Challenges in market-abuse monitoring: Post MiFID

Received (in revised form): 9th April, 2020

Peter Nylén
Head of Outsourced Trading Surveillance, Trapets AB, Sweden

Peter Nylén *has been working in the field of market surveillance since 2000, when he joined the Stockholm Stock Exchange's (NASDAQ Stockholm) surveillance team. Subsequently, he spent five years at Finansinspektionen (the Swedish Financial Supervisory Authority) as a senior market abuse investigator before moving on to the Scandinavian multilateral trading facility (MTF) Burgundy as Head of trading surveillance until that organisation was acquired by Oslo Börs in 2013. Peter is currently Head of Outsourced Trading Surveillance (OTS) at the Swedish RegTech company Trapets AB. The OTS team is an independent department within Trapets that delivers real-time trading surveillance as an outsourced service to a growing number of European investment firms covering investigation and reporting cases of market abuse. Along with heading the OTS team, Peter is also the Product Manager for the trade surveillance system InstantWatch Market and is deeply involved in implementing this system with clients around the globe. This gives him a unique insight into the current market abuse-monitoring challenges for regulators, venues and firms combined with a solid experience of operational monitoring.*

ABSTRACT

Since the implementation of the Markets in Financial Instruments Directive (MiFID) in 2007, the fragmentation and complexity of securities trading have increased and continue to do so even more particularly after the advent of MiFID2/MiFIR in 2018. European regulators, venues and investment firms are all required to contribute to the overall market-abuse monitoring but under widely different conditions and with access to totally different datasets with shifting quality and coverage. The emergence of the market abuse legislation since the first Market Abuse Directive from 2003 has meant that the obligations and expectations have steadily increased, which raises the questions of whether there is a big gap between what is expected and what is possible given the current situation, and what measures can be taken to close this gap. In this paper, the author evaluates the monitoring process with emphasis on the detection phase and the importance of correct trading data to enable the first steps in automating the detection of suspicious market abuse. The different starting points for regulators, venues and firms and their respective access and lack of necessary information result in an apprehension that no single entity is capable of collecting and conducting monitoring of correct, detailed data with full coverage.

Keywords: market abuse, market-abuse monitoring; MiFID2; trade surveillance

INTRODUCTION

The implementation of the Markets in Financial Instruments Directive (MiFID), MiFID II and MiFIR has gradually increased the fragmentation and complexity of today's securities trading landscape, rendering the market hard to monitor.

Parallel to that, the implementation of the Market Abuse Directive (MAD) and the Market Abuse Regulation (MAR) has increased the requirements and the regulator's expectations of those responsible for detecting and reporting suspicious market abuse.

Peter Nylén

Kungsgatan 56,
111 22 Stockholm,
Sweden
Tel: +46 703 06 03 89;
E-mail: peter.nylen@trapets.com

Journal of Securities Operations & Custody
Vol. 12, No. 4, pp. 367–376

1753-1802

A fundamental prerequisite to the ability to detect suspicious market abuse is the access to relevant trading data. This area has clearly moved in the wrong direction over the last few years, which begs the question if it is even possible today to collect the data required to fulfil one's obligations?

The parties responsible for the monitoring of market abuse are National Competent Authorities (NCAs), venues and firms, all of which have access to different types of datasets, but put together, hopefully, will constitute an overall market-abuse monitoring device capable of detecting market manipulation and insider dealing in any form.

Considering the data and conditions for each of these monitoring parties, it is clear that there are gaps in the current solution that need to be identified, addressed and filled.

REGULATORY BACKGROUND

The two regulatory frameworks relevant to this paper, which unfortunately influence in separate directions, are MiFID/MiFID2/MiFIR and MAD/MAR.

MiFID/MiFID2/MiFIR

The first MiFID[1] was implemented in 2007, and its biggest contribution to the area of market-abuse monitoring was the elimination of the stock exchanges monopoly and the fragmentation of securities trading. Prior to MiFID, all trading activity in a certain equity was concentrated to a single exchange, but with the introduction of regulated markets (RMs) and multilateral trading facilities (MTFs), the same equity could be traded on multiple venues.

With this shift in the trading landscape, a more automated and algorithmic trading evolved, which was a direct consequence of, but also a prerequisite for, efficient fragmented securities trading.

With fragmented trading, the possibilities to retain an overview of the trading deteriorated, and the newly introduced competitive situation between the different venues, offering trading in the same instruments and fighting for market shares, hardly encouraged them to cooperate in finding solutions for sharing information and establishing a foundation for cross-venue monitoring.

The NCAs were equipped with the first version of the transaction reporting system (TRS), which should have enabled them to stay in control and still see the whole picture. The TRS then, and still to this date, only covers transactions with no information at all on any order activity, which unfortunately makes it rather inadequate for detecting market abuse. The order activity, rather than the trades executed, is many times the key to assess whether the behaviour constitutes market manipulation or is considered legitimate.

The direction towards a changed trading landscape set out by the first MiFID has continued with its successors MiFID II[2] and MIFIR,[3] which introduce even more variations of venues by adding organised trading facilities (OTFs), systematic internalisers (SIs), liquidity providers (LPs) and market makers (MMs), all of which are referred to as execution venues. RMs, MTFs and OTFs are classified trading venues. (See Figure 1.)

Within the (today) 28 member countries of the European Union, there are approximately 1,500 different venues with their own unique market identifier codes (MICs).[4]

At the time of the publication of this paper, the figures might be 27 countries and 1,100 MICs.

The earlier mentioned fact is manifested in the way that the price of an instrument traded on multiple venues will be kept more or less identical on all venues by a close web of order messages from a large number of high-frequency algorithmic trading firms. Not only is this applicable on a single instrument, but the same applies for all

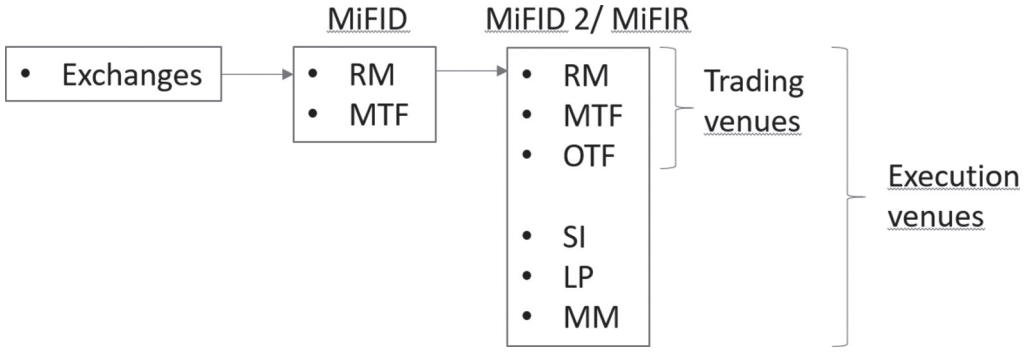

Figure 1 Evolution of various venues
Note: LP, liquidity provider; MiFID, Markets in Financial Instruments Directive; MM, market maker; MTF, multilateral trading facility; OTF, organised trading facility; SI, systematic internaliser; RM, regulated market.

related derivate products with a price that is strictly linked to its underlying instrument(s). Like ripples on water, a price change in an instrument on one venue gives rise to a corresponding, foreseeable price change on multiple other venues and in multiple other instruments. This invisible force glues the price-formation process in all related instruments together and leads to efficient markets but, on the flip side, this opens the door for manipulative trading strategies where a rogue trader enriches itself on one venue by influencing the price on another. Such behaviour will remain undetected if no monitoring organ has the whole picture with access to full covering and highly detailed order and trade data, which is processed by an automated surveillance system capable of processing and analysing all the millions of messages. Such trading strategies require no large price shifts in order to be profitable; a tick or two will be enough to result in a large profit (or loss[5]) if it is repeated over time.

As a counter measure against the increased fragmentation and decreased oversight, it was the legislators' ambition that MiFID2/MiFIR would lower the cost of public market data (which only venues would argue has happened . . .) and to introduce consolidated tape providers (CTPs) that would provide a consolidated post-trade ticker, free of charge and available for all participants.

Not surprisingly, no one has yet jumped onto the idea to become a CTP, and there is still debate between the users and suppliers of public market data whether this is available on reasonable commercial basis, which is why the European Securities and Markets Authority (ESMA) launched a consultation paper this summer with the goal of sorting out these matters.[6] The issue with increased complexity and its potential impact on transparency was raised even before the implementation of MiFID2[7] and is obviously relevant to this day.

The main tool for the NCAs, the TRS, has received a face lift, and the number of fields for each transaction report has increased significantly with several fields identifying the buyer or seller.[8] There is also a new format for order record keeping, which trading venues must relate to, and upon request send to their NCA.[9]

MAD/MAR

The first MAD was implemented back in 2005, and it laid down the foundation for the current market-abuse framework with common definitions of insider dealing,

market manipulation and the newly introduced obligation to report suspicious transactions.

There were a few obvious gaps in this legislation which, since then, have been identified and addressed by the regulator:

- No expressed obligation to monitor

The Directive included an obligation to report suspicious transactions, but no formal requirements regarding the monitoring setup and what actions that need to be taken to actually detect abusive trading, which can be reported.

- Reporting obligation for transactions, not orders

The reporting obligation included only transactions which constituted suspicious insider dealing or market manipulation, not orders.

MAD was implemented differently in different member countries, but in general, the obligation to report suspicious market abuse covered fewer market participants than what was implied in the current legislative framework.

The shift to a more fragmented, automated and high-frequency trading, which stemmed from MiFID, gave rise to the need for an updated guidance from the regulator on how firms and venues should organise and approach their efforts regarding detecting, investigating and reporting suspicious market abuse. In 2012, ESMA published their guidelines on automated trading,[10] which was the first step against filling the gaps in MAD. For example, the guidelines stated that in automated trading, an automated surveillance system is a necessity to be able to detect suspicious activity and also that suspicious orders should be treated and reported just as transactions.

In July 2016, MAR was implemented across the European Union with binding requirements for venues and firms to establish and maintain effective arrangements, systems and procedures in order to be able to detect suspicious orders and transactions and then submit suspicious transaction and order reports (STORs) regarding insider dealing or market manipulation or attempted such.

The selected surveillance solution should certainly be appropriate and proportionate, but looking in the rear view mirror on how the market- abuse legislation has evolved since MAD, it is obvious that the demands and expectations from the regulator of the market participants, widely defined as 'any person professionally arranging or executing transactions', regarding their contribution in the detection and reporting of market abuse have increased drastically.

Over the years, the regulator has added different examples to the growing list of manipulative behaviours for the industry to relate and adapt to. From 'Wash trades/ cross-trades'[11] in the third implementing directive of MAD in 2003, to 'layering, quote stuffing and momentum ignition'[12] in ESMA's guidelines on automated trading from 2012 to the examples listed in the delegated regulation to MAR 2015, 'painting the tape, pump and dump, trash and cash',[13] are a few to mention.

Thus, the implication is that the legislator has, on one hand, created a more complex, fragmented and unmonitorable securities trading landscape and, on the other hand, successively raised the obligations, requirements and scope of those set to detect and report suspicious market abuse, ie two legislative movements going in the opposite directions.

MONITORING PROCESS: DETECT, INVESTIGATE, ESTABLISH SUSPICION

The current situation is that the main task of detecting, investigating and reporting suspicious market abuse lies within the venues and firms. To some extent, the NCAs have

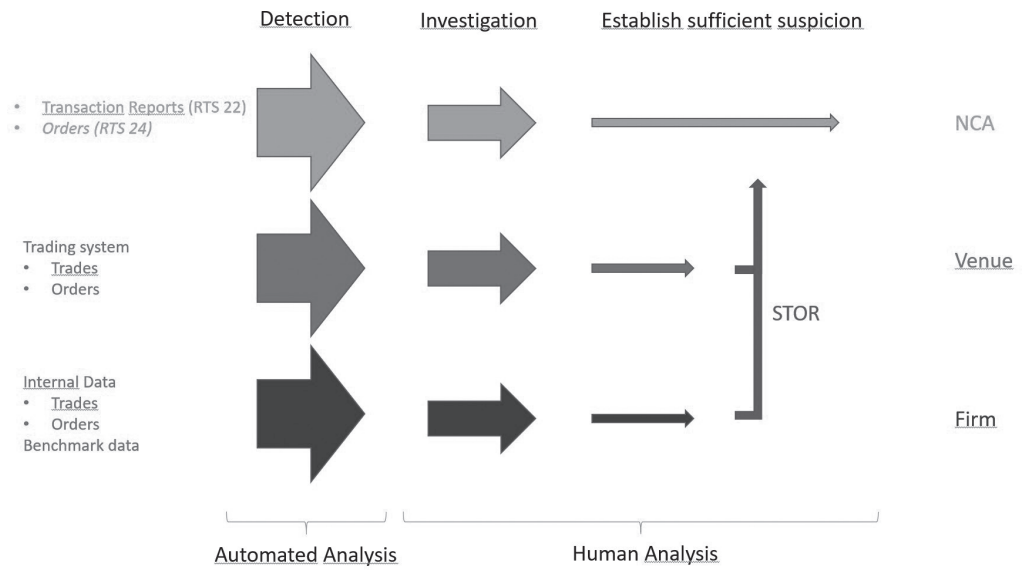

Figure 2 In-data prerequisites and monitoring process for the three different parties
Note: NCA, National Competent Authority; STORs, suspicious transaction and order reports.

their own detection capabilities, but their main task is to handle the submitted STORs and investigate these cases further until they are either closed or lead to a sanction.

(Note: Any part of this process handled by law-enforcement agencies and leading to criminal sanctions is left out of this paper as they do not contribute to the part of detecting and investigating in the early stages.)

As the headline implies, the monitoring process can be divided into three phases: detection, investigation and establishing a 'sufficient' suspicion.

The second phase, investigation, looks rather similar within firms, venues and NCAs as it mainly consists of human analysis based on available information. The three different parties, however, have different information available to them, different authority in obtaining additional information, and different finish lines before reaching the third phase, establishing sufficient suspicion. For venues and firms, their investigation is complete when they either submit a STOR based on reasonable suspicion or close the case as a near miss. An NCA investigates until the case is either closed or handed over for sanction (criminal or administrative).

The main difference in the monitoring processes between firms, venues and NCAs lies in the first phase — the detection phase. The reason for this is that the three parties have totally different data available to base the detection on. (See Figure 2.)

It is a prerequisite in today's trading that the detection phase is fully automated; the time is long gone when manual controls and spread sheets were efficient tools to detect abusive behaviour. The reason for this is the enormous amounts of data that needs to be processed and analysed in order to find that needle in the haystack, which is the starting point of a market-abuse investigation, which eventually might result in a STOR or a sanction case. It might still be an appropriate and proportionate approach for a few firms, but the proportion of the trading that is passing through such firms is negligible.

The common setup is that different types of trading-related data, 'messages', are fed into a surveillance system, which

then analyses the in-data and triggers alerts based on defined deviations and suspicious behaviours.

Different surveillance system vendors have their own twist to their offerings and provide different alert logics and algorithms. There is still, however, somewhat of a prevailing standard and consensus regarding the functionality available in modern surveillance systems. The indicators of manipulative behaviours listed in Annex II to the Delegated Regulation 2016/522, mentioned earlier in this text, must be considered to be some sort of base level for alert scenarios, all of which need to be tweaked and adapted to prevailing trading conditions across different market models and asset classes.

The quality of the alert output is directly dependent of the quality of the in-data, both regarding each message separately but also regarding the synchronisation/correlation between the different messages.

For the sake of simplicity, one can assume that if a surveillance system is fed with 1,000,000 messages, it will create 100 alerts, which will result in one case of suspicious market abuse. Such analysis is obviously extremely sensitive to errors in the in-data. Should only 100 (ie 1 out of every 10,000) messages carry an attribute, which is corrupt or mis-synchronised in any way, this would possibly result in 100 false positives; doubling the number of triggered alerts.

Therefore, it is absolutely key that the quality of the in-data messages is close to flawless for the surveillance officers responsible for the second phase, to investigate the triggered alerts and to focus on anything else but to shovel away false positives.

Another important aspect related to the quality of the in-data is the level of detail in each message as each additional information attached to the message, for example an order or a trade, enables the surveillance system to carry out a more advanced and complex alert logic. This enhances the possibilities to detect more intricate and better camouflaged abusive trading strategies along with enabling a risk-based approach by focusing the detection capabilities to areas where the risk of certain market abuse is higher.

Over time, the alert logics and detection capabilities will be designed based on the existing in-data and, with enhanced quality and level of detail, the functionalities within the surveillance system will improve over time. Unfortunately, it also goes the other way around; when the quality and level of detail deteriorate, the systems will be designed according to bad in-data, resulting in trivial detection capabilities.

Besides having to be of high quality and detailed, the in-data needs to cover all trading activity undertaken by/within/under the jurisdiction of the firm, venue or NCA. The broader the scope of trading activity, the harder to obtain highly detailed messages of good quality.

To fully comply with the monitoring requirements and maintain optimal detection capabilities, the surveillance system must be fed with full-covering, detailed, high-quality in-data, something that is extremely hard, if not impossible given the current legislative framework, to achieve for any firm, venue or NCA.

NCA

The NCAs are the only party with authority and ability to collect full-covering data necessary to detect cross-venue manipulation.

The transaction reports, however, are quite useless from a detection point-of-view for several reasons. First off, the transaction reports submitted are not to be mistaken for trades executed on a venue; they are merely information that a specific instrument has been passed on by an entity from a seller to a buyer. When these transactions form a chain, from an end seller to an end buyer, there is no identifier available to enable these separate links to construct such a

transaction chain. An automated surveillance system will naturally focus on the end clients when detecting abusive behaviour, but with that information missing, the alert logics will fall short.

Secondly, the quality of the submitted transaction reports is in many cases inadequate along with the fact that the reports can be delayed and/or updated at a later stage. As mentioned earlier, surveillance systems are extremely sensitive to errors in the in-data, having disastrous consequences to the alert output if the quality or timing of the in-data is lacking. The process of enhancing the quality of the submitted transaction reports is currently ongoing and will most likely continue to do so for the foreseeable future. Also to be kept in mind is that there are thousands of different parties involved in the whole infrastructure surrounding the TRS, which also raises the risk of errors occurring.

The third obvious deficiency is that order-book data is not submitted continuously to the NCAs in the same way as transaction reports, but only upon request from NCAs. So, even if transaction reports were equal to trades, without corresponding order-book data, the surveillance system lacks the most relevant data for conducting automated analysis and detecting abusive behaviour ie orders.

From an investigative standpoint, the TRS has its advantages as it enables the NCA to identify the end client rather swiftly, provided a potential transaction chain contains all the links necessary to follow it to the end and that there is a limited number of transaction reports for it to be done manually.

Venue

The coverage of data required for the detection phase is considered both extremely good and extremely poor.

It is good due to the fact that the data available in the venues' surveillance systems stems from its own trading system and its own business so, in that sense, the data coverage is 100 per cent. Moreover, as the level of detail in the order and trade data is high and determined by the venue itself along with a perfect synchronisation between orders and trades, this results in very good prerequisites for efficient and accurate detection of abusive trading.

Often, the venues themselves develop both the trading system and the surveillance system used, and when these two systems go hand in hand, it adds additional value to the detection capabilities as the data source (trading system) is in total harmony with the surveillance system.

Another aspect to the quality of the in-data is that it has already gone through rigorous testing before showing up in the surveillance system in the sense that it first must pass through the trading system. As trading system providers regard the security and reliability of their products to what is applicable to the airplane and nuclear industry, this of course ensures the absence of rogue data in the venues' surveillance systems (when comparing this to the submitting model surrounding the transaction reports, the difference in the in-data quality between venues and NCAs becomes quite clear).

But, on the other hand, the venues have no corresponding information on what is happening in the order books of the other venues, which leaves all the venues rather blind when trying to detect cross-venue market manipulation. Any investigation undertaken will be limited to activity only on the current venue with no possibility to obtain additional information from their competitors reminiscent of trying to solve a puzzle with only half of the pieces.

So, looking at the venues' data coverage in the bigger picture, it could be considered extremely poor.

The advantages for venues listed earlier, with surveillance systems fed with in-house

data with perfect quality and order/trade synchronisation and a level of detail determined by themselves, are likely diminishing for the newly added execution venues, which might procure third-party trading and surveillance systems, rendering them a surveillance setup of much lower quality.

FIRM

The firms are in a predicament as they need to synchronise data of rather low quality from multiple data sources to be able to conduct an efficient monitoring of their trading activities.

The first step is to collect all private, in-house data, which is a challenge in itself as this is spread out across several internal trading systems that all have their separate models and formats for extracting data. These internal trading systems are connected to various venues with different trading models and asset classes, which adds extra complexity to the task of compiling the data into one single format. Moreover, add the fact that they are located in different jurisdictions/locations, which results in firms often using several surveillance solutions, each covering a specific part of the firm's trading landscape.

The second step, which can be even more challenging for the firm, is to obtain the public benchmark data, which is a necessity for both the automated and human analyses to be able to assess the trading and conclude whether it is abusive or not. Without public benchmark data, it is practically impossible to analyse one's own trading behaviour as one has nothing to compare it against.

Market manipulation is about manipulating the market, and if the surveillance system or the investigator has no information about the 'market', they cannot function optimally. A case of suspicious order-book layering will naturally be assessed differently if the layering orders are large and placed close to the best bid/offer (BBO) spread compared to if they are of insignificant size and placed way down in the order book. Even the most basic considerations, such as if a price shift is deemed big or if a traded volume is deemed large, are undoable without public benchmark data.

Obtaining public market data that fully corresponds with the whole scope of the firm's own trading is an expensive nightmare due to the technical complexity, and due to the sheer size of the data volumes, it is a practical necessity to throttle and filter the data, for example, in regard to the time-stamp granularity and order-book levels.

As mentioned earlier, MiFID II was supposed to introduce CTPs, but no one has yet to jump on that. Even with established CTPs, one can strongly question what value such an arrangement would add as an in-data source to a firm's automated surveillance system, as the data provided would only constitute post-trade information on the trading activity and no information regarding order activity (BBO, order-book depths etc). So, even if this data was free of charge, most firms would consider the data from the CTP useless as the order-book data is lacking, leaving them with no choice but to continue to procure the public benchmark data from a third-party market data vendor.

GAPS IN THE OVERALL MONITORING

Considering the prerequisites for each of the three parties listed earlier, it is obvious that no single entity is capable of collecting and conducting monitoring on correct, detailed data with full coverage. (See Table 1.)

This leads to clear gaps in the overall pan-European monitoring model, which can be exemplified by cross-product/cross-venue manipulation,[14] an abusive trading strategy where a participant repeatedly influences the price formation in one instrument on one venue (preferably only with orders as that is free of charge) and

Table 1: In-data for the three different parties

	Quality	*Coverage*	*Level of detail*
NCA	Poor	Good	Poor
Venue	Good	Poor	Good
Firm	Medium	Good	Medium

Note: NCA, National Competent Authority.

letting the close web of order messages by other participants create a foreseeable, corresponding and sometimes leveraged price change in other instrument(s) on other venue(s) and then execute profitable trades there.

This manipulation will not be picked up by either venue, provided the trading is somewhat camouflaged and the trader does not get greedy, as each venue will only see their part of the whole scheme. It will not get picked up by any NCA as they lack order-book data, and it will probably not get picked up by the firm as this requires data too detailed and granular for the surveillance system to detect it. To ensure nondetection by the firm, the trader could use multiple firms and then completely fly under the radar.

This does not just apply for cross-product/cross-venue manipulation, but for all types of manipulative trading, where multiple venues (and firms) are utilised, which results in a highly reduced, if not nonexisting, risk of being detected.

By spreading out the trading and leaving smaller footprints within each monitoring entity, the risk of being detected when dealing on insider information can be highly reduced as well. Using several firms, venues and asset classes when dealing with insider information renders it a nearly undetectable crime.

The earlier is exacerbated as each monitoring party would not set their thresholds of their surveillance systems so low that it will detect their part of the abusive trading, as the consequence of that will be that they are flooded with false positives.

POSSIBLE STEPS

Possible steps that could to be taken to improve the overall monitoring scheme should focus on the detection phase and be based on a number of established cornerstones.

First, the data used in the detection phase must be full-covering, detailed and of high quality. Considering the different data-sources available today, this can only be obtained by using the data directly from the venues' trading systems as that is the only data source available with sufficient quality and level of detail. To ensure the data is full-covering, the NCAs need to have the overall responsibility for the detection phase by continuously collecting order-book data from all venues and feeding it into a surveillance system in one single format.

Secondly, the data used in the detection phase must be made available continuously and without delay. Any data that is requested on an ad hoc basis does not contribute in the detection phase.

The idea give earlier is identified by ESMA, which has launched a consultation paper[15] that among other things is evaluating the concept of a cross-market order-book surveillance based on order-book data from trading venues and operated by the NCAs.

The order-book data referred to in the consultation paper (CP) is based on RTS 24,[16] which only affects the trading

venues and not all execution venues, a fact that immediately creates a known gap in the overall monitoring. Also, the suggestion in the CP would certainly enable (almost) all order-book data for a specific instrument traded on multiple venues to be gathered and analysed together. Such an arrangement would, however, not cover all derivatives trading to be analysed with the trading of its underlying asset(s), which would still be a big challenge to overcome.

Concerning today's trading, the question is if it is even possible, even in theory, to create such a massive surveillance system with the data required to detect market abuse or if market-abuse monitoring enthusiasts around Europe will have to look forwards to the long-awaited retraction of MiFID.

References and Notes

(1) DIRECTIVE 2004/39/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 21st April, 2004, on markets in financial instruments.

(2) DIRECTIVE 2014/65/EU OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 15th May, 2014, on markets in financial instruments.

(3) REGULATION (EU) No 600/2014 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 15th May, 2014, on markets in financial instruments.

(4) 'Market identifier codes', available at: https://www.iso20022.org/10383/iso-10383-market-identifier-codes (accessed 11th December, 2019).

(5) 'Knight Capital Group', available at: https://en.wikipedia.org/wiki/Knight_Capital_Group (accessed 11th December, 2019).

(6) 'ESMA launches consultation on cost of market data and consolidated tape', available at: https://www.esma.europa.eu/press-news/esma-news/esma-launches-consultation-cost-market-data-and-consolidated-tape (accessed 12th July, 2019).

(7) Lannoo, K. (2017) 'MiFID II and the new market conduct rules for financial intermediaries: Will complexity bring transparency?', *ECMI Policy Brief*, No. 24.

(8) Commonly known as RTS 22; COMMISSION DELEGATED REGULATION (EU) 2017/590 of 28th July, 2016, Appendix.

(9) Commonly known as RTS 24; COMMISSION DELEGATED REGULATION (EU) 2017/580 of 24th June, 2016, Appendix.

(10) Guidelines on systems and controls in an automated trading environment for trading platforms, investment firms and competent authorities, ESMA/2012/122.

(11) COMMISSION DIRECTIVE 2003/124/EC of 22nd December, 2003, Art. 4 (c).

(12) ESMA/2012/122, GL 5.2.

(13) COMMISSION DELEGATED REGULATION (EU) 2016/522 of 17th December, 2015, Annex II.

(14) COMMISSION DELEGATED REGULATION (EU) 2016/522 of 17th December, 2015, Annex II, p. 2c and 2d.

(15) 'ESMA consults on MAR review', available at: https://www.esma.europa.eu/press-news/esma-news/esma-consults-mar-review (accessed 3rd October, 2019).

(16) Ref. 9 above.