

SPECIAL REPORT

Increasing efficiency in sell-side reference data management

Building a reliable and efficient
data workflow fit for a complex
world

Commissioned by:

Over the past decade, the volume of reference data required to be processed for derivatives trading has increased exponentially, with higher volumes and greater reporting requirements. Simultaneously, the reputational and financial risks of inaccurate data have risen close to the top of compliance department concerns. The challenge for the sell-side is to overcome the traditional challenges with reference data, both internally and externally, to create an efficient and reliable workflow that can reduce operational costs and ensure compliance with increasingly onerous reporting mandates.

Since the introduction of MIFID I, the United Kingdom's (UK) Financial Conduct Authority has issued approximately £100m in fines for breaches of reporting requirements. In 2019, two major international banks were fined a total in excess of £50m for reporting failures relating to MIFID I.

MIFID II brought even more extensive and stringent reporting requirements and, with the honeymoon period for MIFID II reporting now over, the UK regulator and its EU peer ESMA are expected to begin enforcing MIFID II reporting rules more strictly in 2021.

The sheer size of the recent FCA fines and associated reputational damage has forced reporting further up the agenda of focus across the sell-side. However, the increased reporting requirements have also increased the challenge and complexity of managing data in a holistic and secure way and exposed fragmentation and the diversity of taxonomies.

To understand the current state of how reference data is managed and processed across the sell-side, the operational challenges posed to firms by the status quo and to understand what increases efficiency and reliability in data workflows, Acuiti was commissioned by FOW to conduct a study of sell-side reference data operations. The results of that study are set out in this special report.

The problem with reference data

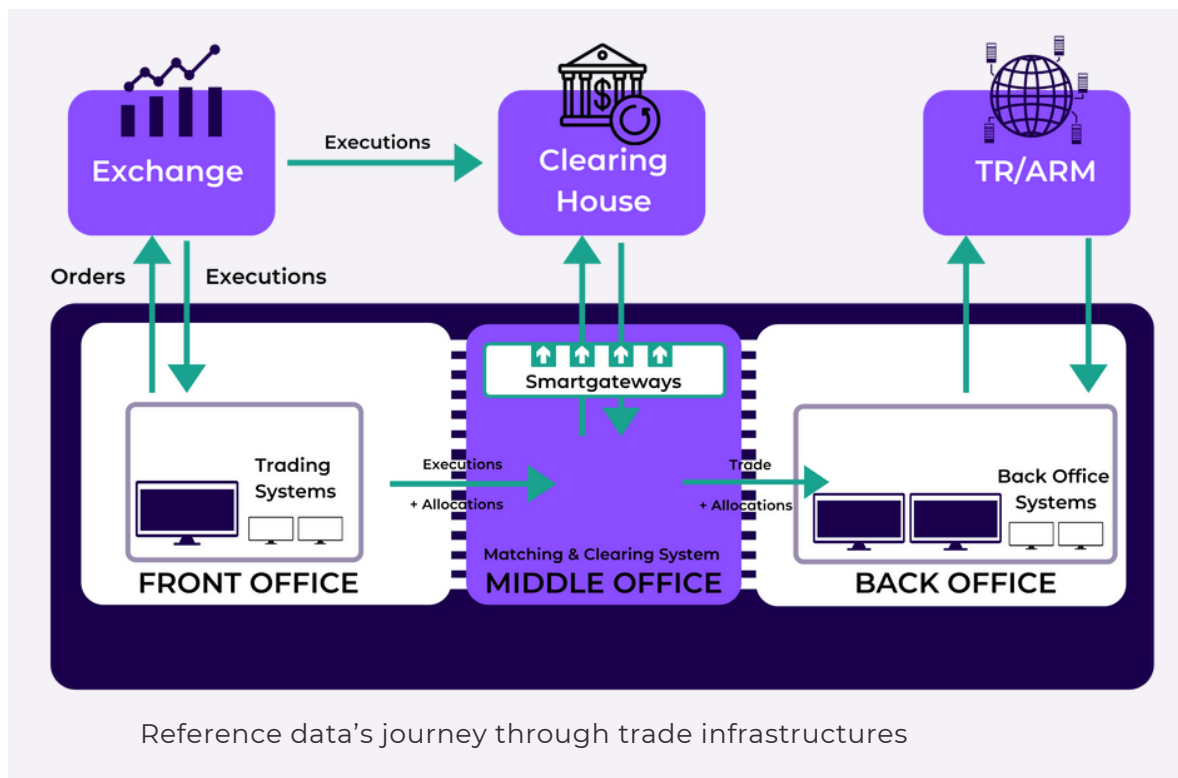
Reference data represent the language of capital markets, making sense of every trade and enabling them to be decoded and interpreted across the trading cycle. However, as language has evolved in relative isolation across thousands of different tongues and hundreds of thousands of dialects, so too has the language of reference data been developed in silos.

Reference data, used in this report to define any static identification data related to exchange-traded or over-the-counter (OTC) derivatives, cover a vast range of underlying instruments, entities and identifiers. The multitude of data points required to identify different trading venues, counterparties, asset classes and the complexity of the products in the derivatives market creates a vast pool of reference data, which even if homogenous would be tricky to manage efficiently. The challenge for sell-side firms to collect and manage reference data internally and across the market has been added to by a swathe of post-financial crisis reporting requirements. EMIR and MIFID II in Europe, the CFTC and 871(m) rules in the US, the Large Open Position Reporting regulation in Canada and a host of local requirements across Asia have added to the administrative burden (and the threat of sanction) for the sell-side.

For the sell-side, sourcing, processing and managing reference data for derivatives is fiendishly complicated. Internally, each instrument has to be referenced across various systems that may each have their own taxonomy and codes and then sent across the market where the fragmentation of data protocols is even more apparent.

The increased requirements of regulatory reporting pose a further challenge because each instrument must be assigned an underlying instrument code (ISIN) based on a different construction methodology to that commonly used across the market. While there have been numerous attempts at standardising identifiers, such as the initiative by CUSIP Global Services, use of the protocols is by no means universal.

Myriad operational issues from trade breaks to risk management failures are caused by this lack of harmonisation and, despite industry-wide initiatives to standardise the language of reference data and millions of dollars spent by sell-side firms to create single taxonomies, the fragmentation of reference data remains one of the greatest inefficiencies in the derivatives markets today.

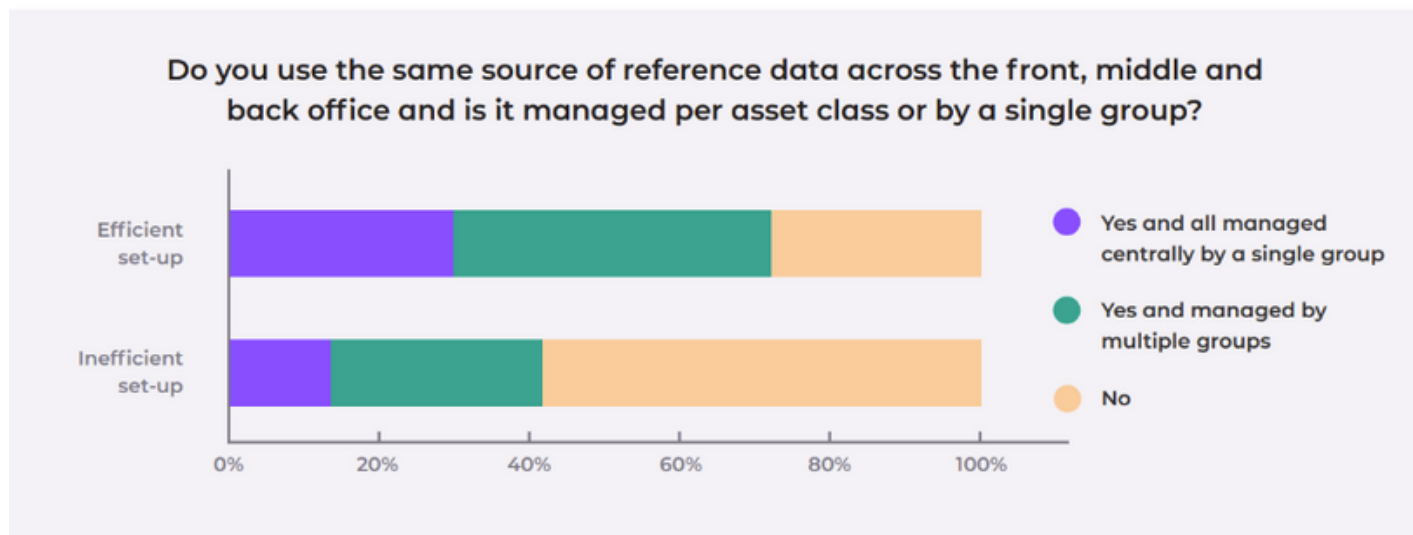


The squeezed middle

To understand the challenges that firms face when managing reference data, Acuiti conducted a survey of senior executives who oversee data operations across the sell-side. Acuiti split respondents into three categories: tier 1 banks, defined as the largest banks by derivatives volumes, tier 2 & 3 banks, and a pool of smaller banks, non-bank futures commission merchants (FCMs) and brokers. The survey found broad differences across the market in the efficiency of their reference data set-ups.

Tier 1 banks generally reported relatively high levels of efficiency, with only 14% of respondents stating that their current operations were inefficient and 19% stating they were very efficient. Smaller firms also reported general satisfaction with current set-ups, with 71% reporting that their management of reference data was either very efficient (38%) or quite efficient (33%).

However, this is in marked contrast to the tier 2 & 3 banks, of which 53% reported that their current set-up was inefficient. Accordingly, tier 2 & 3 banks represent the squeezed middle in the case of reference data, facing similar volume and regulatory challenges as their larger counterparts, along with generally being in the market for more complex instruments, but having not yet gone through the investment required to fully address inefficiencies in their data operations.



The fact that tier 1 banks view their reference data processes as more efficient does not mean they have overcome the challenges faced by the rest of the market. Instead, they have generally found ways of living with the major challenges and inefficiencies, such as investing in teams of people, often offshore, to process and address data - a significant investment in terms of time and money but one that goes some way to solving the issues. However, as this report will argue, there are other, more immediate, ways of addressing data inefficiency.

Challenges firms face

To better understand where the major pinch points were in terms of the collecting and processing of reference data across the sell-side, Acuiti asked respondents about the major challenges they faced.

Again, the survey found significant discrepancies in terms of the challenges that different categories of firms face. While data feeds across internal systems was the greatest challenge for tier 2 & 3 banks and smaller firms, access to external data posed the greatest challenge for tier 1 banks.

Tier 1 banks were also significantly more likely to identify a challenge with data coverage, reflecting the larger scope of their exchange memberships, trading activity, client services and other operations, while aggregation of data across the organisation was a particular challenge for tier 2 & 3 firms. Smaller firms were less concerned with data coverage but were challenged by the cost of data and access to external data.

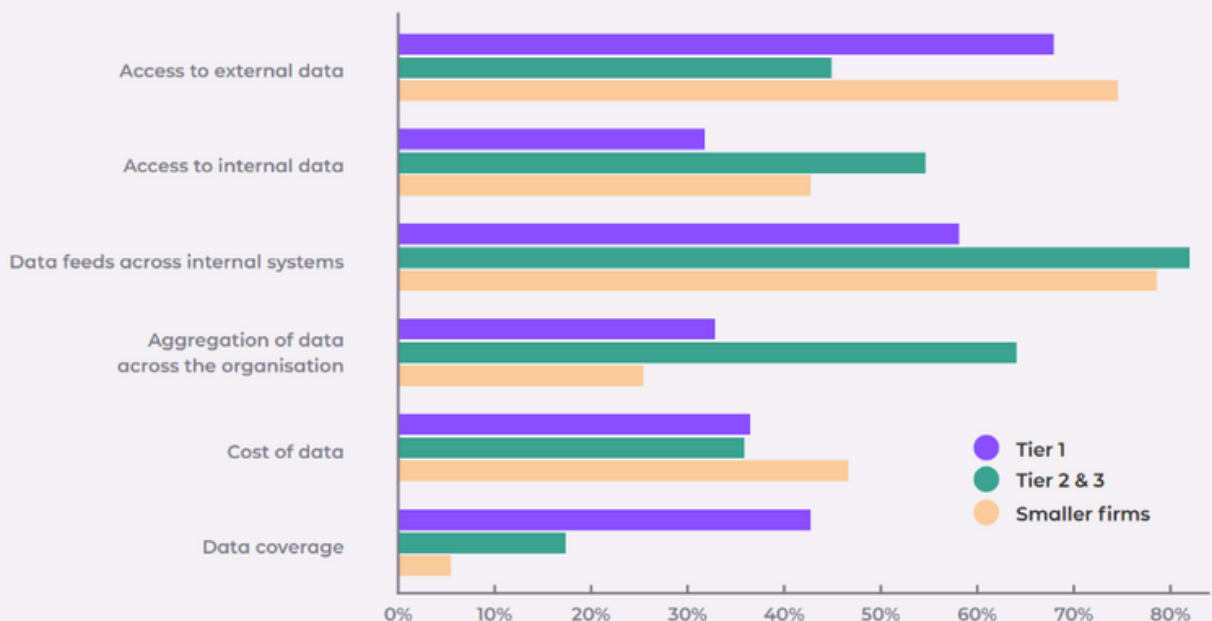
Overall, the evaluation of how much of a challenge is presented by the cost of data depended strongly on the job function of the respondent, with 60% of those on the business side of the firm citing cost as a challenge compared to 35% of those in compliance and 34% of those in operations roles.

Executives overseeing operations, technology and compliance were more likely to identify challenges in accessing external data and data feeds across internal systems.



“The core proposition of tier 1 banks as sell-side providers, is to be a one stop shop for most buy-side firms. They need to offer significant breadth and depth of market coverage to achieve this, which they do by connections to multiple exchanges, index, market and reference data providers. Economies of scale are a prerequisite by winning and retaining minimum threshold business to offset the significant investment in data sources. For tier 2 and 3 firms, the limited price competition in accessing the industry’s key approved data sources, is a major financial hurdle, as these input costs present a more material challenge to be covered by revenues from much lower scale trading. Lower revenues and smaller balance sheets also reduce the scope to invest in sophisticated data management systems, particularly around aggregation to provide a consolidated view across a firm.”

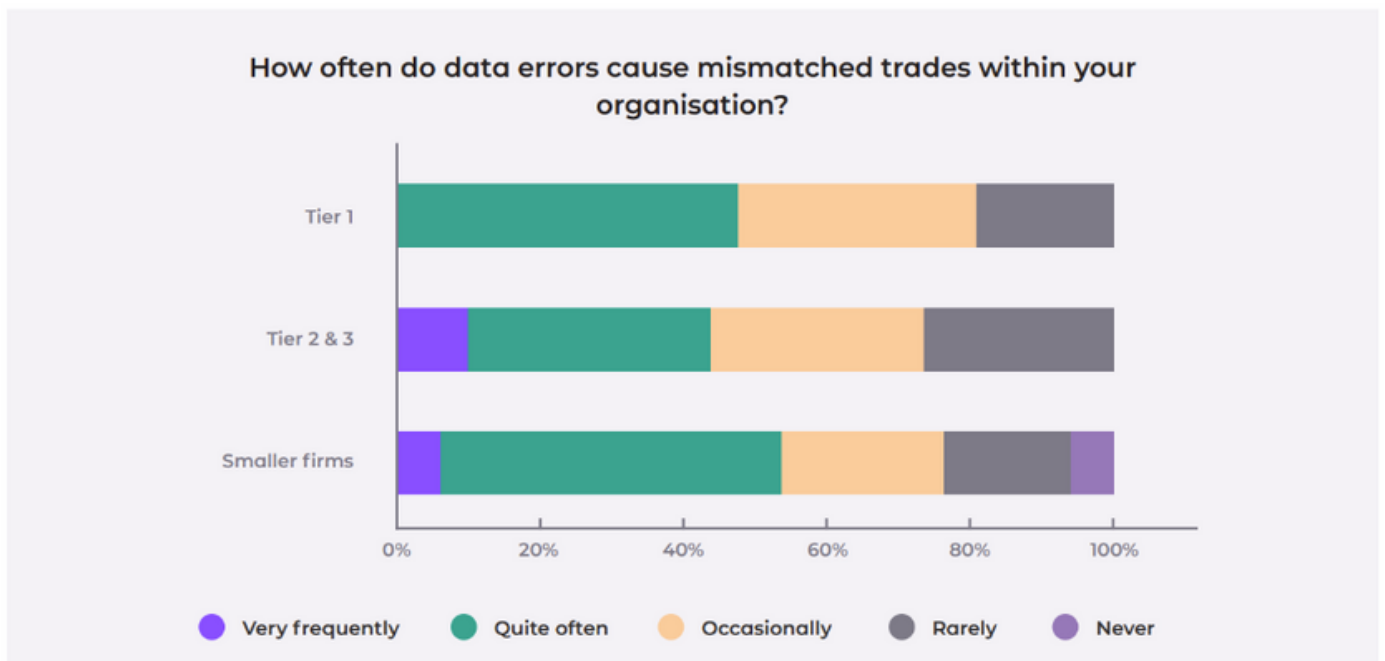
What challenges do you face in the collecting and processing of reference data?



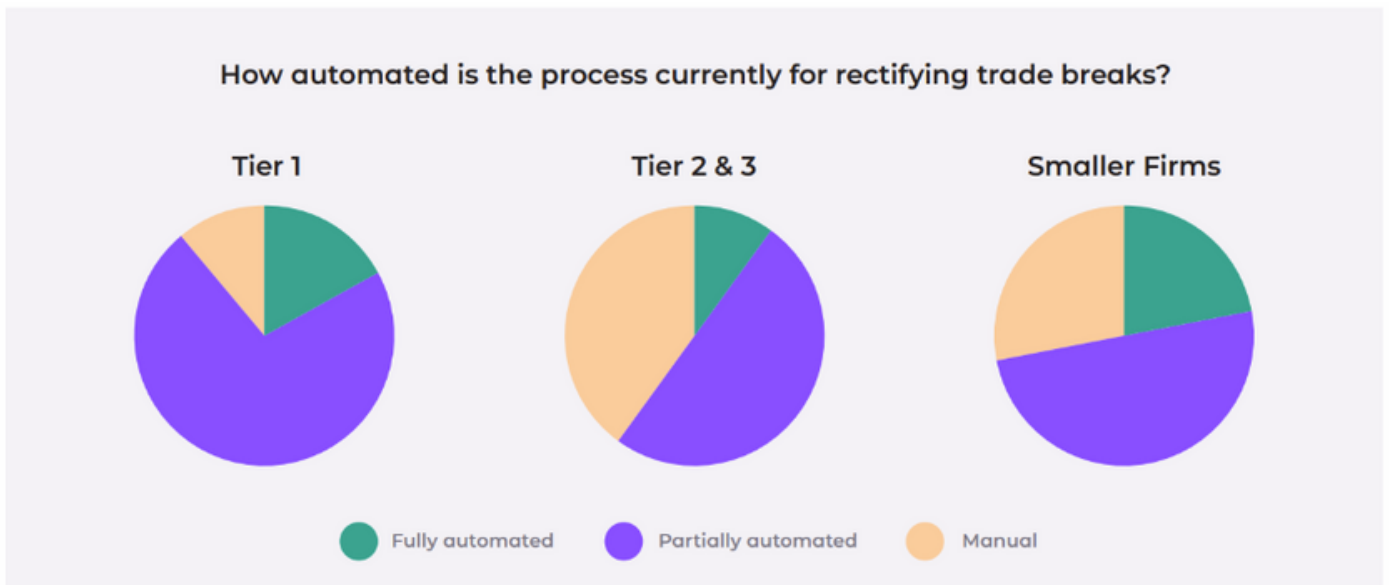
How poor reference data impacts operations

Inaccurate or fragmented reference data can cause a whole range of issues for the sell-side from trade breaks, to reporting failures, to the build-up of unquantified risks. The challenges posed by inaccurate reference data manifest both within an organisation and across the industry.

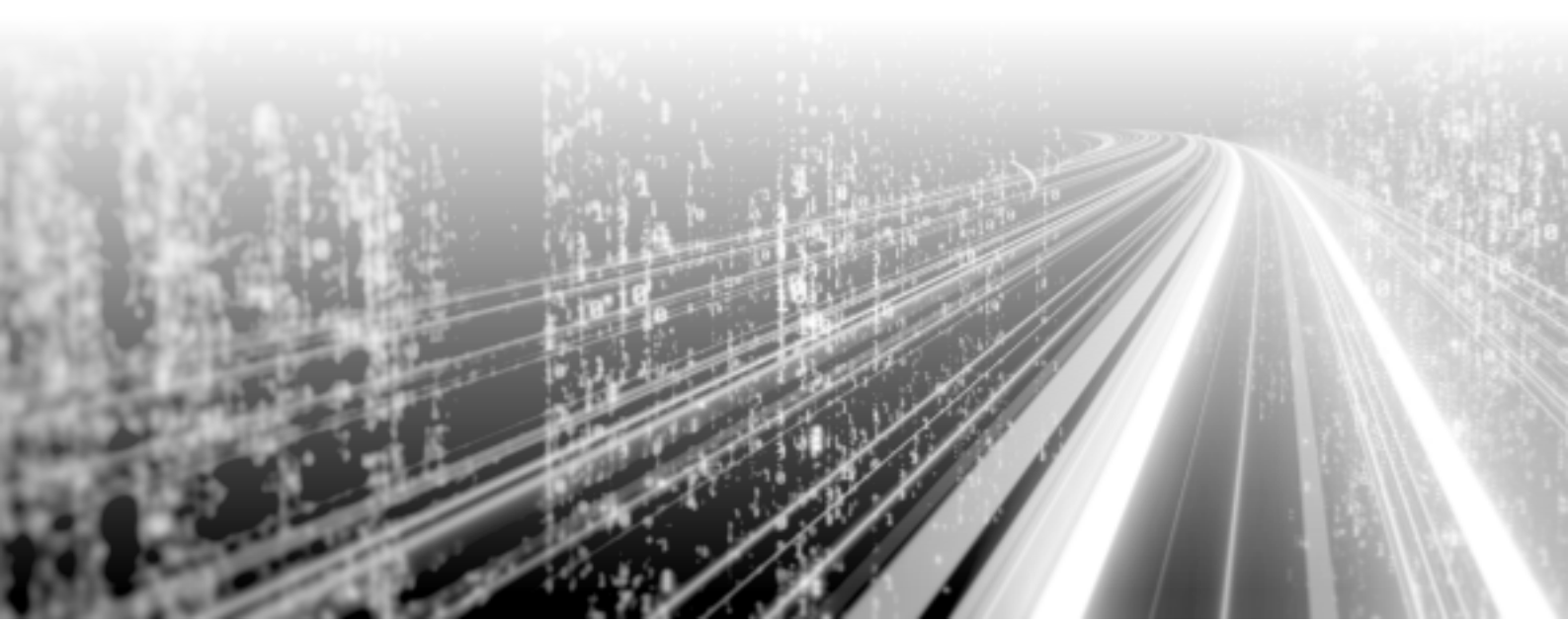
The Acuiti survey found that 47% of smaller firms, 36% of tier 2 & 3 banks and 47% of tier 1 banks experienced mismatched trades caused by data errors 'quite often', with 6% of smaller firms and 9% of tier 2 & 3s reporting that data errors 'very frequently' caused mismatched trades.

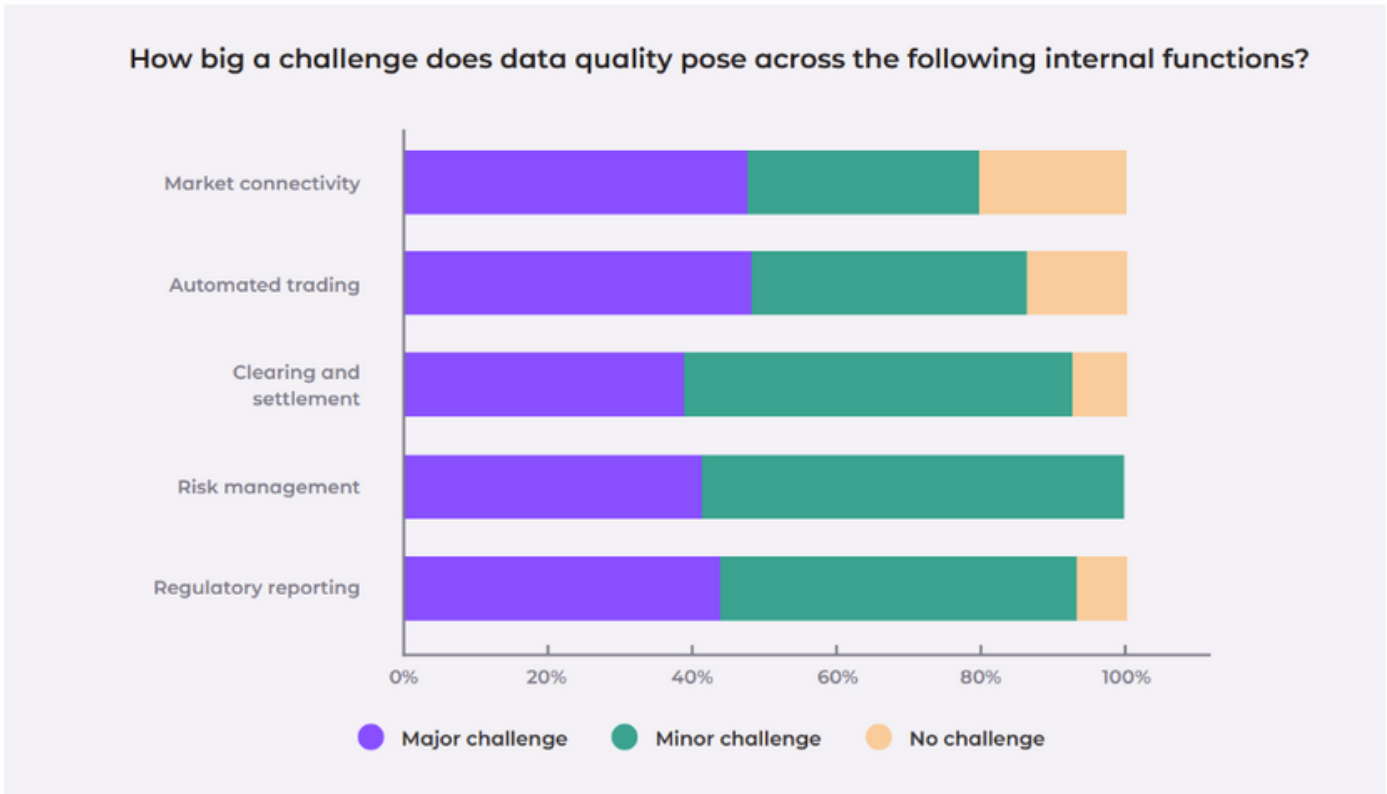


The frequency of trade breaks is one of the major causes of inefficiencies in post-trade operations. While the issue has existed for many years, only a minority of firms have taken steps to fully address the problem. Of those polled by Acuiti, 22% of smaller firms, 17% of tier 1s and 10% of tier 2 and 3s had fully automated the process for rectifying trade breaks with 28%, 11% and 40% respectively still relying entirely on manual processes to fix out trades.



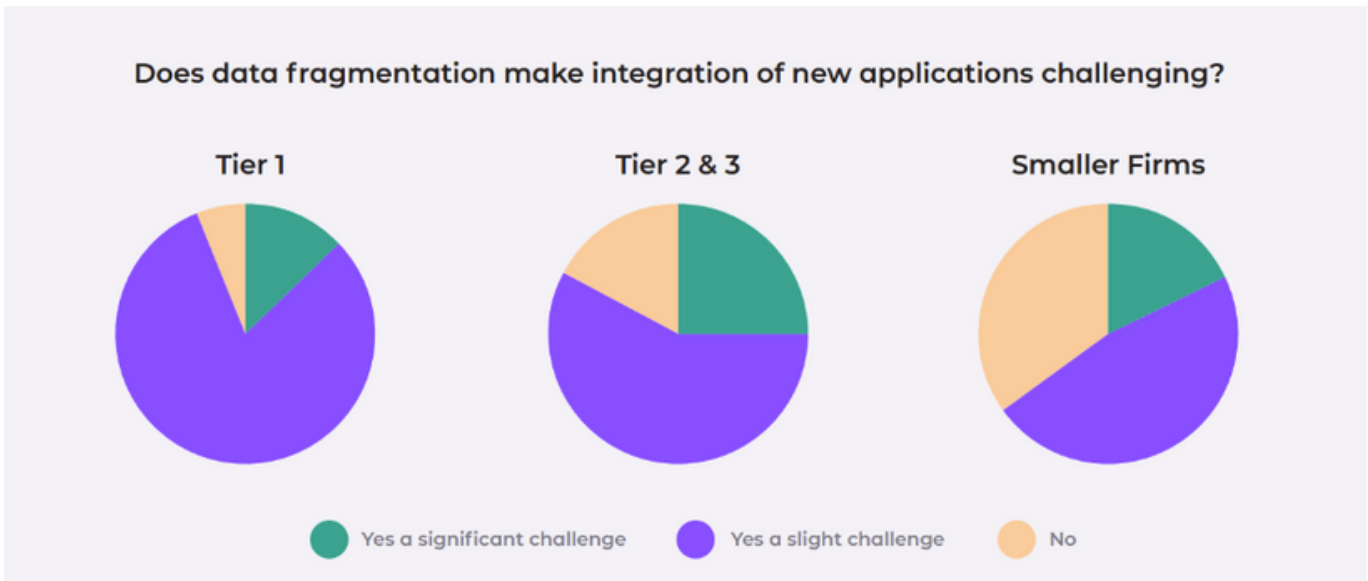
While the fragmentation of reference data is most strongly felt in trade breaks, poor data quality poses challenges across numerous other sell-side operations. Over 90% of respondents to the survey reported that poor data quality caused issues in clearing and settlement, risk management and regulatory report with 80% or more citing challenges in automated trading and market connectivity emanating from inaccurate data. Inaccurate data also exposes a firm's risk to fines for reporting errors.





Fragmented reference data also creates additional challenges for innovation and technology upgrades. As organisations tend to rely on the default taxonomy of their dominant technology stacks, frequently that of the post-trade platform, integrating new applications that use a different taxonomy requires a layer of translation to be built to integrate the new technology.

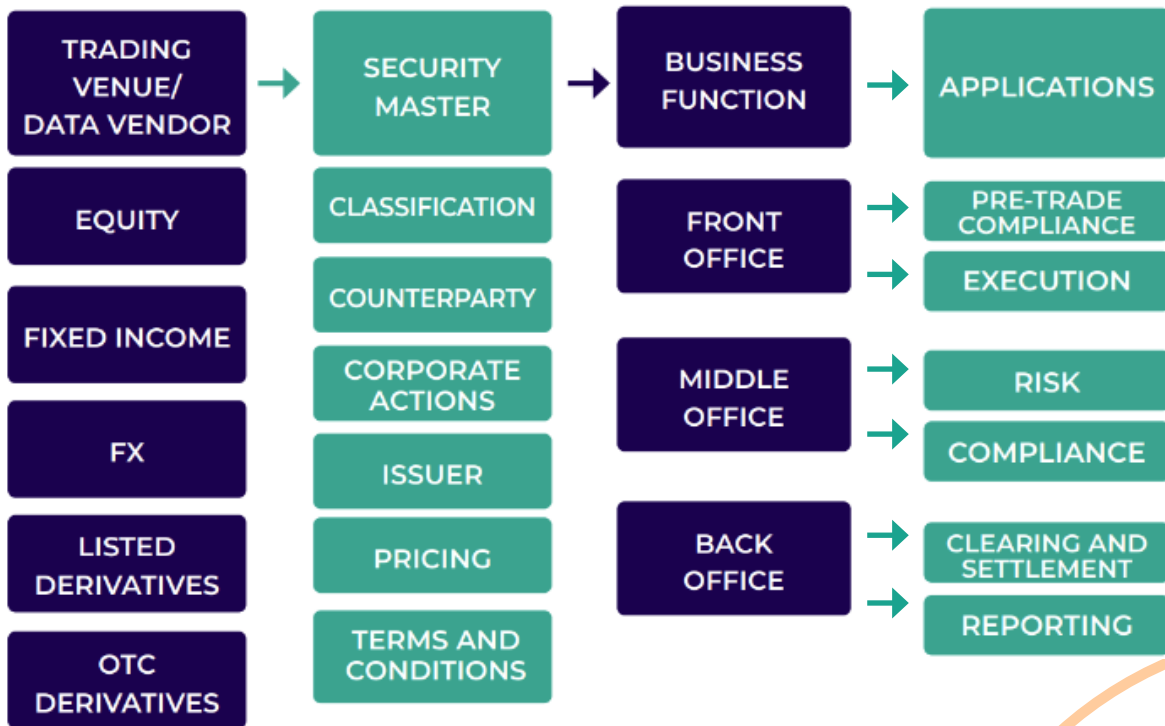
This was a particular issue for tier 1 banks with 94% of respondents saying that data fragmentation made integration of new applications challenging.



How are firms currently managing reference data?

Data management has evolved in most organisations rather than having been created *ex nihilo*. This means that reference data is processed and managed across legacy technology platforms and silos.

Simultaneously, mergers and acquisitions have brought in different data taxonomies to a business, and data harmonisation initiatives in an organisation have often been conducted within silos addressing the fragmentation only for specific asset classes or trading desks.

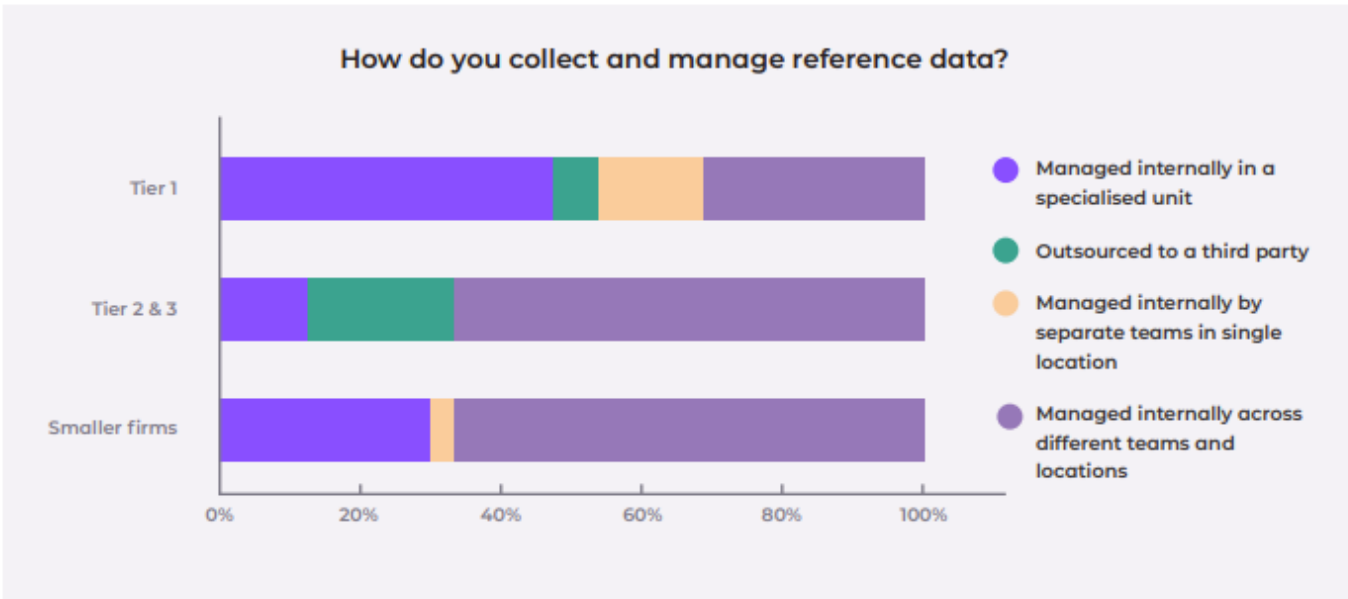


Uses and applications of reference data in the sell-side trade cycle

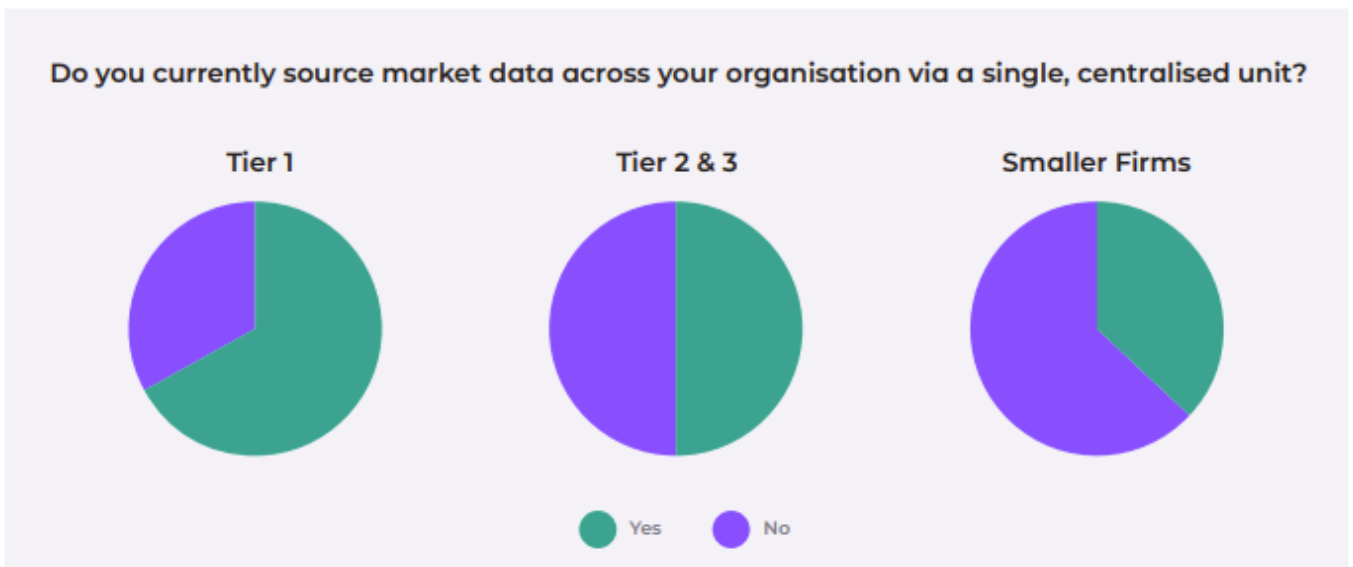
Over the past decade, a debate has raged across the sell-side over the benefits of the centralisation of reference data management in a single unit to mitigate the fragmentation.

Creating a single, central data repository has the advantage of providing a unified source of data across the organisation and a single point of management. However, critics of this model point to its inflexibility and erroneous ‘one-size-fits-all’ approach, arguing instead that data should be handled by local teams.

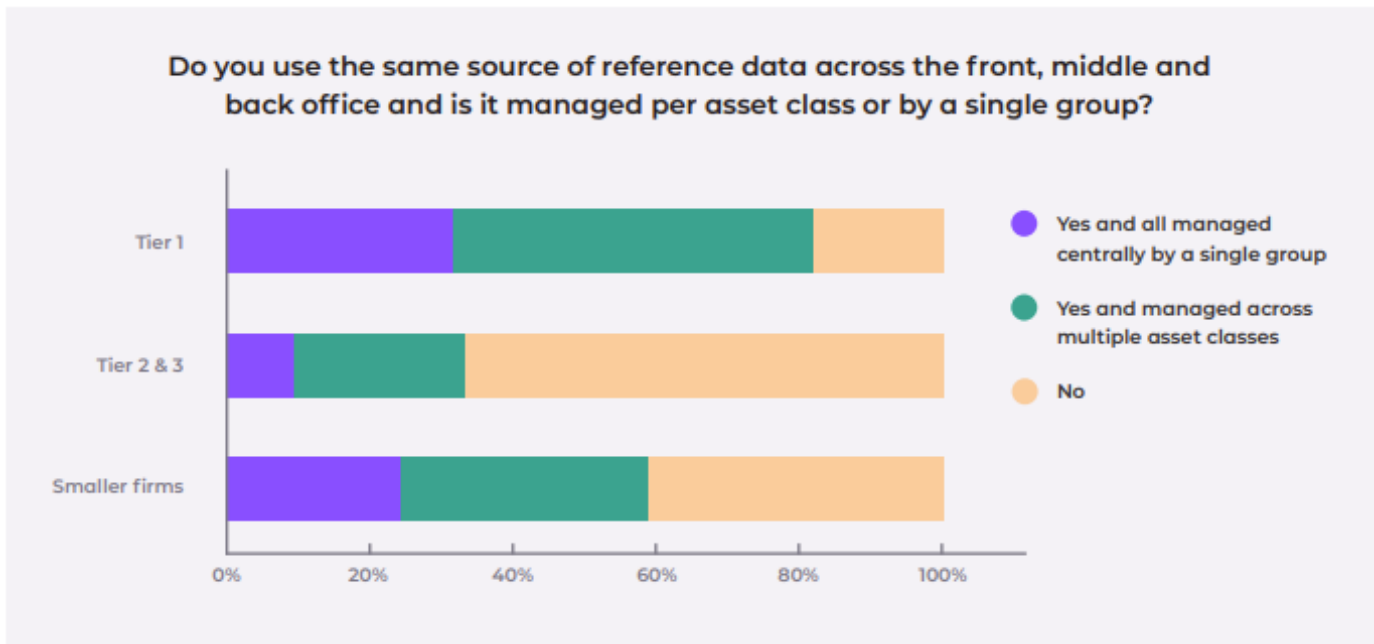
There are three core elements to this centralisation: managing the data, sourcing the data and using the same reference data across the organisation. The Acuiti survey found that 30% of smaller firms, 14% of tier 2 & 3 banks and 50% of tier 1s managed reference data in a specialised unit, while 65% of smaller firms, 30% of tier 1s and 64% of tier 2 & 3s managed reference data internally across different teams and locations.



When it comes to sourcing market data, 67% of tier 1s do so via a single centralised unit compared with 50% of tier 2 & 3 banks and 37% of smaller firms.



The Acuiti survey found that 59% of smaller firms, 81% of tier 1s and 33% of tier 2 & 3 banks used the same source of reference data across the front, middle and back office of an organisation. Moreover, 24% of smaller firms, 31% of tier 1s and 8% of tier 2 & 3 banks managed all reference data in a single group as opposed to per asset class.



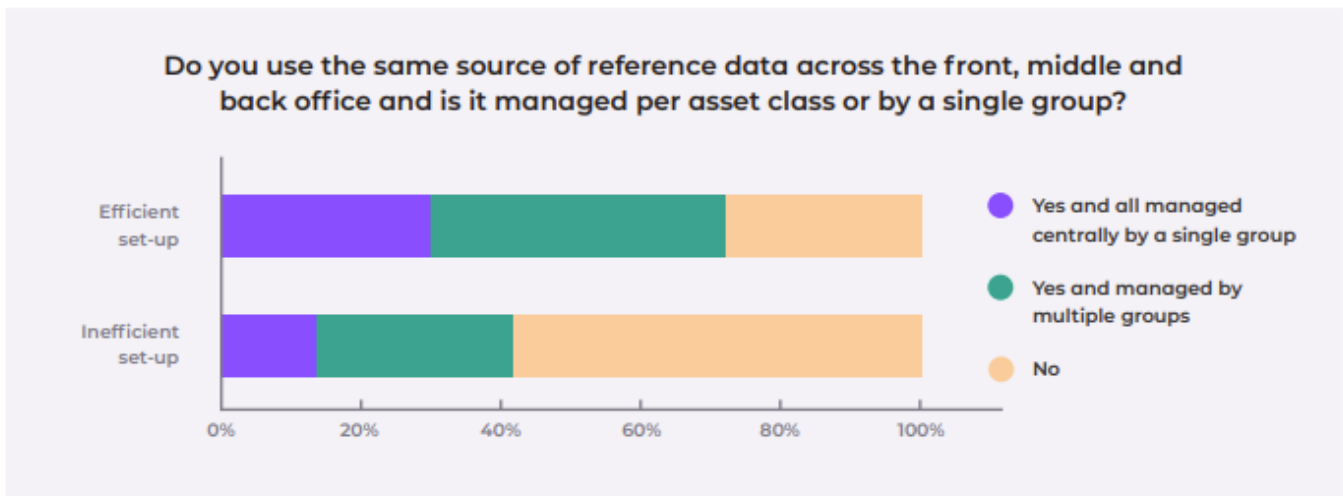
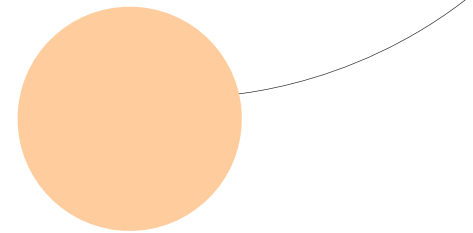


“Even a small to medium-sized investment firm has a number of system components to their trading operations and it is symbol translation that provides the persistent linking of these components, in the form of data passporting, which enables efficient system inter-operability and solution delivery. Without the presence of strong and consistent symbol management, trade breaks are not only more likely but costly, in both financial and reputational terms and so market and reference data procurement regimes must include symbology related quality thresholds in vendor selection criteria.”

What is the most efficient set up?

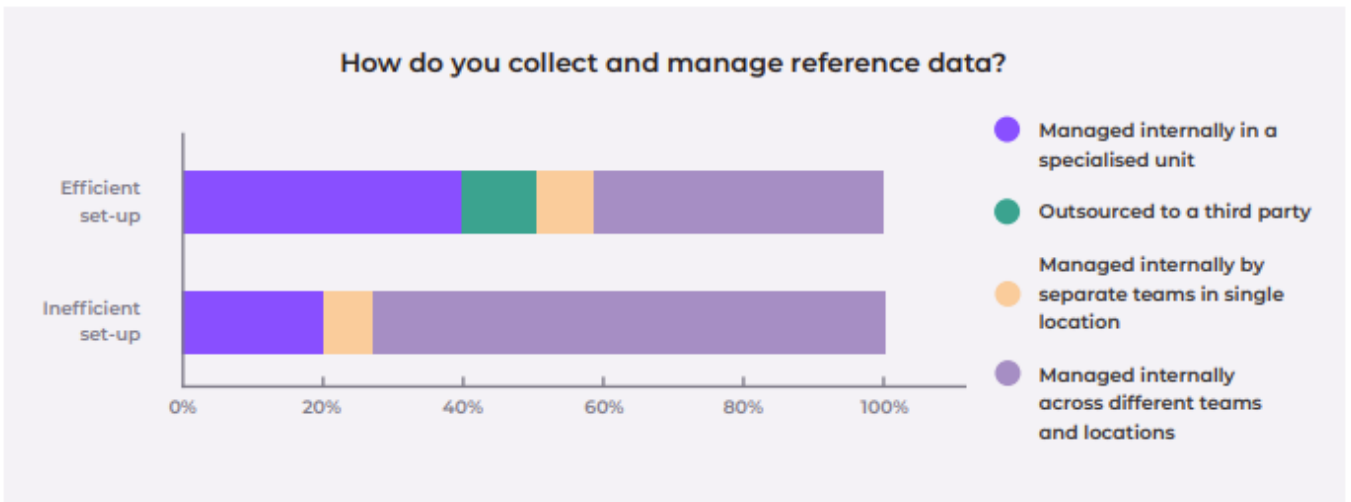
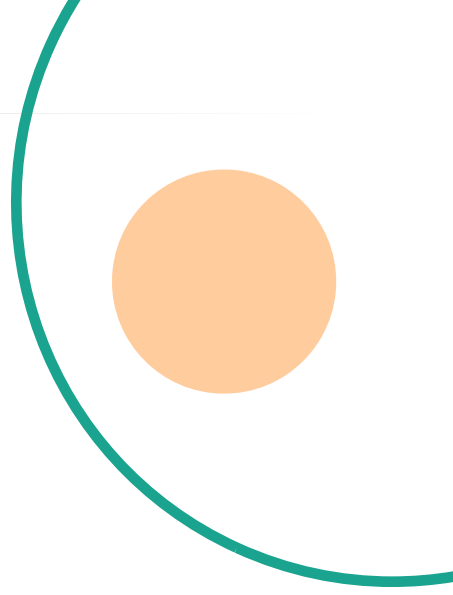
To ascertain the factors that had the greatest impact on the efficiency of a data set-up, Acuiti compared the processes of respondents that reported their operations to be very or quite efficient with those that stated they were inefficient.

This analysis showed that there was no panacea to an efficient set-up. However, using a ‘golden source’ of reference data across the front, middle and back office had the greatest impact with firms reporting efficiency in their operations twice as likely to use a single data source.

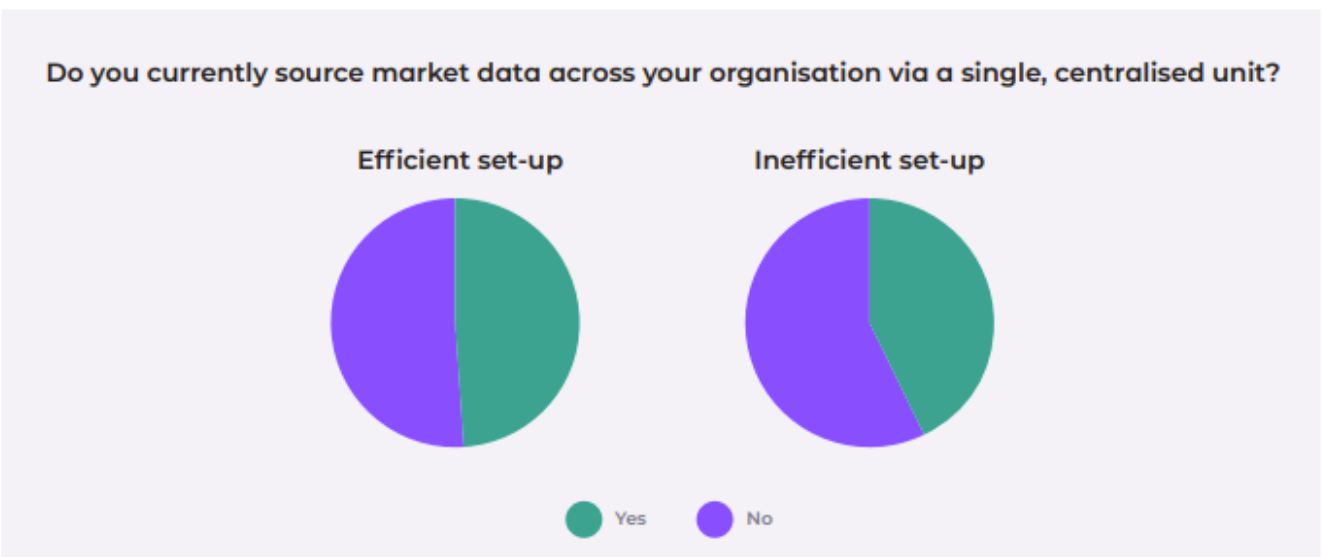


Setting up a single unit to manage reference data also had a significant impact on efficiency, with 40% of those with an efficient set-up structured in this way compared with 20% of those with an inefficient set-up.

Additionally, 73% of those who said their data management was inefficient managed reference data internally across different teams and locations. However, a single unit is not an essential component as 43% of respondents with an efficient set-up managed data across different teams and locations.



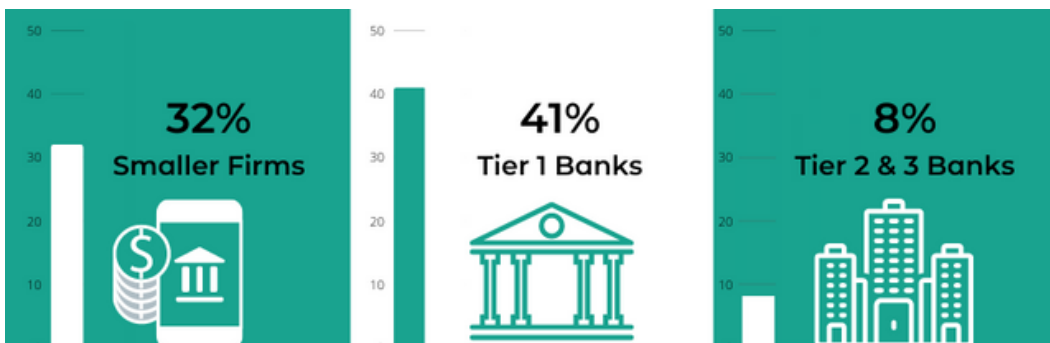
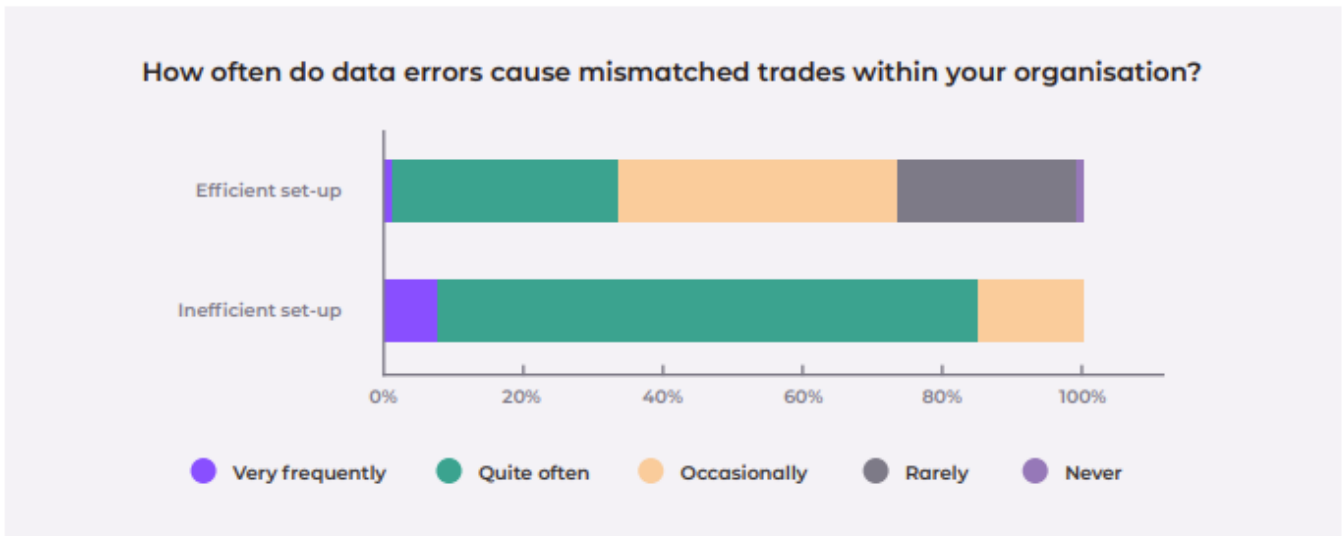
Significantly, the survey found little difference in where data is sourced and acquired with roughly equal percentages of firms sourcing data via a single unit reporting efficient or inefficient data set-ups.





“Issues of siloed data sources within firms were first cited over a decade ago as an impediment to operational efficiency. The sell-side should now be focused on identifying trade break trends from consolidated views of firm-wide trade processing reports, highlighting risk to the business. The listing of full and accurate data set-up for contracts that can trade on markets at any one time is fundamental to supporting any trading service or operation. This applies to all tradeable contracts but moreso to new listings where initial trading volumes can be high due to new market interest and where any trade processing issues will be compounded potentially affecting future trading activity. Sourcing data from a recognised vendor of quality can also provide risk reduction and operational efficiency benefits from being part of a community data model, proven and tested at scale by the industry.”

The benefits of efficiency, however, are stark, with 64% of respondents with an efficient set-up reporting occasional, rare or no mismatched trades caused by data errors compared with 15% of those with inefficient set-ups. This translates into a large operational cost in terms of time and effort in dealing with poor reference data set-ups.



Use of chatbots and Robotic Process Automation in post-trade

Conclusion: incremental steps towards a solution

The problems and inefficiencies caused by the fragmentation of reference data across the derivatives are well-known and tolerated by many firms across the sell-side.

However, this whitepaper concludes that they are not inevitable. Investment by tier 1 firms has shown that many of the challenges can be overcome and, while there is no solution yet that creates a perfect environment for processing reference data, there are many smaller steps that greatly alleviate the challenges.

Emerging technology such as Robotic Process Automation and artificial intelligence is creating the opportunity for firms to significantly improve the process for rectifying trade breaks and other post-trade processes. These are growing in adoption (see infographic on the previous page) and will go a long way to improving efficiencies and reducing the operational burden of poor data.

However, ultimately, greater standardisation of data across the market is essential. This will involve the work of regulators as well as market participants and new reporting requirements must be built around existing protocols and methodologies. The industry is on the path towards an efficient data workflow but there is a long road still to tread.

Choosing a reference data provider

Acuiti asked respondents what factors they considered when sourcing reference data. We found that quality of data was more important than cost, suggesting the central challenge that poor data quality poses operations and the risk of a false economy of a cheaper provider. What factors were the most important?



“The journey toward a data Utopia is valid and desirable for any firm engaged in trading activity and data management strategies founded on zero tolerance are the minimum expectation of clients internally and from their vendors.

“Operational flow analysis can identify the pain points that can be fixed and can lead to improvements to process efficiency. Modern technology can provide micro fixes as part of a trade flow, such as dynamic validation or retrieval of incorrect or missing data, much of which can be automated.

“The devil is still in the detail and still requires the application of subject matter expertise into machine code, to squeeze the last percentage points in the continuous journey toward perfection.”

Kaan Croarkin
Director of Data Platforms,
FOW



About FOW

FOW is the most trusted provider of critical data, news, insights and events for the global futures and options community.

FOW's market-leading data solutions - we source information from the world's trading exchanges and market participants, validating and normalising complex derivatives data and delivering it to the listed derivatives industry as the gold standard of trade reference data. Fully customised around our clients' needs and technology ecosystems and delivered through their workflows, enabling them to achieve seamless trading, clearing and settlement, reduce risk and react faster to changing markets and new industry challenges. Supplying instrument data for over 100,000 contracts on over 110+ exchanges - we are the reference data solution of choice for the world's top financial organisations, exchanges and regulators.

For more information about FOW and our solutions visit:
www.fow.com

About Acuiti

Acuiti is a management intelligence platform designed to provide senior executives with unparalleled insight into business operations and industry-wide performance. Acuiti helps identify market trends, enhance decision-making and benchmark company performance. The platform anonymises and aggregates information from its exclusive network of senior industry figures to provide insightful in-depth analysis.

