EMIR and SFTR data quality report 2021

1. Table of contents

2. Executive summary

This is the second edition of report on data quality under the European Market Infrastructure Regulation (EMIR) and under the Securities Financing Transactions Regulation (SFTR). The objective of the report is to provide a holistic view of state of play of both reporting regimes as regards the quality of the reported data and the actions that the national competent authorities (NCAs') and the European Securities and Markets Authority (ESMA) are taking to improve the quality of the data. EMIR and SFTR data play a pivotal role in the fulfilment of NCAs' and ESMA's supervisory mandates. As such, the data is used extensively for those purposes[1]. Key highlights of the report are as follows:

ESMA strategic priorities on EMIR and SFTR data quality

As regards ESMA's supervision of TRs, key areas of focus are:

i. Timely and complete reporting of regulatory information to the users of TR data,
ii. Accuracy and confidentiality of data reported by counterparties to and stored by TRs, and
iii. Accuracy of regulatory reports submitted to the users of TR data.

As regards the reporting by counterparties, the key common areas of priority for NCAs and ESMA are:

i. Completeness and accuracy of the reported information, in particular with regards to the reporting of valuation and collateral data,
ii. Timely submission of the reports, and
iii. Consistency of reported information reflected in the reconciliation of data submitted by the two counterparties of the same derivative.

Counterparties are strongly encouraged to use the regulatory data in their own internal risk and compliance management processes. In doing so, counterparties incentives to report accurate data will be further aligned.

ESMA, with the cooperation of the NCAs, is and will continue to monitor progress in those areas. ESMA and the NCAs will take actions with the objective to achieve improvement in areas where insufficient quality of the data is identified.

Recent developments impacting EMIR and SFTR data quality

Regarding EMIR, ESMA carried out three supervisory projects focusing on i) the ingestion and processing of data by TRs, ii) application of EMIR access filtering rules for provision of data to NCAs according to their mandates and iii) the assessment of consistency of two key regulatory reports – the trade activity and trade state reports. In all three cases, ESMA found that TRs broadly follow regulatory and supervisory expectations. In certain instances, ESMA found some shortcomings in the quality of the reports provided to regulators[2] and expects that TRs take appropriate remediation steps. As regards SFTR, TRs and reporting counterparties implemented first SFTR XML schema update since the start of reporting in July 2020. The update aimed at removing technical shortcoming that could decrease quality of the information available to the regulators. ESMA also monitored and coordinated with NCAs/TRs all relevant aspect of the wind-down of UnaVista repository services under SFTR as well as the associated porting of SFTR data to other TRs

---

[1] See for example the Trends Risks and Vulnerabilities report (https://www.esma.europa.eu/sites/default/files/library/esma50-165-1842_trv2-2021.pdf) and the Annual Derivatives Statistical Report (https://www.esma.europa.eu/sites/default/files/library/esma50-165-2001_emir_asr_derivatives_2021.pdf)

[2] Such as under-/over-reporting of data in the case NCA access filtering and report accuracy issues in the case of the trade activity and trade state reports.

EMIR and SFTR Data Quality Frameworks

Data quality under EMIR and SFTR relies on an efficient supervision of the reporting counterparties by the NCAs and of the TRs by ESMA. Having in mind these complementary supervisory responsibilities, NCAs and ESMA have established the following dedicated frameworks to coordinate the joint efforts on ensuring high data quality: the Data Quality Action Plan (DQAP) under EMIR and the Data Quality Engagement Framework (DQEF) under SFTR.

Under EMIR DQAP the NCAs analysed and followed up with selected supervised entities on the results of over 30 tests related to different data quality aspects such as completeness, accuracy, or timeliness. In a thematic review focused on the reporting of valuations and collateral, misreporting of valuations was significantly reduced, as compared to the previous year, by around 50% of the targeted entities. Similarly, in a dedicated exercise on timeliness of reporting, most of the targeted entities improved their reporting practices and eliminated or significantly reduced late reports.

SFTR DQEF was launched in 2021 and focused on the timeliness of reporting, rejections, and pairing. While the follow-up on this first iteration of DQEF is still ongoing. Overall, it can be noted that use of the ISO20022 XML end-to-end reporting has brought important benefits in terms of the quality and accessibility of the data from the very beginning of the reporting. Furthermore, some positive trends in key metrics such as rejections and reconciliation can already be observed. Recent developments impacting EMIR and SFTR data quality

Regarding EMIR, ESMA carried out three supervisory projects focusing on i) the ingestion and processing of data by TRs, ii) application of EMIR access filtering rules for provision of data to NCAs according to their mandates and iii) the assessment of consistency of two key regulatory reports – the trade activity and trade state reports. In all three cases, ESMA found that TRs broadly follow regulatory and supervisory expectations. In certain instances, ESMA found some shortcomings in the quality of the reports provided to regulators and expects that TRs take appropriate remediation steps. As regards SFTR, TRs and reporting counterparties implemented first SFTR XML schema update since the start of reporting in July 2020. The update aimed at removing technical shortcoming that could decrease quality of the information available to the regulators. ESMA also monitored and coordinated with NCAs/TRs all relevant aspect of closely the wind-down of UnaVista repository services under SFTR as well as the associated porting of SFTR data to other TRs.

EMIR reporting trends and selected data quality metrics

Brexit has had an important impact on the EU supervisory data reporting landscape as volumes of reported derivatives fell by approximately 50%. In terms of data reporting volumes, equities and futures contracts continue to be the most prominent asset class and contract type respectively. While less than 10% of reported derivatives tend to be reported late by the counterparties, more than 20% do not receive updated valuation on a daily basis as required by EMIR. Non-reporting dropped sharply due to Brexit and is now less than 5%. The sharp drop has been driven by the end of the reporting obligation of UK counterparties and the more limited dual-side reporting which does not allow to detect potential non-reporting issues. TR rejections continue to be low at around 2%. Furthermore, only 1% of records in TR regulatory reports seem to not comply with the applicable validation rules. Volumes of duplicated reporting[3] are negligible. As regards reconciliation, pairing rate continues to be relatively low at 60% while there is on average 5% difference between the number of open derivatives reported between a pair of counterparties. Lastly, in some instances, TRs disagree on the number of derivatives they reconcile against each other. This may be an indication of further enhancements required for the inter-TR reconciliation process. While much attention has already been put to timely reporting, reporting of

---

[3] Reported derivatives are considered duplicated where two or more records have been reported with the same combination of reporting counterparty ID, ID of the other counterparty and trade ID fields. In contrast, double-sided reporting under EMIR is not duplicative since two records on the same derivative should be reported always from the perspective of the respective reporting counterparty.

valuations and reconciliation, clearly much more improvements are needed and those area s will continue to be point of focus of ESMA and NCAs going forward.

SFTR reporting trends and selected data quality metrics

Similar to EMIR, data reporting volumes dropped approximately by 50% following Brexit. In terms of number of open transactions, securities lending and borrowing is the largest SFT type reported with around 70% share at the end of 2021. Credit institutions report most open SFTs (around 50%) while credit institutions share has been increasing (to around 30% at the end of 2021). After merely 1.5 years of reporting, SFTR exhibits comparable results to EMIR across all data quality metrics. Around 10% of SFTs are reported late (after T+1). On the contrary, rejections have been low (around 2%) and duplicated reporting does not pose major issues. As regards reconciliation, pairing rate has been only around 60%. Reconciliation rate of loan and collateral data has been low but increasing to around 40% and 30% respectively. Similar to EMIR, TRs do not agree on the number of records they reconcile against each other, which may be an indication of issues in the inter-TR reconciliation process. Timeliness of reporting, adherence to format and content rules (via rejections) and reconciliation (pairing) has been the point of focus of ESMA and NCAs during 2021. While progress has been made, some areas (particularly reconciliation) need to remain areas of focus also in the future.

3. ESMA's Strategic Priorities on EMIR and SFTR Data Quality

Summary: As regards ESMA's supervision of TRs, key areas of focus are: i) timely and complete reporting of regulatory information to the users of TR data, ii) accuracy and confidentiality of data reported by counterparties to and stored by TRs, and iii) accuracy of regulatory reports submitted to the users of TR data.

As regards the reporting by counterparties, the key common areas of priority for NCAs and ESMA are: i) Completeness and accuracy of the reported information, in particular with regards to the reporting of valuation and collateral data, ii) timely submission of the reports, and iii) consistency of reported information reflected in the reconciliation of data submitted by the two counterparties of the same derivative. Counterparties are strongly encouraged to use the regulatory data in their own internal risk and compliance management processes. In doing so, counterparties incentives to report accurate data will be further aligned.

ESMA, with the cooperation of the NCAs, is and will continue to monitor progress in those areas. ESMA and the NCAs will take actions with the objective to achieve improvement in areas where insufficient quality of the data is identified.

TR supervisory objectives: ESMA is the direct supervisor of TRs under EMIR and SFTR. ESMA sets its supervisory priorities on an annual basis and publishes them in ESMA's annual work programme.

ESMA is a data-driven and risk-based supervisor. Thus, it sets its priorities based on risks it observes which may negatively impact quality of the reported data. The most prominent risks are then included in the list of its annual priorities in the form of a specific project or a supervisory review. Besides one-off projects, ESMA also performs a variety of monitoring activities[4] on an ongoing basis.

Even though ESMA's priorities may evolve from one year to another, there are common themes that remain present over time.

Those themes are:

1. Timely and complete reporting of regulatory reports to the users of TR data;

2. accuracy and confidentiality of data reported by counterparties to and stored by TRs[5]; and

3. accuracy of regulatory reports submitted to the users of TR data[6].

ESMA expects that the TRs pay utmost attention to the above-mentioned aspects and that they have processes, systems and controls in place to monitor and timely identify any issues.

Counterparty reporting supervisory objectives: NCAs are responsible for the supervision of the reporting by the counterparties while ESMA coordinates some key common initiatives in the context of its supervisory convergence mandate[7].

While common priorities are also set annually, there are key areas of permanent focus by NCAs and ESMA:

1. Completeness and accuracy of the reported information, in particular with

---

[4] For example: monitoring timeliness and completeness of TR daily regulatory report submissions.
[5] See Subsection 5.1. on EMIR TR data ingestion review carried out through 2020 and 2021.

[6] See Subsection 5.2. on EMIR regulatory access filtering review and Subsection 5.3. EMIR TAR-TSR consistency both carried out in 2021.
[7] See Section 0.

regards to the reporting of valuation and collateral data;

2. timeliness of the reports; and

3. consistency of reporting reflected in the reconciliation of data reported by the two counterparties of the same derivative.

Reporting counterparties are expected to have processes, systems and controls in place to ensure completeness, accuracy and timeliness of the reported information. Furthermore, they are expected to actively engage in detecting and resolving any identified report rejections, reconciliation breaks and other data quality issues in the already reported data.

Counterparties are strongly encouraged to use the regulatory data in their own internal risk and compliance management processes. In doing so, counterparties will have the appropriate incentives to report accurate data and will be in apposition to better exploit the benefits of consistent data reporting.

What we aim to achieve: ESMA, with the cooperation of the NCAs, is and will continue to monitor progress in those areas. ESMA and the NCAs will take actions with the objective to achieve improvement in areas where insufficient quality of the data is identified.

4. EMIR and SFTR Data Quality Frameworks

Summary: Data quality under EMIR and SFTR relies on an efficient supervision of the reporting counterparties by the NCAs and of the TRs by ESMA. Having in mind these complementary supervisory responsibilities, NCAs and ESMA have established the following dedicated frameworks to coordinate the joint efforts on ensuring high data quality: the Data Quality Action Plan (DQAP) under EMIR and the Data Quality Engagement Framework (DQEF) under SFTR.

Under EMIR DQAP the NCAs analysed and followed up with selected supervised entities on the results of over 30 tests related to different data quality aspects such as completeness, accuracy or timeliness. In a thematic review focused on the reporting of valuations and collateral, misreporting of valuations was significantly reduced, as compared to the previous year, by around 50% of the targeted entities. Similarly, in a dedicated exercise on timeliness of reporting, most of the targeted entities improved their reporting practices and eliminated or significantly reduced late reports.

SFTR DQEF was launched in 2021 and focused on the timeliness of reporting, rejections and pairing. While the follow-up on this first iteration of DQEF is still ongoing, overall, it can be noted that use of the ISO20022 XML end-to-end reporting has brought important benefits in terms of the quality and accessibility of the data from the very beginning of the reporting. Furthermore, some positive trends in key metrics such as rejections and reconciliation can already be observed.

4.1 EMIR DQAP

EMIR Data Quality Action Plan (DQAP) the DQAP is a major project that NCAs and ESMA jointly launched in September 2014. It aims at improving the quality and usability of data that is reported by counterparties and made available by the TRs.

The DQAP encompasses activities related to the policy work, NCAs' supervision of the reporting counterparties and ESMA's supervision of the TRs, to address the potential issues in all areas that are key for the quality of the final data, notably: (i) the comprehensive, detailed, and precise specification of the reporting requirements; (ii) the complete and correct reporting by the counterparties to the TRs; and (iii) the provision of complete and accurate data by the TRs to the authorities.

Data Quality Review (DQR): The DQR is currently the main common exercise performed in the context of the DQAP with regards to the supervision of the reporting by the counterparties. Under the DQR, each NCA, applying a commonly agreed methodology, performs a quantitative assessment of the quality of data reported by selected counterparties in their Member State and follow up with the relevant entities on the identified issues.

NCAs provide subsequently to ESMA information on the results of the DQR and on the follow-up supervisory actions. Based on this feedback ESMA prepares a summary report that is subsequently shared with the NCAs. High-level outcomes are also provided to the Board of Supervisors as part of the annual update on the execution of the DQAP.

DQR 2021: Similarly to the previous year, the 2021 DQR contained a series of over 30 data quality tests grouped into three broad areas: (i) analysis of pairing and matching of the reports, (ii) analysis of completeness, accuracy, timeliness, and rejections of reports made by significant reporting entities, and (iii) thematic review: analysis of reporting of valuation and collateral data.

In 2021, 19 NCAs participated in the DQR. Given that the DQR analysis is based on limited samples of counterparties[8] and that each counterparty may face different reporting issues, the results of the data quality checks are not representative for the full EMIR dataset and vary across the different tests and between the participating Member States. However, based on the samples considered in the DQR, overall, a slight improvement has been noted in some of the analysed areas as compared with the 2020 DQR, notably the pairing and matching rates as well as the consistency of number of reported derivatives with the entities' internal records.

Thematic review: Furthermore, it is worth mentioning the outcomes of the thematic review, focused on the reporting of valuations and margins, which is a key information for the monitoring of systemic risks. In addition to analysing the aggregate results for the samples from each Member State, ESMA has looked also into the evolution of one basic measure, notably the number of trades with empty/zero valuations, at entity level for all the entities that were selected for the thematic review based on this measure in the 2020 DQR. The comparison of the statistics computed for the purpose of 2020 DQR and 2021 DQR, revealed that approximately 50% of the entities have significantly reduced the misreporting of valuations[9] and further 20% recorded some reduction in the number of trades impacted by this data quality issue.

These outcomes confirm the conclusions from the previous year that adequate supervisory pressure and close monitoring of the implementation of remedial actions are needed to ensure a material long-term impact on the improvement of the quality of data by all relevant entities. They also show that targeted actions directed at the main misreporting entities having the highest impact on a given data quality aspect constitute an efficient approach to resolve the data quality problems.

Framework for provision of information on data quality issues to NCAs and the follow-up with supervised entities: In line with these findings NCAs and ESMA have also established a common framework, applicable whenever a significant problem impacting the data quality of the EMIR data at EU level is identified. The framework sets up a procedure for an efficient resolution of such most significant data quality issues, specifying, among others:

- the responsibilities of NCAs and ESMA,
- the timelines for the exchange of information between NCAs and ESMA,
- the format and minimum content of the statistics to be shared by ESMA,
- the criteria to decide which reporting should be addressed,
- the feedback information to be provided by the NCAs to ESMA,
- the steps to ensure that the data quality problem has been mitigated (incl. reassessment of the data).

The important feature of the framework is that the follow-up is focused on a limited subset of entities with the highest share of incorrect reports in the total number of impacted reports at EU level, thus ensuring the most efficient use of the NCAs resources. The framework was first launched in practice in 2021 for the follow-up on the timeliness of reporting and resulted in a material improvement in the reporting practices of the targeted entities. (see Subsection 4.3 for more details).

What have we achieved: EMIR DQAP is a comprehensive data quality framework based on the harmonised data assessment methodology. Implementation of EMIR DQAP by the NCAs allowed them to detect and follow up on several data quality issues with their supervised entities.

Furthermore, ESMA shared with relevant NCAs information on entities with highest numbers of late reports at EU level. The targeted follow-up with those entities resulted in an improvement of their reporting practices.

4.2   SFTR DQEF

SFTR Data Quality Engagement Framework (DQEF) The reporting under SFTR started in July 2020, following a three-month delay due to the Covid-19 pandemic. SFTR, as a new transaction-level reporting regime, has required dedicated efforts for its supervision and data quality over the year 2021.

Compared to EMIR, SFTR is ISO20022 XML end-to-end reporting regime and this has brought important benefits in terms of the quality and

---

[8] Each Member State selects 5 entities per each of the three areas of analyses

[9] Reduction of the number of trades with empty/missing valuation by halve or more was considered as a significant reduction

accessibility of the data from the very beginning of the reporting.

To ensure a consistent and efficient approach to data quality assurance and supervision, ESMA is systematically developing and implementing various convergence tools across all relevant reporting regimens and data systems including the SFTR DQEF.

The SFTR DQEF was agreed in 2021 and leverages on the EMIR DQAP setting out the SFTs data quality work to be undertaken jointly by NCAs and ESMA. It defines the necessary coordinated procedures to verify, communicate and prioritise the data quality findings detected in the SFT data submitted by the reporting counterparties and to subsequently apply the relevant corrective measures leveraging on the agreed best practices to foster the SFTs data quality and to enforce the supervisory actions.

DQEF 2021: 2021 has been the first year of data quality assessments performed by NCAs and ESMA on SFTR data. Some national authorities indicated that due to the novelty of the SFTR reporting regime and to the complexity of the activities to be implemented, they were not able to contribute to the first data quality exercises. To allow NCAs to focus on building their systems as well as to facilitate the deployment of adequate resources for their supervisory activities and engagement with entities, ESMA has centrally performed, on behalf of the NCAs, a targeted set of checks and reported the detected data quality issues to the NCAs. In turn, during the first year of activities, the NCAs focused mostly on the remedial actions and follow-ups with the counterparties and entities responsible for reporting under their direct supervision.

The implementation and performance of the data quality checks is based on an incremental approach. The 2021 data quality exercise was performed in two separate rounds of tests on weekly datasets (June and November) and had a very targeted nature focusing on the aspects, which under other reporting regimes have proven to be the cornerstones of data quality, namely: (i) timeliness of reporting, (ii) rejected reports due to incorrectness or inaccuracy of SFT records according to the validation rules (iii) unsuccessful paired status of the reported records.

Considering the importance of the SFTR regime on the one hand and its relative complexity on the other, it is crucial to perform regular data quality

assessments and expand in the near future the existing data quality activities with a view to ensure the usability of the data for monitoring of financial stability risks.

What have we achieved: The activities related to the data issues of November 2021 cycle are not yet completely terminated as the participating NCAs are finalising the outputs of their interactions with the supervised entities and communicate them to ESMA. Therefore, it is premature to draw conclusions on the overall outcome of their remedial actions in this report. It is important to highlight that the trends of the data quality findings of the two cycles are also reflected in the relevant outcomes of the SFTR data quality analyses that are covered in Section 7. In some areas such as rejections and reconciliation rates, these indicators already begin to show positive trend.

4.3 Cooperation with data users

EMIR Timeliness Analysis: ESMA performed an analysis of the timeliness of reports under EMIR based on a time series constructed for several dates across 5 months and using both Trade State Reports (TSR) and Trade Activity Reports (TAR). Thanks to the interactions with the NCAs, the analysis was further enhanced to eliminate certain false positives or legacy trades.

In line with the criteria specified in the framework for provision of data by ESMA to NCAs (incorrect reports by a given entity exceeding the 1% of all incorrect reports in the EU), ESMA has identified for the follow-up 15 counterparties from 7 jurisdictions. Furthermore, as envisaged in the framework, all relevant NCAs have provided ESMA with the feedback on the follow-up with the entities explaining the reasons for late reporting and indicating whether the issue has been resolved.

In order to assess the actual impact of the exercise, ESMA staff have rerun the timeliness analysis for other dates in July, i.e., after the NCAs finalised the follow-up with the entities.

No improvement was observed in the TSR, which was an expected outcome given the particularity of late reporting, i.e., once a given derivative is reported late, it is not possible to 'correct' the time of its initial submission. Therefore, such derivative until further life-cycle event is reported, will continue to appear as reported late in the TSR. However, the approached entities have enhanced their reporting practices with regards to

timeliness, as an improvement has been observed in the TAR submitted after the follow-up. In particular, among the 10 entities which were identified for the follow-up based on the number of late reports in the TAR data, 4 entities had no more late reports in the TAR of July and a reduction in the number of late reports has been observed also for the remaining 6 entities.

This example showed that targeted ad-hoc exercises directed at the most relevant entities are very efficient to reduce the most significant data quality issues. ESMA plans to continue engaging with NCAs in this way on a broader number of issues.

Abnormal values: In August 2020 ESMA implemented a new data quality process to identify abnormal values on the numerical fields reported by the counterparties under EMIR regime. The data were then shared with the NCAs to support them in their supervisory activities. Notably, once the data are shared, NCAs can verify if the detected outliers are due to data quality issues in the reporting of the counterparties under their supervision and follow up accordingly, if needed.

The data quality analysis identifies irregular numerical values (e.g., too high or too low) reported for EMIR fields such as value of the contract, margins, notional, fixed-rate legs of the contract, price rate and quantity[10]. The focus on these values is driven by the impossibility to automatically detect and reject outliers through the validation rules and the need for specific soft checks.

The abnormal value analyses were shared with the NCA monthly from August 2020 to August 2021. This section summarises the main findings.

Firstly, the abnormal values that ESMA identified represent a small proportion of the outstanding open derivatives. However, given the high impact on data quality of the derivatives with abnormal values, it is important to perform such data quality checks on an ongoing basis to swiftly detect any abnormal values and to inform the NCAs responsible for the supervision of the counterparties reporting such values. For instance, a recurrent example concerns the reported value of the contract above 100 billion EUR, which may impact significantly results of analyses based on EMIR data.

From August 2020 to August 2021, ESMA notified NCAs on 84 occasions, reporting 3493 potential outliers to fifteen jurisdictions. ESMA then received 19 responses from NCAs clarifying the reasons for the outlier or committing to contact the counterparty responsible for the misreporting. **Chart 1** provides an overview of the notified potential abnormal values detected from August 2020 to August 2021 per jurisdiction. **Chart 2** shows the EMIR fields with most outliers identified in the same period.

Chart 1
Cooperation with data users
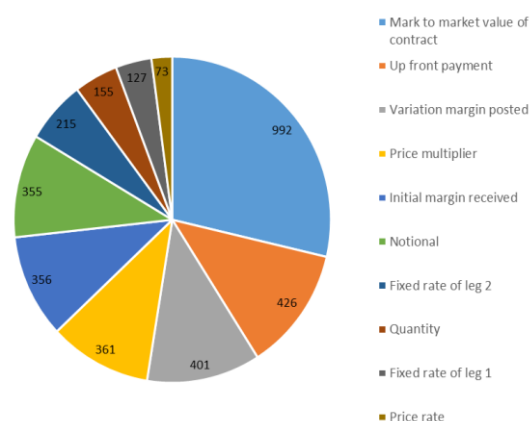**Number of abnormal values, per country**
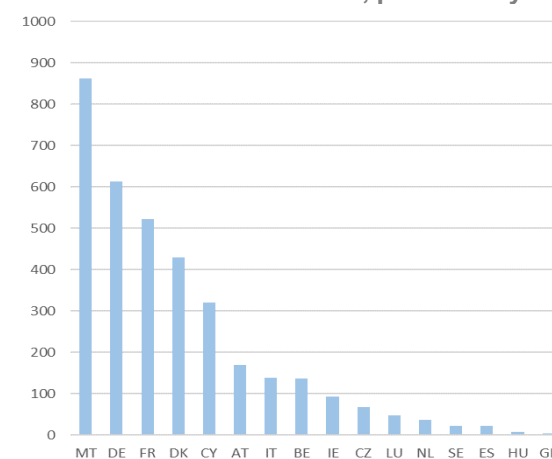
Source: ESMA data & calculations

Chart 2
Cooperation with data users
Number of abnormal values by EMIR field

Source: ESMA data & calculations

EMIR and SFTR log of data quality issues: Since 2016, ESMA has established a structured framework for the users of data to be able to

---

[10] ESMA uses various internal methodologies to identify and treat abnormal values. For example, for the purposes of economic analysis, ESMA is using a statistical approach (identifying outliers in notional values) elaborated in detail here:

https://www.esma.europa.eu/system/files_force/library/esma50-165-639_esma-rae_asr-derivatives_2018.pdf?download=1

report any encountered issues and, in turn, to receive feedback. ESMA receives issues pertaining to TRs (for example, incorrect generation of regulatory reports) as well as counterparties (for example, implausible notional/collateral values). When a counterparty reporting issue is identified, the issue is channelled to the responsible NCA. TR issues are addressed by ESMA.

Chart 3 shows the breakdown of all reported issues since the inception of the log by their status, i.e., TR issues closed and open, and counterparty reporting issues. In 2021, ESMA processed around 30 issues overall. Most TR issues are also being closed during the same year. However, ESMA prioritizes its follow-ups based on urgency and impact, thus not all the issues may be addressed during the same year.

What have we achieved: Data quality issues can be of structural nature, but significant issues can also appear from one day to another. Thus, it is important for ESMA and NCAs to maintain agility to be able to react to issues that were not planned to be addressed through structural projects such as the EMIR DQAP and SFTR DQEF.

Through ad-hoc sharing of DQ issues (such timeliness reporting, abnormal values, and any other data quality issues) and their immediate prioritisation, identified problems are being address in an agile fashion.

Chart 3
Cooperation with data users
Data quality issues reported by NCAs\CBs
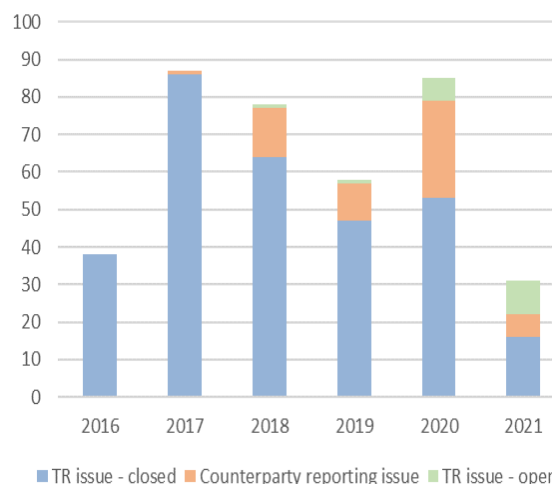
Note: Number of issues reported data quality issues since inception and broken down by status.
Source: ESMA data & calculations

5. Recent developments impacting EMIR and SFTR data quality

Summary: Regarding EMIR, ESMA carried out three supervisory projects focusing on i) the ingestion and processing of data by TRs, ii) application of EMIR access filtering rules for provision of data to NCAs according to their mandates and iii) the assessment of consistency of two key regulatory reports – the trade activity and trade state reports. In all three cases, ESMA found that TRs broadly follow regulatory and supervisory expectations. In certain instances, ESMA found some shortcomings in the quality of the reports provided to regulators[11] and expects that TRs take appropriate remediation steps. As regards SFTR, TRs and reporting counterparties implemented first SFTR XML schema update since the start of reporting in July 2020. The update aimed at removing technical shortcoming that could decrease quality of the information available to the regulators. ESMA also monitored and coordinated with NCAs/TRs all relevant aspect of the wind-down of UnaVista repository services under SFTR as well as the associated porting of SFTR data to other TRs

5.1   EMIR - TR data ingestion review

In 2020, ESMA identified a need to verify whether and to what extent EMIR data quality issues arise during the data ingestion processes[12] of TRs. A thematic review was initiated to assess the data ingestion processes of three out of four registered TRs using a data-driven supervisory approach.

The sample consisted of trade activity reports for two consecutive dates (25-26 November 2020) from eleven major financial counterparties. ESMA engaged with the French (AMF), Dutch (AFM) and German (BaFin) authorities to obtain proprietary trade activity data from the selected counterparties. Without the collaboration of these NCAs and the effort made by the counterparties to extract, prepare, and submit the data to ESMA, it would have not been possible to fully assess the data ingestion processes of TRs.

A methodology and algorithms were developed to assess and identify concrete data integrity issues stemming from TRs' data ingestion processes. In broad terms, it consisted of comparing data reported by counterparties with data stored by

TRs in their internal databases before any subsequent data transformation, aggregation, filtering, or report generation processes were performed. The end-to-end process was also verified by comparing data reported by counterparties with data received through TRACE.

Over 20 million records were processed and analysed both from a data completeness (paring) and accuracy (matching) perspective. While in some cases the ingestion process remains inherently complex and issues were detected, our analysis showed a robust EMIR data ingestion process and a good level of data integrity for all the three TRs included in our sample. This implies that the information stored in the TRs' internal database matches the information reported by the counterparties.

Chart 4 summarises the results from the data completeness analysis. Perfect or nearly perfect pairing rates (>99%) were obtained for eight out of eleven TR-counterparty sets. This implies that

---

[11] Such as under-/over-reporting of data in the case NCA access filtering and report accuracy issues in the case of the trade activity and trade state reports.

[12] Data ingestion refers to the part of TR data processing after it the data is received from the reporting participants and before it is loaded to TR databases. TRs typically queue, validate and perform other pre-processing tasks on the incoming data before it is loaded TR databases.

all or almost all records in the counterparty dataset were successfully identified in the TR dataset for these cases.

EMIR – TR Data Ingestion Review Comparative analysis of CP and TR data
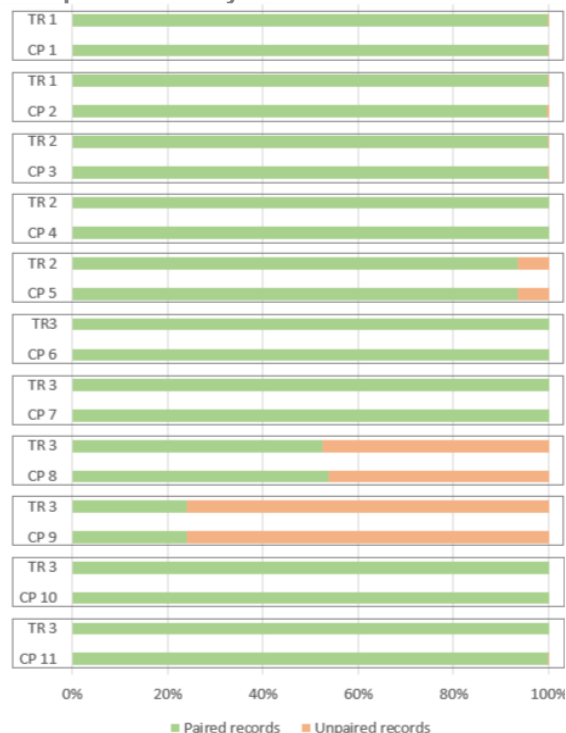
Note: Trade repositories and counterparties in the sample are anonymized.
Source: Trade Repositories & ESMA calculations

On the contrary, lower completeness rates were observed for three out of eleven TR-counterparty sets. The main reason for these discrepancies is not directly attributed to deficiencies with the TR's data ingestion process but rather to the submission of out-of-scope data by these counterparties for this review (i.e., records with reporting timestamp not being part of the subset defined in the request for information).

What have we achieved: In terms of data accuracy, ESMA discovered a range of discrepancies between the information submitted by counterparties and how it had been stored in the TRs' internal databases. Most were of non-critical nature and could be explained by the way the TRs' have implemented their internal IT systems. For example, rounding errors of decimal values, date/time formats and other misalignments which did not have a critical

impact on data quality when TRs generate outbound reports for regulatory authorities.

However, a few critical issues caused by inappropriate modification of counterparty data were detected. ESMA is liaising with the affected TRs to rectify these issues which can have an adverse impact on data quality.

The outcome of this review will mainly be used as input to ESMA's data quality risk assessment, by eliminating risks that could arise from the data ingestion process and focusing the supervisory efforts on other EMIR data reporting process.

It is also worth mentioning that this analysis was carried out under the current reporting framework. The entry into force of EMIR Refit could bring significant changes that could adversely impact the TR's data ingestion processes. Going forward, ESMA will continue to monitor incidents and complaints that are linked to TR's data ingestion processes to ensure that those processes are adequate and resilient to regulatory changes.

5.2 EMIR regulatory access filtering review

When providing data to the authorities, TRs need to apply filtering rules to make the data available to the authorities based on their respective mandates. This is important to avoid that an authority receives data which it is not entitled to and to ensure that each authority receives all the data that is necessary to fulfil its mandates.

Mandates of authorities are set out in the Article 81(3) of EMIR and further developed in technical standards. In order to verify that TRs are providing data according to the regulatory requirements, ESMA has cooperated with four NCAs: CBoI (IE), CNB (CZ), CSSF (LU) and MFSA (MT). These authorities shared with ESMA the regulatory reports submitted to them by the TRs. In parallel, ESMA applied the expected filtering rules applicable to each authority to its own reports received from the TRs[13]. Then ESMA compared its filtered report with those that were actually received by the NCAs.

Through this assessment, ESMA confirmed that TRs seem to broadly follow the regulatory

---

[13] ESMA receives all derivatives data reported under EMIR.

requirements with regards to provision of data to the authorities.

ESMA also identified some shortcomings at TRs which led to either underreporting or overreporting of EMIR data to NCAs. Such issues may affect the data completeness and thus the ability of the NCAs to effectively supervise all entities under their mandates.

The project identified the following key findings:

- On average, 4.5% of expected derivative reports are not provided by TRs (underreporting). This may prevent NCAs from fulfilling their supervisory mandate due to the missing information[14].
- On average, 1.7% more derivative reports than expected are provided by TRs (overreporting). This poses a confidentiality issue as the NCAs receiving these additional derivative reports are not entitled to have access to them.

Chart 5

EMIR ESMA regulatory access filtering review
Derivatives overreported to NCAs

TR/NCA	NCA1	NCA2	NCA3	NCA4	TR total
TR1	240,952	285	80	-	241,317
TR2	3,303	100,668	1,001	296,466	401,438
TR3	12,766	308,597	35,938	47	357,348
TR4	27,577	35,640	23,044	N/A	86,261
NCA total	284,598	445,190	60,063	296,513	1,086,364

TR/NCA	NCA1	NCA2	NCA3	NCA4	TR weighted average
TR1	7.1%	0.0%	0.0%	0.0%	2.5%
TR2	35.3%	96.0%	30.3%	10.2%	13.3%
TR3	1.1%	8.0%	4.4%	0.0%	5.4%
TR4	1.7%	2.2%	1.3%	N/A	1.7%
NCA weighted average	4.6%	6.2%	1.2%	4.9%	4.5%

Note: The tables show number of open derivatives for one reference date that were present in the NCA report but shouldn't. The first table shows the total number of overreported records. The second table shows the number of overreported records as a percentage of the total number of records that should have been included in each NCA report (this number is not shown in the table). The total number of records that should have been included in each NCA report also serves as a weight to calculate weighted averages by NCAs and TRs.
TR2 confirmed to ESMA that the problem is related to a one-off issue with the generation of the regulatory report in question. The issue has been now remediated.
Source: Trade Repositories & ESMA calculations

Chart 6

EMIR ESMA regulatory access filtering review
Derivatives underreported to NCAs

TR/NCA	NCA1	NCA2	NCA3	NCA4	TR total
TR1	9,936	9,545	9,996	10,252	39,725
TR2	4	2	1	562	569
TR3	34,003	34,785	1,145	41,340	111,277
TR4	22,199	189,195	25,831		237,226
NCA total	66,142	233,527	36,973	52,154	388,797

TR/NCA	NCA1	NCA2	NCA3	NCA4	TR weighted average
TR1	0.3%	0.6%	0.4%	0.4%	0.4%
TR2	0.1%	0.0%	0.0%	0.0%	0.0%
TR3	2.9%	1.0%	0.4%	4.7%	2.0%
TR4	1.3%	10.7%	1.5%	N/A	4.6%
NCA weighted average	1.1%	3.4%	0.9%	0.9%	1.7%

Note: The tables show number of open derivatives for one reference date that weren't present in the NCA report, but should have been. The first table shows the total number of underreported records. The second table shows the number of underreported records as a percentage of total number of records that should have been included in each NCA report (this number is not shown in the table). The total number of records that should have been included in each NCA report also serves as a weight to calculate weighted averages by NCAs and TRs.
Source: Trade Repositories & ESMA calculations

What have we achieved: With this project, ESMA has verified that NCAs are broadly receiving data in accordance with their mandates as specified in EMIR.

As outlined above, ESMA identified several shortcomings in the provision of data to authorities. ESMA will follow up on those issues with TRs and ensure that they are appropriately remediated.

5.3   EMIR TAR-TSR review

The incorrect incorporation of the information contained in the Trade Activity Reports (TAR) into the Trade State Reports (TSR) is one of the key TR data quality issues identified by the NCAs during the execution of the EMIR DQR. ESMA has been conducting in 2021 a dedicated project for the comparison of the TAR and TSR data, in particular, for the assessment of the completeness and accuracy of the TSR as an aggregation of individual TARs.

---

[14] As shown in Chart 5, TR2, exhibiting the worst results, confirmed that the problem is related to a one-off issue with the generation of the regulatory report in question. The issue has been now remediated.

The project methodology is based on the comparison of two consecutive EMIR TSRs. The information collected in the first TSR is dynamically updated with the successive TARs received in the period between them. This results in a calculated TSR which is compared with the second TSR submitted by the TR. By comparing the two files, it is possible to detect quality problems related to the incorrect incorporation of the information into the TSR.

The analysis was carried out during 4 consecutive weeks in November 2021, the information presented in this report represents the average of the individual results obtained in each of these weeks. Considering all TRs, more than 300 million TAR and 80 million of TSR records were processed for each week of analysis.

The metrics obtained allow the identification of two types of data quality issues: first, completeness issues related to the incorrect presence or absence of records in the TSR; second, accuracy issues related to the incorrect update of a selected group of TSR fields. ESMA is currently checking and following up with the TRs on the results presented below.

Completeness of the TSR: **Chart 7** shows the percentage of missing records in the TSR over the total volume of records (i.e., derivatives present in the internally calculated TSR and not found in the TSR received from the TRs). These are records that were erroneously deleted or not included in the TSR by the TRs. Although in general terms the results can be considered positive, there are certain divergences in the figures obtained for each TR as one of the TRs presents results close to 2%. These, together with the high volumes of operations that constitute TSR makes a further analysis of the root causes of this issue necessary. It is equally relevant to note that some TRs show very positive results, with error levels close to 0%.

Analogously, as shown in **Chart 8**, data on redundant records have also been obtained (i.e. derivatives present in the TSR received from the TRs but not in the internally calculated TSR). These are records that were erroneously included or not removed from the TSR by the TRs. The results again diverge when comparing the different TRs, although it is important to highlight that one of the entities presents very positive results with virtually no errors in this

metric. As in the previous test, the results require further analysis to understand the possible explanations for these findings.

Chart 7
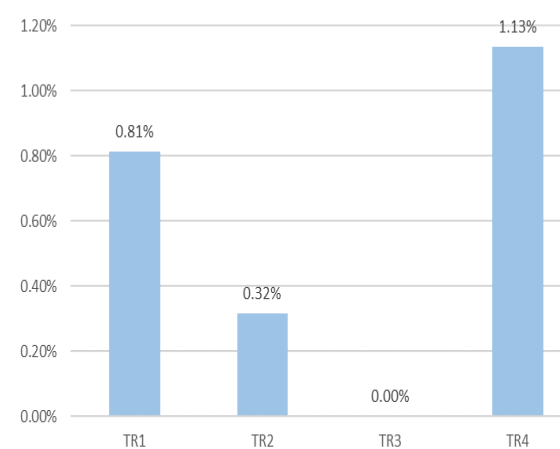Completeness of the TSR
Missing records in the TSR
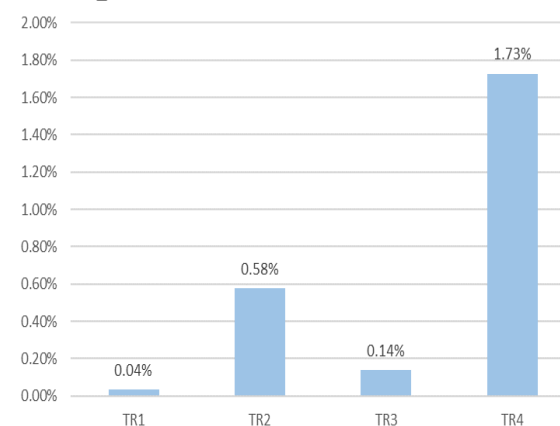
Note: Trade Repositories are anonymized and presented as TR1, TR2, TR3 & TR4. The repositories included in this chart are DDRIE, KDPW, UNAVISTA B.V. & REGIS, the anonymized aliases do not correspond to the order of the TRs named.
Source: Trade Repositories & ESMA calculations

Chart 8
Completeness of the TSR
Redundant records in the TSR

Note: Trade Repositories are anonymized and presented as TR1, TR2, TR3 & TR4. The repositories included in this chart are DDRIE, KDPW, UNAVISTA B.V. & REGIS, the anonymized aliases do not correspond to the order of the TRs named.
Source: Trade Repositories & ESMA calculations

Accuracy of the TSR: The second part of the project consisted in comparing a number of key fields (36 different fields were selected) of the TSR to verify whether the information contained in them has been correctly updated or not. For this purpose, the records of the internally calculated TSR and the TSR provided by the TRs were compared. The results are shown in **Chart 9**.

In general terms, it can be observed that all TRs are close to 100%, which implies that a vast majority of the fields in the report are correctly updated. However, in this first cycle of analysis a selection of 36 fields of the TSR has been used, so the results may vary if the number of fields analysed is extended in future iterations.

Chart 9
Completeness of the TSR
Equality of the fields analysed
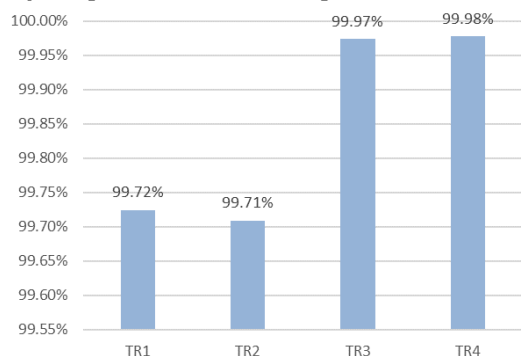
Note: Trade Repositories are anonymized and presented as TR1, TR2, TR3 & TR4. The repositories included in this chart are DDRIE, KDPW, UNAVISTA B.V. & REGIS, the anonymized aliases do not correspond to the order of the TRs named.
Source: Trade Repositories & ESMA calculations

Consideration will be given to broadening the scope of fields included in the accuracy test to verify whether the problems identified can be extrapolated to the rest of the fields in the TSR.

Overall, for the two types of analyses included in the test (completeness and accuracy), the observed values differ between the different TRs and the metrics obtained do not show significant issues in terms of the correct generation of the two key regulatory reports.

What have we achieved: It has been found that, in general terms, there is a correct configuration of the TSR as a result of the messages reported in TAR. This further enhances the usefulness of both reports for subsequent analysis by the NCAs. On the other hand, the project has allowed the implementation of a verification framework for this process that can be replicated and increased in the future. Finally, the project has allowed the detection of certain data quality issues that, once solved, will increase the overall quality of the EMIR reporting framework.

5.4 SFTR – Implementation of new XML schemas

Leveraging on the experience with the implementation of EMIR, SFTR reporting regime relied since the beginning on the end-to-end reporting in a standardised ISO 20022 XML schema. Such design of the reporting framework allowed to mitigate many data quality issues and improve the usability of the data from the start.

What have we achieved: Since the beginning of reporting under SFTR in July 2020, ESMA has identified or has been made aware of some limitations and inconsistencies in the XML schemas used in SFTR. ESMA collected and thoroughly analysed all the identified issues and, basing on this assessment, prepared an updated version of the schemas. The amendments to the schemas are aimed to ensure that there are no technical limitations to the accuracy of the reports submitted by the counterparties to the TRs or the reports provided by the TRs to the authorities.

The go-live of the updated schemas took place on 31 January 2022.

5.5 SFTR – UnaVista wind-down

In August 2021, UnaVista initiated the wind-down process of its SFTR TR as a result of a decision to not continue to provide these services.

Although the Guidelines on data transfer between trade repositories under SFTR were still under consultation at the time[15], the four SFTR TRs began the implementation of the porting infrastructure for SFTs. Leveraging on the already existing portability framework and infrastructure under EMIR and based on the guidance provided by ESMA, TRs were able to quickly adapt the existing porting infrastructure and implement the necessary functionalities to enable porting of SFT data.

UnaVista started the porting out process of outstanding SFTs in November 2021. By late January 2022, all outstanding SFTs were ported out to the new TRs. The data transfer process of the remaining SFT data is to be finalised by March 2022.

---

[15] See esma74-362-1941_consultation_paper_guidelines_on_portability_emir_sftr.pdf (europa.eu)

What have we achieved: Through continuous monitoring, consequent follow-ups with the involved TRs, and quick resolution of the encountered issues, ESMA ensured that the wind-down activities did not lead to any interruptions in the continuity of the provision of regulatory reports to all data users.

6.EMIR reporting trends and selected data quality metrics

Summary: Brexit has had an important impact on the EU supervisory data reporting landscape as volumes of reported derivatives fell by approximately 50%. In terms of data reporting volumes, equities and futures contracts continue to be the most prominent asset class and contract type respectively. While less than 10% of reported derivatives tend to be reported late by the counterparties, more than 20% do not receive updated valuation on a daily basis as required by EMIR. Non-reporting dropped sharply due to Brexit and is now less than 5%. The sharp drop has been driven by the end of the reporting obligation of UK counterparties and the more limited dual-side reporting which does not allow to detect potential non-reporting issues. TR rejections continue to be low at around 2%. Furthermore, only 1% of records in TR regulatory reports seem to not comply with the applicable validation rules. Volumes of duplicated reporting[16] are negligible. As regards reconciliation, pairing rate continues to be relatively low at 60% while there is on average 5% difference between the number of open derivatives reported between a pair of counterparties. Lastly, in some instances, TRs disagree on the number of derivatives they reconcile against each other. This may be an indication of further enhancements required for the inter-TR reconciliation process. While much attention has already been put to timely reporting , reporting of valuations and reconciliation, clearly much more improvements are needed and those area s will continue to be point of focus of ESMA and NCAs going forward.

6.1 Data reporting – Key trends

EMIR reporting trends can be viewed through various dimensions of the data, such as the number of life-cycle events, contract types, or by asset class.

Submissions are the life-cycle event reports received by the TRs representing the conclusion, modification, and termination of a derivative (as specified in the EMIR action type field) throughout its life. When market volatility is high, it may affect reporting counterparties' trading behaviour, leading to an increase in traded volumes, which in turn will result in an increase in the number of reported submissions.

Key trends: Most notable in **Chart 10** is the decrease in reported volumes post-Brexit, as UK counterparties ceased to report under EMIR. There has been a significant uptick in the volumes reported by REGIS under EMIR since

the dissolution of CME and ICE Trade Repositories.

Chart 10
EMIR reporting key trends
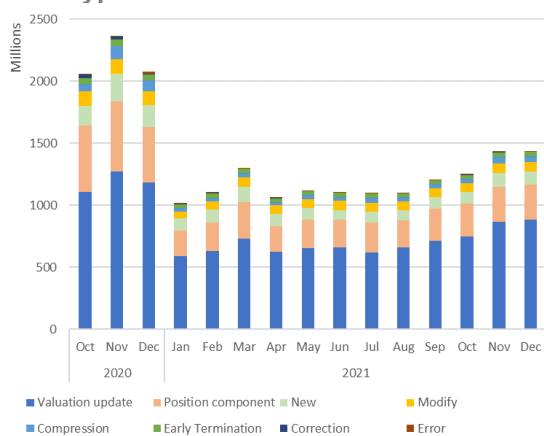**Monthly submissions per TR**
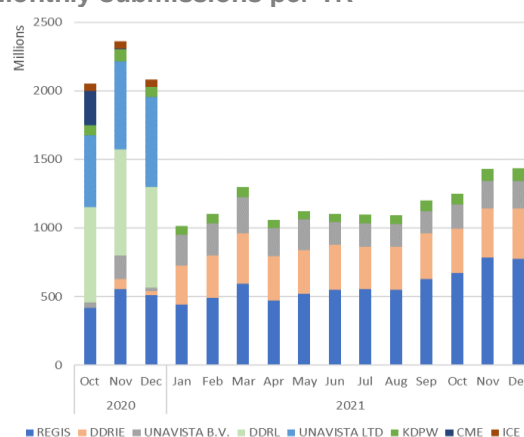
Note: Total number of submissions per month and TR. DDRIE is former DDRL and UNAVISTA B.V. is former UNAVISTA LTD.
Source: Trade Repositories & ESMA calculations

---

[16] Reported derivatives are considered duplicated where two or more records have been reported with the same combination of reporting counterparty ID, ID of the other counterparty and trade ID fields. In contrast, double-sided reporting under EMIR is not duplicative since two records on the same derivative should be reported always from the perspective of the respective reporting counterparty.

Chart 11 includes a breakdown of the total number of submitted EMIR reports by action type. The action type refers to the type of derivative life-cycle event reported in the submission by the reporting counterparty. In order of reported frequency, the most common submissions are valuation updates, position components, new derivatives, modifications, compressions, early terminations, corrections and errors. Valuation updates remain the most frequent action type submitted by the counterparties.

Chart 11
EMIR reporting key trends
Valuations are the most common submitted action type

Note: Total number of submissions per month and action type.
Source: Trade Repositories & ESMA calculations

Chart 12 shows that the share of each asset class has remained relatively stable during the year 2021. Notably there was a shift post Brexit with equites representing a higher portion amongst other asset classes, while commodity and emission allowances as well as credit occupy a lesser portion than previously.

Chart 13 provides a breakdown of EMIR reporting by contract type: despite a relative reduction after the end of 2020, futures remained in 2021 the most reported EMIR contract type (on average 40% of total reporting), followed by financial contracts for difference (30%), options (25%) and other types (5%).

Chart 12
EMIR reporting key trends
Equities are the most common EMIR asset class
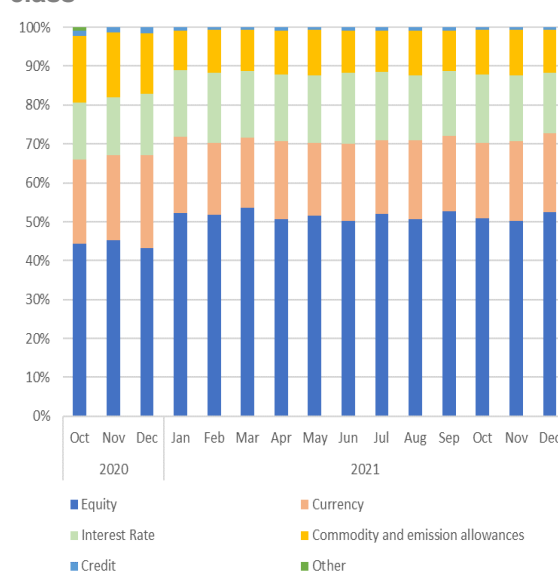
Note: Total number of daily submissions per month and asset class. As a percentage of total.
Source: Trade Repositories & ESMA calculations

Chart 13
EMIR reporting key trends
Futures are the most common EMIR contract type
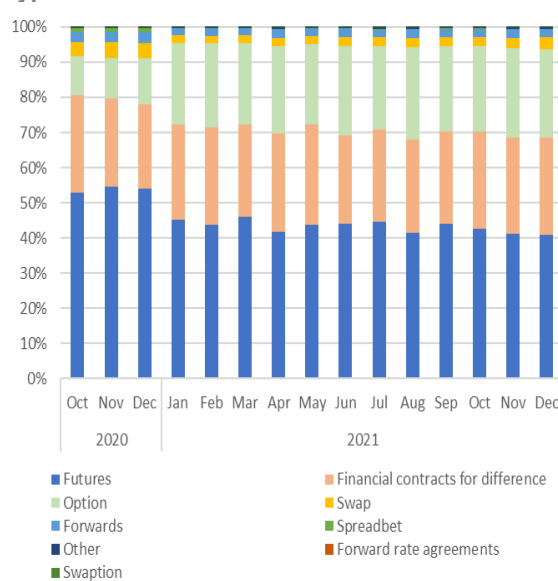
Note: Total number of daily submissions per month and contact type. As a percentage of total.
Source: Trade Repositories & ESMA calculations

6.2   Data completeness, timeliness and availability

EMIR Timeliness Analysis: Counterparties are required to report newly concluded derivative contracts by the end of the following working day to a TR of their choice. To assess the timeliness of reporting by the counterparties, ESMA considers the difference between the "Execution timestamp", reflecting the date and time of a derivative contract's conclusion, and the "Reporting timestamp", reflecting the date and time of reporting to the TR.
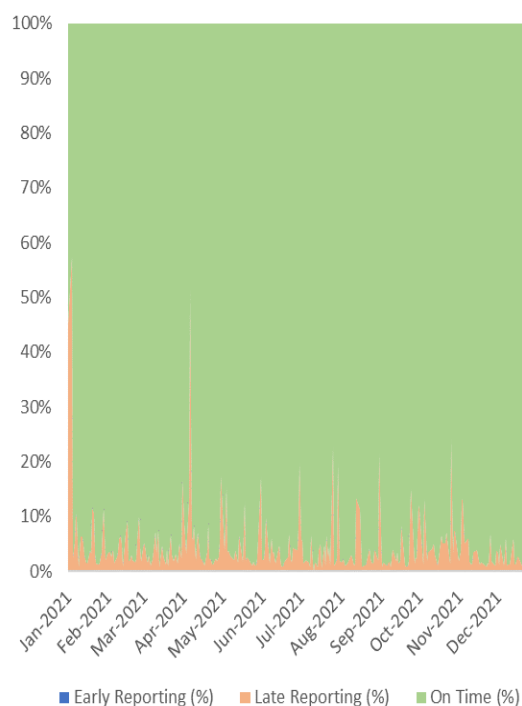
A derivative is considered "reported on time", if it is reported by the working day following the day on which the contract was concluded, at the latest. A derivative is considered "late reported", if it is reported later than the working day following the day on which the contract was executed. A derivative is considered "early reported", if it is reported earlier than the date specified in the "Execution timestamp" field.

A derivative concluded on a Friday or Saturday and reported on the consecutive Monday is subject to a "weekend effect" which is accounted for in the calculation and correctly classified as "on time".

On the contrary, public-, national- and bank holidays (i.e. "calendar effect") are not accounted for in the calculation, nor is the conversion of Coordinated Universal Time (UTC) to local time made. These approximations simplify and speed up the calculation but could give rise to some degree of inaccuracy (i.e., records wrongly classified as "Late Reporting" due to UTC vs local time differences, or due to calendar effect) impacting the overall results. Despite these methodology limitations, the analysis depicts a fair representation of reporting behaviour by the reporting entities.

Chart 14 shows the results obtained from the analysis of daily Trade Activity Reports (TAR) for 2021. The proportion of late reporting remained on average below 10% while early reporting is negligible or non-existent.

Chart 14
EMIR data completeness, timeliness and availability
EMIR TAR - Timeliness Analysis (daily)

Note: A derivative executed at time T and reported at T+1 at latest, is considered "On Time". A derivative executed at time T and reported after T+1 is considered "Late Reporting". A derivative executed at time T and reported before T is considered "Early Reporting". A derivative subject to "weekend effect" is classified as "On time". Source: Trade Repositories, and ESMA calculations using daily trade activity reports for 2021.
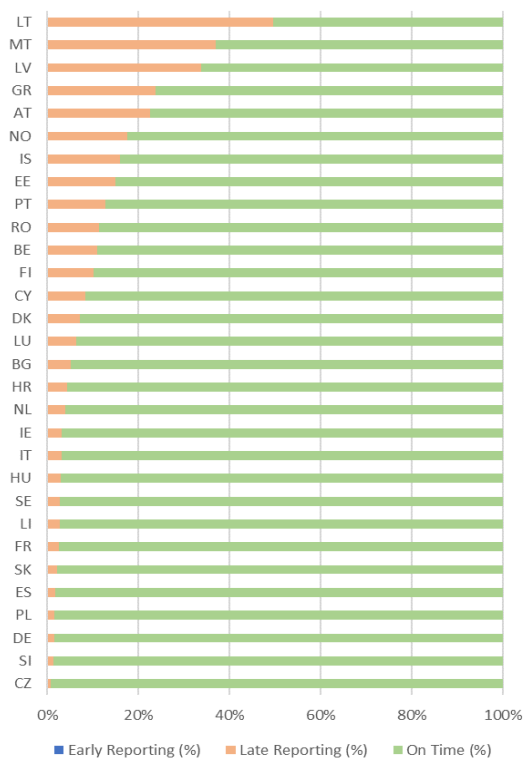
A few spikes in late reporting observed around New Year, Easter and other public holidays should be considered as merely a calendar effect impacting the accuracy of the results.

Chart 15 shows the aggregated results split by the jurisdiction of reporting counterparties. While in most jurisdictions the occurrence of late reporting remains limited and low, there are a few jurisdictions like Lithuania, Malta, Latvia, Greece, Austria, amongst others, where reporting entities tend to systematically report late.

It is often one or a few reporting entities that drives the build-up of late reporting which is why it is relevant for NCAs to intervene appropriately as targeted effort from their side can often rectify the problem quickly and bring down a significant portion of late reporting within its jurisdiction[17].

---

[17] For example, Maltese MFSA has engaged with two most problematic counterparties in their jurisdiction and is seeking remediation of the issue.

EMIR data completeness, timeliness and availability

EMIR TAR - Timeliness Analysis by jurisdiction

Note: A derivative executed at time T and reported at T+1 at latest, is considered "On Time". A derivative executed at time T and reported after T+1 is considered "Late Reporting". A derivative executed at time T and reported before T is considered "Early Reporting". A derivative subject to "weekend effect" is classified as "On time". Source: Trade Repositories, GLEIF and ESMA calculations using daily trade activity reports for 2021.

may also be partially caused by batches of back-dated reporting of previously non-reported trades.

However, monitoring of the evolution of timeliness reporting using TSR data can help to identify persistent patterns caused by the frequent as well as the sporadic misreporting events that would otherwise be difficult to identify.

Chart 16
EMIR data completeness, timeliness and availability

EMIR TSR - Timeliness Analysis (weekly)
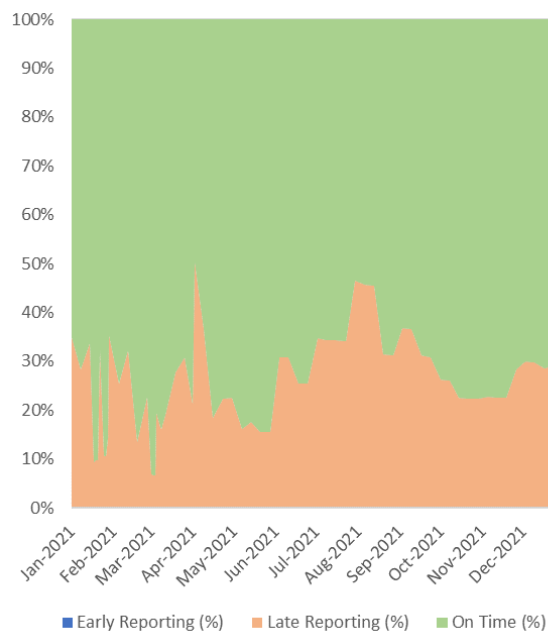
Note: A derivative executed at time T and reported at T+1 at latest, is considered "On Time". A derivative executed at time T and reported after T+1 is considered "Late Reporting". A derivative executed at time T and reported before T is considered "Early Reporting". A derivative subject to "weekend effect" is classified as "On time – weekend effect".
Source: Trade Repositories, and ESMA calculations using weekly trade state reports for 2021.

To complement and expand the previous analysis, ESMA applied the same calculation using a different dataset, namely the Trade State Report (TSR), which contains the latest state of all outstanding derivative contracts.

Chart 16 shows the results obtained from the analysis of the weekly TSR data for 2021. The proportion of late reporting is significantly higher compared to the results obtained using TAR data. This was somewhat expected as TSR data reflects the cumulative effect of daily reporting events.

Once a given derivative contract is executed and reported late, it is not possible to 'correct' the time of its initial submission. Therefore, such derivative contract will continue to appear as reported late in the TSR (at least until another lifecycle event is reported for the same contract). Moreover, the share of late reports in the TSR

Chart 17 shows the aggregated results split by the jurisdiction of reporting counterparties. Also, here the effect of late reporting is more pronounced compared to the results obtained using TAR data (see **Chart 15**). Thirteen out of thirty jurisdictions have a late reporting rate above 20% and a few are even above 40%.

Although the late reporting is perceived as a marginal issue when looking at TAR data, it becomes more relevant when analysing the cumulative effect of such misreporting behaviour in the TSR data.

Identifying reporting entities who consistently report late newly executed derivatives and rectifying such behaviour remains important for
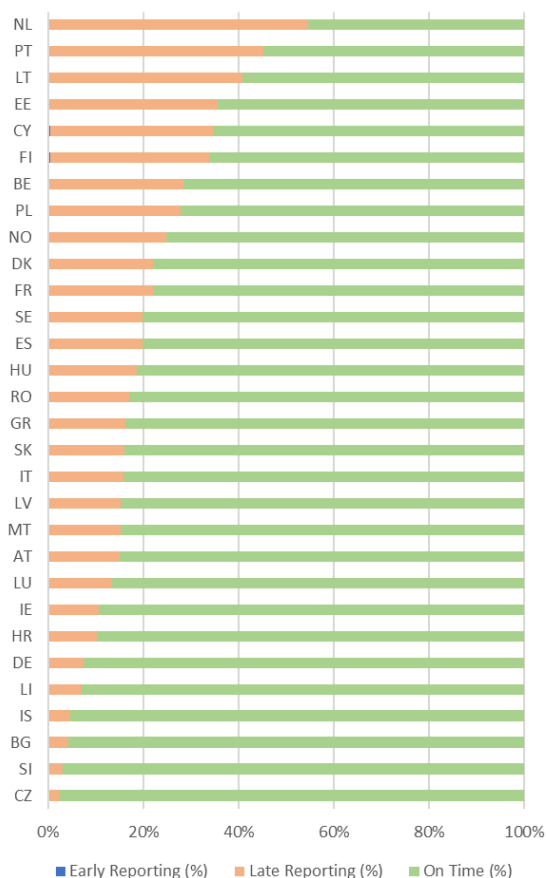
achieving an accurate picture of the derivatives' market at all times and thus for the efficient surveillance of the systemic risk.

Timeliness of reporting by counterparties has been the focus of NCAs during 2021 (see Subsection 4.3 Cooperation with data users).

Chart 17
EMIR data completeness, timeliness, and availability
EMIR TSR - Timeliness Analysis by jurisdiction

Note: A derivative executed at time T and reported at T+1 at latest, is considered "On Time". A derivative executed at time T and reported after T+1 is considered "Late Reporting". A derivative executed at time T and reported before T is considered "Early Reporting". A derivative subject to "weekend effect" is classified as "On time – weekend effect".
Source: Trade Repositories, GLEIF and ESMA calculations using the trade state report of 2021/21/31.
:

Non-reporting of valuations by counterparties: EMIR requires that financial and non-financial counterparties above the clearing threshold report daily the valuation and collateral data relating to their open derivatives.

To assess the timeliness of reporting of the valuation updates, ESMA computes the difference in number of days between the reference date of a Trade State Report (TSR) and the "Valuation timestamp" of a record, which reflects the date and time of a valuation update.

While a stricter reading of EMIR would mean that valuations that are older than one working day are outdated, for the purposes of this analysis it is considered that valuation updates older than 15 calendar days are outdated and should have been subject to new valuation updates. Four distinct buckets are used to group each record that is in scope for this analysis and measure how frequent valuation updates occur.

Chart 18 shows that around 80% of open derivatives have received valuation updates that are not older than 15 days. It also shows that around 20% of open derivatives subject to daily valuation have not received updates for several days, months and even years.

Chart 18
EMIR data completeness, timeliness, and availability
EMIR TSR - Valuation updates
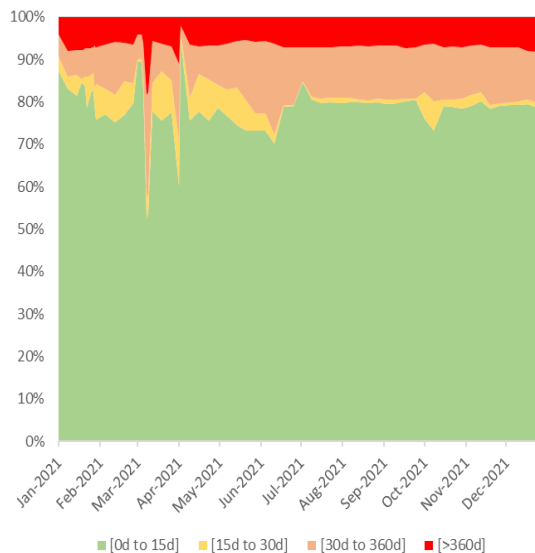
Note: The analysis uses all open derivatives from Trade State Reports (TSR) with action type = "N" (new) or "V" (valuation) and clearing threshold = "Y". The difference in number of days is computed between the reference date of the TSR and the valuation timestamp of a record. Each record is grouped into one of the four buckets to measure the magnitude of number of outstanding trades that have or have not received a valuation update between a certain number of days.
Source: Trade Repositories and ESMA calculations

A significant number of open derivatives with outdated valuation timestamps could indicate misreporting practices by counterparties and/or open derivatives that have not been properly terminated (i.e., "dead" trades).

Chart 19 shows the aggregated results split by the jurisdiction of reporting counterparties. While in most jurisdictions the stock of open derivatives is frequently receiving valuation updates, in a few other jurisdictions there is a significant portion of open derivatives that are not being updated. The problem appears to be mostly prominent in Bulgaria, Estonia, Latvia and Malta based on the

results obtained from the trade state report of 31 December 2021[18].

Considering the importance of the valuation data for economic and financial risk analysis, NCAs and ESMA will continue focusing on the completeness and timeliness of valuation reporting going forward.

Timeliness of valuations reporting by counterparties has been the focus of NCAs during 2021 during the EMIR DQR (see Subsection 4.1 EMIR DQAP).

Chart 19
EMIR data completeness, timeliness, and availability
EMIR TSR - Valuation updates by jurisdiction
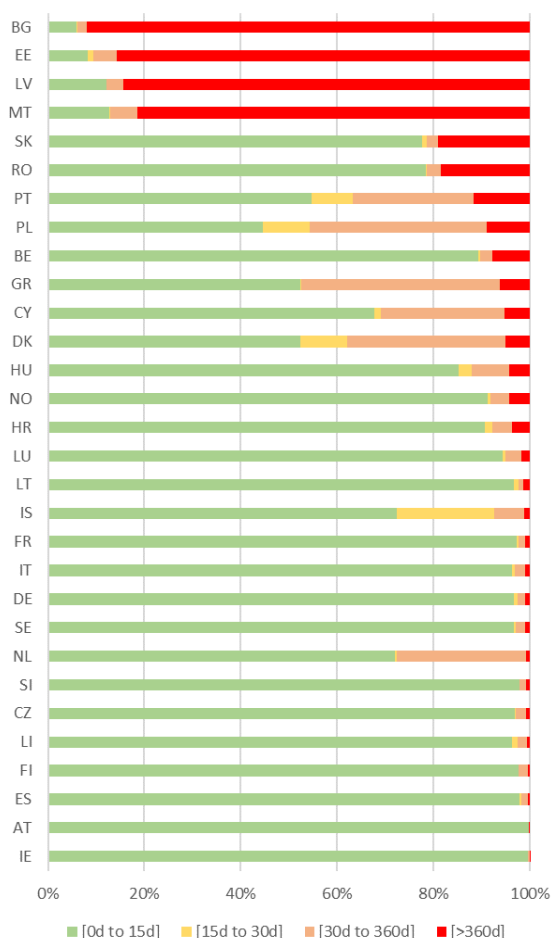
Note: The analysis uses all open derivatives from Trade State Reports (TSR) with action type = "N" (new) or "V" (valuation) and clearing threshold = "Y". The difference in number of days is computed between the reference date of the TSR and the valuation timestamp of a record. Each record is grouped into one of the four buckets to measure the magnitude of number of outstanding trades that have or have not received a valuation update between a certain number of days.
Source: Trade Repositories, GLEIF and ESMA calculations using the trade state report from 2021/12/31.

Non-reporting of derivatives: **Chart 20** shows the scale of potential non-reporting. It is not possible to estimate the non-reporting fully, however, due to the double-sided reporting obligation ESMA has estimated the potential scale of the problem by identifying derivatives where only a report in one direction was submitted and a report from the other direction is expected, i.e., the other counterparty is in the EEA, and it has a reporting obligation. Prior to Brexit, the number of potentially non-reported derivatives stood at around 3.5 million (approximately 5% of open reconcilable derivatives at the end of 2020). After Brexit, with the end of reporting obligation of UK counterparties the number has sharply dropped to around 0.5 million open derivatives (around 1% of open reconcilable derivatives).

Importantly, the drop in number of non-reported derivatives is driven by the cease of reporting obligation in the UK. While the reporting obligation of UK counterparties may have stopped, the EU side of the trade still has its reporting obligation, and it is expected to report. Thus, with Brexit, potential non-reporting has become somewhat harder to detect.

Chart 20
EMIR data completeness, timeliness, and availability
**Potential non-reporting sharply dropped after Brexit**
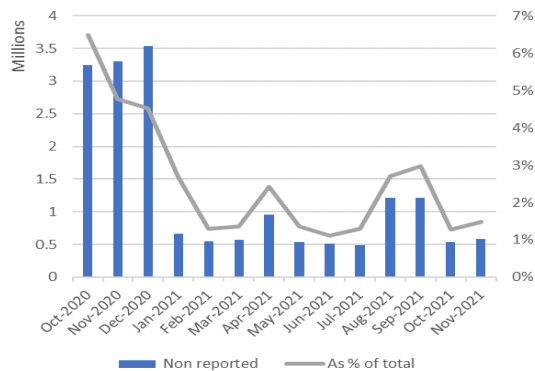
Note: Number of potentially nonreported open derivatives by date. The estimate is based on the number of reports where only one leg of the derivative was reported and the second leg is expected, i.e., the second counterparty is EEA, and it has a reporting obligation (legal entity).
Source: Trade Repositories & ESMA calculations

Chart 21 shows a breakdown of open derivatives where the report of the second counterparty to a derivative is missing grouped by the jurisdiction of the reporting counterparty. The largest jurisdictions in terms of open derivatives also tend to have the highest number of potentially non-reported derivatives.

---

[18] For example Maltese MFSA confirmed that it engaged with the most problematic counterparties in their jurisdiction and they are now actively working on the remediation of the issue.

Chart 21

EMIR data completeness, timeliness, and availability

Top 5 European jurisdictions take above 75% of all potentially non-reported derivatives
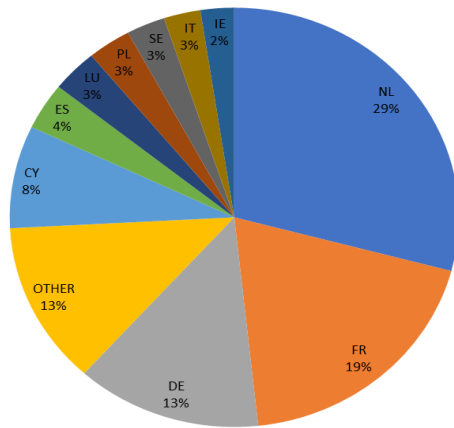

Note: The designation "OTHER" includes countries with less than 2% values. Those are, in descending order, DK, EE, NO, FI, LV, AT, BE, LI, HU, CZ, MT, PT, LT, BG, GR, HR, SK, RO, SC, IS, SI, GB, PA, KY, GI and JE.
Source: Trade Repositories & ESMA calculations

6.3 Data integrity – Adherence to format and content rules

As shown in the **Chart 22,** the rejection statistics provided by TRs pointed to a transitory uptick in the rejection rates from April to August. A further breakdown of the data indicates that the elevation in the rejection rate affected TR2 in April and TR4 in June to July, as illustrated in the **Chart 23**.

Chart 22

EMIR data Integrity – Adherence to format and content

Total volume and rejections as % of total
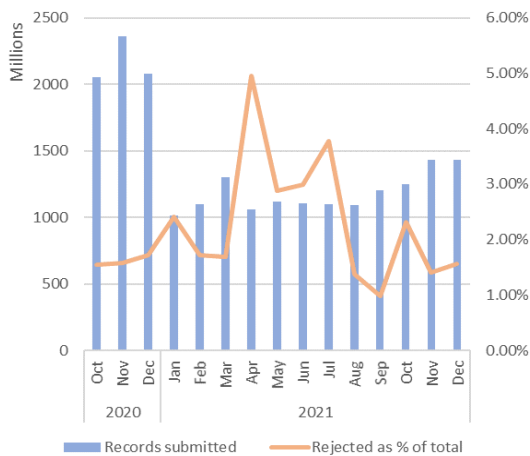
Source: Trade Repositories & ESMA calculations

Chart 23

EMIR data Integrity – Adherence to format and content

Rejection rate by TR
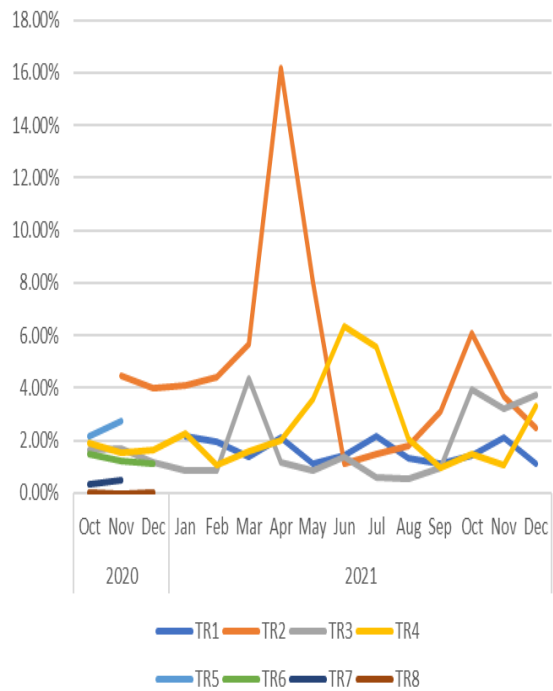
Note: Rejection % of total submissions received by each TR. The repositories included in this chart are ICE, CME, DDRIE, DDRL, KDPW, UNAVSITA B.V., UNAVISTA LTD & REGIS, the anonymized aliases do not correspond to the order of the TRs named.
Source: Trade Repositories & ESMA calculations

Duplicated reporting: Unique derivative is identified based on three EMIR fields, i.e., reporting counterparty ID, ID of the other counterparty and trade ID. To avoid undue double-counting when using the data for economic/financial analysis, it is essential that counterparties report in a way that the uniqueness of each derivative is respected. TRs are expected to verify the uniqueness of reported new derivatives and to reject those that have been reported with the same triplet of IDs in the past. TRs are, however, unable to identify duplicates when a counterparty reports to two different TRs.

To assess the volumes of duplicate reports, ESMA performed analysis of the uniqueness of the derivatives in the Trade State Report.

Chart 24 shows a percentage of duplicates at TR level ("intra-TR"). While from time to time, the number of duplicates can increase, the overall number is less than one percent. Thus, at TR level, duplicated reporting does not pose significant issues.

Chart 24

EMIR data Integrity – Adherence to format and content

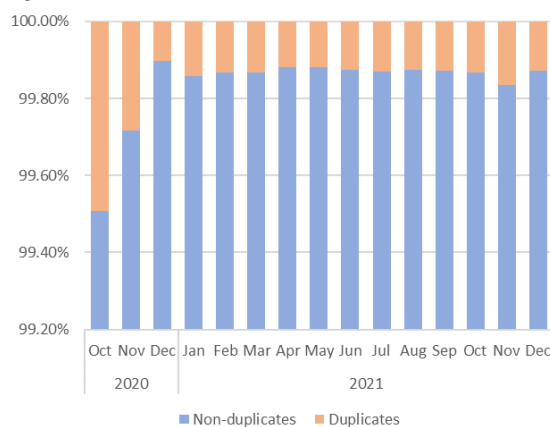**Intra TR duplicate records as % of total reported volume**
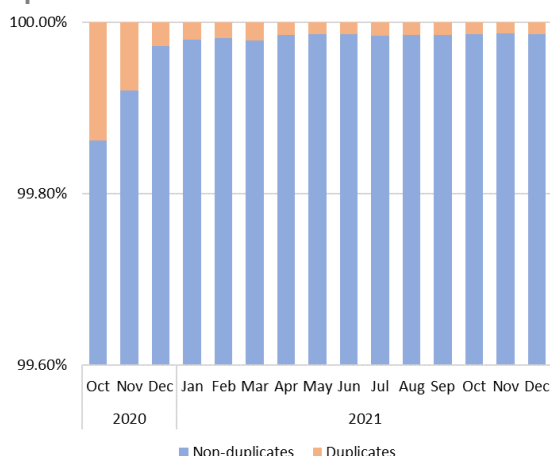
Note: Duplicated open derivatives as % of all open derivatives. A unique derivative is defined at the level of three EMIR fields: reporting counterparty ID, ID of the other counterparty, and trade ID. Intra-TR duplicates are detected in individual TR reports submitted to ESMA.
Source: Trade Repositories & ESMA calculations

Chart 25 shows a percentage of duplicates across TRs ("inter-TR"). Similarly, to the results shown above, the duplicated records do not seem to pose significant issues even at the inter-TR level.

Chart 25

EMIR data Integrity – Adherence to format and content

Inter TR duplicate records as % of total reported volume

Note: Duplicated open derivatives as % of all open derivatives. A unique derivative is defined at the level of three EMIR fields: reporting counterparty ID, ID of the other counterparty, and trade ID. Inter-TR duplicates are detected across all TRs reporting to ESMA.
Source: Trade Repositories & ESMA calculations

Revalidation: When a counterparty submits reports to a TR, the latter needs to validate whether the incoming data is in line with the regulatory reporting requirements. For this purpose, TRs have implemented the ESMA's validation rules against which they check the incoming data. TRs are expected to reject reports that are not adhering to the validation rules.

Since the introduction of the validation rules in December 2014, ESMA regularly performs a revalidation of the data made available by the TRs with a view to assess whether TRs have implemented the validation requirements correctly. In ESMA's analysis, a randomly selected data sample extracted from one daily Trade Activity Report per month is used. Each data point is checked against the ESMA validation rules in force at the time of the verification performed by the TRs. Following the identification of an issue, e.g., a specific field that causes unduly rejections, ESMA engages with the relevant TR to remediate the issue at hand.

Chart 26 shows the number of records analysed and the percentage of errors. On average, each iteration of the analysis processed 8 million records. The proportion of records containing errors remained low and stable, fluctuating around 1% during 2021. ESMA will continue to perform the revalidation analysis for monitoring the evolution and liaise with TRs when material issues are detected.

Chart 26

EMIR data Integrity – Adherence to format and content

Number of failures/breaking records as % of total
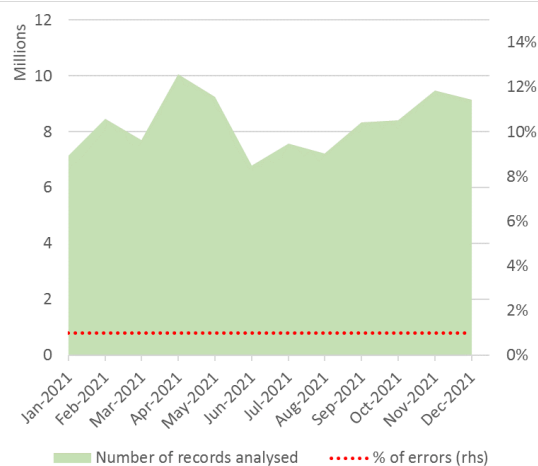
Note: The analysis uses a randomly selected data sample (~15%) extracted from one daily Trade Activity Report per month. Each data point is checked against the current ESMA validation rules.
Source: Trade Repositories and ESMA calculations

6.4 Data integrity – Reconciliation

Reconciliation: Under EMIR, both counterparties to the derivative are required to report their side, to the extent that they are subject to the reporting obligation. TRs are then required to reconcile the

data. The TRs provide the results to the reporting participants (so that any reconciliation breaks can be addressed) and to the authorities (so that they can monitor reporting of the counterparties in their jurisdiction).

Chart 27 shows the results of pairing[19]. The pairing rate stood at around 60% at the end of 2021. The rate has remained relatively stable throughout the year. Considering that pairing is performed by comparing three EMIR fields only, its current level is not satisfactory. Unfortunately, there are several reasons for lack of pairing such as lack of agreement on the trade ID between counterparties, under- and overreporting, wrong identification of the other counterparty or lack of agreement on the number of reports that should be submitted in relation to a given derivative.

Chart 27
EMIR data integrity – Reconciliation
Reconciliation: Pairing has stabilised at around 60%
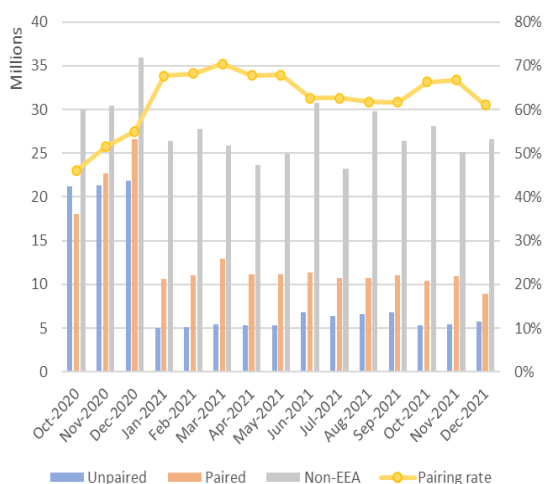
Note: Pairing is performed based on three fields: Reporting counterparty ID, ID of the other counterparty and trade ID. Pairing rate is calculated by paired derivatives by sum of paired and unpaired (excluding non-EEA derivatives).
Source: Trade Repositories & ESMA calculations

Chart 28 provides some potential further insights as to the reasons for lack of pairing. The chart depicts the net difference between the number of derivatives reported by the two counterparties. Prior to Brexit, the difference stood at around 8 million (10% of all open derivatives). After Brexit, the number has dropped as UK counterparties no longer have a reporting obligation (and the number of derivatives with expected two legs of a trade has dropped). Throughout 2021, the number was around 2 million representing around 5% of all open derivatives where 2 reports

are expected, i.e., excluding derivatives where the other counterparty is non-EEA.

Chart 28
EMIR data integrity – Reconciliation
Difference in number of records reported by the two sides
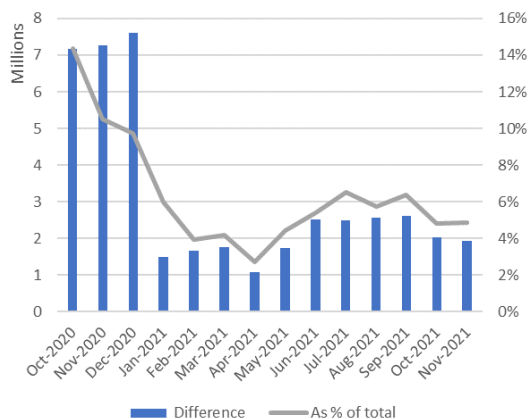

Note: The metric is calculated by taking a difference between the number of derivatives reported by leg 1 and leg 2.
Source: Trade Repositories & ESMA calculations

Chart 29 shows a breakdown of open derivatives where the report of the second counterparty to a derivative is missing grouped by the jurisdiction of the counterparty. Top 3 jurisdictions represent nearly 70% of all issues.

Chart 29
EMIR data integrity – Reconciliation
Difference in number of records reported by the two sides by jurisdiction
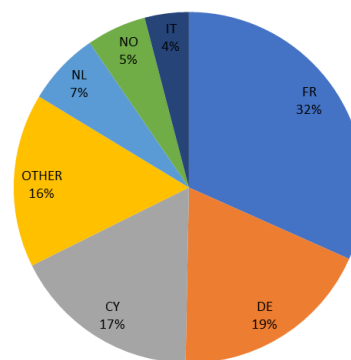
Note: The designation "OTHER" includes countries with less than 2% values. Those are, in descending order, ES, DK, LU, SI, PL, SE, IE, BE, LV, FI, LI, EE, AT, HU, PT, CZ, SK, MT, BG, LT, GR, HR, RO, IS and GB.
Source: Trade Repositories & ESMA calculations

Given that EMIR reports are characterized by a large volume of data, reconciliation could be for TRs a complex process. To assess the reliability
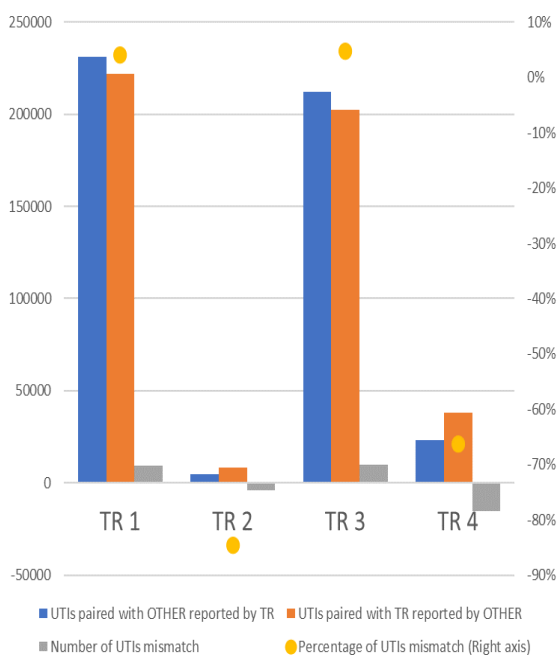
---

[19] Pairing is performed on the basis of three fields: Reporting counterparty ID, ID of the other counterparty and trade ID.

of reconciliation statistics, ESMA started in 2021 to collect and analyse periodic information requested from TRs in the context of its supervisory activities. As part of this information, TRs must provide the number of UTIs paired and reconciled internally (when both sides of the trade are reported to the same TR) or with other TRs (when the two sides of the trade are reported to two different TRs).

While reconciliation has been the focus of NCAs during 2021 during the EMIR DQR (see Subsection 4.1 EMIR DQAP), the results do not yet show sufficient progress. Reconciliation thus needs to remain a point of focus going forward.

Chart 30 and **Chart 31** provide the preliminary results of an analysis of the discrepancies in the inter-TR reconciliation statistics: the number of UTIs reported by each TR as paired/reconciled versus the other TRs has been compared with the number of UTIs reported by the other TRs as paired/reconciled versus each TR.

The main finding of this analysis is that, although pairing statistics do not seem to be affected by significant inter-TR discrepancies, there are relevant divergences in the number of UTIs reported as reconciled among the TRs. On top of the complexity of the reconciliation process, discrepancies in the inter-TR statistics could be caused also by the different times at which each TR submits its data to reconciliation.

ESMA will continue to collect such information and to investigate potential TR-specific issues.

Chart 31
EMIR data integrity – Reconciliation
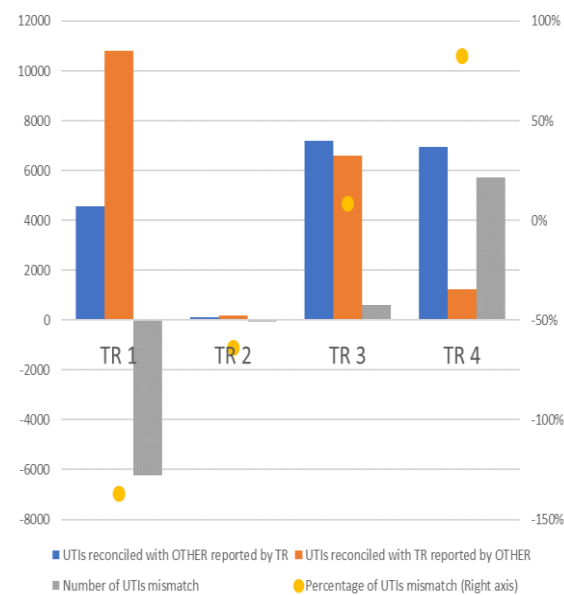**Discrepancies in number of UTIs reconciled versus Other TRs**

Note: The charts are based on information provided monthly by TRs to ESMA in the context of periodic information (Item 36 – Reconciliation Statistics). The figure in the charts refers to outstanding trades and are computed as the monthly average for the reference period June 2021 – December 2021.
Source: Trade Repositories & ESMA calculations

.

Chart 30
EMIR data integrity – Reconciliation
Discrepancies in number of UTIs paired versus Other TRs

Note: The charts are based on information provided monthly by TRs to ESMA in the context of periodic information (Item 36 – Reconciliation Statistics). The figures in the charts refer to outstanding trades and are computed as the monthly average for the reference period June 2021 – December 2021
Source: Trade Repositories & ESMA calculations

7.SFTR reporting trends and selected data quality metrics

Summary: Similar to EMIR, data reporting volumes dropped approximately by 50% following Brexit. In terms of number of open transactions, securities lending and borrowing is the largest SFT type reported with around 70% share at the end of 2021. Credit institutions report most open SFTs (around 50%) while credit institutions share has been increasing (to around 30% at the end of 2021). After merely 1.5 years of reporting, SFTR exhibits comparable results to EMIR across all data quality metrics. Around 10% of SFTs are reported late (after T+1). On the contrary, rejections have been low (around 2%) and duplicated reporting does not pose major issues. As regards reconciliation, pairing rate has been only around 60%. Reconciliation rate of loan and collateral data has been low but increasing to around 40% and 30% respectively. Similar to EMIR, TRs do not agree on the number of records they reconcile against each other, which may be an indication of issues in the inter-TR reconciliation process. Timeliness of reporting , adherence to format and content rules (via rejections) and reconciliation (pairing) has been the point of focus of ESMA and NCAs during 2021. While progress has been made, some areas (particularly reconciliation) need to remain areas of focus also in the future.

7.1   Data reporting – Key trends

SFTR reporting started in July 2020 and was followed by a phased-in period - concluded in January 2021 – during which reporting requirements were gradually extended to different types of counterparties. Given a more extensive availability of data reported under SFTR in 2021, this section of the report contains more elaborated analyses than in the previous edition. Since SFTR data have a similar structure to EMIR, reporting trends can be analysed from similar perspectives (i.e., types and volumes of life-cycle events, open contracts and reporting counterparties).

TRs are required to submit daily to NCAs and ESMA a set of 4 reports providing a thorough overview on SFTR reporting activity:

1) the Trade Activity Report (TAR), in which TRs provide all the life-cycle events reported by SFT counterparties on the reference date.
2) the Trade State Report (TSR), which provides a snapshot of all the outstanding SFTs at the reference date (i.e., incorporating all the reported life-cycle events and applying them to the respective SFT records).
3) the Rejection Report, which contains statistics – at file and SFTs level – on the

acceptance/rejection of the reports received by the TRs on the reference date (see Subsection 7.3 Data integrity – Adherence to format and content).
4) The Reconciliation Report, which provides information – at a counterparties-pair level – on the reconciliation activity performed by TRs on expired and outstanding trades at the reference date (see Subsection 7.4 Data integrity – Reconciliation).

Key trends: Due to the removal of reporting requirements for UK counterparties after Brexit, SFTR reporting volumes significantly decreased in 2021, falling to an average of 49 million life-cycle events reported per month.

As shown in **Chart 32,** there were 4 TRs providing SFTR reporting services in 2021: DDRIE (DDRL before Brexit), which received the largest share of submissions (on average 41 million records per month in 2021), followed by UnaVista (5 million), Regis-TR (3 million) and KDPW (0.1 million). It is worth mentioning that the volumes reported by UnaVista in the TAR started to decrease in December 2021 as a consequence of the voluntary withdrawal of the TR from the provision of reporting services for SFTR, which is being finalised in the first half of

2022 (see Subsection 5.5 SFTR – UnaVista wind-down).

Chart 32
SFTR data reporting key trends
SFTs reported by Trade Repository
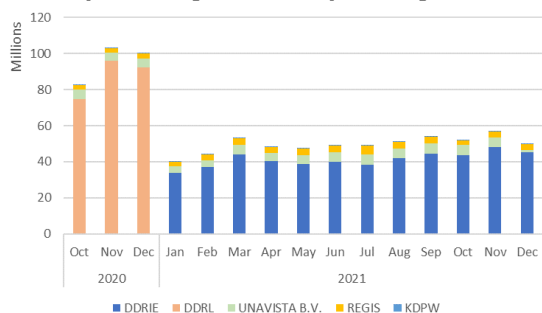
Note: Until Brexit, DDRL was the TR (UK entity part of the DTCC group) authorized by ESMA for EMIR and SFTR reporting services. After Brexit, ESMA authorization of DDRL was withdrawn and granted to DDRIE (DTCC-entity based in EU)
Source: Trade Repositories & ESMA calculations

Chart 33 provides the figures of SFTR reporting volumes broken down by action type. Likewise for EMIR, the action type field refers to the type of life-cycle event reported for the SFT. The most reported action types are the ones referring to contract modification and valuation updates.

Chart 33
SFTR data reporting key trends
SFTs reported by action type
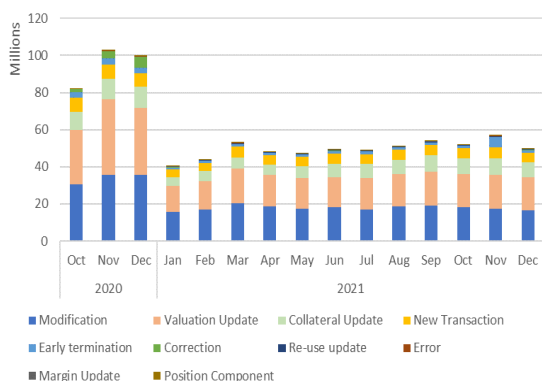
Note: Total number of submissions per month and action type.
Source: Trade Repositories & ESMA calculations

Through the periodic monitoring of SFTR reports, ESMA detected a relevant data quality issue inflating the number of outstanding SFTs caused by one counterparty failing to report maturity date on its transactions. As shown in **Chart 34** the number of open SFTs has been constantly increasing until November 2021, when – as a result of an action taken by the relevant competent authority – the counterparty corrected

the data reporting errors by submitting backloaded early terminations.

Chart 34
SFTR data reporting key trends
Inflation in open SFTs reports caused by one reporting counterparty
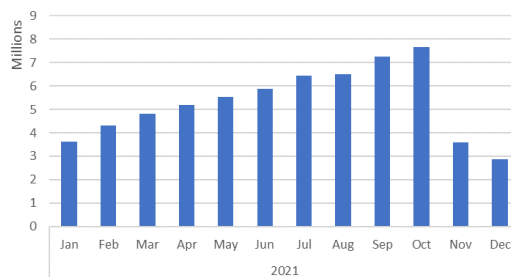
Note: Extract of the TSR submitted by the TR contracted by one SFT counterparty missing to report early termination life-cycle events.
Open SFTs, are transactions that have not matured, or which have not been the subject of reports with action types 'Error', 'Termination/Early termination', or 'Position component'.
Source: Trade Repositories & ESMA calculations

According to the TSR submitted in 2021, it results that Securities and Commodities Lending or Borrowing (SLEB) is the most common type of outstanding SFTs as shown in **Chart 35**. The notable drop in the percentage of SLEB at the end of 2021 is attributable to the data quality issue mentioned in the previous paragraph.

Chart 35
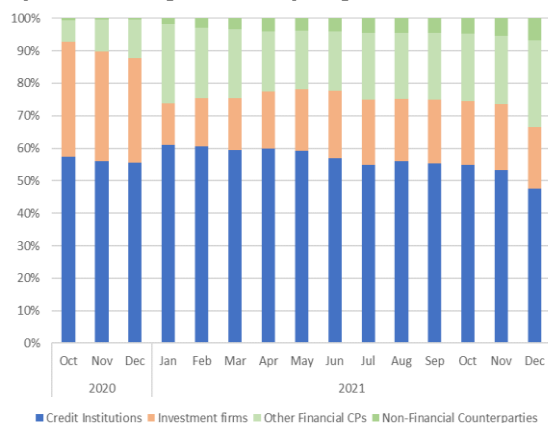SFTR data reporting key trends
Open SFTs by type
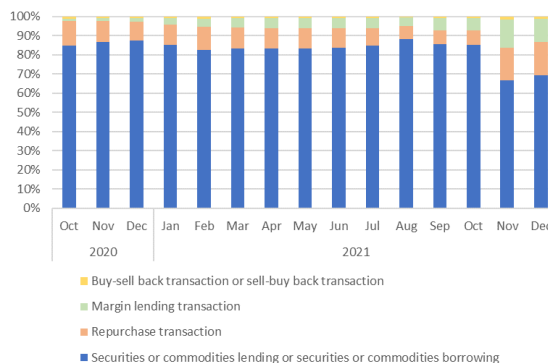
Note: Total number of open SFTs per month and SFT type, as percentage of the total.
Open SFTs, are transactions that have not matured, or which have not been the subject of reports with action types 'Error', 'Termination/Early termination', or 'Position component'.
The chart contains all open SFTs including those that were 'overreported' as shown in Chart 35.
Source: Trade Repositories & ESMA calculations

Chart 36 shows the distribution of outstanding SFTs providing a breakdown by type of reporting counterparty: most of the open SFTs have been

reported by credit institutions (on average 60% of the outstanding SFTs), followed by Investment Firms (20%). It is notable how in January 2021 the distribution of types of counterparties changed (less Investment firms due to Brexit and more Other Financial CPs and Non-Financial CPs after the SFTR phase-in).

Chart 36
SFTR data reporting key trends
Open SFTs by Counterparty Sector

Note: Total number of open SFTs per month and reporting counterparty sectors, as percentage of the total. Other Financial CPs are Insurance and Re-insurance firms, AIFMs, Pension Funds and UCITs.
Open SFTs, are transactions that have not matured, or which have not been the subject of reports with action types 'Error', 'Termination/Early termination', or 'Position component'.
The chart contains all open SFTs including those that were 'overreported' as shown in Chart 35.
Source: Trade Repositories & ESMA calculations

7.2 Data completeness, timeliness, and availability

SFTR Timeliness Analysis: The timeliness of reporting under SFTR was performed using the same concept and methodology as for EMIR (see section 6.2).

Chart 37 shows the results obtained from the analysis of daily Trade Activity Reports (TAR) for 2021. The proportion of late reporting remained on average below 10% while early reporting is negligible or non-existent.

Chart 37
SFTR data completeness, timeliness, and availability
SFTR TAR - Timeliness Analysis (daily)
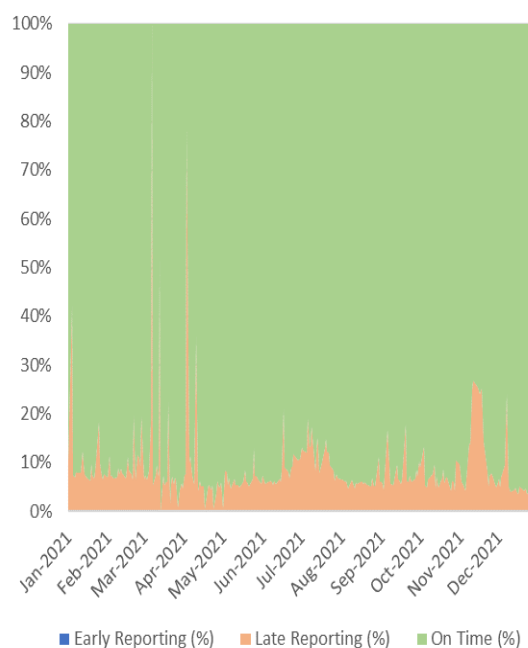
Note: An SFT executed at time T and reported at T+1 at latest, is considered "On Time". An SFT executed at time T and reported after T+1 is considered "Late Reporting". An SFT executed at time T and reported before T is considered "Early Reporting". An SFT subject to "weekend effect" is classified as "On time – weekend effect".
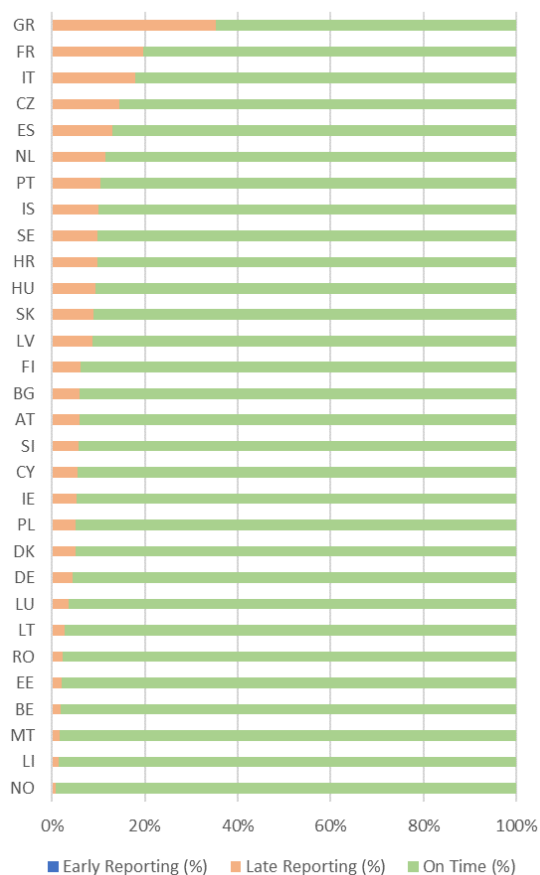Source: Trade Repositories, GLEIF and ESMA calculations using daily trade activity reports for 2021

A few spikes in late reporting observed around New Year, Easter and other public holidays should be considered as merely a calendar effect impacting the accuracy of the results. In other cases, a more persistent upward trend accompanied by a drop in late reporting is observed. This could be an indication of misreporting behaviour by reporting entities.

Chart 38 shows the aggregated results for the full 2021 split by the jurisdiction of the reporting counterparties. While in most jurisdictions the occurrence of late reporting remains limited and low, there are a few jurisdictions like Greece, France, and Italy where reporting entities have higher volumes of late reports on an aggregated level.

Chart 38

SFTR data completeness, timeliness, and availability

SFTR TAR - Timeliness Analysis by jurisdiction

Note: An SFT executed at time T and reported at T+1 at latest, is considered "On Time". An SFT executed at time T and reported after T+1 is considered "Late Reporting". An SFT executed at time T and reported before T is considered "Early Reporting". An SFT subject to "weekend effect" is classified as "On time – weekend effect". Source: Trade Repositories, GLEIF and ESMA calculations using daily trade activity reports for 2021

It is often one or a few reporting entities that drive the build-up of late reporting, which is why it is relevant for NCAs to intervene timely and appropriately. Therefore, a targeted effort from NCA's side can often rectify the problem swiftly and bring down a significant portion of late reporting within its jurisdiction.

To complement and expand the previous analysis, ESMA applied the same calculation using a different dataset, namely the Trade State Report (TSR) which contains the latest state of all outstanding SFTs.

Chart 39 shows the results obtained from the analysis of daily TSR data for 2021. The proportion of late reporting is slightly lower compared to the results obtained using TAR data,

while early reporting, which was non-existent in TARs, is an apparent issue in this analysis.

Chart 39

SFTR data completeness, timeliness, and availability

SFTR TSR - Timeliness Analysis (daily)
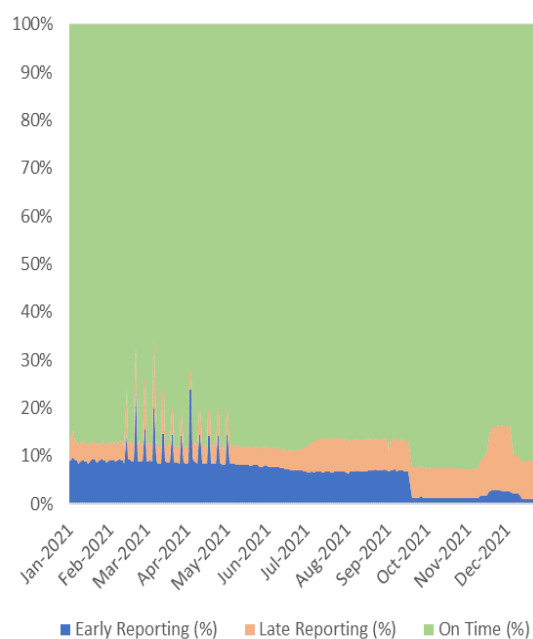
Note: An SFT executed at time T and reported at T+1 at latest, is considered "On Time". An SFT executed at time T and reported after T+1 is considered "Late Reporting". An SFT executed at time T and reported before T is considered "Early Reporting". An SFT subject to "weekend effect" is classified as "On time – weekend effect". Source: Trade Repositories, GLEIF and ESMA calculations using daily trade state reports for 2021

Once an SFT is concluded and reported early or late, it is not possible to 'correct' the time of its initial submission. Therefore, such SFT will continue to appear as reported early or late in the TSR.

Consequently, the significant stock of early reported SFTs visible in the TSR occurred before 2021 as the TARs of 2021 does not show such reporting behaviour. The issue has been encapsulated in the TSR, vanishing only once those SFTs reach maturity or are terminated, as it can be observed in the sharp drop that occurred towards the end of 2021. Moreover, the share of late reports in the TSR may also be partially caused by batches of back-dated reporting of previously non-reported trades.
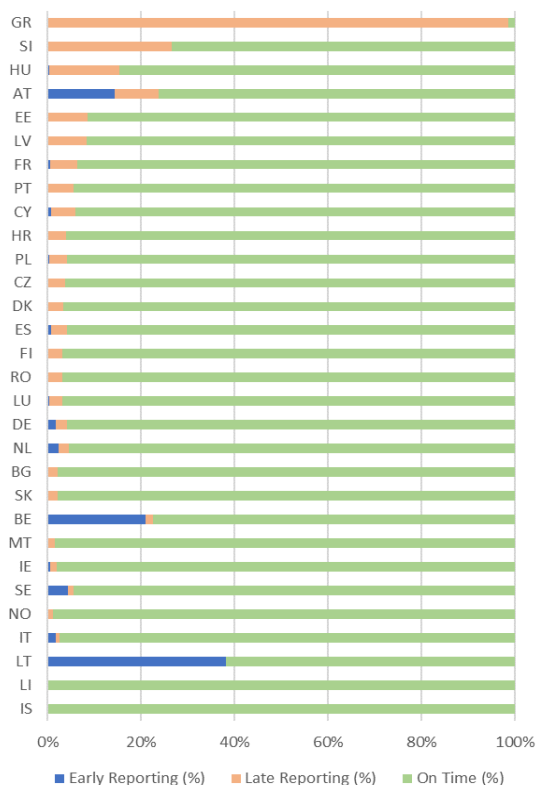
Monitoring of the evolution of timeliness reporting using TSR data can help to identify patterns caused by frequent as well as sporadic misreporting events occurring both in the present as well as in the past which would otherwise be difficult to identify.

Chart 40 shows the aggregated results split by the jurisdiction of reporting counterparties. The effect of early reporting is more pronounced compared to the results obtained using TAR data (see **Chart 38**).

SFTR TSR - Timeliness Analysis by jurisdiction

While early reporting is evident in Lithuania, Belgium, and Austria, late reporting is more apparent in Greece, Slovenia, and Hungary.

Early reporting becomes more relevant when analysing the cumulative effect of such misreporting behaviour using TSR data. In some jurisdiction above 20% of open SFTs have been reported with the event date being later than the reporting timestamp.

Identifying reporting entities who consistently misreport newly executed SFTs and rectifying

such behaviour remains important for achieving an accurate picture of SFTs' market at all times and thus for the efficient surveillance of the systemic risk.

Timeliness of reporting by counterparties has been the focus of NCAs during 2021 during the SFTR DQEF (see Subsection 4.2 SFTR DQEF).

7.3 Data integrity – Adherence to format and content rules

As the SFTR regime settles in, the rejection rate for SFT submissions continues a downward trend. Furthermore, **Chart 41** displays the significant decrease in the volume of rejected records. This trend has been noticeable mainly since May-June and has become more pronounced throughout the rest of the series. The positive decreasing trend is the result of the interaction with TRs and the joint work of NCAs and ESMA in the context of the SFTR DQEF on the quality of reporting by the counterparties (see Subsection 4.2 SFTR DQEF).

Total volume and rejection as % of total
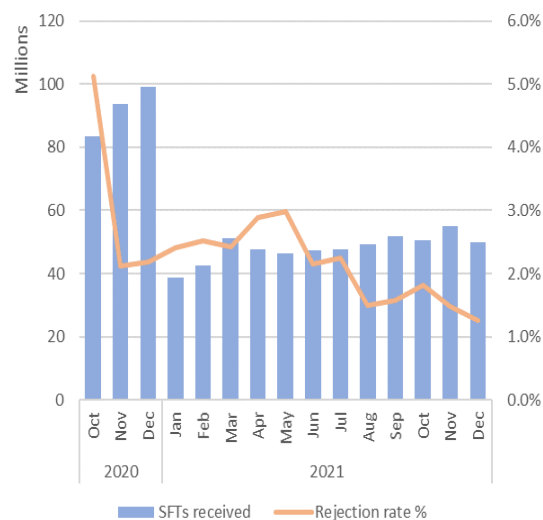
Duplicated reporting: When validating transaction messages submitted by reporting participants, TRs are asked among other things, to validate the incoming data and ensure that no duplicates[20] are being reported.

---

[20] Unique transaction under SFTR is defined on the basis of three fields: Reporting counterparty ID, ID of the other

counterparty and trade ID. In the case of collateral data submissions, master agreement type is also used.

Having said that, TRs can validate uniqueness of reporting on data that is reported to them, i.e., a TR cannot validate whether another TR receives the same record.

To verify uniqueness of reporting, ESMA identifies duplicated records on, both, TR level (duplicates that TRs are expected to identify are reject) and across TRs (duplicated records which cannot be identified by TRs, and which are pure counterparty misreporting issue).

Chart 42 and **Chart 43** depict split between duplicated and unique open SFT transactions on monthly basis on a TR level (intra-TR) and across all TRs (inter-TR). While there have been one-off incidents at one TR where large number of duplicates have appeared in the regulatory report (particularly during end of 2020 and early 2021), the results currently show that only a small fraction of records is duplicated, and this issue does not, now, pose any real problems for end users.

Chart 42
SFTR data integrity – Adherence to format and content
Intra TR duplicate records as % of total reported volume
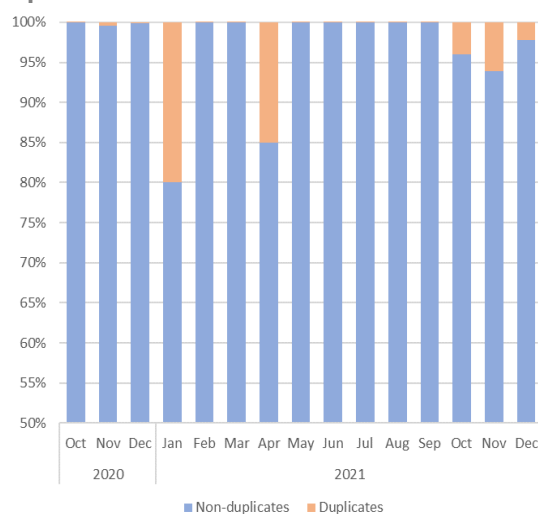
Note: Duplicated transactions are identified based on three fields: Reporting counterparty ID, ID of the other counterparty and trade ID. In case collateral report, master agreement type is used as well. The analysis has been performed on open SFT transactions on a given reference date.
Source: Trade Repositories & ESMA calculations

Chart 43
SFTR data integrity – Adherence to format and content
Inter TR duplicate records as % of total reported volume
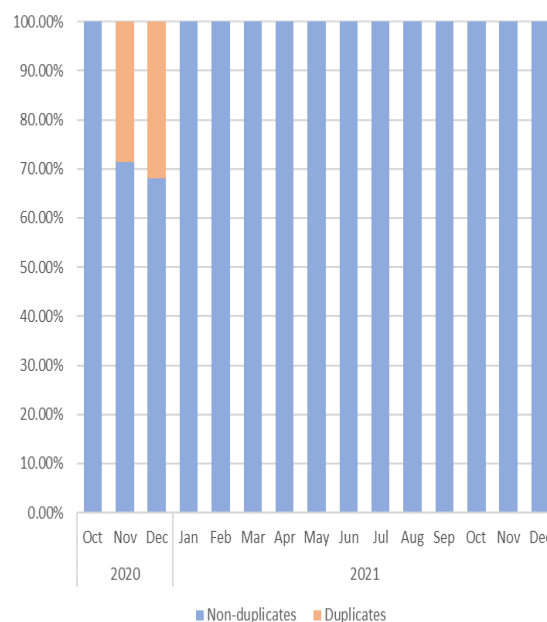

Note: Note: Duplicated transactions are identified based on three fields: Reporting counterparty ID, ID of the other counterparty and trade ID. In case collateral report, master agreement type is used as well. The analysis has been performed on open SFT transactions on a given reference date.
Source: Trade Repositories & ESMA calculations

7.4   Data integrity – Reconciliation

Reconciliation: Reconciliation is one of the key data quality processes performed by TRs. TRs provide information on the results of the process to all key stakeholders, i.e., reporting participants and NCAs/ESMA. To visualize the results, ESMA relied on the information provided in the reconciliation reports for outstanding SFT transaction. These reports contain, among others, detailed breakdown of reconciliation status for each open SFT.

Chart 44 shows the result of pairing[21] over time. In the 4th quarter of 2020, the pairing rate stood at about 45-50%. There has been a positive trend throughout 2021 and by the end of the year the pairing rate increased to 60%. Considering that the SFTR is a relatively young reporting regime,

---

[21] Pairing is performed on the basis of three fields which define uniqueness of an SFT transaction: Reporting counterparty ID, ID of the other counterparty and trade ID.

these are not entirely disappointing results (indeed EMIR pairing rate stands at around 60% as well after 8 years of reporting). Having said that, ESMA's intention is to continue to focus on pairing (and reconciliation more broadly) as part of its work with the NCAs (see Subsection 4.2 for more details).

Chart 44
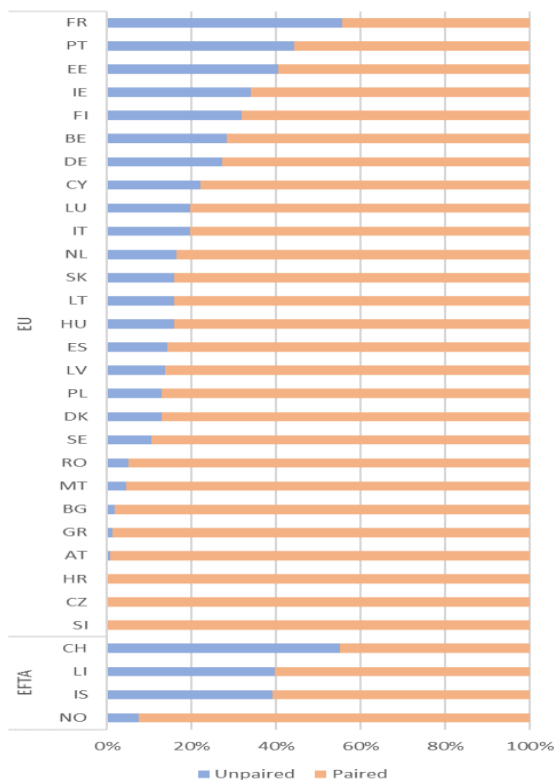SFTR data integrity – Reconciliation
Paired records
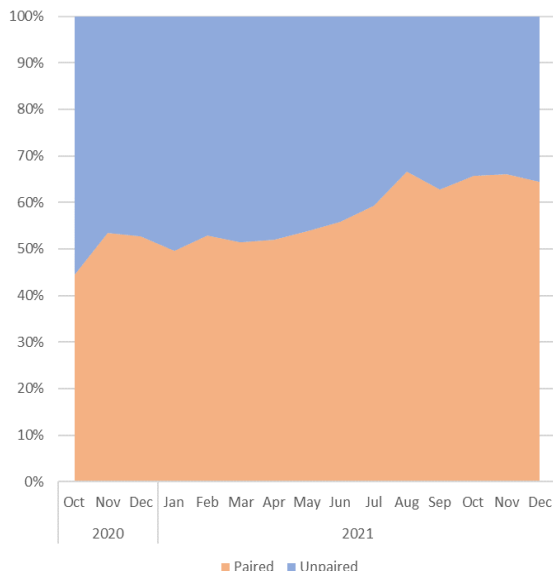
Note: Pairing is performed based on three fields: Reporting counterparty ID, ID of the other counterparty and trade ID. The results show pairing rate on open SFTs.
Source: Trade Repositories & ESMA calculations

Chart 45 shows the distribution of pairing rate across jurisdictions. Clearly, there is a substantial variation in the results as counterparties tend to display varying performance levels in different jurisdictions.

Chart 46 once a transaction is successfully paired, TRs proceed with an attempt to match the remaining fields[22] between the two reported sides of a transaction. TRs perform matching both on the loan as well as on the collateral information of each transaction.

As regards loan matching, through 4th quarter 2020 and until end 2021 there has been a positive trend as the matching rate increased from around 30% to nearly 50%.

Chart 45
SFTR data integrity – Reconciliation
Paired records by country of the reporting counterparty for the period Q4 2021

Note: Pairing rates by jurisdiction.
Source: Trade Repositories & ESMA calculations

Chart 46
SFTR data integrity – Reconciliation
Reconciled loan components
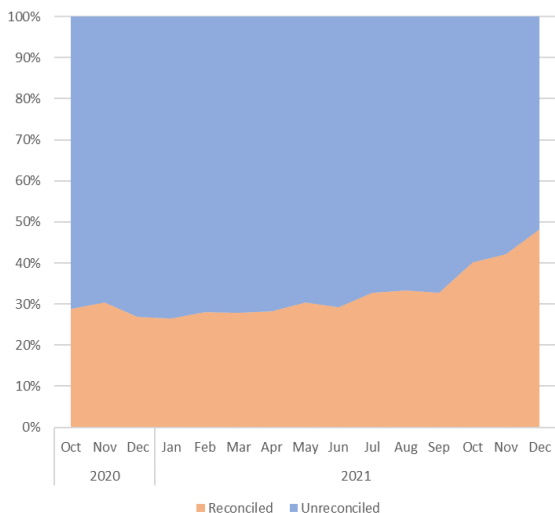
Note: Reconciliation rate on loan components of open SFTs.
Source: Trade Repositories & ESMA calculations

---

[22] The list of reconcilable fields that TRs use, including, where applicable, tolerances can be found in the RTS on data quality

Chart 47 shows loan reconciliation rate by jurisdiction. Once again, there is variation in the performance of the results across countries.

Chart 48 shows reconciliation of collateral components. Considering that collateral reconciliation is the most complex step in the entire process it is not surprising to see that successful reconciliation exhibits the lowest rates – only around 20% at the end of 2021. Importantly, like in the previous cases, there has been an increasing trend throughout the displayed period.

Chart 49 shows breakdown of collateral reconciliation rates by jurisdiction. Consistently with the findings in the previous charts, there is a substantial variation in reconciliation rates across different jurisdictions.

While SFTR reconciliation (namely pairing) has been the focus of NCAs during 2021 during the SFTR DQEF (see Subsection 4.2 SFTR DQEF) and some progress has been made, it is reconciliation still needs to be point of focus going forward.

Chart 47
SFTR data integrity – Reconciliation
Reconciled loan components by country of the reporting counterparty for the period Q4 2021
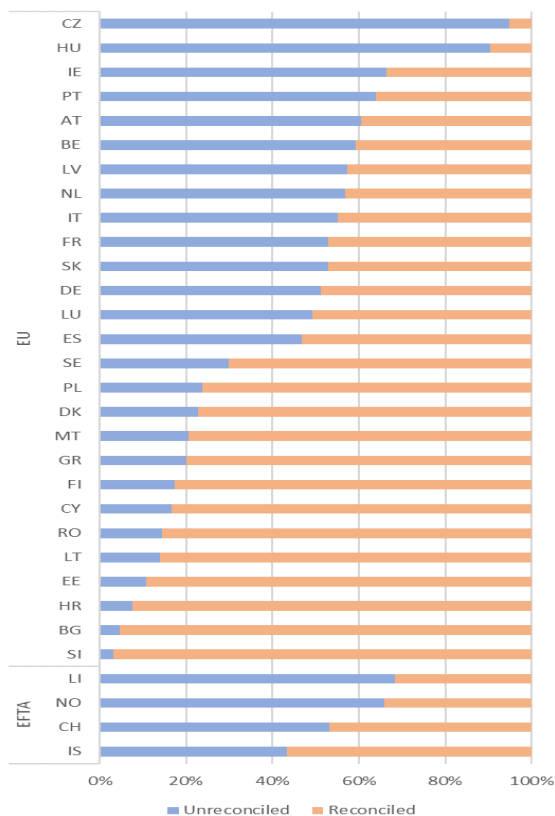
Note: Loan reconciliation by jurisdiction.
Source: Trade Repositories & ESMA calculations

Chart 48
SFTR data integrity – Reconciliation
Reconciled collateral components
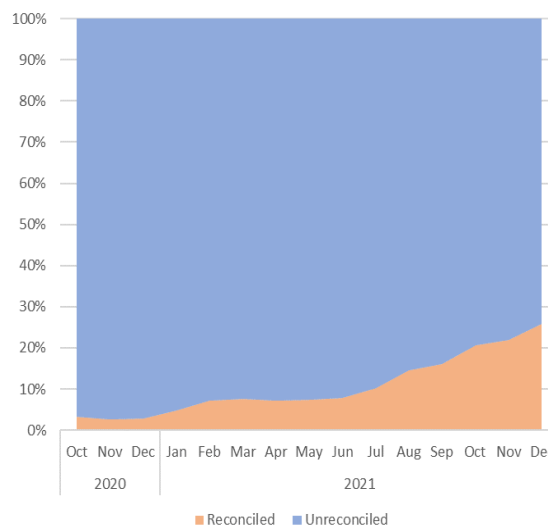
Note: Collateral component reconciliation of open SFTs.
Source: Trade Repositories & ESMA calculations

Chart 49
SFTR data integrity – Reconciliation
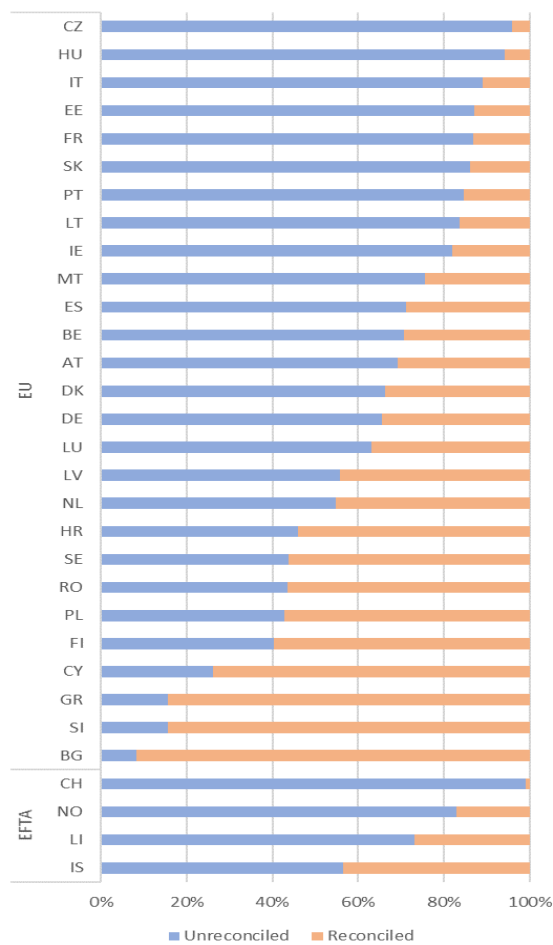**Reconciled collateral components by country of the reporting counterparty for the period Q4 2021**

Note: Collateral reconciliation by jurisdiction.
Source: Trade Repositories & ESMA calculations

A similar analysis to EMIR inter-TR reconciliation process discrepancies (see Subsection 6.4 Data integrity –) has been carried out under SFTR. **Chart 50** shows that there are significant divergences in the number of UTIs that a TR considers that it has paired against other TRs. Such discrepancies are an indication that the reconciliation process performed by TRs may not function appropriately and that information provided by TRs on the outcomes of the reconciliation process to the reporting participants and NCAs may not correctly reflect the correct state of reconciliation between two legs of any SFT.

In case of SFTR, the reconciliation is somewhat more complex than in EMIR ad TRs are required to reconcile both the Loan and Collateral component) which results in more evident discrepancies (see **Chart 51**).

Chart 50
SFTR data integrity – Reconciliation
Discrepancies in number of UTIs paired versus Other TRs
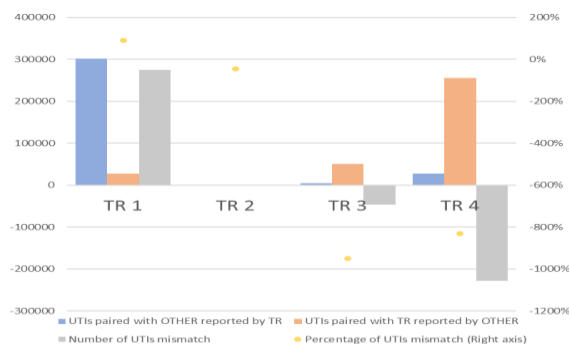
Note: The charts are based on information provided monthly by TRs to ESMA in the context of periodic information (Item 36 – Reconciliation Statistics). The figures in the charts refer to outstanding trades and are computed as the monthly average for the reference period June 2021 – December 2021
Source: Trade Repositories & ESMA calculations

Chart 51
SFTR data integrity – Reconciliation
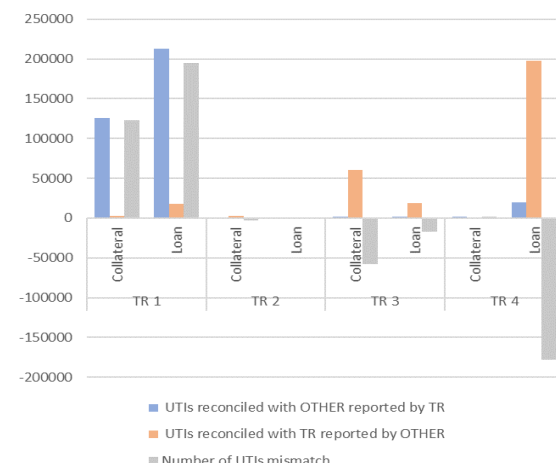**Discrepancies in number of UTIs reconciled versus Other TRs**

Note: The charts are based on information provided monthly by TRs to ESMA in the context of periodic information (Item 36 – Reconciliation Statistics). The figures in the charts refer to outstanding trades and are computed as the monthly average for the reference period June 2021 – December 2021
Source: Trade Repositories & ESMA calculations

8. Methodological Annex

EMIR reporting trends and data quality metrics

Data reporting – key trends: ESMA monitors key trends in the reporting volumes by performing a count of all daily submissions and open derivatives for a given reference date by action type, asset class and contract type.

Data completeness, timeliness, and availability – execution vs. reporting timestamps: ESMA measures the timeliness of reporting by counterparties by applying the following four assumptions: (1) derivatives concluded at time T and reported at T+1 at the latest, are considered "On Time", (2) derivatives concluded at time T and reported after T+1 are considered "Late Reporting", (3) derivatives concluded at time T and reported before T are considered "Early Reporting", and (4) derivatives concluded on a Friday or Saturday and reported on the consecutive Monday are subject to a "weekend effect" which is accounted for in the calculation and correctly classified as "on time". Submissions with action type N (New) or P (Position component) reported at transaction level (Level = T) are used for this analysis. For each submission in the sample, we compute the difference between the "Reporting Timestamp" and the "Execution Timestamp" expressed in days. Public-, national- and bank holidays (i.e., "calendar effect") are not accounted for in the calculation, nor is the conversion of Coordinated Universal Time (UTC) to local time made. These approximations simplify and speeds up the calculation but could give rise to some degree of inaccuracy (i.e., records wrongly classified as "Late Reporting" due to UTC vs local time differences, or due to calendar effect) impacting the overall results.

Data completeness, timeliness, and availability – non-reporting: ESMA estimates the number of non-reported derivatives by counting a number of open derivatives reported between a counterparty pair (i.e., EMIR fields 'Reporting counterparty ID' and 'ID of the other counterparty') in both directions (i.e., CP1 vs. CP2 and CP2 vs. CP1) and taking a difference in those instances where open derivatives were reported only in one direction. Non-EEA counterparties and open derivatives with non-LEIs in ID of the other counterparty are excluded from the calculation. Member State of non-reporting is identified by the country of LEI in 'ID of the other counterparty' using the GLEIF reference data.

Data completeness, timeliness, and availability – non-reporting of valuations by counterparties: ESMA measures non-reporting of valuations by counterparties by analysing all open derivatives with action type = N (new) or V (valuation) and clearing threshold = Y. The difference in number of days is computed between the reference date of the TSR and the valuation timestamp of a record. Each record is grouped into buckets (0-15 days, 16-30 days, 30-365 days, and >365 days) to measure the frequency of valuation updates.

Data accuracy – adherence to format and content – revalidation and rejection rates: ESMA performs a data revalidation process on the daily submissions to detect data quality issues linked to the validation process of TRs. The analysis uses a randomly selected data sample (~15%) extracted from one daily submission report per month, per TR. Each data point is checked against the current ESMA validation rules.

Rejection statistics produced by TRs are aggregated by ESMA and used to monitor how many reports are being rejected by TRs due to misreporting by CPs.

Data integrity – reconciliation: ESMA performs reconciliation process on open derivatives on a given reference data by replicating the process applied by the TRs. Firstly, non-EEA open derivatives are excluded from reconciliation.

Pairing is performed by finding second leg of each derivative by using a unique key (i.e., EMIR fields 'Reporting counterparty ID', 'ID of the other counterparty', and 'Trade ID'). The second leg of a derivative

is found by looking CP1-CP2-TradeID vs CP2-CP1-TradeID. Both sides of each derivative are counted towards the aggregate values.

The difference in the number of reported derivatives is calculated by counting open derivatives reported between a counterparty pair (i.e., 'Reporting counterparty ID' and 'ID of the other counterparty') in both directions (i.e., CP1 vs. CP2 and CP2 vs. CP1) and taking a difference.

The analysis on inter-TR reconciliation issues is based on data submitted by TRs monthly according to the Guidelines on Periodic Information[23]. The scope of this analysis includes only information related to open derivatives. The reference period of the analysis provided in this report is June'2021-December'2021. The relevant values that are considered are the following:

- The number of UTIs each TR reports to have paired/reconciled with all other TRs (A)
- The number of UTIs all other TRs report to have paired/reconciled with that specific TR (B)
- The difference between the number of UTIs reported by the TR and the other TRs (A – B)
- The percentage difference [(A – B)/A]

Those values have been calculated for each month of the reference period and have been averaged out.

SFTR reporting trends and data quality metrics

Data reporting – key trends: ESMA monitors key trends in the reporting volumes by performing a count of all daily submissions and open SFTs for a given reference date by action type, type of SFT and sector of the reporting counterparties.

Data accuracy – adherence to format and content: Total number of accepted and rejected SFTs is computed from dedicated TR regulatory reports containing aggregated as well as SFTs level information on rejected and accepted SFTs submitted to TRs by counterparties.

Duplicated records are identified using counterparty ID, ID of the other counterparty and trade ID. In the cases of collateral components master agreement type is used as well. The analysis is performed on the trade state report containing all open SFTs on a given date.

Data completeness, timeliness, and availability – event date vs. reporting timestamps: ESMA measures the timeliness of reporting by counterparties by applying the following four assumptions: (1) new SFTs concluded at time T and reported at T+1 at the latest, are considered "On Time", (2) new SFTs concluded at time T and reported after T+1 are considered "Late Reporting", (3) new SFTs concluded at time T and reported before T are considered "Early Reporting", and (4) SFTs concluded on a Friday or Saturday and reported on the consecutive Monday are subject to a "weekend effect" which is accounted for in the calculation and correctly classified as "on time". For each submission in the sample, we compute the difference between the "Reporting Timestamp" and the "Event Date" expressed in days. Public-, national- and bank holidays (i.e., "calendar effect") are not accounted for in the calculation, nor is the conversion of Coordinated Universal Time (UTC) to local time made. These approximations simplify and speeds up the calculation but could give rise to some degree of inaccuracy (i.e., records wrongly classified as "Late Reporting" due to UTC vs local time differences, or due to calendar effect) impacting the overall results.

Data integrity – reconciliation: Pairing and reconciliation flags are calculated using dedicated fields with respective reconciliation statuses in the trade state report. Only records where both counterparties

---

[23] The item analysed is "Item 36a – Reconciliation Statistics". For more information, please consult the "Guidelines on periodic information" and the Data Reporting templates on the ESMA web page.

have a reporting obligation are considered. Information on the jurisdiction of the two counterparties to the transaction (reporting counterparty ID and ID of the other counterparty) is obtained from GLEIF.

The analysis on inter-TR reconciliation issues is based on data submitted by TRs on a monthly basis according to the Guidelines on Periodic Information[24]. The scope of this analysis includes only information related to outstanding SFTs. The reference period of the analysis provided in this report is June'2021-December'2021. The relevant values that are considered are the following:

- The number of UTIs each TR reports to have paired/reconciled with all other TRs (A)
- The number of UTIs all other TRs report to have paired/reconciled with that specific TR (B)
- The difference between the number of UTIs reported by the TR and the other TRs (A – B)
- The percentage difference [(A – B)/A]

Those values have been calculated for each month of the reference period and have been averaged out.

---

[24] The item analysed is "Item 36b – Reconciliation Statistics". For more information, please consult the "Guidelines on periodic information" and the Data Reporting templates on the ESMA web page.

9.List of abbreviations

AFM	Autoriteit Financiële Markten
AMF	Autorité des Marchés Financiers
BaFin	Bundesanstalt für Finanzdienstleistungsaufsicht
BIS	Bank for International Settlements
CBoI	Central Bank of Ireland
CCP	Central Counterparty
CNB	Česká národní banka
CSSF	Commission de Surveillance du Secteur Financier
CD	Credit Derivatives
CDS	Credit Default Swap
CFD	Contract for Difference
CM	Clearing Member
CME	CME Trade Repository Ltd. (CME TR)
CO	Commodity Derivatives
CSD	Central Securities Depositories
CP	Counterparty
CU	Currency Derivatives
DDRIE	DTCC Data Repository (Ireland) Plc
DDRL	DTCC Derivatives Repository Plc
DQAP	Data Quality Action Plan
DQEF	Data Quality Engagement Framework
DQR	Data Quality Review
EEA	European Economic Area
EMIR	European Markets Infrastructure Regulation
EQ	Equity Derivatives
ESMA	European Securities and Markets Authority
ETD	Exchange Traded Derivatives
FC	Financial Counterparty
FSB	Financial Stability Board
GLEIF	Global Legal Entity Identifier Foundation
HHI	Herfindahl-Hirschman Index
ICE	ICE Trade Vault Europe Ltd. (ICE TVEL)
IORP	Institutions for Occupational Retirement Provision
IRD	Interest Rate Derivatives
IRS	Interest Rate Swaps
ISDA	International Swaps and Derivatives Association
KDPW	Krajowy Depozyt Papierów Wartosciowych S.A.
LEI	Legal Entity Identifier
MIC	Market Identifier Code
MiFIR	Markets in Financial Instruments Regulation
MFSA	Malta Financial Services Authority
NCA	National Competent Authority
NFC	Non-Financial Counterparty
OTC	Over-the-Counter
REGIS	REGIS-TR
REPO	Repurchase Agreement
RTS	Regulatory Technical Standard
SFT	Securities Financing Transaction
SFTR	Securities Financing Transactions Regulation
SLEB	Securities and Commodities Lending or Borrowing
TAR	Trade Activity Report
TR	Trade Repository
TSR	Trade State Report
UCITS	Undertakings for Collective Investment in Transferable Securities
UNAVISTA B.V	UnaVista TRADEcho B.V. (The Netherlands)
UNAVISTA LTD	Unavista limited
UTC	Coordinated Universal Time
UTI	Unique Trade Identifier

| XML | Extensible Markup Language |

Countries abbreviated according to ISO standards
Currencies abbreviated according to ISO standards