Financial stability: New, detailed datasets allow for innovation of stress tests

By Justus Inhoffen and Iman van Lelyveld

- To better identify risks in the financial services industry, data reported to financial supervisors need to be more detailed
- Such granular data allows authorities to conduct better stress tests to assess the stability of the financial system
- Empirical analysis suggests that less detailed datasets may lead to biased results and incorrect policy recommendations
- Market participants find effort required to comply with data reports too large in relation to their benefits
- There seems to be potential for efficiency gains in the harmonization and standardization of reporting obligations that can alleviate compliance costs

Granular data make the interconnectedness more transparent which allows to better identify the sources of financial stress, schematic representation

FROM THE AUTHORS

“The financial crisis has shown that it is inevitable to increase the transparency of financial markets. However, lawmakers should act with prudence in order not to overstrain especially small market participants.”

— Justus Inhoffen —
ABSTRACT

The 2008-2010 crisis has shown that authorities were missing crucial information necessary to identify risks to the financial system in an accurate and timely manner. To be prepared for future crises, a range of legislation in Europe and beyond was passed. The scope and depth of information being reported from across the financial system, including previously disregarded segments, have thus significantly increased. With a focus on stress tests, we examine how granular datasets increase the transparency of the financial system’s interconnectedness as well as the health of its market participants. We show that risks propagate primarily in the asset class of equities and via strong domestic links. This makes the data collection necessary to identify previously opaque risk drivers and derive effective policy recommendations despite large costs. The oft-bemoaned compliance costs for market participants may be alleviated if reporting requirements can be harmonized and data collection processes can be standardized.

More than one decade after the crisis, the reform of the financial services industry is still underway. Results of stress tests for the banking industry and the financial system raise awareness about the vulnerability of specific market participants. New European financial regulations seek to provide a remedy by strengthening financial stability. Nevertheless, it is still debated whether the collected data is adequate to sufficiently identify all risk drivers.

Besides the positive effect of new regulations on the transparency of financial systems, these structural interventions fundamentally affect market participants on nearly every level of their organizations, especially if they are related to the reporting obligations. This diverts needed resources away from day-to-day activities.

At the same time, the financial services industry experiences tremendous pressure on profitability in times of low interest rates, diminishing availability of profitable investment opportunities, and competition from technology companies. Stakeholders frequently lament the disproportion between the benefit of new regulations and the cost of compliance.

In this article, we discuss how more detailed data enable better stress tests and more transparency. In an illustration, we empirically investigate the interconnectedness among market participants in the euro area using a new dataset from the European System of Central Banks (ESCB). This is based on a network of common asset holdings between economic sectors through which shocks can propagate in the financial system. We analyze which asset classes are most prone to risk propagation, whether domestic links or links across euro-area countries drive propagation, and which country-sector pairs are most vulnerable. Finally, we examine how less detailed pre-crisis datasets may lead to biased stress test results, as they miss many crucial interactions.

1 Views expressed are those of the authors and do not necessarily reflect official positions of the Dutch central bank, De Nederlandsche Bank, where the author Iman van Lelyveld works.

DOI: https://doi.org/10.18723/diw_dwr:2020-3-1
Opportunities and challenges of granular data

Together with the scope of the data submitted to national authorities, the level of detail, or the granularity, of information has increased. Its most important property is that it allows for made-to-measure aggregations and thus can serve multiple purposes.

Traditionally, it was the task of the reporter to classify a particular data point. However, it is now more common to ask for information that is much more detailed and leave the classification to the authorities. For example, when reporting a transaction under EMIR, reporters no longer need to determine the sector of the counterparty themselves, but only supply the Legal Entity Identifier (LEI), a unique identification number for legal entities. Using the LEI, data can, for instance, be aggregated to a financial group or a sector. This ensures a consistent handling of the sector allocation across datasets, for example even when a sector assignment of a particular firm changes.

Challenges in data collection

A challenge in the collection of granular data is the legal framework that governs them. Depending on the institutional framework of a country, the legal mandates are assigned to different agencies. Despite the global nature of financial markets, sharing data among them or across jurisdictions is difficult. Given the oft-complex governance, it is no wonder that reporting requirements are a patchwork of overlapping or differing components. Accordingly, metadata are often not well coordinated.

Opportunities for supervision and stress testing

With data on the entire financial system, authorities will be able to deliver the raw material for stress tests of the financial system (macroprudential) or individual financial institutes (microprudential). Using the very same dataset, they can seamlessly zoom from the macro level all the way down to individual transactions of individual market participants. Moreover, they can also connect markets and see what market participants or sectors are doing in different securities. A microprudential supervisor will be looking at the solvency of a particular firm while a macroprudential authority will look for vulnerabilities at the macro level. With granular data, views for each of these, also in different time scales, can be generated from a single data source.

Changes to the European regulatory framework have many positive effects, but at high costs

Since the financial crisis, the European Commission has adopted more than 40 legislative acts in accordance with G20 partners to establish a sustainable financial industry in Europe.2

The breadth of reporting increased especially in the area of over-the-counter (OTC) markets, a segment of the financial industry that was previously opaque to the supervisor because counterparties transact directly with each other outside of regulated exchanges. One far-reaching regulation was the European Market Infrastructure Regulation (EMIR) from 2010, which requires parties with significant derivative dealings to centrally report their transactions to trade repositories. Other examples are the Analytical Credit Datasets (AnaCredit), which provides a basic version of a European credit register, and the Money Market Statistical Reporting (MMSR), which provides deal level information of large banks’ activities in the money market.

In light of the increasing complexity of the financial services industry, supervisors stress the need for this level of information to achieve the regulatory objectives. Market surveillance and risk assessment in particular have greatly improved for national authorities.3

Furthermore, these new granular datasets (Box 1) have positive effects for policy work. For example, the MMSR, which is actually collected to construct interest rate benchmarks, also allows for a better assessment of the monetary policy transmission mechanism and thus enables policy makers to identify more effective policy measures. In times of the unequal distribution of excess liquidity within and across euro area jurisdictions, the central bankers’ job has become more reliant on such high-level information.

Positive effects potentially also arise for reporting agents. Data collected for reporting purposes can augment their own risk management models and practices. It can further lead reporting agents to think of their IT as a profit instead of a cost center. Building on a new infrastructure, they can systematically store the breadth of their business data, often scattered in siloed systems, and make it available internally to improve the conduct of business in a range of use cases.4

Box 1

Opportunities and challenges of granular data

Together with the scope of the data submitted to national authorities, the level of detail, or the granularity, of information has increased. Its most important property is that it allows for made-to-measure aggregations and thus can serve multiple purposes.

Traditionally, it was the task of the reporter to classify a particular data point. However, it is now more common to ask for information that is much more detailed and leave the classification to the authorities. For example, when reporting a transaction under EMIR, reporters no longer need to determine the sector of the counterparty themselves, but only supply the Legal Entity Identifier (LEI), a unique identification number for legal entities. Using the LEI, data can, for instance, be aggregated to a financial group or a sector. This ensures a consistent handling of the sector allocation across datasets, for example even when a sector assignment of a particular firm changes.

Challenges in data collection

A challenge in the collection of granular data is the legal framework that governs them. Depending on the institutional framework of a country, the legal mandates are assigned to different agencies. Despite the global nature of financial markets, sharing data among them or across jurisdictions is difficult. Given the oft-complex governance, it is no wonder that reporting requirements are a patchwork of overlapping or differing components. Accordingly, metadata are often not well coordinated.

Opportunities for supervision and stress testing

With data on the entire financial system, authorities will be able to deliver the raw material for stress tests of the financial system (macroprudential) or individual financial institutes (microprudential). Using the very same dataset, they can seamlessly zoom from the macro level all the way down to individual transactions of individual market participants. Moreover, they can also connect markets and see what market participants or sectors are doing in different securities. A microprudential supervisor will be looking at the solvency of a particular firm while a macroprudential authority will look for vulnerabilities at the macro level. With granular data, views for each of these, also in different time scales, can be generated from a single data source.

Changes to the European regulatory framework have many positive effects, but at high costs

Since the financial crisis, the European Commission has adopted more than 40 legislative acts in accordance with G20 partners to establish a sustainable financial industry in Europe.2

The breadth of reporting increased especially in the area of over-the-counter (OTC) markets, a segment of the financial industry that was previously opaque to the supervisor because counterparties transact directly with each other outside of regulated exchanges. One far-reaching regulation was the European Market Infrastructure Regulation (EMIR) from 2010, which requires parties with significant derivative dealings to centrally report their transactions to trade repositories. Other examples are the Analytical Credit Datasets (AnaCredit), which provides a basic version of a European credit register, and the Money Market Statistical Reporting (MMSR), which provides deal level information of large banks’ activities in the money market.

In light of the increasing complexity of the financial services industry, supervisors stress the need for this level of information to achieve the regulatory objectives. Market surveillance and risk assessment in particular have greatly improved for national authorities.3

Furthermore, these new granular datasets (Box 1) have positive effects for policy work. For example, the MMSR, which is actually collected to construct interest rate benchmarks, also allows for a better assessment of the monetary policy transmission mechanism and thus enables policy makers to identify more effective policy measures. In times of the unequal distribution of excess liquidity within and across euro area jurisdictions, the central bankers’ job has become more reliant on such high-level information.

Positive effects potentially also arise for reporting agents. Data collected for reporting purposes can augment their own risk management models and practices. It can further lead reporting agents to think of their IT as a profit instead of a cost center. Building on a new infrastructure, they can systematically store the breadth of their business data, often scattered in siloed systems, and make it available internally to improve the conduct of business in a range of use cases.4
The impact on the financial industry

For a recently published study, the European Commission consulted around 400 respondents from 15 member states regarding the fitness and suitability of EU supervisory reporting requirements. The results show, among others, that reporting requirements account for 30 percent of respondents’ total compliance costs on average. The European Commission estimates that the ongoing annual supervisory reporting costs of the entire industry ranges north of four billion euros.

Major challenges arise from the high level of data granularity in regulatory reportings. The inadequate IT infrastructure is a particularly strong cost driver: more than half of the respondents state that they are not very well adapted to IT developments which poses challenges to automating reporting processes. As a consequence of, for example, merger waves in the banking sector, reporting agents are facing legacy systems that lock them into inefficient solutions like manual reporting. Another contributing factor is the need for qualified employees in a tight labor market.

Granular data make interconnectedness more transparent

A major advantage of data on a granular level is that it allows the supervisor to pinpoint the source of stress and the location of the most vulnerable parts of the financial system. Once governance challenges, technical difficulties, and conceptual issues are resolved, it is possible to paint a detailed picture of the financial system and develop needs-oriented stress test models.

To exemplify this point, we focus on the interconnectedness in the financial ecosystem. A novel granular dataset (Box 1) allows us to empirically identify an overlapping portfolio

---

**Table 1**

**Sector activities across asset classes, Q1 of 2018**

In billion euros

<table>
<thead>
<tr>
<th>Sector</th>
<th>Equities</th>
<th>Investment fund shares</th>
<th>Short-term debt</th>
<th>Long-term debt</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insurance corporations</td>
<td>161</td>
<td>1,515</td>
<td>27</td>
<td>2,742</td>
<td>4,445</td>
</tr>
<tr>
<td>Investment funds</td>
<td>1,048</td>
<td>1,147</td>
<td>72</td>
<td>1,986</td>
<td>4,252</td>
</tr>
<tr>
<td>Monetary financial intermediaries (banks)</td>
<td>141</td>
<td>233</td>
<td>114</td>
<td>2,637</td>
<td>3,115</td>
</tr>
<tr>
<td>Households</td>
<td>566</td>
<td>1,129</td>
<td>3</td>
<td>218</td>
<td>1,916</td>
</tr>
<tr>
<td>Non-financial corporations</td>
<td>853</td>
<td>334</td>
<td>14</td>
<td>84</td>
<td>1,285</td>
</tr>
<tr>
<td>Pension funds</td>
<td>38</td>
<td>481</td>
<td>5</td>
<td>349</td>
<td>874</td>
</tr>
<tr>
<td>Governments</td>
<td>153</td>
<td>193</td>
<td>4</td>
<td>260</td>
<td>610</td>
</tr>
<tr>
<td>Other financial intermediaries</td>
<td>797</td>
<td>78</td>
<td>11</td>
<td>107</td>
<td>494</td>
</tr>
<tr>
<td>Money market funds</td>
<td>0</td>
<td>31</td>
<td>214</td>
<td>94</td>
<td>339</td>
</tr>
<tr>
<td>Total</td>
<td>3,257</td>
<td>5,141</td>
<td>465</td>
<td>8,466</td>
<td>17,329</td>
</tr>
</tbody>
</table>

Sources: ECB; authors’ own calculations.

**Table 2**

**Characteristics of investments by asset class**

In absolute figures; graph density in percent, Q1 of 2018

<table>
<thead>
<tr>
<th></th>
<th>Equities</th>
<th>Investment fund shares</th>
<th>Short-term debt</th>
<th>Long-term debt</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country-sector pairs</td>
<td>129</td>
<td>135</td>
<td>100</td>
<td>143</td>
<td>146</td>
</tr>
<tr>
<td>Outstanding securities</td>
<td>3,669</td>
<td>47,848</td>
<td>45,020</td>
<td>239,879</td>
<td>336,416</td>
</tr>
<tr>
<td>Held securities (on average)</td>
<td>468</td>
<td>1714</td>
<td>530</td>
<td>3496</td>
<td>5,786</td>
</tr>
<tr>
<td>Investments</td>
<td>60,147</td>
<td>23,1392</td>
<td>53,024</td>
<td>499,952</td>
<td>844,715</td>
</tr>
<tr>
<td>Graph density</td>
<td>12.75</td>
<td>3.58</td>
<td>1.18</td>
<td>1.46</td>
<td>1.72</td>
</tr>
</tbody>
</table>

1 The graph density is the ratio of observed to potential investments.

Sources: ECB; authors’ own calculations.

---


**Dataset and methodology**

**Dataset**

We use the Securities Holdings Statistics by Sector (SHS-S) dataset, collected by the European System of Central Banks (ESCB), which covers 83 percent of all euro-denominated securities issued in euro area jurisdictions and thus is well suited for analyzing overlapping asset holdings.1 This dataset provides the euro amount of holdings in securities, identified by their International Security Identification number (ISIN), for all possible country-sector pairs. The ISIN is an example of a well-established standardized metadata for securities whose coverage is large yet not universal. The dataset classifies securities into shares listed on a recognized stock exchange (listed shares or equities), shares issued by investment funds that entitle the holder to a principal and interest (debt securities), subdivided into short-term and long-term debt.

**Methodology**

We measure the importance of common asset holdings among euro area sectors. To this end, we construct a network for each asset class separately. In this network, nodes represent the portfolio of a holding country-sector pair. If at least one security is common to two pairs’ portfolio, a link is established.

We assign each link a similarity measure that ranges from 0 to 1. In particular, we measure the portfolio overlap between country-sector pair $i$ and $j$ as the number of securities in the intersection divided by the number of securities in their union

$$\frac{|P_i \cap P_j|}{|P_i \cup P_j|}$$

where $P_i$ is the portfolio of sector $i$. This measure is commonly known as the Jaccard similarity index. The results are qualitatively robust to other definitions of the portfolio overlap measure.

To measure the risk exposure of every individual country-sector pair to which it is exposed in the network, we use eigenvector centrality. The risk measure of every market participant

$$c_i(j) = \frac{1}{|A|} \sum_{j=1}^{|A|} A_{ij} c_j(j)$$

is determined recursively by the weighted sum of risk of all connected market participants. Consequently, the measure accounts for the position in the entire system. $A$ denotes the adjacency matrix which contains the information on all links and the portfolio similarity measure. An eigenvalue decomposition yields the solution to the system of equations.

---


---

**Portfolio overlap strongest in equity segment**

We analyze the portfolio overlap based on the Securities Holdings Statistics by Sector (SHS-S) dataset. This dataset provides the euro amount of holdings in securities for all possible country-sector pairs. The dataset contains nine sectors and 19 euro area countries, resulting in a total of 171 possible country-sector pairs.20

Euro area sectors as well as euro area countries vary significantly in terms of the market value of their asset holdings, their degree of diversification, and their activity across asset classes. The largest euro area sector by size (insurance companies) is larger than the smallest sector (money market funds) by a factor of 12 (Table 1). Although all sectors are active in each asset class—except for money market funds,
FINANCIAL STABILITY

which do not hold equities—their importance varies. This is primarily due to differences in business models, but is also a consequence of, for example, regulatory requirements and risk considerations. Asset holdings aggregated to euro area countries show a similarly large spread but cannot be disclosed due to confidentiality reasons.

These descriptive statistics show that there is a potentially strong overlap among portfolios of country-sector pairs. The more securities a country-sector-pair holds, the higher the likelihood that a link establishes to another market participant. The number of observed investments as a share of all potential investments (graph density) shows that the risk of overlaps is highest on the equities-segment (Table 2).

A link in the sense of the above mechanism is only established between two country-sector pairs through a common security holding. The larger the similarity in their portfolio, the stronger the shock propagation (Box 2).

The results show that the intensity of portfolio overlaps is very heterogenous (Figure 1). Within an asset class, there are relatively many weak links while there is still a small number of strong links. Across asset classes, the concentration of strong links increases from the segments short-term debt over investment fund shares and long-term debt to equities. For the mechanism described above, equities are thus most prone to risk contagion.

We can compare the intensity of overlaps within the same country (domestic) as well as across countries (cross-border). The comparison indicates that country-sector pairs are more likely to coordinate in their security selection domestically, which creates stronger overlaps (Figure 2). Thus, the shock propagation may be stronger within a country than across countries.

Risk exposure of market participants is largest on equities segment

We can further analyze how exposed individual market participants, i.e., country-sector pairs, are in each asset class. The number as well as the intensity of overlaps that a country-sector pair faces determines how exposed it is to shock propagation. The susceptibility of all market participants in the financial system are contingent upon each other. Consequently, a suitable measure of a specific country-sector pair’s risk must consider its position in the entire network structure (Box 2).

We measure every country-sector pair’s exposure separately for all holdings in a particular asset class and plot the measure’s distribution (Figure 3). The comparison indicates that the median susceptibility is largest in the market segment of listed shares and smallest in short-term debt. This suggests that market participants are most exposed on the segment of equities.

However, portfolios typically include multiple asset classes, allowing for shocks to spread across market segments as well. This means that market losses in a specific segment are not necessarily due to incidents therein, but may have been caused elsewhere in the market. The sheer complexity requires us to zoom into specific relations using granular data to identify the source of a market-wide phenomenon.

To illustrate this notion, we visualize the risk exposure from a sector perspective. This approach would be relevant for a macro-prudential approach. Because each national central bank collects individual security holdings from relevant entities in their jurisdiction, it would even be possible to take a firm-level perspective. The result shows that the exposure of financial industry sectors is largest, in particular for the investment fund and insurance sector (Figure 4). These are especially relevant among those sectors connected to the Other Financial Intermediaries sector, which is attracting increasing attention given its limited regulatory scrutiny.

Pre-crisis data ignores security-specific dependencies and would yield different results

Before the crisis, this analysis would have only been possible on a coarser dataset. It is questionable whether this dataset...
would have been sufficient to generate consistent results in this risk analysis. To mimic such a dataset, we aggregate all security holdings by the issuer country and sector.

We compare the distribution of risk exposures of country-sector pairs based on the granular dataset (Figure 3) and the coarser dataset (Figure 5) and find that the pre-crisis dataset ignores a critical number of interactions. While the coarser dataset melts linkages in individual securities that are issued by the same country-sector pair, together, we can distinguish between them in the granular dataset. As a consequence, the ordering of asset types in terms of their susceptibility to shock propagations changes. In the hypothetical pre-crisis dataset, the most vulnerable market segment are investment fund shares while it is the equities segment in the granular dataset. The comparison of the median values also shows an overestimation for the coarse dataset.

**Conclusion: more coordination could reduce costs**

Our analysis exemplifies how granular data benefits monitoring and assessing financial stability. The results show that portfolio overlap and the risk exposure it creates for sector-country pairs is most pronounced for the equities segment. Coarse datasets, like the ones that would have been available before the crisis, ignore important interactions that are capable of biasing the results. Policy measures would have borne on the investment fund shares segment instead of the equities segment.

---

**Figure 2**

**Intensity of domestic and cross border portfolio overlaps**

In percent\(^1\), Q1 quarter of 2018

- **Equities**
  - Increasing density
  - Domestic
  - Cross border

- **Investment fund shares**
  - Increasing density

1 Jaccard index

Sources: ECB; authors’ own calculations.

---

**Figure 3**

**Risk exposure for country-sector pairs by asset class based on the granular dataset**

Median value (orange) and deviations, Q1 of 2018

- **Equities**
- **Investment fund shares**
- **Short-term debt**
- **Long-term debt**
- **Total**

Sources: ECB; authors’ own calculations.

© DIW Berlin 2020

The risk exposure of equities is the highest.
Hence, granular data facilitate the effective fulfillment of regulatory mandates and the identification of risk drivers for authorities. However, it seems necessary to explore avenues to alleviate the costs for market participants in order not to create a disadvantage in the competition within the euro area and beyond.

A lack of coordination in the reporting obligations across regulatory regimes drives costs. There are various regulations that require very similar yet different information. If data requirements were to be sufficiently coordinated across frameworks and reported at a sufficiently granular level, then the national authorities could aggregate the data depending on their need. Further, some regulations have dual supervision regimes, which create overheads for reporting agents. EMIR is a prominent example: both the European Securities and Markets Authority (ESMA) as well as national authorities are responsible for it.

More stable rules with fewer adjustments would also significantly reduce costs because fewer resources would be necessary for costly adjustments to already implemented processes. However, this needs to be weighed against the changing data needs in a rapidly evolving financial landscape.

Once the relevant data is processed, reporting agents face yet another costly bottleneck in the transmission of data. Currently, most national authorities follow a template-based approach, which imposes a very specific format on the transmission that is different across jurisdictions. This creates unnecessary overhead, especially for large market participants with cross-border activities.

Some authorities are looking for alternative ways to acquire data. If disadvantages, for example in terms of greater risks to confidentiality, can be controlled, some could eventually serve as a blue print for the entire euro area. In the Austrian system, direct reporting to the central bank has been replaced by reporting to a joint venture, the Austrian Reporting Services GmbH (known as AuRep), which transmits the data to the central bank. It is owned by seven of the largest banking groups in the country and is tasked with data collection as well as transmission. This structure allows for greater harmonization across banks as well as the integration between IT systems of reporting agents and the supervisor. The Dutch Central Bank investigates a more radical approach as it experiments with ways to plug directly into the systems of private institutions to gather the regulatory information itself.

According to the coarser pre-crisis dataset, the risk exposure of investment funds is the highest. This is in contrast to the granular dataset, where equities have the highest risk exposure.