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 authors and do not necessarily reflect official positions of De Nederlandsche Bank.

Working Paper No. 757

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

December 2022

A Literature Review of Securities Holdings Statistics Research and a Practitioner's Guide

Martijn A. Boermans

De Nederlandsche Bank

DNB Working Paper No. 757

December 2022 [1]

Abstract: Securities Holdings Statistics (SHS), compiled by the European System of Central Banks (ESCB) have spurred research over the past decade. SHS provide high-quality security-by-security data on portfolios. SHS benefit from very high coverage across euro area investors, relying on harmonized reporting and data preparation by the ECB since 2013-Q4. This paper provides a literature review of SHS research by surveying all published journal articles and working papers using granular SHS data. We demonstrate a rising popularity of SHS with 69 studies so far, advancing most prominently three research fields: (i) the banking and finance literature, mostly on interconnectedness and contagion, (ii) the international investment literature, and, (iii) monetary policy research on quantitative easing. Still, this review argues that SHS research is in its infancy, yet quickly growing. We highlight a surge of new studies, notably in sustainable finance. We provide a practitioner's guide with code, cleaning procedures and common specifications illustrated with home currency bias regressions. Finally, this review discusses avenues of future research.

Key words: Securities holdings statistics, portfolio investment, literature review, Eurosystem data, home currency bias.

JEL-classifications: E52, E58, F14, F3, G11, G2, G51, Q56.

---

[1] First version circulated on 14 November 2022. Please contact the author if there are any omission in this literature review. The document will be updated in the near future, detailed documentation can be found on the DNB website (link). I thank Maurice Bun, Iman van Lelyveld, Robert Vermeulen and participants of the DNB SHS workshop and DNB research seminar 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. E-mail: m.a.boermans[@]dnb.nl

A literature review of Securities Holdings Statistics research and a practitioner's guide

Research highlights:

- Ten years of Securities Holdings Statistics (SHS) have spurred academic research in a wide range of fields, advancing the literature from a macroeconomic view to granular analyses, most prominently on banking and finance, international investment and monetary policy.
- SHS provide high-quality security-by-security data on portfolios with very high coverage for euro area investors across different countries and sectors, relying on harmonized reporting and data preparation by the ECB with almost 1,000 million observations, covering more than 40 trillion euros of euro area holdings by 2022-Q2.
- We provide a literature review of all 69 journal articles and working papers so far, showing that more than 180 researchers published with granular portfolio holdings SHS data.
- We suggest that SHS research is still in its infancy but strongly growing, highlighted by the recent rapid expansion of SHS research with 30 working papers between 2021 and 2022. This surge is especially notable in sustainable finance research.
- We offer a practitioner's guide using portfolio holdings data, illustrated by home currency bias regressions.

1. Introduction

The financial crisis in 2008-2009 highlighted the urgency for granular portfolio holdings. During the fall of Lehman Brothers in September 2008 contagion channels were largely unclear for euro area financial institutions. Aggregate information could not adequately be applied to calibrate direct and indirect exposures and wider systemic risks in the financial system. To fill this data gap, based on the principles of the international G-20 Data Gaps Initiative under the Financial Stability Board (FSB) and IMF, in October 2012 the European Central Bank (ECB) published a new legal framework to advance the harmonized collection of securities holdings statistics (SHS) at the security level in the euro area. Before SHS, policy makers and researchers mostly relied on macroeconomic country level data such as the IMF Coordinated Portfolio Investment Surveys (CPIS).

SHS comes in two modules: SHS-*Sectoral* (SHS-S) and SHS-*Group*. SHS-S comprises security-by-security holdings and transactions aggregated at the level of investor sector for each investor country, comprising a full economy view of financial and non-financial investors. SHS-G is now available for all +120 large banking groups that fall under supervision of the European Single Supervisory Mechanism (SSM), and recorded at either individual entity or group level of the bank. Securities data comprise money market paper, bonds, listed stocks and investment funds. Data is available as of reference period 2013-Q4 and updated regularly on a quarterly basis with about only a lag of two months' time, covering almost 1,000 million observations and 40 trillion euros in assets from euro area investors by 2022-Q2. The main benefit of *granular* SHS data is that it allows supervisors, policy makers and researchers to combine portfolio flows with other sources, allowing users great flexibility to look into a "kaleidoscope" in different topics related to portfolio investments. This has become a gamechanger compared to earlier country level data.

After two years of data collection, in 2015 researchers within the Eurosystem gained access to the SHS database. Within less than a decade, SHS data collection and research has led to new insights on portfolio flows, mostly on interconnectedness, international investment and unconventional monetary policy. The ECB has merged the SHS data with the ECB Centralised Securities Database, a reference database on securities and issuers of securities. In this way researchers have detailed security data for about 100 variables, including price data and yields, asset size indicators like amount outstanding and market capitalization of stocks and funds, issuer country and sector information.

This literature review provides an overview of this academic journey so far using SHS. Our aim is to review all journal publications and working papers, organizing by different research themes (Section 4). Taken together these papers provide us with insights on best practices, allowing us to outline a practitioner guide (Section 5) and to distill a roadmap for future research using granular SHS data (Section 6).[2] We argue that SHS research is still in it's infancy, albeit growing rapidly. Most of the work is based internally by central banks for policy purposes and most visible in external policy work such as annual reports of central banks, financial stability review and macroeconomic outlooks that increasingly rely on SHS data (see e.g. ECB, 2015, 2022). In this review however we focus on academic output.

Our analysis yields 69 papers using SHS, contributing to chiefly to three research agendas. Of these 69 papers, 20 are published in journals. Of these papers, first and foremost, the data gaps related to interconnectedness and contagion channels that provoked the SHS data collection have spurred most research in banking and finance (see Squartini et al. 2017; Anand et al. 2018; Boermans & Van Wijnbergen, 2018; Hüser et al. 2018; Hüser & Kok, 2020; Martino, 2021; Aldasoro et al. 2022 and Jourde, 2022). Second, the international investment literature has been revisited with granular portfolio insights (Boermans & Vermeulen, 2016, 2020; Boermans & Burger, 2021, *forthcoming*; De Haan & Vermeulen, 2021 and Carvalho, 2022). Third, as SHS data became accessible for research in 2015, the ECB started quantitative easing with the asset purchase program. Benefitting from this timing, an important stream of research has focused on monetary policy effectiveness and transmission channels using SHS data (see Arrata et al. 2020; Bubeck et al. 2020; Albertazzi et al 2021; Koijen et al. 2021; Kabaca et al. 2022, *forthcoming*).

Most research using SHS data is work in progress but surging with an increasing number of recent papers. Our literature review suggests that there are currently 49 working papers, of which 30 released between 2021 and 2022. Hence, research using SHS is rapidly growing as more working papers have appeared between 2021-2022 than the combined number of journal articles (20) over the past decade. While the research on banking and finance, international investments and monetary policy is still rapidly growing, the number of articles on sustainable finance stands out with nine papers between 2020-2022. Also work on financial markets is quickly expanding, with eight working papers.

[2] There are other granular portfolio holdings dataset used in research, see Appendix 1 for highlights. Some of these datasets have also benefitted from literature reviews, see e.g. Dick-Nielson (2014) on TRACE, Hu et al. (2018) on Abel Noser data. See also ECB (2015) on SHS.

Merging portfolio holdings with other datasets has initiated a kaleidoscopic view in different research agendas. For financial stability the data of individual banking groups (SHS-G) has been merged with supervisory data such as COREP, FINREP and IBSI, but also commercial sources on bank's balance sheets such as BankScope and Orbis. Bank portfolios are also linked with granular loan information for AnaCredit. Also, the Secured Financing Transaction Regulation (SFTR) has provided daily data on repos, securities lending activities and margin lending which can be linked with counter-party information. Studies have also used EMIR data on derivatives, mostly to understand banking exposures on interest rate risks and currency swaps.[3] In the international investment literature relying on SHS-S data on the issuers have been merged including credit ratings, CDS spreads, industry classifications and distinctions by issuer sectors including sovereigns but also for security characteristics including currency of denomination, residual maturity, collateral eligibility under ECB operations, yields and stock performance. For monetary policy, difference-in-difference methods have thrived to allow for proper construction of a bond control group of ineligible assets under the various ECB programs including the PSPP and CSPP. Ongoing work in sustainable finance tracks energy-intensive firms by industry and CO_2 emission data.

After our literature review we provide a practitioner's guide with best practices for future research. We focus on SHS-Sectoral holdings, which was used in 75% of published articles and 83% of working papers. We suggests ways to exclude and regroup certain data based on common practices in the literature. Our practitioner's guide is informed by previous studies and aims to further harmonize the cleaning steps and facilitate future research, of which we discuss several new avalanches.

This literature study is structured as follows. Section 2 presents background information on SHS data. Section 3 explains how articles were selected for this literature review, leading to a sample of 62 papers so far. Section 4 provides an overview of this academic journey with SHS data by reviewing all papers grouped by research theme. Section 5 presents a practitioner guide with a case study with over 10 million observations on home currency preferences and Section 6 distills a roadmap for future research. Section 7 concludes.

[3] Ullersma and Van Lelyveld (2022) explain how granular data in general and portfolio holdings in particular can easily be linked with reference data, i.e. using the Legal Entity Identifier (LEI) of issuers and holders and the International Security Identification Number (ISIN) of individual securities, with supervisory data and other granular sources on loans and derivatives, especially for stress-testing purposes. The ECB has widely shared coding linking SHS other granular data sources, e.g. AnaCredit, SFTR and EMIR using LEI and ISIN.

2. Background information on SHS

The securities holdings statistics come in two main formats, sectoral level and individual banking group. First, a Sectoral module provides security-by-security flows and positions at the level of investor sector and investor country. By 2022-Q2 total holdings by euro area investors were about 40 trillion euro in the SHS-S data, covering almost 1,000 million observations over the period 2013Q4-2022Q2.[4] For research purposes so far this database has been extensively used because it allows a general European investor perspective due to its broad coverage of different sectors of the economy. Second, an individual banking Group module of SHS features bank-level holdings, currently of 123 large banking groups, covering a total of 4.3 trillion euro of investment positions by 2022-Q2. Studies using the bank-level have mostly focused on financial intermediation and risk, linking SHS to bank characteristics.

In the SHS, portfolio positions and flows are provided at a quarterly frequency and cover holdings of money market paper, bonds, listed stocks and mutual funds. The vast majority of securities are collected with an International Securities Identification Number (ISIN), which facilitates comparability across reporting agents and linking with other reference data (see ECB, 2015). Data consists of two levels: the security-by-security reference data, including issuer level information and the investor country and investor sector details, which are aggregated over the individual institutions in the SHS-S module.

The SHS-Sectoral data is most extensively 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. 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.

[4] Each instrument class has a single raw data file ("ECB clean copy") which for all securities includes 30.2 million observations for a single quarter: 12.5 million observations for bonds, plus 6.9 million for stocks, 9.5 million for investment funds and 1.3 million for money market paper. Currently there are 35 quarters of data, which yields about 1 billion observations.

Access to the granular data is currently restricted to the European System of Central Banks, however, detailed series are published at the ECB website (see also Radke et al. 2021). For an overview of the history of the SHS database, see work by the Irving Fisher Committee of (i) Sola and Strobbe (2010) who discuss the data gaps initiative, (ii) Lavrador et al. (2012) who discuss the data gathering of the SHS in the context of national accounts and Fache Rousová and Rodríguez Caloca (2014) who provide applications to financial integration data series with informative graphs comparing SHS data with other sources and highlighting the breakdowns available only with SHS data.[5]

Data on SHS are merged with about 80 key items from the Centralised Securities Database (CSDB) facilitating researches to work with a common dataset. This merging with reference data on securities has greatly facilitated research as the CSDB benefits from multiple data providers as well as input directly from national central banks based on reporting from domestic institutions on security reference data. Clearly, use of yields, country of issuer classifications and instrument types etc have all been provided for in the SHS dataset so as to further harmonize common practices within the ESCB.

3. Data on SHS research for this review

3.1 Sample selection

We obtain our sample of journal publications and working papers through online search. First, we use Google Scholar with broad search key words "SHS securities holdings statistics" and select articles among the first 250 searches that use the Eurosystem SHS data. Next, we do the same for the first 100 searches using the Science Direct database from Elsevier. Only studies benefiting from using granular SHS data are included. Only articles in academic journals as well as working papers are included.[6] We belief this methodology should suffices as external publications of any analysis using SHS data has to be cleared for non-disclosure of confidential data, whereby the source has to explicitly be acknowledged.

[5] The Eurosystem collects SHS data under Regulation ECB/2012/24. The ECB website provides further details on the SHS database: Securities (europa.eu).

[6] Various articles rely on national SHS data, which are not included in this overview. The searches were completed on October 1st 2022 and will be updated. Note that the 250 and 100 items were chosen because the majority of output at that stage did not include any useful articles.

3.2 Findings of SHS research overview

Our analysis yields 20 published journal articles for the survey, but highlights strong growth of SHS research with another 49 recent working papers. Table 1 presents an overview of these published articles by year. Figure 1 shows the rapid growth in the number of studies in the past few years. In 2016 the first two published journal articles came out. Since then there is a significant growth in the number of studies, both in terms of published articles and working papers. More than half (65%) of the journal papers were published in the last three years. The publication cycle in economics is relatively long, therefore it is not surprising that there are still many working papers from a few years ago that are not published. The rapid rise is therefore mostly concentrated in the number of working papers since 2019. In the first nine months of 2022 a record of 20 working papers were released, along with 5 journal articles. These data suggest a surge of SHS research recently.

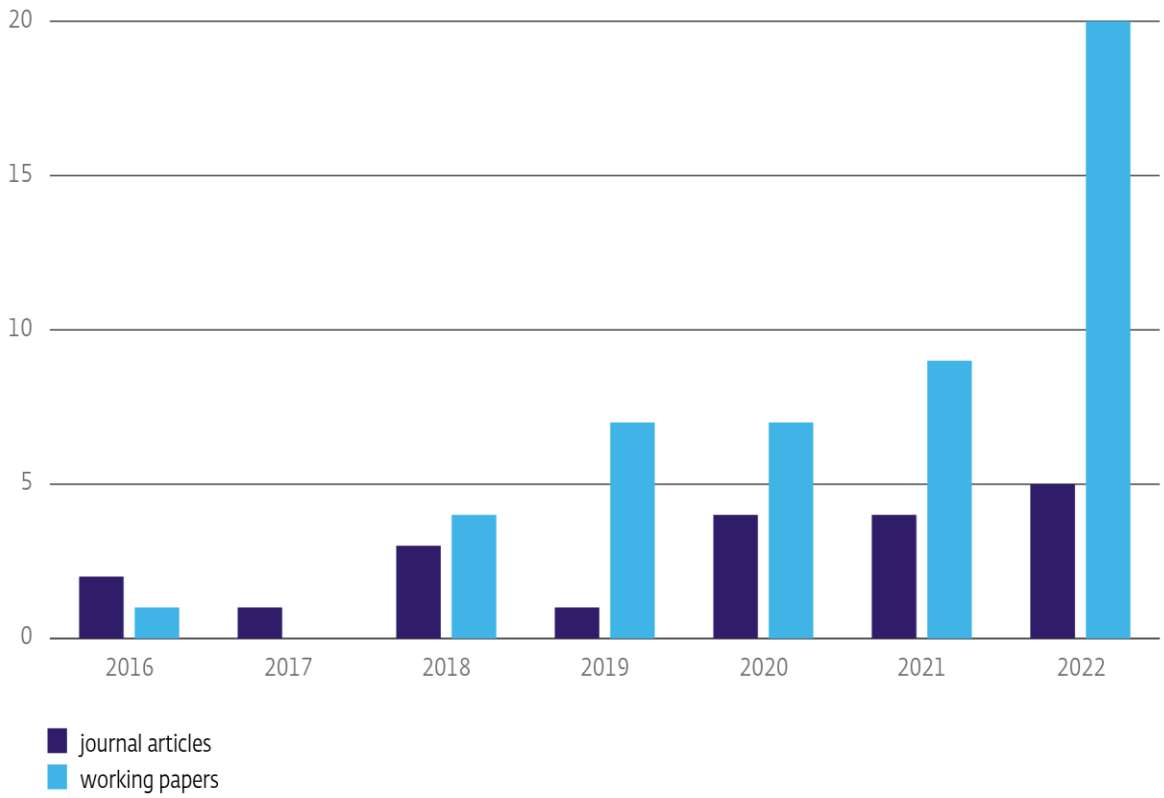
Table 1: Main findings on publications (journal articles only), by theme and year

Theme	No.	Authors	Year	Journal
Banking and finance	1	Squartini, Almog, Caldarelli, van Lelyveld, Garlaschelli & Cimini	2017	PhysR E
	2	Anand, van Lelyveld, Banai, Friedrich (…) & de Souza	2018	JFS
	3	Boermans & van Wijnbergen	2018	AEL
	4	Hüser, Hałaj, Kok, Perales & van der Kraaij	2018	JFS
	5	Bubeck, Maddaloni & Peydro	2020	JMCB
	6	Hüser & Kok	2020	JNTF
	7	Martino	2021	JCLS
	8	Aldasoro, Hüser & Kok	2022	JEDC
	9	Jourde	2022	JRI
International investment	10	Boermans & Vermeulen	2016	FRL
	11	Boermans & Vermeulen	2020	RoIE
	12	de Haan & Vermeulen	2021	JIMF
	13	Carvalho	2022	JIMF
	14	Boermans & Burger	FC	JIE
Monetary policy	15	Arrata, Nguyen, Rahmouni-Rousseau & Vari	2020	JFE
	16	Albertazzi, Becker & Boucinha	2021	JFI
	17	Koijen, Koulischer, Nguyen & Yogo	2021	JFE
	18	Kabaca, Maas, Mavromatis & Priftis	FC	EER
Financial markets	19	Boermans, Frost & Steins Bisschop	2016	EL
Sustainable finance	20	Alessi & Battiston	2022	IRFA

Notes: All published journal articles using granular SHS data up to 1 October 2022, organized by research theme and year. The year FC stands for forthcoming and available online. Journal abbreviations are implied in the reference list. Updates will be provided on the DNB website (link).

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. Our results show that 75% uses SHS-S data and 30% of the papers uses SHS-G data (as two papers employ both). 65% uses time-series data while 35% only analysis a cross-section. Bonds and stocks are the most widely used instrument classes, where only 12% uses all four instruments. Almost all data rely on the portfolio holdings and only a few papers utilize the transactions data.

Figure 1: The growth of SHS research over time



Notes: Number of journal articles & working paper using granular SHS data, by year.

As the number of articles is rapidly expanding, so is the SHS community in numbers of researchers. Currently, the number of different researchers working with SHS stands at 180. While a few years ago most researchers were affiliated with a central bank, more and more academics from universities are joining the list of authors. However, due to confidentiality restraints the number of central bankers in the top list of authors is still 100% (see Table 2).

Table 2: Overview of top-10 authors by number of papers, with affiliations

Author	No. of papers	Current affiliation
Boermans, M.A.	11	DNB
Lelyveld, I.	5	DNB
Vermeulen, R.J.G.	5	DNB
Breckenfelder, J.	4	ECB
Schmitz, M.	4	ECB
Carvalho, D	3	BdP
Fache Rousová	3	ECB
Hüser, A.	3	BoE
Kok, C.	3	ECB
Nguyen, B.	3	BdF
Sydow, M.	3	ECB

Notes: The number of papers include both journal articles and working papers combined. Abbreviations stand for De Nederlandsche Bank (DNB), European Central Bank (ECB), Banco du Portugal (BdP), Bank of England (BoE) and Banque du France (BdF).

4. Literature overview

Ten years of Securities Holdings Statistics (SHS) has mostly advanced the literature on banking and finance, international investment and monetary policy. In this section we discuss key insights of SHS research grouped by research theme, focusing on journal publications mostly in chronological order, specifically in Subsection 4.1 on banking and finance, Subsection 4.2 on the international investments literature and Subsection 4.3 on monetary policy. We also discuss all work in progress based on the available working papers by theme. In Subsection 4.4 we discuss the rapid growth of papers in the sustainable finance literature as well as on financial markets in Subsection 4.5. Subsection 4.6 highlights other research themes including household decisions, government finances and supervision.

4.1 Banking and finance

Our review shows there are 23 studies in the banking and finance theme. These papers fall within three clusters: (i) CoCos, (ii) risk taking and (iii) interconnectedness. We discuss the journal articles and then place the working papers within these research clusters under the item (iv) work in progress. Table 3 gives an overview, which we discuss in this section.

Table 3: Overview of SHS research on banking & finance, by year

Year	No.	First author**	Journal	SHS type	Security type	Amount type	Sample period
2017	1	Squartini	PhysRE	SHS-S	stocks	positions	09Q4-15Q1
2018	2	Anand	JFS	SHS-G	bonds & stocks	positions	13Q4-15Q4
	3	Boermans	AEL	SHS-S	bonds	positions	15Q4
	4	Hüser	JFS	SHS-G	bonds & stocks	positions	15Q1
	5	Wang	WP*	SHS-G	all		13Q4-16Q2
2019	6	Bekaert	WP*	SHS-S	all		13Q4-18Q2
	7	Fache Rousová	WP*	SHS-S	bonds		13Q4-16Q4
	8	Hüser	JNTF	SHS-S & SHS-G	bonds & stocks	positions	15Q4
	9	Ringe	WP*	SHS-S	bonds		13Q4-18Q3
2020	10	Boermans	WP*	SHS-S	bonds		13Q4-19Q4
	11	Bubeck	JMCB	SHS-G	bonds	positions	13Q4-14Q4
	12	Mink	WP*	SHS-G			
2021	13	Attina	WP*	SHS-S	bonds		13Q4-20Q2
	14	Martino	JCLS	SHS-S	bonds	positions	13Q4-17Q4
	15	Montagna	WP*	SHS-G	all	positions	13Q4-18Q4
	16	Sydow	WP*	SHS-G	all	positions	19Q4
2022	17	Aldasoro	JEDC	SHS-G	all	positions	16Q1
	18	Del Vecchio	WP*	SHS-G	all	positions	20Q1-20Q2
	19	di Iasio	WP*	SHS-S	bonds		22Q2
	20	Fukker	WP*	SHS-S	bonds & stocks		18Q1-20Q1
	21	Henricot	WP*	SHS-G	bonds	positions	16Q1-19Q4
	22	Jourde	JRI	SHS-S	stocks	positions	14Q4-20Q1
	23	Maddaloni	WP*	SHS-S			

Notes: all journal articles and working papers in the Banking & Finance research theme by year. Journal abbreviations are given in the reference list, where WP* stands for working paper. The author lists is the first author**, where in economics alphabetical order is often applied. Line markings are applied in the table rows by year to delineate the development over time.

4.1.1 Contingent convertible capital

We identify five papers on contingent convertible capital (CoCo), of which two published. 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. While literature on the issuance of these new financial instruments developed, little was known about the investors in these relatively risky assets. Using the SHS-S data Boermans and Van Wijnbergen (2018) show that cross-holdings of these CoCo-instruments within the euro area banking sector are virtually non-existent. Foreign investors outside the euro area and investment funds located in Ireland and Luxembourg hold the large majority of contingent convertibles. Building on this, Martino (2021) shows how certain investors are better suited to hold CoCos from a risk perspective. He highlights a trade-offs between corporate governance and the threat posed by different investors.

4.1.2 Risk taking

Our analysis yields nine studies on risk taking, of which only one is published. Bubeck et al. (2020) use a difference-in-differences setup to study how the holdings by 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.

4.1.3 Interconnectedness

We categorize nine papers on banking and finance within the interconnectedness cluster, of which six published, mixing SHS-S and SHS-G data. Squartini et al. (2017) argue that reconstructing patterns of interconnections from partial information is one of the most important issues in the statistical physics of complex networks. In fact, the spreading and amplification of financial distress in capital markets are strongly affected by the interconnections among financial institutions. Using SHS-S data, Squartini et al. (2017) are able to better capture the topology of the networks using an enhanced CAPM framework to estimate the systemic risk due to fire-sales and spillovers. In similar spirit, Anand et al. (2018) highlight the importance of granular data to capture how network linkages drive contagion among individual banks. Using the SHS-G banking data, Anand et al. (2018) construct bilateral exposures and link these interbank portfolio networks to other markets at

the bank-level, including payments systems, overnight interbank loans, repo transactions and interbank deposits to better understand what type of spillover models fit best with the network structure. Similarly, Aldasoro et al. (2022) link the banking groups holdings to loans and derivates data to study interbank networks contagion channels. Their stress-test highlights a prominent role of overlapping portfolios and associated negative price externalities due to fire sales. Hüser et al. (2018) also use SHS-G banking group data to construct networks linking these to reporting on balance sheet data, own funds and capital requirements. They focus specifically on cross-holdings of CoCo-instruments because these instigate direct contagion when being bailed in. In their multi-layered networks analysis, Hüser et al. (2018) find that due to low levels of securities cross-holdings in the interbank network there is no direct contagion in terms of creditor banks failing as a result of another bank being bailed in in line general findings from Boermans and Van Wijnbergen (2018) based on SHS-S. Using a similar multi-layer model Hüser and Kok (2020) further show that global systemically important banks (G-SIBs) are most interconnected in general, however, several non-G-SIBs also display similar levels of interconnected. Finally, for insurance corporations Jourde (2022) suggests that insurance corporations have become more interconnected. While insurers are more vulnerable to shocks stemming from non-financial sectors, banks have stronger links with the rest of the financial industry.

Work in progress: Apart from the eight published articles, we collected 13 working papers related to the field of banking and finance, including work within the themes on CoCos and bail-in (Ringe & Patel, 2019; Attinà & Bologna, 2021; Maddaloni & Scardozzi, 2022), risk taking (Bekaert & Breckenfelder, 2019; Fache Rousová & Giuzio, 2019; Boermans & van der Kroft, 2020; Mink et al. 2020; Montagna et al., 2021; Fukker et al., 2022; Del Vecchio et al., 2022; Henricot & Piquard, 2022) and interconnectedness (Wang et al., 2018; Sydow et al. 2021; diIasio et al. 2022).

4.2 **International investment**

Our survey shows there are seven paper related to the research theme on international investment, of which five published. Table 4 provides an overview of the journal articles and working papers which we discuss briefly in this subsection under work in progress. Detailed information on international investment patterns have revisited the literature. Boermans and Vermeulen (2020) compare the drivers of portfolio choice established in the macroeconomic international finance literature and test if those factors are still relevant when allowing for investor heterogeneity and security characteristics. Boermans and Vermeulen (2016) show that the preference to invest within the euro area is mostly explained by the tendency to hold assets denominated in the home currency. De Haan and Vermeulen (2021) analyze cross-border holdings responses to changes in sovereign debt ratings. They find investors from core euro area countries respond stronger to credit ratings than investors from the euro area periphery. Carvalho (2022) uses a look-through approach for mutual fund holdings. The results show that exposures to foreign bonds are much larger than aggregated data sources suggest because these positions are taken via investment funds. Finally, Boermans and Burger (2021, *forthcoming*) focus on the currency dimension of fickle flows to emerging markets, highlighting amplification by the broad US dollar exchange rates yet stabilization by the Euro-denominated debt.

Table 4: Overview of SHS research on international investment, by year

No.	Year	First author**	Journal	SHS type	Security type	Amount type	Sample period
1	2016	Boermans	FRL	SHS-S	bonds	positions	14Q4
2	2020	Boermans	RoIE	SHS-S	all	positions	13Q4–18Q4
3	2021	Boermans	JIE	SHS-S	bonds	transactions	13Q4-22Q2
4	2021	Carvalho	WP*	SHS-S	all	positions	20Q1-20Q2
5	2021	de Haan	JIMF	SHS-S	bonds	positions	09Q4-16Q1
6	2022	Carvalho	JIMF	SHS-S	all	positions	14Q1-29Q3
7	2022	Carvalho	WP*	SHS-S	all	positions	13Q4-20Q4

Notes: all journal articles and working papers in the International Investment literature by year. Journal abbreviations are given in the reference list, where WP* stands for working paper. The author lists is the first author**, where in economics alphabetical order is often applied.

Work in progress*:* with four published articles there are also three working papers in this field marking significant progress. First, Carvalho and Schmitz (2021) apply a look-through approach to study portfolio reallocations of euro area investors during the Covid crisis.

Second, Carvalho and Schmitz (2022) analyze the impact of the Brexit on euro area investment patterns in the United Kingdom.

4.3 Monetary policy

We document that SHS-S has been extensively used for monetary policy research with 12 papers analyzing the various unconventional policy measures, mostly using SHS-S bond holdings. Our survey identifies five journal articles. When the SHS data became available to researchers, the Eurosystem in early 2015 initiated quantitative easing (QE) to lower interest rates by purchases long-term government bonds. Given the coincidental timing, several important studies on unconventional monetary policy effects have been published but with many ongoing as working papers (see Table 5).

Table 5: Overview of SHS research on monetary policy, by year

No.	Year	First author**	Journal	SHS type	Security type	Amount type	Sample period
1	2018	Boermans	WP*	SHS-S	bonds		13Q4-16Q3
2	2018	Boermans	WP*	SHS-S	bonds		13Q4-16Q4
3	2019	Eser	WP*	SHS-S	bonds		13Q4-18Q2
4	2020	Arrata	JFE	SHS-S	bonds	transactions	15Q1-17Q2
5	2020	Bergant	WP*	SHS-S	bonds		
6	2020	Ferdinandusse	WP*	SHS-S	bonds		13Q4-18Q4
7	2021	Albertazzi	JFI	SHS-S & SHS-G	bonds & stocks	positions	14Q1-15Q2
8	2021	Jasova	WP*	SHS-S	bonds	positions	
9	2021	Koijen	JFE	SHS-S	bonds & stocks	positions	13Q4-17Q4
10	2022	Greppmair	WP*	SHS-S	bonds		19Q4-21Q4
11	2022	Hudepohl	WP*	SHS-S	Bonds	transactions	13Q4-19Q4
12	2022	Kabaca	EER	SHS-S	bonds	positions	13Q4-18Q4

Notes: all journal articles and working papers on monetary policy by year. Journal abbreviations are given in the reference list, where WP* stands for working paper. The author lists is the first author**, where in economics alphabetical order is often applied.

The majority of papers study 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 effect portfolio

rebalancing if 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. (2022, *forthcoming*) explore optional allocations of sovereign bond purchases, they show that given frictions, relatively more purchases in the periphery instead of the core are optimal to induce greater yield impact. 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: Boermans and Vermeulen (2018) explain that QE in the euro area depends on preferred habitat investors who are relatively unwilling to sell assets to the ECB. Boermans and Keshkov (2018) study how QE drives asset scarcity, leading to greater portfolio concentration of sovereign debt in the euro area. Relatedly, Ferdinandusse et al. (2020) highlight how QE affects market liquidity. Eser et al. (2019) 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. Jasova et al. (2021) study the central bank role of lender of last resort from a haircut gap channel in relationship to bank interconnectedness and systemic risk. They show these monetary operations increase bank bond holdings. Greppmair and Jank (2022) also look at the central bank lending facilities and focus on the regulatory chock in November 2020 when the Eurosystem changed pricing conditions and fees. They show that these policy shifts increased the use of securities lending facilities surged, in particular for bonds with otherwise inelastic supply to the repo market, alleviating scarcity in the repo market and enhancing liquidity. Two papers use the transactions data to analyze portfolio rebalancing effects related to the ECB asset purchase program. First, Bergant et al. (2020) highlight that euro area investors actively rebalanced their bond portfolios away from euro area sovereign bonds targeted by the ECB quantitative easing program. They show that euro area investors bought more foreign bonds in response to QE. Second, in a related empirical specification to Bergant et al. (2020), Hudepohl (2022) shows that euro area investors bought more bonds Euro-denominated debt from vulnerable euro area countries and emerging market local currency bonds, suggesting greater risk-taking in the bond market after quantitative easing.

4.4 Sustainable finance

Our survey highlights a recent growth on SHS research in sustainable finance with only one published article, versus another nine working papers. Table 6 tabulates these studies related to carbon footprints, energy intensity and climate stress testing. Specifically, the published work by Alessi and Battiston (2022) categorize European portfolio holdings of non-financial firms' bonds and equity based on NACE classifications to estimate the 'greenness' and climate transition risk. The granularity of the data allows them to calibrate alignment with the EU Taxonomy for sustainable activities. Alessi and Battiston (2022) highlight that the greenness of investors' portfolios is lower than their exposure to transition risk and that these are especially relevant for institutional investors like investment funds and insurance corporations.

Table 6: Overview of SHS research on sustainable finance, by year

No.	Year	First author**	Journal	SHS type	Security type	Amount type	Sample period
1	2020	Boermans	WP*	SHS-S	stocks	positions	13Q4-18Q4
2	2020	Mésonnier	WP*	SHS-S	bonds & stocks	positions	13Q4-19Q3
3	2021	Alogoskoufis	WP*	SHS-G	bonds & stocks	positions	18Q4
4	2022	Aghion	WP*	SHS-S	bonds & stocks		13Q4-20Q3
5	2022	Alessi	IRFA	SHS-S	bonds & stocks	positions	22Q1
6	2022	Belloni	WP*	SHS-G	all	positions	13Q4-17Q4
7	2022	Dubiel-Teleszynski	WP*	SHS-S & SHS-G	all		20Q4
8	2022	Lucia	WP*	SHS-S	stocks	positions	15Q1-20Q3
9	2022	Papoutsi	WP*	SHS-S	bonds	positions	17Q4
10	2022	Pietsch	WP*	SHS-S	bonds	positions	16Q1-21Q3

Notes: all journal articles and working papers in the sustainable finance literature by year. Journal abbreviations are given in the reference list, where WP* stands for working paper. The author lists is the first author**, where in economics alphabetical order is often applied.

Work in progress: Boermans and Galema (2020) introduce the concept of carbon home bias, showing that investors not only overweight domestic stocks, but especially those with greater carbon intensity. They find that carbon home bias is associated with a carbon premium. Also using firm carbon emission data, Lucia et al. (2022) focus on holdings of EU stocks and provide a simple difference-in-difference model to study how portfolios adjusted after the Paris Agreement end-2015. They find minor reductions in the exposures to carbon-intensive stocks attributed in response to the Paris Agreement. Mésonnier and Nguyen (2020) also test the impact of a regulatory shock, specifically a new French institutional investor regulation,

on divestment by focusing on energy intensive stocks grouped by industry data. They also find that the regulatory shock decreased positions in energy intensive industries. In a related approach, Aghion et al. (2022) also rely on industry-classifications to distinguish high and low carbon sectors. They illustrate that banks significantly reduced their share of non-financial securities in high carbon industries, mostly in equities. Papoutsi et al. (2022) bridged the sustainable finance literature with unconventional monetary policy. They show that the ECB asset purchase program for corporate bonds overweights energy-intensive industries relative to market portfolio along the carbon intensity dimension. Alogoskoufis et al. (2021) provide a methodology to conduct climate stress tests. Dubiel-Teleszynski et al. (2022) analyze the impact of a climate change stress scenario (disorderly transition) on banks, funds, and insurance companies simultaneously with multiple interconnections between the sectors. Their stress test highlights contagion through second-round effects and the largest losses in the non-bank financial sector. In related work, Belloni, Kuik and Mingarelli (2022) analyze how changes in carbon price affect European banks. They find that early and gradual changes in carbon prices have no negative effect on the banking sector, especially if firms reduce emissions efficiently. Pietsch and Salakhova (2022) analyze how shifts in investor demand for green bonds affect the greenium using retail share and euro area share of a bond. The findings suggest that more puchases by households increase the green bond greenium.

4.5 Financial markets

We find one published journal article on financial markets so far, plus 13 working papers. Table 7 tabulates these studies. The main focus is on crisis periods like the Taper Tantrum, Bund Tantrum and Covid shock early 2020. Much work is complementary to the banking and finance literature on how banks affect financial markets (see Section 4.1). In addition, work in this research cluster provides a bridge between the monetary policy literature (see Section 4.3) and its impact on asset scarcity and liquidity. Boermans, Frost & Steins Bisschop (2016) analyze the effects of market liquidity and ownership structure on European bond price volatility. They show that during the 2013 Taper Tantrum and 2015 Bund Tantrum market illiquidity aggravates price shock of sovereign bonds, however, concentrated bond holdings only affected volatility during the Bund Tantrum.

Work in progress: Della Corte and Federico (2016) analyze the holders of Italian government bonds, showing the importance of euro area investment funds as investor sector. Kliatskova and Savatier (2020) find that better insolvency regimes encourage portfolio

investments. Fricke et al. (2022) study the effects of a US money market fund reform on money market funds in the euro area. As US-based prime MMFs became less money-like due to the reform, euro area-based prime MMFs received large inflows from foreign investors, reducing euro area fund risk taking, yet making the funds more vulnerable to foreign investor runs. Breckenfelder and Ivashina (2021) work on bank balance sheet constraints and bond liquidity falls in between the banking and finance literature and financial markets theme. They show that the introduction of a leverage ratio for European banks had large effect on bond liquidity. During the Covid shock early 2020, mutual funds with the larger exposures to dealer banks' balance sheet constraints where affected more strongly. Fache Rousová et al. (2022) combine SHS data with daily data from EMIR on derivatives to study margin calls for money market funds. They show that the variation margin payments faced by some investors holding these funds drive flows of funds in the euro area during the Covid-19 outbreak.

Table 7: Overview of SHS research on financial markets, by year

No.	Year	First author**	Journal	SHS type	Security type	Amount type	Sample period
1	2016	Boermans	EL	SHS-S	bonds	positions	13Q4-15Q1
2	2016	Boermans	EL	SHS-S	bonds	positions	13Q4-15Q1
3	2018	Accornero	WP*	SHS-S	bonds	positions	13Q4-17Q2
4	2019	Brand	WP*	SHS-S	bonds	positions	15Q1-17Q4
5	2019	Dötz	WP*	SHS-S	mutual funds	positions	13Q4-16Q2
	2019	Grandia	WP*	SHS-S	bonds	positions	17Q3
6	2020	Kliatskova	WP	SHS-S	bonds & stock	positions	13Q4-17Q4
7	2021	Breckenfelder	WP*	SHS-S	bonds	positions	13Q4-20Q2
8	2022	Ahmed	WP*	SHS-S	bonds	positions	13Q4-20Q4
9	2022	Breckenfelder	WP*	SHS-S	money market	positions	13Q1-20Q2
10	2022	Darmouni Fache	WP*	SHS-S	bonds	positions	14Q4-20Q2
11	2022	Rousová	WP*	SHS-S	bonds	positions	20Q1
12	2022	Faia	WP*	SHS-S	bonds	positions	13Q4-21Q1
13	2022	Fricke	WP*	SHS-S	money market	positions	13Q4-17Q1

Notes: all journal articles and working papers on financial markets by year. Journal abbreviations are given in the reference list, where WP* stands for working paper. The author lists is the first author**, where in economics alphabetical order is often applied.

Corporate bonds markets are an important theme. Dötz and Weth (2019) study German mutual fund holdings of corporate bonds in the context of liquidity dry-ups and bond redemptions. They use SHS data to obtain a variable that proxies the share of euro area institutional investor holdings and retail investor holdings. Dötz and Weth (2019) show that illiquid funds dominated by retail investors are more exposed to outflows than illiquid funds

primarily owned by institutional investors, because institutional investors internalize the fire-sale-driven loss that a withdrawal inflicts. Accornero et al. (2018) study the corporate bond holdings of Italian issuers. They show that larger firms have a more international investor clientele, whereas smaller firms are held mostly by Italian investors, where domestic households are important financers. Darmouni and Papoutsi (2022) study the rise of bond financing in Europe. They show that mostly large firms issued bonds and during the Covid-19 crisis the bond market flight by investors affected primarily the largest, rated issuers.

Several studies relate monetary policy to financial market outcomes. Bridging literature on financial markets and monetary policy, Brand et al. (2019) analyze repo markets and utilize SHS data to derive a proxy of asset scarcity. They suggest that safe asset scarcity under quantitative easing appears to mostly play a role once holdings exceed about 10-15% of outstanding amounts of a country's bonds. Faia et al. (2022) explore corporate bond holdings in the euro to study currency pricing in the context of market segmentation and euro-dollar return differentials. They show that these differentials decline following the increase in ECB asset purchases of euro securities: the scarcity induced by the drain in supply reduced the required yields on euro securities, eroding dollar convenience yield. Grandia et al. (2019) analyze the availability of euro-denominated debt instruments that qualify as high-quality liquid assets (HQLA) in the euro area, mostly government bonds. They show that besides banks and the Eurosystem, non-euro area investors are the largest holders of euro-denominated HQLA. Breckenfelder and Schepens (2022) analyze the factors driving liquidity dry-ups in money markets during the early Covid shock in 2020 and central bank interventions. They show that the dry-up was driven by money market funds (MMFs) and affected corporate funding. Ahmed et al. (2022) analyze US corporate bonds and combine SHS-S data with eMAXX data on US investors and look-trhough funds to uncover reaching for yield tendencies.

4.6 Other fields

4.6.1 Household decisions

Research using SHS data on household investments is scarce. While the investment literature on household portfolio decision is consider a separate field in the finance literature, only two working papers focus specifically on the household portfolios. Lamas and Martinez-Miera (2021) household stock positions of banks and non-financial corporates and suggest that households are stock market liquidity providers. Boermans et al. (2022) study home bias and foreign bias in relationship to active portfolio management and excess returns, applying a very detailed look-through approach for positions held through mutual funds.

4.6.2 Government finance

Currently there are no papers focusing specifically on government finances. However, the ECB (2015) highlights how SHS provides new details on government bond holders and interest rate channels. Indirectly this field has benefits from work on quantitative easing in the euro area, which analyses the holdings of sovereign bonds (see Section 4.3).

4.6.3 Supervision

There are relatively few SHS papers focusing on supervision. While a range of journal publications using SHS in the banking and finance literature have provided tools for supervisors to analyze interconnectedness (see Section 4.1), appreciating possible new instruments for supervision and oversight could further benefit from ongoing and future SHS research on financial regulation and risk financial institutions face through portfolio holdings. One study of direct use to supervisors is from Abidi et al. (2021) how study the effect of (private) information disclosures by supervisors. In 2020 the ECB published for the first time bank-by-bank details on their Pillar 2 requirements, information on creditworthiness used by supervisors but not disclosed by financial institutions nor supervisors to the market. Abidi et al. (2021) show that bond prices and international bond portfolio positions are sensitive to new regulatory information as well as to rating gaps between the ECB and private credit rating agencies. Hence, they suggest that supervisors have "specific, distinctive, and valuable knowledge of the banks they supervise". Another area of particular interest for supervision and potential capital requirements comes from climate stress-testing as discussed in Section 4.4 on sustainable finance.

5 Practitioner guide for SHS researchers

In this section we provide guidance to facilitate use of SHS data for research using common practices. We highlight common practices for positions, starting with data cleaning up to reduced form regression models which aim to explain investment behavior. We focus on holdings instead of transactions because all journal articles except one rely on portfolio holdings data and among the working papers there are only three out of 43 benefiting from the transactions data. One serious challenge for several studies is potential simultaneity bias. For example, in a simple OLS regression where a researcher aims to explain certain bond positions while controlling for the bond yield, one ignores that holding or purchasing a bond can have an upward pricing effect, driving down bond yields. Similarly, once a researcher analyzes particular asset positions by currency, exchange rate effects might drive this relationship. Such endogeneity biases are typically resolved with instrumental variables (including GMM) estimation techniques, not further addressed in this practitioner's guide (see also Balazsi et al. 2018 on multidimensional panel data techniques).

Our practitioner's guide highlights choices to be made when preparing the data as well as cleaning methods. We explain the many choices researchers face and propose several cleaning procedures useful for most areas, with a step-by-step analysis provided. To illustrate the multitude of choices and these impact the results, we illustrate our setup with a study on currency bias of euro area investors in a simplified format.

5.1 Data cleaning procedures

Researcher working with SHS-S data often face similar decisions with regard to the cleaning of the data. This practical guide informs users on the most commonly taken steps and provide code in Stata which easily transfers to other software packages to ensure common practice in the literature using SHS-S data.[7] The example in terms of number of observations tracked is based on bond and stock data for the period 2022-Q2 but the procedure of cleaning extends to the whole time-series from 2013-Q4 onwards and can also be applied to money market paper and mutual fund data. Raw data files (clean copy) includes 12.5 million observations for bonds and 6.9 million observations for stocks.[8]

[7] Code in Stata is available on the DNB website (link).
[8] For investment funds this is 9.5 million observations and for money market paper 1.3 million observations.

Step 1: positions or transactions?

So far, all published articles have either used positions or transactions, while the SHS-S data also includes information on price and currency changes and other mutations. Most work relies on **positions** so you must select the variable "amount_type" with value "LE" while for transactions you take the value "T". This filter typically reduces the file size by half. There are also other " reconciliations" available, such as price effects (see Appendix 2).

Step 2: market values or nominal?

Depending on your topic you can take the positions at market value or nominal value. For stocks and investment fund data nominal values are not useful so we suggest to always take market values. In most circumstances for bonds market values are also of interest. This filter typically reduces the file size again by half.

Step 3: focus on euro area investors?

So far, most published articles have focused on euro area investors. The benefit of this approach is that the coverage is generally above 90 percent for each country so very representative. Non-euro area data is (very) incomplete, but through custodian data these may also include euro area positions, so-called third-party holdings, to be identified through the reporting country (*ref_area*). Because of potential double counting with these custodian data we suggest to only include these for households (S.14) and non-profit organizations (S.15) because these cannot suffer from double reporting. Therefore commonly 20 euro area investor countries are selected.[9] After this essential procedure the total holdings in 22-Q2 at market values by euro area investors are for bonds EUR 14.1 trillion, with 1.4 million observations and for stocks EUR 8.97 trillion, with 0.8 million observations.

Depending on the research question one better choose differently, however, key point then is to take one alternative route as a starting point to make processing feasible and separate data storage if you want to analyze both flows and positions (and merge them later if required, see e.g. Boermans & Burger, 2021, *forthcoming*, who take the gross flows divided by the previous quarter positions data to obtain a flow ratio dependent variable measure).

[9] On 1 January 2023 Croatia (country code "HR") joined the European Monetary Union.

Step 4: cleaning rules (part 1)

A best practice in the literature is to apply several cleaning rules. Focusing on euro area investments gives us a benchmark of EUR 14.1 trillion in bonds and EUR 8.97 trillion for stocks, against which we can compare our coverage after further data enhancements. We propose to adopt the following cleaning rules that are relevant to most research theme. First, focus on portfolio holdings. Second, exclude (aggregate) short positions. Third, exclude non-active securities (e.g. those in default status or already redeemed). Fourth, exclude investment in tax havens. The issue is that many of these small destination countries are not the final recipients of the capital flows, but only pass-through as issuer countries in SHS are based on residency principle and not group-based (see Lane & Milesi-Ferretti, 2008; Coppola et al., 2021).[10] For bonds the coverage is 97.8% and for stocks 94.4%. Depending on the instrument class and research topic additional cleaning rules are advised which are later discussed in Step 6.

One effective further cleaning rule is choosing carefully which issuers one required for the unit of analysis. Specifically, various studies focus only on government bonds, allowing one to select only issuer_sector "S13". Other variations for specific issuer sectors can be applied, e.g. only non-financials or just banks. On similar lines, several researchers focus only on euro area issuers, allowing one to drop all issuer countries from the rest of the world.

Step 5: aggregate relevant holder sectors or euro area investor?

There are 24 different holder sector instances 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. 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. In this way we define new investor sector definitions and reduce potential noise in the data (without affecting the total holdings).

We propose to focus on eight distinct investor sectors which have clear economic interpretation and sizeable holdings as a group.[11] The regrouping of sectoral data and

[10] For our definition of tax havens, see the provided code in Stata, available on the DNB website (link).
[11] The investor sector we define are (i) banks, (ii) insurance corporations, (iii) pension funds, (iv) mutual funds (including money market funds and other investment funds), (v) the household sector including non-profit institutions serving households, (vi) other financial institutions, including securitization vehicles, (vii) government and, (viii) non-financial corporations.

exclusion of undefined sectors is based on Boermans and Vermeulen (2020) who focus on investor sector heterogeneity. Other studies also rely of a similar investor-sector aggregate, but often go a few steps further to consolidate even further, e.g. labelling insurance corporations and pension funds as a single sector (e.g. Koijen et al. 2021) or putting all non-financial investor sectors in a single category.

For many papers the investor country dimension may not be of relevance and aggregation over the "holder_area" is advised to obtain an euro area investor perspective. This significantly reduced the size of the dataset by about half (without affecting the total holdings/ data coverage). Various researcher typically start with an euro area perspective analysis and later for robustness purposes explore the investor country level dimension to test for holder country specific effects that may drive the results and similarly provide a sensitivity analysis with the investor-sector dimension salient. Others even include the combination of investor country-sector dimensions for further inspections to make effective use of the granularity of the dataset while keeping large dataset restraints in mind (see e.g. Hudepohl, 2022).

Step 6a: cleaning rules (part 2)

After aggregation we can identify holding positions that are larger than the amount outstanding or market capitalization as a data quality management procedure (affecting generally less than one percent of observations). Next we can drop very small holding positions, where we suggest a threshold of EUR 10,000 as a minimum about held by euro area investors from a given investor sector in a single security at the relevant time period. For bonds it reduces the coverage from 97.8% to 96.5% while for stocks dropping small positions has no impact on the coverage, which stays at 94.4%. while still dropping 150,000 observations to 300,000 stock observations.

Step 6b: very specific cleaning rules for debt and shares

6b - Debt

For bonds we propose to include only bonds after primary issuance and not redeemed. Bonds without recent market price are dropped because these are often not securities. Similarly, some bonds have quoted market prices in numbers instead of percentage, which we also drop because of potential difficulties in determining market prices of the holdings. Certain debt types including in the database are more equity or derivate like in a hybrid sense, these are

also dropped from the analysis. More debatable is the inclusion of bonds with highly unlikely prices – unless that is a study focus – which we propose to drop. Because market values of bonds can be murky, we propose to winsorize all price information at one percent (two-sided, so effectively at 0.5% on each side after dropping extreme values fully from the sample). For bonds it reduces the coverage from 96.5% to 95.7% and the same time lowers the number of observations to 450,000 bonds at the level of euro area investor-sector.

6b - Shares

For stocks we only include stocks quoted in numbers (CCY). Next, only stocks with a quarterly price available are included, otherwise these are not really listed and traded stocks. Finally, we propose to drop ultra-penny stocks with a price below 0.05 euros always because of dubious trading and stocks with prices above 50,000 euros as these positions in a single stock are also not readily traded for portfolio purposes. With these quotation and pricing filters it reduces the coverage from 95.7% to 93.4% while decreasing the number of observations to 95,000 stock observations at euro area investor-sector stock-level.

5.2 Setting up a standard regression model with positions as regressor

Most articles focus on portfolio position data, specifically holdings at market value. A wide range of papers aims to explain the holdings as dependent variable in (panel) regressions. A typical empirical specification is the following:

(1a) $Hold_{i,s,j} = \alpha + \beta' X_{i,s,j} + \varepsilon_{i,s,j}$

(1b) $Hold_{i,s,j,t} = \alpha + \beta' X_{i,s,j,t} + \varepsilon_{i,s,j,t}$

where *Hold* represents the positions (at market value), preferably *in logs* to obtain a distribution closer to a normal distribution by smoothing outliers (e.g. 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, *j* the investor country (in principle all countries worldwide) and *t* the time period, quarterly, in Equation (1b) time-series variant (where Equation 1(a) is a cross-sectional analysis).

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.[12] When *X* are general macroeconomic variables that change on a quarterly basis, time fixed effects are not feasible die 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 analysis 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.

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:

$$(2)\ Hold_{i,t} = \alpha + \beta' X_{i,t} + \varepsilon_{i,t}$$

In this way one treats all euro area investors as a single homogenous 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. 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

[12] In addition, clustering standard errors at investor country-sector level is advised in this baseline setup.

another factor of about 10 for the investor sector s dimension, leaving about 6 million observations.

We propose researchers to consider investor heterogeneity based on the "aggregated" investor sector *s* information. Again, SHS-S data has no information on individual investors, while SHS-G data covers individual banking group data. With SHS-S, to account for true investor heterogeneity at the sectoral level, the coefficients of the explanatory variables for each investor sector *s* should be allowed to vary. The SHS data has 24 different, very detailed investor sectors based on ESA-2010 classifications, yet most important investor groups fall in certain ESA-2010 buckets that can be grouped together. We propose to focus on eight distinct investor sectors which have clear economic interpretation and sizeable holdings as a group.[13]

Researcher using security-by-security regressions carefully select proper fixed effects. When investment patterns towards a set of destination countries are analyzed without further modelling, 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. Included issuer sector fixed effects further capture unobserved issuer sector characteristics, which are often referred to as *multilateral resistance* terms in gravity models (see Anderson & van Wincoop, 2003) when combined with holder sector-country and issuer sector-country dummies (see Boermans & 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:[14]

$$(3)\ Hold_{i,s,j,t} = \beta_s \ X_{i,s,j,t} + FE_{s,j} + FE_{h,c} + \varepsilon_{s,j,t}$$

[13] The investor sector we define are (i) banks, (ii) insurance corporations, (iii) pension funds, (iv) mutual funds (including money market funds and other investment funds), (v) the household sector including non-profit institutions serving households, (vi) other financial institutions, including securitization vehicles, (vii) government and, (viii) non-financial corporations.

[14] In many settings a cross-sectional setup suffices and the stability of the estimated coefficients can be tested by choosing difference references periods from the SHS data.

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.[15] Instead of destination sector classifications one may use industry (NACE) classifications instead. We drop the constant alpha" 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 & Vermeulen, 2020). Here standard error 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. Appendix 3 outlines how *EA_Share* can be applied as alternative robustness test.

5.3 Case study: Home currency bias in the bond market

We illustrate the modelling choices regarding the level of aggregation, fixed effects settings and specific cleaning procedures. In this section we analyze how holdings are affected by the currency of denomination, highlighting the importance of investor sector heterogeneity and currency preference for most studies with SHS-S data. Let us start with the cross-sectional model following Equation (1a) using a EUR-denomination dummy, issued by a euro area issuer dummy and the market value of the amount outstanding (size) as explanatory variables in the vector *X* for reference period 2022-Q2. For each separate column we perform the following (dis)aggregations: (i) at the euro area level, (ii) at the investor-

[15] However, while inclusion of interaction terms is preferable in most contexts, the size of the number of variables included in any such regressions typically becomes too 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).

country level, (iii), at the investor-sector level and (iv), at the investor-country and sector level, where we introduce investor sector-specific interaction terms, highlighting the importance of investor heterogeneity (see Boermans & Vermeulen, 2020).

Table 8: Investor heterogeneity and home-currency preferences (*ln Hold,* cross-section)

	(1)	(2)	(3)			(4)		
	EA	EA_j	EA_s	banks	omfis	insur	pfund	hhold
EUR	1.43***	1.50***	1.39***	1.71***	0.52***	2.10***	0.71***	1.22***
	[0.011]	[0.011]	[0.009]	[0.020]	[0.023]	[0.019]	[0.024]	[0.020]
ln Size	0.84***	0.58***	0.51***	0.59***	0.37***	0.59***	0.47***	0.43***
	[0.001]	[0.001]	[0.001]	[0.003]	[0.004]	[0.003]	[0.007]	[0.002]
Observations	177,854	284,110	355,474			355,474		
R-squared	0.756	0.507	0.547			0.440		
Investor country FE	N/A	Yes	N/A			N/A		
Investor sector FE	N/A	N/A	Yes			Yes		
Issuer country FE	Yes	Yes	Yes			Yes		

Notes: 2022-Q2 cross-section. Robust standard errors in brackets (clustered) *** p<0.01, ** p<0.05, * p<0.1.

Our cross-section findings highlight the importance of home currency for investors. Table 8 Column (1) shows that at the security-level both currency and asset size are key drivers of the portfolio allocations in bond markets. Generally, euro area investor hold 143 percent more in home currency bonds. In Columns (2) and (3) these effects are robust at the investor country-level and sector-level including fixed effects. In Column (4) we further allow for the home-currency preference for EUR-denominated debt to vary across investor sectors. This is our preferred specification as these findings underscore significant differences across investor sector. For example, keeping all else constant, other financial institutions hold on average 52 percent more of a EUR-denominated bonds compared to non-EUR debt, while for insurance corporations the home-currency denomination is associated with 210 percent higher investment. In the aggregated specifications the effects for homogeneous investors are between 139 percent and 150 percent, thus hiding the variance at investor sector level. This result shows that all euro area investors have strong home currency preferences (see

Boermans & Burger, 2021, *forthcoming*) but with notable differences across investor sectors (see Boermans & Vermeulen, 2020).

In a multidimensional time-series panel setting home currency preference are also persistent. Now we move to the full-panel model following Equation (1b) where we retain all fixed effects and include the interaction terms for investor-sector specific effects for the EUR-denomination and Size, as well as time fixed effects for 2013Q4-2022-Q2. Our panel settings are at *i,s,j* for a given quarter *t* with standard errors clustered at the investor sector. Table 9 Column (1) shows that euro area investors have a tendency to invest 140 percent more in home currency bonds, very consistent with the cross-sectional analysis for 2022-Q2 in Table 8 Column (1). However, Table 9 Column (2) demonstrates the importance of the euro is not equal across different investors, showing the importance of investor heterogeneity (similar to Boermans & Vermeulen, 2020). Again, in this preferred specification especially insurance corporations have a strong preference for home currency bonds, but also households, banks and to still a great extent pension funds, whereas for other financial institutions this effect is absent.[16] The results highlight that for this case study the cross-sectional results in Table 8 are close to those from the time-series in Table 9 and robust.

Table 9: Investor heterogeneity and home-currency preferences (*ln Hold*, panel)

	(1)			(2)		
	all	banks	omfis	insur	pfund	hhold
EUR	1.40***	1.49***	0.20	2.33***	1.07***	1.99***
	[0.248]	[0.159]	[0.232]	[0.268]	[0.136]	[0.209]
ln Size	0.56***	0.66***	0.40***	0.65***	0.52***	0.53***
	[0.034]	[0.035]	[0.035]	[0.020]	[0.029]	[0.023]
Observations	11,783,139			11,783,139		
R-squared	0.541			0.436		
Time FE	Yes			Yes		
Investor country FE	Yes			Yes		
Investor sector FE	Yes			Yes		
Investor Country*Sector FE	Yes			Yes		
Issuer country FE	Yes			Yes		
Weighted (lag dependent)	No			No		

Notes: Time period 2013Q4-2022Q2. Robust standard errors in brackets (clustered) *** p<0.01, ** p<0.05, * p<0.1.

6 A SHS research: an agenda for the future?

[16] These results are robust when we apply a weighted regression with the lagged dependent variable as analytical weight. In Appendix 4 we discuss these findings. One notable downside of applying such weights is that one losses many observations, which is not random because newly issued bonds are consistently dropped with a lagged dependent variable weight as it is impossible to hold prior positions before issuance.

Granular data allows much flexibility required to understand unique crisis periods. One of the main lessons of the financial crisis 2008-2009 from a data perspective, the need for granular data, has initiated the Securities Holdings Statistics harmonized within the euro area. Clearly, ongoing research and future research will benefit from analyzing the next crises with this granular portfolio holdings data. For example, the onset of the Covid-19 pandemic in 2020 has already spurred research with SHS data (e.g. Breckenfelder & Ivashina, 2021; Carvalho & Schmitz, 2021; Breckenfelder & Schepens, 2022; di Iasio et al. 2022). Similarly, new work on spillovers from the recent interest rate hikes by the Fed and ECB from mid-2022 onwards will follow (see e.g. Boermans et al. 2016 on the Taper Tantrum in 2013). Correspondingly, the literature on international investment patterns is able to analyze Brexit (Carvalho & Schmitz, 2022) and emerging market crises (Boermans & Burger 2021, *forthcoming*).

New contributions on monetary policy will be able to evaluate new programs impacts on financial markets and securities holdings. These can also analyze the side-effects of quantitative easing and other unconventional monetary operations in the future. While SHS data collection officially started in 2014, only a year later in January 2015 the ECB announced their asset purchases program of government bonds. The available of SHS data has greatly benefitted policy work on monetary policy and also provoked wide ranging research (see e.g. Albertazzi et al 2021; Koijen et al. 2021). Still, the ESCB also has information on the Eurosystem holdings for monetary operations, which are not available for research. Allowing researchers to work with those data could increase the precision of the identification of program effectiveness and pricing implications of the PSPP and CSPP among others. This could greatly encourage more research on monetary policy as other central banks around the world do not have such granular data available (even though some scholars already use Eurosystem daily transactions data, see Appendix 1).

Another take-away from this literature review is that the transaction data is underutilized. Researchers working with the gross flows data have highlighted the importance of net transactions data, for example for testing portfolio rebalancing (Bergant et al. 2020) and short-term responses to global shocks (Boermans & Burger, 2021, *forthcoming*) and monetary policies (Arrata et al. 2020; Hudepohl, 2022). Future research will not only expand using the transactions data but studies on the financial performance of portfolio investments are still at their infancy (for exceptions, see Boermans et al. 2022; Boermans & Galema, 2022). The SHS data includes the price revaluations and exchange rate effects,

which have to date not been applied in research. This module in SHS enables direct evaluations of price impacts and foreign exchange rate effects on portfolio positions.

So far the full dimensionality of securities data has not been fully analyzed for research purposes. For example, looking at different debt types a multitude of areas can been explored such as green bonds, asset-backed securities and covered bonds in all its detail (think of residential mortgage-backed securities or car loan securitizations), holdings of bonds that are in distress, Islamic financed bonds or bonds with negative yields. Similarly, conceptualizations around specific issuer industries are of interest, as have been done only for energy intensive stocks in the sustainable finance theme (see Section 4.4) or with focus on government holdings. Here, analyses of particular industries are almost endless, such as a focus on aircraft industry, tech-giants, car firms, health providers, bank sector ownership or positions in supranational organizations such as the newly issued EU bonds under the NGEU and SURE programs.

While SHS data can be linked to other databases, this area is still underutilized while showing great promises. We anticipate more and more integration on International Security Identifiers (ISIN) and Legal Entity Identifiers (LEI). We can learn from researchers how to efficiently merge SHS with granular loan information for AnaCredit, the Secured Financing Transaction Regulation (SFTR) and EMIR data on derivatives, supervisory data such as COREP, FINREP and IBSI, but also commercial sources on bank's balance sheets such as BankScope and Orbis or credit ratings and CO_2 emission data. At he moment such project are only undertaken partially and individually. Decisions in joining datasets need to be transparant to create a data community with feedback loops to learn for experiences. For example, grouping entity and issuer level data to consolidated firm level data is becoming increasingly important (see Coppala et al. 2021). Enhanced programming power, data sharing of common practices and coding will facilitate such large scale data initiatives for better research on portfolio investments.

One area that overshadows this literature review is all the internal and external work by policy makers and supervisors with SHS which is not labelled research. While academic research in the area of financial regulation and supervisory is in a developmental stage, regulatory authorities have conducted a multitude of analyses based on SHS data not reviewed in this academic literature review (see Table 1). These works for most part have not been published publicly, e.g. think how resolution authorities like the SRB benefit from

understanding the scope of a resolution decision when able to also analyze the first-order and second-order impacts when a banking group fails. Still, various reports from authorities such as the ECB, EIPOA and the ESRB include a wide range of studies with SHS for the public policy debate. To illustrate this with a very recent example, the ECB Financial Stability Report of November 2022, a ECB flagship publication twice a year, includes 9 graphs that were based on SHS data, highlighting various topics on credit risk, short positions in sovereign debt markets, non-bank financial sector transactions, losses from debt revaluations, downgrade effects, energy industry exposures and the investor base of sovereign debt in relation to monetary policy (see ECB, 2022). As researchers are catching up using the full scope of the SHS data for research, this will greatly synergize with ongoing policy work, making it easier for both to swiftly analyze new questions at bay.

7 Conclusion

This literature review has studied 62 studies using Securities Holdings Statistics at granular portfolio levels between 2016-2022. While these SHS papers fill important data gaps, we argue that the rise of studies employing SHS still are likely to be in an early stage of development as signified by more than half of these paper (32) published between 2021-2022. This paper has discussed these important contributions by classifying each paper by six different research theme. Most prominently we find that the field most rapidly developing is on sustainable finance. Overall we expect that in a few year from now that number of studies can further multiply most strongly in the research themes we proposed in Section 4, but also in new fields highlighted for future research in Section 6 due to the kaleidoscopic nature and flexibility of granular data. Our practitioner guide can further facilitate harmonized conduct among researchers in the field. For example, our case study in Subsection 5.3 on home currency preferences highlights the importance of currency and investor sector heterogeneity for portfolio choice models. provides new researchers an opportunity to quickly adapt and apply this large and growing dataset to their advantage using insights from previous research and best practices in the future.

References

Abidi, N., Amato, L., Miquel-Flores, I., & Vandeweyer, Q. (2021). The Bright Side of Transparency: Evidence from Supervisory Capital Requirements. *Mimeo*.

Accornero, M., Finaldi Russo, P., Guazzarotti, G., & Nigro, V. (2018). Missing investors in the Italian corporate bond market. *Bank of Italy Occasional Paper 450.*

Aghion, P., Boneva, L., Breckenfelder, J., Laeven, L., Olovsson, C., Popov, A. A., & Rancoita, E. (2022). Financial Markets and Green Innovation. *ECB Working Paper No. 2686*.

Ahmed, A., Hofmann, B., & Schmitz, M. (2022). Foreign institutional investors, monetary policy, and reaching for yield. *Mimeo*.

Albertazzi, U., Becker, B., & Boucinha, M. (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., Hüser, A. C., & Kok, C. (2022). Contagion accounting in stress-testing. *Journal of Economic Dynamics and Control,* 137, 104354.

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

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

Anand, K., van Lelyveld, I., Banai, Á., Friedrich, S., Garratt, R., Hałaj, G., ... & de Souza, S. R. S. (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., & Van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review,* 93(1), 170-192.

Aragon, G. O., Hertzel, M., & Shi, Z. (2013). Why do hedge funds avoid disclosure? Evidence from confidential 13F filings. *Journal of Financial and Quantitative Analysis,* 48(5), 1499-1518.

Arrata, W., Nguyen, B., Rahmouni-Rousseau, I., & Vari, M. (2020). The scarcity effect of QE on repo rates: Evidence from the euro area. *Journal of Financial Economics,* 137(3), 837-856.

Attinà, C. A., & Bologna, P. (2021). TLAC-eligible debt: who holds it? A view from the euro area. *Bank of Italy Occasional Paper 604.*

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

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

Bergant, K., Fidora, M., & Schmitz, M. (2020). International capital flows at the security level–evidence from the ECB's asset purchase programme. *IMF Working Paper No. 20/46.*

Boermans, M.A., & van der Kroft, B. (2020). Capital regulation induced reaching for systematic yield: Financial instability through fire sales. *DNB Working Papers No. 673*.

Boermans, M. A., & Van Wijnbergen, S. (2018). Contingent convertible bonds: Who invests in European CoCos?. *Applied Economics Letters,* 25(4), 234-238.

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

Boermans, M. A., & Vermeulen, R. (2020). International investment positions revisited: Investor heterogeneity and individual security characteristics. *Review of International Economics,* 28(2), 466-496.

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

Boermans, M., & Burger, J. D. (2021, *forthcoming*). Fickle Emerging Market Flows, Stable Euros, and the Dollar Risk Factor. *Journal of International Economics.*

Boermans, M., & Galema 2020. Carbon home bias of European investors, *Mimeo*.

Boermans, M., & Keshkov, V. (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., & Vermeulen, R. (2018). Quantitative easing and preferred habitat investors in the euro area bond market. *DNB Working Paper No. 586.*

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

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

Brand, C., Ferrante, L., & Hubert de Fraisse, A. (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., & Ivashina, V. (2021). Bank balance sheet constraints and bond liquidity. *ECB Working Paper No. 2589.*

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

Bubeck, J., Maddaloni, A., & Peydró, J. L. (2020). Negative monetary policy rates and systemic banks' risk-taking: Evidence from the euro area securities register. *Journal of Money, Credit and Banking*, 52(S1), 197-231.

Carletti, E., De Marco, F., Ioannidou, V., & Sette, E. (2021). Banks as patient lenders: Evidence from a tax reform. *Journal of Financial Economics*, 141(1), 6-26.

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

Carvalho, D., & Schmitz, M. (2021). Shifts in the portfolio holdings of euro area investors in the midst of COVID-19: looking-through investment funds. *ECB Working Paper No. 2526.*

Carvalho, D., & Schmitz, M. (2022). Brexit, what Brexit? Euro area portfolio exposures to the United Kingdom since the Brexit referendum. *ECB Working Paper No. 2734*.

Coppola, A., Maggiori, M., Neiman, B., & Schreger, J. (2021). Redrawing the map of global capital flows: The role of cross-border financing and tax havens. *Quarterly Journal of Economics*, 136(3), 1499-1556.

Darmouni, O., & Papoutsi, M. (2022). The rise of bond financing in Europe. *ECB Working Paper No. 2663*.

de Haan, L., & Vermeulen, R. (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.

Della Corte, V., & Federico, S. (2016). Foreign holders of Italian government debt securities: new evidence. *Bank of Italy Occasional Paper No. 363*.

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

Dick-Nielsen, J. (2009). Liquidity biases in TRACE. *Journal of Fixed Income*, 19(2), 43-55.

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

di Iasio, G., Alogoskoufis, S., Kordel, S., Kryczka, D., Nicoletti, G., & Vause, N. (2022). A model of system-wide stress simulation: market-based finance and the Covid-19 event. *Bank of Italy Occasional Paper 687.*

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

Dubiel-Teleszynski, T., Franch, F., Fukker, G., Miccio, D., Pellegrino, M., & Sydow, M. (2022). System-wide amplification of climate risk. *ECB Macroprudential Bulletin*, 17.

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

ECB (2022). Financial Stability Review, 14 November.

Edison, H. J., & Warnock, F. E. (2004). US investors' emerging market equity portfolios: a security-level analysis. *Review of Economics and Statistics*, 86(3), 691-704.

Elton, E. J., Gruber, M. J., & Blake, C. R. (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(6), 2415-2430.

Eser, F., Lemke, W., Nyholm, K., Radde, S., & Vladu, A. (2019). Tracing the impact of the ECB's asset purchase programme on the yield curve. *ECB Working Paper No. 2293*.

Fache Rousová, L. F., Ghio, M., Salakhova, D., & Bauer, G. V. (2022). Derivative margin calls: A new driver of MMF flows? *Mimeo*.

Fache Rousová, L.F., & Giuzio, M. (2019). Insurers' investment strategies: pro-or countercyclical?. *ECB Working Paper No. 2299*.

Fache Rousová, L.F. and Rodríguez Caloca, A. (2014). The use of Securities Holdings Statistics (SHS) for designing new euro area financial integration indicators, *IFC Bulletin, No. 39*.

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

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

Franzoni, F., & Schmalz, M. C. (2017). Fund flows and market states. *Review of Financial Studies*, 30(8), 2621-2673.

Fricke, D., Greppmair, S., & Paludkiewicz, K. (2022). You can't always get what you want (where you want it): Cross-border effects of the US money market fund reform. *Deutsche Bundesbank No. 03/2022*.

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

Grandia, R., Hänling, P., Lo Russo, M., & Åberg, P. (2019). Availability of high-quality liquid assets and monetary policy operations: an analysis for the euro area. *ECB Occasional Paper 218*.

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

Henricot, D., & Piquard, T. (2022). Credit Default Swaps and Credit Risk Reallocation. *Banque du France Working Paper No. 860*.

Hu, G., Jo, K. M., Wang, Y. A., & Xie, J. (2018). Institutional trading and Abel Noser data. *Journal of Corporate Finance*, 52, 143-16

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., & Kok, C. (2019). Mapping bank securities across euro area sectors: comparing funding and exposure networks. *Journal of Network Theory in Finance*, 5(2), 59-92.

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

Jasova, M., Laeven, L., Mendicino, C., Peydró, J. L., & Supera, D. (2021). Systemic risk and monetary policy: The haircut gap channel of the lender of last resort. *CPER Discussion Paper No 16240.*

Jourde, T. (2022). The rising interconnectedness of the insurance sector. *Journal of Risk and Insurance*, 89 (2), 397-425.

Kabaca, S., Maas, R., Mavromatis, K., & Priftis, R. (*forthcoming*). Optimal quantitative easing in a monetary union. *European Economic Review*.

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

Lamas, M., & Martinez-Miera, D. (2021). Sectorial holdings and stock prices: The household-bank nexus. *Banco de España Working Paper No. 2130*.

Lane, P., & Milesi-Ferretti, G.M. (2008). International investment patterns. *Review of Economics & Statistics*, 90(3), 538-554.

Lavrador, I., Peronaci, R. and Silva, N. (2012). Security-by-security data on holdings of securities: the importance for national and euro area accounts, *IFC Bulletin, No. 36.*

Lucia, A., Battiston, S., & Kvedaras, V. (2022). Over with carbon? Investors' reaction to the Paris Agreement and the US withdrawal. *Mimeo*.

Maddaloni, A. M., & Scardozzi, G. (2022). The New Bail-In Legislation. Studies in Banking and Financial Institutions. *Palgrave Macmillan.*

Martino, E. D. (2021). Towards an optimal composition of bail-inable debtholders? *Journal of Corporate Law Studies*, 21(2), 321-364.

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

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

Montagna, M., Torri, G., & Covi, G. (2021). On the origin of systemic risk. *Bank of England Staff Working Paper No. 906.*

Papoutsi, M., Piazzesi, M., & Schneider, M. (2021). How unconventional is green monetary policy. *JEEA-FBBVA Lecture at the ASSA.*

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

Puy, D. (2016). Mutual funds flows and the geography of contagion. *Journal of International Money and Finance*, 60, 73-93.

Radke, T., Caloca, A. R., & Schmitz, M. (2021). The more the merrier: Enhancing traditional cross-border portfolio investment statistics using security-by-security information. *Proceedings of 63rd ISI World Statistics Congress*, 11.

Ringe, W. G., & Patel, J. (2019). The dark side of bank resolution: counterparty risk through bail-in. *European Banking Institute Working Paper No. 31*.

Schwarz, C. G., & Potter, M. E. (2016). Revisiting mutual fund portfolio disclosure. *Review of Financial Studies*, 29(12), 3519-3544.

Sola, P. and Strobbe F. (2010). Addressing data gaps revealed by the financial crisis: European Central Bank statistics on holdings of securities, *IFC Bulletin, No. 34*.

Squartini, T., Almog, A., Caldarelli, G., Van Lelyveld, I., Garlaschelli, D., & Cimini, G. (2017). Enhanced capital-asset pricing model for the reconstruction of bipartite financial networks. *Physical Review E*, 96(3), 032315.

Sydow, M., Schilte, A., Covi, G., Deipenbrock, M., Del Vecchio, L., Fiedor, P., ... & Tente, N. (2021). Shock amplification in an interconnected financial system of banks and investment funds. *ECB Working Paper No. 2581*.

Ullersma, C., & van Lelyveld, I. (2022). Granular data offer new opportunities for stress testing. *Handbook of Financial Stress Testing*, 185.

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

Wilson, C., & Caldecott, B. (2021). Breaking the bond: Primary markets and carbon-intensive financing. *Mimeo*.

Appendix 1: Other portfolio holdings databases – in brief

There are several other datasets with granular portfolio holdings that are widely used in research. In this part, we list a selection frequently used in research (see also Hu et al. 2018). First, the Trade Reporting and Compliance Engine (TRACE) database includes with bond-level transactions and pricing data. TRACE includes reporting of over-the-counter transactions of bonds as mandated by the Securities and Exchange Commission (SEC) rules, yielding valuable bond-level trade data since 2002 (see e.g. Dick-Nielsen, 2009, 2014). One advantage of TRACE is that it record individual transactions, whereas in SHS one only observes netted gross transactions at the security level over the quarter. However, the scope of TRACE is focused on bonds and US-centric for only large investors. Second, in the US key investors must file 13-F forms to the SEC with line-by-line data on portfolio positions, also spurring research on international investment (e.g. Edison & Warnock, 2004) and financial markets (e.g. Aragon et al. 2013). The advantage of these 13-F filings is that they cover most assets classes and various investor sectors, but like TRACE are focused on US investors.

While TRACE and 13-Filings are US-centric, international data collection has been available mainly for mutual funds. A large share of global mutual funds provide information on their holdings which are collected by various commercial data providers. Security-by-security holdings of mutual funds are provided by Center for Research in Security Prices (CRSP) (Franzoni & Schmalz, 2017), Morningstar (Elton et al. 2001), LionShares, Bloomberg (e.g. Wilson & Ben Caldecott, 2021), Lipper database (e.g. Borgers et al. 2015; Schwarz & Potter, 2016; Boermans et al. 2022) and EPFR focusing on bond flows (Puy, 2016). The advantage over SHS-S is again individual investor data but more fragmented coverage for only specific sectors in the economy.

It is beyond scope of this review to enlist all securities database but two others stand out. First, over 50 papers use Abel Noser data which provides high-frequency data from a broker which includes transaction costs data from institutional clients (see Hu et al. 2018). Several researchers have also used data from central bank portfolios including those from Canada, Japan and Switzerland (see e.g. Brand et al. 2019 for an ECB perspective).

Appendix 2: Reconciliation in SHS data

Most studies use holdings data but more reconciliation posts are available. While most papers rely on positions data, the full flow components, including the net transactions are available in the SHS dataset which leads to a full reconciliation (as indicated with the plus+ and minus- signs and the " amount type" codes in the SHS).

Table A.1: Positions, transactions and full reconciliation in the SHS

Amount type	Flow components	Description
LE_{t-1}	position at t-1	Total amount held in euro of a security at the end of period t-1.
T_t	+/- transactions	Net (gross) transactions. Sum of purchases minus sales of a security, recorded at transaction value in euro including accrued interest.
$K7A_t$	+/- currency movements	Revaluations due to exchange rate changes, refering to movements in exchange rates against the euro that occur between end-period reporting dates.
$K7B_t$	+/- price effects	Revaluations due to other price changes being the price revaluations that include changes in the value of end-period positions that occur in the reference period because of changes in the reference value at which they are recorded, i.e. holding gains or losses.
KA_t	+/- other changes in volume	Other changes in the volume of assets refer to changes on the investor side or on the issuer side. Changes on the investor side can be caused among others by reclassification of institutional units and assets. Changes on the issuer side (also depicted as PICV = Pseudo identifier for other changes in volume) refer to changes in reference data attributes such as issuer sector or issuer country.
LE_t	position at t	Total amount held in euro of a security at the end of period t.

Notes: information based on ECB internal documentation.

Appendix 3: Regression results using *EA_SHARE*

While analyzing holdings is the standard while controlling for asset size, one can also define the investment share as dependent variable. One alternative approach for ease of interpretation of the estimated coefficients in relation to a I-CAPM model is to reshuffle this right hand side explanatory variable Size at market value to the left hand side to obtain the share an investor has in the particular security (as alternative to add size as explanatory variable in Equation 3):

$$(4)\ \frac{HOLD_{i,s,j,t}}{MarketValue_{i,t}} = \beta_s \ \ X_{i,s,j,t} + FE_{s,j} + FE_{h,c} + \ \varepsilon_{s,j,t}$$

where the dependent variable can be multiplied with a factor 100 to obtain the share in percentages (see e.g. Lucia et al. 2022). When taking a euro area investor perspective, following Boermans and Burger (2021, *forthcoming*), this approach to make holdings relative to the size of the security translates in aggregate form in the share of a security held by the euro area (*EA_Share*) building on Equation (2):

$$(5)\ EA\_Share_{i,t} = \beta' \ \ X_{i,s,j,t} + FE_{s,j} + FE_{h,c} + \varepsilon_{s,j,t}$$

In this section we show the consistency when using *EA_SHARE* as an alternative, as long as one controls for the size of the asset, e.g. stock market capitalization on bond amount outstanding at market value. The findings presented here follow Section 5.3, Table 8 with the holdings dividend by the size of the bond at market values. Table A.3.1 Column (1) uses euro area investor aggregated data and show that the *EA_SHARE* increases by 44 percent if the bond is denominated in the home euro currency. This effects becomes smaller in economic terms once allowing for investor heterogeneity but is still significant in Columns (2) and (3). Finally, our preferred specification in Table A.3.1 Column (4) confirm that insurance corporations display the stronger home-currency preference, indicating that the share held by insurances from a given investor country tends to increase by 26 percent if the bond is in the home currency denomination.

Table A.3.1: Investor heterogeneity and home-currency preferences

	(1)	(2)	(3)			(4)		
	EA	EA_j	EA_s	banks	omfis	insur	pfund	hhold
EUR	0.44***	0.21***	0.13***	0.22***	0.01**	0.26***	0.10***	0.24***
	[0.003]	[0.002]	[0.002]	[0.004]	[0.004]	[0.003]	[0.003]	[0.003]
Observations	177,854	284,110	355,474			355,474		
R-squared	0.514	0.394	0.306			0.351		
Investor country FE	N/A	Yes	N/A			N/A		
Investor sector FE	N/A	N/A	Yes			Yes		
Issuer country FE	Yes	Yes	Yes			Yes		

Notes: Period 2022-Q2. Dependent variable is *ln HOLD* over *ln Size* (at market values). Robust standard errors in brackets (clustered) *** p<0.01, ** p<0.05, * p<0.1.

Appendix 4: Weighted regressions

Given the large differences in the size of the holdings one may consider using weighted regressions. In this section we show that while weighting might be appropriate, there are no strict indicators suggesting one must apply weighted regressions. Table A.4.1 is similar to Table 9, except that the lagged dependent variable, *ln* $HOLD_{t-1}$, is used to weight each observation in terms of it's contribution to the averaged estimated coefficients.

Table A.4.1: Weighted time-series panel regressions for *ln HOLD*

	banks	omfis	insur	pfund	hhold
EUR	1.23***	0.17	2.25***	0.87***	2.20***
	[0.162]	[0.216]	[0.235]	[0.146]	[0.235]
ln Size	0.68***	0.38***	0.58***	0.44***	0.53***
	[0.033]	[0.037]	[0.022]	[0.028]	[0.025]
Observations			7,825,089		
R-squared			0.419		
Time FE			Yes		
Investor country FE			Yes		
Investor sector FE			Yes		
Investor Country*Sector FE			Yes		
Issuer country FE			Yes		
Weighted (lag dependent)			Yes		

Notes: Time period 2013Q4-2022Q4. Robust standard errors in brackets (clustered) *** p<0.01, ** p<0.05, * p<0.1.

DeNederlandscheBank

EUROSYSTEEM

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