNB Working Paper No*, 797.

Boermans, M. A. and J. D. Burger (2023): "Fickle emerging market flows, stable euros, and the dollar risk factor," *Journal of International Economics*, 142, 103730.

Boermans, M. A., T. Carrera de Souza, and R. Vermeulen (2024b): "Quantitative easing and preferred habitat investors in the euro area bond market," *DNB Working Paper No*, XX.

Boermans, M. A., I. Cooper, P. Sercu, and R. Vanpée (2022): "Foreign bias in equity portfolios: Informational advantage or familiarity bias?" *DNB Working Paper No*, 742.

Boermans, M. A., J. Frost, and S. S. Bisschop (2016): "European bond markets: Do illiquidity and concentration aggravate price shocks?" *Economics Letters*, 141, 143–146.

Boermans, M. A. and R. Galema (forthcoming): "Carbon home bias of European investors," *Journal of Corporate Finance*.

Boermans, M. A., R. Galema, A. Plantinga, and B. Scholtens (2024c): "The Zelensky Moment: Investments in the defense industry and the Russian invasion," *Investments and the Russian Invasion (November 24, 2023)*.

Boermans, M. A. and E. Jacobs (2024): "Green bond, green promise?" *Mimeo*.

Boermans, M. A. and V. Keshkov (2018): "The impact of the ECB asset purchases on the European bond market structure: Granular evidence on ownership concentration," *DNB Working Paper No*, 590.

Boermans, M. A. and B. van der Kroft (2024): "Capital regulation induced reaching for systematic yield: Financial instability through fire sales," *Journal of Banking & Finance*, 158, 107030.

Boermans, M. A. and S. Van Wijnbergen (2018): "Contingent convertible bonds: Who invests in European CoCos?" *Applied Economics Letters*, 25, 234–238.

Boermans, M. A. and R. Vermeulen (2016): "Newton meets Van Leeuwenhoek: Identifying international investors' common currency preferences," *Finance Research Letters*, 17, 62–65.

——— (2020): "International investment positions revisited: Investor heterogeneity and individual security characteristics," *Review of International Economics*, 28, 466–496.

Borgers, A., J. Derwall, K. Koedijk, and J. Ter Horst (2015): "Do social factors influence investment behavior and performance? Evidence from mutual fund holdings," *Journal of Banking & Finance*, 60, 112–126.

Brand, C., L. Ferrante, and A. Hubert de Fraisse (2019): "From cash-to securities-driven euro area repo markets: The role of financial stress and safe asset scarcity," *ECB Working Paper No*, 2232.

Breckenfelder, J. and V. De Falco (2023): "Investor Heterogeneity and Large-Scale Asset Purchases," *Mimeo*.

Breckenfelder, J. and V. Ivashina (2021): "Bank balance sheet constraints and bond liquidity," *ECB Working Paper No*, 2589.

Breckenfelder, J. and G. Schepens (2022): "Non-bank liquidity provision to firms: Fund runs and central bank interventions," *Mimeo*.

Bubeck, J., A. Maddaloni, and J.-L. Peydró (2020): "Negative monetary policy rates and systemic banks' risk-taking: Evidence from the euro area securities register," *Journal of Money, Credit and Banking*, 52, 197–231.

Calvet, L. E., J. Y. Campbell, F. Gomes, and P. Sodini (2021): "The cross-section of household preferences," *NBER Working Paper No*, 28788.

Carvalho, D. (2022): "The portfolio holdings of euro area investors: Looking through investment funds," *Journal of International Money and Finance*, 120, 102519.

Carvalho, D. and M. Schmitz (2023): "Shifts in the portfolio holdings of euro area investors in the midst of COVID-19: Looking-through investment funds," *Review of International Economics*, 31, 1641–1687.

——— (2025): "Brexit, what Brexit? Euro area portfolio exposures to the United Kingdom since the Brexit referendum," *International Finance*.

Coppola, A., M. Maggiori, B. Neiman, and J. Schreger (2021): "Redrawing the map of global capital flows: The role of cross-border financing and tax havens," *Quarterly Journal of Economics*, 136, 1499–1556.

Craig, B. R., M. Karamysheva, and D. Salakhova (2023): "Do market-based networks reflect true exposures between banks?" *ECB Working Paper No*, 2867.

Cremers, K. M., J. A. Fulkerson, and T. B. Riley (2019): "Challenging the conventional wisdom on active management: A review of the past 20 years of academic literature on actively managed mutual funds," *Financial Analysts Journal*, 75, 8–35.

Darmouni, O. and M. Papoutsi (2022): "The rise of bond financing in Europe," *Journal of Economic Dynamics and Control*.

de Haan, L. and R. Vermeulen (2021): "Sovereign debt ratings and the country composition of cross-border holdings of euro area sovereign debt," *Journal of International Money and Finance*, 119, 102473.

Del Guercio, D. and P. A. Tkac (2008): "Star power: The effect of monrningstar ratings on mutual fund flow," *Journal of Financial and Quantitative Analysis*, 43, 907–936.

Del Vecchio, L., C. Giglio, F. Shaw, G. Spanò, and G. Cappelletti (2022): "A sensitivities based CoVaR approach to assets commonality and its application to SSM banks," *ECB Working Paper No*, 2725.

Della Corte, V. and R. Santioni (2023): "The performance of household-held mutual funds: Evidence from the euro area," *Bank of Italy Working Paper No*, 1426.

Di Iasio, G., S. Alogoskoufis, S. Kordel, D. Kryczka, G. Nicoletti, and N. Vause (2022): "A model of system-wide stress simulation: market-based finance and the Covid-19 event," *Bank of Italy Occasional Paper No*, 687.

Dick-Nielsen, J. (2014): "How to clean enhanced TRACE data," *Mimeo*.

Dötz, N. and M. A. Weth (2019): "Redemptions and asset liquidations in corporate bond funds," *Deutsche Bundesbank Discussion Paper No*, 11.

Du, W., A. W. Huber, et al. (2024): "Dollar asset holding and hedging around the globe," *Mimeo*.

Dubiel-Teleszynski, T., F. Franch, G. Fukker, D. Miccio, M. Pellegrino, and M. Sydow (2022): "System-wide amplification of climate risk," *Mimeo*.

ECB (2015): "Who holds what? New information on securities holdings," *ECB Economic Bulletin*, 2, 72–84.

Edison, H. J. and F. E. Warnock (2004): "US investors' emerging market equity portfolios: A security-level analysis," *Review of Economics and Statistics*, 86, 691–704.

Ehrenbergerová, D., S. Malovaná, and C. Mendicino (2023): "How Do Climate Policies Affect Holdings of Green and Brown Firms' Securities?" *Czech National Bank Working Paper No*, 11.

Elsayed, M., D. Henricot, J. Idier, and T. Piquard (2023): "Habitat sweet habitat: The heterogeneous effects of Eurosystem asset purchase programs," *Mimeo*.

Elton, E. J., M. J. Gruber, and C. R. Blake (2001): "A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases," *Journal of Finance*, 56, 2415–2430.

Emiris, M., J. Harris, and F. Koulischer (2024a): "The effect of environmental preferences on investor responses to ESG disclosure," *Mimeo*.

——— (2024b): "Regulating ESG Disclosure," *Mimeo*.

Eser, F., W. Lemke, K. Nyholm, S. Radde, and A. Vladu (2023): "Tracing the impact of the ECB's asset purchase programme on the yield curve," *International Journal of Central Banking*, 19.

Fache Rousová, L. and A. R. Caloca (2015): "The use of securities holdings statistics (SHS) for designing new euro area financial integration indicators," *IFC Bulletin*, 39.

Fache Rousová, L., M. Ghio, D. Salakhova, G. V. Bauer, and A. Patozi (2023): "Derivative margin calls: A new driver of MMF flows?" *Mimeo*.

Fache Rousová, L. and M. Giuzio (2019): "Insurers' investment strategies: pro-or counter-cyclical?" *Mimeo*.

Faia, E., K. K. Lewis, and H. Zhou (2024): "Do investor differences impact monetary policy spillovers to emerging markets?" *NBER Working Paper No*, 32986.

Faia, E., J. Salomao, and A. V. Veghazy (2022): "Granular investors and international bond prices: Scarcity-induced safety," *CEPR Discussion Paper No*, 17454.

Fang, X., B. Hardy, and K. K. Lewis (2022): "Who holds sovereign debt and why it matters," *NBER Working Paper No*, 30087.

Fay, C. and A. Ghiselli (2023): "Insurers' investment behaviour and the coronavirus (COVID-19) pandemic," *ESRB Occasional Paper No*, 22.

Ferdinandusse, M., M. Freier, and A. Ristiniemi (2020): "Quantitative easing and the price-liquidity trade-off," *ECB Working Paper No*, 2399.

Fourel, V. and A. Schwenninger (2024): "The impact of the PEPP on the corporate commercial paper market," *Banque de France Working Paper No*, 946.

Franklin, J. (2020): "PRIMER: Schuldschein," *International Financial Law Review*.

Fratzscher, M. (2012): "Capital flows, push versus pull factors and the global financial crisis," *Journal of International Economics*, 88, 341–356.

Fricke, D., S. Greppmair, and K. Paludkiewicz (2024): "You can't always get what you want (where you want it): Cross-border effects of the US money market fund reform," *Journal of International Economics*, 147, 103846.

Fricke, D., S. Jank, and C. Meinerding (2022a): "Who pays the greenium?" *Mimeo*.

Fricke, D., S. Jank, and H. Wilke (2022b): "Who creates and who bears flow externalities in mutual funds?" *Deutsche Bundesbank Working Paper No*, 41.

Fukker, G., M. Kaijser, L. Mingarelli, and M. Sydow (2022): "Contagion from market price impact: A price-at-risk perspective," *ECB Working Paper No*, 2692.

Gabaix, X., R. S. Koijen, R. Richmond, and M. Yogo (2024): "Asset embeddings," *Mimeo*.

Gil-Bazo, J. and R. Santioni (2024): "Geographic shareholder dispersion and mutual fund flow risk," *Mimeo*.

Gomes, F., C. Peng, O. Smirnova, and N. Zhu (2022): "Reaching for yield: Evidence from households," *Mimeo.*

Greppmair, S. and S. Jank (2022): "Securities lender of last resort: On the causal effects of central banks' securities lending facilities," *Mimeo.*

Hartung, B. (2024): "Liquidity transformation and Eurosystem credit operations," *ECB Working Paper No.*, 2933.

Henricot, D. and T. Piquard (2022): "Credit default swaps and credit risk reallocation," *Mimeo.*

Hu, G., K. M. Jo, Y. A. Wang, and J. Xie (2018): "Institutional trading and Abel Noser data," *Journal of Corporate Finance*, 52, 143–167.

Hudepohl, T. (2022): "The rebalancing channel of QE: New evidence at the security level in the euro area," *DNB Working Paper No*, 756.

Hüser, A.-C., G. Hałaj, C. Kok, C. Perales, and A. van der Kraaij (2018): "The systemic implications of bail-in: A multi-layered network approach," *Journal of Financial Stability*, 38, 81–97.

Jansen, K. A. (2024): "Long-term investors, demand shifts, and yields," *Review of Financial Studies*, hhae071.

Jansen, K. A., H. S. Shin, and G. von Peter (2023): "Which exchange rate matters to global investors?" *Mimeo.*

Jasova, M., L. Laeven, C. Mendicino, J.-L. Peydró, and D. Supera (2021): "Systemic risk and monetary policy: The haircut gap channel of the lender of last resort," *Mimeo.*

Jochem, A. and E. Lecomte (2024): "Risky sovereign bond holdings by commercial banks in the euro area: Do safe assets availability and differences in bank funding costs play a role?" *Deutsche Bundesbank Discussion Paper.*

Jourde, T. (2022): "The rising interconnectedness of the insurance sector," *Journal of Risk and Insurance*, 89, 397–425.

Jourde, T. and K. Koné (2023): "Climate change and financial stability: A risk assessment of investment funds," *Mimeo*.

Kabaca, S., R. Maas, K. Mavromatis, and R. Priftis (2023): "Optimal quantitative easing in a monetary union," *European Economic Review*, 152, 104342.

Kaufmann, C., M. Storz, and J. Leyva (2024): "Insurance corporations' balance sheets, financial stability and monetary policy," *ECB Working Paper No*, 2892.

Kliatskova, T., L. B. Savatier, and M. Schmidt (2023): "Insolvency regimes and cross-border investment decisions," *Journal of International Money and Finance*, 131, 102795.

Koijen, R. S., F. Koulischer, B. Nguyen, and M. Yogo (2021): "Inspecting the mechanism of quantitative easing in the euro area," *Journal of Financial Economics*, 140, 1–20.

Kubitza, C., J.-D. Sigaux, and Q. Vandeweyer (2024): "The implications of CIP deviations for international capital flows," *Fama-Miller Working Paper, Chicago Booth Research Paper*, 24–18.

Lamas, M. and D. Martínez-Miera (2021): "Sectorial holdings and stock prices: The household-bank nexus," *Banco de Espana Working Paper*.

Lambert, C., L. M. Vivar, and M. Wedow (2024): "Is home bias biased? New evidence from the investment fund sector," *ECB Working Paper*, 2924.

Lane, P. R. and G. M. Milesi-Ferretti (2008): "International investment patterns," *Review of Economics and Statistics*, 90, 538–549.

Lavrador, I., R. Peronaci, and N. Silva (2012): "Security-by-security data on holdings of securities: The importance for national and euro area accounts," *IFC Bulletin*, 28, 434.

Levels, A., C. Lambert, and M. Wedow (2023): "Green bond home bias and the role of supply and sustainability preferences," *DNB Working Paper No*, 767.

Liberati, D. and G. Marinelli (2024): "Was Covid-19 a wake-up call on climate risks? Evidence from the greenium," *Bank of Italy Occasional Paper No*.

Maddaloni, A. M. and G. Scardozzi (2022): *New Bail-In Legislation*, Springer.

Marques-Ibanez, D., G. Santilli, and G. Scardozzi (2024): "Bail-in in action," *Economics Letters*, 111764.

Menkveld, A. J., A. Dreber, F. Holzmeister, J. Huber, M. Johannesson, M. Kirchler, S. Neusüss, M. Razen, U. Weitzel, D. Abad-Díaz, et al. (2024): "Nonstandard errors," *Journal of Finance*, 79, 2339–2390.

Mésonnier, J.-S. and B. Nguyen (2020): "Showing off cleaner hands: Mandatory climate-related disclosure by financial institutions and the financing of fossil energy," *Banque de France Working Paper No*, 800.

Mink, M., R. Ramcharan, and I. van Lelyveld (2020): "How banks respond to distress: Shifting risks in Europe's banking union," *DNB Working Paper No*, 669.

Montagna, M., G. Torri, and G. Covi (2020): "On the origin of systemic risk," *Mimeo*.

Nagengast, A. and Y. V. Yotov (2023): "Staggered difference-in-differences in gravity settings: Revisiting the effects of trade agreements," *Deutsche Bundesbank Discussion Paper No.*, 30.

Nguyen, B., D. Tomio, and M. Vari (2023): "Safe asset scarcity and monetary policy transmission," *Mimeo*.

Nicoletti, G., J. Rariga, and C. R. d'Acri (2024): "Spare tyres with a hole: investment funds under stress and credit to firms," *ECB Working Paper*.

Papoutsi, M., M. Piazzesi, and M. Schneider (2021): "How unconventional is green monetary policy," *Mimeo*.

Parida, S. and T. Teo (2018): "The impact of more frequent portfolio disclosure on mutual fund performance," *Journal of Banking & Finance*, 87, 427–445.

Pietsch, A. and D. Salakhova (2022): "Pricing of green bonds: Drivers and dynamics of the greenium," *ECB Working Paper No*, 2728.

Radke, T., A. R. Caloca, and M. Schmitz (2021): "The more the merrier: Enhancing traditional cross-border portfolio investment statistics using security-by-security information," *ISI World Statistics Congress*, 11.

Ringe, W.-G. and J. Patel (2019): "The dark side of bank resolution: Counterparty risk through bail-in," *Mimeo*.

Schumacher, D. (2018): "Home bias abroad: Domestic industries and foreign portfolio choice," *Review of Financial Studies*, 31, 1654–1706.

Schwarz, C. G. and M. E. Potter (2016): "Revisiting mutual fund portfolio disclosure," *Review of Financial Studies*, 29, 3519–3544.

Sola, P. and F. Strobbe (2010): "Addressing data gaps revealed by the financial crisis: European Central Bank statistics on holdings of securities," *IFC Bulletin*, 25, 141.

Sydow, M., T. Dubiel-Teleszynski, F. Franch, G. Fukker, S. Gallet, H. Gründl, S. Kotronis, D. Miccio, M. Pellegrino, S. Schlütter, et al. (2024a): "Banks and non-banks stressed: Liquidity shocks and the mitigating role of insurance companies," *Mimeo*.

Sydow, M., A. Schilte, G. Covi, M. Deipenbrock, L. Del Vecchio, P. Fiedor, G. Fukker, M. Gehrend, R. Gourdel, A. Grassi, et al. (2024b): "Shock amplification in an interconnected financial system of banks and investment funds," *Journal of Financial Stability*, 101234.

Tzamourani, P. (2021): "The interest rate exposure of euro area households," *European Economic Review*, 132, 103643.

Ullersma, C. and I. van Lelyveld (2021): "Granular data offer new opportunities for stress testing," *Handbook of Financial Stress Testing', Cambridge University Press, Cambridge*.

Wang, D., I. van Lelyveld, and J. Schaumburg (2018): "Do information contagion and business model similarities explain bank credit risk commonalities?" *DNB Working Paper No*, 619.

Notes

1. Berninger et al. (2024) analyze commercial datasets used in finance between 2000 and 2016 across 14,000 articles, highlighting investor holdings data features prominently. Data is available from various vendors such as Center for Research in Security Prices (CRSP), EPFR Global (e.g. Fratzscher, 2012), FactSet LionShares (e.g. Schumacher, 2018; Gabaix et al., 2024), Lipper (e.g. Borgers et al., 2015), Morningstar (e.g. Del Guercio and Tkac, 2008), and, Thomson-Reuters (e.g. Ben-David et al., 2021). However, due to voluntary reporting of granular holdings data by mutual funds, there have been challenges to the data quality of these sources, see for instance (Schwarz and Potter, 2016) on CRSP, Thomson and SEC 13F filings and Elton et al. (2001) on CRSP and Morningstar. Abel Noser data provides high-frequency data from a broker which includes transaction costs data from institutional clients (see Hu et al., 2018, for a literature review). Other examples of granular portfolio holdings data include analysis of central bank portfolios, data from specific custodians with client-level positions (e.g. Gomes et al., 2022), or panel data from household portfolios, (see e.g. Calvet et al., 2021, for Sweden).

2. For an historical overview on the SHS database, see the works published for the BIS Irving Fisher Committee of Sola and Strobbe (2010) who discuss the data gaps initiative, Lavrador et al. (2012) who discuss the collection of SHS data in the context of national accounts, and Fache Rousová and Caloca (2015) who analyze financial integration and compare SHS data with other sources. Radke et al. (2021) outline how official statistics on international investment benefit from holdings data at the security level.

3. In general, over time the SHS-S file size has gradually increased hence to estimate the total number of observations of raw files between 2013Q4 to 2024Q1 of 2.3 billion (23.8+7.6+9.8+8.3 * 45 periods) somewhat overstates the expected value.

4. When dropping bonds without an amount outstanding, already 495,445 observations are removed, representing 2.6% of the total holdings after cleaning. Setting the threshold to 1 million euros would additionally remove 661,914 observations, but now only 0.7% additional holdings at market value are removed, representing 94.1% coverage compared to the prepared SHS data without cleaning.

5. In an early version of this practitioner's guide (up to December 2022), we proposed setting a filter on holdings of at least 10,000 euros to exclude small positions at the investor sector by investor country level for a given individual security. This threshold amount is, however, sensitive to the data aggregation over the investor sectors and countries, a topic discussed in Section *D*. For example, several papers aggregate all euro area holdings at the security-time level, thus here a suitable minimum threshold will be larger to have any material impact on the number of observations dropped compared to our proposed cleaning which retains investor sector and investor country details. Still, for many applications a threshold is useful to ensure that small holdings do not dominate the sample in terms of number of observations giving highly non-linear distributions of holdings. One may set a threshold on the issuance size in terms of market capitalization or amount outstanding of the individual security (which of course, needs to be at least larger than any holdings amount threshold to be consequential). For example,

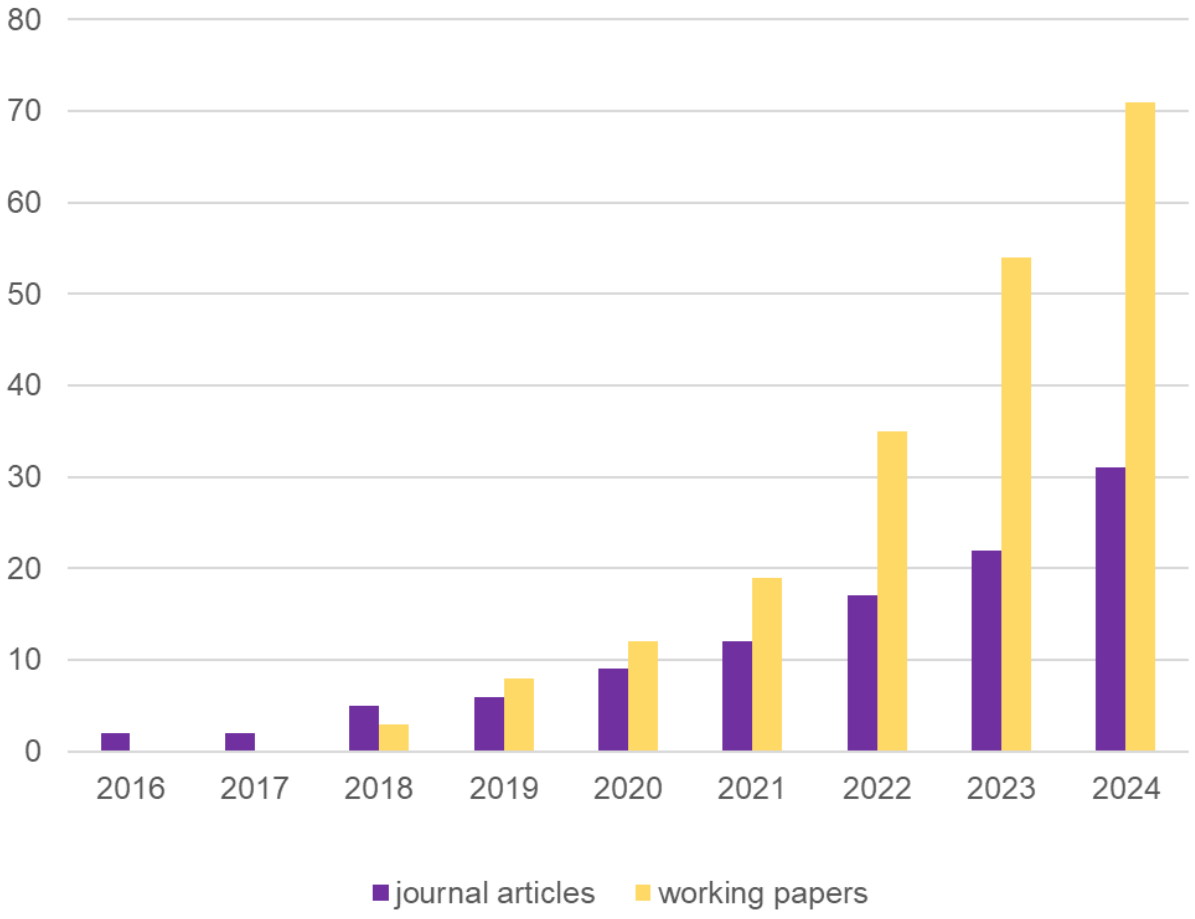
Boermans and Burger (2023) exclude securities with market values of less than 5 million euros, Boermans and van der Kroft (2024) use a 10 million euros threshold and Boermans et al. (2016) set a floor of 100 million euros.

6. While inclusion of interaction terms between the explanatory variables and the investor sectors is preferable in most contexts, the size of the number of variables included in any such regressions typically becomes large. If not, one may further include issuer FE at entity level (firm/government) or even security FE (see also Balazsi et al., 2018, on multidimensional panel data) and Nagengast and Yotov (2023) on biases in the empirical literature.

7. In Appendix A we repeat the same analysis for the raw SHS-S file for holdings at market value across euro area investors, without further cleaning. While Table A2 provides robust estimates with similar signs, it highlights how strongly small holdings shape the estimated coefficients.

8. Our main time-series panel regressions from Table 6 are similar when we apply weighted regressions with the lagged dependent variable as analytical weight. One notable downside of applying such weights is that one loses many observations not randomly because newly issued bonds are consistently dropped.

Figure 1 – The growth of SHS research over time (cumulative count by category)



Figures

Tables

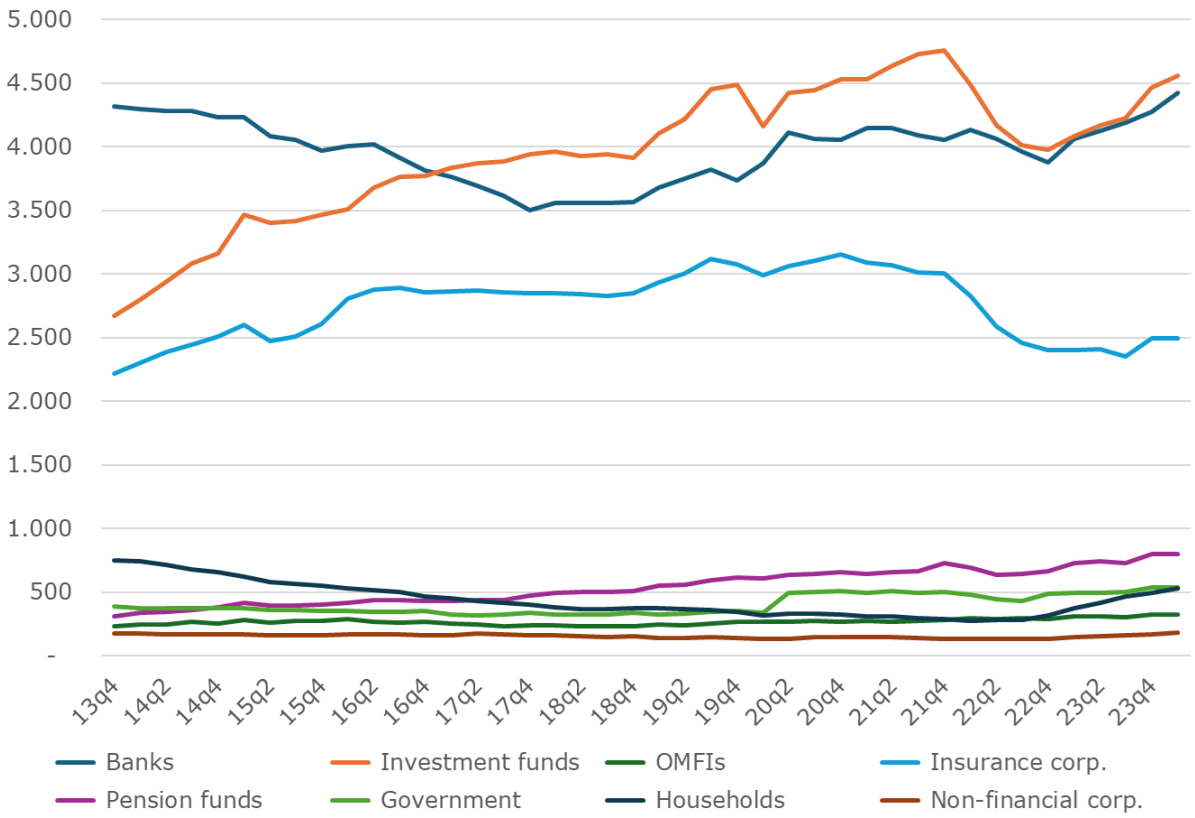
Table 1 – Overview of SHS journal publications, sorted by ABS score

ID	Author(s)	Journal	ABS	Impact	Cite	Theme	Year
1	Arrata et al. (2020)	*JFE*	4*	10.4	73	MP	2020
2	Koijen et al. (2021)	*JFE*	4*	10.4	196	MP	2021
3	Barbiero et al. (2024)	*JF*	4*	7.6	2	BF	2024
4	Jansen (2024)	*RFS*	4*	6.8	33	IIP	2024
5	Bubeck et al. (2020)	*JMCB*	4	2.9	115	BF	2020
6	Albertazzi et al. (2021)	*JFI*	4	3.1	49	MP	2021
7	Boermans and Burger (2023)	*JIE*	4	3.8	8	IIP	2023
8	Fricke et al. (2024)	*JIE*	4	3.8	1	FM	2024
9	Boermans and Galema (forthcoming)	*JCF*	4	7.2	20	SF	2024
10	Boermans et al. (2016)	*EL*	3	2.1	19	FM	2016
11	Anand et al. (2018)	*JFS*	3	6.1	153	BF	2018
12	Hüser et al. (2018)	*JFS*	3	6.1	92	BF	2018
13	de Haan and Vermeulen (2021)	*JIMF*	3	2.8	9	IIP	2021
14	Carvalho (2022)	*JIMF*	3	2.8	11	IIP	2022
15	Aldasoro et al. (2022)	*JEDC*	3	1.9	11	BF	2022
16	Jourde (2022)	*JRI*	3	2.1	12	BF	2022
17	Alessi and Battiston (2022)	*IRFA*	3	7.5	70	SF	2022
18	Kabaca et al. (2023)	*EER*	3	2.8	16	MP	2023
19	Darmouni and Papoutsi (2022)	*JEDC*	3	1.9	69	FM	2022
20	Eser et al. (2023)	*IJCB*	3	1.3	148	MP	2023
21	Kliatskova et al. (2023)	*JIMF*	3	2.8	9	FM	2023
22	Boermans and van der Kroft (2024)	*JBF*	3	3.6	7	BF	2024
23	Sydow et al. (2024b)	*JFS*	3	6.1	33	BF	2024
24	Marques-Ibanez et al. (2024)	*EL*	3	2.1	1	BF	2024
25	Tzamourani (2021)	*EER*	3	2.8	45	BF	2024
26	Alessi et al. (2024)	*JFS*	3	6.1	35	SF	2024
27	Boermans and Vermeulen (2016)	*FRL*	2	7.4	19	IIP	2016
28	Boermans and Vermeulen (2020)	*RoIE*	2	1.9	35	IIP	2020
29	Boermans (2023)	*JCLEP*	2	9.7	20	SF	2023
30	Carvalho and Schmitz (2023)	*RoIE*	2	1.9	15	IIP	2023
31	Boermans and Van Wijnbergen (2018)	*AEL*	N/A	1.6	33	BF	2018

Notes: Sorted by ABS scores, year and first author. Journal abbreviations can be found in the References in Section VII of the paper. ABS refers to ABS scores or Academic Journal Guide 'AJG' and are taken from the Chartered Association of Business Schools rankings as on June 2024 (based on 2021 reports from JournalRanking). The ABS list groups journals into "fields", such as accounting and finance, and then rates journals on a scale from 1 (low quality) to 4 (highest quality). Impact refers to the impact factors taken from Clarivate Analytics, June 2024 reports. Cite reflects the number of citations retrieved from Google Scholar on 31 December 2024. The Themes are based on the following groups: 'MP' = Monetary Policy, 'B&F' = Banking & Finance, 'FM' = Financial Markets, 'IIP' - International Investment Positions, 'SF' = Sustainable Finance.*

Figure 2 – SHS research word cloud (from 102 paper titles in www.wordclouds.com)

Figure 3 – Euro area bond holdings by investor sector m (in EUR bln)



Notes: Investor sector groupings are based on Table 4 and euro area bond holdings are taken from the cleaned SHS-S data based on the practitioner's guide procedure from Section V.

Table 2 – Thematic review of SHS journal publications (By theme, chronological order)

ID	Author(s)	Short title	SHS Module	Journal
I. Banking & Finance:				
11	Anand et al. (2018)	A global study on uncovering financial network structures from partial data	SHS-G	*JFS*
31	Boermans and Van Wijnbergen (2018)	Contingent convertible bonds: Who invests in European CoCos?	SHS-S	*AEL*
12	Hüser et al. (2018)	The systemic implications of bail-in: a multi-layered network approach	SHS-G	*JFS*
5	Bubeck et al. (2020)	Negative monetary policy rates and systemic banks' risk-taking	SHS-G	*JMCB*
15	Aldasoro et al. (2022)	Contagion accounting in stress-testing	SHS-G	*JEDC*
16	Jourde (2022)	The rising interconnectedness of the insurance sector	SHS-S	*JRI*
22	Boermans and van der Kroft (2024)	Capital regulation induced reaching for systematic yield	SHS-S	*JBF*
23	Sydow et al. (2024b)	Shock amplification in an interconnected financial system	SHS-G	*JFS*
3	Barbiero et al. (2024)	Liquidation value and loan pricing	SHS-G	*JF*
24	Marques-Ibanez et al. (2024)	Bail-in in action	SHS-S	*EL*
25	Tzamourani (2021)	The interest rate exposure of euro area households	SHS-S	*EER*
II. International Investment:				
27	Boermans and Vermeulen (2016)	Identifying international investors' common currency preferences	SHS-S	*FRL*
28	Boermans and Vermeulen (2020)	International investment positions revisited	SHS-S	*RoIE*
13	de Haan and Vermeulen (2021)	Sovereign debt ratings and the country composition of cross-border holdings	SHS-S	*JIMF*
14	Carvalho (2022)	The portfolio holdings of euro area investors: Looking through investment funds	SHS-S	*JIMF*
7	Boermans and Burger (2023)	Fickle emerging market flows, stable euros, and the dollar risk factor	SHS-S	*JIE*
30	Carvalho and Schmitz (2023)	Shifts in the portfolio holdings of euro area investors in the midst of COVID-19	SHS-S	*RoIE*
4	Jansen (2024)	Long-term investors, demand shifts, and yields	SHS-S	*RFS*
III. Monetary Policy:				
1	Arrata et al. (2020)	The scarcity effect of QE on repo rates: Evidence from the euro area	SHS-S	*JFE*
6	Albertazzi et al. (2021)	Portfolio rebalancing and the transmission of large-scale asset purchase programs	both	*JFI*
2	Koijen et al. (2021)	Inspecting the mechanism of quantitative easing in the euro area	SHS-S	*JFE*
20	Eser et al. (2023)	Tracing the impact of the ECB's asset purchase programme on the yield curve	SHS-S	*IJCB*
19	Kabaca et al. (2023)	Optimal quantitative easing in a monetary union	SHS-S	*EER*
IV. Financial Markets:				
10	Boermans et al. (2016)	European bond markets: Do illiquidity and concentration aggravate price shocks?	SHS-S	*EL*
19	Darmouni and Papoutsi (2022)	The rise of bond financing in Europe	SHS-S	*JEDC*
21	Kliatskova et al. (2023)	Insolvency regimes and cross-border investment decisions	SHS-S	*JIMF*
8	Fricke et al. (2024)	Cross-border effects of the US money market fund reform	SHS-S	*JIE*
V. Sustainable Finance:				
17	Alessi and Battiston (2022)	Green Taxonomy alignment versus transition risk in financial portfolios	SHS-S	*IRFA*
29	Boermans (2023)	Preferred habitat investors in the green bond market	SHS-S	*JCLEP*
26	Alessi et al. (2024)	Over with carbon? Investors' reaction to the Paris Agreement	SHS-S	*JFS*
9	Boermans and Galema (forthcoming)	Carbon home bias of European investors	SHS-S	*JCF*

Notes: ID based on ordering by ABS scores as in Table 1. SHS Module refers to the SHS-S for the Sectoral data and SHS-G for the individual banking group data.

Table 3 – Data selection criteria to be reported with SHS research

ID	SHS Module	I. Amount type	II. Valuation	III. Euro area	IV. Third-party
1	SHS-S	T	M	NR	NR
2	SHS-S	LE	NR	X	NR
4	SHS-S	LE	NR	NR	NR
7	SHS-S	LE+T	M	X	NR
8	SHS-S	LE	M	X	X
9	SHS-S	LE	M	X	X
10	SHS-S	LE	N	X	NR
13	SHS-S	LE	N	X	NR
14	SHS-S	LE	M	X	NR
16	SHS-S	LE	NR	X*	NR
17	SHS-S	NR	NR	NR	NR
18	SHS-S	LE	NR	NR	NR
19	SHS-S	LE	NR	X	NR
20	SHS-S	LE	N	NR	NR
21	SHS-S	LE	M+N	X	NR
22	SHS-S	LE	M	X	NR
24	SHS-S	LE	N	NR	NR
25	SHS-S	LE	NR	NR	NR
26	SHS-S	LE	M	NR	NR
27	SHS-S	LE	M	X	NR
28	SHS-S	LE	M	X	X
29	SHS-S	LE	M	X	NR
30	SHS-S	LE	M	NR	NR
31	SHS-S	LE	M	NR	NR
6	both	LE	M	NR	NR
3	SHS-G	LE	NR	N/A	N/A
5	SHS-G	LE	N	N/A	N/A
11	SHS-G	LE	NR	N/A	N/A
12	SHS-G	LE	NR	N/A	N/A
15	SHS-G	LE	NR	N/A	N/A
23	SHS-G	LE	NR	N/A	N/A

Notes: Column 'ID' is based on ordering by ABS scores as in Table 1 and sorted by SHS Module. 'SHS Module' refers to the SHS-S for the Sectoral data and SHS-G for the individual banking group data. Column 'I. Amount type' includes positions 'LE' and net financial transaction 'T'. Column 'II. Valuation' includes 'M' for market values and 'N' is nominal values. 'NR' means Not Reported in the journal publication. The Column 'III/ Euro area' includes 'X' meaning the paper explicitly states that it filters on euro area investors (and 'X' focuses on 14 of the 20 euro area countries), the 'Y' means the paper implicitly uses euro area investors only. The Column 'IV. Third-party' shows if the paper has addressed any data selection for the third-party holdings. 'N/A' means not applicable.*

Table 4 – Data aggregations by Investor sector for SHS-S research

Investor sector	**Guide**	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Financial:																																
Banks	X	X	X	X	X	X	X			X			X	X	X	X				X	X	X	X	X	X		X		X	X	X	X
Money market fd.							X														X											
Investment funds		X	X		X									X	X					X	X		X				X		X	X	X	X
OMFI	X						X			X					X					X							X			X	X	X
Insurance corp.	X						X			X				X	X		X						X				X			X	X	X
Pension funds	X						X			X				X	X								X				X			X	X	X
Non-financial:																																
Households		X	X				X								X						X				X		X		X	X	X	X
Non-profit							X																									
NFC	X	X								X					X							X			X					X	X	X
Government	X						X			X										X										X		
Groupings:																																
MMF+IF	X									X																						
Banks + MMF									X																							
IC+PF		X	X		X															X	X								X			
Instit. invest.																						X										
IF+IC+PC									X																							
Non-banks																									X							
HH+NPO	X									X												X										
Other investors:																																
ECB		X	X																	X												
RoW (resid.)		X	X		X		X		X											X	X											X
Other			X		X			X	X												X								X			
Not reported											X	X						X	X									X				
N sectors	8	7	7	1	5	1	9	1	4	8	?	?	1	4	7	1	1	?	?	7	7	4	4	1	4	1	6	?	5	8	7	8

*Notes: Column numbers refer to the study IDs based on Table 1, with the Column '**Guide**' denoted the sector aggregation based on this practitioner's guide. Investor sectors are based on the European System of Accounts (ESA 2010), covering Monetary Financial Institutions (MFIs) or simply 'Banks' (S.122), Money market funds 'MMF' (S.123), Investment funds 'IF', mostly mutual funds (S.124), Insurance corporations 'IC' (S.128), Pension funds 'PF' (S.129), Other Financials (S.125+S.126+S.127), Non-financial institutions 'NFC' (S.11), Households 'HH' (S.14), Non-profit organizations 'NPO' (S.15), Institutional Investors (S123-129), Rest of the world (ROW) as residual.*

Table 5 – Cross-sectional regressions: Bond demand functions by investor sector

	(1)	(2)	(3)	(4)				(5)		
	EA	EA_j	EA_s	EA_sj	main	bank	omfi	insr	pfnd	hhld
EUR	2.51***	2.16***	2.17***	1.21***	1.13***	0.69*	-0.35	0.67**	-0.30	-0.28
	[0.023]	[0.010]	[0.012]	[0.107]	[0.185]	[0.376]	[0.350]	[0.292]	[0.234]	[0.291]
$lnSIZE$	0.99***	0.63***	0.72***	0.39***	0.44***	0.17	-0.21**	-0.09	0.02	-0.20**
	[0.007]	[0.004]	[0.004]	[0.040]	[0.071]	[0.112]	[0.095]	[0.079]	[0.116]	[0.098]
Constant	-2.39***	0.81***	0.89***	6.45***				1.51		
	[0.170]	[0.098]	[0.112]	[0.837]				[1.982]		
Obs.	57,900	306,092	206,226	635,453				635,453		
Adj. R^2	0.556	0.407	0.464	0.357				0.366		
HC FE	N/A	Yes	N/A	Yes				Yes		
HS FE	N/A	N/A	Yes	Yes				Yes		
IC FE	Yes	Yes	Yes	Yes				Yes		

*Notes: Dependent variable $lnHOLD$ with modeling based on Equation 6. Reference period 2024-Q1 with 68,373 unique bonds across 20 holder countries and 8 investor sectors. Robust standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. Fixed effects (FE) are based on a combination of HolderCounty (HC), HolderSector (HS) and IssuerCountry (IC), where N/A means not applicable given the level of aggregation of the respective column. Columns (2), (4) and (5) include 20 holder countries and Columns (3) to (5) include eight different investor sectors, with security specific interactions with dummies of investor sectors Monetary Financial Institutions (MFIs) or 'bank', Other Financials 'omfi', insurance corporations 'insr', pension funds 'pfnd' and households 'hhld', see also Table 4 and section D.1 on grouping investor sectors.*

Table 6 – Multidimensional panel time-series regression results

	(1)			(2)			
	EA_sjt	main	bank	omfi	insr	pfnd	hhld
EUR	1.100***	0.983***	0.642*	-0.603**	0.881***	-0.118	-0.251
	[0.112]	[0.188]	[0.355]	[0.273]	[0.310]	[0.208]	[0.293]
$lnSIZE$	0.379***	0.457***	0.007	-0.262***	-0.098	0.032	-0.291***
	[0.039]	[0.064]	[0.100]	[0.085]	[0.069]	[0.098]	[0.084]
Constant	6.746***			4.510***			
	[0.879]			[1.700]			
Observations	20,112,593			20,112,593			
R-squared	0.339			0.371			
HC FE	Yes			Yes			
HS FE	Yes			Yes			
IC FE	Yes			Yes			
Time FE	Yes			Yes			

*Notes: Dependent variable $lnHOLD$, with the specifications based on Equation 4. Reference period 2013-Q4 to 2024-Q1. Robust standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. Fixed effects (FE) are based on a combination of $HolderCounty$ (HC), $HolderSector$ (HS) and $IssuerCountry$ (IC), where N/A means not applicable given the level of aggregation of the respective column and also include quarterly fixed effects. All estimations include 20 holder countries (filter on euro area countries) and across eight different investor sectors, with security specific interactions with dummies of investor sectors Monetary Financial Institutions (MFIs) or 'bank', Other Financials 'omfi', insurance corporations 'insr', pension funds 'pfnd' and households 'hhld', see also Table 4 and section D.1.*

Table 7 – Sustainability preferences of European investors revisited

	(1)	(2)	(3)
	Raw data	Cleaned data	Time-series clean
Green bonds (GB)	-0.01	0.15**	0.15***
	[0.080]	[0.071]	[0.053]
GB * Banks	-0.17	-0.39***	-0.03
	[0.195]	[0.135]	[0.175]
GB * Insurance	-0.16	-0.32***	-0.24***
	[0.113]	[0.102]	[0.061]
GB * Pension funds	0.03	-0.07	-0.05
	[0.087]	[0.114]	[0.077]
GB * Households	-0.18*	-0.22**	-0.29***
	[0.107]	[0.096]	[0.092]
Sustainability-linked bonds (SLB)	-0.07	0.16**	0.18**
	[0.097]	[0.80]	[0.084]
SLB * Banks	-0.74***	-1.21***	-0.16
	[0.283]	[0.234]	[0.162]
SLB * Insurance	-0.29*	-0.44***	-0.61***
	[0.157]	[0.141]	[0.142]
SLB * Pension funds	0.01	-0.06	0.02
	[0.170]	[0.180]	[0.193]
SLB * Households	0.01	-0.07	-0.20
	[0.128]	[0.128]	[0.149]
EUR	0.67***	1.12***	1.13***
	[0.196]	[0.186]	[0.179]
EUR * Banks	1.72***	0.72*	0.27
	[0.329]	[0.376]	[0.399]
EUR * Insurance	0.77***	0.70**	0.99***
	[0.221]	[0.292]	[0.320]
EUR * Pension funds	-0.46*	-0.30	-0.13
	[0.233]	[0.232]	[0.208]
EUR * Households	-0.14	-0.27	-0.02
	[0.247]	[0.292]	[0.309]
lnSIZE	0.41***	0.54***	0.61***
	[0.036]	[0.072]	[0.134]
$lnSIZE$ * Banks	-0.06	0.16	-0.34**
	[0.045]	[0.114]	[0.171]
$lnSIZE$ * Insurance	-0.19***	-0.10	-0.15
	[0.051]	[0.080]	[0.147]
$lnSIZE$ * Pension funds	-0.19***	0.02	-0.07
	[0.066]	[0.117]	[0.153]
$lnSIZE$ * Households	-0.13***	-0.24**	-0.32**
	[0.028]	[0.098]	[0.150]
Constant	3.78***	1.70	9.07***
	[0.823]	[2.030]	[2.252]
Obs.	1,652,180	635,379	6,487,957
R-squared	0.769	0.367	0.396
HC FE	Yes	Yes	Yes
HS FE	Yes	Yes	Yes
IC FE	Yes	Yes	Yes
Time FE	-	-	Yes

*Notes: Dependent variable lnHOLD, with the specifications based on Equation 4. Reference period 2024-Q1 for the cross-sectional analysis in Columns (1) and (2) and the full period 2013-Q4 to 2024-Q1 for Column (3). Robust standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. Fixed effects (FE) are based on a combination of HolderCounty (HC), HolderSector (HS) and IssuerCountry (IC). All estimations include 20 holder countries (filter on euro area countries) and across eight different investor sectors, as in Table 4 and section D.1, with the exception of Column (1) which relies on the raw data for the investor sectors.*

Appendix A Working papers by theme

In this Section we list all working papers using granular SHS data. Table A1 lists 71 papers across the themes Banking & Finance, International Investments, Monetary Policy, Financial Markets, Sustainable Finance and Others. All these papers are included in the reference list.

Table A1 – Working papers by theme (Continues on the next page)

No.	First author	Year	Short title	SHS Module
I. Banking and finance				
A1	Wang	2018	Information contagion and business model similarities	SHS-G
A2	Bekaert	2019	The (re) allocation of bank risk	SHS-S
A3	Fache Rousová	2019	Insurers' investment strategies: pro-or countercyclical?	SHS-S
A4	Ringe	2019	Counterparty risk through bail-in	SHS-S
A5	Mink	2020	Shifting risks in Europe's banking union	SHS-G
A6	Attina	2021	TLAC-eligible debt: Who holds it	SHS-S
A7	Montagna	2021	On the origin of systemic risk	SHS-G
A8	Del Vecchio	2022	A sensitivities based CoVaR approach to assets commonality	SHS-G
A9	di Iasio	2022	A model of system-wide stress simulation	SHS-S
A10	Fukker	2022	Contagion from market price impact	SHS-S
A11	Henricot	2022	Credit default swaps and credit risk reallocation	SHS-G
A12	Allaire	2023	Fund fragility: The role of investor base	SHS-S
A13	Altavilla	2023	Bank bond holdings and bail-in regulatory changes	SHS-G
A14	Fay	2023	Insurers' investment behaviour and the coronavirus	SHS-S
A15	Altavilla	2024	Determinants of bank performance: Evidence from replicating portfolios	SHS-G
A16	Craig	2024	Do market-based networks reflect true exposures between banks?	SHS-G
A17	Hartung	2024	Liquidity transformation and Eurosystem credit operations	SHS-G
A18	Jochem	2024	Risky sovereign bond holdings by commercial banks	SHS-S
A19	Kaufmann	2024	Insurance corporations' balance sheets, financial stability and monetary policy	SHS-S
A20	Nicoletti	2024	Investment funds under stress and credit to firms	SHS-S
A21	Sydow	2024	Liquidity shocks and the mitigating role of insurance companies	both
II. International investment				
A22	Carvalho	2022	Brexit, what Brexit?	SHS-S
A23	Bergant	2023	Cross-border investment in emerging market bonds	SHS-S
A24	Jansen	2023	Which exchange rates matter to global investors	SHS-S
A25	Beck	2024	Geography of capital allocation in the euro area	SHS-S
A26	Kubitza	2024	The implications of CIP deviations for international capital flows	SHS-S
A27	Lambert	2024	Is home bias biased?	SHS-S

Table A1 – Working papers by theme (continued, continues on the next page)

No.	First author	Year	Short title	SHS Module
III. Monetary policy				
A28	Boermans	2018	The impact of the ECB asset purchases on the European bond market structure	SHS-S
A29	Bergant	2020	International capital flows at the security level	SHS-S
A30	Ferdinandusse	2020	Quantitative easing and the price-liquidity trade-off.	SHS-S
A31	Jasova	2021	Systemic risk and monetary policy	SHS-S
A32	Greppmair	2022	Securities lender of last resort	SHS-S
A33	Hudepohl	2022	The rebalancing channel of QE	SHS-S
A34	Breckenfelder	2023	Investor heterogeneity and large-scale asset purchases	SHS-S
A35	Elsayed	2023	The heterogeneous effects of Eurosystem asset purchase programs	both
A36	Nguyen	2023	Safe asset scarcity and monetary policy transmission	SHS-S
A37	Boermans	2024	Quantitative easing and preferred habitat investors in the euro area bond market	SHS-S
IV. Financial markets				
A38	Brand	2019	From cash-to securities-driven euro area repo markets	SHS-S
A39	Dötz	2019	Redemptions and asset liquidations in corporate bond funds	SHS-S
A40	Breckenfelder	2021	Bank balance sheet constraints and bond liquidity	SHS-S
A41	Ahmed	2022	Foreign institutional investors, monetary policy, and reaching for yield	SHS-S
A42	Breckenfelder	2022	Non-bank liquidity provision to firms	SHS-S
A43	Faia	2022	Granular investors and international bond prices: Scarcity-induced safety	SHS-S
A44	Bagattini	2023	Liquidity support and distress resilience in bank-affiliated mutual funds.	SHS-G
A45	Fache Rousová	2023	Derivative margin calls: A new driver of MMF flows	SHS-S
A46	Fricke	2023	Who creates and who bears flow externalities in mutual funds?	SHS-S
A47	Fourel	2024	The Impact of the PEPP on the corporate commercial paper market	SHS-S
A48	Gil-Bazo	2024	Geographic shareholder dispersion and mutual fund flow risk	SHS-S

Table A1 – Working papers by theme (continued)

No.	First author	Year	Short title	SHS Module
V. Sustainable finance				
A49	Mésonnier	2020	Showing off cleaner hands	SHS-S
A50	Alogoskoufis	2021	ECB economy-wide climate stress test	SHS-G
A51	Aghion	2022	Financial markets and green innovation	SHS-S
A52	Belloni	2022	Euro area banks' sensitivity to changes in carbon price	SHS-G
A53	Dubiel-Teleszynski	2022	System-wide amplification of climate risk	both
A54	Papoutsi	2022	How unconventional is green monetary policy	SHS-S
A55	Pietsch	2022	Pricing of green bonds: Drivers and dynamics of the greenium	SHS-S
A56	Alessi	2023	Taxonomy-alignment and transition risk: A country level approach	SHS-S
A57	Ehrenbergerová	2023	How do climate policies affect holdings of green and brown firms' securities	SHS-S
A58	Fricke	2023	Who pays the greenium?	SHS-S
A59	Jourde	2023	Climate change and financial stability: A risk assessment of investment funds	SHS-S
A60	Jourde	2023	Investor exposure to climate risk through investment funds	SHS-S
A61	Levels	2023	Green bond home bias and the role of supply and sustainability preferences	SHS-S
A62	Boermans	2024	Funding the fittest	SHS-S
A63	Boermans	2024	Green bond green promise	SHS-S
A64	Emiris	2024	Effect of environmental preferences on investor responses to ESG disclosure	SHS-S
A65	Emiris	2024	Regulatig ESG disclosure	SHS-S
A66	Liberati	2024	Was Covid-19 a wake-up call on climate risks	SHS-S
VI. Other				
A67	Abidi	2021	Bright side of transparency	SHS-S
A68	Lamas	2021	Sectorial holdings and stock prices: The household-bank nexus	SHS-S
A69	Boermans	2022	Foreign bias in equity portfolios	SHS-S
A70	Della Corte	2023	The performance of household-held mutual funds	SHS-S
A71	Boermans	2024	The Zelensky Moment: Investments in the defense industry and the Russian invasion	SHS-S

Note: Working papers, updated in December 2024.

Table A2 – Cross-sectional regression for bond demand functions (raw data)

	(1)	(2)	(3)	(4)				(5)		
	EA	EA_j	EA_s	EA_sj	main	bank	omfi	insr	pfnd	hhld
EUR	1.41***	1.17***	1.44***	0.92***	0.53***	1.85***	-0.29	0.95***	-0.25	0.08
	[0.011]	[0.007]	[0.009]	[0.108]	[0.189]	[0.318]	[0.277]	[0.245]	[0.247]	[0.243]
$lnSIZE$	0.76***	0.53***	0.47***	0.32***	0.40***	-0.06	-0.22***	-0.19***	-0.19***	-0.12***
	[0.004]	[0.002]	[0.002]	[0.024]	[0.043]	[0.051]	[0.051]	[0.058]	[0.070]	[0.030]
Constant	2.43***	3.60***	3.74***	4.86***				3.92***		
	[0.098]	[0.066]	[0.068]	[0.662]				[0.944]		
Obs.	263,224	1,063,595	568,744	1,652,180				1,652,180		
Adj. R^2	0.774	0.889	0.691	0.865				0.869		
HC FE	N/A	Yes	N/A	Yes				Yes		
HS FE	N/A	N/A	Yes	Yes				Yes		
HC*HS FE	N/A	N/A	N/A	Yes				Yes		
IC FE	Yes	Yes	Yes	Yes				Yes		

*Notes: Dependent variable lnHOLD with modeling based on Equation 6. Reference period 2024-Q1 with 810,578 unique bonds, held by 20 holder countries across 18 investor sectors. Robust standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. Fixed effects (FE) are based on a combination of HolderCounty (HC), HolderSector (HS) and IssuerCountry (IC), where N/A means not applicable given the level of aggregation of the respective column. Columns (2), (4) and (5) include 20 holder countries and Columns (3) to (5) include 18 different investor sectors, with security specific interactions with dummies of investor sectors Monetary Financial Institutions (MFIs) or 'bank', Other Financials 'omfi', insurance corporations 'insr', pension funds 'pfnd' and households 'hhld', see also Table 4 and section D.1 on grouping investor sectors.*

Table A3 – Cross-sectional regression with ownership shares

	(1)	(2)	(3)			(4)			
	EA	EA_j	EA_sj	main	bank	omfi	insr	pfnd	hhld
EUR	0.35***	0.07***	0.02***	0.01	0.05***	-0.01	0.02	-0.01	-0.01
	[0.002]	[0.001]	[0.005]	[0.008]	[0.014]	[0.009]	[0.016]	[0.008]	[0.008]
Constant	0.41***	0.00	0.03***			-0.00			
	[0.015]	[0.004]	[0.012]			[0.015]			
Obs.	272,616	306,092	635,453			635,453			
Adj. R^2	0.609	0.157	0.0825			0.0901			
HC FE	N/A	Yes	N/A			Yes			
HS FE	N/A	N/A	Yes			Yes			
HC*HS FE	N/A	N/A	N/A			Yes			
IC FE	Yes	Yes	Yes			Yes			

*Notes: Dependent variable Share defined as the holdings of an investor over the total size of the bond, indicating the ownership share in a bond. Reference period 2024-Q1 held by 20 holder countries across 8 investor sectors. Robust standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. Fixed effects (FE) are based on a combination of HolderCounty (HC), HolderSector (HS) and IssuerCountry (IC), where N/A means not applicable given the level of aggregation of the respective column. Columns (2), (4) and (5) include 20 holder countries and Columns (3) to (5) include eight different investor sectors, with security specific interactions with dummies of investor sectors Monetary Financial Institutions (MFIs) or 'bank', Other Financials 'omfi', insurance corporations 'insr', pension funds 'pfnd' and households 'hhld', see also Table 4 and section D.1 on grouping investor sectors.*

DeNederlandscheBank

EUROSYSTEEM

De Nederlandsche Bank N.V.
Postbus 98, 1000 AB Amsterdam
020 524 91 11
dnb.nl