The Study on the Creation of an EU Consolidated Tape

FINAL REPORT

Market Structure Partners
September 2020

EUROPEAN COMMISSION

Directorate-General for Financial Stability, Financial Services and Capital Markets Union
Directorate C — Financial Markets
Unit C.3 — Securities Markets

E-mail: FISMA-C3@ec.europa.eu

European Commission
B-1049 Brussels

EUROPEAN COMMISSION

The Study on the Creation of an EU Consolidated Tape

FINAL REPORT

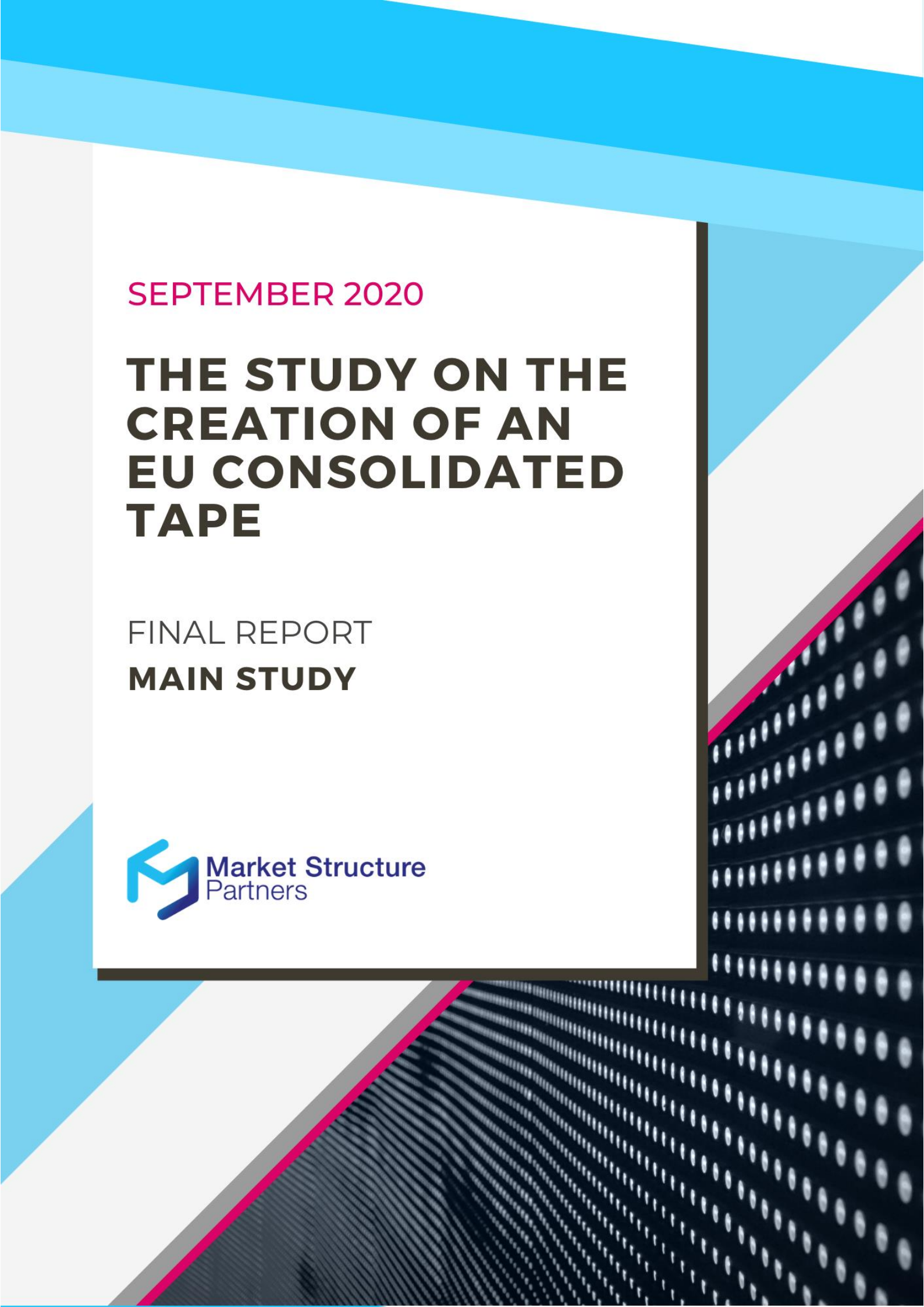
Directorate-General for Financial Stability, Financial Services and Capital Markets Union

SEPTEMBER 2020

THE STUDY ON THE CREATION OF AN EU CONSOLIDATED TAPE

FINAL REPORT

MAIN STUDY



Market Structure Partners

ABOUT MARKET STRUCTURE PARTNERS

Market Structure Partners is an independent consulting firm specialising in global capital markets structure.

We aim to deliver real benefit to society by providing valuable, non-biased, strategic advice and consultancy services with the utmost integrity to help enhance and develop capital market structures.

The Market Structure Partners (MSP) team undertaking this work is entirely made up of experienced industry practitioners who have broad geographic, asset class and cross-functional knowledge. Team members are based in Europe and North America.

ABSTRACT / L'ABSTRAIT

This study analyses the demand for consolidated European financial market data in equities and bonds. It examines the benefits that would arise from such data, the challenges to creating it and recommends an architecture for future data consolidation.

The study is based on the input of 200 data stakeholders and research into North American consolidation frameworks.

It finds that:

- There is high demand for consolidated data and it would be of significant benefit to market participants, including investors, issuers, and regulators.
- Consolidation is prevented by legal, organisational and economic challenges. Technology is not a constraint.
- These challenges could be overcome by the creation of an exclusive consolidated data provider that is overseen by the regulator, run by data stakeholders and empowered to set rules for data contributors.
- The benefits of creating consolidated data outweigh the costs of implementing the organisational and technical frameworks needed to deliver it.
- Significant progress can be made towards a solution by regulators and stakeholders working together under the current legal framework. However, full data consolidation and resolution of all the challenges require further legislative change.

It is recommended that the European Commission pursues its current powers to create an exclusive consolidated tape provider that is run as a utility whilst seeking to introduce further legislative changes to support the full solution.

Cette étude analyse la demande de données consolidées pour les actions et obligations sur les marchés financiers européens. Elle examine les avantages qui résulteraient de telles données, les défis pour les créer et recommande une architecture pour le futur de la consolidation des données.

L'étude est basée sur l'apport de 200 professionnels et études nord-américaines des méthodes de consolidation.

Elle constate que:

- Il existe une forte demande pour les données consolidées qui bénéficieraient considérablement aux professionnels, y compris les investisseurs, émetteurs et régulateurs.
- La consolidation fait face à des obstacles juridiques, organisationnels et économiques. L'aspect technologique n'est pas une contrainte.
- Ces défis pourraient être surmontés par la création d'un fournisseur exclusif de données consolidées supervisé par le régulateur, géré par les parties prenantes et habilité à fixer les règles pour les fournisseurs de données.
- Les avantages de la création de données consolidées sont supérieurs aux coûts opérationnels et techniques nécessaires.
- Des progrès significatifs peuvent être accomplis ensemble par les régulateurs et les parties prenantes travaillant sous la juridiction actuelle. Cependant, la consolidation complète des données et la résolution des enjeux nécessite des changements législatifs.

Il est recommandé que la commission européenne exerce ses pouvoirs actuels pour créer un fournisseur exclusif de compilation des données géré comme un service tout en cherchant à introduire certaines modifications législatives pour soutenir une solution intégrale.

MAIN STUDY

Project Number: 2019.5426				
Title: Study on the creation of an EU consolidated tape				
Linguistic version	Media/Volume	Catalogue number	PDF ISBN	DOI
EN PDF	PDF/Volume_01	EV-03-19-794-EN-N	978-92-76-12113-8	10.2874/434465 EV-03-19-794-EN-N

CONTENTS

GLOSSARY

APA	Approved Publication Arrangement
ATS	Alternative Trading System
BBO	Best Bid and Offer
CCP	Central Counterparty
CLOB	Central Limit Order Book
CMU	Capital Market Union
CSD	Central Securities Depository
CTA	Consolidated Tape Association
CTA Plan	Consolidated Tape Association Plan
CTP	Consolidated Tape Provider
DVC	Double Volume Cap
EBBO	European Best Bid and Offer
EC	European Commission
ECTP	Exclusive Consolidated Tape Provider
EMIR	European Market Infrastructure Regulation
ESMA	European Securities and Markets Authority
ETF	Exchange-Traded Fund
FCA	Financial Conduct Authority (UK)
FINRA	Financial Industry Regulatory Authority
FRTB	Fundamental Review of the Trading Book
IIROC	Investment Industry Regulatory Organization of Canada
IP	Information Processor
ISIN	International Securities Identification Number
MAR	Market Abuse Regulation
MiFID II	Markets in Financial Instruments Directive II
MiFIR	Markets in Financial Instruments Regulation
MTF	Multilateral Trading Facility
NASD	National Association of Security Dealers
NBBO	National Best Bid and Offer
NCA	National Competent Authority
OPR	Order Protection Rule
OSC	Ontario Securities Commission
OTF	Organized Trading Facility
QIF	Qualified Investment Firms
RM	Regulated Market
RTS	Regulatory Technical Standard
SEC	Securities and Exchange Commission
SI	Systematic Internaliser
SIP	Securities Information Processor
SRO	Self-Regulatory Organization
TOTV	Traded on a Trading Venue
TRACE	Trade Reporting and Compliance Engine
TRF	Trade Reporting Facility
TV	Trading Venue

1 / STUDY OVERVIEW

This Chapter explains the background, scope and approach to the study and how the participants and stakeholders were identified and engaged with.

CHAPTER SUMMARY

1. The study is set against the background of the Capital Markets Union's objectives and the underlying regulations (mainly MiFID II/MiFIR and MAR) which serve to meet those objectives by increasing market transparency and integrity and improve competition.
2. MiFID II aimed to improve the quality and availability of market data and included provisions and arrangements for the consolidation of data through competing consolidated tape providers. It also provided for a public procurement process to be put in place if a consolidated tape provider does not materialise, which it has not.
3. The aim of the study is to consider the demand for consolidated pre- and post-trade data in equities and post-trade data in equities and bonds and the demand for that data, as well as to make a recommendation for the future architecture of a consolidated tape for Europe.
4. The study had to take into account at least the 11 largest markets in Europe and consider both US and Canadian experiences in data consolidation.
5. The study methodology included desk and field research using interviews and group meetings in North America and Europe.
6. Market data stakeholders were identified in a functional framework and interviews were sought with representatives from a broad representation of stakeholders.

1.1. STUDY AIMS AND BACKGROUND

The requirements for financial market participants to acquire and maintain complete and accurate data records upon which to base their decisions is becoming more urgent due to new regulations and the increasing proliferation of data.

This study considers the possible establishment of European consolidated tape data (CT data) in the context of the European Union's (EU's) strategic objectives, and the underlying financial markets regulation that has been implemented to achieve these objectives. The aims of the study are to:

- *Provide a clear picture of the demand for an EU consolidated tape/s by determining potential use cases across equity, equity-like instruments and corporate and government bond instruments including the need for pre- and post-trade data and any additional requirements.*
- *Design and recommend a future tape architecture for European consolidated data, taking into account current data aggregation issues, field research of North American markets, consideration of academic and industry literature and European participant feedback.*

1.1.1. EU Strategic Objectives: Capital Markets Union

The Capital Markets Union (CMU)[1] is a major initiative to help stimulate economic growth across Europe and has recently been deepened to enhance access to finance for EU businesses with 5 key objectives:

1. remove structural and legal barriers to increase cross-border capital flows.
2. provide incentives and remove obstacles for well-informed retail savers to invest.
3. support the transition to sustainable economies.
4. embrace technological progress and digitalization.
5. strengthen global competitiveness.

1.1.2. MiFID II, MiFIR & MAR

The Markets in Financial Instruments Directive II (MiFID II)[2] and Markets in Financial Instruments Regulation (MiFIR)[3] (hereafter jointly referred to as MiFID II) are part of a long-term plan to harmonise financial markets across the EU, which includes improving pre- and post-trade transparency and the protection of investors and ensuring overall market integrity.

Market Transparency under MiFID I and II

MiFID I enabled cross-border competition between trading venues in equities. It also harmonised equity pre- and post-trade transparency rules and created the concept of data aggregators of trade reports. The result was the fragmentation of markets for trading and also of trade reporting of data, as well as an increase in the amount of data being generated. Despite some attempts to create it, CT data that the whole equity market could refer to did not emerge.

MiFID II introduced a similar pre-and post-trade data transparency regime for non-equity asset classes and a more stringent definition of trading venues across all asset classes. As a result, more data is being created and processed across all asset classes than before, and market participants are now grappling with increased data fragmentation and quality issues.

Data Consolidation under MiFID II

Regulators recognised that CT data should be the principal tool for addressing these issues and for providing a more transparent and complete picture of equity and non-equity market activity. As a result, MiFID II formally introduced data consolidation into the law by recognising consolidated tape providers[4] (CTPs) as entities to consolidate data and providing rules about the data that should be sent to them. It assumed that competing commercial solutions for consolidating equity instruments would emerge and that other asset classes would learn from that experience and follow suit.[5]

However, to date, no solution has materialised. In the event that this occurred, the Directive allowed for steps[6] to be taken by the European Commission (EC) to request ESMA to launch a negotiated procedure for the appointment, through a public procurement process run by ESMA, of a commercial entity operating a consolidated tape. The EC can also specify the conditions for this procedure. This study will contribute to the Commission's considerations on consolidated tape requirements and potential development across a broad set of asset classes.

MAR

The Market Abuse Regulation (MAR)[7] is designed to work in tandem with MiFID II by bringing integrity and public confidence to markets across all asset classes. It aims to establish a stronger framework for the markets, avoid potential regulatory arbitrage and ensure accountability in the event of attempted market abuse. Market participants across asset classes are very reliant on good data to enforce it.

1.1.3. Other Legislation Linked to Data

Other legislation, which is heavily reliant on data, has been introduced to strengthen capital markets. Examples of such legislation are the Fundamental Review of the Trading Book (FRTB)[8] and the initial margin rules in the European Market Infrastructure Regulation (EMIR).[9] These require clear, unambiguous trade prices and accurate historical data to calculate all required sensitivities, correlations and valuations. Such critical calculations will only be as good as the data from which they are derived.

1.2. STUDY SCOPE

1.2.1. Instruments and Type of Data Covered

The scope of the study was to review use cases for:

- Pre- and post-trade consolidated data for equity and equity-linked instruments (hereafter summarised as equities unless otherwise stated).
- Post-trade data for corporate and government bonds (hereafter summarised as bonds unless otherwise stated).

The study also explored whether it would be useful to include other information alongside the CT data that would be of value for analytical purposes.

1.2.2. Geographical Coverage

The study was required to cover a representative sample of Member States (inc UK) that:

- Constitutes at least 80% of the total amount outstanding of equity listed on regulated markets or MTFs at the end of 2018 (see Appendix 1), and,
- Includes the following 11 countries: France, Germany, Italy, the Netherlands, Belgium, Ireland, Poland, Sweden, Spain, Denmark and the UK.[10] The outcome of the UK's departure from the EU is unknown but has been considered further in Chapter 10.

It also required an analysis of existing consolidated data frameworks in the US and Canada.

1.3. APPROACH AND METHODOLOGY

The study took place over a five-month period between October 2019 and March 2020.

Four high-level steps to the overall project approach were used in the study (described in Appendix 2).

- Collecting background information about European and North American data consolidation frameworks.
- Creating a structured framework for data gathering and analysis from a cross-section of the identified data stakeholders (see Figure 1 below) in Europe and North America.
- Identifying, screening, and obtaining data from relevant stakeholders
- Defining and validating recommendations.

1.4. IDENTIFICATION OF MARKET DATA STAKEHOLDERS

A stakeholder framework of potential data users (Figure 1) was developed at the outset of the study and was used throughout to help ensure that the inputs of all potential users of consolidated data were captured.

High-level groups of stakeholders were mapped into industry groups and then into granular entity types along with an estimate of the numbers behind each European industry group. More detailed descriptions of what each stakeholder entity type does are given in Appendix 3 and the sources and calculations behind the numbers of users are given in Appendix 4.

The analysis showed that data is used by a broad set of stakeholders performing many different industry functions. Interviews were sought with representatives from each group to verify the functions and use cases associated with the stakeholders and refinements and additions were made during the study.

Figure 1: Market Data Stakeholder Group Descriptions

Stakeholder Group	Industry Grouping	Estimated Minimum Number of Entities that use Data in Europe, inc. UK, (Source of Data)	Stakeholder Entity Type
Issuers	**Issuers**	40 (est.)	Government Issuers Supranational Issuers Agency Issuers
		9,321 (WFE)	Corporate Issuers
		69 ETF (ETFGI)	Issuers of Funds including ETFs
End Investors	**End Investors**	15,500,000 (users not entities) (ECB+UKHMRC)	Direct Retail Customers
Financial Intermediaries	**Buy-Side**	4,366 (ESMA)	Asset Managers & Portfolio Managers OEICs/SICAVs (inc. ManCos and Fund Administrators) Insurance Companies Pension Funds Endowments Wealth Managers
	Sell-Side	10,576 (ESMA)	Investment Banks (inc. SIs) Commercial / Retail Banks Development Banks Institutional Agency Brokers Inter-dealer Brokers Retail Brokers Proprietary Traders (inc. SIs) Originators / Advisors
	Trading Venues/APAs	136 (ESMA)	Regulated Markets
		220 (ESMA)	MTFs (Multilateral Trading Facility)
		74 (ESMA)	OTFs (Organised Trading Facility)
		20 (ESMA)	APAs (Approved Publication Arrangements)
Custodians/ Risk Managers	**Custody, Clearing & Settlement**	20 (est.)	Custodian Banks
		15 (ESMA)	CCPs (Central Counterparty)
		Not sourced or estimated	3rd Party Clearers (GCMs)
		30 (ECSDA)	CSDs/ICSDs (Central Securities Depository / International Central Securities Depository)
Regulators	**Regulators**	57 (FCA)	Regulators
Data Analytics & Benchmark Providers	**Data Analytics & Benchmark Providers**	Not sourced or estimated	Benchmark Providers
		Not sourced or estimated	Data Analytics Providers
Other	**Other**	Not sourced or estimated	Research Providers
		Not sourced or estimated	Chartered Financial Planners
		Not sourced or estimated	Independent Financial Advisors

1.5. STAKEHOLDER ENGAGEMENT

In total 189 entities and 11 trade associations were consulted with using one on one interviews, small group discussions with industry associations and some group sessions held at the Alpha Forum Conference in February 2020 for European buy-side traders in both equities and bonds.

39 interviews took place in North America and the remainder were in Europe. Most of the interviews in Europe took place with European headquartered organisations.

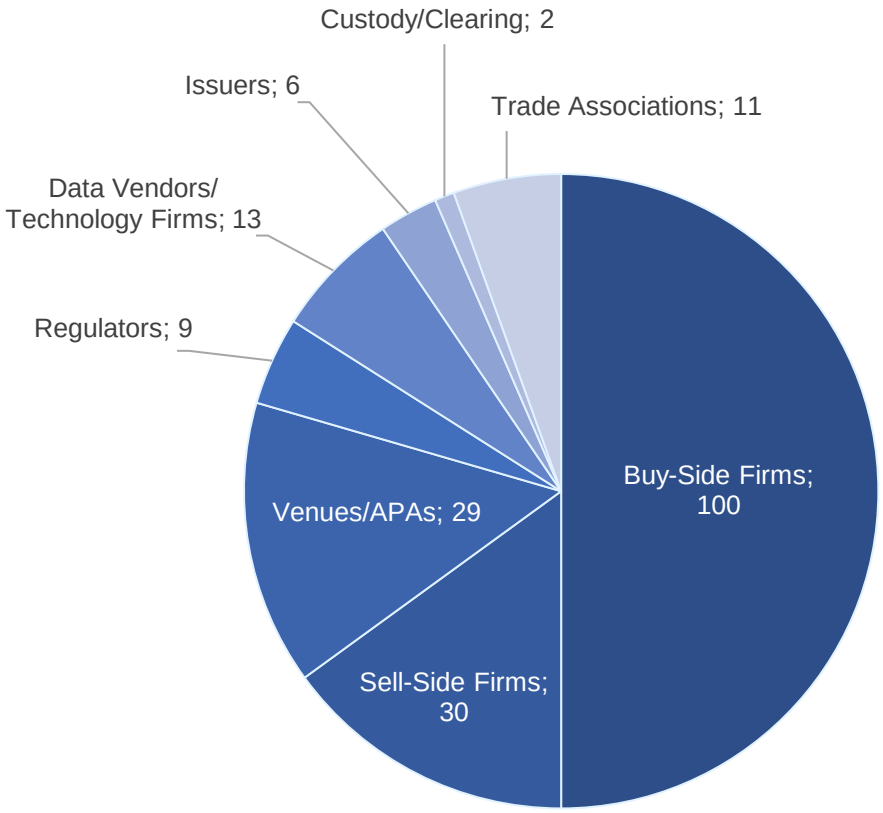
The majority of the secondary market interviewees in bonds, regardless of their location in Europe, covered investment processes and trading in the complete set of pan-European markets.

Interviews with secondary markets equity participants regardless of their location covered investment processes and trading in an extensive range of the largest 11 European markets and many covered most of the smaller markets.

Interviews with primary market participants and issuers took place in Holland, France, and Ireland.

A final workshop was held with some of the participating stakeholders to validate the feedback. Some of these stakeholders had been previously interviewed and some had not. This was held with a cross-section of stakeholders from all over Europe.

Figure 2: Categories of Stakeholders (By Industry Group) Engaged with For This Study



2 / WHAT IS MARKET DATA AND HOW IS IT CONSOLIDATED?

This Chapter discusses how pre- and post-trade data is determined, who generates that data and how the MiFID II rules apply to the transparency, aggregation, and consolidation of that data.

CHAPTER SUMMARY

1. **Consolidated data** is the collection of pre- and post-trade data from multiple sources of trading data across an entire financial market that is disseminated through a single, standardised, market data feed.

2. **Pre-trade data** comprises the visible prices and volumes of orders placed in order-driven markets or visible quotes advertised in quote-driven markets.

3. **Orders** are firm and participants can interact with every order that they can see. **Quotes** may only be firm up to the advertised size but, even then, may have to be negotiated. Participants are not guaranteed to be able to interact with orders and quotes that are not visible to the market such as for large-sized trades or trades that are subject to other special conditions.

4. **Post-trade data** comprises the prices and volumes of trades that have been executed against those visible orders or quotes, as well as trades executed against orders and quotes that were not visible to the entire market. It also includes end-of-day statistics.

5. Participants define the orders and quotes that are visible and that they can interact with as **"addressable liquidity"**. Post-trade data is expected to hold sufficient information in the form of flags to ascertain whether the liquidity was addressable at the time of the trade.

6. Order and quote data is generated by **Qualified Investment Firms (QIFs) and Systematic Internalisers (SIs)**. These are all regulated entities managing orders, quotes, and trade executions either for investors or dealing on their own accounts.

7. **Trading Venues (TVs)** aggregate on-venue pre- and post-trade data captured from their systems and **Approved Publication Arrangements (APAs)** aggregate off-venue post-trade data for equities and bonds from QIFs and SIs that must report to them.

   The law envisages **multiple competing aggregators** and there are already hundreds of TVs and APAs across asset classes in Europe.

8. MiFID II only considers the consolidation of post-trade data which must be collected from the TVs and APAs and consolidated by Consolidated Tape Providers (CTPs). The law envisages **multiple competing consolidators** but there are none.

9. TVs, APAs and CTPs are all regulated entities but TVs can self-regulate by setting rules and sanctioning their members without requiring intervention from the regulator whereas APAs and CTPs are dependent on their clients to adhere to regulatory standards and enforcement.

2.1. WHAT IS CONSOLIDATED DATA?

The term consolidated data or a consolidated tape (CT data) is broadly used to describe the collection of various data feeds from multiple sources associated with trading activity across an entire financial market, and the dissemination of that data, through a single, market data feed in a "standardised" format.

It can include both pre-trade and post-trade information and may be provided and consumed in real-time, or on a delayed time basis.

2.2. WHY IS IT IMPORTANT?

Information is the lifeblood of financial markets. Each piece of information creates data, which has a rippling effect throughout the market. It drives behaviour and leaves a permanent record or footprint that derives from the original data itself. Incomplete or incorrect data drives behaviours and outcomes that are different and may be suboptimal compared to those that would have prevailed had full and accurate data been available. Incomplete data also means that a true picture of the market at any point in time is lost and hence the value of any subsequent analysis or comparison will be reduced.

A complete and accurate picture of the entire market at all times is therefore important for all investors and those who act on their behalf. If information asymmetries exist, then some participants or investors will be at a disadvantage.

2.3. HOW IS MARKET DATA GENERATED?

The trading models and mechanisms used in financial markets directly impacts the data that is delivered to users and their ability to manage and use the data in real-time and after the event. The market models in European and North American markets share similarities and tend to be classified either as order-driven markets or quote-driven markets.

Trading Venues (TVs) exist to facilitate most trading. In Europe, under MiFID II, there are different types of TVs which are defined as:

1. Regulated Markets (RMs) which offer primary and secondary markets trading. In both Europe and North America these are more simply known as exchanges.
2. Multilateral Trading Facilities (MTFs) which offer secondary markets trading only. In North America, the equivalent would be Alternative Trading Systems (ATSs).
3. Organised Trading Facilities (OTFs), which are a new venue for non-equities only and are similar to MTFs except that they can have more discretion over the management of orders. North American markets do not make such distinctions with their ATS systems.

These markets may be order or quote-driven or a hybrid of both.

2.3.1. Pre-Trade Data

Market data begins its journey when an investment firm, which could be a buy-side or a sell-side firm as described in Figure 1, Chapter 1, creates and/or handles orders on behalf of investors or other investment firms. Under MiFID II, investment firms are known as Qualified investment firms (QIFs), which are defined as "any legal person whose regular occupation or business is the provision of one or more investment services to third parties and/or the performance of one or more investment activities on a professional basis".[11]

Order-Driven Markets

Order-driven markets are most prevalent in liquid instruments, such as equities and, under MiFID II, QIFs must send all equity orders to TVs so that they contribute to the price discovery process. TVs must

make the current bid and offer prices of orders continuously available during normal trading hours, along with the depth of trading interests at the prices advertised.

However, some exception waivers exist which allow QIFs or TVs to withhold orders from publication. For example, if an order is large in scale, it may require more careful management because the market impact would be too great if it were published immediately.

TVs operate Central Limit Order Books (CLOBS) where buyers' and sellers' orders are gathered and displayed to all participants. CLOBs allow multilateral interaction and continuous matching between all participants' orders, subject to each TV's pre-defined rules. If the right conditions are present, two orders can result immediately in a matched trade, thus all orders are considered "firm" when they are sent to the CLOB, as no further intermediation is required. This is also facilitated by clearing through a central counterparty (CCP), which serves to eliminate concern about the risk of a counterparty defaulting on the trade, meaning that a counterparty's identity and creditworthiness is not a consideration prior to trading.

A number of other trading systems and protocols have been developed by equity TVs as an extension or a complement to their CLOBs to help manage orders. For example, the system sometimes moves from offering continuous matching of orders to offering different types of auction processes such as frequent batch or a closing auction. When viewing the market, participants need to know which system or segment a TV is operating at any time so that they know what trading protocols and choices that are available to them. TVs transmit trading status codes to identify the system and protocol being used

However, under certain protocols, orders and prices are not always transparent to the market prior to execution or during the auction period and so are not contributing to price discovery. (A list of market models and underlying protocols is given in Appendix 5). Pre-trade market prices in order-driven markets are therefore determined only by the visible prices and orders that all participants can interact with and which are published by TVs.

Quote-Driven Markets

Quote-driven markets are most prevalent in less liquid instruments such as bonds. This is because there may not be enough natural liquidity for orders to be matched immediately and the markets need to be intermediated by firms that are willing to risk their capital to supplement liquidity in the market.

Under MiFID II, QIFs can deal on their own account by risking their capital and QIFs that do this are usually sell-side firms only, as described in Figure 1, Chapter 1. If a QIF deals on its own account on an organised, frequent systematic and substantial basis when executing client orders outside of a TV then it must become a Systematic Internaliser (SI).[12]

SIs, in both equities and bonds, are obliged to make the prices at which they are willing to deal for trades up to a certain standard size public. They can advertise their quotes either through a TV or through a private channel.

However, participants in quote-driven markets are not always guaranteed of being able to execute a trade even when a quote is visible. Quotes are disseminated electronically but automatic execution is not always offered and, if it is, it is usually restricted to trades up to the quoted size. Otherwise further manual or electronic negotiation is required because, for example, the price for a larger size trade will be different from the advertised price.

Quote-driven markets also tend to be bilateral with no centralised clearing and settlement, thus the identity of the counterparty is very important, and firms will likely undertake credit assessments and establish credit limits for each trading counterparty before trading. Most firms maintain lists of approved counterparties with corresponding credit limits and will only trade outside these limits on an exceptions basis.

Pre-trade quote data is therefore limited by the visibility of the on-screen quotes that dealers are obliged or willing to make public to the entire market through a variety of TVs or private channels. It is these quotes that constitute pre-trade data, but counterparty considerations may further reduce the pool of liquidity available to each participant.

TVs that operate quote-driven markets also offer different trading protocols, for example where two-sided quoting is not mandatory then one-sided quoting is possible, or participants can send different

types of requests for quotes to certain counterparties. (A list of bond market models and underlying protocols is given in Appendix 5).

Hybrid Markets

In practice, even the more liquid markets, such as equities, operate a hybrid system of order-driven and quote-driven markets. For example, quoting is frequently used to support liquidity in small and medium capitalised firm's shares which are less liquid. Conversely, there are examples of more liquid bonds (e.g. treasuries) starting to migrate to CLOBs.

In each instance, orders and quote data can be made publicly available via the same market data feeds but would be consolidated separately because they have different characteristics.

2.3.2. Post-Trade Data

Once a trade has been executed, whether as a result of an order or a quote that is visible to the market or as a result of non-visible negotiations that may occur through a different TV protocol or outside of a TV, it simply becomes an executed trade.

TVs, SIs and QIFs must publish the volume and price of any trade in a financial instrument that is defined as traded on a trading venue (TOTV) within one minute of execution for equity and equity-like products. For non-equity products publication must occur within fifteen minutes of execution, falling to five minutes in 2020. Both equity and bond trade publications are subject to deferral regimes where trade publication may be deferred depending on the size or liquidity of the instrument.

Addressable Liquidity

Market participants call the liquidity that is visible, and that they can fully interact with, "addressable liquidity". When looking at post-trade data they need to separate out the addressable liquidity, that they could have interacted with, from the non-addressable liquidity, that they could not have known about because it was not public at the time. As discussed above, non-addressable liquidity could include trades that were negotiated away from the quote advertised, special technical trades or possibly trades with counterparties that are not approved by the firm wishing to execute the trade.

This is especially important information when measuring the success of a trading strategy or when making decisions about the potential liquidity of an instrument. Post-trade data, therefore, needs to hold qualifying data to help participants determine whether the trade contributed to addressable liquidity that the participant could have accessed. This data should include indicators such as the trading protocol that was used for the trade, where the trade was executed, why it was executed, when it was executed, how large it was and whether it was above the size being advertised.

2.4. DATA AGGREGATION UNDER MIFID II/MIFIR

MiFID II only addresses the aggregation of post-trade data. The intention is for data to be gathered into larger pools of data for ease of access. Aggregation is undertaken either by:

1. TVs, which naturally capture the post-trade data resulting from the trading undertaken on their systems, **or,**
2. Approved Publication Arrangements (APAs), which are responsible for aggregating and publishing details of trades executed off-venue by QIFs and SIs. In bond markets, even when quotes are advertised on a TV, they are often executed bilaterally and reported to APAs.

Both types of aggregators must publish data as close to real-time as possible, on a reasonable commercial basis. Data should be made available for free 15 minutes after publication and should be disseminated in a manner that ensures fast market-wide access on a non-discriminatory basis. Aggregators should check trade messages for accuracy and completeness. Some APAs calculate and manage the publication of deferred trades as an added value service for their clients.

MiFID II envisages multiple competing data aggregators[13] and there are many already functioning.

According to ESMA, in European equity markets, there are 178 TVs (67 RMs, 111 MTFs) that aggregate pre- and post-trade information and a further 17 APAs that aggregate post-trade data (see Figure 3). In bond markets, there are 96 RMs, 147 MTFs and 50 OTFs and 20 APAs, although much of the post-trade reporting would be routed to the APAs (see Figure 4).

2.4.1. Regulatory Characteristics of Data Aggregators

TVs and APAs are authorised and regulated by their National Competent Authorities (NCAs). However, there are organisational differences between the two entity types. TVs have an additional layer of their own regulation that supplements any formal legislation and does not require the direct involvement of the regulator.

TVs achieve this through the self-regulation of their members or participants upon whom they can enforce their own rules and standards. If a member violates the rules, a TV can restrict or fully retract its membership. This may have serious consequences for a member's business; if a trading member found itself excluded from a venue, particularly one with large market share, then it may not be able to fully conduct its own or its clients' business because that TV cannot be simply swapped for another.

APAs do not have members and so cannot easily enforce their rules in the same way as TVs. A firm that decides to move from one APA to another can do so with little consequence to its business.

TVs and APAs can both set data standards but they both have a commercial incentive to set their own proprietary standards. Neither has the legal authority nor obligation to enforce European data laws and rules.

2.5. DATA CONSOLIDATORS UNDER MIFID II/MIFIR

Data consolidators were introduced under the rules of MiFID II/MiFIR as Consolidated Tape Providers (CTPs). CTPs are authorised to collect post-trade data from the data aggregators, namely TVs and APAs, but not directly from QIFs or SIs.

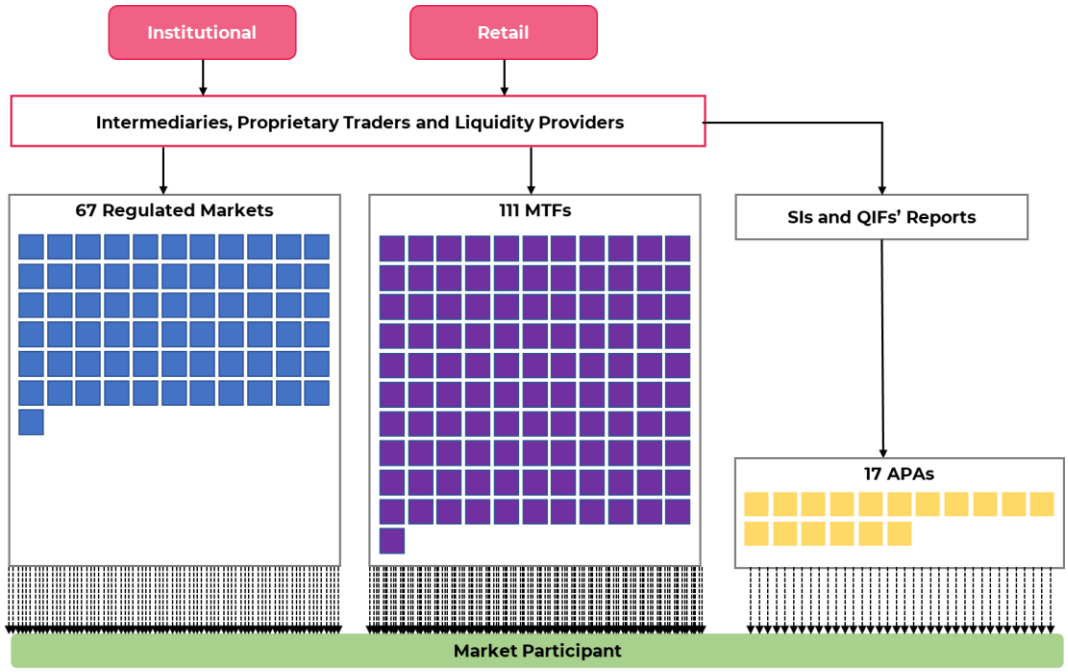
CTPs can specialize in different asset classes. An equity CTP is expected to consolidate 100% of the market and a non-equity CTP must consolidate at least 80% of the total volume and transactions for each asset class published by all trading venues and APAs in the EU. Any CTP would have to consolidate data from hundreds of aggregators as illustrated in Figure 3 and Figure 4 below. There is currently no consolidator in equities or bonds.

2.5.1. Regulatory Characteristics of Consolidators

MiFID II envisaged that multiple CTPs would exist and compete with each other whilst regulated and overseen by their NCAs.

Like APAs, CTPs were conceived in the legislation with no concept of membership, ability to set rules or apply sanctions, nor the ability to set market-wide standards or enforce data laws. Such responsibilities would not fit easily with the idea of competition.

Figure 3: European Equities Reporting Flow Post-Trade Aggregators

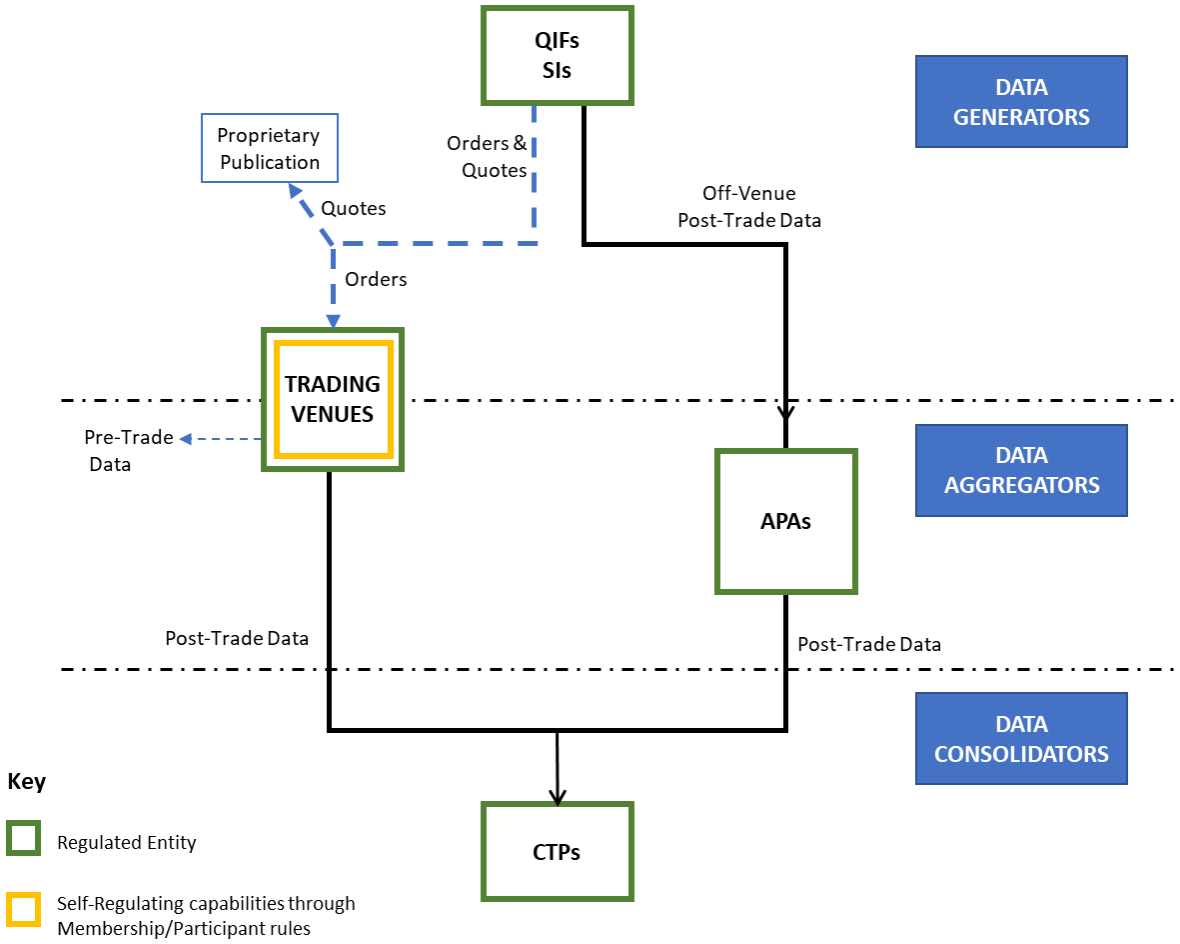


Source: MSP Research, ESMA

Figure 4: European Bond Trade Reporting Flow Post-Trade

Source: ESMA - (Note the number of venues includes derivatives venues, but the number of bond markets alone is not provided)

Figure 5: Flow of Data and Regulatory Characteristics of Data Generators and Aggregators under MiFID II/MiFIR



Source: Market Structure Partners

2.6. CONCLUSIONS

Pre- and post-trade data are important components of consolidated market data. Both types of data need to carry sufficient, accurate and complete information for correct interpretation, and this is particularly important to help market participants accurately assess what is addressable liquidity.

MiFID II considers the aggregation of both pre- and post-trade data in equities, but it only considers the consolidation of post-trade data in equities and bonds. Multiple firms generate data, but only APAs and TVs can aggregate data. APAs and TVs are the only entities that can send data directly to CTPs. There are many aggregators from which a CTP would need to consolidate data, but no CTP exists.

TVs, APAs and CTPs have different characteristics. Only TVs can set rules and sanction their members without relying on regulatory support.

3 / DATA: FEEDS, INTERFACES AND CONSUMPTION

This Chapter explains the data feeds and interfaces, the constituents of data feeds and how end-users consume market data.

CHAPTER SUMMARY

1. Data engagement occurs via three main types of data interfaces: 1) order or quote entry interfaces, 2) market data feeds and 3) reference data feeds. These are operationally independent of each other.

2. Order/quote entry feeds are used to send a firm's orders or quotes to a trading venue. Market data feeds disseminate all the pre-trade and post-trade data for that market and are the most important in relation to consolidation. Reference data feeds are static and used to communicate important information such as instrument data at the start of the day.

3. Pre-trade data can be shown at individual price levels or aggregated to show the interest at each price level. If data is aggregated, then the depth of information to be shown needs to be decided. The simplest approach is to take the best bid and best offer, which is known as the BBO. The most detailed approach is to show the entire depth of the book. Alternatives are to take a subset of data down to a certain level e.g. 3 or 5 levels.

4. Other data that is disseminated in the market data feed includes session statistics for order-driven markets and important administrational event information about the trading session e.g. a trading halt.

5. There are different ways of transmitting market data depending on the volumes that may have to be disseminated. Some TVs and APAs have different delivery mechanisms depending on the requirements of the recipients. Data in very liquid markets can be delivered at high speed, often within nanoseconds, which gives rise to the term 'low latency'.

6. Historical market data is data that is stored and may be available for analysis on an ex-post basis.

7. Firms that provide data will often do so for economic profit and put contracts and licenses in place to govern how data is paid for and who can access and use the data. The pricing and contractual terms have many variations depending, for example, on whether data is used by a human or machine or a professional or non-professional investor.

   This leads to many contractual complexities. Any CTP would have to negotiate multiple contracts and licenses with each of the hundreds of aggregators that exist in Europe.

3.1. DATA INTERFACES AND FEEDS

Data engagement occurs through three main types of data application programming interfaces (APIs) and feeds; order or quote entry interfaces, market data feeds and reference data feeds. These are described below.

Figure 6: Data Feeds and Interfaces

Order Quote or Entry Feeds	Market Data Feeds	Reference Data Feeds
• Orders or quotes originate at an individual participant that uses the order or quote entry interface to transmit buy/sell interest to a trading venue. The trading venue then transmits execution messages and acknowledgements back to the member. • These feeds are also called "private" feeds because they carry data that is specific only to the individual member.	• These feeds disseminate all the order, quote and trade activity emanating from the buy/sell interest received via the order entry interfaces, along with the associated prices and volumes. • They are also called "public" feeds because they are based on all the aggregated orders and quotes available for execution at a venue. • They may also include other messages that are required for the correct interpretation of the information, such as administrational update data, statistics or intra-day reference data updates.	• Venues disseminate data relating to their universe of tradeable instruments, e.g. trading session-related parameters, (start and end times, static and dynamic price bands), connectivity settings (IP address and ports) and business-related parameters (ISIN, quoting currency, etc.). • These feeds are static in nature and uploaded at the beginning of the day. It is important data that is fed to front, middle & back-office systems of the participants interacting with a venue.

Source: Market Structure Partners

The most important of these feeds and interfaces for consolidation purposes is the market data feed and this chapter mostly focuses on this feed. However, it is important to understand that all three feeds are closely related to each other but from an operational point of view, they can be independent of each other. For example, an order entry interface may have technical issues, whilst a market data interface is operating perfectly.

3.2. THE COMPONENTS OF THE MARKET DATA FEED

Market data feeds have different components which are all important when assessing the viability and use cases for CT data. There are five key components: pre-trade data, post-trade data, statistics, session administration messages and intraday reference data updates. Each is examined in turn below.

3.2.1. Pre-Trade Data: Order and Quote Events

This refers to the dissemination of the buy and sell interest resting in an order book. The word "interest" is used because order book or quote information can be disseminated in multiple ways, which may or may not include a breakdown of individual orders.

Order Book and Quote Events (includes new orders and quotes being inserted into the order book or quoting venue, as well as cancellations and amendments.).

- Order book and quote dissemination are defined by two key attributes: 1) "Aggregation" (Y/N) and 2) "Depth" (number of levels).

- **Aggregation**: Events may be disseminated to show individual orders or aggregated to show total interest by individual price level.
- **Depth:** Different levels of depth can be defined for aggregated feeds. A full depth-of-book feed includes all available buy/sell interest at each available price level on either side of the book but not everyone viewing data needs to see the full depth. The simplest alternative is to take only the first level (Level 1) which generates a single best bid and best offer, known as the BBO. This is the basic offering that most TVs offer. Alternatively, participants may take a number of levels of best orders available, such as 3 or 5 levels of depth. In this case, a volume-weighted BBO could be calculated if required.

A more detailed explanation of the different types of order event feeds and their pros and cons are given in Figure 7 below.

3.2.2. Post-Trade Data: Trade Events

Trade events (executions, cancellations/amendments, trade reports, etc.) are usually included in the same "channel" as the order and quote book feed, meaning that regardless of the type of order and quote book feed being consumed, all executions and other trade-related event data will be included.

3.2.3. Statistics

This refers to key statistics related to the current trading session for each instrument. These usually include:

- Previous day's closing price and current day's opening price which remain static during the day.
- The current session's high and low prices (updated dynamically as they are breached).
- The cumulative volume (updated dynamically every time a trade has been executed).
- The current day's closing price (which will be updated immediately after the closing auction or based on the trading venue's alternative methodology, as the case may be).

The inclusion of session statistics is usually inversely related to the complexity of the market data feed:

- Low latency feeds (discussed below) may exclude this functionality because it adds latency due to a higher volume of traffic and also consumers of those feeds typically have the technical knowledge to track the statistics they require.
- Less complex feeds (such as Level 1 or Aggregated Level 2) are likely to include the additional statistics to facilitate the consumption of all data to clients that are less concerned with latency.

3.2.4. Session Administration Messages

This refers to messages that may be required for the proper interpretation or handling of the market data and could include, for example:

Order book status updates	*Trading halts*	*Declaration of fast markets*	*"Other"*

The exact scope of the messages required will depend, to a large extent, on factors such as trading systems architecture, the type of products traded, product offering etc. These administration messages are usually included in the same "channel" as the order book updates and the post trade data. It is not uncommon for different types of order book feeds to contain different subsets of these administration messages.

Figure 7: Most Common Types of Order Book Feed (By Depth and Aggregation)

Red Box denotes User Requirements explained in Chapter 4.2.1.

	Type of Pre-Trade Data Feed Design				
	Market by Order	**Market by Level Full Depth**	**Market by Level e.g. 5 Levels**	**Level 1**	**VWAP**
Key Features	• A breakdown of all individual orders on each side of the order book; includes all orders and all price levels.	• Sum of all liquidity available at all individual price levels on each side. • Also includes the number of orders at each price level.	• Sum of all liquidity available per individual price level and side, up to the 5th best price level on each side. • Also includes the number of orders at each price level.	• Sum of all available liquidity per at the best price level of each side of the book • Also includes the number of orders at each price level.	• The weighted average price which a "typical" order with a pre-defined value (e.g. 7K EUR) could be filled on each side of the book, based on the liquidity available at each price point, as well as the number of shares that would be included in the trade.
Depth	Full Depth	Full Depth	5 Levels	1 Level	Example 7000 EUR size
Aggregation	Dis-aggregated	Aggregated	Aggregated	Aggregated	Aggregated
Hypothetical Example of Rendering Based on Number of Shares Available at Different Prices	**Bid** 100 x 10.00 300 x 10.00 100 x 9.98 150 x 9.98 100 x 9.95 200 x 9.91 200 x 9.90 200 x 9.90 200 x 9.90 200 x 9.89 200 x 7.88 200 x 5.00	**Offer** 100 x 10.10 200 x 10.10 50 x 10.10 100 x 10.10 100 x 10.15 100 x 10.15 100 x 10.20 200 x 10.25 100 x 10.30 100 x 12.00 100 x 18.50	**Bid** (2) 400 x10.00 (2) 250 x 9.98 (1) 100 x 9.95 (1) 200 x 9.91 (3) 600 x 9.90 (1) 200 x 9.89 (1) 200 x 7.88 (1) 200 x 5.00	**Offer** 450 x 10.10 (4) 200 x 10.15 (2) 100 x 10.20 (1) 200 x 10.25 (1) 100 x 10.30 (1) 100 x 12.00 (1) 100 x 18.50 (1)	**Bid** (2) 400 x 10.00 (2) 250 x 9.98 (1) 100 x 9.95 (1) 200 x 9.91 (3) 600 x 9.90	**Offer** 450 x 10.10 (4) 200 x 10.15 (2) 100 x 10.20 (1) 200 x 10.25 (1) 100 x 10.30 (1)	**Bid** (2) 400 x 10.00	**Offer** 450 x 10.10 (4)	**Bid** 700 x 9.9892	**Offer** 691 x 10.1204

	Type of Pre-Trade Data Feed Design				
	Market by Order	**Market by Level Full Depth**	**Market by Level e.g. 5 Levels**	**Level 1**	**VWAP**
Pros/Cons of the Different Types of Feeds	+ Price-time priority of each order is known, so allows for more accurate routing and modelling. - Requires individual monitoring of all order-related events on an order-by-order basis. - May include irrelevant information (i.e. orders priced very far from touch levels).	+ Simpler processing of order book events as they are disseminated at an aggregated level. - Offers the same view of overall liquidity as market by order but offers less granular routing and modelling. - May include irrelevant information (i.e. liquidity priced very far from touch levels).	+ Simpler processing of order book events as they are disseminated at the aggregated level + Provides assurance regarding price formation. - Assumes that only the first *n (in this case 5)* levels of liquidity are relevant for pricing and discards any "noise" from price levels that are too far from the Best Bid/Offer.	+ Very low bandwidth and processing requirements. + Provides an indication of current Best Bid/Offer prices. - Can be misleading, especially for less liquid instruments, where meaningful liquidity may be at price levels that are further down the order book. - Reliability decreases in markets where there are no minimum lot size requirements (i.e. best price can be available only for 1 share).	+ Addresses the weakness of a Level 1 feed by calculating Best Bid and Offer prices that will be available for a "typical" order size.
Latency and Technical Implications	• Lowest latency. • Highest level of granularity with the most market detail. • Usually enriched with non-functional details such as latency timestamps. • High bandwidth requirements due to the volume of information.	• Increased latency as a result of the aggregation process. • Lower granularity but still detailed • Usually excludes any details that are not functional in nature (e.g. latency measurements) • Slightly lower complexity • Lower bandwidth requirements	• Added latency as a result of the aggregation process. • Less granularity with some detail. • Usually excludes any details that are not functional in nature (e.g. latency measurements). • Lower complexity than Market by Order. • Lower bandwidth requirements.	• Added latency as a result of the aggregation process. • Very low bandwidth requirements as less volume. • Very low complexity as no underlying detail is included.	• Added latency as a result of VWAP calculation. • Very low bandwidth requirements. • Low complexity.

3.2.5. Intraday Reference Data Updates

Reference data updates may or may not be included in a market data feed. For example, TVs that offer derivatives trading tend to have more advanced capabilities for broadcasting reference data updates via a market data API. This is because intraday instrument creation is a relatively frequent event in derivatives markets, whereas equities and bond trading platforms tend to follow a more static approach towards intra-day reference data creation.

3.3. TECHNICAL CONSIDERATIONS REGARDING DATA APIS

3.3.1. Latency

Data delivered in real-time should be delivered instantly to the market, but some participants have different requirements even in "real-time". Latency is the speed at which a trader can source market information, place an order on the market and execute an order. It can be especially important in very liquid markets, and TVs have adapted by providing multiple market data and order entry APIs aimed at different types of users based on their latency sensitivity. The lower the latency the higher the speed.

As trading technology has evolved, the focus on whether a trader is gaining a competitive advantage or disadvantage due to latency issues has intensified. However, from the use case analysis, discussed in the next chapter, it is most important to note that only a very small sub-section of market makers who are taking risk in the market, and hoping to profit from that risk, are focused on latency.

3.3.2. Time-Delayed Data

Market data can be provided on a time-delayed basis, typically anything from a few seconds to fifteen minutes. The information provided on this basis may not include the full data set that is provided in a real-time data feed. For these reasons, a delayed time data feed will be less valuable and cheaper.

3.3.3. Transmission Protocols

Due to the large volumes of data involved, market data APIs are based on **multicast transmission protocols**. The advantage of these protocols is that a stream of data can be broadcast simultaneously to a large number of consumers (as opposed to streaming the data to each consumer individually). This allows much faster transmission and fewer resources to handle it on the side of the broadcaster.

The disadvantage of this technology is that it is asynchronous, meaning that there is no concept of acknowledgement and that receipt of data packages by all listeners is not guaranteed. Neither is the concept of a "retransmission" or "re-send request".

Data packages can and do, get "lost". Market data protocols include certain features to deal with this:

- **Snapshot and Incremental Channels:**
  - Snapshot channels: broadcast a "snapshot" of the order book at pre-set frequencies (e.g. every few minutes).
  - Incremental channels: broadcast individual order updates as they are taking place. It can also include other data such as "admin" messages (i.e. session status changes, declaration of fast markets etc.), reference data updates, etc.
- **Incremental Feed Arbitration Mechanism:**
  - For additional safety, the incremental feed is broadcast via "A" and "B" services, so that in the event of a package is lost in the "A" service, the same package can be obtained in from the "B" service. If the data package is lost in both, then a recovery process would start via the Snapshot Channel.

3.4. HISTORICAL MARKET DATA

Historical market data can also be sourced from a market data feed but is different from real-time data, as it is not continuously streamed to the market. It refers to past market data, including pre-trade order and quote information, and post-trade information that needs to be stored if it is to be used for analysis. This data may include information about orders that were in an order book but were never transacted or cancellations and amendments that were made to order book data.

3.5. MARKET DATA CONSUMPTION

Market data is used in many different ways as shown in the use case analysis in this study (Chapter 4).

Those firms that provide data will often do so for economic profit, so when selling market data they will put contracts and licenses in place to govern how it is paid for and who can access and use the data. The ways in which data is used or consumed impacts upon contractual agreements and data management.

- Data feeds can be taken as 'displayed' (for human consumption) or 'non-displayed' (for machine consumption) data. There may be different fees for each type of use.
- Users may be defined as 'professional' or 'commercial' versus 'non-professional' or 'private' users. These users are usually subjected to different terms under data provider contracts and have to be identified.
- Data may be provided under an enterprise license or charged to each individual user.
- Data feeds may be combined by vendors into a bundled monthly fee for a data terminal.
- Data redistribution to other entities, e.g. clients, can require permission and incur a fee.
- Each data provider creates its own unique contracts to govern the use of data and imposes audits on the firms that use the data to ensure that they are making the right declarations of usage.

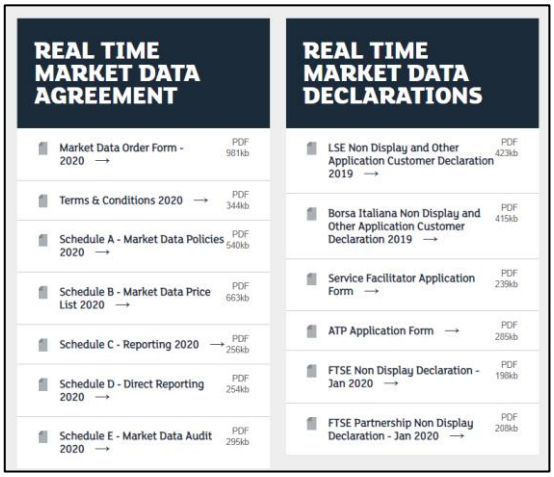
Whether data is consumed direct or via a third party, the individual provider or consumer of data will incur its own additional costs to store, transport and format that data.

As the sources, uses and complexities of data consumption increase, a firm or user may take the same data through multiple systems and thus may be charged for the same data multiple times. To counter this, a voluntary industry principle of '**Multiple Instance Single Use'** (MISU) has been established but few European TVs apply it.

The factors above contribute to making market data contracts particularly complex to manage. ESMA states that an equity CTP would have to negotiate with over 170 entities[14]. However, each venue may have multiple contracts and declarations that have to be managed and monitored (see the example below in Figure 8). If this is multiplied by the number of aggregators in Europe there are thousands of contracts or declarations that must be managed by a CTP. Even the very large firms interviewed for this study report on the difficulties in managing these contracts and controlling the resultant costs.

Figure 8: Examples of the Number of Data Contract Documents To Be Negotiated and Monitored With Individual EU Trading Venues

London Stock Exchange



Deutsche Börse

Source: London Stock Exchange Group and Deutsche Börse

3.6. CONCLUSIONS

The design of CT data needs to take many aspects of data constituents and feeds into consideration and it is important that, in assessing the viability and use cases for CT data, consideration be given not just to order, quote book and trade events, but to any other messages that are required for a correct interpretation of those events.

Latency is only an issue for a small subset of market participants, but large volumes of data need to be processed as efficiently as possible.

Data consumption is complex and requires contractual arrangements that expose potential consolidators to legal risk and additional monitoring responsibilities which must be managed for each entity that the data is taken from.

4 / USE CASES AND REQUIREMENTS FOR CT DATA

This chapter summarises the overall demand for CT data including the high-level functional use cases for a consolidated tape, the type of data required by each function and the likely numbers of users. It also provides some background context from the interviews.

CHAPTER SUMMARY

The study identifies multiple use cases for CT data across many different functions:

1. In most financial markets firms, groups of users can be broadly split into functions that align with the three lines of defence risk management principles. CT Data is therefore required by these different functions in each organisation's lifecycle:

- Functions that own and manage risk.
- Functions that oversee risks and challenge the front line.
- Functions that provide independent assurance.

2. Functions are similar across asset classes and there are multiple underlying use cases for CT data within each function:

- Issuing
- Asset Allocation
- Portfolio/Investment Management
- Pre-Trade Analysis
- In-flight Monitoring of trades
- Post-Trade Analysis/ Best Execution
- Middle and Back Office Processes/Valuations
- Funding and Collateral Management/Securities Lending
- Market Surveillance
- Risk Management
- Performance Measurement
- Regulatory Oversight
- Audit

3. The key data required is:

- Pre-trade order data with 5 levels of depth and session administration data in equities.
- Real-time post-trade data across all asset classes.
- Historical data with pre-trade and post-trade event information.

Ultra-low latency data is not required.

4 The scope of the data required depends on the user. Most participants want a comprehensive view of all trades TOTV, but some would like a subset.

5. It is estimated that tens of thousands of market participants would use the data if it were available.

6. Some context to the use cases is given from the interviews.

4.1. USE CASES

4.1.1. Three Lines of Defence

The study finds that financial market firms typically arrange their functions along the three lines of defence risk management principle:

1. **Functions that own and manage risk.**
2. **Functions that oversee risks and challenge the front line.**
3. **Functions that provide independent assurance.**

As a result, multiple functions across organisations require data. Each function and each line of defence may use different sets of data or different calculations, at different times, to create, monitor and challenge processes or results. CT data would reduce the need for each function and then each line of defence to process and clean data for their use, which currently requires considerable resources.

Figure 9: Use of Consolidated Data by Functions Organised by the Three Lines of Defence

Role		Consolidated data would be used to:	
Front-Line Roles	Own and manage risk	• Help make better investment and trading decisions. • Increase the accuracy of information to investors. • optimise capital usage.	Reduce reliance by **all functions** on the need for multiple, expensive, technical solutions or cumbersome manual processes to consume, manipulate, clean, and manage data across pan-European markets. This takes skill and resources that many stakeholders do not have.
Second-Line Roles	Oversee and challenge the front line	• Support challenge of the front line through risk management models and improved cross-market surveillance.	
Third-Line Roles	Provide wider independent assurance to the market	• Improve regulatory oversight. • Support policy decisions. • Simplify audit processes.	

Source: Market Structure Partners

4.1.2. Functional Uses

A framework to identify the functions within each stakeholder firm that would use CT data was established at the start of the study and then validated and refined during the interviews. Figure 10 below shows those functional uses, which is found to be the same across asset classes and the table is structured as follows:

- **Use Case Function**: The high-level functions making use of CT data are detailed in each row, along with their role in providing first, second and third-line risk management.
- **Stakeholder Type**: the stakeholders making use of CT data are detailed in each column according to their organisation type.

Multiple underlying use cases exist within each function and the use cases described by all the interviewees have been mapped into a normalised set of use cases per function with more detailed granularity given to the different types of requirements in equities and bonds within each use case function. A generic description of each function and detail of the underlying use cases are described in Appendix 6.

Figure 10: High-Level Descriptions of Use Case Functions for CT Data and Likely Usage by Stakeholders Based on Interview Feedback

●	Green: Large number of users that might subscribe to the tape (actual end-users not the number of firms) = High
◑	Blue: Limited number of users that might subscribe to the tape = Low
○	Grey: No users that would subscribe to the tape = None

#	Use Case Function	No. Use Cases Identified	Risk Management Line of Defence	Stakeholder Type: Issuers	End Investors	Buy-Side	Sell-Side	Trading Venues	Custody, Clearing & Settlement	Regulators	Data Analytics & Benchmark Providers*	Other
				Indicative Number of Entities Europe inc. UK)								
				9,430	15,500,000	4,366	10,576	430	64	57	Not estimated	Not estimated
1	Issuance	5	1st Line	●	○	◑	●	◑	○	◑	●	◑
2	Asset Allocation	2		○	◑	●	◑	○	○	○	●	○
3	Portfolio / Invest Management	3		○	◑	●	●	○	○	○	●	○
4	Pre-trade Analysis	6		○	●	●	●	○	○	○	●	○
5	In-flight Management	5		○	◑	●	●	●	○	●	●	○
6	Post-trade Analysis	3		○	●	●	●	●	○	●	●	●
7	Middle/Back Office & Valuations	4		○	○	●	●	○	●	○	●	○
8	Funding & Collateral Management	5		○	○	●	●	○	●	○	●	○
9	Market Surveillance	3	2nd Line	○	○	●	●	●	○	●	●	●
10	Risk Management	7		◑	◑	●	●	●	●	●	●	●
11	Performance Measurement	5		○	○	●	●	○	○	○	●	●
12	Regulatory Oversight	6	3rd Line	○	○	●	●	●	●	●	○	○
13	Audit	1		◑	○	●	●	●	●	○	○	○
14	Helping to achieve environmental action policies – reduced data processing	1	N/A	○	○	◑	◑	○	○	○	◑	○

*It is noted that some data analytics firms and vendors would simply distribute the data whilst others may use the data to create products to support the different stakeholder functions.

4.2. TYPE OF DATA REQUIRED

Users cited the following elements as critical for the success of CT data:

4.2.1. Pre-Trade Tape Design Requirements (Equities)

Participants who require pre-trade data are mostly focused on the consolidation of order event information. Consolidating quotes may be useful in the future but is not currently a priority.

- **Depth of Book to 5 Levels (explained in more detail in Chapter 3, Figure 7)**
  - This supports the ability to assess the liquidity profile of a particular security.
  - Liquidity is increasingly provided in small sizes at an increasingly fast pace and so a Level 1 BBO as a gauge for liquidity/ price is increasingly meaningless.
  - Individual firms or the CTP can calculate a volume-weighted BBO as required**.**
  - If 5 levels of depth are not given, then consideration should be given to defining minimum lot sizes for inclusion in the data to avoid a crossed book[15] of one share.
- **Detailed Auction Data:**
  - CT data should include the indicative auction price, plus the indicative volume and imbalance data during any scheduled auction call processes.
- **Inclusion of Session Administration messages:**
  - CT data should include any other types of messages that are required for the proper interpretation or handling of the consolidated data. For example, order book status updates (e.g. continuous trading, auction call, etc.), declaration of trading halts, fast markets, etc. All session data codes should be normalised.
  - Firms also require an indicator showing what type of halt is occurring. For example, a technical outage versus another reason.

4.2.2. Post-Trade Tape Design Requirements (Equities and Bonds)

- **A Real-Time Trade Events Feed** (applicable to equities and bonds)
  - This should include all trade events regardless of their source (i.e., TVs, SIs, APAs) and/or any other OTC trades subject to post-trade transparency under MiFID II.
- **Session Statistics for Exchange-Traded Instruments** *(applicable to liquid instruments that trade on a CLOB)*
  - Dissemination in real-time of a number of session-related statistics including the previous day's closing price, the opening price, the closing price, the high and low prices, the cumulative volume, and values traded. This data should be disseminated on a per venue basis as well as on a consolidated basis.
- **Historical Market Data** *(applicable to all asset classes)*
  - Data captured, stored, and made available for analysis on an ex-post basis as required. This should include pre-trade order event tick data for CLOB traded instruments.

4.2.3. Speed of Delivery and Access Requirements to the CT Data

The very low latency of equity CT data is not an issue for the use cases identified. If 5 levels of order book depth are provided, then a speed in the order of 100-200 milliseconds for equity data is sufficient. For bonds, it is 5 minutes.

Different functions identified different needs for real-time pre- and post-trade consolidated data at different times. Access to consistently available historical data is critical for many functions.

For example, pre-trade data is mostly required by users responsible for the trading functions in pre- and post-trade analysis and in-flight execution management. Other functions such as issuing and asset allocation mostly rely on historical data but at the time of an issue, capital restructuring or asset transfer, real-time data can become more important. Participants who require pre-trade data are mostly focused on the consolidation of order event information. Consolidating quotes may be useful in the future but not currently a priority.

Figure 11 below summarises the emphasis each function puts on the need for pre- and post-trade CT data.

4.3. SCOPE OF DATA REQUIRED

Overall, participants would like as comprehensive a view of the relevant market data for all the transactions in all the instruments related to equity and bond markets as possible. This includes both EU and non-EU instruments that are traded on a trading venue (TOTV). This is also important for regulators who have to oversee the markets. However, some other participants acknowledge that non-EU instruments can only be a subset of what is traded globally and so the data is less meaningful.

Some equity market participants put greater priority on the data for the most liquid instruments in the 11 largest markets where fragmentation has occurred across the EU. However, equity issuers, particularly smaller ones, want to ensure that their data is easily found in CT data, as it is not easily visible today. Bond participants are more focused on the entire European market.

In each case, participants would like to be able to delineate between the data for EU and non-EU instruments, which is currently not possible due to the lack of an official EU instrument list.

4.4. NUMBER OF USERS

Publicly available statistics, interview findings and numbers of known data users provided by data vendors during this study suggest there are hundreds of thousands of users of data across the various functions within stakeholder firms. Given the demand expressed by interviewees, it seems reasonable to estimate that tens of thousands of professional users would use a consolidated tape/s if it existed. Many more retail investors are also likely to use it. This estimate is based on three considerations:

1. **The publicly available high-level numbers about each of the stakeholders** (given in the table above and explained in Appendix 4). Even if only one individual at each of the entities above used the data, there would be 25,000 or more users of the data.
2. The information gathered from the interviews suggests that **many stakeholder entities have multiple users of data**, which can sometimes be hundreds or even thousands of individuals across one entity.
3. **The publicly available data about the US consolidated tape and the number of data users of the tape provided by data vendors.** In the US there are over 280,000 professional users of one consolidated tape (Tape A) and 2,669,941 non-professional users of data of the same tape. It is estimated that the tape reaches 145,776,063 households. It, therefore, seems a conservative estimate to assume that 25-50,000 professional users of European consolidated tape data may exist and that a greater number of non-professional investors may use the tape.[16]

4.5. CONTEXT TO THE USE CASES

Figure 12 below provides some additional context to the use cases that were identified during the interviews.

Figure 11: Demand and Priorities for The Establishment of Pre-, Post-Trade and Historical CT Data by Use Case Function (Based on Interview Feedback)

●	**Red: Functions that have the greatest need of the data on a regular basis = Critical**
◕	**Orange: Functions where the data can be important = Important**
◑	**Yellow: Functions that may occasionally use the data or where CT data is a nice to have = Low**
○	**Grey: Functions that do not require the data = None**

#	Use Case Function	No. Use Cases	Risk Management Line of Defence	Stakeholder Type: Pre-Trade Order Events (Equities Only)	Pre-Trade Quotes (Equities Only)	Administrational Events	Post-trade Real-time	Post-Trade EOD/Session Statistics	Historical Data (including pre-trade order event data)
1	**Issuance**	4	1st Line	◕	◕	◑	◑	◑	●
2	**Asset Allocation**	2		◑	◑	◑	◑	●	◕
3	**Portfolio / Invest Management**	3		◑	◑	◑	◑	●	◕
4	**Pre-trade Analysis**	6		●	◑	●	●	●	◕
5	**In-flight Management**	4		●	◑	●	●	◕	◑
6	**Post-trade Analysis**	4		◕	◑	◕	●	●	●
7	**Middle/Back Office & Valuations**	3		○	○	○	◑	●	◕
8	**Funding & Collateral Management**	2		○	○	○	○	●	◕
9	**Market Surveillance**	4	2nd Line	◑	◑	◑	◑	●	●
10	**Risk Management**	8		○	○	○	○	●	●
11	**Performance Measurement**	5		○	○	○	○	●	◕
12	**Regulatory Oversight**	6	3rd Line	◑	◑	◕	◕	●	●
13	**Audit**	1		○	○	○	○	◕	◕
14	**Helping to achieve environmental change policies – reduced data processing**	1	N/A	◕	◕	◕	◕	○	●

Figure 12: High-Level Use Case Contextual Feedback from Interviewees

Ref.	Use Case Function	Use Cases Described	Contextual Comments from Interviews
1	**Issuance**	a. Pre-Issuance Analytics & Pricing b. Improved visibility of Small and Mid-Cap Issuer Instruments c. Publication of Issuer Yield Curves d. Decisions and Disclosures e. Cash and Capital Structure Management	• Most issuers rely on their advisors and do not have access to data to challenge the advice they receive. • Corporates monitor prices for performance, employee share schemes and share buy-backs. • Some corporates are unaware of the underlying equity market structure; two large issuers in the small sample for this study were completely unaware that their shares were traded on more than one market. • Accurate data is critical for pricing of bond new issues where both the issuer and the buyer are heavily reliant on accurate pricing for benchmark bonds and credit spreads. • Post-issue, debt issuers generally rely on data provided by their primary dealers. Some would like broader market-wide information and identifiers to understand where liquidity is and who is providing it. • Smaller issuers' data is not easily discoverable as it is often excluded from data feeds for economic reasons. • ETF and fund issuers need detailed historical data to analyse product costs (rebalancing, creating/redeeming units) and the lifespan of a fund. Significant complexities exist in obtaining and cleaning the data for this. • Portfolio managers emphasise the importance of good data to ensure that product information (e.g. Key Investor Documents (KIDS – for UCITS wrappers) gives an accurate assessment of a fund's risks.
2	**Asset Allocation**	a) Asset Allocation b) Transition Management	• The lack of good underlying data for benchmarks makes it very hard to assess the performance of asset classes and asset managers. • Poor data makes it hard to estimate the cost of moving large amounts of assets or indeed whether they should be moved and/or re-allocated. Errors can result in poor outcomes for investors. • If a fund must be liquidated, then the cost must be estimated, and current data sets make it very difficult. • Good historical consolidated data over a 12-18-month period is needed. • Many assets are now being switched into ETFs – due to poor data it is very hard to know that you are getting the correct price when you submit assets in return for an ETF.
3	**Portfolio/ Investment Management**	a) Portfolio Construction b) On-Going Monitoring c) Portfolio Rebalancing	• Portfolio managers are some of the largest consumers of real-time data feeds in order to manage and monitor their portfolios. • CT data would help funds to define and achieve their investment objectives and reduce build and maintenance costs. • Understanding liquidity – how quickly positions can be unwound – is essential. A lack of information means that funds will be conservative with the allocation of capital, particularly to small and mid-cap

Ref.	Use Case Function	Use Cases Described	Contextual Comments from Interviews
			businesses. • Most portfolio management work is done using historical data. However real-time data becomes much more important during times of market stress.
4.	**Pre-Trade Analysis**	a) Pre-Trade Analytics b) Pre-Trade Analytics for Smart Order Routers (Sell-side Agency) c) Price Formation and Transparency (Buy and Sell-side trading desks) d) Block-size liquidity provision (Sell-side Trading Desk) e) Trading Strategy Research f) Trade idea generation	• One of the biggest data issues is the information asymmetry that exists between buy-side trader calculations and the calculations of brokers due, particularly in relation to lack of CT data volume information. • This mismatch leads to under or overestimations in the trading strategy by either party and a different result to what was expected and incorrect trade forecasting calculations. • The larger the order, the more important the volume information to help determine how to manage the order and where to trade it. Current processes are very cumbersome – most of the industry does not have a good picture of all the liquidity. Moving large blocks or asset transfers requires constant calculations about the realised and unrealised costs. • Firms committing capital also need very good volume data to know how quickly they can unwind risk or how much capital to commit. Without good volume data, they may over or under commit. • Retail investors only see the equity prices their retail brokers give them, which is usually the domestic exchange price, even if their order is then executed on a different venue. Retail size orders for bonds are very hard to find and must be requested manually. • Price discovery of ETFs is at the index level vs the level of the underlying constituents which is not good enough.
5.	**In-Flight Execution Management**	a) Investor's In-Flight Execution Management (Buy-side Trading Desk) b) Sell-Side In-Flight Execution Management (Sell-side Agency and Proprietary Trading Desks) c) Harmonised taxonomies d) Utility data for monitoring and risk checks e) Front Line Support	• Retail investors cannot monitor the execution of their orders as they only see the domestic market, but the trade may be being executed on other markets or against SI prices. • A CT is essential for controlling the execution and delivering best execution in real-time. Without good data, it is hard to challenge brokers whilst it is still possible to change the outcome. • Liquidity spikes that would change the trading strategy to speed up or slow down orders are often missed. • Algorithms need standardised administrative session statistics to prevent errors that come from each recipient having to normalise them. • Bond markets have no confidence for knowing when to halt a market in a bond. • It would be extremely helpful to have flags indicating reasons for a trading outage that can be disseminated to the whole market as today they are sent to direct members of an exchange but often have no information about outage reasons.

Ref.	Use Case Function	Use Cases Described	Contextual Comments from Interviews
6.	**Post-trade Analysis**	a) Transaction Cost Analysis (TCA) b) Post-Trade Analytics, block trades c) Provision of Best Execution	• Most retail investors appear to be unaware of the concept of best execution and rarely challenge their brokers or have the right information to do so. • Retail brokers do not provide their clients with the depth of information they would need to challenge execution because it is too expensive to do so. • Institutional investors are struggling to get the data they need to build a consolidated view of both the order and trade events across the entire market. This makes it hard for them to challenge their brokers about transaction costs. • Consolidated data would help venues to publish their quarterly reports on the execution quality achieved and investment firms, including retail brokers, to publish their annual disclosures on why they have selected their top 5 execution venues. • Closing prices in the bond market often bear no relationship to the trade price. • Issuers undertake best execution analysis on share buy-backs. Data vendors see frequent price reversions during these times.
7	**Middle and Back Office Processing and Administration**	a) Valuations and Fair Price Adjustments b) Reconciliations c) CSDR Penalties Calculation d) Initial Consistency checks/Product Improvements	• Consolidated data would be used every day to calculate Net Asset Values of funds and position management for firms that commit capital. Everyone has different valuation data. The same portfolio can have three different valuations as the custodian and asset manager have different numbers. A bond asset manager's administrative department can spend up to a month trying to reconcile the differences. • The ManCo/Depository challenge on NAV pricing data is evidenced by the recent Woodford / Way Fund Manager scenarios where funds were suspended because of exposure to illiquid stocks. • There is concern that the data being used to validate prices is not accurate and/or independent. • The potential scale of valuation errors was highlighted. It was also flagged that these errors may not be identified until the assets are offered for sale and the resulting costs can run into billions. • Firms want to validate their CSDR Penalties invoices. It is not yet always clear which price they should use to match that being used by the CSD calculating the fine and once this is resolved this could be a source that they are not currently paying for. A standard price would make the validation process much simpler and avoid the need to buy new price feeds.
8	**Funding and Collateral Management (including non-cash collateral)**	a) Initial and Variation Margin calculations b) Securities Lending and Collateral Management c) Less Liquid Instruments e.g. ETFs for lending and use as collateral	• CCPs use instrument price and liquidity for their initial and variation margin calculations. OTC counterparties also use the same data to determine collateral requirements. A full set of post-trade data for all instruments cleared would provide better data and improve the valuations for less liquid assets. It would also help to assess other instruments that could potentially also be centrally cleared. • Securities lending processes are designed for equities, not ETFs. Lenders use data from primary listing venues to provide a proxy for the total market. ETFs do not work this way – EMEA ETF liquidity is distributed across multiple trading venues. The current infrastructure does not work effectively for ETFs.

Ref.	Use Case Function	Use Cases Described	Contextual Comments from Interviews
		d) Standardised Collateral Agreements e) Credit / Counterparty Risk, OTC Derivatives	• The same applies to the use of ETFs as collateral. The amount of an asset that can be used as collateral is determined in part by the daily trading volume. If the collateral receiver is only seeing a subset of the volume this will significantly restrict the amount that can be deployed, thereby potentially increasing funding costs for the asset class. And this assumes that the ETF has been approved for use as collateral - minimum liquidity levels may be a prerequisite and only met if full access to data is available.
9.	**Market Surveillance**	a) Detection of Insider Trading b) Detection of Market Abuse/Manipulation by the market c) Detection of Insider Trading/ Manipulation (Regulatory Bodies)	• Market surveillance needs confidence in the data it is using – it cannot just replicate what has happened in the front office and needs to challenge through its own data sets. It is extremely hard to replay cross-market scenarios without consolidation of data. Reconstruction of historical order and event data is also critical. • Market abuse rarely occurs in one venue. It is usually across multiple venues and often in multiple, linked instruments. Financial intermediaries and regulators need a complete set of data to discharge their obligations. CT data would help by providing a single source of data for Europe. • Good historical data is needed to be able to reconstruct scenarios, but many exchanges cannot provide good historical data when requested or keep data in inconsistent formats. Comparing current data sets is hard.
10	**Risk Management**	a) Liquidity Risk Management (Buy-side) b) Fund Manager Oversight c) Liquidity Risk Management (Sell-Side) d) Liquidity Risk Management (Buy-Side) e) Market Risk Management f) Market Risk - Fundamental Review of the Trading Book (FRTB) g) Credit / Counterparty Risk, OTC Derivatives h) Operational Risk - Back-up Source of Market Data	• The industry increasingly runs on models. Independent risk oversight of these models depends on access to independent, complete, and accurate data sets. • Models increasingly need years of detailed historical data. Historical data is often not available or consistently stored by TVs. • Liquidity risk management is critical and feeds back into portfolio construction and monitoring processes or into the capital commitment processes of the sell-side. • Second-line challenge requires good data to ensure errors are identified or operational behaviours can be challenged. • FRTB is a huge cost to market firms and regulators. If CT data existed now it would significantly reduce the project cost and complexity for both firms and regulators. • Without good, easily accessible cross-market data, managing risk is often a manual process. This becomes very difficult in stressed market conditions.

Ref.	Use Case Function	Use Cases Described	Contextual Comments from Interviews
11	**Performance Measurement, Evaluation, and Attribution**	a) Calculation of the rate of return b) Index/benchmark creation and pricing c) Macro Performance Attribution d) Micro Performance Attribution - Equities e) Micro Performance Attribution – Fixed Income	• Benchmarking data, particularly in non-equities is considered very poor. Equity data is considered reasonable. • CT data would be an additional data constituent to give more confidence in a benchmark. • Some providers collate market prices all day long from their sponsors, but the quality of the data is bad. Consolidated data for credit markets would be particularly helpful. • A portfolio can often have three different valuations - a custodian can have one value and a fund manager can have a different value and a different return based on prices from a different set of data. • A bond manager's portfolio accounting department can spend months reconciling data differences. • CT data would contribute to assessments of a fund manager's skills and help improve measurement.
12.	**Regulatory Oversight / Policy**	a) Update of regulatory metrics/ thresholds and obligations b) Ongoing monitoring of regulatory requirements c) Cross-market scenarios involving NCAs d) Forthcoming regulatory proposals e) Harmonised implementation of regulatory requirements f) Regulatory oversight	• It would have been helpful if ESMA had access to CT data before it implemented the liquidity thresholds for bonds and restrictions on dark volumes. • CT data would better help ESMA perform calculations for the many RTSs and make improvements in future. • NCAs rely on phone calls between each other for regulatory trading halts and have to manually request data for cross-market issues. The amount of data they can request from each other may be limited due to time and resources. • It is virtually impossible for regulators to fully undertake cross-market surveillance on a pan European basis. Market abuse across venues in multiple jurisdictions is very hard to detect. • NCAs find it hard to access and/or create a complete picture of the market
13	**Audit**	a) Audit oversight	• CT data would simplify the audit process, particularly for reconciling data using an independent source.
14	**Environment**	a) Reducing Duplication of Processing	• A CT would reduce duplication of processing across multiple organisations and data centres across Europe.

4.6. CONCLUSIONS

The use cases described by participants make clear that there is an urgent need for both pre- and post-trade consolidated data across asset classes in many different functions. Front line users note that it would help them to make well-founded and better investment decisions, provide accurate information to investors and optimise the use of capital. Second-line defence personnel, such as surveillance and risk managers, speak of the benefits arising from having more accurate and consistent data to challenge the front line. Third line functions report that access to better data is key to performing oversight roles, whilst regulators would benefit from having improved and comprehensive data to aid policy decisions and help monitor the market.

Demand is therefore high, and tens of thousands of users are likely to use CT data if it is delivered to meet their specified requirements. The design of CT data needs to take into consideration all of the user requirements outlined in this chapter or they may not be able to use it, which would undermine the value of any consolidation efforts.

Consolidated post-trade data and good historical data are an absolute priority but consolidated pre-trade equity data is also very important. This CT data would provide users with better data and result in improved processes and outcomes for all the use cases identified in this chapter. Low latency data, which is only relevant to equities, is not critical for most participants.

5 / THE BENEFITS OF CONSOLIDATED DATA

This Chapter describes the benefits that would accrue to a broad set of stakeholders, including end investors, if CT data were readily available in Europe. This is based on user feedback and research.

CHAPTER SUMMARY

1. The total benefits of CT data cannot be quantified but the **firms that are responsible for managing and trading trillions of euros of assets on behalf of investors report that the lack of CT data means they currently have to rely on sub-optimal data** when seeking to do their jobs for those investors.

   Examples of the benefits of CT data that could result in savings to end investors and improve the functioning of capital markets are:

   Improved trade outcomes: More accurate forecasting of trading costs and broadening awareness of liquidity options.

   a) **Independent data for detecting errors and misdeeds**. Helping uncover issues, which may otherwise remain undetected.

   b) **Increased pricing accuracy:**
      - Improved **portfolio valuations** for investors who are often misled by inaccurate data.
      - Improved **benchmark calculations**, which are used to judge performance.
      - Better **derivative pricing**.

   c) **Improved liquidity risk management and capital allocation** processes:
      - Underestimates of liquidity risk can cause harm to end investors who may get trapped in failing funds.
      - Overestimates of liquidity risk may also come at the expense of the capital raiser if capital is allocated too conservatively.

   d) **Promotion of innovation, competition and lit markets** would all be greatly facilitated by CT data providing visibility of all available liquidity:
      - The liquidity provision and listing models of incumbent trading venues are being forced on the market and alternative liquidity provision and listing models cannot emerge.
      - Brokers are not being held accountable for their smart order routing decisions and innovative venues may lose out.
      - Lack of clean data is misleading participants to trade away from lit markets.
      - Data processing costs are leading to industry concentration.

5.1. THE IMPORTANCE AND MATERIALITY OF GOOD DATA

As described in the previous chapter, market participants and end-users have highlighted a multitude of use cases where they would benefit from CT data.

Quantifying the total value of these benefits to investors is impossible. However, it is known that the firms that are responsible for managing and trading trillions of euros of assets on behalf of investors report that the lack of CT data means they currently have to rely on sub-optimal data when seeking to do their jobs for those investors.

Some key examples that further illustrate and quantify the benefits of CT data are set out below.

5.2. IMPROVED TRADE OUTCOMES

A key use case identified for CT data is the pre-trade analysis function, which is used to forecast the estimated cost of a trade and hence helps to determine the appropriate trading strategy.

5.2.1. Benefits of CT data for Pre-Trade Analysis in Equities

Liquidity is not always immediately available in the market, particularly for larger orders (i.e. for most institutional investor orders), basket orders (i.e. orders in multiple instruments) or even small orders for illiquid instruments. In this case, the trader or investor will use pre-trade analysis to determine how best execute the order.

The analysis draws upon a number of key data points. These include the current price and bid-offer spread (pre-trade data) plus a view of the prices and volumes that have been recently executed in the market and where they have been executed (post-trade data). A sophisticated trading strategy will also consider the risk of the instrument(s) to be traded, which is typically measured by comparing the prices at which the instrument has historically traded relative to a benchmark index to determine its volatility (historic data).

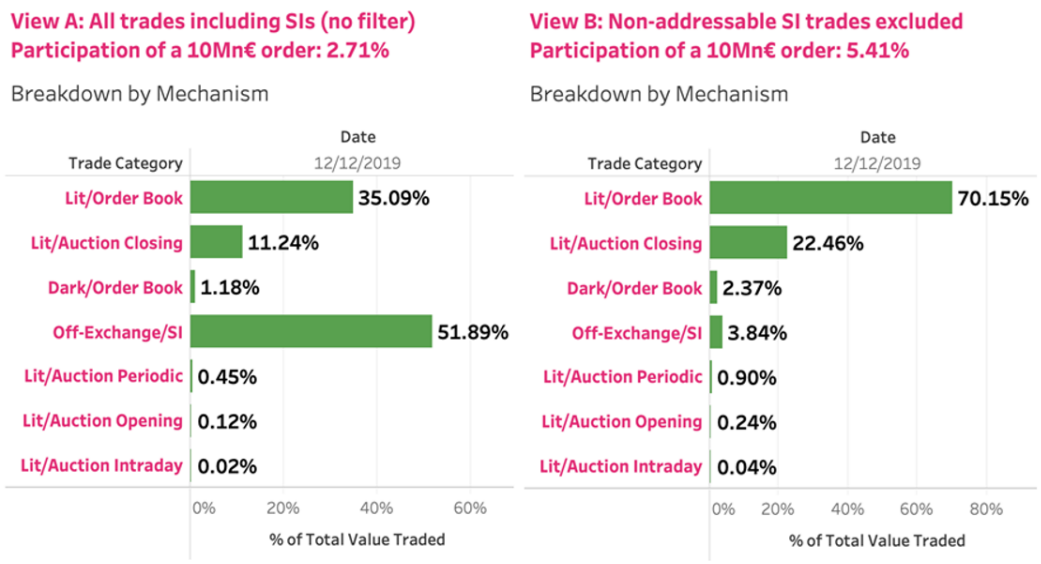
Therefore, pre-trade analysis requires a complete and accurate consolidated data set. With this information, the trader or investor can forecast, and thus optimise, the expected cost and risk of the trade versus the timeframe required to complete the trade. Traders also report increasing their use of automated and algorithmic trading strategies to determine their trading strategy and complete their orders. These strategies are almost wholly dependent on data and therefore are only as good as the data that goes into them. When the data is poor or incomplete then the forecast will be inaccurate. The likely outcome will be a sub-optimal result as the trade is executed at the wrong speed; price and/or volume and the additional cost or risk is borne by the end investor.

An example of the impact of sub-optimal data is given in Figure 13 below. This examines the impact on a single German stock, Munich Reinsurance Company, on 12/12/2019 which was considered a relatively “normal” day in the markets.

Under View A, an investor with a typical institutional sized order of €10 million is presented with a gross (uncleaned and unfiltered) set of data that show liquidity available to the market. The order appears to represent 2.71% of the day’s reported turnover in the stock and the data suggests that the trading strategy should be to route 35.09% of an order to the lit markets continuous trading, 11.24% to the auction, 1.18% into a dark book and 51.89% to a SI.

For View B, billions of rows of data have been filtered, cleaned, and manipulated under a series of subjective interpretations (described in Chapter 6) to establish a truer picture of where the instrument typically trades. Under this view, the order represents 5.41% of the adjusted volumes and the investor is presented with a very different picture that suggests that 70.15% of the order (vs. 35.09%) should be routed to the lit order book and only 3.84% (vs. 51.89%) to an SI.

Figure 13: Analysis of Total vs. Adjusted Volumes Traded for Munich Reinsurance Company on 12/12/2019



Source: Big XYT

In this example, an investor with View A would likely find that a large portion of the trade routed to the SI has not completed and they may have missed good liquidity on the lit market.

The use of a poorly informed trading strategy has also increased the risk of the trade. The reasons are two-fold. First, liquidity providers with more efficient data processing and cleaner data may detect and trade against the order. Second, by taking longer to complete there is an increased risk that the price moves against the order. This is a common complaint that traders say occurs daily.

This example also shows that complete and accurate CT data would increase the likelihood that more trading activity is channelled to lit markets and thus more investors could participate in the price discovery process.

5.2.2. Estimates of Cost Savings to the End Investor from Accurate Equity Data

Interviewees at equity asset managers were asked to estimate the cost to their annual trading strategies of not having complete and accurate CT data to forecast trading costs. Whilst it is not possible to give an exact number most were happy to give a range estimate. The majority estimated costs or slippage in the range of 0 to 1.0 basis points although some rated it even higher with one or two putting the figure above 5.0 bps.

Significantly, each respondent's perspective was based on **how clean they thought their own data was**. The larger the firm and the more extensive the resources to compile and clean the data, the lower the estimate might be.
By applying these estimates to the annual traded value of European equities, the total adverse cost can be calculated. The table below shows that there are substantial amounts that can be potentially saved by giving investors and traders access to complete and accurate CT data.

Figure 14: Estimated Cost in Basis Points of not having a Consolidated Tape in Equities

Basis Point Estimate of Slippage	**Number of Asset Management Trading Desks Responding in Each Category**	**Cost In €billions** (Based on Annual Value Traded In 2019)
Impossible to estimate	1	-
0 – 0.25 bps	15	0 - 0.53
0.25 – 0.5 bps	12	0.53 - 1.06
0.5 – 1 bps	5	1.06 - 2.12
1 -1.5 bps	3	2.12 - 3.18
1.5 bps – 2 bps	4	3.18 - 4.24
2-3 bps	0	4.24 - 6.36
3-5 bps	1	6.36 - 10.61
Above 5 bps	2	10.61

Source: MSP Interviews and Alpha Forum Conference, Feb 2020

5.2.3. Benefits of a Consolidated Tape for Pre-Trade Analysis in Bond Markets

The pre-trade analysis process for bonds follows a very similar process as for equities.

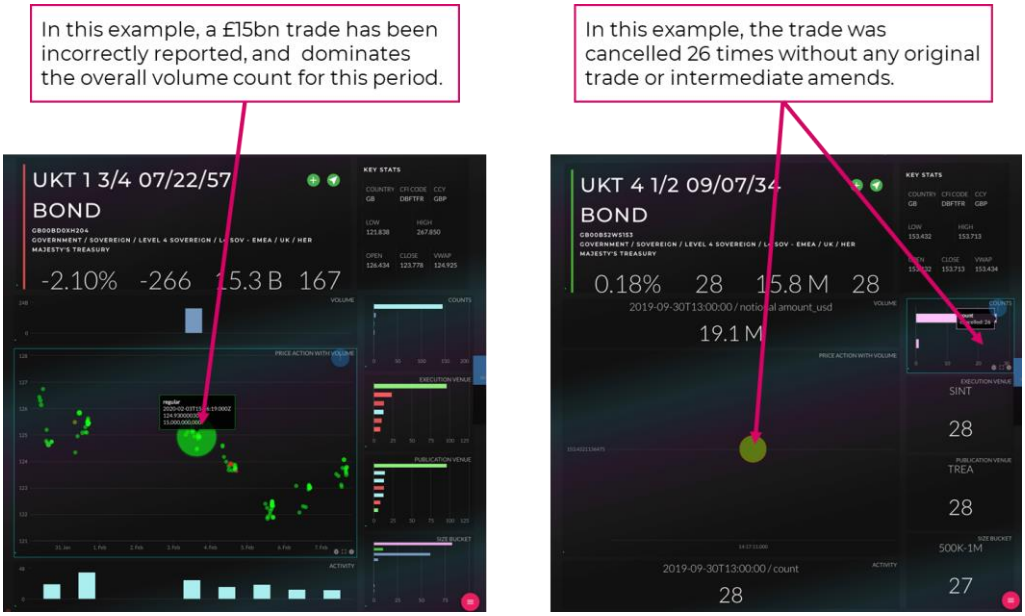
Generally, there is less data available for a bond than for an equity as they can be less liquid, and the market is mainly quote-driven where quotes are not firm. This means that the prices reported 'on-screen' or in a data tape may not represent the prices at which the instrument can be traded.

However, just because there is less data, this does not mean that there is less need for it. Respondents to this study frequently commented on the difficulties in sourcing good bond data.

The analysis for bonds often looks at an asset class or the sector level. For example, the price of a bond can be estimated by looking at the credit spread of similar bonds over the equivalent benchmark bond. Additionally, as the 'on-screen' prices may not represent the price at which the bond can be traded, there can be less focus on pre-trade data and more reliance on post-trade data. The analysis for bonds is also very sensitive to the accuracy of the data and **a single data inaccuracy has the potential to completely distort the picture of the market**.

For example, an independent technology provider noted that they regularly see trade reports that are ten times the actual volume due to the wrong notional flag being used. Further examples are highlighted in the diagram below. The trade on the left was incorrectly reported and displayed to the market as a £15 billion trade. Whereas the trade on the right was cancelled 26 times without any original underlying trade report being submitted. These examples show how incorrectly reported trades significantly distort the real volumes in the market and the data used for forecasting by traders. CT data with stricter data quality controls and enforced technical protocols could avoid such costly errors.

Figure 15: Examples of Incorrect Trade Reports or Trade Cancellations that have Impacted the Market Volumes and Value Traded



Source: Ediphy

5.2.4. Estimates of Adverse Costs to the End Investor from Inaccurate Bond Data

Interview respondents at bond asset managers were also asked to estimate the cost in basis points to their annual trading strategies of not having complete and accurate data to properly size orders. These estimates are included in the table below although the absence of any reliable data for the traded value in bonds only serves to highlight the need for CT data and makes it impossible to quantify the cost in absolute terms.

Figure 16: Estimated Cost in Basis Points of Not Having a Consolidated Tape in Bonds

Basis Point Estimate of slippage	Number of Asset Management trading desks responding in each category
Impossible to estimate	0
0 – 5 bps	10
5-10 bps	4
30-50 bps	2
50 -1 bps	1

Source: MSP Interviews and Alpha Forum Conference

5.3. MORE ACCURATE PRICING, VALUATIONS AND BENCHMARKING

5.3.1. Portfolio Valuations

Reconciliations are a daily requirement across the industry. For example, asset managers will reconcile their positions and prices to their custodians and fund administrators for good governance and to ensure that fund market values are fair and accurate. However, these reconciliations are only as good as the data used to reconcile against. If data is poor or inaccurate then fund values could be calculated incorrectly, and differences can take weeks to resolve.

This is particularly problematic if investors have invested or disinvested from a fund at the wrong market value, which is very likely if, for example, the fund is priced daily.

The examples below highlight the potential impacts on both retail and institutional investors arising from incorrect bond prices.

Figure 17:

Example of The Impact on Retail Investors

Two brokers *that aggregate and execute retail orders were interviewed for this project. They both* ***identified an issue with accurate end-of-day bond valuations for retail customers*** *and both independently stated that the order of magnitude of* ***the error on valuations could be around 5%.*** *This means that a bond investor with bonds that appear worth €100,000 may be able to only get €95,000 euros or could get €105,000 euros on liquidating their portfolio.*

One of these retail aggregators holds €40 billion euros of assets for retail customers. Even if only 10% of the holdings were in bonds, the 5% variation on the value of their total investments could be €200 million.

Figure 18:

Example of The Impact on Institutional Investors

An investment advisor *that was interviewed for this study* ***cited an actual transition of assets*** *that was undertaken* ***where the wrong valuations had a significant impact*** *on the end value of the fund. At the outset, the total fund was valued at approximately €15 billion in assets which included a tranche of corporate bonds valued at €2 billion based on the data at the time.*

These corporate bonds turned out to be completely illiquid, were not saleable and thus the value was in fact zero. *The fund was therefore only worth €13 billion when it was actually realised.*

These impacts could be avoided if robust CT data were available to support timely and accurate pricing, analysis, and valuations.

5.3.2. Better Quality Benchmark Calculations

Benchmarks are used to measure and assess the performance of investment strategies and investment products. These benchmarks are often viewed at the top, or index level, but many users would also like CT data to provide greater visibility of the data for the underlying components of the index. This would allow them to better assess, or understand the reasons for the benchmark index's performance, and the scope for potential improvements.

Ultimately the benchmarks are only as good as the data that is used to create them. Many users consider this poor, particularly in bonds.

CT data would greatly improve the pricing of benchmarks and increase their usefulness in assessing investment returns.

5.3.3. Improved Derivative Pricing

Derivative pricing is based on the price and certain key data points of the instruments or securities underlying the derivative. Complete and accurate CT data would, therefore, improve pricing processes, transparency, and accuracy. This applies particularly to OTC derivatives where counterparties need to agree on derivative pricing to determine any collateral requirements.

5.4. INDEPENDENT DATA FOR DETECTING ERRORS & MISDEEDS

A robust system of checks and balances is in place across financial services firms to minimise potential errors and identify any misdeeds. This activity is typically undertaken by second and third-line functions that are partially or fully independent of the front-line asset management or trading process. However, these functions are one step removed from any direct involvement in financial markets and, as such, they are fully reliant on sourcing good data to perform their checks. If the data is poor, then errors or misdeeds may not be spotted until a later date or in extreme cases may not be spotted at all.

The example below relates to an enforcement action taken by a European NCA. In this example, a transition manager, trading on behalf of multiple clients, over a period of time, applied unauthorised mark-ups (i.e. charges) to certain bond transactions. The clients were not aware that the transition manager was applying these mark-ups and ultimately the misdeed was only discovered when a third party compared the prices being applied to the bond transactions against the independent CT data that is available in the United States. As there is currently no European equivalent of this CT data it is possible that the misdeed would not have been discovered if the transition manager had only traded European bonds.

Figure 19: Example of the Benefits of Independently Prepared Consolidated Data – UK Transition Manager fined £32 million+ by National Regulator[17]

In 2014, the FCA fined a large UK-based transition manager £32,692,800 *(which was discounted to £22,885,000 for early settlement)* ***for overcharging six clients a total of US$20,169,603 during a 14-month period.***

Transition managers assist clients to support structural changes to asset portfolios with the intention of managing risk and increasing portfolio returns. They may be required when a client needs a large portfolio of securities to be restructured, or when a client decides to remove or replace asset managers. In this case, the transition manager in question applied a series of substantial and unauthorised mark-ups to the transitions which included assets in both equities and bonds.

The discovery of this overcharging was only made possible through the existence of the US consolidated bond data tool TRACE. *When one of the six clients used an independent consultant to verify the US trades, the consultant could see the differential between the publicly available bond pricing information in TRACE and the mark-ups on the client's trades which had not been expressly agreed. It was impossible to reconcile such data in Europe because there was no consolidated tape of trades but once the TRACE data had highlighted the problem, the investigation showed that the mark-ups had been applied on both US and European transactions.*

The transition manager's second-line defence systems and controls were also subject to criticism. The findings were that the annual Compliance Monitoring and Testing Programme failed to identify any of the control issues in respect of the transition management business during that period. Also, the risk management function did not identify the operational control weaknesses which existed during the relevant period, such as the lack of adequate processes overseeing the correct application of commission rates and spreads for TM trading.

A lack of accessible, external independent data may have made the internal information available harder to challenge. *Compliance and risk managers interviewed for this report have repeatedly highlighted the importance of easy access to independent data to enable them to do their jobs more effectively.*

Similarly, a recent US class-action case[18] alleges that a number of banks unreasonably restrained the trade of odd-lots of corporate bonds in the secondary market in order to maintain wider odd-lot spreads, and thus higher profits for themselves collectively. It also alleges that these banks deliberately boycotted market evolutions that would increase pre-trade pricing transparency for retail investors through electronic platforms. The evidence being used to support this case is from TRACE data.

5.5. IMPROVED LIQUIDITY RISK AND CAPITAL ALLOCATION

This study especially highlights liquidity risk management as a priority use case for CT data.

Firstly, managers in front line functions should assess the short and long-term liquidity of the assets in a fund or on their balance sheet to understand how quickly they may be able to buy and sell assets. For asset managers, this trading is in response to inflows and outflows or a change in their investment views. The analysis requires a good knowledge of both the historic and projected liquidity in each individual instrument to ensure the right calibration at a total fund level. If these calculations are wrong, it can have significant consequences for investors or result in opportunity costs for issuers.

Furthermore, this analysis is dynamic because asset managers constantly need a clear and on-going picture of liquidity. For example, if a fund experiences unexpectedly large withdrawals, the impact can be immediate and severe. In some cases, the fund may be forced to sell its most liquid assets first so that the remaining investors are left owning the least liquid assets in the fund. In an extreme scenario, this can even threaten the viability of the fund.

One example is the UK-based £3 billion Woodford Equity Income Fund. This fund was ultimately suspended in June 2019 after being overwhelmed by a series of large withdrawals, which left the remaining investors trapped in the fund. Prior to this, the fund was considered one of the most successful launches since the financial markets crash in 2008 and had attracted billions of assets from both retail and institutional investors. However, in an attempt to generate the returns expected on the fund, the asset manager had increasingly invested in less liquid securities. This quickly became problematic when the inflows turned to outflows and the asset manager was forced to sell positions.

The internal and external risk functions failed to mitigate this liquidity risk. One of the reasons given for this failure was the apparent difficulty in sourcing good pricing data for the illiquid holdings in the fund, given their very large size.

The cost of this failure was significant. In the period between the fund's launch in June 2014 and the announcement of its wind-up, hundreds of thousands of investors lost 18%, whilst the FTSE All-Share rose 33%. Moreover, the fund's value continued to fall post-suspension, until specialists were called in to close the fund. This closure is not yet complete but modelling, conducted by private equity specialists PJT Park Hill in November 2019 forecast a base case loss of 32.5% and a worst-case loss of 42.6%.[19]

Whilst CT data would not have prevented the fund's collapse, this example highlights the importance of good data to understand and manage liquidity and mitigate risk. Operators of venues that cover less liquid markets highlight the difficulty of getting their data seen without CT data.

If Woodford is an example of an asset manager being too aggressive then it is also possible that asset managers can be too conservative and thus miss out on investment returns. For example, some asset managers report a lack of confidence in liquidity data that leads to lower levels of investment in small and mid-cap securities. This not only deprives end investors of returns but ultimately leads to a lower allocation of capital to small and mid-cap issuers.

5.6. THE PROMOTION OF INNOVATION, COMPETITION AND LIT MARKETS

5.6.1. Remove Barriers to Entry for Innovators and Reduce Concentration

For equities, interviewees generally noted that valuations of liquid instruments are less of an issue as firm prices are constantly available. As such, many only take their data from the domestic incumbent TV as a reasonable, if incomplete, source of prices and volumes for the market as a whole.

These TVs are, therefore, in the advantageous position of being able to sell their data to end-users as a proxy for the entire market, which effectively means that their liquidity provision, listing and data provision models are by default also being forced onto the market.

Competitors who offer, or would like to offer, alternative models for listings and secondary market liquidity struggle to get their prices and volumes seen and may lose the incentive to invest resources to innovate. To compound the issue, data vendors may not distribute their data even if it is free because the data vendors still have to spend money on the data interfaces and are not guaranteed a return for their efforts.

This dynamic is a barrier to competition and innovation and the overall effect is to perpetuate the existing structures as the only meaningful liquidity provision, listing and data models.

Investors and issuers are often not aware of the underlying market structure and the choice of primary and secondary market options available. During this study, some extreme examples were given, whereby two issuers were unaware that their shares were traded on other markets away from the main exchange. Additionally, retail brokers report that their customers are also unaware that shares can be traded away from the main exchanges even though any reasonable definition of best execution would require that other trading venues be in scope for execution.

5.6.2. Promoting Lit Markets

The examples given in Figure 13 above and Figures 24 and 25 in Chapter 6 show that current market statistics can provide a distorted and misleading view of market liquidity. In these examples, the liquidity provided by the SI was significantly overstated when in fact there was more liquidity available on the lit venues. Based on current market statistics, if data is cleaner and CT data was more readily available then more trades would likely be routed to lit books.

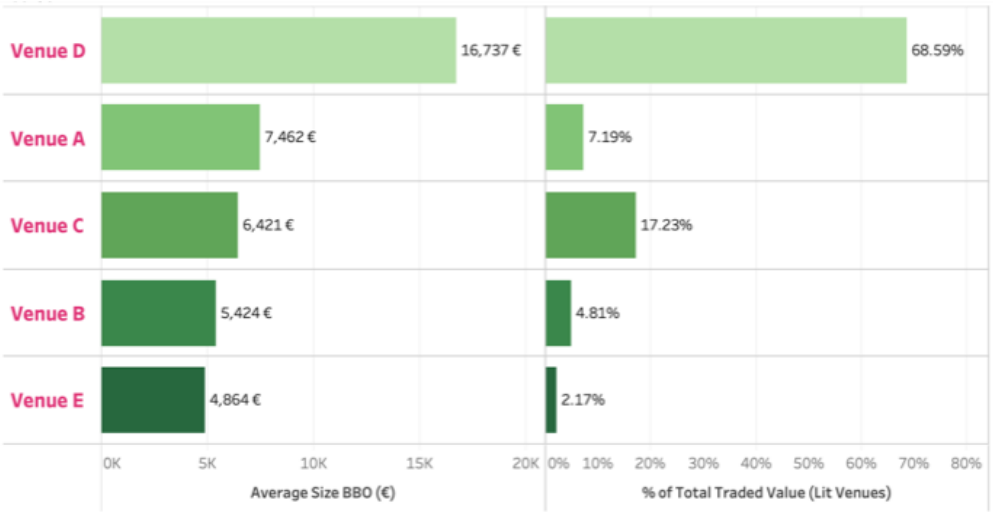
5.6.3. Stronger Data and Accountability for Smart Order Routing Decisions

Institutional and retail investors are increasingly likely to execute their orders through a broker or a broker's automated or algorithmic trading platform. As a result, almost all investors depend heavily on their broker to seek out the best prices and venues to achieve best execution.

However, even large institutional investors report that they cannot see all markets and therefore cannot audit or challenge brokers to ensure that their order routing decisions are in the investor's best interest.

For example, Figure 20 below shows the average liquidity available at the BBO at various TVs versus the actual turnover at those TVs for the main Danish index in December 2019 across completely fungible venues (i.e. where clearing and settlement are exactly the same). Venue D had the highest average liquidity available at the BBO and consequently, the largest value of transactions was executed there. However, Venue A, which had the next highest average liquidity available, accounted for a much smaller share of the turnover than Venue C. In other words, Venue C had less liquidity but secured considerably more turnover. There can be other factors that brokers consider when making smart order routing decisions but without good CT data clients are unable to challenge or assess these decisions or outcomes.

Figure 20: Liquidity vs. Turnover for DK25 Index in Dec 2019



Source: Big XYT

5.7. INCREASED ACCESS TO LOW-COST INVESTMENTS FOR SAVERS

Exchange-Traded Funds (ETFs) have become increasingly popular with both institutional and retail investors in recent years and now represent one of the fastest-growing segments of the investment and savings market. Originally developed in the US, they were initially focussed on equities but have since expanded to include government, investment grade and high yield bonds across all major financial markets. Additionally, because the underlying securities are usually managed on a passive basis, the fees to end investors are typically much lower than for similar actively managed products.

An ETF is a hybrid instrument that can be traded in two ways.

1) **Primary liquidity:** similar to a unit trust, the units are created or redeemed in response to investors either investing or disinvesting from the ETF. This process is managed by Authorised Participants (e.g. broker-dealers) who effectively buy or sell baskets of the securities held by the ETF in coordination with the creation or redemption of shares in the ETF.

2) **Secondary liquidity:** similar to equities, the shares in an ETF are traded on an exchange or other trading venue.

The combination of both primary and secondary liquidity is innovative and critical to the success of an ETF. It not only provides investors with two avenues for sourcing liquidity but trading in the secondary market effectively acts as an arbitrage mechanism to ensure that the price moves in line with the value of the underlying securities throughout the trading day.

Despite their success, the absence of CT data impacts the secondary market pricing and liquidity of ETFs. By their nature, ETFs are broad investment products that are made available across multiple exchanges and trading venues to improve their accessibility to end investors. As such they do not have a link to a home exchange in the same way that a listed company is usually linked to the main exchange in its country of domicile.

This can be problematic if participants economise by only taking data from the main domestic exchange as a proxy for the whole market. The result is that different investors may have a very different picture of the pricing and liquidity available to them. This not only reduces the attractiveness of these products to potential investors but also creates inefficiencies in trading costs and the arbitrage mechanism that helps keep the price of the ETF in line with its underlying securities.

The information below shows how the average spread (from daily time-weighted aggregations) varied across venues for the same iShares Stoxx 50 ETF from the start of 2020 to May 2020. It also shows how it changed dramatically with volatility quarter on quarter. Investors in Italy, the UK, Germany, or Switzerland would have very different information to inform their investment decisions.

Figure 21: Example of ETF Spread Difference in Basis Points for iShares Stoxx 50 ETF, EUR.DIST across European Trading Venues during Q1 2020 and Q2 (to May 2020)

Spreads Across Trading Venues

Venue	AtTouch VWAS in bps 2020-Q1	2020-Q2
Aquis Exchange	8.58	24.23
Borsa Italiana	5.33	8.34
Cboe BXE	10.76	28.42
Cboe CXE	10.28	25.44
Euronext Amsterdam	6.67	12.38
LSE	7.39	14.61
SIX Swiss Exchange	14.33	27.21
Turquoise	14.15	28.58
Xetra	3.75	7.25

Source: Big XYT

The absence of CT data also contributes to concentration in ETF providers. This can result from individual ETFs appearing to be less liquid and more expensive to trade (as shown in the example above) and from the large-scale data processing capabilities required to be an ETF provider. The result is that market share is highly concentrated. Of the companies that issue ETFs globally almost 75%[20] of the market is captured by 3 US ETF providers. In Europe, there are other issuers offering ETF products, but the market is still concentrated with the 3 largest providers accounting for 65.9% market share.

5.8. IMPROVED REGULATORY CALCULATIONS AND REGULATORY POLICY MAKING

To fulfil its supervisory obligations ESMA has to source, cleanse and manipulate significant amounts of data to make certain regulatory calculations. This requires considerable resources, effort, and time delay to complete. Alternatively, if this information were available from CT data then ESMA would not have to cleanse the data and could focus its limited resources elsewhere.

The example below highlights the effort required to cleanse non-consolidated data and the discrepancies that currently exist.

The RTSs[21] in MiFID II specify that a non-equity CTP must cover both:

- 80% of the cumulated volume of transactions as defined within Annex II of RTS 2 reported by all APAs and/or trading venues in the Union over the previous 6-month for the relevant non-equity asset class; and
- 80% of the cumulated number of transactions reported by APAs and/or trading venues in the Union over the previous 6-month period for the relevant non-equity asset class.

ESMA does not currently calculate the CTP threshold requirements but recently voluntarily aggregated data[22] from the NCAs to help market participants in the performance of the SI test[23]. This calculation is integral to the SI regime, but it is also the only data available to assess whether a potential consolidator could meet the 80% coverage requirement for non-equity CTP.

Ediphy, an independent data analytics provider in bond markets, interviewed for this study, is interested in providing a CT service. It compared the aggregate numbers as published by ESMA to its own aggreged data, which was taken directly from the data sources. Its analysis found the following discrepancies:

- ESMA's data showed that 57,000 ISINs out of a total of 334,000 ISINs in the ESMA file appear to have some activity but Ediphy's results suggest only 41,000 of those 57,000 ISINs had activity during the same period.
- Ediphy also noted trading activity in 1,300 ISINs where ESMA did not note any activity
- To calculate whether it could consolidate 80% of the data for the CTP coverage, Ediphy sums the number of trades it records for individual ISINs and divides that by the number of trades that ESMA records in the same ISIN. The result is that Ediphy believes they can consolidate approximately 70% of the market on 50% of the ISINs. However, it does not know where to get the remaining 30% of data from on the instruments where it already has 70% of the data or where to get the data for the 50% of instruments that it has less than 70%. ESMA says they would need to ask each individual NCA for more information about where the data came from to build the picture and ascertain what is missing.

On the analysis above **it is possible that ESMA's numbers are not accurate and are not capturing all the data**. In the absence of CT data, this is understandable. Nevertheless, it is therefore unlikely that any CTP can be confident that they have met the 80% threshold requirement. There is also a risk that other important regulatory calculations that ESMA is required to undertake may not be accurate.

Where there are discrepancies in the data there is also a danger that, without official data, interested parties can fill the vacuum with their own interpretations of data, particularly those with the largest resources for lobbying and data manipulation.

Finally, poor data can also lead to badly designed legislation or regulatory policy. For example, regulators may seek to amend existing legislation because the available data suggests it is not achieving its intended purpose or regulators may lack objective data to help decide between competing interests and outcomes when framing new legislation.

5.9. CONCLUSIONS

There are many direct and indirect benefits to a broad array of market participants and stakeholders that would result from the establishment of CT data. The direct cost savings from CT data could potentially save end-users billions of euros but it could also bring benefits to market transparency, competition and regulatory oversight and provide greater confidence and market integrity for issuers and investors.

There are small subsets of financial intermediaries who would be disadvantaged by the establishment of CT data. These are primarily those firms that have sufficient economic and technical power to process and profit from the current information asymmetries in European market data and the large incumbent data vendor firms who profit from the data that they sell.

6 / CHALLENGES TO DATA CONSOLIDATION

This Chapter summarises the issues identified with respect to consolidating data in the current environment, whether consolidating data for own use or consolidating it for commercial reasons. A CTP would have to deal with these issues to create a viable tape.

CHAPTER SUMMARY

The study identified the following barriers to data consolidation:

1. **Price of Data Not Determined by Market Forces Making Consolidation Unviable:**
   - Legacy market structure and cost drive users to economise by taking a subset of market data as a proxy, usually from the dominant venue. This reinforces the pricing power and liquidity provision models of the largest venues.
   - These venues have no incentive to price data at the value that the market would put on it when the data is consolidated, and the quality of each venue's contribution is assessed.
   - Unless this is addressed, consolidators will pass the costs from TVs and APAs directly to participants who will continue to economise by taking subsets of data, rendering consolidation efforts a waste of time for the CTP.
2. **Data Quality and Complexity**
   - Poor data requires considerable effort and resources to clean so that it can be properly compiled and used effectively. Issues result from ambiguities and inconsistencies in the rules, subjective interpretation of the rules, abuse of the rules or misuse of flags and a lack of mandated technical operating standards for data submissions.
3. **Poor Governance and Lack of Enforcement of Rules**
   - Issues arising need an immediate resolution for CT data to be of value to the market. ESMA and the NCAs are not close to the technical interfaces in the market where the data is submitted and cannot immediately identify and address issues.
   - NCAs must manually seek cross-market information from each other resulting in time lags for identifying and fixing data issues. A true picture of the market may be lost for a considerable time, possibly months or years.
   - There is also no centrally agreed penalty mechanism for poor data submissions or agreement about how one should be implemented.
   - TVs have better quality data than APAs because they have members who must follow their rules. APAs have worse quality data because they who do not have members, are not empowered to enforce penalties on their clients and do not always see both sides of a trade.
4. **Other Factors That Would Make CT Data More Useful and Viable**
   - Harmonising rules such as the deferral regime for bonds, the treatment of SIs as trading venues, more clearly defining what is an EU instrument and increasing the number of bonds available for publication would also improve the viability of the data.

6.1. PARTICIPANT FEEDBACK

The main concerns that participants raised as data consolidation issues were:

- Cost
- Data quality and complexity
- Governance and enforcement of rules
- Lack of harmonised rules on data
- The population of instruments that are currently included in the data

The following table highlights the significance that different stakeholders put on different issues. It demonstrates that the further away the stakeholder is from the point of execution, the fewer concerns they raised. However, the interviews also revealed that these stakeholders had a lower grasp of the potential incompleteness or inaccuracies of data and that they rely more heavily on advisors or data vendors for the accuracy of that information.

Figure 22: Matrix of Stakeholders and Barriers to Data Consolidation Raised by Interview Participants

●	Red: Raised often and seen as a critical issue impeding consolidation	◑	Yellow: Raised occasionally but not seen as a big issue
◕	Orange: Raised often and seen as an important issue to add value to the data	○	Grey: Stakeholders did not raise this problem or it is not relevant to them

		Barriers to Data Consolidation Raised by Participants				
Stakeholder Group	**Industry Grouping**	**Cost**	**Governance and Enforcement**	**Data Quality**	**Harmonisation of the Deferral Regime**	**Population of Instruments**
Issuers	**Issuers** (Corporate and Government/Funds)	◑ / ●	◑ / ●	◑ / ●	◕	◑
End Investors	**End Investors***	●	●	●	◕	◕
Financial Intermediaries	**Buy-Side**	●	●	●	◕	◕
	Sell-Side	●	●	●	◕	◕
	Trading Venues (Largest) Equities/Bonds	○	◕** / ●	◕ / ●	◕	◑ / ●
	Trading Venues (Alternative competitors) Equities/Bonds	●	●	●	○ / ◕	◕
Custodians/ Risk Managers	**Custody, Clearing & Settlement**	◑	◑	◑	◑	◑
Regulators	**Regulators**	●	●	●	◕	◕
Data Analytics & Benchmark Providers	**Data Analytics & Benchmark Providers**	●	●	●	◕	◕

Based on retail broker feedback

***Equity venues mainly raised issues about OTC data and adoption of standards rather than on venue data. Bond venues were concerned with the broader issues impacting OTC data.*

6.2. EXAMINATION OF THE CHALLENGES TO CONSOLIDATION

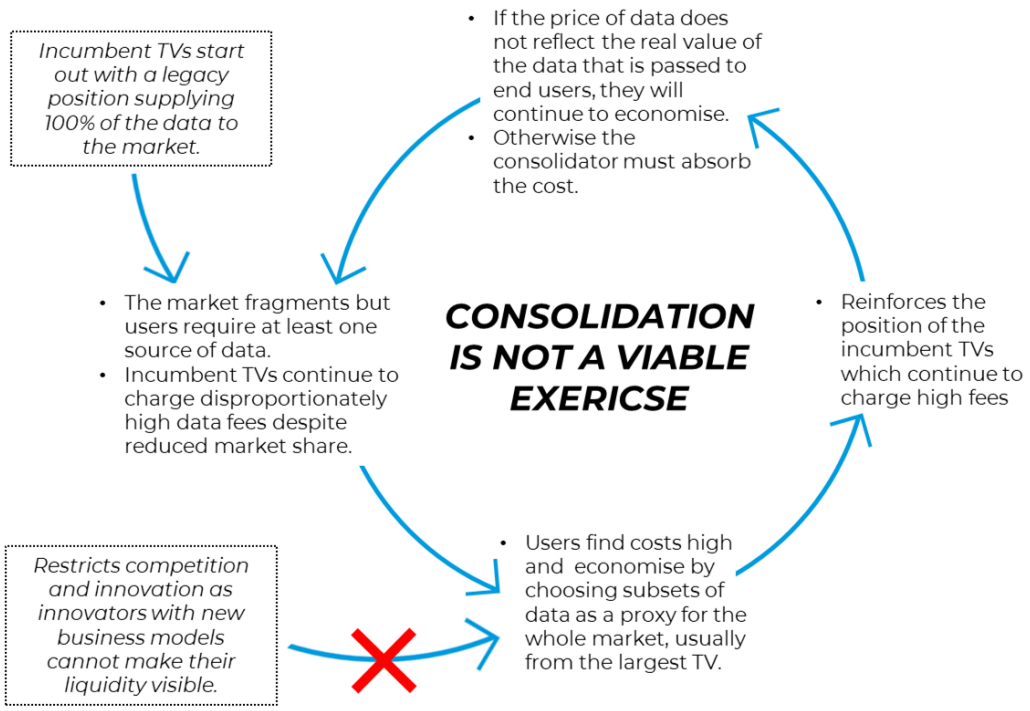
The tables below explain the background to the challenges and the issues that current participants and potential consolidators describe when considering consolidation of data.

6.2.1. Price of Data

Problem: The price of market data is not being shaped by competitive forces.	
Background	**Impact**
• Most incumbent (i.e. pre-existing) equity TVs already had a dominant position when MiFID I was implemented. As the market fragmented, the share of market liquidity captured by the incumbent TVs reduced. • Nevertheless, these TVs could still impose their legacy contractual terms and prices for data because the participants still needed at least one market reference point, even if they traded elsewhere. • In effect, these TVs continued to price their data according to what participants had previously been paying despite these TVs now accounting for a smaller market share of the total market. • As a result, participants that want to take data for the whole market (i.e. from other TVs and APAs) now have to find more budget plus they also have to expend resources on legal negotiations and cleaning of data across multiple venues. Alternatively, they can elect to receive less data (in terms of market share) from the incumbent TVs. • APAs, particularly those with TVs, are now behaving in a similar way in bond markets.	• The perceived high direct and indirect cost of taking and managing data feeds drives users to economise by taking a subset of market data as a proxy, usually from the dominant trading venue. • The pricing power and primary and secondary market models of the largest venues are reinforced. • The competition provided by smaller or new primary and secondary market models is restricted even if they provide data for free (because it takes time and effort to compile it and many consolidators will not or cannot expend resource on this

Problem: A consolidator has no control over the cost being ascribed by each TV or APA versus the value of the overall CT data to the consumer.	
Background	**Impact**
• The price that the dominant TVs and APAs put on their data may not reflect the true worth of their data, or the proportionate market share of their data when data is consolidated and the value of all data contributions is across the market is properly assessed. • TVs and APAs that own TVs and that currently profit from data revenues and the reinforcement of their primary and secondary market models have no desire to allow a competitor's data to be seen or to be consolidated as it might reduce their profit.	• Data consolidators must either absorb the data costs set by aggregators or pass them on to users. • If each aggregator's costs are simply passed on then users will continue to economise by taking a subset of the data, which effectively reinforces the pricing and business models of incumbent TVs and APAs. • Data vendors are not incentivised to compile data that users are not willing to pay for. • In these circumstances a viable CTP cannot be established, rendering any consolidation effort a waste of time.

Figure 23: The Impact of The Current Pricing of Data on Consolidation - Equity Example



6.2.2. Data Quality and Complexity

Problem: There are no technical operating standards for data submissions, data delivery formats and data storage.	
Background	**Impact**
• Significant efforts have been made by the industry to get voluntary standards/flags adopted to help reporting interpretations. For example, the Market Model Typology (MMT) which is an offshoot of the FIX Protocol industry association. MMT is meant to be an operational solution/tool to help organizations fulfil their trade flagging requirements and began in 2011. • Trading venues have their own bespoke codes for different trading systems, segments, and sessions such as opening or closing auctions, frequent batch auctions etc. These are not harmonised across venues making interpretation of cross-market data difficult. • Some aggregators are not supplying data in machine-readable formats. • Historical data requirements were not envisaged in the law. TVs currently store data in different formats over different periods of time.	• Pre-, post-trade and historical data is extremely hard to normalise and clean, making it hard to analyse and compile across the entire market. • Efforts to implement voluntary industry standards cannot easily move forward and time and effort are wasted. • Key information about trading status of a venue is not communicated to all market participants at the same time (e.g. trading halts) and TVs often obfuscate the reason for a trading halt, particularly if it is a technical issue. • It is very expensive and difficult to download data. Members of venues might have more access to data because they pay for a screen but non-members who wish to consolidate do not.

Problem: Poor quality data of off-venue data	
Background	**Impact**
Subjective interpretations of the RTSs vary from firm to firm. Original Level 2 text needed further guidance to correct misinterpretations: • Bonds and derivative data requirements are all in one RTS, creating confusion (e.g. a futures or a derivative contract would use "quantity", but a bond would use "notional" - notional figures are meaningless in certain contexts. • Specifications do not include reference data for identifying the instrument/issuer. • Confusion exists over the 'issuer and operator of the trading venue identifier'. **Flags are being used incorrectly:** • Amendments occur without cancellations of the original trade, which results in valid transactions being discarded and invalid transactions being included. • Equity flags are being used for bond transactions e.g. 'TNCP'. • Widespread misuse of deferral flags for equities and bonds is occurring. **Abuse of the rules includes:** • Double counting or no reporting. • Incremental bond data information that is due for publication post-additional deferral periods expiring is not being published (mainly by APAs). • Banks may systemically withhold trade reports until the absolute last moment of their obligation, rather than publishing at the moment of execution. **Ambiguities and inconsistencies allow:** • Firms can avoid being SIs (which brings undesired transparency requirements) by being registered liquidity providers on TVs and using the TV quoting mechanisms to provide very wide quotes on-venue whilst dealing inside the quote off-exchange. Then they report the trade to the exchange so that the post-trade data is flagged as on-venue. • Different interpretations of the "free after 15 minutes" rule. It is not clear if data can be stored and used once it becomes free and TVs charge for the data.	• Data fields are not being used as expected from firm to firm, formats are not being followed and reporting times are not being adhered to. • It is impossible to correctly calculate volume information. • It is impossible to build a true picture of the market. • Quality market transparency is unattainable. • Post-trade data is not easily interpreted as on or off-venue. • Data often disappears immediately after 15 minutes and venues charge for historical data. • Data cannot be consistently interpreted, even when from the same group company. • Firms are not able to consistently and easily able to source historical data from exchanges. This has implications for regulators, risk managers, compliance, and market surveillance teams as well as traders and analysts. • It is not possible to scroll back a few minutes or a few days and see what the prices were.

6.2.3. Governance and Accountability

Data governance needs centralised oversight without which the current data quality threatens the viability of a CTP.

Problem: There is no single body to oversee data submissions and mandate standards.	
Background	**Impact**
ESMA is expected to enforce data rules but ESMA relies on a federated model of regulation: • NCAs oversee APAs and TVs in their home markets. TVs and APAs collect subsets of data from participants, often in other markets. CTPs consolidate data from all these different markets. • Each TV can enforce its own rules. APAs cannot enforce rules. • No central penalty mechanism exists for poor data submissions or agreement about how a regime may be enforced, particularly under competing CTPs. • ESMA and the NCAs have no technical interfaces to see the data that is submitted so that they could immediately identify and address issues. • ESMA and the NCAs also economise and choose subsets of data to see - they do not receive CT data as it does not exist, particularly in real-time. • If issues occur in another market, NCAs must manually seek data from other NCAs. (Some regulators get their free domestic market information free from TVs, but they do not get consolidated cross-market data). • TVs can deny and restrict other regulators' access to their data, irrespective of formal requests. • TVs can reconcile trades because they have both sides of a trade. APAs do not. • MiFID II/MiFIR only specified single-sided trade reporting. Reconciliation is therefore difficult. • Many firms already consolidate data voluntarily, but each must make its own subjective interpretation of what the right volume numbers are and what might constitute addressable liquidity.	• Errors are undetected or take too long to resolve. • The manual nature of the oversight is not aligned with real-time data submissions • TV data is of a higher quality than APA data. • Market abuse oversight is hard to implement and detect across different jurisdictions and investigations are hindered. • No single entity can take responsibility for assuring data accuracy as the official source of data. • APAs have no incentive to penalise clients that could move their business elsewhere without repercussions. APAs also have no incentive to penalise clients who may also be clients of a TV or vendor business owned by the APA's parent company. • Mismatches between buy-side and sell-side data are frequent and often contribute to unexpected results and arguments as a result of slippage of trading strategies and costs for the end investor. • Whilst cost and centralisation of data governance remain an issue, it is preferable for data vendors/potential consolidators to remain unofficial consolidators of certain subsets of data without incurring other liabilities in relation to the accuracy or completeness of the data or to be subjected to regulatory oversight. • Different lobby groups use data sets selectively to their advantage, as there is no official source of data.

Problem: Conflicts of interest incentivise poor behaviour.	
Background	**Impact**
• The parent companies of APAs also run other businesses such as TVs and/or data vending businesses. • Some APAs, which are also running OTFs or MTFs, are withholding data from their OTF/MTF competitors and from other data vendors but freely incorporating the data into their own offerings. This is a particular issue in bond and ETF markets. • Group entities are not applying the same data standards to their APAs, MTFs and OTFs.	• It is very difficult for a CTP to access all the data needed on agreeable terms. • Currently, a CTP would need to speak to the regulator where each APA resides in order to find a resolution. • APAs, MTFs and OTFs often report information about the same instruments in different ways.

6.2.4. Lack of Harmonised Rules

The lack of harmonised rules applied to TVs and SIs and across jurisdictions for deferral publications makes data consolidation less viable.

Problem: SIs are treated differently to trading venues	
Background	**Impact**
• SIs are exempt from the tick size and clock synchronisation regime that trading venues must follow. • (Note that the SI tick size regime is changing in June 2020)	• SIs can attract flow by quoting just inside the spread of a trading venue. However, this price improvement may be *de minimis* whilst creating a significant amount of data that needs to be managed and may not, on the whole, be very valuable to the market. • It is possible that SI quotes and trade reports may not be synchronised with the TVs.

Problem: Bond deferrals are not harmonised	
Background	**Impact**
• Under the MiFID II rules, NCAs can determine their own rules for deferring publication about bonds.	• This creates regulatory arbitrage as participants may select where to report their trades, giving some EU countries an advantage over others. It may also hinder competition between TVs in different markets as liquidity may shift to markets where deferrals are longer and liquidity providers have more protection. • Real-time data cannot easily be used for data comparisons as bonds information becomes available at different times. This prevents useful consolidation.

6.2.5. Population of Instruments

The population of instruments that might be included in a tape is not easily defined and is also low in numbers for some asset classes.

Problem: The population of bonds eligible for the tape is very small.	
Background	**Impact**
• ESMA decided to take a phased approach for the liquidity assessment of bonds, gradually decreasing the average daily number of transactions in a bond needed for determining a market as liquid (and thus reportable - only the most liquid bonds are reportable). This is assessed on a quarterly basis taking into account the daily average trading activity (trades and notional amount) and percentage of days traded per quarter.	• Consolidation of bond data is less viable because there is very little data that can be consolidated and of use to participants

Problem: There is no way to clearly define an EU instrument.	
Background	**Impact**
• All trades that are traded on a trading venue TOTV must be reported regardless of whether they are for EU instruments or non-EU instruments.	• A significant amount of trades for non-EU instruments are being reported but cannot easily be filtered out of the data due to the resources required to adjust substantial numbers of records with incohesive identifiers.

6.3. THE CUMULATIVE PROBLEM

These issues create significant barriers to consolidation but also, they contribute to the problems that market participants have in dealing with data every day.

Each entity that compiles a set of data makes multiple discretionary decisions about whether it will pay for all data or a subset of that data, whether they will take the data from the source or from a vendor and how much additional cleaning of the data they will do. Data vendors that consolidate data pass on the indirect costs for managing and cleaning the data as well as the direct cost of the data feeds. Each customer of a data vendor may see different data, depending on what they are willing to pay for.

Many interpretations of data exist depending on the ability of each firm to clean and interpret data. The results, depending on the technical resources available to clean the data, vary significantly.

As a result, no one has the same set of data for any instrument. When one party refers to its own set of data, it is likely to be different from another's.

6.4. PRACTICAL IMPLICATIONS OF THE CHALLENGES THAT ARISE

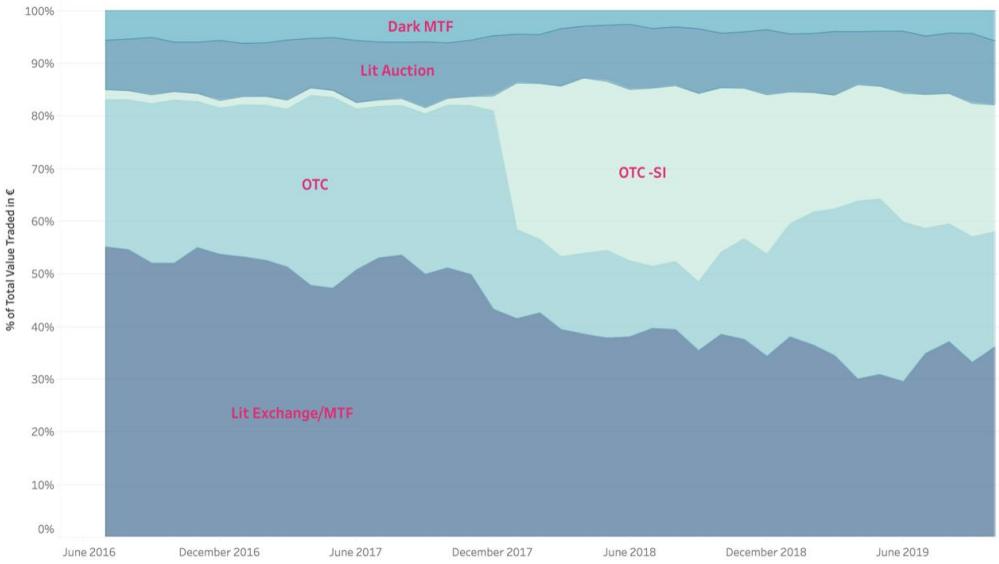
One of the key data points users need is an average daily trading volume figure for each instrument.

Big XYT is a sophisticated, independent data analytics firm interviewed for this study. The question Big XYT's customers want to know is, 'what was the actual addressable liquidity in the market at the time of trading? As described in Chapter 2, this is the liquidity that customers can actually interact with, rather than trades that are non-price forming or do not contribute to market liquidity, such as technical trades, duplicates, give-ups, or special dividend trades and after-hours trades. Without an official source of data, vendors and firms must try and establish this for themselves.

To answer this question, Big XYT must first assemble the data. In one month, it processes over 13 billion rows of trades from TVs, which is on average 15,000 rows of data per second. These must be matched with over 800 bespoke trade condition codes that are not standardised by the TVs.

At first, Big XYT takes the unfiltered view of the gross volumes of the entire market data (See Figure 24 below). At this stage, it shows that off-exchange data accounts for approximately 45% of the total market volumes, with SIs accounting for 30% and the rest of the off-venue trades accounting for 15%. Lit exchange markets appear to be declining. This is the picture of the market that anyone who takes uncleansed data or who cannot afford to clean the data would see.

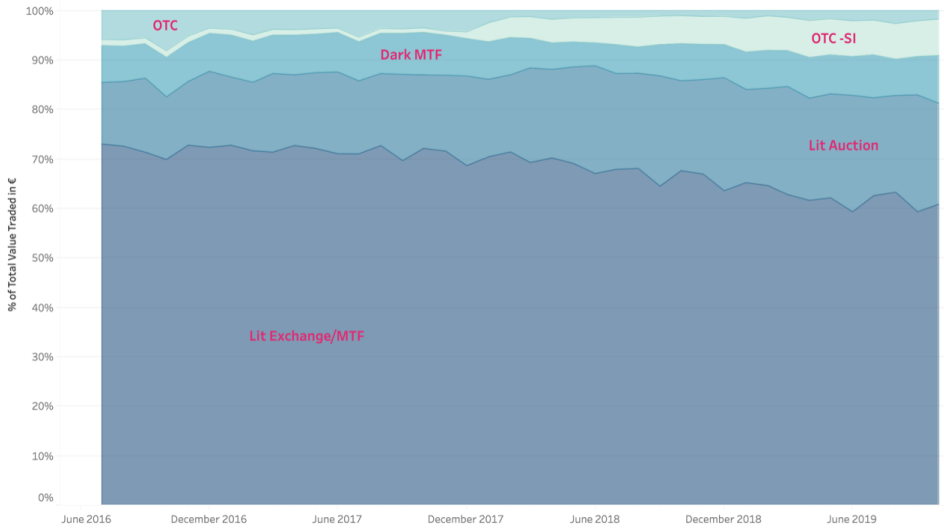
Figure 24: Unfiltered View of Gross Post-Trade Volumes (for instruments TOTV) by Trading Protocol over a 3 Year Period.



Source: Big XYT

Once it has this data, Big XYT goes through a series of steps. First, it filters out non-price forming trades and then it filters out block trades on the basis that anything large in scale is most likely a result of a bilateral negotiation that was a specific bespoke trade for a customer and not addressable. At this point, the picture of liquidity appears completely different. Lit trading is a much larger proportion of the market and appears relatively steady as an overall percentage of the market (see Figure 25).

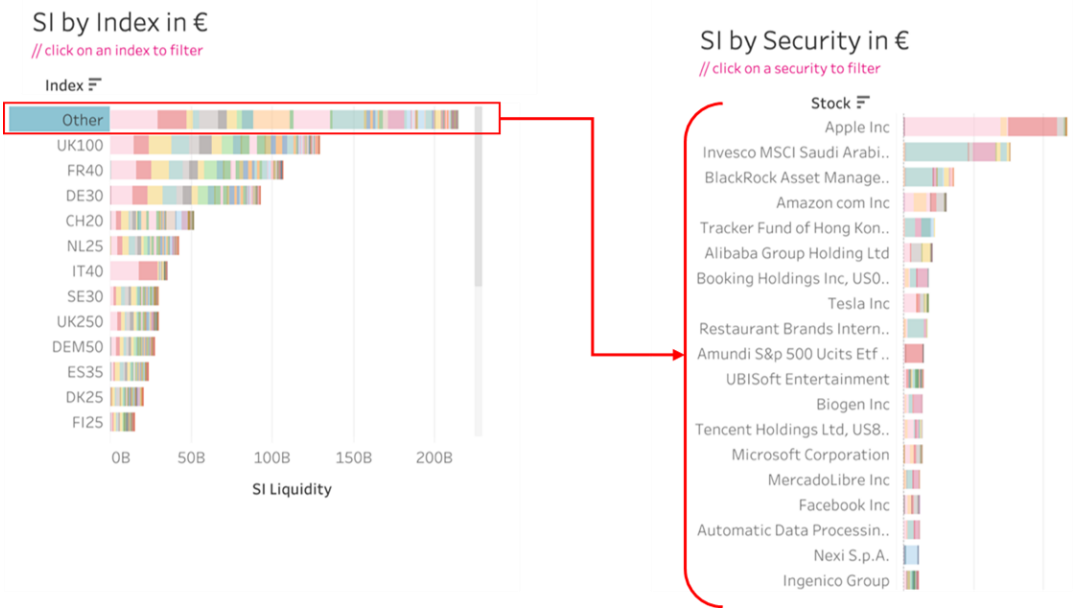
Figure 25: View of European Post-Trade Volumes (for instruments TOTV) by Trading Protocol with Filters Applied.



Source: Big XYT

Even then this data is further complicated by the scope of what must be reported under MiFID II. This is defined as any financial instrument that is traded on a trading venue (TOTV), which includes non-EU listed equities that further distort numbers. The above numbers include vast numbers of trades in non-EU instruments such as Apple or Microsoft. Further analysis by Big XYT shows that this category of "other" stocks is larger by notional traded value than any other EU index.

Figure 26: Number of SI Reported Trades by Index including Non-EU Instruments



Source Big XYT

This experience of Big XYT demonstrates the complexities of data interpretation for European investors; how time-consuming it is and how much processing power it takes to compile useful data about European equity markets without an official source of CT data.

6.5. CONCLUSIONS

Successful data consolidation is impeded by legal, structural and economic issues. This is because the underlying data that needs to be consolidated resides across many competing commercial entities, some with disproportionate economic leverage and conflicts of interest, as well as with inconsistent quality checks, data models and interfaces, all under a federated model of supervision and enforcement in multiple jurisdictions. Technology is not a problem.

The proposition for a European CTP is less commercially attractive and more technically burdensome than the unsupervised role that unofficial data consolidators such as data vendors play today. Without changes to the way that data is valued and governed, data providers (whether individual or aggregators) will continue to pass on the direct and indirect cost of data feeds, which may not reflect the true worth of the data when consolidated and participants will continue to economise by selectively choosing which feeds they take, thus rendering consolidation a pointless exercise.

7 / LESSONS LEARNT FROM LITERATURE AND FIELD RESEARCH

This Chapter explains the lessons learnt from the field research undertaken in North America and a review of the globally available literature with some observations about how Europe compares.

CHAPTER SUMMARY

1. There is currently no competition in data aggregation and consolidation in the North American markets. Multiple data aggregators exist, but each has exclusive responsibility to aggregate sub-sets of data. There is a single data consolidator in each asset class.
2. Optimum data quality occurs when there is an exclusive self-regulated aggregator and/or consolidator for all on and off-venue data.
3. An entity that has the exclusive responsibility for enforcing rules and standards for CT data should be recognised in the law and be able to enforce data-related laws.

   The regulatory authorisation of any data contributor to that entity should depend on its membership and willingness to abide by the entity's rules.
4. Entities with exclusive responsibilities for CT data provision should not be conflicted and their governance must require the balanced representation of data stakeholders. Otherwise, they may pursue business models that are not in the best interests of the broader market.
5. The widest use of official consolidated data occurs when:
   - A single technical, contractual, and pricing interface exists for receiving data from the aggregator and disseminating it to the consumer,
   - The underlying data is not acquired at a price set by each TV but is based on the value of each data set to the overall CT data revenue.
6. A single consolidator with balanced governance may be a priority (for Europe) because it can work with all stakeholders in the market whereas aggregators only work in the interest of a subset of stakeholders.
7. The information to be provided in the CT data should be defined in the law. For equities, this includes the depth of data and auction imbalance information.
8. The design of shared revenue allocation models can be used to facilitate competition and drive changes in trading behaviour.
9. Mandated use of a consolidated tape for best execution depends on the underlying market structure, so may not be appropriate for Europe and can be problematic for institutional participants.
10. A vendor display rule is beneficial and should be available for free to independent retail investors to prevent substitute products from being used.
11. The design and constituents of CT data can incentivise both positive and negative changes in trading behaviour and need to be clearly understood.
12. Aggregators and/or consolidators need to have strong reconciliation processes and capabilities to be able to clean data.
13. Consolidated data increases transparency and reduces costs although some participants are concerned about the impact on large, illiquid bond trades.

7.1. A BRIEF DESCRIPTION OF NORTH AMERICAN DATA CONSOLIDATION FRAMEWORKS

7.1.1. US Equities

(A full background to US equity data consolidation is given in Appendices 7, 8, 9)

In the US there are 17 exchanges, 32 alternative trading systems (ATSs) and numerous dealers reporting trades. There are **three aggregators** responsible for different data sets, i.e. they **do not compete to aggregate the same data**.

The management of the consolidated tape, also known as the SIP (Securities Information Processor), is overseen by the Consolidated Tape Association (CTA). The CTA is run by representatives from the exchanges and The Financial Industry Regulatory Authority (FINRA), a self-regulated organisation, which is owned and funded by dealers. Nasdaq and NYSE, the most historically prominent commercial stock exchanges, were originally appointed to run the technical aggregation of different sets of listed securities pre and post-trade data from across the TVs into one official set of data for each underlying instrument and continue to do so today, although this could be changed by the CTA. FINRA aggregates and cleans off-venue post-trade data. FINRA acts as the single technical consolidator of the on and off-venue data.

FINRA also undertakes cross-market surveillance of transactions for the whole market.

Use of the tape is mandated for best execution and for display to retail investors at the point of trade.

See Figure 27 below.

7.1.2. Canadian Equities

(A full description of Canadian equity data consolidation is given in Appendix 10)

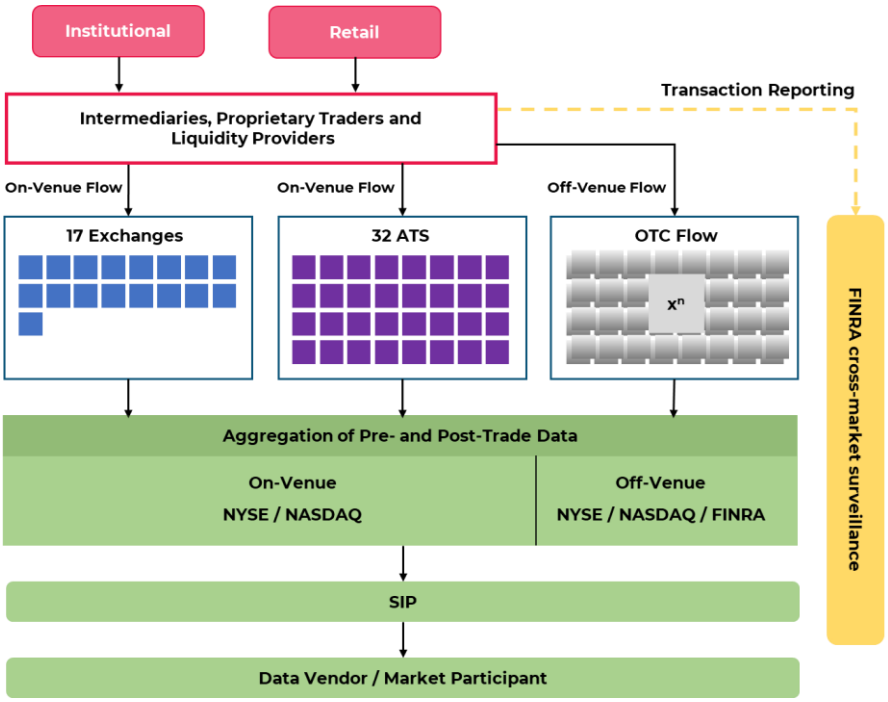
In Canada, there are 6 exchange groups and 5 ATSs.

Each TV acts as the aggregator of its own data. The legislation provides for the existence of consolidators, known as Information Processors (IP). They must satisfy certain criteria and be approved by the regulator. One IP, run by an exchange, currently exists as a consolidator of pre- and post-trade data but the use of its data by the market for best execution or other purposes is not mandated through regulation and this IP does not mandate technical standards.

The Investment Industry Regulatory Organisation of Canada (IIROC), which is a dealer-funded, self-regulated entity, undertakes cross-market surveillance for the whole market.

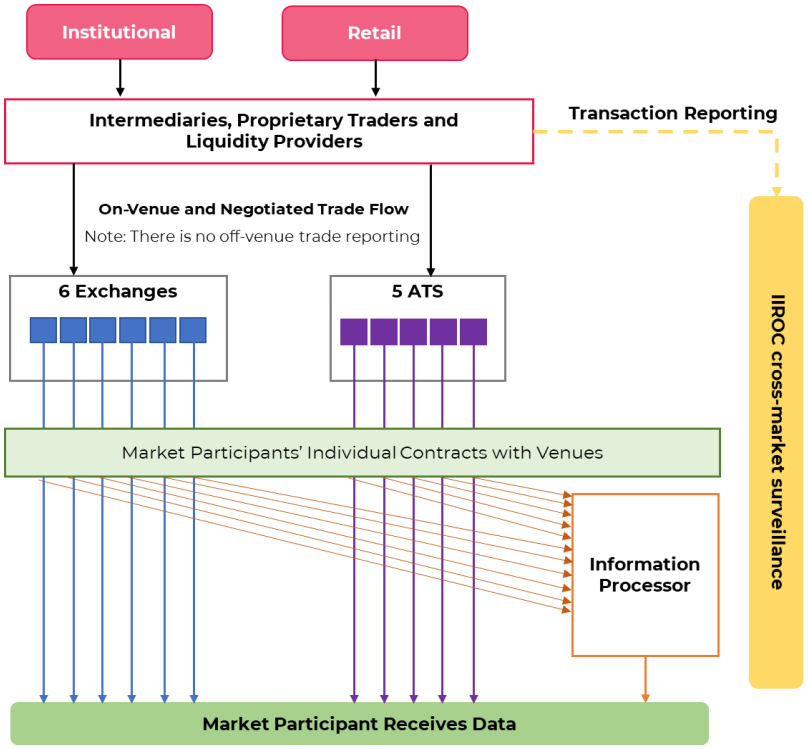
See Figure 28 below.

Figure 27: Schematic of US Equity Data Aggregation and Consolidation



Source: MSP research, FINRA, CTA

Figure 28: Schematic of Canadian Equities Data



Source: MSP research, FINRA, IIROC, IP

7.1.3. US Bonds

(A full description of US Bond Market Consolidation is given in Appendices 7, 11)

In the US bond market, dealers are responsible for reporting to TRACE, which is run and governed by FINRA. All data is aggregated, consolidated and disseminated by FINRA through its TRACE system.

FINRA also undertakes market surveillance for the whole market.

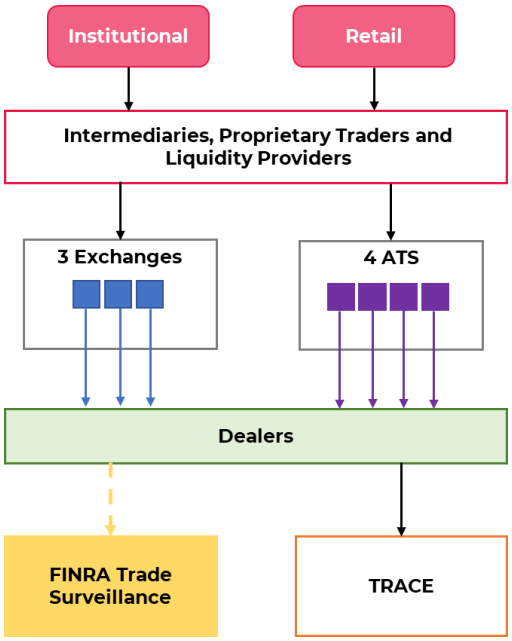


Figure 29 (right):

US Bonds Trade Reporting and Aggregation Flow Post-Trade

Source: MSP Research, FINRA

7.1.3. Canadian Bonds

(A full description of Canadian Bond Market Consolidation is given in Appendix 12)

In the Canadian bond market, dealers are responsible for reporting to IIROC, which is a dealer funded self-regulated entity. All data is aggregated, consolidated, and disseminated by IIROC.

IIROC also undertakes market surveillance for the whole market.

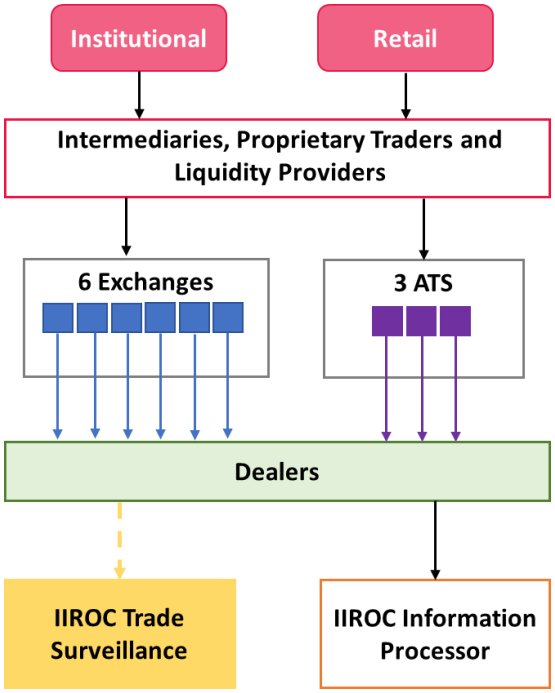


Figure 30 (right):

Canadian Bonds Trade Reporting and Aggregation Flow Post-Trade

Source: MSP Research, IIROC

7.2. LESSONS LEARNT

1 Optimum data quality occurs when there is a single, self-regulated, aggregator for each sub-set of data and/or consolidator for all data.

North American Field Research Findings

- Users cite few issues with data quality in North American markets.
- There is no competition for data aggregation or consolidation. Multiple aggregators exist but with exclusive rights to aggregate subsets of data that are then consolidated by a single consolidator. These exclusive aggregators and consolidators can establish and enforce rules and define the technical operating standards for its members through self-regulation. Data is cleaned before it is consolidated and disseminated.
- In Canadian equities, competition is allowed but only one consolidator exists (participants attribute this to the underlying costs and governance of the data). TVs have their own technical standards, which require translation in order to be consolidated. However, there are only 11 TVs from which to consolidate data.
- The Securities and Exchange Commission (SEC), the US equity market regulator, is proposing to introduce competing consolidators, responsible for technically collecting the data from the aggregators (e.g. the SROs) and then consolidating and disseminating the data. The proposal would also allow brokers/ dealers to self-aggregate and consolidate data for their own use.
- It is very important to note that this proposal still requires the sole aggregators of the data (FINRA and the CTA) to act in the same exclusive manner using the same methods as they do currently. This ensures that standards and rules are applied to a member's data submissions which are cleaned prior to the data being issued to consolidators.

Global Literature Research Findings

- Europe's intention to allow competing consolidators under a market-led approach to data consolidation could risk poor data quality due to the lack of mandated technical standards.[24]
- Lack of standardisation, inconsistencies in data formats and data errors are some of the key data aggregation issues for participants post-MiFID II/MiFIR.[25]
- European participants want the confidence that there is a neutral and reliable source of the current market price.[26]
- Self-regulation may be a valuable complement[27] to regulators in achieving their objectives but some question the value of it.
- There is no universal definition of self-regulation.[28] The term may refer to formal self-regulatory organizations (SROs) or simply to standards set by industry associations.
- Self-regulation in Europe (except in the UK) was never extensive because of Europe's civil law system and cultural approach to government supervision of financial business.[29]
- The efficacy of self-regulation depends on each market's circumstances. Self-regulation should neither be adopted automatically nor rejected out of hand.[30]

Observations About European Market Structure and User Feedback

- TVs have members and can self-regulate. Users have few issues with the TV data quality.
- APAs do not have members and cannot self-regulate. Users have many issues with off-venue data quality sourced from APAs.
- There is no single technical operating standards body and no official source of data.
- Europe (ex. UK) has not had a dealer culture for off-venue flow and regulation of this flow has been evolving.

2	An entity that has the ultimate responsibility for enforcing rules and standards for CT data should be recognised in the law so that it is empowered to enforce its rules as well as relevant laws. The regulatory authorisation of any data contributor to that entity should depend on its membership and willingness to abide by its rules.

North American Field Research Findings

- Self-regulatory entities that are also recognised in the law are known as Self-Regulated Organisations (SROs) (further detail is given in Appendix 7).
- Aggregators and consolidators that are recognised as SROs have a clear mandate in the law to consolidate data under defined terms. Under regulatory supervision, they can enforce their own rules as well as related laws. They can also issue fines.
- Regulatory authorisation of dealers requires membership of the SROs, FINRA and IIROC.

Global Literature Research Findings

- SROs have shown that they can work successfully with regulators and stakeholders to drive change in data provision (see Appendix 7) but issues arise when SROs have commercial interests.

Observations About European Market Structure and User Feedback

- The law does not provide for the delegation of regulatory responsibility to a TV, APA or CTP and has not provided them with a role to play in enforcing relevant data laws. TVs can enforce their own rules.
- A firm's willingness to provide quality data to a CTP is not currently linked to its authorisation and there are no sanctions for poor data submissions.

3	Entities with exclusive responsibilities for CT data should not have conflicting interests and their governance requires a balanced representation of data stakeholders. Otherwise, they may pursue business models that are not in the best interest of the broader market.

North American Field Research Findings

- Users cite few issues with governance or the cost of off-venue bond data in markets where the aggregation and/or consolidation is run by a mutualised entity.
- TVs are mostly 'for profit' entities and as such must generate returns for shareholders. Stakeholders cannot influence data governance and cost at these entities.
- The prevailing liquidity provision model at the for-profit equity venues where certain market makers are very active and have been willing to pay more to receive faster data creates a two-tier system of speed where some participants are prepared to pay for access to faster data and then create high volumes of small-sized quotes in the market. These liquidity provision models generate huge amounts of data which all data users are forced to process regardless of its usefulness.
- The recent consolidation of TVs has resulted in the emergence of fewer but more powerful TV groups, which can block or sway decisions and prevent change at the CTA.
- Other venues with wider stakeholder

Global Literature Research Findings

- The pursuit of profit maximisation through high-speed data sales may deter TVs from innovating. This can have a negative impact on price discovery and capital allocation.[31]
- Where intense competition exists between multiple informed participants to obtain faster data to compete against others, the TVs are incentivised to charge for the different speed of signals. This can drive markets to fragment in structural ways that favour a subgroup of well-situated market participants at the expense of others.[32]
- This also drives concentration as the twin effects of the increased cost of data and the significant scale needed to process the data can outweigh the benefits of innovation at competing venues.[33]
- Exchanges have historically been shown to use weaknesses in regulation to their own economic benefit. When they demutualise, consideration should be given to creating balanced or mutual governance structures for market data particularly in relation to cost.[34] Political intervention will be needed to enable

representation are trying to compete with different liquidity provision models where tiered systems of speed are less relevant or where data is free, but they have little voting power at the CTA.

- The SEC cites concerns about the commercialisation of SROs as a driver of its proposed changes to allow competition between consolidators. It also intends to cap the voting rights of each TV group at the CTA.

this.[35]

- There is a tension between creating CT data and promoting competition between TVs. [36] This involves trade-offs which incumbents are unlikely to willingly accept and will likely require affirmative and compulsory action by the market regulator.

Observations About European Market Structure and User Feedback

- Similar liquidity provision models and tiered systems of speed prevail in European equity markets and users cite issues with the cost of data and processing requirements.
- Some users question the usefulness of the data generated by some TVs liquidity provision models.
- Data stakeholders do not have influence on TVs and APAs and complain about their costs.

---

4 Consolidation and dissemination work best when:

a) there is single technical, contractual and pricing interface for the users to receive the data.

b) the underlying data does not have to be acquired at a cost determined by each aggregator and the revenue generated by the consolidated data is shared between the members based on the value of each data set to the overall consolidated data.

North American Field Research Findings

- For US equities and Canadian and US bonds, the CTPs provide one technical interface, one price and one contract for users to consume the data.
- US equity data is submitted freely to the aggregators in return for a share of the revenue. The CTA sets the price and manner in which TVs get paid for their data through a revenue allocation model and one price is set for the consumption of the data. (A full description of the US equity revenue allocation model is given in Appendix 9).
- In Canada, the IP does not set the price of data or deal with contractual issues so that the participants must negotiate their own contracts with each underlying venue. Consequently, the Canadian consolidated tape offering is less attractive for users and many participants take subsets of data as a proxy for the whole market rather than pay for all the data. However, the smaller nature of the market makes it easier for participants to consolidate the data.
- Equally a key takeaway from research into the IP suggests that running the IP is not a particularly viable or profitable business (because there is considerable effort required but little use of the data)

Global Literature Research Findings

- Direct and indirect costs (e.g. contractual complexities) are considered an issue by most market participants globally and these costs are a primary reason why a European consolidated tape has not yet emerged.[37]
- The debate about the cost of data is extensive. Some participants advocate the free provision of data or marginal cost pricing as an alternative to current pricing models.[38]
- TVs argue that market data is a by-product of trading and execution, which means costs and revenues are shared and hard to separate.[39] They also say that costs for consumption of data feeds have barely changed and that their own costs are only a small portion of overall intermediary costs for the end investor.[40 & 41]
- In Canada, the cost of data has increased[42] and restrictive legal agreements exist for market data.[43] Even when differences in market characteristics and scale are taken into account, CT data fees are still significantly higher than the US.[44]
- Lack of CT data in Europe means that traders are unable to get information from anywhere but the primary exchange data products which is a worse situation than the US.[45]

Observations About European Market Structure and User Feedback

- There are hundreds of TVS and APAs that any CTP must interface with.

- Potential CTPs (and users) without any control over data costs and contractual arrangements combined with unknown revenues find the commercial proposition of consolidation unattractive.

5 A single consolidator with balanced governance may be a priority (for Europe) because it can work at the highest level with all stakeholders in the market, and aggregators only work in the interest of a subset of stakeholders.

North American Field Research Findings

- US participants complain about both the cost and latency of CT data, both of which are governed by the CTA.
- The latency of the CT data is not sufficient for many participants to use for trading, but they must use the NBBO for best execution in the law (see below).
- TV aggregators are believed to be investing less in the technology used to deliver data between data centres to the consolidator than on their own proprietary data feeds.
- The SEC's proposal to introduce competing consolidators is aimed at improving latency and addressing conflicts of interest about investment in the CT data technology, but it does not give the stakeholders influence over the cost of CT data which is determined by the aggregators of the data and influenced by a small number of TV groups at the CTA.
- The SEC has begun to increase oversight of data fees through review processes with each TV.

Global Literature Research Highlights

- The US consolidated tape processor has not benefitted from the same level of technology investment as the direct data feeds with less superior fibre optic cables for transmission, less streamlined setup, and additional processing requirements, thus causing additional latency.[46]
- In regions where there are multiple legal jurisdictions, such as Europe, there are likely to be conflicting bodies of law that will make it impossible for regulators or courts to resolve data issues alone.[47] This means that both regulators and stakeholders will need to work together to create change through practical alternative solutions.

Observations About European Market Structure and User Feedback

- Cost, not latency, is one of the main issues cited by participants.
- ESMA is increasing focus on market data costs but change is harder to manage and implement under a federated model of supervision and when most TVs are being run for profit.

---

6 Sufficient detail must be provided in the CT data for users to ascertain current liquidity and trading intentions. This data should be defined in the law but with the flexibility to provide for future enhancements. For equities, it includes:

- Depth of data, which is more important than a BBO and over time the concept of a market-wide BBO may breakdown.
- Administrational information and auction imbalance data.

Without depth of data, the introduction of lot sizes must be considered in Europe.

North American Field Research Findings

- In US equity markets, orders are increasingly in smaller sizes and using the NBBO, which flickers faster than can be seen by a naked eye, as a gauge for liquidity or even price has become meaningless.
- Many users including investment banks, asset managers and retail brokers, would prefer to see up to 5 levels of depth of the order book as

Global Literature Research Findings

- The focus that users have on latency is dependent on the underlying market structure, e.g. mandated rules to use the data, and it may not be an issue in jurisdictions outside the US[48].
- Latency issues in the US appear unavoidable because the use of CT data is mandated, and information cannot be propagated instantaneously across a fragmented market

an offset to the speed.

- The CT data is considered the official version of administrative events even if firms receive data faster through other feeds.
- The core data and administrational event data that must be submitted freely by the TVs for equities is defined in the law but currently, only the NBBO and a limited range of auction information is provided.
- The SEC proposes to expand core data in the law to include 5 levels of depth of book data and information about opening and closing auctions. It will also introduce several new defined terms including "consolidated market data," "core data,", "regulatory data," "administrative data," and "exchange-specific program data."
- It also intends to include any information specified by SRO rules or effective Consolidated Tape Plans that are generated by an SRO leading up to and during an auction, including opening, reopening, and closing auctions and information disseminated during the time periods and at the time intervals provided in such rules and plan.

with spatially separated matching engines. Due to this, bids and offers on different feeds may vary[49] and the concept of a consolidated BBO may break down[50]

- Time stamping precision is therefore important in creating an accurate picture of the market at any point[51 & 52], regardless of the latency experienced by each user.
- Broker-dealers, investment banks and asset managers in the US, back the SEC proposals.[53]
- A TV has proposed three different tape offerings with three different levels of depth that might appeal to the different needs of users.[54]

Observations About European Market structure And User Feedback

- Users would like to have similar detail in EU CT data (e.g. five levels of depth and auction imbalances).
- Pre-trade data consolidation is not envisaged in the current law. Core data is not defined.
- Trade sizes are typically divisible by one and lot sizes do not exist. However, if depth of market is not introduced to CT equity data then crossed[55] books may become an issue.

---

7 Revenue allocation models, based on the value of data, can be used to facilitate competition and drive changes in trading behaviour and liquidity provision. These must be carefully calibrated and monitored to incentivise the right behaviours.

North American Field Research Findings

- In the US equity market, any TV has the immediate opportunity to earn revenue from the consolidated tape based on the quality of its contribution to the overall market liquidity (see explanation in Appendix 9).
- The addition of new TVs does not by itself drive up the cost of data for the end-user. Instead, it simply means that the revenue shared by the data providers is divided or pro-rated to those that contribute to the best liquidity (subject to agreement as to what is "good liquidity").

Global Literature Research Findings

- The model's formulae, which allocates payment for data in the form of a share in the revenue from data sales, can change behaviour. It can have both positive and negative effects depending on how it is calibrated, e.g. revenue based only on the number of trades may result in larger numbers of smaller trades.[56]

Observations About European Market Structure and User Feedback

- The price of market data is not shaped by market forces.
- New TVs in Europe are very hard to establish because:
  - The market cannot see its liquidity.
  - They cannot monetise their data.

8	Mandated use of a tape for best execution depends on the underlying market structure (e.g. homogenous clearing and settlement across the region), which may not suit Europe and can be problematic for institutional participants.

North American Field Research Findings

- In the US and Canada, equity clearing and settlement are handled by one single entity in each market. Additionally, the price for clearing and settlement paid by a participant is the same regardless of where they trade. This means that a strong emphasis can be put on price as the main factor under best execution rules.
- The US and Canadian equity markets have similar requirements on order protection rules. These rules mandate that orders must be routed to the marketplace with the best-priced orders available or executed at that price. In the US, this is linked to the consolidated NBBO.
- Many US market participants say that the order protection rule and use of the tape for best execution is problematic for institutional investors who need to deal in large sizes but are forced to access small-sized quotations and reveal their trading intentions due to the rule.
- In Canada, the core data for equities is not mandated in the law, as the data does not have to be freely given to the IP.

Global Literature Research Findings

- A one size fits all (both institutional and retail investors) approach to best execution may be unfair.[57]
- A threshold could be introduced into US markets (e.g. 1.5% of market share for a venue) to help reduce the fragmentation in the market and the forced connection costs.[58]
- Global market structures are not uniform and regulatory frameworks in each jurisdiction need to evolve accordingly to find the right consolidated data solution.[59]

Observations About European Market Structure and User Feedback

- Best execution in a European context takes into consideration many factors, including the cost of clearing and settlement, and not just price.
- Market participants typically determine their own best execution policies and TVs do not have any responsibility to onward route orders to other markets.
- Users do not want CT data to be mandated for best execution.

9	Mandated use of a tape to display CT data to independent retail investors is beneficial and should be available for free to prevent substitute products from being used.

North American Field Research Findings

- In the US, the Vendor Display Rule ensures that brokers display CT data to retail investors when a trading decision is being made. However, if the investor is using the tape for non-trading purposes (e.g. valuations), the use of the tape is not mandated.
- TVs have a key role in governing CT data and are therefore conflicted when developing substitute products of their own data to retail investors for non-trading purposes. They offer their own Level 1 data (which is a subset of the NBBO) as a proxy of the NBBO. They make this cheaper and easier to manage contractually, thus undercutting the SIP (i.e. CT data).
- Retail brokers cite the extraordinary complexities of managing the data used by

Global Literature Research Findings

- Various exchanges are filing programs with the SEC to allow small retail brokers to purchase their top of book market data at discounted fees.[60]

their clients. Incorrect classifications of retail investors can lead to legal risks and costs. As a result, it can often take months to onboard clients.

Observations About European Market Structure and User Feedback

- Some users consider a vendor display rule for independent retail investors a good idea.
- However, retail brokers are concerned that such a rule could force expenditure on smart order routers and will require clear explanations about all the factors taken into consideration in seeking best execution.

---

10 The design and constituents of CT data can have positive and negative effects that need to be well understood. Any tape design should allow flexibility should remain appropriate for the market it serves.

North American Field Research Findings

- In the US market, trades in odd lots are not included in the CT data. (Round lots in equities are units of at least 100 shares and odd lots for bonds are amounts that are lower than or not divisible by $1 million).
- Because only one level of depth is currently used for the equity BBO in the US, odd lots (orders below 100 shares) are excluded from the NBBO to ensure that the NBBO reflects a meaningful number of shares.
- Odd lots were historically thought to be insignificant retail orders. However, professional algorithmic traders took advantage of the odd-lot rules to avoid detection by placing a larger volume of small orders in the market.
- Additionally, some trading in heavily traded stocks is no longer part of the NBBO because the prices of stocks such as Apple and Amazon are so high that a retail investor might not be able to afford to buy 100 shares and so will be buying in odd-lots.
- SEC statistics[61] show that odd lot trading is now at historic highs and they are proposing to include odd lot information in the consolidated tape.

Global Literature Research Findings

- The underlying characteristics of data and how it drives market behaviour must be well understood when designing CT data. For example, the reporting of orders versus trades at different TVs can limit the usefulness of CT data, leading to erroneous conclusions and undermining the empirical integrity of the CT feed.[62]
- Inclusion or exclusion of odd lots can change trading behaviour. A recent class action complaint in the US asserts that odd-lot investors persistently pay bid-offer spreads that are 25% to 300% wider than investors trading in round-lots of the same underlying bonds.[63]

Observations About European Market Structure and User Feedback

- Trade sizes in European bond markets do not formally recognise odd lots but institutional orders in corporate bonds are typically traded in minimum or round quantities (e.g. 100,000 nominal).
- Retail brokers say there are no prices at a retail size in bond markets.
- Different TVs have different practices in reporting orders versus trade data.

11	Aggregators and/or consolidators need to have strong reconciliation processes and capabilities to effectively monitor and cleanse the data for reporting.

North American Field Research Findings

- In the US and Canada, FINRA and IIROC aggregate both the post-trade data and the equivalent of transaction reporting data. This gives them an additional set of data to reconcile against.
- In bond markets, double-sided reporting is required if two dealer member firms are involved in the trade. This provides an extra set of data that can be used to help reconcile and clean data.

Global Literature Research Findings

- There was no literature discovered on this point.

Observations About European Market Structure and User Feedback

- In Europe, the transaction reporting data is sent to each domestic NCA in a format specified by each NCA. This format is not uniform across Europe, which means that it is not readily available or easily usable to a CTP provider.
- At present only single-sided reporting is mandated in Europe and participants are not in favour of double-sided reporting.

12	Consolidated data increases transparency and reduces costs. However, consideration must also be given to the effect of transparency on the ability to transact large orders.

North American Field Research Findings

- Most equity and bond market participants in the US and Canada advocate the benefits of a consolidated tape.
- However, in bond markets, institutional market makers and asset managers are concerned that the immediate transparency of large trades can damage their ability to source or provide liquidity, particularly since bank capital requirements have increased and the introduction of the Volcker Rule,[64] which prohibits banking entities from engaging in proprietary trading.
- Conversely, ETF market makers plus other market makers that provide liquidity in smaller sizes on CLOBs believe that the current transparency is not enough.
- Retail brokers also report too few price points for retail investors.

Global Literature Research Findings

- Lack of CT data in fragmented equity markets may result in higher trading costs, difficulties in using smart order routing for effective best execution, difficulties in measuring best execution, reduced transparency for buy-side traders and retail investors, and limitations on the longer-term adoption of electronic trading.[65]
- In the US, TRACE has had an impact on lowering transaction costs for investors, particularly retail investors.[66] However, both institutional buy-side and sell-side participants say that trading is more difficult under TRACE and this can decrease willingness to commit capital and increase the time and the cost to locate or sell bonds.[67]
- This transparency provided by TRACE has impacted investment firm profit margins and trader compensation leading some to shift of their focus to trading less liquid products.[68]

Observations About European Market Structure and User Feedback

- Most users are advocates of CT data but some express concerns that increased transparency, particularly in bond markets, may impact brokers' willingness to commit capital.
- Retail brokers also report that it is difficult to find prices in retail sizes in bonds, which make trading and valuations of portfolios difficult.
- Most users across asset classes say that costs are higher without CT data.

7.3. CONCLUSIONS

North American market structure is much less complex than that of Europe. The starting point for successful data consolidation has been through the use of well-governed, self-regulated entities with exclusive aggregating and/or consolidating responsibilities. This allows them to mandate rules and standards for the market.

Pre- and post-trade data that is clearly defined in the law and legal recognition of these entities also gives them additional powers to enforce related laws on their members. Additionally, the US equity experience demonstrates that revenue allocation models, based on the value of each participant's contribution to the consolidated data, can significantly facilitate consolidation and the use of CT data. The Canadian equity experience shows that not resolving this issue results in less CT data usage.

The market structure continues to evolve and there are many legacy issues that need to be addressed; in particular, a lack of balanced governance and stakeholder representation at certain entities in equity markets that have exclusive responsibilities for consolidation and price setting. This is particularly important where some large TVs appear to have undue influence and possible conflicts of interest.

8 / THE OPTIMAL ARCHITECTURE FOR SUCCESSFUL DATA CONSOLIDATION IN EUROPE

This Chapter consolidates the feedback from participant interviews, the evaluation of available literature and the field research into other market models and recommends optimal architecture for European data consolidation based on the requirements identified by users in Chapter 4.

CHAPTER SUMMARY

1. The following are considered pre-requisites for successful CT data in Europe:

- A single, exclusive consolidated tape provider (ECTP) that is not subject to competition and is run as a utility should be created across both asset classes.
- All data aggregators of data should be obligatory members of this ECTP and should contribute towards its funding.
- The ECTP must be regulated and empowered by ESMA to enforce operating standards and rules via a harmonized set of rules, including penalties and other sanctions to maintain good data quality and behaviour.
- The ECTP should be recognised in the law and able to enforce market data law, as long as it has no conflicting interests.
- Balanced governance, with representation from different stakeholders, is required and no single stakeholder or stakeholder group should have undue influence over the ECTP.
- All pre-, post-trade and end-of-day data should be acquired and stored freely by the ECTP without contractual obligations. A revenue sharing mechanism to fairly share the revenue between contributors based on the value of each of their data sets to the consolidated data should be established.
- Sufficient pre-trade order and administrational event data must be provided to the tape and this must be defined in the law.
- The ECTP must have the appropriate tools to reconcile data. As such, reporting to the ECTP should be in real-time (no deferral management elsewhere) and double-sided trade reporting for non-matched trades should be introduced.

2. There are a further number of ancillary issues, which could also be addressed to enhance the use and viability of the data such as harmonising bond deferral regimes, increasing the population of bonds eligible for publication, separating RTS data formats for bonds and derivatives and introducing a vendor display rule.

8.1. PRE-REQUISITES FOR CT DATA

There are a number of pre-requisites which are considered critical for the development of EU CT Data. These should be viewed as the optimal and holistic foundation in order to successfully deliver CT data.

Figure 31: Pre-Requisites for Consolidated Tape Data

#	Pre-Requisites	Detail	Rationale
1	**An exclusive consolidated tape provider (ECTP) that is not subject to competition and is run as a utility should be created.** **It should have no conflicting interests.**	A single entity should be put in place across equities and bonds. It should not be subject to competition. In order to ensure the entity has no conflicting interests, it should most likely be run on a not-for-profit basis, but the final agreement on this would be established between stakeholders through the agreed corporate structure.	Technology alone cannot deliver CT data. To bridge the gap between the market requirements, current challenges and the technology available requires a single, organisational layer as the official manager of CT data governance and standards. This ECTP should be the undisputed authority and trusted source of EU CT data. Competition does not achieve this. Without the support and drive from this ECTP, the existing issues relating to collation, management and dissemination of data will only exacerbate as the industry's reliance on data increases.
2	**All data aggregators of MiFID II/MiFIR pre- and post-trade data should be obligatory members of this ECTP and contribute to its funding.**	The members of the ECTP must include all aggregators. QIFs and SIs should also be able to elect to self-aggregate and report data directly. If they do so they should pass a conformance test. These firms should also become members. All members should contribute to funding the entity. Membership categories with different rights and funding obligations for other data stakeholders could be introduced. The members should work together to establish the rules and ensure that all members follow those rules.	The ECTP will only be able to fully hold its members (the firms that aggregate data) to account through the concept of membership and enforcement of rules. Market participants should fund the ECTP as an incentive to improve the quality of the data they supply to the market. Aggregators are likely to pass on penalties to the underlying data generators so that market-wide behaviour changes.
3	**The ECTP must be regulated and suitably empowered by ESMA to enforce technical and operating standards via a harmonized set of rules, including penalties and other sanctions that have a sufficient impact on behaviour.**	This should include: • Defining the standards, inputs, outputs, obligations, enforcement rules, responsibilities and technical requirements for the working of the CT. • Determining the reasonable commercial price at which all data should be sold. • Determining the appropriate allocation and contractual mechanism for revenue sharing with contributors. • Defining, evolving and enforcing the rules, including controls and a penalty mechanism. • Selecting the outsourced infrastructure provider/s for data consolidation and dissemination through a competitive tender process. Translation to a common technical operating standard mandated by the ECTP must be required and should be the responsibility of the data generator or aggregator.	Without such powers, the ECTP cannot clean data or impose the standards that will improve data quality. There must be appropriate sanctions for not following the rules. The ECTP must be able to establish and enforce a penalty regime that fully holds its members to account. A common technical and operating standard must be followed in order to aid the compilation and cleaning of data.

#	Pre-Requisites	Detail	Rationale
4	**The ECTP should be recognised in the law.** **The regulatory authorisation of firms that are data contributors should be linked to membership of the ECTP.**	If the ECTP is a recognised body it should be able to enforce not only its laws but also have the powers (under the oversight of ESMA) to enforce European laws related to market data and the collection of it so that it can resolve issues quickly. A covenant should exist that states that if the entity becomes commercial or conflicted then it should lose its status. Any member of the entity must show that it has the systems and procedures in place to follow the ECTP requirements upon authorisation and on an on-going basis.	If the regulatory authorisation of firms that are data aggregators is linked to membership of the ECTP, then the ultimate sanction for not complying with the rules can be to prevent them from doing their broader regulated business. This will be particularly important for resolving difficult issues where aggregators and other data generators are resistant to change. The recognition of the ECTP in the law should allow for the fact that the role of the entity can evolve over time in which case its powers may need restraint in future.
5	**Balanced governance of the ECTP is required and no single stakeholder or stakeholder group should have undue influence.**	All types of data stakeholders should have input and be represented in the governance of the ECTP. This may include other stakeholders who are not data aggregators. Stakeholder categories would need to be determined and the rights of any one stakeholder type should be capped. A majority of independent directors will be required at Board level to ensure decisions are made. Stakeholders would also be represented through advisory committees.	All stakeholders should be represented to agree on principles, resolve the key issues and, where necessary, to find a compromise to deliver the successful launch and on-going management of the consolidated tape. However, to ensure resolution, a majority of independents at Board level will be needed to fairly evaluate issues and push through difficult decisions. Stakeholders can also be represented at different specialist sub-committees.
6	**The ECTP must be able to acquire and store all pre-, post-trade and end-of-day data freely without contractual obligations.** **A revenue sharing mechanism to fairly share the revenue between data aggregators based on the value of each of their contributions should be established.**	The data required is everything specified by the users in sections 4.2.1 and 4.2.2. Data should be provided to the ECTP at the same speed as it is made available to other recipients. See Appendix 17 for the full information required across asset classes. The value of the various data sources should each be assessed by the quality of their contribution to the CT data. This does not mean data should be free but that the mechanism for determining its value is run by the ECTP. It should be clear that the ECTP can store data for long-term historical use. A revenue allocation model is required to fairly assess the data in recognition of the value it has to the formation of the price. The revenue allocation mechanism used in the US equity markets (discussed in Appendix 9) is considered as a good foundation for liquid instruments.	Data aggregators should not be able to deliberately delay the data that they send to the ECTP in order to make their own proprietary feeds more attractive. To operate on a sound financial footing the ECTP needs certainty about the cost of data and so the ability to determine the price at which it can be sold. Ultimately it will be up to the stakeholder governance of the ECTP to determine the value of the overall data and how they sell that data in the market.

#	Pre-Requisites	Detail	Rationale
7	**Sufficient pre-trade order event, administrational event data and session statistics must be provided to the tape and this must be defined in the law.**	The data required is everything specified by the users in section 4.2.1. An official volume-weighted BBO could be calculated as required. Data will be stored and used for the CTP's and end-users' checks, reconciliations, reporting, modelling and other historical analysis as required. Data that takes out any non-public information would be available for end-users.	Without this pre- and post-trade information, a consolidated tape will not fully replace the information within the data feeds that users take today, and consolidation will be a waste of time. If 5 levels of depth are not given, then latency may be more of an issue and the introduction of lot sizes may be needed.
8	**Reporting to the ECTP should be in real-time (no self-determined deferral management elsewhere) and double-sided reporting should be introduced for non-matched trades.**	APAs or individual firms should no longer be able to determine and apply deferrals. Real-time reporting of bond data could allow ESMA to instigate a dynamic deferral regime for temporary spikes in bond liquidity. Two-sided trade reporting should be mandatory for trades executed outside of a matching system/CLOB with trade identifiers attached. A trade identifier should be generated by either the buyer or the seller (to be agreed) and attached to both sides of the trade report.	This allows the ECTP to immediately identify erroneous data and reporting discrepancies as soon as possible and will also help identify any systemic withholding of data. It is noted that two-sided trade reporting raised some concerns when tested with market participants in the workshop but without it, data reconciliation will be difficult.

8.1.1. Alternative Models Evaluated

This study also considered several alternative organisational models to provide oversight, governance and management of a consolidated tape. These alternative possibilities along with some relevant comments are as follows:

1. **TVs could act as exclusive aggregators of all on and off-venue flow for their listed instruments and/or establish a CTP under their combined jurisdiction.**
    - This is not recommended, as it would reinforce the position of the incumbent exchanges at the expense of TV competition and innovation and they are already deemed to have conflicts of interest. Additionally, the exchanges are already both aggregators and vendors of data. To position them as the exclusive aggregators of data gives them even more power than they have today at the expense of other stakeholders who would not have a role in the governance.
2. **Expanding the role of APAs to allow membership and fining capabilities.**
    - This is not recommended, as APAs are commercial businesses with potential conflicts of interest, making it unlikely that they would be able to enforce the rules effectively or would want to fine clients. Additionally, it is not clear how competing APAs would be able to set a single market-wide set of rules and standards.
3. **Creating exclusive aggregators for each instrument in pre- and post-trade and allowing multiple CTPs.**
    - This is not recommended as there will be no single body to mandate market-wide standards. It is also not recommended as it does not bring all the stakeholders together at the highest level to resolve governance issues and the aggregators may only work in the interests of a subset of stakeholders.
4. **The roles of the standards body and technical consolidators could be divided so that there is a single body that mandates standards and enforces rules and a technical layer of consolidators that compete to disseminate the data.**

- Under this model, some improvements to data quality could be achieved, but it would still require the creation of a single, exclusive body to create the standards and rules.
- However, this will not resolve governance issues, mostly in relation to cost, and consolidation will, therefore, remain unviable.

5. **ESMA could be the entity that runs the ECTP with internal resources.**
   - This is not recommended as it does not have significant resources or data expertise.
   - There is a risk that if market stakeholders do not have some "skin in the game" over issues such as cost, consolidation will be a waste of time.

8.2. OTHER IMPORTANT ENHANCEMENTS TO IMPROVE DATA USABILITY AND QUALITY

Figure 32: Additional Enhancements to the Recommendations to Improve Data Usability and Quality.

#	Enhancement	Detail	Rationale
1.	**Bond deferral harmonisation.**	Deferral regime of government bonds needs to be harmonised.	Without harmonisation, comparing government bonds will be very difficult and government bond data will not be very useful.
2.	**RTS definitions of fields and formats must be enriched and separated for each asset class.**	A separate RTS for the fields/reportable information for bonds and derivatives must be created. A flag denoting the trading system, segment and session is also required. A non-price forming and price-forming flag should be added. (Noted that this means having a negotiated trade waiver for fixed income).	Bonds and derivative data requirements are all in one RTS, creating confusion (e.g. notional figures are meaningless in certain contexts). Users need to know under which market trading session and segment protocol a trade occurred and translating from bespoke standards is difficult. A price forming and non-price forming flag would be of immediate assistance in identifying addressable liquidity.
3.	**Ambiguities should be removed as to what is on-venue and off-venue reporting.**	Firms that make quotes both on-venue and off-venue and then execute off-venue at a price inside the spread being quoted on-exchange should not be allowed to report the trade as on-exchange.	The data generated by the TVs that report these trades is confusing and it allows firms to avoid being classified as SIs.
4.	**The population of the bonds eligible for inclusion in the consolidated data tape should be increased.**	The population of bonds that are available to be published needs to be broadened.	This makes the data more viable for the ECTP and more useful for the market.
5.	**Introduce a vendor display rule and mandate the data to be used for free.**	Retail brokers should display CT data to end investors prior to and at the point of trading. Ultimately a compilation of quotes may be helpful for display to retail investors, as many retail trades are executed off-venue.	This will educate investors about the underlying market structure and help them to challenge best execution. Data should be free to prevent aggregators from creating substitute products. Otherwise, retail brokers may be incentivised to economise by finding cheaper but poorer quality data that could be supplied by the TVs.

#	Enhancement	Detail	Rationale
6.	**An official definition of an EU instrument.**	This would help the ECTP to delineate EU versus non-EU activity to provide different data sets to customer segments.	Some participants may not need to see and process all market activity.

8.3. OTHER ADVANTAGES OF A SINGLE ECTP ARCHITECTURE

An ECTP acting as the sole, independent, official source of data could deliver other advantages such as:

1. Improved market transparency management e.g. A **dynamic deferral regime** could be calculated for temporary liquidity spikes in bonds.
2. Undertaking regulatory calculations on behalf of ESMA.
3. Disseminating reference data for ESMA.

8.4. CONCLUSIONS

The optimal architecture to develop EU consolidated data at this point in European financial market evolution is one that rejects the possibility of competition between consolidators and moves to the adoption of an exclusive CTP that is not conflicted and has self-regulating capabilities.

The pre-requisites outlined in this chapter must be adopted holistically in order to successfully deliver the data requirements described in Chapter 4 and any compromises are likely to limit the successful development of CT data. The viability of the data could be further improved by additional enhancements that would certainly make the data more useful and a single official source of data may be used to harness other efficiencies for participants such as helping with regulatory calculations.

9 / OPERATIONAL AND TECHNICAL DESIGN: FEASIBILITY, COSTS AND FUNDING

This Chapter assumes that the recommendations in Chapter 8 are adopted and describes the high-level organisational design of the ECTP and the technical design required to deliver the CT data. The designs and budgets are based on extensive discussions with vendors, technology providers and organisations with similar mandates and technical requirements.

CHAPTER SUMMARY

1. CT data for both equities and bonds can be delivered via the same high-level organisational and technical designs. This allows for economies of scale and the flexibility required to phase in different instruments and functionality over time.

2. Organisational Design:

- There would be a Board, and a permanent Executive and staff with advice given through stakeholder committees advising on product and technical requirements and rules.
- The key decisions and activities of the ECTP would be to define and implement:
  - The strategy and business plan and risk management policies and processes.
  - The standards, inputs, outputs, obligations, enforcement rules, responsibilities, and technical requirements for its members.
  - The cost at which to sell data and the revenue sharing mechanism.
  - The tender process for the outsourcing of the technology and the on-going relationship with the supplier/s.

3. Technical Design:

- The technology provided by a third-party supplier, which would allow for equity data to be delivered in tens of milliseconds, is expected to include:
  - "CT Engines" with core processing logic would be used to interact with other components and provide "plug and play" flexibility to allow for new functional elements to be added and allow for scalability.
  - Machine learning algorithms to identify data anomalies.
  - Operations in two data centres in Europe, which could be expanded as required. ECTP stakeholders would decide their locations.
  - The cloud for storage of historical data.

4 Costs and Funding

- The setup costs of the ECTP, including the technology for both asset classes, are expected to be €11 million. Start-up funding could be raised by levying a one-off joining fee from data aggregators with an average contribution of €25,000.
- The annual running costs of the entity for all asset classes in scope, including the running of the on-going technology, are estimated in the range of €6 million to €7 million. This could be obtained with a membership fee levied on data aggregators of **an average of €16,000** per entity per annum.
- Revenues from data sales would be allocated back to data contributors. Profits generated from other activities could be returned to members.

9.1. ORGANISATIONAL DESIGN

The key objectives of the ECTP entity would be to bring cultural change to the management of market data in Europe and fundamentally improve the governance, preparation, and distribution of the CT data.

The entity would be owned, run and funded by members and other stakeholders. The organisational detail and operating model of the ECTP would be determined through the governance structure and the scope of the work would include:

- Setting up the organisation together with its governance structure, relevant committees, terms of reference, etc.
- Hiring and establishing the Executive.
- Defining the CT data offering for pre- and post-trade products.
- Overseeing the procurement process to select optimal technology and other support provider(s) and subsequently negotiating their contract(s), finalising project plans, managing SLAs, etc.
- Determining fair and optimal pricing and reward structures in agreement with all stakeholders.
- Harmonizing data contracts and definitions with data contributors and for the redistribution of data.
- Undertaking the audits of data usage for itself and possibly offering a similar service for TVs and APAs that already use the services of other third parties to undertake audits on their behalf.
- Holding historic data that could be interrogated by the market in one place.
- Monitoring compliance, enforcing rules and levying penalties or fines as appropriate.

9.1.1. Governance

Under the oversight of ESMA, the ECTP will exercise effective operating authority over the entity through a balanced governance structure, whereby all key stakeholders have a say as to how the ECTP should be run.

A Board would govern the entity with a majority of independent directors. Material stakeholders would be represented in the governance of the entity even if they are not direct members (discussed below) of the entity. Voting rights of any one stakeholder type would be capped. The type and numbers of stakeholders to be represented and how directors are appointed would need to be addressed in the company's Articles of Association and any shareholder agreements.

ECTP members would be directly represented through the election of some of the directors. ESMA may wish to assert some authority and input into the management of the ECTP and it could approve the Chair, the Independent Directors and the CEO or even appoint them.

A committee structure, made up of industry representatives and under the Board's oversight, should ensure that there is subject matter expertise focussed on the key functions of the ECTP and that the input and feedback provided to the Board results from good cross-sector representation with no particular group having undue influence.

Figure 33: Proposed ECTP Organisational Structure and Responsibilities

Functional Area	Activities
Board	Strategic direction, business planning, performance monitoring, priorities, policies, risk, regulation, oversight of Executive and Sub-committees, stakeholder engagement.
Executive	Day-to-day management & control, legal, finance and business management functions, contributor/user/provider contracts and management, PR, general compliance, human resources and staff welfare, secretariat, and support for Board & sub-committees.
Product Advisory Committee	Per Asset Class (or sub-asset class): product development priorities, product design, pricing, and revenue allocation recommendations. The Product Advisory Committee could advise on cost and pricing of data or this could be dealt with by a separate committee or at the Board. The ECTP must be in a position to determine the price at which it receives the data and the price at which it sells it. It will do this in consultation with the parties involved. An appropriate methodology for revenue sharing and cost allocation will be established and approved by the Board.
Technical Advisory Committee	Per Asset Class (or sub-asset class): Technical operating standards, architecture, record structures, technical development priorities, capacity/scaling/latency, selecting and monitoring an out-sourced provider for the technology, SLAs/KPIs and monitoring thereof.
Compliance and Oversight	Rules required to ensure data integrity: completeness, accuracy, timeliness & consistency; audit oversight; penalty regime; appeals & arbitration process.

9.1.2. Staffing

The organisation would be run by a permanent staff of approximately 22 people (staff estimates and budgets are given in Appendix 14), which would cover:

- *The Executive*
- *Product Development*
- *Client and Member Services*
- *Operational Support/Helpdesk*
- *Technology liaison**
- *Compliance, Audit and Risk*
- *Financial Control & Admin*
- *HR*

**It is assumed that the technology would be outsourced so the cost of people to develop and run the technology is included in the cost of the technology.*

9.1.3. Membership of The Entity

Membership categories would need to be established and managed for the data aggregators. There is also an argument for having separate membership categories for other material stakeholders, such as data users and data vendors.

Each category could have different rights and obligations such as the right to nominate and elect representatives for its category onto the Board and sub-committees of the ECTP. Other data generators could have their memberships managed and handled through the APAs, e.g. the APAs could handle the membership costs or pass-through queries and fines.

9.1.4. Compliance and Oversight

At an operational level, it is important to ensure that the membership and constitution of the Compliance and Oversight Committee are robust with suitable access to legal resource and regulator(s) support where necessary. There should also be a clear and well-defined appeals process to support the integrity of this function.

To help ensure compliance with the rules and good behaviour, the ECTP could either levy fines or other penalties or, where relevant, hold back revenue share payments if the miscreant is a data aggregator. These policies would need to be submitted to ESMA for approval.

9.1.5. Procurement of IT Services

The ECTP is not expected to undertake any significant IT development in-house. There are many potential third-party providers for the required capture, processing, dissemination, and storage of data who are likely to be interested in tendering for the business. The procurement process would follow the usual rules for public procurement within the EU. The quality and completeness of the specified requirements and the efficacy of the procurement process will be critical to the success of the project and will be highly dependent on the quality of the work in defining the ETCP.

A single technology provider for the ECTP would, however, be preferable as it avoids any contention as to who is responsible for what but it may be decided that having specialist providers for different aspects may be desirable, such as equity versus non-equity or sub-asset classes.

Once the provider(s) is selected contract negotiations must be completed and service level agreements (SLAs) will be required. The information obtained during the procurement process will enable the project plan to be finalised and the budget to be updated. Developing and testing will be needed with stakeholders playing an active part in testing and rehearsals.

Monitoring the performance will be critical, both to ensure that the provider(s) is fulfilling its contract and SLAs and also as a basis for further developing further products and services going forward.

9.2. HIGH-LEVEL TECHNICAL ARCHITECTURE OVERVIEW

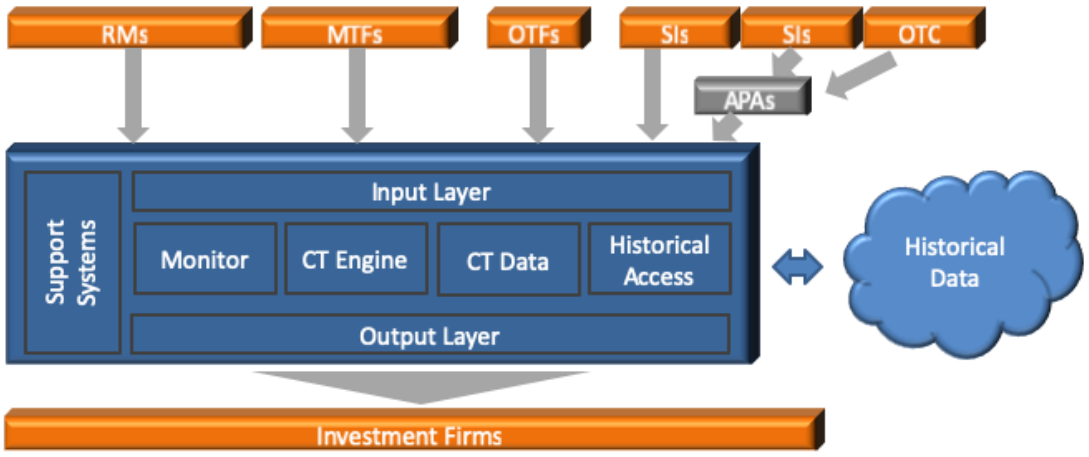
It is assumed that the technology required to support and manage the consolidated tape would be outsourced to a third-party provider. However, the following is an outline of the main technical components of the proposed tape, based on the research and interviews with data companies and market participants and the MSP team's own knowledge of building similar systems.

Although they capture different data points, the overall design of any pre- and post-trade consolidated tapes will be the same.

A key aspect of this design is the concept of "CT Engines". These are the core processing units, running the logic for taking the inputs, and assessing and creating the outputs. These engines will interact with the other components identified but would be designed to be interoperable ("plug and play") with other components, allowing multiple engines to be deployed to make the architecture highly scalable. This enables new functional behaviour to be added as required. It also allows for different parts of the tape to be delivered over time, e.g. equities post-trade, bonds post-trade, equities pre-trade, bonds post-trade.

The image below shows the overall technical architecture for a Consolidated Tape.

Figure 34: Overall Technical Architecture for a Consolidated Tape



Source: Market Structure Partners Research

Figure 15: Overall Technical Design Descriptions of a Consolidated Tape

Key Components	Function
Input Layer	Connectivity into the CT supported by a FIX API to receive different types of messages: *Pre-Trade* *Post-Trade* *Market Events*
Monitor	Machine learning engine to detect patterns that indicate erroneous reporting including the ability to match reports where both sides of the trade are reported.
CT Engine	Core processing logic that generates the pre- and post-trade tapes. This engine will consolidate and sequence the reports across the multiple data sources. Note that multiple engines can exist, processing data in parallel – e.g. separate engines for bonds and equities. This component would also ensure that only one side of the trade is reported.
CT Reference Data	Data needed to support the technology including data such as TV identifiers, firm identifiers and mappings, access permissions, asset class and instrument identifiers and other reference data. Intra-day tape and events would also be stored in a fast-access storage medium within the core system.
Historical Access	This component controls the data that consumers can obtain from the historical data source.
Output Layer	Connectivity to CT data consumers. A standard API to provide access to different types of data. *Pre-Trade* *Post-Trade* *Historical* *End-of-Day* *Venue Events* The API would support specification of the type of data that a subscriber wishes to receive – e.g. equities or bonds for the pre- and post-trade CT, day, or time intervals for historical and end-of-day etc.
Support Systems	Ancillary systems that the CT organization will need to support it – e.g. HR systems, General Ledger, Audit etc.
Historical data	Use of a cloud-based service to store data for historical purposes.

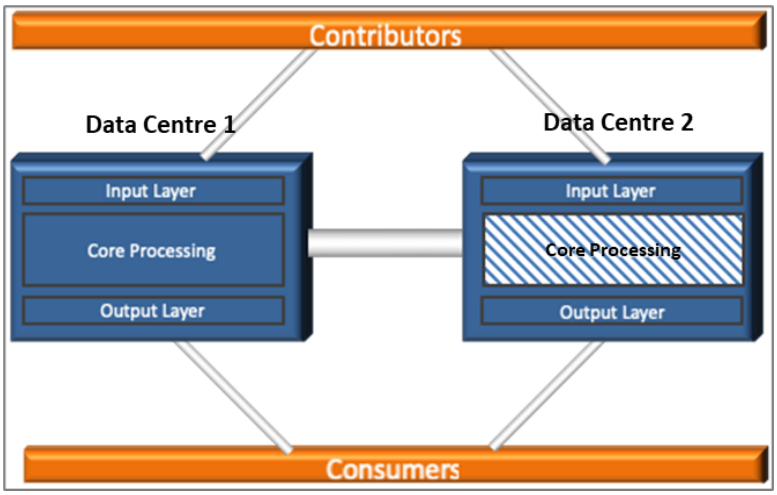
The output layer would take the form of a multicast distribution (see Chapter 3 for further explanation). As this is a non-guaranteed protocol, the CT would need to support a request/refresh snapshot to enable organisations to detect and "repair" data loss. This approach is consistent with the currently existing consolidated tape providers in the US.

The design and architecture would provide scalability and readily support additional capacity by increasing the number of CT engines as needed. This design can process data for both equities and bonds but can also be deployed for a single asset class at a time, adding more engines as needed.

The majority of the cost incurred for this architecture would be for the hardware to operate the system and for the storage of data. However, once a base infrastructure is in place, any additional costs become incremental (linear step function) in relation to the message throughput from the market.

Taking into consideration the use case requirements, which suggest that nano-second latency is not required, a single data centre would technically be sufficient. However, using two data centres is believed to be more prudent for resiliency.

Figure 36: High Availability Design Over Two Data Centres



Source: MSP Research

- *Each data centre (represented by the blue boxes) would hold the same data but each would be located in a different region within Europe. The two data centres would be linked with a high-speed connection to route traffic between the sites.*
- *Data generators and other contributors would connect to both data centres. However, they can contribute their data to either and should manage business continuity between the sites per their own recovery procedures.*
- *End-users would be recommended to connect to both data centres however this would not be mandatory. End-users can take messages and other data from one site or both. Taking from both sites would improve resiliency but would also require end-users to manage the same data from both sites.*
- *Regardless of the end-user's choice of data centre, the core processing for the CT would be carried out in one primary site with the secondary acting as a hot failover. This means that there is a seamless transition for the CT calculations from one site to the other.*

9.2.1. Cloud Storage

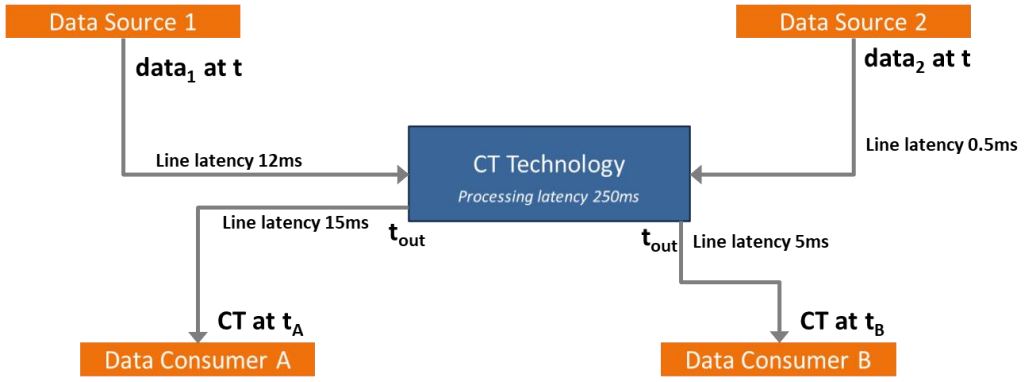
Pre- and post-trade historical data will be stored and archived in a cloud storage medium and available to users through an output layer via a standard API to download for their own use.

9.2.2. Data Centre Choices and Platform Latency

The platform latency will be subject to physical network limitations but the proposed technology, which is already available and widely used should allow the data to be processed in the order of tens of milliseconds.

To illustrate the practical implications of latency, the example below assumes that two venues, Data Source 1 and Data Source 2 generate data at time t such that $data_1$ and $data_2$ should appear in an aggregated form (or as a contiguous set of data).

Figure 37: Illustrative Latencies Example



		$data_1$	$data_2$
①	Time CT receives data	t+12ms	t+0.5ms
②	Data available at the output	t_{out} = t+12+250ms	t_{out} = t+12+250ms
		Consumer A	Consumer B
③	Data available at consumers at	$t_A = t_{out}$+15ms = t + 277ms	$t_B = t_{out}$+5ms = t+267ms

① The CTP would receive data from all data sources per their connectivity latencies. However, the CTP will need to allow time to "wait" for the sequencing of all data for a given time, t, to be received before processing data for the consolidated tape at that time

② In this example the tape will need to "wait" 12ms to before it can process like for like on both inbound streams - data will be available at an outbound time taking into account additional processing within the core technology – e.g. aggregation, stats etc. Time at the outbound layer is "normalized" by having to wait for all inbound messages to be at the same point in time i.e. at the time t_{out}.

③ Consumption of the data will again be dependent on communications latencies. In the example above, two end-users, one close to a distributor versus one further away would receive data 10ms apart. This is directly a consequence of the end-user's latency difference to distribution points.

This example highlights the following key design features for a consolidated tape:

1. The ECTP will need to account for the different latencies on inbound messaging to correctly align and sequence the data for the consolidated tape.
2. The end-users will receive the outbound messages resultant from the consolidated tape at different times based on the network latency from the distribution centre to each respective end-user.
3. Due to the multicast nature of the outbound messaging, end-users may also experience further delays if they have lost or are missing data packets and need to invoke a refresh and restart to their data feeds.

There are multiple data centres available within Europe where the ECTP could be based. A decision on the best locations would be made under the governance of the ECTP taking account of the outcome for all stakeholders. Those further from the data centre may experience higher latency but if they have access to five levels of depth then this should not be an issue. Timestamping will, however, be critical to allow all participants to construct the same view of the consolidated tape across the market.

Further improvements to latency could be achieved by using additional data centres in but ultimately the overall time delay from an order book change or trade execution is subject to overall network latencies across Europe.

Any further improvements in latency must also consider whether the benefits outweigh the increased monetary and environmental costs. Whilst these environmental costs are difficult to quantify it should be noted that the greenhouse footprint would be linear with the number of data centres utilised whereas the improvements in latency may only be marginal and superfluous to the requirements of some end-users.

9.3. BUDGETS AND FUNDING

9.3.1. Assumptions

- The ECTP is likely to be a not-for-profit organisation. Any profits not required for future investment can be returned to members.
- The ECTP should effectively stand in the middle of the revenue being paid by the data users to the data contributors, potentially extracting a small fee for its services where appropriate. The ECTP should not be in the position of taking a commercial risk as regards the cost or price of the data.
- The ECTP would cover its initial set up cost and on-going operational funding from a number of sources such as:
    - Membership fees.
    - Administration fees for handling revenues from data users to data contributors.
    - Integrity audit fees, fines for non-compliance, etc.
    - Trade reporting fees for SIs wishing to report data directly to the ECTP.

If the above sources of funding were not sufficient then the ECTP could retain a percentage of data sales revenue to help cover its costs.

An alternative or complementary approach would be for the ECTP to levy a small charge for each transaction it has to process.

9.3.2. Source of Funding

Applying the membership concept that was outlined in the Governance section (9.1.3), the primary source of funding for the ECTP would be from membership fees of the data aggregators.

The initial set up costs would be funded by a one-off ‘entry fee’ to become a member and the on-going operational costs would be funded by annual membership fees.

To demonstrate how this revenue could be raised, consider that in Europe there are 482 potential providers of aggregated data across all asset classes. This includes all data aggregators and SIs that may elect to self-aggregate. The table below illustrates how relatively modest membership fees could be used to cover both the initial set and annual funding costs of the ECTP.

In this example, a one-off entry fee of approximately €25,000 per member would raise over €11 million to cover set up costs, whilst an annual membership fee of €16,000 per annum would cover annual operational costs. Note that these figures are averages only. The actual membership fees would be

determined by the ECTP and could be based on the size and nature of each member and the benefit they are likely to derive from the allocation of revenue from the sale of CT data. An alternative source of set up funding could be sought through a loan..

Once the ECTP is up and running then the membership fees should be reviewed on annual basis to ensure that they are fair to each member and sufficient to support the ECTP in the future expansion and development of the consolidated tape in Europe.

Figure 38: Potential ECTP Number of Members and Annual Fees

INPUTS (Data contributors)	RM	MTF	OTF	APA	SIs, OTC*	Total
All Assets Classes	#	#	#	#	#	**#**
EU	121	139	25	16	15	**316**
UK	15	81	49	6	15	**166**
Total	**136**	**220**	**74**	**22**	**30**	**482**
	€k	€k	€k	€k	€k	**€k**
Average Annual Membership Fee €k	16	16	16	16	16	**16**
Total Annual Membership Fees €k	**2,176**	**3,520**	**1,184**	**352**	**480**	**7,712**
	€k	€k	€k	€k	€k	**€k**
Average One Off Set up Fee €k	**25**	**25**	**25**	**25**	**25**	**25**
Total Set Up Fees €k	**3,400**	**5,500**	**1,850**	**550**	**750**	**12,050**

**Estimated direct contributors to ECTP*

9.3.3. Conclusions

CT data for both equities and bonds can be delivered via the same high-level organisational and technical designs. This allows economies of scale to be achieved and provides for the flexibility required for the phasing in of different instruments and functionality over time.

The proposed organisational and technical design to support the ECTP and deliver the CT data requirements is feasible and could be implemented.

An ECTP that would meet user requirements could be established for €11 million with on-going running costs of approximately €7.6 million, all of which could be funded by the industry for a relatively modest membership fee.

10 / CONSOLIDATED DATA IN THE CONTEXT OF THE EU 27 MARKETS (EX. UK)

This Chapter considers the requirements for CT data in the context of the EU 27 markets (i.e. post the departure of the United Kingdom ('UK') from the EU, where equivalence is not granted to the UK).

CHAPTER SUMMARY

1. The impact of the UK's departure from the EU has been part of the backdrop to the analysis undertaken for this study but cannot yet be fully assessed as the terms of that departure have not been finalised. However, the findings from this study suggest that the **need for CT data within the EU 27 markets** remains compelling and will be **equally important** without the UK because:

- There is no discernible difference between the requirements and issues raised by participants within the EU 27 markets and outside of them. All participants in EU 27 markets, regardless of their location, still require quality CT data to manage their portfolios, risk and regulatory obligations and perform the other functions as identified in the use cases in this study.
- In a scenario where passporting rights or equivalence is not given to the UK, there will be a **proliferation of additional TV, SI, QIF and APA entities** set up within the EU, as UK based firms seek to continue doing business within the EU.
- The results will be twofold. Firstly, there will be a significant **increase in the number of entities** from which participants (or a CTP) will need to collect data. Secondly, there could also be an **increase in the number of overall trades** because QIFs and SIs are likely to undertake back-to-back trading between their EU and non-EU entities in order to manage client orders and risk.
- The overall effect is that it is likely to **introduce more complexities** and issues with data quality and is also **likely to further increase costs,** which participants already complain about.
- Specifically, for equities, MiFID II's Share Trading Obligation (STO), means that **equity liquidity pools are likely to be split** between the UK and the EU. This may generate **arbitrage opportunities** and **reduce trade sizes on lit markets** if data cannot be relied upon to give a complete picture of the market. This will contribute to more fragmentation and **larger volumes of trade reports** and put an even greater emphasis on the importance of **quality pre- and post-trade data**.
- If a third country creates **competing offshore pools** of liquidity in EU instruments for non-EU 27 investors, CT data will help the EU compete for capital and liquidity from those investors.

10.1. THE IMPACT OF THE UK'S DEPARTURE FROM THE EU ON CT DATA

The impact of the UK's departure from the EU has been part of the backdrop to the analysis undertaken for this study but cannot yet be fully assessed as the terms of that departure have not been finalised. It has not, therefore, been possible to fully examine the impact of the UK's withdrawal on these recommendations.

However, it is possible to outline the possible high-level scenarios under which the UK may leave and how the different outcomes attached to each of these scenarios will impact data stakeholders including TVs, APAs, investment firms, investors, regulators and other participants.

These high-level scenarios are as briefly described as follows:

1. The UK **retains its passporting privileges** similar to other EEA countries, which means that firms operating their business from the UK can continue to do so without the need to establish entities within the EU.
2. The UK **loses its passporting rights but is granted third country equivalence** which allows non-EU based firms to offer a limited number of services into the EU if their home country regulatory regime is accepted by the EU as being 'equivalent' to EU standards. This would mean that some investment firms and TVs may still be able to operate outside of the EU without the need to establish a presence in the EU.
3. The UK **loses passporting rights and is not granted equivalence** in which case UK based entities would need to establish a business presence within the EU.

This study incorporated feedback from EU 27 and UK investors who describe their desire for CT data as being "pan-European". This includes UK data and it can be reasonably assumed that in Scenarios 1 and 2, the EU and the UK would work together to progress the regulatory, operating and technical framework required to establish and develop a consolidated set of data across Europe. In this case, the analysis and recommendations in this study are applicable to the combined EU 27 markets and the UK.

However, in Scenario 3, the UK and EU may not co-operate to build combined CT data and the EU will consider whether it needs to develop CT data for the EU 27 markets and whether the analysis and recommendations in this study remain relevant.

10.2. CONSOLIDATED DATA REQUIREMENTS FOR EU 27 PARTICIPANTS AND INVESTORS

This study found no discernible difference between the use cases, requirements and issues raised in relation to data consolidation by investment firms and market participants that operate their European operations from the UK and investment firms that operate their European operations from within the EU 27 markets. After the UK's departure, the EU 27 markets will continue to compete to attract capital and investment from global investors and, as such, will benefit from setting high standards of market transparency and efficiency.

Most stakeholders, apart from possibly the incumbent trading venues, see CT data as an important tool for enhancing their business processes and fulfilling their use case requirements. Ease of access to good quality consolidated data with respect to both UK and EU 27 markets will remain a critical issue for all participants, both within and outside of the EU, post the departure of the UK.

10.3. THE IMPACT OF SCENARIO 3 ON EU 27 MARKETS AND DATA

UK-based market participants have been preparing for Scenario 3 and have either already established, or are preparing to establish, a presence for their businesses in the EU 27 markets. This presence typically takes the form of a separate EU domiciled entity within the firm's corporate structure. In the event that passporting or equivalence is not granted these entities will step up their activities and there will be a proliferation of QIFs, SIs, TVs and APAs within the EU from which participants will need to source data in order to manage portfolios, risk and best execution obligations. This burden will fall on all global participants who currently invest in EU 27 and 28 markets.

The results will be twofold. Firstly, there will be an increase in the number of sources from which to collect data and second, there may also be an increase in the number of overall trade reports This is because QIFs and SIs are likely to undertake simultaneous trading (known as back-to-back trading) between their EU and non-EU entities in order to manage client orders and risk.

10.3.1. Share Trading Obligation

Specifically, for equities, MiFIR's Share Trading Obligation (STO)[69] and more recent ESMA guidance mean that EU investment firms must ensure that the trades they undertake in shares that are admitted to trading on a regulated market, or traded on a trading venue in the EU, take place on a regulated market, MTF or systematic internaliser within the EU or a third-country trading venue that is considered to be equivalent to a regulated market by the European Commission.

As a result, equity liquidity pools are likely to be split between the UK and the EU and EU firms will only be able to trade EU shares on an EU venue regardless of where liquidity or best price may be achieved. Non-EU investors could trade in EU instruments in liquidity pools based outside of the EU[70]. This may generate arbitrage opportunities and reduce trade sizes on lit markets if data cannot be relied upon to give a complete picture of the market liquidity. It may also increase the number of trades being reported and data that needs to be managed.

Exchange-Traded Funds (ETFs) will also be impacted even though they are not subject to the trading obligation. Related participants including brokers, liquidity providers and fund managers that are subject to MiFID II will need to consider the STO when trading in the underlying securities that are subject to the STO. This will need careful data management.

Additionally, in the event that the UK, or any other non-EU jurisdiction, creates competing offshore pools of liquidity in EU instruments for non-EU 27 investors, CT data will be very important for the EU. Easy access to a complete and accurate picture of liquidity in the EU will help the EU compete for capital and liquidity from non-EU investors with large amounts of capital to invest that could be drawn to pools of liquidity elsewhere, particularly if there is a more complete and accurate picture of an alternative market. A failure to provide good data could, in turn, impact primary markets by encouraging EU domiciled companies to seek listings or capital raising outside of the EU 27 markets.

10.3.2. Two Consolidated Tapes?

It is likely that an ECTP in each jurisdiction would need to be created to solve the current data issues that exist both in the EU 27 markets and the UK but the creation of these entities and their recognition in the law may be treated differently. This would clearly increase the costs to the market and if the proposed funding model is adopted, these costs are likely to be borne by similar sets of stakeholders who are active in both markets.

Although this would be less efficient, if the costs described in this study were charged for both tapes, it would still outweigh the benefits for investors. The most important focus of data users would be to ensure that the technical and operating standards of both CT consolidators were as closely aligned as possible so that the inefficiencies could be reduced, particularly when cleaning and compiling the data. This may be achieved by the similar stakeholders being represented in the governance of both ECTPs.

Whichever jurisdiction commenced the development of CT data first could take the lead in establishing these standards.

10.4. CONCLUSIONS

The requirements and benefits identified for investors and regulators in this study continue to apply to the EU 27 markets regardless of the direction that the UK takes and there is still a very compelling case for CT data.

The most likely effect of the UK's departure from the EU is that more data will be generated from within the EU 27 markets than there is today and, in the absence of a consolidated tape, this will increase the complexities, issues with the quality of data and costs that participants already complain about. It may also be even more critical to develop EU 27 CT data if the UK leaves without equivalence in order to ensure that the capital of non-EU investors' capital is attracted to EU 27 markets and not to third countries.

Even in the less desirable scenario that two tapes must be developed, based on the findings in this study, the benefits of building two consolidated tapes would still far outweigh the costs of building them.

11 / LEGAL ANALYSIS

This Chapter considers which aspects of a consolidated tape can be implemented within the current legislative framework versus which aspects would require further legislative amendments before they can be implemented. The full detail behind this legal analysis is set out in Appendix 19.

CHAPTER SUMMARY

1. **The EC has sufficient powers under current MiFID II legislation to direct ESMA to establish a consolidated post-trade tape by using its procurement process to:**
   - Establish an exclusive CTP (ECTP) to run a consolidated tape utility with self-regulating capabilities under ESMA's oversight.
   - Require that TVs and APAs, as key aggregators, become mandated members of this ECTP. Other key stakeholders could be non-mandated members.
   - Empower the ECTP to develop its rule book (including standards) and compliance framework (including sanctions and penalties).
   - Ensure that:
     - post-trade data for the consolidated tape would be provided free to the ECTP, which then determines the price for that data to be sold to end-users.
     - revenue from the consolidated post-trade tape would be shared via a revenue allocation model as agreed by the stakeholders.
     - post-trade data would be available to end-users in real-time and could also be accessed on a historical basis.
2. **Under MiFID II, many other market data issues could also be resolved as follows:**
   - With political goodwill, bond deferral rules could be harmonised.
   - ESMA could increase the population of bonds in scope for post-trade reporting.
   - RTSs could be amended to improve data quality and ease the ECTP's task of reconciling the data.
3. **The following cannot be achieved without new, or amendments to, legislation:**
   - A consolidated pre-trade tape is not possible because TVs are not currently mandated to submit pre-trade data to any CTP.
   - Recognition of the exclusivity of the ECTP in law thus empowering it to enforce European market data laws and align membership of the ECTP with authorisation of the data generators and aggregators.
   - Mandating other data generators such as SI's to be ECTP members.
   - Consistency in the data submitted by TVs and SIs, as they are not subject to the same pre- and post-trade transparency regimes.
   - Formal delineation between EU and non-EU instruments.
4. **To move forward the EC is faced with two possible alternatives. These are:**
   - **Option 1:** Use the powers already provided by MiFID II to implement an initial post-trade consolidated tape and introduce further legislation to support a pre-trade consolidated tape and the enhancements noted above, or,
   - **Option 2**: Defer any development on a consolidated tape until the full legislative framework is in place to support a complete pre- and post-trade consolidated tape.
5. **The pros and cons of each alternative are discussed in more detail within this chapter.**

11.1. ANALYSIS OF CURRENT POWERS VERSUS THE PROPOSED REQUIREMENTS

The following table outlines what is possible under the current law and where amendments would be required. The assessment of what can be achieved depends on a robust interpretation of the law and for the EC to provide clear specifications to ESMA as to what is expected of the ECTP and how it should be structured.

Figure 39: An Assessment of What Can Be Achieved Under Current Legislation.

Key: ● Yes ◐ Partial Achievement ○ No

Requirements	Can it be Achieved Now?	What is Achievable under Current Legislation	Proposals for Future Legislation
TO ESTABLISH AND EMPOWER A SINGLE, NOT-FOR-PROFIT SELF-REGULATED ENTITY WITH CENTRALISED OVERSIGHT FROM ESMA			
1. A single non-conflicted layer between the regulators and the technological solution.	●	The EC has the power to request ESMA to set up an exclusive CTP, answerable to ESMA. The EC can set certain parameters that ESMA must consider and ESMA can use its current procurement process to establish an ECTP following the EC's desired specifications. ESMA and a number of stakeholders would need to work together to effectively create the entity that meets the stipulated requirements. Through the procurement process, the ECTP can be empowered to determine its own suitable rules, standards, and behaviours for the CT data to succeed. However, it cannot enforce the related European laws or fully sanction members. Balanced governance and lack of conflicts of interest is possible to achieve and should be stipulated as a requirement in the procurement process. It can also establish that a "commercial entity" means it can be a not-for-profit entity. The organisation should be able to undertake audits of members and consumers. An ECTP should be able to be directly authorised and regulated by ESMA (note that this is possible from Jan 2022[71]).	The ECTP needs to be recognised in the law and empowered to enforce not only its rules on its own members but also to enforce the European laws directly relating to the consolidated tape and reporting of data. There should be a covenant to ensure that if the ECTP becomes conflicted that its powers under the law can be curbed or retracted. When any firm that aggregates MiFID II data (either as a self-aggregator or a market aggregator) becomes authorised, it should be required to show that it has arrangements in place to provide data to the ECTP. These arrangements must be ongoing with the ultimate sanction being that an authorised firm could lose its authorisation if it is not willing to abide by ECTP rules.
2. The ECTP must be suitably empowered by ESMA, to enforce operating standards and rules via a harmonized set of rules, including penalties and other sanctions that have a sufficient impact on behaviour.	◐		
3. The ECTP should be recognised in the law, as long as it has no conflicting interests.	○		
4. Balanced governance of the ECTP is required.	◐		

Requirements	Can it be Achieved Now?	What is Achievable under Current Legislation	Proposals for Future Legislation
MANDATED MEMBERSHIP FOR ALL DATA GENERATORS AND CONTRIBUTORS			
1. All data generators and aggregators of MiFID II/MiFIR pre- and post-trade data should be obligatory members of this ECTP and will contribute to funding.		The ESMA procurement process could ensure that there is a compulsory membership of the ECTP for data aggregators. However, APAs and TVs are the only entities that could currently be mandated to become members of the ECTP. QIFs and SIs do not have the option to make data directly available to a CTP (but they could become voluntary members or it could be resolved it the ECTP became an APA, allowing them to report directly.	Allow all data generators/contributors of MiFID II pre- and post-trade data to report directly to the ECTP. .
EQUITY PRE-TRADE DATA CONSOLIDATION			
1. To collect pre-trade order data including 5 levels of depth, auction imbalance data and session statistics.		The European Commission will not be able to use its existing delegated powers to mandate a pre-trade consolidated tape or to mandate that firms or entities other than APAs and trading venues (TVs) submit data to it. It may be possible to ask for voluntary sets of data from TVs but unlikely that all TVs would agree.	A core set of data and administrational event data that must be given up by exchanges has to be stipulated in the law and it must be clear that this data should be freely given up to the ECTP. The concept of "Historical Data" needs to be introduced, defined and mandated as data that must be given to the ECTP for long-term storage and it must be clear that this should be freely given up to the ECTP. Once this data is clearly defined in the law, the ECTP should be mandated to provide pre-trade equity order data in real-time and have historical data stored for analysis. The ECTP must be able to establish: • the reasonable commercial basis for which pre-trade data should be charged to consumers. • the revenue allocation mechanism under which pre-trade data revenue would be shared with its members. Although quote data is currently not a "must-have", flexibility should be allowed for quote data to be included for consolidation at some point and similar rules should apply.
2. To publish the collected pre-trade data as soon as technically possible.			
3. To collect order event data to full depth for historical analysis and allow it to be available for analysis.			
4. Acquire the data freely in return for a revenue-sharing mechanism of all CT pre-trade data with contributors determined by the ECTP.			
5. Determine the reasonable commercial cost at which all the CT data should be sold.			

Requirements	Can it be Achieved Now?	What is Achievable under Current Legislation	Proposals for Future Legislation
POST-TRADE DATA CONSOLIDATION FOR EQUITIES AND BONDS			
1. To collect post-trade data.		ESMA can use its procurement rules to ensure that the ECTP can: • publish real-time post-trade data submitted by contributors at a cost (a "reasonable commercial basis") until the point, 15 minutes after publication, at which it will be provided at no charge. The EC can define the (parameters of) a "reasonable commercial basis" and the "reasonable cost" at which TVs and APAs shall make their data available to the CTP [NOTE: this does not permit the EC to set the cost at which other contributors of data may make their data available to the CTP]. It could either choose to do this by requiring the ECTP stakeholders to establish and recommend the reasonable cost for data (and the mechanism for establishing this cost) or it could interpret its obligations as the right to directly define the reasonable cost.	The concept of Historical Data needs to be introduced, defined, and mandated as data that must be given to the CTP for long-term storage and it must be clear that this should be freely given up to the CTP. It needs to be clear that post-trade data is not always real-time and can include end of day/session data. It should be clarified that the ECTP can establish: • the revenue allocation mechanism under which post-trade data revenue for CT data would be shared with its members and that aggregators must comply with these rules.
2. To publish the data as soon as technically possible.			
3. To collect and publish end-of-day data/session statistics.			
4. To collect and make available the data for historical analysis.			
5. Acquire the data (both real-time and historical) freely in return for a revenue-sharing mechanism of all CT post-trade data with contributors determined by the ECTP.			
OTHER PRE-REQUISITES			
1. Harmonising Bond Deferrals.		This is for the NCAs and politicians to resolve.	
2. Separate RTS fields for each asset class (bond and derivatives) and enrichment of RTS data to include additional information about trading protocols and systems.		This can be done under the current law by amending or supplementing current Articles.	
3. Reconciliation Capabilities.		Two-sided trade reporting for dealer-to-dealer trades can be introduced.	

Requirements	Can it be Achieved Now?	What is Achievable under Current Legislation	Proposals for Future Legislation
ANCILLARY CHANGES TO MAKE DATA MORE VIABLE			
1. SIs are treated differently from trading venues.		SIs will have to adhere to the Tick Size Regime from 26 June 2020.	All trading venue rules should also apply to SIs. Sufficient consideration should be given to the future need for SIs to publish quotes to the ECTP for consolidation.
2. Increasing the population of bonds eligible for the tape.		ESMA can resolve this and it may be helped by greater transparency through the CT data.	
3. Improving the definition of an EU instrument.		This is not possible without an official instrument list.	An official list of EU instruments needs to be created.

11.1.1. Choices available

The creation of a regulated entity with a post-trade and historical data offering should, therefore, be possible under current legislation. However, it is not possible to create a pre-trade tape without changes to the legislation, as the data that must be given up is not specified in the law.

Therefore, the EC is faced with a choice. It may either:

- Commence the establishment of an ECTP pursuant to the existing delegated authority (which can be used as a proof-of-concept for an ECTP with wider responsibilities in the future) and follow through with amendments to the Level 1 text.
- Wait and seek political agreement to amend Level 1 text (or introduce a new “exchange act”).

The pros and cons of each choice are evaluated below.

Figure 40: Pros and Cons of Choices Available.

Option 1: Use the current legislative powers to create an ECTP to establish a post-trade tape for both equities and bonds and then follow through with amendments to Level 1 text or new legislation.	
Pros	**Cons**
• A framework for an exclusive CTP entity is established that can be recognised in the law at a later stage. • Through mandated membership and stakeholder governance, the industry has to work together to find solutions and this can help inform further changes at Level 1. • Finding a solution is funded by the industry and means that it has "skin in the game" to resolve the problems. • A significant part of the post-trade solution for non-venue generated data could be achieved. • It can be implemented relatively quickly. • There is currently significant impetus from market participants to work together (although not always with the same interests) to establish a consolidated tape. • The end investor may accrue benefits earlier. • Sufficient stakeholders appear ready to engage. • Under the UK's departure from the EU, the need for CT data may be more urgent and the same stakeholders are still likely to be involved.	• ESMA's public procurement process needs to define the ECTP's role clearly and carefully to ensure that it follows the ECs parameters and the ECTP can achieve everything proposed. • The ECTP will not have the power to enforce European laws with respect to data. • The ECTP can only have APAs and TVs as members (two of the stakeholders most potentially disrupted by a CT) who may try to unduly assert themselves in the early governance and formation of the ECTP and CT data. (Other data contributors could become voluntary members) • Any stakeholder with a vested interest not to have CT data could slow the process down or not co-operate. • Equity participants would have to adapt to a hybrid state and may find it hard to consume and value post-trade CT data whilst pre-trade data is still being taken via direct feeds. • It is possible that the proposed Level 1 legislative changes may not be achieved. • Brexit may fragment the interests of different stakeholders

Option 2: Wait for wholesale legislative change.	
Pros	**Cons**
• The role of the ECTP can be more clearly defined in the law. • The full consolidated pre- and post-trade tape can be designed with greater regulatory certainty. • The full pre- and post-trade consolidated tape can be delivered together.	• The practical experience and knowledge gained from Option 1 are foregone. • The industry may not work together to find solutions and continued lobbying for vested interests may not achieve a workable solution. • Regulators will have to continue using sub-optimal data sets and dedicate resources to managing data or funding the development of a consolidated tape. • Any benefits to end investors are deferred for a number of years whilst a small number of financial intermediaries continue to extract profit from information asymmetries.

11.2. CONCLUSIONS

It is legally feasible under the current legislation to create an ECTP with a post-trade tape solution. There are pros and cons to moving forward now or waiting for more legislative change. To achieve a solution in the near term will require the EC to be robust in its interpretation of the law and for ESMA to carefully use its procurement process to work closely with stakeholders to bring the proposed ECTP to fruition.

12 / CONCLUSIONS & RECOMMENDATIONS

There is an urgent need for CT data to improve the transparency and efficiency of European equity and bond markets. Multiple use cases exist across the industry to show that it would bring considerable benefits to investors who could save billions of euros through the existence of such data. The establishment of CT data could also improve the functioning of European capital markets by increasing competition and supporting policy-making decisions.

The only firms who may be potentially disrupted by the development of CT data, are a small number of financial intermediaries which either have sufficient economic and technical power to process and gain advantage from current information asymmetries, or, which can disproportionately profit from the price of the data that they sell.

These findings remain applicable to the EU 27 markets after the UK's exit from the EU. CT data should become a fundamental cornerstone to achieving the Capital Market's Union's objectives, not only to serve the current needs of all investors and participants, but to preserve the competitiveness of the EU 27 financial markets in the event that pools of liquidity develop in third-country markets.

The technology required to build CT data is proven and widely available. The key issues to be overcome in creating such data are legal, structural, organisational and economic. It is impractical to expect multiple competing commercial consolidators of data to emerge when the underlying data that needs to be consolidated resides across many competing commercial entities, some with disproportionate economic leverage and conflicts of interest. Additionally, pre-trade CT data cannot be achieved under current legislation because it is not specified in the law.

The consolidation and aggregation experiences in North America, although not an exact model for Europe, show that, even in less fragmented markets, the optimum structure to commence creating CT data is via well-governed, self-regulated entities with exclusive aggregating and/or consolidating responsibilities. This allows them to mandate rules and standards for the market. Clearly defining pre- and post-trade data in the law combined with the legal recognition of these entities also gives them additional powers to enforce related laws on their members. The US equity experience also demonstrates that revenue allocation models, based on the value of each participant's contribution to the consolidated data, can significantly facilitate consolidation and the use of CT data.

Many similar characteristics are desirable in Europe but with adjustments that allow for the different market characteristics and to avoid some of the legacy issues that exist in both the US and Canada. In particular, an exclusive, non-conflicted consolidator across asset classes that has the role of enforcing standards, rules and related laws for market-wide CT data on all aggregators of data is more important for Europe at this point in its market evolution. A balanced governance structure and stakeholder representation are also critical, particularly to resolve the issue of cost and pricing of CT data. However, as in North America, such a consolidator should be self-regulating and recognised in the law. All of these recommendations are considered pre-requisites and should be viewed holistically if EU CT data is to be successfully delivered. Compromises are likely to result in less optimal CT data.

Economies of scale can also be achieved by using the same high-level technical and organisational framework to deliver both equity and bond data consolidation. The organisation and technology required to deliver a consolidated tape that meets the user requirements and resolves many of the issues identified in this study can be set up for a cost of less than €11 million with annual running costs of €7.6 million. The costs of setting up the CT data, should therefore not be an impediment as the benefits clearly outweigh the costs.

Whilst a consolidated pre-trade data tape cannot be implemented under current legislation, the establishment of an ECTP with a post-trade solution that meets most of the design requirements could be achieved. The EC could use its existing powers to set clear parameters for the creation of an ECTP and request ESMA to use its procurement process to follow those parameters to achieve the desired outcome. This would start to deliver benefits to investors and participants, many of whom believe that the development of CT data is long overdue and who appear ready to engage to help bring it to fruition. ESMA will need to carefully manage the procurement process, as well as to set milestones to measure progress and encourage the resolution of issues between stakeholders.

Further improvements to the post-trade consolidated tape, overall strengthening of the ECTP's role and the creation of a pre-trade consolidated tape for equities would need to be addressed in new legislation.

Recommended Actions:

- The EC should take a robust approach to the interpretation of its existing powers and pursue its power in MiFID II to request ESMA to launch a public procurement process that creates the specified conditions for the establishment of an exclusive CTP for post-trade data as soon as possible. This would be in anticipation that the CTP is a self-regulatory organisation, that it is exclusive and that a "commercial entity" means that it is not conflicted and can be a not-for-profit self-regulatory body with compulsory membership for data aggregators.

- ESMA and the NCAs should be encouraged to review government bond deferral rules and to reconsider and accelerate the phased approach on assessing liquidity in bonds.

- The EC should also seek to introduce further legislative changes to fully support an exclusive CTP for consolidation of both pre- and post-trade data and to enhance the viability of the data.

BIBLIOGRAPHY TO THE MAIN REPORT

ACADEMIC LITERATURE

Angel J (1998)

Consolidation in the Global Equity Market: An Historical Perspective

Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.200.5944&rep=rep1&type=pdf

Asquith, P., Covert, T., Pathak, P.A. (2019 revision of 2013 original)

The Effects of Mandatory Transparency in Financial Market Design: Evidence from the Corporate Bond Market

National Bureau of Economic Research (NBER) Working Paper No. 19417
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2320623

Balp G., Strampelli G. (2018)

Preserving Capital Markets Efficiency in the High-Frequency Trading Era

Bocconi Legal Studies Research Paper No. 3097723; Journal of Law, Technology and Policy, Forthcoming
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3097723

Bartlett R.P., McCrary J. (2017 revision of 2016)

How Rigged Are Stock Markets? Evidence from Microsecond Timestamps

University of California, Berkeley Public Law Research Paper No. 2812123
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2812123

Bessembinder H., Maxwell W., Venkataraman K. (2006)

Market Transparency, Liquidity Externalities, and Institutional Trading Costs in Corporate Bonds

Journal of Financial Economics, Volume 82, Issue 2, November 2006, Pages 251-288
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=644624

Biais B., Declerck F., Dow J., Portes R., von Thadden E-L (2006)

European Corporate Bond Markets: Transparency, Liquidity, Efficiency

Centre for Economic Policy Research, City of London
Available at: http://faculty.london.edu/rportes/TT%20CorporateFULL2.pdf

Boehmer E., Saar G., Yu L. (2004)

Lifting the Veil: An Analysis of Pre-Trade Transparency at the NYSE

Leonard N. Stern School of Business Dept. of Finance Working Paper No. 03-002
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=368102

Boulatov A., Dierker M. (2007)

Pricing Prices

European Finance Association 2007 Ljubljana Meetings Paper.
Available at: https://ssrn.com/abstract=967363

Cespa G., Foucault T. (2013 revision of 2008)

Insiders-Outsiders, Transparency and the Value of the Ticker

Working Papers 628, Queen Mary University of London, School of Economics and Finance
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1117581

Cespa G., Foucault T. (2013)

Sale of Price Information by Exchanges: Does it Promote Price Discovery?

Management Science, published 2014
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1875001

Colby R.L.D., Sirri E.R. (2010)

Consolidation and Competition in the US Equity Markets

Capital Markets Law Journal Vol 5, No 2, 2010
Available at: https://doi.org/10.1093/cmlj/kmq004

Ding S., Hanna J., Hendershott T. (2014)

How Slow is the NBBO? A Comparison with Direct Exchange Feeds

Eastern Finance Association – The Financial Review 49 Issue 2 (2014) 313-332
Available at: http://faculty.haas.berkeley.edu/hender/NBBO.pdf

Easley D., O'Hara M., Yang L. (2016 revision of 2013)

Differential Access to Price Information in Financial Markets

Journal of Financial and Quantitative Analysis (JFQA), 2016, 51(4), 1071–1110; Johnson School Research Paper Series No. 11-2011
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1787029

Edwards, A. Harris L., Piwowar M.S. (2007),

Corporate Bond Market Transaction Costs and Transparency Costs

Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=982654

Hamilton J. L. (1978)

Marketplace Organization and Marketability: NASDAQ, the Stock Exchange, and the National Market System

The Journal of Finance Vol. 33, No. 2 (May 1978), pp. 487-503
Available at: https://www.jstor.org/stable/2326565

Hasbrouck J. (2018)

Price Discovery in High Resolution

Stern School of Business, New York University
Available at: http://people.stern.nyu.edu/jhasbrou/Research/HRVAR/HRVAR08.pdf

Hasbrouck J, Saar G. (2013)

Low-Latency Trading

Johnson School Research Paper Series No. 35-2010, AFA 2012 Chicago Meetings Paper
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1695460

Holden C. W., Jacobsen S. E., Subrahmanyam A. (2017 revision of 2014)

The Empirical Analysis of Liquidity

Foundations and Trends in Finance 8, No 4, 263-365, 2015, Kelley School of Business Research Paper No. 2014-09
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2402215

Lee, R. (2002)

The Future of Securities Exchanges

Brookings-Wharton Papers on Financial Services 2002, 1-33.
Available at:
http://www.oxfordfinancegroup.com/media/10341/lee%20future%20securities%20exchanges%202002.pdf

Madhavan A., Porter D., Weaver D. (2005)

Should Securities Markets Be Transparent?

Journal of Financial Markets 8 (2005) 266-288
Available at:
http://www.coba.unr.edu/faculty/liuc/files/RUC/topic_markettransparency/Madhavan_Porter_Weaver_2005.pdf

Maxwell W. F., Bessembinder H. (2008)

Transparency and the Corporate Bond Market

Journal of Economic Perspectives, 2008
Available at: https://ssrn.com/abstract=1082459

O'Hara, M., Yao C., Ye, M. (2014)

What's Not There: Odd Lots and Market Data

The Journal of Finance, Vol. 69, No. 5 (October 2014), pp. 2199-2236,
Available at: https://www.jstor.org/stable/43612955

O'Hara, M., Ye, M. (2009)

Is Market Fragmentation Harming Market Quality?

Journal of Financial Economics 100.3 (2011): 459-474.
Available at:
https://www.law.columbia.edu/sites/default/files/microsites/capitalmarkets/Market%20Fragmentation%20Paper.pdf

Roseman B., Van Ness B. F., Van Ness R. A. (2018 revision of 2016)

Odd-Lot Trading in U.S. Equities

Quarterly Review of Economics and Finance, Vol. 69, 2018
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2651975

Seligman J. (2002)

Rethinking Securities Markets: The SEC Advisory Committee on Market Information and the Future of the National Market System

The Business Lawyer, Vol. 57, No. 2 (February 2002), pp. 637-680 (44 pages)
Available at: https://www.jstor.org/stable/40688042

Shkilko A. V. *, Van Ness B. F., and Van Ness R. A. (2008)

Locked and Crossed Markets on NASDAQ and the NYSE

Journal of Financial Markets, Volume 11, Issue 3, August 2008, Pages 308-337
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=610584&download=yes

Tivnan B.F., Dewhurst D.R., Van Oort C.M., Ring J.H. IV, Gray T.J., Tivnan B.F., et al. (2018)

Fragmentation and inefficiencies in US equity markets: Evidence from the Dow 30

The Mitre Corporation (US not-for-profit)
Available at: https://arxiv.org/pdf/1902.04690.pdf

Upson J., McInish T., Hardy Johnson IV, B. (2018)

Orders versus Trades on the Consolidated Tape

Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2625401

Wah E. (2016)

How Prevalent and Profitable are Latency Arbitrage Opportunities on U.S. Stock Exchanges?

Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2729109

ACADEMIC LITERATURE - SPONSORED

Angel J. (2018)

Retail investors get a sweet deal: The cost of a SIP of stock market data

Sponsored by NASDAQ

Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3268916

Jones C.M. (2018)

Understanding the Market for U.S. Equity Market Data

Sponsored by NYSE

Available at:

https://www0.gsb.columbia.edu/faculty/cjones/papers/2018.08.31%20US%20Equity%20Market%20Data%20Paper.pdf

Macey J. R., O'Hara M. (1999)

Regulating Exchanges and Alternative Trading Systems: A Law and Economics Perspective

Financial support from Instinet (providers of a U.S. alternative trading system).

Journal of Legal Studies, vol. XXVIII January 1999, University of Chicago
Available at: https://www.jstor.org/stable/10.1086/468045?seq=1

REGULATORY AND GOVERNMENT PAPERS

Alberta Securities Commission (1999)

Discussion Paper: Consolidation Plan for a Consolidated Canadian Market

Available at: https://www.albertasecurities.com/-/media/ASC-Documents-part-1/Regulatory-Instruments/2018/10/Discussion-Paper.ashx

Bank of Canada: Barker W., Pomeranets A. (2011)

The Growth of High-Frequency Trading: Implications for Financial Stability

Bank of Canada, Financial System Review
Available at: https://www.bankofcanada.ca/wp-content/uploads/2011/12/fsr-0611-barker.pdf

Bank of Canada: Garriott C., Pomeranets A., Slive J., Thorn T. (2013)

Fragmentation in Canadian Equity Markets

Bank of Canada Review Autumn 2013, Financial Markets Department, Bank of Canada
Available at: https://www.bankofcanada.ca/wp-content/uploads/2013/11/boc-review-autumn13-garriott.pdf

The Board of the International Organisation of Securities Commissions OICV-IOSCO (2013)

Regulatory Issues Raised by Changes in Market Structure (Consultation Report)

CR03 March 2013
Available at: https://www.iosco.org/library/pubdocs/pdf/IOSCOPD407.pdf

The Board of The International Organization of Securities Commissions (2003)

Objectives and Principles of Securities Regulation

Available at: https://www.iosco.org/library/pubdocs/pdf/IOSCOPD154.pdf

European Capital Markets Institute (ECMI): Karel L. (2007)

ECMI Policy Brief: Financial Market Data and MiFID

Issue No 6, 30 March 2007

Available at: https://www.ceps.eu/ceps-publications/financial-market-data-and-mifid/

European Securities and Markets Authority (ESMA) (2019)

MiFID II/MiFIR Review Report No. 1 , On the development in prices for pre- and post-trade data and on the consolidated tape for equity instruments

ESMA70-156-1606, 5 December 2019

Available at: https://www.esma.europa.eu/sites/default/files/library/mifid_ii_mifir_review_report_no_1_on_prices_for_market_data_and_the_equity_ct.pdf

Federal Reserve Board: Caglio C., Mayhew S. (2012)

Equity Trading and the Allocation of Market Data Revenue

Available at: https://www.federalreserve.gov/pubs/feds/2012/201265/201265pap.pdf

Financial Conduct Authority (UK): Aquilina M., Foley S., O'Neill P., Ruf T. (2016)

Asymmetries in Dark Pool Reference Prices

Occasional Paper 21
Available at: https://www.fca.org.uk/publication/occasional-papers/occasional-paper-16.pdf

Financial Conduct Authority (UK): Aquilina M. and Ysusi C. (2016)

Are High-Frequency Traders Anticipating the Order Flow? Cross-Venue Evidence from the UK Market

Occasional Paper 16, Chief Economist's Department of the Financial Conduct Authority
Available at: https://www.fceligmana.org.uk/publication/occasional-papers/occasional-paper-16.pdf

Ontario Securities Commission (2012)

CSA Staff Consultation Paper 21-401: Real-Time Market Data Fees

Available at: https://www.osc.gov.on.ca/en/SecuritiesLaw_csa_20121108_21-401_real-time-data-fees.htm

U.S. Department of Treasury (2017)

A Financial System That Creates Economic Opportunities – Capital Markets

Available at: https://www.treasury.gov/press-center/press-releases/Documents/A-Financial-System-Capital-Markets-FINAL-FINAL.pdf

World Bank: Carson, J. (2011)

Self-Regulation in Securities Markets

Policy Research Working Paper 5542, Financial and Private Sector Development, Global Capital Markets Department, Securities Market Group
Available at: https://elibrary.worldbank.org/doi/pdf/10.1596/1813-9450-5542

INDUSTRY PARTICIPANT PAPERS

Aequitas Innovations Inc. (2015)

Breaking the Virtual Canadian Market Data Monopoly

Available at: https://aequitasneo.com/documents/en/publications/market-data-position-paper.pdf

Aite Group (2013)

Market Fragmentation and Its Impact: A Historical Analysis of Market Structure Evolution in the United States, Europe, Australia, and Canada

Sponsored by BM&F Bovespa
Available at:
http://www.cvm.gov.br/export/sites/cvm/audiencias_publicas/ap_sdm/anexos/2013/sdm0513manifestacaoBMFBovespaAnexoII_09-08-2013.pdf

Bank of Canada: Anderson L., Andrews E., Devani B, Mueller M., Walton A. (2018)

Speed Segmentation on Exchanges: Competition for Slow Flow

Investment Industry Regulatory Organization of Canada and Financial Markets Department, Bank of Canada

Bank of Canada Staff Working Paper 2018-3

Available at: https://www.bankofcanada.ca/wp-content/uploads/2018/01/swp2018-3.pdf

Copenhagen Economics (2018)

Pricing of Market Data

Commissioned by the Danish and Swedish Security Dealers Associations

Available at: https://www.copenhageneconomics.com/publications/publication/pricing-of-market-data

FIX Global (2019)

Say Goodbye to the Order Protection Rule?

1st May 2019

Available at: https://www.fixglobal.com/home/say-goodbye-to-the-order-protection-rule/

ICMA: Callaghan E. (2020)

EU Consolidated Tape for Bond Markets. Final report for the European Commission

ICMA MiFID II Consolidated Tape Taskforce, International Capital Market Association. April 2020

Available at:

https://www.icmagroup.org/assets/documents/Regulatory/MiFID-Review/EU-Consolidated-Tape-for-Bond-Markets-Final-report-for-the-European-Commission-290420v2.pdf

ICMA: Callaghan E. (2019)

ICMA Quarterly Report Assessment of Market Practice and Regulatory Policy Secondary Markets; MiFID II/R: post-trade data

International Capital Market Association, 11 July 2019 Third Quarter. Issue 54. Pg 40 - 42

Available at:

https://www.icmagroup.org/assets/documents/Regulatory/Quarterly_Reports/ICMA-Quarterly-Report-Third-Quarter-2019-Final.pdf

Oxera (2019)

The Design of Equity Trading Markets in Europe – An Economic Analysis of Price Formation and Market Data Services

Sponsored by FESE

Available at:
https://fese.eu/app/uploads/2019/03/190321-The-design-of-equity-trading-markets-in-Europe-full-report.pdf

Oxera (2014)

Pricing of Market Data Services, An Economic Analysis

Available at: https://www.oxera.com/publications/pricing-of-market-data-services-an-economic-analysis/

Securities Litigation & Consulting Group Inc. (SLCG) (2011)

An Economic Study of Securities Market Data Pricing by Canadian Trading Venues

Commissioned by Investment Industry Association of Canada (IIAC)
Available at:
https://iiac.ca/wp-content/themes/IIAC/resources/1580/original/SLCG-Canadian-Market-DataJune%207.pdf

Securities Litigation & Consulting Group Inc. (SLCG) (2008)

An Economic Study of Securities Market Data Pricing by the Exchanges

Sponsored by the Securities Industry and Financial Markets Association (SIFMA)
Available at:
https://www.sifma.org/wp-content/uploads/2017/05/an-economic-study-of-securities-market-data-pricing-by-the-exchanges.pdf

REFERENCES

1 https://ec.europa.eu/info/business-economy-euro/growth-and-investment/capital-markets-union_en

2 Directive 2014/65/EU

3 Regulation (EU) No 600/2014

4 MiFID II Article (4)(1)(53) means a person authorised under the directive to provide the service of collecting trade reports for financial instruments listed in Articles 6, 7, 10.

5 MiFIR Recitals 117 and 118.

6 MiFIR II Recital 117

7 Regulation (EU) No 596/2014

8 Regulation (EU) 2019/876 and Regulation (EU) No 648/2012

9 Regulation (EU) No 648/2012. Commission Delegated Regulation (EU) 216/2251 is a Regulatory Technical Standard (RTS) that establishes prescriptive requirements regarding the Risk Mitigation techniques referred to in Article 11 of EMIR.

10 See Appendix 1: European Market Capitalisation; A.1. The Share of Value (Market Capitalization) of Equities in EU Countries".

11 MiFID II Article 4 (1). QIF in this document refers specifically to an EU MIFID II qualifying investment firm.

12 MiFID II Article 4 (1) (20)

13 Directive 2014/65/EU, Article 4(1)(63)

14 MiFID II/MiFIR Review Report No. 1 On the development in prices for pre- and post-trade data and on the consolidated tape for equity instruments, ESMA, Dec 2019.

15 A crossed market occurs when a market's bid price exceeds its ask price.

16 https://www.ctaplan.com/publicdocs/CTAPLAN_Population_Metrics_4Q2019.pdf

17 https://www.fca.org.uk/news/press-releases/state-street-uk-fined-%C2%A3229m-financial-conduct-authority-transitions-management

18 United States District Court, Southern District of New York. Case 1:20-cv-03154 Document 1 Filed 04/21/20 Page 1 of 81 ISABEL LITOVICH, on Behalf of Herself and All Others Similarly Situated,

19 https://citywire.co.uk/investment-trust-insider/news/revealed-woodford-investors-forecast-to-lose-33-in-wind-up/a1293136

20 Refinitiv, ETF.com, Cetfa.

21 RTS 13, MiFID II Articles 61(4), 64(6),(8), 65(6), (8), and 66(5).

22 https://www.esma.europa.eu/data-systematic-internaliser-calculations

23 According to Article 4(1)(20) of Directive 2014/65/EU (MiFID II) investment firms dealing on own account when executing client orders over the counter (OTC) on an organised, frequent, systematic and substantial basis are subject to the mandatory systematic internaliser (SI) regime.

24 ECMI Policy Brief (2007), "*Financial Market Data and MiFID*".

25 ICMA (2020), "*EU Consolidated Tape for Bond Markets. Final report for the European Commission*"

26 ESMA (2019), "*MiFID II/MiFIR Review Report No. 1 – On the development in prices for pre- and post-trade data and on the consolidated tape for equity instruments*"

27 IOSCO Board (2003), "*Objectives and Principles of Securities Regulation*", Section 7.2, p12.

28 World Bank: Carson, J. (2011), "*Self-Regulation in Securities Markets*"

29 Ibid.

30 Ibid.

31 Boulatov A., Dierker M. (2007), "*Pricing Prices*".

32 Colby R.L.D., Sirri E.R. (2010), "*Consolidation and competition in the US equity markets*"

33 *Ibid.*

34 *Ibid.*

35 Lee, R. (2002), "*The Future of Securities Exchanges*"

36 Colby R.L.D., Sirri E.R. (2010), "*Consolidation and competition in the US equity markets*"

37 ESMA (2019), "*MiFID II/MiFIR Review Report No. 1. On the development in prices for pre- and post-trade data and on the consolidated tape for equity instruments*"

38 Copenhagen Economics (2018), "*Pricing of Market Data*"

39 Angel J. (1998), *"Retail investors get a sweet deal: The cost of a SIP of stock market data"*; Boulatov A., Dierker M. (2007), *"Pricing Prices"*

40 Jones C.M. (2018), *"Understanding the Market for U.S. Equity Market Data"*

41 Oxera for FESE (2019), *"The Design Of Equity Trading Markets In Europe – An Economic Analysis Of Price Formation And Market Data Services"*

42 Bank of Canada: Garriott C., Pomeranets A., Slive J., Thorn T (2013), *"Fragmentation in Canadian Equity Markets"*

43 Aequitas Innovations Inc. (2015), *"Breaking the Virtual Canadian Market Data Monopoly"*

44 Ontario Securities Commission (2012), *"CSA Staff Consultation Paper 21-401: Real-Time Market Data Fees"*

45 Easley D., O'Hara M., Yang L. (2016 revision of 2013), *"Differential Access to Price Information in Financial Markets"*

46 Tivnan B.F., Dewhurst D.R., Van Oort C.M., Ring J.H. IV, Gray T.J., Tivnan B.F., et al. (2018), *"Fragmentation and inefficiencies in US equity markets: Evidence from the Dow 30"*

47 Lee, R. (2002), *"The Future of Securities Exchanges"*

48 FCA (UK): Aquilina M., Foley S., O'Neill P., Ruf T. (2016), *"Asymmetries in Dark Pool Reference Prices"*

49 Tivnan B.F., Dewhurst D.R., Van Oort C.M., Ring J.H. IV, Gray T.J., Tivnan B.F., et al. (2018), *"Fragmentation and inefficiencies in US equity markets: Evidence from the Dow 30"*

50 Holden C. W., Jacobsen S. E., Subrahmanyam A. (2017 revision of 2014), *"The Empirical Analysis of Liquidity"*

51 Hasbrouck J. (2018), *"Price Discovery in High Resolution"*

52 FCA (UK): Aquilina M., Foley S., O'Neill P., Ruf T. (2016), *"Asymmetries in Dark Pool Reference Prices"*

53 File No. 4-729: SIFMA Comment Letter on Improvements to the Market Data Structure, Sept 18, 2019

54 Notice of Proposed Order Directing the Exchanges and FINRA to Submit a New National Market System Plan Regarding Consolidated Equity Market Data (File No. 4-757), Feb 5 2020

55 A crossed market occurs when a market's bid price exceeds its ask price.

56 Federal Reserve Board: Caglio C., Mayhew S. (2012), *"Equity Trading and the Allocation of Market Data Revenue"*

57 FIX Global (2019), *"Say Goodbye to the Order Protection Rule?"*

58 https://www.nasdaq.com/articles/what-does-totalmarkets-do-market-structure-debate-2019-05-02

59 IOSCO Board (2013), *"Regulatory Issues Raised by Changes in Market Structure Consultation Report"*

60 SEC Release No. 34-87295 October 2019 (https://www.sec.gov/rules/sro/cboeedgx/2019/34-87295.pdf)

61 https://www.sec.gov/marketstructure/datavis/ma_stocks_oddlotvolume.html#.XqlapmRKhsM

62 Upson J., McInish T., Hardy Johnson IV, B. (2018), *"Orders verses Trades on the Consolidated Tape"*

63 Case 1:20-cv-03154. Document 1, Filed 04/21/20, Page 1 of 81, United States District Court, Southern District of New York

64 79 Fed. Reg. 5536, Pub. L. 111-203, July 21, 2010, section 619, codified at 12 U.S.C. § 1851

65 Aite Group for BM&F Bovespa (2013), *"Market Fragmentation and Its Impact: a Historical Analysis of Market Structure Evolution in the United States, Europe, Australia, and Canada"*

66 Bessembinder H., Maxwell W., Venkataraman K. (2006), *"Market Transparency, Liquidity Externalities, and Institutional Trading Costs in Corporate Bonds"*; Edwards, A. Harris L., Piwowar M.S. (2007), *"Corporate Bond Market Transaction Costs and Transparency Costs"*

67 IOSCO (2017), *"Examination of Liquidity in the Corporate Bond Market"*

68 Maxwell W.F., Bessembinder H. (2008), *"Transparency and the Corporate Bond Market"*

69 MiFIR Article 23(1)

70 https://www.fca.org.uk/news/statements/fca-update-share-trading-obligations

71 REGULATION (EU) 2019/2175 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 18 December 2019 amending Regulation (EU) No 1093/2010 establishing a European Supervisory Authority (European Banking Authority), Regulation (EU) No 1094/2010 establishing a European Supervisory Authority (European Insurance and Occupational Pensions Authority), Regulation (EU) No 1095/2010 establishing a European Supervisory Authority (European Securities and Markets Authority), Regulation (EU) No 600/2014 on markets in financial instruments, Regulation (EU) 2016/1011 on indices used as benchmarks in financial instruments and financial contracts or to measure the performance of investment funds, and Regulation (EU) 2015/847 on information accompanying transfers of funds.

LEGAL NOTICE

Luxembourg: Publications Office of the European Union, 2020

PDF ISBN 978-92-76-12113-8 doi: 10.2874/434465 EV-03-19-794-EN-N

GETTING IN TOUCH WITH THE EU

In person

All over the European Union, there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact/meet-us_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by Freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 2 299 96 96, or
- by email via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications from: https://publications.europa.eu/en/publications.

Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact/meet-us_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: http://eur-lex.europa.eu

Open data from the EU

The EU Open Data Portal (http://data.europa.eu/euodp/en) provides access to datasets from the EU. Data can be downloaded and reused for free, for both commercial and non-commercial purposes.

Publications Office
of the European Union