Understanding Information Flow and Price Discovery in the Decentralized Foreign Exchange Market: An Information-Theoretic Approach

Aleksander Janczewski[*,1,2], **Ioannis Anagnostou**[1,4], **and Drona Kandhai**[1,2,3]

[1]Computational Science Lab, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands
[2]Quantitative Analytics, Financial Markets, ING Bank, Bijlmerdreef 98, 1102 CT Amsterdam, The Netherlands
[3]Korteweg de Vries Institute, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands
[4]Models and Portfolio Analysis, European Investment Bank, 98-100 Boulevard Konrad Adenauer, L-2950, Luxembourg, Luxembourg

ABSTRACT

The foreign exchange (FX) market is an integral component of the global economy with over $6 trillion in daily turnovers. The market has undergone substantial decentralization and fragmentation driven by the adoption of electronic trading and the influx of new market participants. Consequently, it has evolved into a complex system where locally created information percolates throughout the dealer network via high-frequency interactions before it is ultimately impounded into the asset prices. A significant piece of information such as a major economic announcement or a geopolitical event can spread through the network and trigger a cascade of trades or reactions. This may lead to extreme price volatility, liquidity disruptions and potential contagion effects on other financial markets. Yet limited research has sought to understand the mechanics of information flows in this market and address this question in a quantitative manner. We present a novel information-theoretic approach which, when applied on a unique, high-resolution dataset of bid and ask prices, reveals that better-informed dealers share more information and have a larger impact on the price discovery process. By employing conditional transfer entropy, we uncover a granular view of the flow of information among market participants, providing invaluable insights into the underlying dynamics of financial markets, particularly regarding flash crashes and other market disruptions. This paves the way for effective risk management, informed policy-making, and ensuring market efficiency.

Introduction

The FX market is a vital pillar of the global financial system, underpinning international trade by facilitating currency conversion for businesses worldwide. It provides the means to mitigate currency risk by granting market participants with liquidity, thus also helping to keep exchange rates stable. Furthermore, it stands as a barometer of global economic conditions, reflecting global economies' relative strength and health. The importance of the FX market extends far beyond its role in international trade and economic health. Price volatility in the FX market has far-reaching implications, affecting the purchasing power of individuals and stability of international trade, due to potential depreciation of the domestic currency and subsequent rise in import prices. This volatility can also introduce ripple effects in financial institutions, possibly resulting in bank failures, and thereby directly impacting the state of investments and retirement savings. While the FX market's key role is widely recognized, it is also imperative to consider its evolving state, especially in the light of contemporary structural transformations.

In recent years, the FX market has undergone considerable decentralization and fragmentation and thus, it has evolved into a highly complex system. The market structure transformation has been driven by the widespread adoption of electronic trading, which enabled faster and more efficient trading, but also increased the risk of flash crashes and other market disruptions. A growing number of participants, including retail investors and non-traditional actors such as hedge funds and high-frequency traders, has contributed to a more interconnected and intricate market structure. Additionally, the use of financial derivatives, such as FX futures and options, has further added to the complexity of the market.

Significant research efforts have been dedicated to unraveling complex systems and their underlying interactions in recent years[1–7]. Various approaches, including information-theoretic methods, have been proposed to better understand the underlying dynamics, interactions, and information flow within these intricate systems. One influential study by Bardoscia et al.[8] investigated pathways towards instability in financial networks. Their work highlighted the importance of understanding the interactions between components in complex systems and presented a framework to assess the stability of financial systems and the contagion mechanisms that can lead to cascading failures. In a different context, a study that introduced multivariate transfer entropy as a model-free measure of effective connectivity in the neurosciences is "Large-scale directed network inference

with multivariate transfer entropy and hierarchical statistical testing"[9]. Their work showcased the broad applicability of information-theoretic approaches in various fields, and through the utilization of transfer entropy, they were able to identify the directed flow of information within neural networks, providing a deeper understanding of the interactions and communication within these systems. Building on the potential of information-theoretic methods, Quax, Kandhai, and Sloot[10] employed metric called information dissipation to detect early-warning signals for financial market instability, specifically the Lehman Brothers collapse. Their findings demonstrated the potential of information-theoretic methods in identifying systemic risks, which could be crucial in preventing future financial crises and enhancing regulatory measures.

The study of information flows in the FX market is necessary to gain a deeper understanding of the underlying dynamics of this complex system. It can be a powerful tool for assessing the efficiency of the market, improving trading strategies, managing risk, making policy decisions, identifying opportunities for innovation in financial technology, and understanding flash crashes. For instance, by quantifying and analysing flows of information, market participants could potentially evaluate the risks associated with diverse market conditions and events, thereby enhancing risk management efficacy. The relationship between information flows and market volatility or liquidity disruptions could offer valuable insights for risk exposure assessments. Additionally, analysing patterns in information transmission could assist in identifying periods of increased stress and uncertainty, facilitating proactive risk management decisions and actions. For policymakers, understanding how information percolates across the market could highlight potential vulnerabilities and contagion risks, therefore contributing to systemic risk assessment. This could further guide the development of strategies to mitigate systemic risks and strengthen financial stability. In the context of market efficiency, assessing the speed and timing of information processing can inform how promptly information is integrated into prices, thereby aiding the assessment of price efficiency. Finally, studying information flows can also enhance our understanding of how new information influences prices, and provide insights into market depth and liquidity. In essence, the study of information flows in the FX market extends its utility beyond understanding the underlying dynamics of the system, with a broad range of potential applications in risk management and policy making.

Despite the promising potential of the information-theoretic approach, exposing the information flows in the FX market presents several challenges. First of all, the market is composed of multiple market participants, with different objectives and strategies, and the data that is used to inform prices are often heterogeneous and noisy. Secondly, the interactions between market participants and between different markets can be complex, making it difficult to understand how information flows through the market and how it affects prices. Finally, the FX market is often non-linear, meaning that small changes in one part of the market can have large effects elsewhere, which makes it difficult to predict how prices will respond to changes in information.

The literature presents various econometric approaches for characterizing the price discovery process in decentralized markets. Notable contributions include Gonzalo and Granger[11], Hasbrouck[12], and Harris et al.[13], who proposed methods and metrics to quantify the extent of each market's contribution to the price discovery process. Hasbrouck's information share and Harris-McInish-Wood component share metrics are based on the assumption that a market dominates the price discovery process if it has a dominant influence on the change of the long-term cointegration equilibrium price, also known as the common efficient price[12–14]. On the other hand, Granger and Gonzalo's permanent-transitory decomposition provides a divergent approach, wherein contribution is defined in terms of the market's error correction coefficients[11,15]. Nonetheless, all methods utilize the vector error correction model and rely on the assumptions of cointegration and common trend among decentralized markets, as set by Stock and Watson[16].

A novel approach to the price discovery process has been recently proposed by Hägstromer and Menkveld[17], who build upon the framework proposed by Hasbrouck and extend it with new metrics. Hägstromer and Menkveld not only investigate the influence of individual markets on the changes in the long-term efficient price of the asset, but also zoom in on the short-term dynamics of information flows in the dealer-network. By accounting for the short-term dynamics, Hägstromer and Menkveld are able to infer the flow of information in the network and also characterize the process of information revelation. To do so, the authors propose a multivariate impulse-response function-based approach to mapping information revelation empirically. Undoubtedly, cointegration-based approaches have received a lot of attention from the research community, as they are widely used to study the dynamics of various financial instruments and markets[18–21]. However, the agreement on which cointegration-based approach to the price discovery is the most appropriate has not been reached until this day[22]. Moreover, cointegration-based approaches are limiting as they as they can only be applied to financial instruments and markets characterized by cointegration relations. And finally, these approaches do not directly quantify the information flows between dealers in the dealer-network, but infer them instead. Given the above-mentioned limitations, there is a clear need for a new approach to studying information flow in the FX market, which is precisely what our study aims to address.

In this paper, we contribute to the literature on the price discovery process in decentralized markets by providing an information-theoretic approach that resolves the limitations posed by cointegration-based approaches and offers the means to directly infer the information flows between the markets. A relatively new information-theoretic metric called transfer entropy has been recently proposed by Schreiber[23], building upon the concept of entropy introduced by Shannon[24]. Transfer

entropy allows for detecting and quantifying observational causal, in a probabilistic sense[23,25], relationships between variables in complex systems, which, in the context of the FX market and price discovery, can be treated as quantifying information flows between dealers[26]. As Schreiber states, the attractiveness of transfer entropy is rooted in the fact that there is no need to make any assumptions about the dynamics of the system that one wants to investigate, thus making it a truly generic metric[23]. Our research utilizes a unique dataset of FX bid and ask prices, enabling us to directly measure the information flows between different market participants. The advantage of this approach over existing methodologies is that, rather than focusing on cointegration, transfer entropy is a directional measure which can infer causality and the flow of information between dealers in the FX market. Furthermore, transfer entropy captures both linear and non-linear relationships, providing a more comprehensive understanding of the complex interactions within the FX market, which is particularly important in financial settings where non-linear relationships are often prevalent.

To the best of our knowledge, this study represents the first investigation of FX dealer-network interactions using an information-theoretic network inference algorithm. Our study is also the first to analyze such an extensive, high-frequency data set, with over one million transfer entropy estimations performed. Furthermore, we present the first openly available implementation of this parallelized information-theoretic network inference algorithm in a low-level language. By applying this algorithm to the FX dealer-network interactions during the 2019 USD/JPY flash crash and an announcement related to the European Central Bank (ECB)'s asset purchase program in March 2020, we gained valuable insights into market information processing during these events. Our findings shed light on how FX dealers process information and interact with each other, which can have significant implications for market efficiency and stability. Importantly, while our study focuses on the FX market, the proposed network inference algorithm can be applied to a wide range of complex systems where time series data is available, serving as a powerful tool for uncovering interdependencies.

Results

Baseline USD/JPY

In this section, we present the results from our investigation of the flow of information in the USD/JPY dealer-network during the period between January 2nd to January 17th, 2019. Figure 1 displays the baseline information flow map generated for this time period. The flows exposed in the USD/JPY dealer network represent the average of family-wise, statistically significant (at a 10% level) conditional transfer entropies (CTEs) computed for each pair of dealers. In the baseline network, we observe numerous directional connections between dealers, with varying degrees of strength. One notable observation is that the most significant information flows appear to be directed from the more price efficient dealers towards the less price efficient dealers.

	Information share [%]		TE [nats]		CTE [nats]		Price inefficiency
Dealers	Upper	Lower	(Inflow)	(Outflow)	(Inflow)	(Outflow)	$\tau = 0$
B1	17.274	1.895	0.072	0.003	0.012	0.001	0.629
B2	54.526	9.585	0.003	0.034	0.003	0.009	0.175
B3	43.451	7.104	0.003	0.026	0.003	0.004	0.318
B4	48.874	10.581	0.002	0.033	0.002	0.009	0.110
B6	46.669	8.222	0.008	0.021	0.004	0.004	0.255
E1	10.353	1.573	0.030	0.002	0.005	0.001	0.429

Table 1. Summary of key information-theoretic and econometric metrics characterizing the USD/JPY baseline presented in Figure 1.

Moreover, we observe that dealers B2 and B4 appear to play a crucial role in shaping the information flows in the network. efficient dealers, with price inefficiencies of 0.175 and 0.110 respectively, they exhibit the highest total information outflows, with total CTE outflows of 0.009 nats each. Furthermore, our results indicate that B2 and B4 hold the largest information shares in the network, at 9.585% and 10.581%, respectively. Hence, we observe a clear alignment of econometric and information-theoretic metrics, suggesting that dealers B2 and B4 are dominant in terms of the information they hold and share

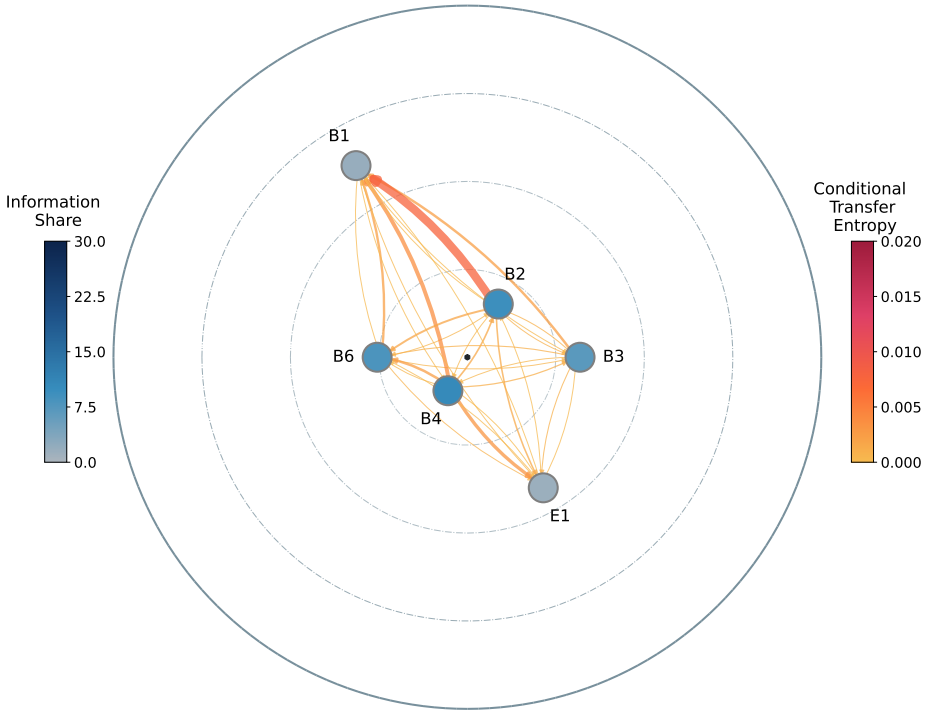

Figure 1. Information flow map illustrating mean information flows in the USD/JPY dealer-network for period between 2nd January and 17th January 2019. The position of each dealer (represented as nodes) is determined by the price inefficiency metric, with the closest dealers to the center of the unit circle being the most price efficient. The thickness and color of the directional edges between dealer pairs are proportional to the conditional transfer entropy. The color of each vertex corresponds to the Hasbrouck's information share, presented in percentage units, while the conditional transfer entropy is expressed in nats.

in the USD/JPY dealer-network. In addition to our observations on dealers B2 and B4, we also observe interesting information dynamics associated with bank B1 - the most price inefficient dealer in the network. In Figure 1 we can clearly see that all dealers in the network have relatively strong information flows directed towards bank B1. Further inspection of Table 1 reveals that bank B1 is characterized by the highest total information inflow of 0.012 nats, while also having the lowest information outflow of 0.001 nats. Moreover, bank B1 holds the second lowest information share in the network, and has the highest price inefficiency of 0.629.

Baseline EUR/USD

In Figure 2, we present the baseline information flow map generated for the EUR/USD currency pair for the time period between February 27th and March 27th, 2020. In this figure, we notice that market maker M1, the most price efficient dealer, is the source of very strong flows of information to other dealers.

Table 2 reveals that M1 has the highest total CTE outflow of 0.033 nats, while having the lowest total CTE inflows of 0.003 nats. Moreover, M1 holds the highest information share in the network at 15.489%. On the other hand, we note that very little information is flowing out of electronic trading platform E1, which is the most price inefficient dealer. The total CTE outflow for E1 is determined to be 0.001 nats, whereas the total CTE inflow is 0.010, the highest out of all dealers. Our results show a clear alignment of econometric and information-theoretic metrics that highlights the unique roles of both M1 and E1 in shaping the information flows of the EUR/USD dealer-network, with M1 being the dominant source and E1 being the dominant recipient of information.

	Information share [%]		TE [nats]		CTE [nats]		Price inefficiency
Dealers	Lower	Upper	(Inflow)	(Outflow)	(Inflow)	(Outflow)	$\tau = 0$
B1	2.454	47.812	0.076	0.070	0.009	0.002	0.472
B2	0.933	43.602	0.068	0.062	0.008	0.002	0.452
B3	2.699	50.183	0.023	0.129	0.009	0.009	0.385
B4	1.712	60.013	0.007	0.166	0.005	0.011	0.347
B5	0.508	34.494	0.102	0.048	0.007	0.001	0.543
B6	0.779	43.911	0.060	0.084	0.009	0.003	0.437
M1	15.489	78.717	0.004	0.183	0.003	0.033	0.131
E1	3.932	22.000	0.414	0.010	0.010	0.001	0.603

Table 2. Summary of key information-theoretic and econometric metrics characterizing the EUR/USD baseline presented in 2.

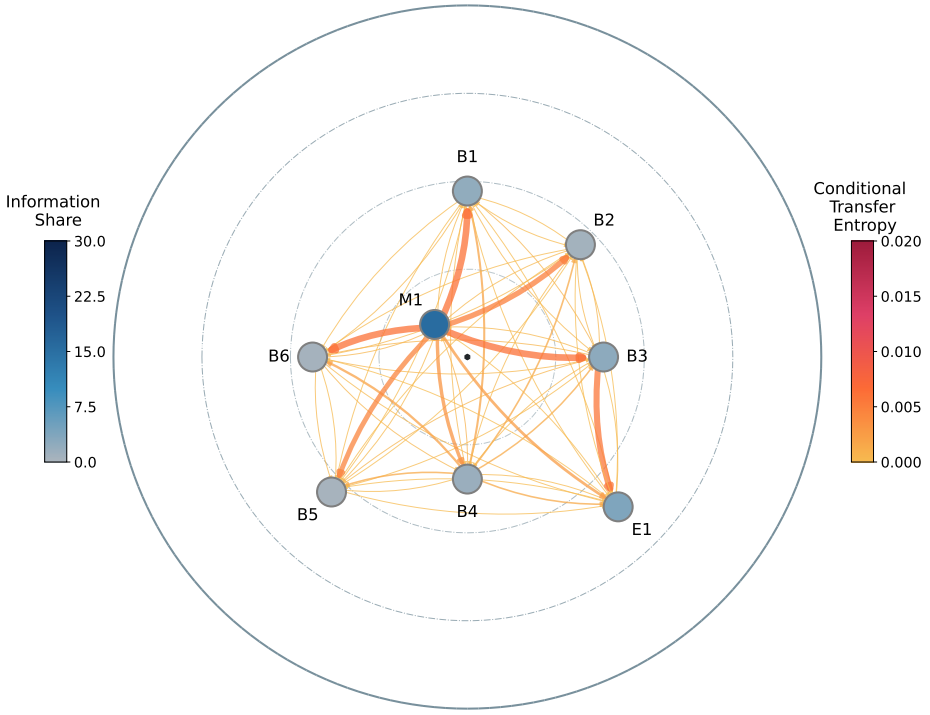

Figure 2. Information flow map illustrating mean information flows in the EUR/USD dealer-network for period between 27th February and 27th March 2020.

Linking information flows to price discovery

We find that information flows inferred with conditional transfer entropy provide a statistically significant explanatory power for fundamental price discovery metrics. For the analysis, we consider both the lower and upper bounds of Hasbrouck's information share. The upper bound information share accounts not only for the market's unique contribution to the changes in the efficient price, but also its correlation with other markets. On the other hand, the lower bound information share is considered to only account for the direct and unique contribution to changes in the efficient price that is uncorrelated with any other market[15]. In Figure 3, scatter plots are presented to show the relationship between the lower and upper bounds of the

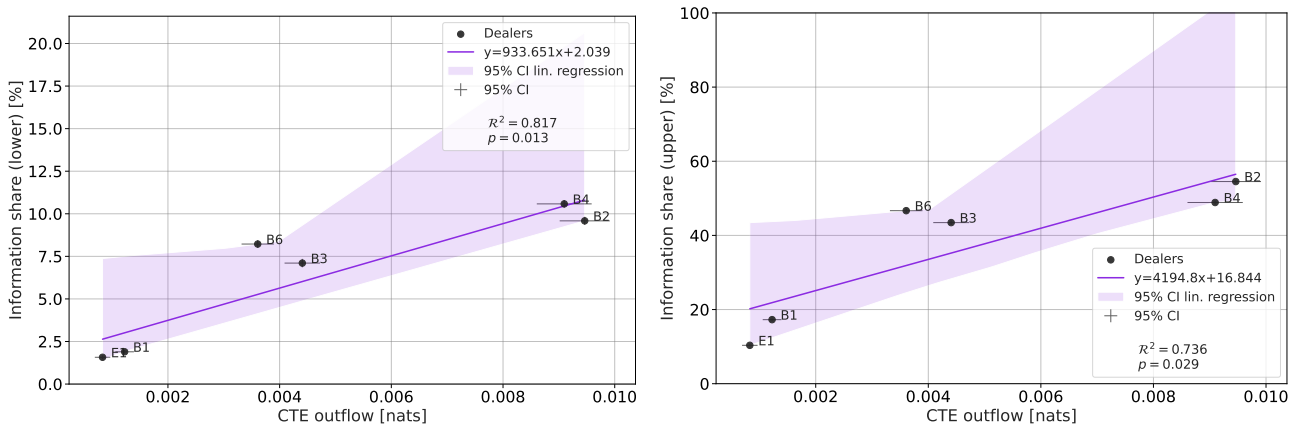

Figure 3. Scatter plot of CTE outflows versus Hasbrouck's lower (LHS) and upper (RHS) bound information shares for USD/JPY baseline.

information share and the total CTE outflow for each dealer. The left-hand side figure illustrates the relationship between the lower bound information share and CTE outflow. This relationship is statistically significant at a 5% level, with a p-value of 0.013 and an $\mathscr{R}^2$ of 0.817. The right-hand side figure visualizes the relationship between the upper bound information share and CTE outflow. This relationship is also statistically significant at a 5% level, with a p-value of 0.029 and an $\mathscr{R}^2$ of 0.736.

The observed trend suggests that dealers with the highest information share, i.e., the highest contributions to innovations in Hasbrouck's efficient price, exhibit the most information outflows. This aligns with expectations, as information flows estimated with conditional transfer entropy quantify the direct contributions of a dealer to changes in quotes of other dealers. The common efficient price relies on the prices of all dealers since it reflects all public and private information available in the dealer network. Consequently, the dealer with the most significant impact on the efficient price changes is the one that serves as the primary source of information for other dealers.

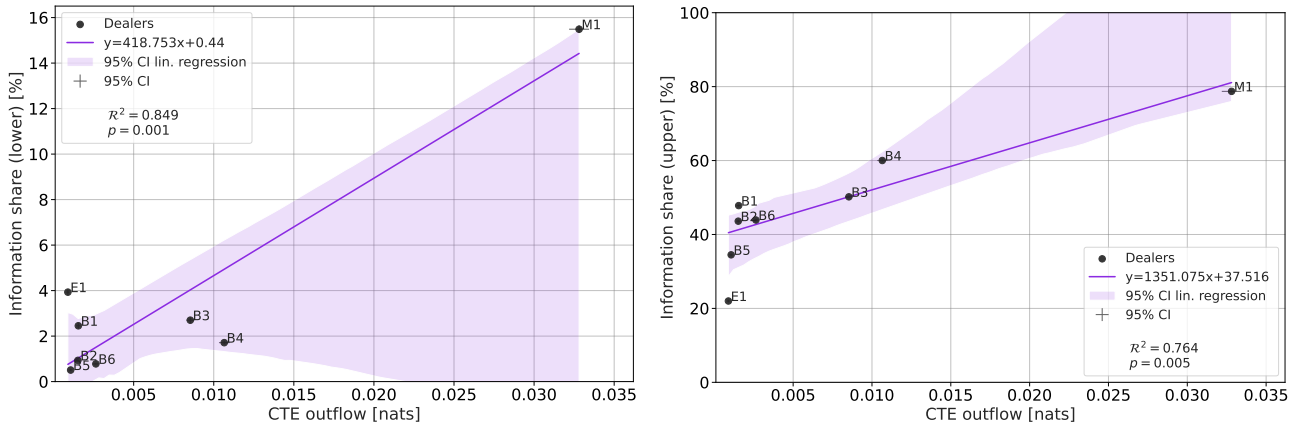

Figure 4. Scatter plot of CTE outflows versus Hasbrouck's lower (LHS) and upper (RHS) bound information shares for EUR/USD baseline.

The scatter plots for the EUR/USD dealer-network are shown in Figure 4. The relationship between CTE outflow and lower bound information share is significant at the 1% level, with a p-value of 0.001 and an $\mathscr{R}^2$ of 0.849. As a result, CTE outflows seem to explain a considerably larger portion of the variance in Hasbrouck's lower bound information share compared to TE outflows or dealer centrality. Moreover, the scatter plot reveals that market participants E1, B1, B2, B5, and B6 contribute very little unique information to the dealer-network. In contrast, banks B3 and B4 contribute significantly more information to the dealer-network. Market maker M1 is the dominant source of information for the dealer-network. The relationship between the upper bound information share and CTE outflows is also significant at the 1% level, with a p-value of 0.005 and a slightly weaker explanatory power of $\mathscr{R}^2$ of 0.764.

Analysis of the USD/JPY flash crash

On January 2nd, 2019, at 22:36 GMT time[27] the USD/JPY currency pair experienced a so-called flash crash, when the US Dollar depreciated by approximately 3% against the Japanese Yen in a mere 30 seconds[27,28]. The causes of the flash crash are not entirely understood but are believed to be a combination of factors, including significant liquidations of carry positions by Japanese retail investors, reduced liquidity due to the time of year and day, and the safety mechanisms of automated market making systems potentially amplifying the event[27–29].

In the remainder of this section, we focus on studying the flow of information between FX dealers during the flash crash event, as this analysis can offer valuable insights into the factors contributing to the crash and inform future risk management strategies. To investigate this, we generated informational maps for each hour leading up to, during, and after the flash crash using conditional transfer entropy (CTE) estimations. These maps depict the interactions and flow of information between dealers during each one-hour time window between 16:00 on January 2nd and 8:00 on January 3rd, as presented in Figures 5 and 6.

From 16:00 to 21:00, as shown in Figure 5, we observed minimal and relatively weak information flows between dealers. The strongest information flow occurred from bank B2 to bank B1, similar to the steady-state information map. Price inefficiency for all dealers, except E1, was roughly the same as in our baseline.

Panel F depicts the average information maps from 21:00 to 22:00, approximately 30 minutes before the flash crash. We observed notable changes in the dealer-network, such as all dealers becoming substantially more price inefficient and a slight decrease in the information share of the most informed banks. Interestingly, these changes occurred even before the flash crash, potentially serving as an early warning.

Panel G presents the dealer-network structure between 22:00 and 23:00, capturing the changes during the flash crash and the first 25 minutes after the event. We noticed an increase in price inefficiencies for all dealers, with bank B2 being the exception. Furthermore, we observed the emergence of more information flows compared to Panel F and an increase in information share for electronic trading platforms. This snapshot is unique, as it is the only one in which we can observe a relatively large information share for E1.

In subsequent panels, we observed that dealers gradually become more price inefficient. By Panel P, banks B1, B2, B3, B4, and B6 re-established their steady-state price inefficiencies, while the electronic trading platform E1 had not yet returned to its steady-state price inefficiency. The distribution of Hasbrouck's information share closely aligned with the baseline, suggesting that the dealer-network converged back to its steady-state structure.

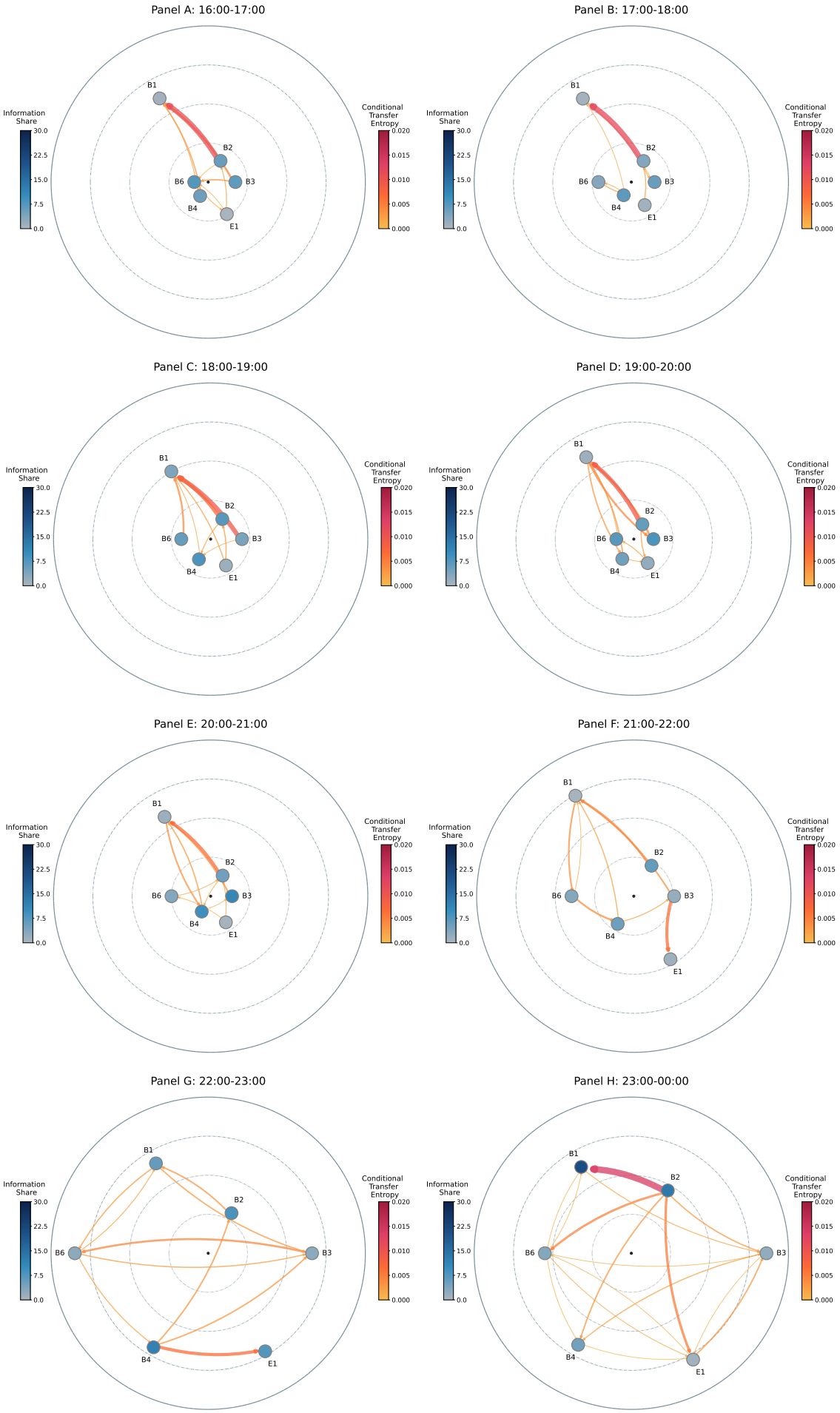

Figure 5. CTE information maps illustrating changes in the USD/JPY dealer-network between 16:00 and 24:00 on the 2nd January 2019 - the day of flash crash. Each panel is generated from 12 of five-minute-long subsamples.

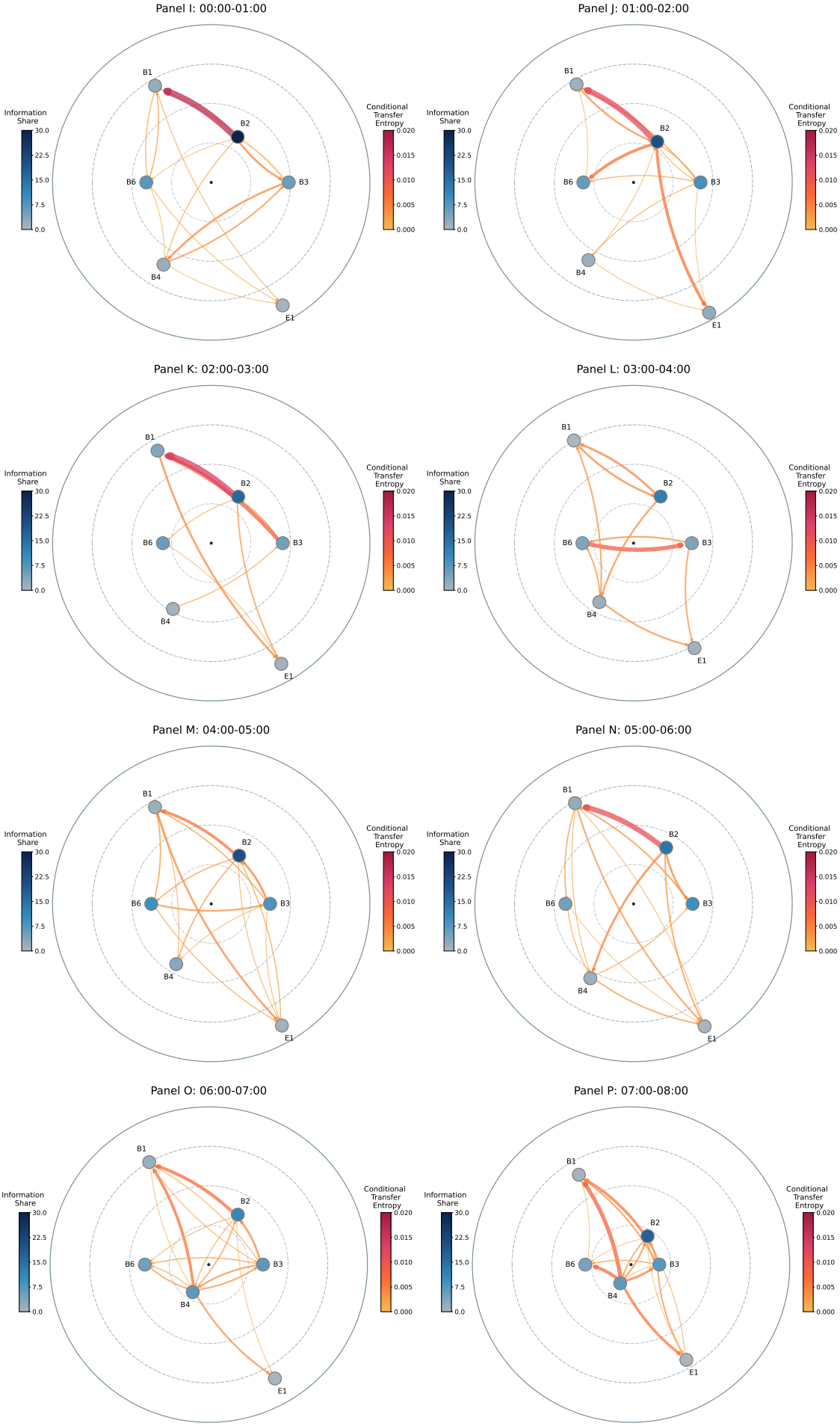

Figure 6. CTE information maps illustrating changes in the USD/JPY dealer-network between 24:00 and 8:00 on the 3rd January 2019 - the next day after the flash crash. Each panel is generated from 12 of 5-minute-long subsamples.

Analysis of the Impact of Asset Purchases Announcement

In January 2015, the Governing Council of the ECB announced an expanded asset purchase programme (APP) as part of its monetary policy measures to stimulate economic activity and increase inflation in the Eurozone. The program started in March 2015 and since that has been adjusted several times. (More information about the history and evolution of the ECB's Asset Purchase Programme can be found on ECB's website[30].) The APP can be considered a form of Quantitative Easing, a monetary policy strategy commonly employed by central banks globally[31,32]. Announcements of asset purchases may have an impact on exchange rates, with evidence pointing to announcement an expectation effects that tend to depreciate the Euro[33–35]. In this paper, we focus on obtaining an understanding of how information on asset purchases is getting processed by the market rather than studying the long term effects on the level of the exchange rates.

On 12 March 2020 at 13:30 GMT time, the ECB announced an effective expansion of the APP in the form of a temporary envelope of additional net asset purchases of €120 billion[36,37]. The aim of this measure was to promote favorable financing conditions for the actual economy during a period of elevated uncertainty, as the COVID-19 pandemic was spreading rapidly around the world and causing widespread economic disruption. In the rest of this section, we present the results of our analysis of the information flows in the EUR/USD dealer-network during the period surrounding the ECB's announcement. Previous research conducted by Hagströmer and Menkveld[17] has shown that public announcements can reduce the incentive for dealers to seek information. Consequently, we postulate that the ECB announcement would result in a reduction in information flows in the dealer-network. Moreover, we hypothesize that due to all dealers reacting to the same event, we should observe a temporary synchronization of changes in their quotes, reflected in a uniform reduction in price inefficiency across all dealers.

To examine this phenomenon, we analyze the CTE information maps generated for each hour between 8:00 and 20:00 GMT time, presented in Figures 7 and 8. These figures help us to visualize the changing dynamics in the dealer-network during the period of the ECB announcement.

In Panels A and B of Figure 7, we observe that between 8:00 and 10:00 hours the CTE information network closely resembles the baseline structure. The non-bank market maker is located close to the center of the map and is a source of many strong information outflows to other dealers. Additionally, the price inefficiency of all dealers is roughly the same as the baseline, with the only exception being dealer E1, which is significantly more price inefficient than average.

In Panels C and D, the information flows between dealers are weaker compared to the baseline and the previous hourly snapshots. In Panel C, dealers B3, B4, and E1 experience a slight decrease in their price inefficiencies, while in Panel D the non-bank market maker, M1, experiences a slight increase in price inefficiency.

In Panel E, approximately 30 minutes before the ECB announcement, we observe the emergence of many strong information flows stemming from M1 and a significantly lower information share of M1 compared to the baseline. All dealers' price inefficiencies are approximately the same as in the baseline.

In Panel F, which captures the changes that happened before and during the announcement between 13:00 and 14:00, all dealers except B5 become much more price efficient, located closer to the center of the map, with weak information flows primarily flowing from M1. The information share of M1 remains relatively low compared to the baseline.

In Panel G, approximately one hour following the announcement, all dealers are even more price efficient, with the largest increase observed in B1, B3, B5, and M1. The characteristic strong information flows from M1 are still present. This pattern continues in Panel H, where the dealer-network reaches its overall highest price efficiency.

In the following hours, represented in Panels I and J, dealers gradually move away from the center of the map as they become more price inefficient. In Panel K, the information map closely aligns with the steady-state structure observed in the baseline, with the only difference being that E1 is more price inefficient. The information flows between dealers are weaker than the baseline. This pattern is further verified in Panel L, where very little change is observed in the dealer-network, indicating that it has reached its steady-state structure.

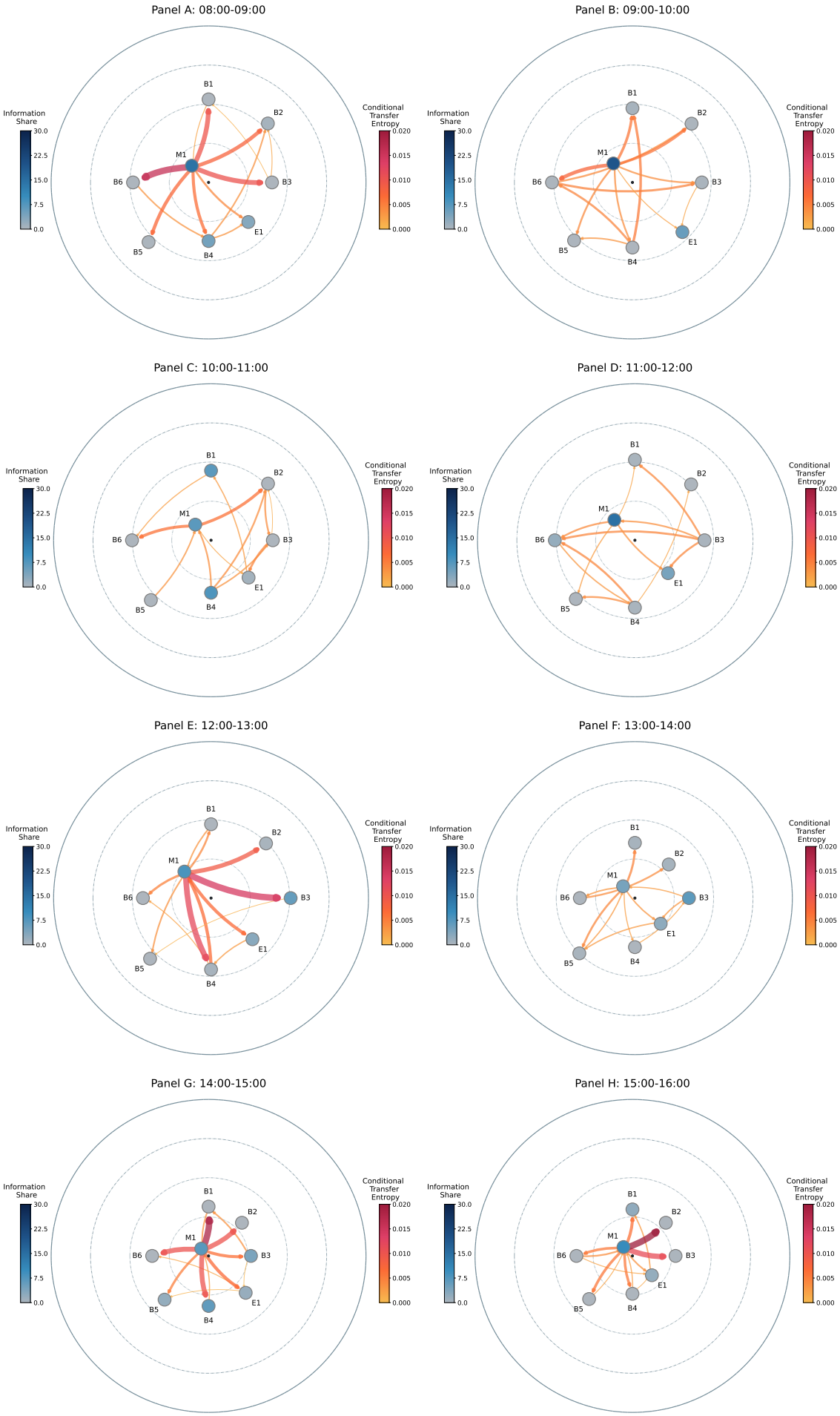

Figure 7. CTE information maps illustrating changes in the EUR/USD dealer-network between 8:00 and 16:00 on the 11th March 2020 - the day of ECB's announcement. Each panel is generated from 12 of five-minute-long subsamples.

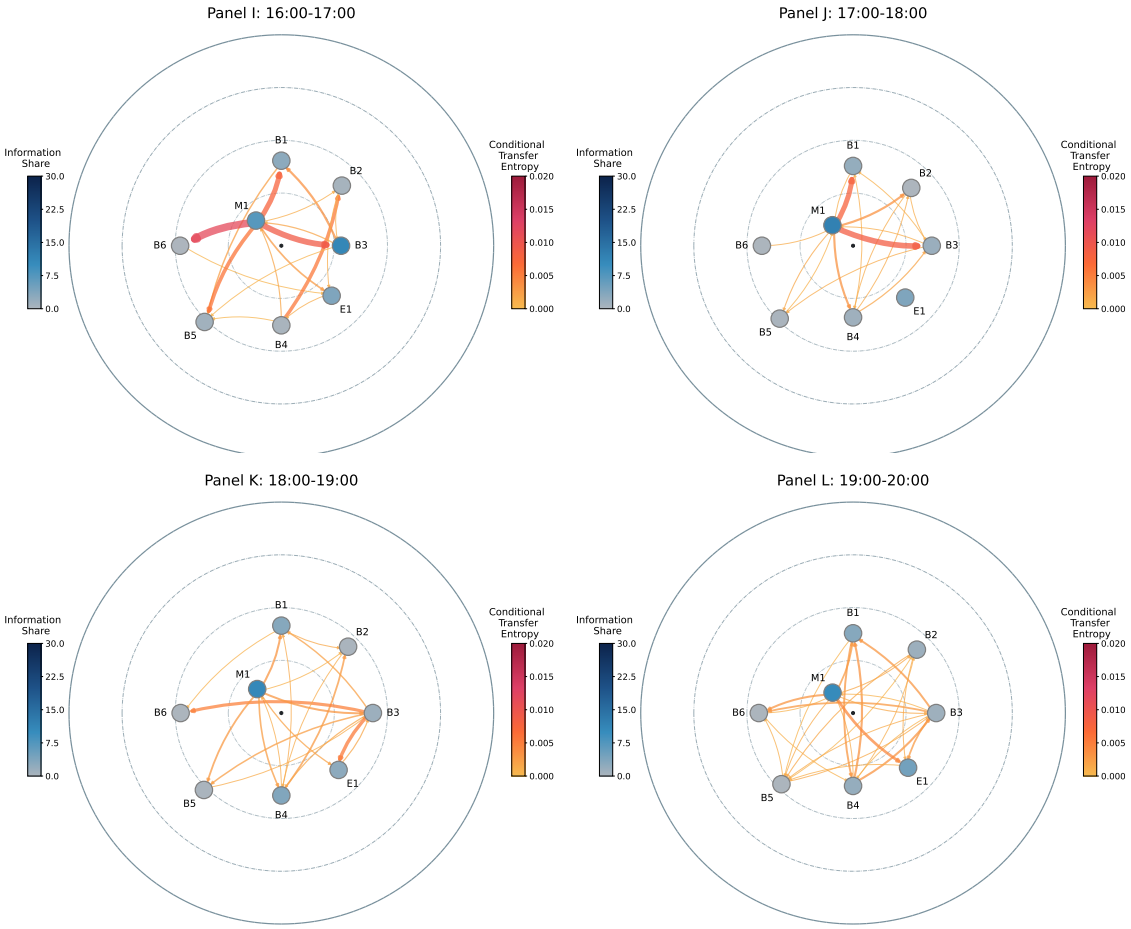

Figure 8. CTE information maps illustrating changes in the EUR/USD dealer-network between 16:00 and 20:00 on the 12th March 2020 - the day of ECB's announcement. Each panel is generated from 12 of five-minute-long subsamples.

Discussion

We observed that dealers with the highest contributions to the innovations in the efficient price had the most information outflows. Additionally, we demonstrated that information flows computed with information-theoretic metrics largely align with the variance-decomposition-based metric proposed by Hasbrouck. We also determined that CTE information flows explain a considerable portion of the variance of the lower and upper bounds of the information share metric.

As observed in Figures 5 and 6, the structure of the dealer-network changed drastically during and after the flash crash. Interestingly, even before the flash crash, we noticed considerable changes in the dealer-network structure, as some dealers became substantially more price inefficient compared to the previous one-hour snapshots and the baseline information map. In the information map capturing the flash crash, we observed a sudden, substantial increase in price inefficiency of all dealers. Additionally, during the flash crash, more information flows were present. Therefore, our investigation presents strong evidence supporting the hypothesis that dealers become more inclined to interact during extreme events like flash crashes. The maps show the formation of more information flows between dealers as they all became more price inefficient due to an abrupt change in the market.

We also examined the impact of public announcements on the FX dealer-network by analyzing structural changes in the network on the day of the ECB's asset purchase announcement, as shown in Figures 7 and 8. Interestingly, our results did not reveal any evidence supporting the idea that the exchange of private information between dealers is reduced at the times of public announcement. However, we made a notable observation that after the ECB's announcement, all dealers became more price efficient. Moreover, we found that the dealer-network structure converged back to its steady-state structure about five hours after the announcement.

FX market regulators can leverage insights into the behavior of FX dealers during high volatility events. By analyzing the flow of information and interactions among dealers during flash crashes or announcements, regulators can better understand the risks associated with different market participants, detect potential market misconduct, and inform policy-making decisions aimed at enhancing market transparency and stability. Such insights can ultimately help regulators ensure fair and efficient markets.

Methods

Data

The research was conducted using two distinct high-frequency FX spot rate data sets: EUR/USD, and USD/JPY. Each data set is comprised of privately collected, irregularly spaced, temporal data of FX rates from a selection of different dealers, provided to us by courtesy of ING Netherlands. Each data point includes two distinct time stamps; the original timestamp that is assigned at the origin by the dealer posting the quote (Timestamp1) and the timestamp set by ING when the quote update is received in the system (Timestamp2). Both timestamps are collected at a precision of 1 millisecond.

The EUR/USD data set is comprised of 109,449,980 high-frequency observations of the best bid and ask prices quoted by eight parties between February 27 2020, and March 27 2020. This time period encompasses 22 trading days and 8 non-trading days (weekends). Although trading is not conducted on Saturdays, it resumes on Sundays at 23:00 GMT. The USD/JPY data set is comprised of 14,284,491 observations of bid and ask prices quoted by six parties between January 1 2019, and January 17 2019. In this period, there are 12 trading days and 4 non-trading days.

In the investigation, we distinguish three types of dealers; the market makers (M), banks (B), and electronic trading platforms (E). Note that while banks often serve as market makers as well, we differentiate between banks and non-bank market makers in this context.

Data processing

To prepare the high-frequency FX spot rate data for the analysis, we transform it into a regularly spaced time series with a consistent sampling frequency across all subsamples using the resampling and forward-fill approach. We employ the forward-fill approach to ensure that no information from the future is propagated backward in time.

As our objective is to accurately reveal the information flow and the price discover process from dealers' perspective, we choose to analyze the data using timestamp assigned at the origin by the dealer who posts the quote (Timestamp1). To ensure that our results are not distorted by the electronic signal latency (the time it takes for an electronic signal to travel from its origin to destination) or ING and dealers' clocks asynchronization, we analyze the differences between the two timestamps assigned to each quote. By comparing the time evolution of the differences between Timestamp1 and Timestamp2, we ensured that the latencies stay approximately constant over time in all data sets. Given that the highest latency observed is no larger than 100 milliseconds, we conclude that the sampling period of 100 milliseconds is sufficient for all dealers to observe other dealers' updates and take action. Therefore, we resample the data to 100 milliseconds - the same sampling frequency used by Hagströmer and Menkveld[17]. For a more detailed discussion of the data preprocessing methods and latency adjustment considerations, please refer to section S1.1 in the supplementary material[38].

With the aim of determining the optimal interval for daily sampling, we investigate the average intraday quote frequency for each dealer within each dataset independently. By doing so we uncovered how the average intraday quote frequency of dealers evolves throughout the day. The highest average quote frequency is observed during the overlapping period of the London and New York sessions. Based on these observations, we opt to sample from each day eight-hour-long time windows between 8:00 and 16:00 on consecutive days, as this interval exhibits on average the highest trading activity in both EUR/USD and USD/JPY data sets. Further information on the intraday quote frequency is available in section S2.3 of the supplementary material[38].

Afterwards, each time series is differenced to ensure stationarity and split into five-minute-long subsamples. Effectively, an 8-hours-long time window yields 96 five-minute-long subsamples. Consequently, the resampling process yields 2,112 subsamples for the EUR/USD baseline, and 1,152 subsamples for USD/JPY baseline. From 2,112 subsamples, a total of 95 subsamples were excluded; in 67 subsamples the cumulative impulse response function did not converge, and in 28 subsamples all of the dealers had less than 10 midpoint updates within five-minute time window. Additionally, only in 12 subsamples not all dealers were included in the analysis due to some of them having insufficient number quote updates.

In the following steps, the subsamples are tested for unit test and cointegration relations. Stationarity is an essential assumption for both the econometric and information-theoretic models, whereas cointegration is crucial assumption for the econometric model alone. To test subsamples for stationarity we performed the augmented Dickey-Fuller test to verify stationarity for each 5-minute-long time series, resampled to 100ms time intervals. The results show that 99.4% and 98.3% of subsamples in EUR/USD and USD/JPY datasets, respectively, are integrated of order one at 5% confidence level, satisfying the stationarity assumption. Additionally, we conducted the augmented Engle-Granger two-step cointegration test for each dealer-dealer pair within each five-minute-long subsample to ensure that cointegration relations are present. The results demonstrate that 96.3% and 95.8% of subsamples in EUR/USD and USD/JPY datasets, respectively, are cointegrated at 5% confidence level, satisfying the cointegration assumption. Note that we employ the Bonferonni correction to account for the large number of hypothesis tests performed. For a more detailed account of the data preprocessing methods and latency adjustment considerations, please refer to section S1.1 in the supplementary material[38].

For the information-theoretic network inference algorithm, apart from transforming data into stationary time series, Kraskov et al.[39] suggest to standardize the time series to zero mean and unit variance; this approach is widely accepted by the scientific

community[40,41]. After the time series is standardized, we also introduce very low-amplitude noise to the data, since for double-precision floating-point operations, Kraskov et al.[39] suggests to add noise of order 10^{-10}. This treatment is essential when one works with empirical data with limited precision, potentially resulting in many points having equal values. Consequently, this would lead to breaking the assumption of continuously distributed points and result in spurious estimates[39]. Adding noise to the time series introduces very slight stochasticity to the KSG estimator. However, if the information transfer is significant, it will remain significant after the addition of the noise[42].

Econometric tools and techniques

To develop the econometric model and calculate the key econometric metrics, we meticulously follow the methodology of Hagströmer and Menkveld outlined in their publication entitled *Information Revelation in Decentralized Markets*[17]. In their publication, Hagströmer and Menkveld build upon the vector-error correction model-based framework proposed by Hasbrouck[12], extending it with multivariate impulse response function-based metrics.

In line with their approach, we employ the Bayesian Information Criterion (BIC) to determine the optimal number of lags in the VECM model for each subsample considered. The maximum lag that is considered is lag 20. Although Hagströmer and Menkveld's work does not explicitly mention this maximum lag selection, it is retained as a default parameter in their publicly available R code implementation of the model. Additionally, to ensure that cointegration relations can be uncovered, we require that the subsample of data from a specific dealer has at least 10 changes in their quote midpoints, otherwise, the dealer is excluded from the subsample. The same approach is employed by Hagströmer and Menkveld.

Additionally, the common trend representation introduced by Stock and Watson[16] is employed to uncover the common trend component for the cointegrated processes which is assumed to represent the efficient price. The first econometric metric that we employ is the Hasbrouck's information share metric, which quantifies the contribution of each market to the innovations in the efficient price. To estimate the lower bound of the information share we follow the methodology of Hagströmer and Menkveld:

$$\text{Lower bound InfoShare}_i = \min_{x \in \text{Aut}(X)} \frac{\left([\psi F_x]_{x(i)} \right)^2}{\psi \Omega \psi^\intercal} \tag{1}$$

where Ω stands for the covariance matrix of innovations captured with vector-error correction model, and ψ denotes the long-run cumulative price change common to all markets. X is the sequence of markets used in VECM modelling, $\text{Aut}(X)$ stands for the automorphism with all possible combinations of market sequences[17]. Hence, x simply stands for one particular combination of market sequence, the F_x is the Cholesky factor of Ω which was resequenced according to x and $x(i)$ stands for the position of market i in the sequence x[17]. The upper bound of the information share is estimated by simply taking the maximum of all market sequence combinations.

Another econometric metric that is employed in the model is the price inefficiency introduced by Hägstromer and Menkveld[17]. Price inefficiency is defined as a fraction of the information that the dealer has yet to impound into his prices. To compute this metric, first we need to determine the so-called beta coefficient associated with jth market:

$$\beta_{j,\tau} = \frac{\text{Cov}[\tilde{\Psi}_{j,\tau}\varepsilon_t, \psi\varepsilon_t]}{\text{Var}(\psi\varepsilon_t)} = \frac{\tilde{\Psi}_{j,\tau}\Omega\psi^\intercal}{\psi\Omega\psi^\intercal} \tag{2}$$

where $\tilde{\Psi}_{j,\tau}$ is the matrix with cumulative price responses until τ of dealers to the multivariate unit shocks and ε_t captures vector-error correction model residuals. The price inefficiency of dealer j at τ is defined in the following manner:

$$\text{Price inefficiency}_{j,\tau} = \mid 1 - \beta_{j,\tau} \mid \tag{3}$$

Information-theoretic metrics

To quantify the information flows in the FX dealer network we developed an information-theoretic network inference algorithm that utilizes two metrics; transfer entropy (also known as apparent entropy) and conditional transfer entropy. Transfer entropy, introduced by Schreiber[23], quantifies the amount of information that the past states of a source process contribute to predicting the future state of a target process, while accounting for the information about the future already ingrained in target's past states. In the context of Markov chains, transfer entropy can be interpreted as a measure of deviation from generalized Markov property that can be computed with Kullback-Leibler divergence. Hence, the transfer entropy between two continuous processes from X_t (source process) to Y_t (target process) with time-lag of 1 is defined as Kullback-Leibler divergence between two transitional probabilities; the one conditioned on the past states of both target and source processes $f(y_{t+1}|y_t^{(d_y)}, x_t^{(d_x)})$ and the one that only accounts for the past states of the target process $f(y_{t+1}|y_t^{(d_y)})$. Continuous transfer entropy can be computed with the following

formula

$$\mathrm{TE}_{X_t \to Y_t}^{(d_y,d_x)} \triangleq D\left(f(y_{t+1}|y_t^{(d_y)}, x_t^{(d_x)}) || f(y_{t+1}|y_t^{(d_y)})\right) \tag{4}$$

$$= \int_{\mathbb{R}^{d_x}} \int_{\mathbb{R}^{d_y}} \int_{\mathbb{R}} f(y_{t+1}, y_t^{(d_y)}, x_t^{(d_x)}) \log\left(\frac{f(y_{t+1}|y_t^{(d_y)}, x_t^{(d_x)})}{f(y_{t+1}|y_t^{(d_y)})}\right) dy_{t+1} dy_t^{(d_y)} dx_t^{(d_x)} \tag{5}$$

where $f(y_{t+1}, y_t^{(d_y)}, x_t^{(d_x)})$ represents the joined probability density function, and the $\int_{\mathbb{R}^D}$ denotes the D-dimensional integral over the support of the variable[43]. Vector $y_t^{(d_y)}$ represents past d_y states of a process y_t, where d_y is the history length of the process also called the embedding dimension. By choosing some finite history length of the process, we are making an assumption that the process considered is a Markov process of order d, hence it is conditionally independent of the states of history further than the past d states. Analogously, d_x represents the history lengths of processes X_t. The methodology for choosing the optimal embedding dimension is further outlined in the Algorithm section.

To compute the conditional transfer entropy, we need to extend the transfer entropy formula to account for potential contributions from another causal information contributor process, say Z_t. Consequently, the formula for continuous transfer entropy is as follows:

$$\mathrm{CTE}_{X_t \to Y_t | Z_t}^{(d_y,d_x,d_z)} \triangleq D\left(f(y_{t+1}|y_t^{(d_y)}, x_t^{(d_x)}, z_t^{(d_z)}) || f(y_{t+1}|y_t^{(d_y)}, z_t^{(d_z)})\right) \tag{6}$$

$$= \int_{\mathbb{R}^{d_z}} \int_{\mathbb{R}^{d_x}} \int_{\mathbb{R}^{d_y}} \int_{\mathbb{R}} f(y_{t+1}, y_t^{(d_y)}, x_t^{(d_x)}, z_t^{(d_z)}) \log\left(\frac{f(y_{t+1}|y_t^{(d_y)}, x_t^{(d_x)}, z_t^{(d_z)})}{f(y_{t+1}|y_t^{(d_y)}, z_t^{(d_z)})}\right) dy_{t+1} dy_t^{(d_y)} dx_t^{(d_x)} dz_t^{(d_z)} \tag{7}$$

where vector $z_t^{(d_z)}$ represents the past d_z states of a process Z_t - the process which could potentially also reduce the uncertainty about the future state of the target process. By conditioning on process Z_t we can filter out its contributions, and thus determine the unique information flow from process X_t to Y_t. We note that while conditioning does eliminate redundancy, it may also introduce synergistic contributions from multivariate interactions between the conditional variables and the target variable. In Tables 1 and 2, we observe a substantial reduction in the average weights of the information flow when CTE is considered instead of TE. This reduction suggests that conditioning effectively filtered out influences from other information contributors, leaving unique contributions from each dealer. Although it is uncertain whether any synergistic interactions were captured, the fact that all CTE edges are significantly smaller than the TE edges, implies that overall more redundant information was filtered out than synergistic information introduced.

Both transfer entropy and conditional transfer entropy can be also represented as the conditional time-delayed mutual information, which takes the following form:

$$\mathrm{TE}_{X_t \to Y_t}^{(d_y,d_x)} = I(Y_{t+1}; X_t^{(d_x)} | Y_t^{(d_y)}) \tag{8}$$

$$\mathrm{CTE}_{X_t \to Y_t | Z_t}^{(d_y,d_x,d_z)} = I(Y_{t+1}; X_t^{(d_x)} | Y_t^{(d_y)}, Z_t^{(d_z)}) \tag{9}$$

where I represents the mutual information.

Estimation of apparent and conditional transfer entropies requires the approximation of two transitional and one joint probability densities of the involved processes, typically based on a single realization of these processes. In our network inference algorithm, we estimate continuous apparent and conditional transfer entropies using nearest-neighbor-based Kraskov, Stögbauer, and Grassberger (KSG) algorithm I, following the methodology of Kraskov[39] and Frenzel and Pompe[44], respectively. KSG algorithm has recently received much attention, especially in neuroscience, given that it is numerically unbiased for finite samples and relatively robust in high-dimensional settings[40,45,46]. With the developments in KSG algorithms, for example, an extension of the algorithm to conditional mutual information by Frenzel and Pompe[44], the application of transfer entropy to network inference algorithm has recently gained much attention, especially in bioinformatics and neuroscience[47].

Transfer entropy is estimated with KSG algorithm I, which takes the following form:

$$\widehat{I}(Y_{t+1}; X_t^{(d_x)} | Y_t^{(d_y)})_{KSG(1)} = \psi\left(K_{Y_{t+1} X_t^{(d_x)} Y_t^{(d_y)}}\right) + \left\langle \psi\left(n_{Y_t^{(d_y)}} + 1\right) - \psi\left(n_{Y_{t+1} Y_t^{(d_y)}} + 1\right) - \psi\left(n_{X_t^{(d_x)} Y_t^{(d_y)}} + 1\right) \right\rangle \tag{10}$$

where, $\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$ is the digamma function - the derivative of the log of the gamma function[39]. Furthermore, $n_{Y_t^{(d_y)}}$ represents the number of neighbours found in space $(Y_t^{(d_y)})$, $\langle \cdots \rangle$ denotes the average over all data points, and $K_{Y_{t+1} X_t^{(d_x)} Y_t^{(d_y)}}$ represents the K-th nearest-neighbor that is used to determine the distances in the joint space $\left(Y_{t+1}, X_t^{(d_x)}, Y_t^{(d_y)}\right)$.

Algorithm

The information-theoretic network inference algorithm involves many steps.

Step 1 Data preprocessing: Initially, the data in a subsample is preprocessed as detailed in the Data preprocessing section.

Step 2 Optimal dimension and delay embedding: Afterward, for each dealer (source) and dealer (target) pair in the subsample, we jointly optimize both the embedding dimension and delay parameters using the Ragwitz criterion[48]. Ragwitz criterion relies on a locally constant predictor of the future state (w_{t+1}) of embedding vector w_t^k, with the future state prediction estimated from the nearest neighbors of the variable after embedding is applied[49]. For a given neighborhood diameter ε, the neighborhood of w_t^k, denoted by $\mathscr{U}_n$, is defined as $\mathscr{U}_n = \{w_n^k : ||w_n^k - w_t^k|| \leq \varepsilon\}$. Using this neighborhood, the Ragwitz criterion computes a locally constant estimate $\widehat{w}_{t+1}$ of the future state w_{t+1} as follows:

$$\widehat{w}_{t+1} = \frac{1}{|\mathscr{U}_n|} \sum_{w_n^k \in \mathscr{U}_n} y_{n+1} \tag{11}$$

This estimation simply amounts to taking the mean value of the future states (w_{n+1}) of the nearest neighbors of the embedding vector w_t^k. In practice, ε is replaced by a specific number of K nearest neighbors taken into account for the estimation, which serves as a "natural" substitute when the KSG algorithm is employed[40]. Next, the squared error of the local predictor is computed for each time index in the embedding vector, and the mean squared error is determined. The mean squared error is computed for each combination of embedding dimension and embedding delay parameters that one chooses to investigate. Based on Ragwitz's criterion, the parameters that yield the smallest mean squared error are optimal. In our investigation, the maximal embedding dimension and delays are set to 4, simply because higher embedding dimensions would be computationally unfeasible given the number of subsamples considered.

Step 3 Optimal source-target delay: In the following step, for each dealer (source) and dealer (target) pair in the subsample, we determine the true delay between source and target processes following the method proposed by Wibral et al.[50]. The optimal source-target delay is determined to be, the delay for which the largest transfer entropy is determined. Namely,

$$\delta = \underset{u \in \mathscr{U}}{\operatorname{argmax}} \left(I(Y_{t+1}; X_{t-u}^{(d_x)} | Y_t^{(d_y)}) \right) \tag{12}$$

where u represents source-target delay, $\mathscr{U}$ is the investigated delay-space, i.e. $\mathscr{U} = \{0, 1, 2, \ldots, u_{\max}\}$, and δ represent the true delay between source and target processes. Our algorithm examines source-target delays up to a maximum of 4, which, considering the resampling interval, is equivalent to a 400 ms delay. From our latency investigation and the literature, it is apparent that 400ms is enough time for dealers to react to quote updates of other dealers. Furthermore, the analysis of the distribution of true source-target delays revealed that most of the true delays recovered were equal to 1.

Step 4 Estimating transfer entropy: In the fourth step of our approach, the transfer entropy from the source dealer to the target dealer is estimated using the KSG algorithm I. First, we concatenate the vectors of all processes $(Y_{t+1}, X_t^{(d_x)}, Y_t^{(d_y)})$ to create a joint embedding space. Next, we determine the Chebyshev distance to the K-th nearest neighbor (in our case, the 8th nearest neighbor) for each time point using a modified Scipy's KD-tree algorithm[51]. Then we use the obtained distance to identify the number of nearest neighbors, i.e. $n_{Y_t^{(d_y)}}$, $n_{Y_{t+1}Y_t^{(d_y)}}$ and $n_{X_t^{(d_x)}Y_t^{(d_y)}}$ in the respective spaces $(Y_t^{(d_y)})$, $(Y_{t+1}, Y_t^{(d_y)})$, $(X_t^{(d_x)}Y_t^{(d_y)})$. This process is repeated for all vectors in the subsample to calculate the average number of nearest neighbors found for each distance in each space, as in Equation 10. This yields one transfer entropy estimate for a single pair of (source) dealer and (target) dealer.

K effectively controls the bias of the estimation, which here scales with a factor K/N[39]. Based on the results from the literature and numerical experiments with the FX data, we choose $K = 8$ which very significantly reduced the variance of the estimate, while maintaining negligible systematic errors. This choice was also motivated by large variance of transfer entropy estimates for $K = 4$ observed for the time windows that included EUR/CHF and USD/JPY flash crashes.

Step 5 Statistical testing: Afterward, we perform permutation testing to determine the statistical significance of the transfer entropy estimates. When performing the statistical assessment with permutation technique, we are essentially assessing whether the estimated Kullback-Leibler divergence between transitional probabilities $f(y_{t+1}|y_t^{(d_y)})$ and $f(y_{t+1}|y_t^{(d_y)}, x_t^{(d_x)})$ is indeed statistically significant. We test for the null hypothesis that the state changes $y_t^{(d_y)} \to y_{t+1}$ have no temporal dependence on the source process $x_t^{(d_x)}$[52], hence:

$$\mathbf{H_0} : f(y_{t+1}|y_t^{(d_y)}) = f(y_{t+1}|y_t^{(d_y)}, x_t^{(d_x)}) \implies \widehat{\mathrm{TE}}_{X_t \to Y_t}^{(d_y, d_x)} = 0 \tag{13}$$

$$\mathbf{H_1} : f(y_{t+1}|y_t^{(d_y)}) \neq f(y_{t+1}|y_t^{(d_y)}, x_t^{(d_x)}) \implies \widehat{\mathrm{TE}}_{X_t \to Y_t}^{(d_y, d_x)} > 0 \tag{14}$$

The one-sided alternative hypothesis is motivated by the fact that transfer entropy is a non-negative measure. To test the above-presented hypotheses, we generate an empirical distribution of transfer entropy estimates under null hypothesis. We achieve this by generating a large number of source process surrogates (X_t^s), which preserve transitional probability $p(y_{t+1}|y_t^{(d_y)})$, but destroy the dependence in $p(y_{t+1}|y_t^{(d_y)}, x_t^{(d_x)})$[40]. We generate the S of source process surrogates (X_t^s) by shuffling the vectors of past states $x_t^{(d_x)}$ among the set of $\{y_{t+1}, y_t^{(d_y)}, x_t^{(d_x)}\}$ tuples - i.e. interchanging their time indices within the time-series, as proposed by Lizier[40]. Then, the p-value is determined by simply counting the number of cases when $\widehat{\mathrm{TE}}_{X_t^s \to Y_t}^{(d_y,d_x)} > \widehat{\mathrm{TE}}_{X_t \to Y_t}^{(d_y,d_x)}$, i.e.:

$$\text{p-value} = \frac{1}{S}\sum_{i=1}^{S} 1\left(\widehat{\mathrm{TE}}_{i\ X_t^s \to Y_t}^{(d_y,d_x)} > \widehat{\mathrm{TE}}_{X_t \to Y_t}^{(d_y,d_x)}\right) \tag{15}$$

where $1(\cdot)$ denotes the indicator function, and S is the number of surrogate source processes. For a given significance level α, we reject $\mathbf{H_0}$ if p-value $< \alpha$[52]. Additionally, one should note that since we are going to conduct multiple hypothesis tests, we need to employ Bonferroni correction. In our research, S is set to 500 because this is the largest number of permutations that was feasible given our computational limitations. Note that we compute a total of 236,544 entropy estimates (only for the EUR/USD data set). Thus, with 500 permutations for each estimate, we are looking at computing approximately up to 118 million entropy estimations, which is already quite a large number. Steps 2-5 are repeated for all possible source and target dealer pairs in a given subsample.

Step 6 Estimating conditional transfer entropy: For all statistically significant transfer entropy pairs between source and target dealers, we recompute the transfer entropy while conditioning on all other statistically significant information contributors to the source process. To accomplish this, we essentially repeat Steps 2-5 for statically significant information flows, with the key difference being that we now also account for all statistically significant information contributors identified through the initial transfer entropy calculation. This approach allows us to filter out redundant information flows and focus on the unique information that flows from source and target dealers. For more details on the network inference algorithm, please refer to Sections S1.3 and S1.4 in the supplementary material[38].

Validation

We validated the econometric model using the EUR/CHF data set, which consists of 96 subsamples, each with a length of 5 minutes. First, we independently validated each metric used in the econometric model against the openly provided model by Hagströmer and Menkveld[17]. Next, we validated the transfer entropy metric with the analytical solution provided by Kaiser and Schreiber[53]. Additional details regarding the validation process and its results can be found in section S3 of the supplementary material[38].

Code availability

The code associated with this research, including the econometric model and the information-theoretic network inference algorithm, has been developed in C++ and is open-source. It can be accessed through the following link: Github.

References

1. Bonanno, G., Lillo, F. & Mantegna, R. N. Levels of complexity in financial markets. *Physica A: Statistical Mechanics and its Applications* **299**, 16–27 (2001).
2. Bouchaud, J.-P. & Cont, R. A Langevin Approach to Stock Market Fluctuations and Crashes. *The European Physical Journal B-Condensed Matter and Complex Systems* **6**, 543–550 (1998).
3. Bouchaud, J.-P. The (unfortunate) complexity of the economy. *Physics World* **22**, 28 (2009).
4. Battiston, S. *et al.* Complexity theory and financial regulation. *Science* **351**, 818–819 (2016).
5. Bardoscia, M. *et al.* The Physics of Financial Networks. *Nature Reviews Physics* **3**, 490–507 (2021).
6. Squartini, T., Van Lelyveld, I. & Garlaschelli, D. Early-warning signals of topological collapse in interbank networks. *Scientific reports* **3**, 3357 (2013).
7. Cimini, G., Squartini, T., Garlaschelli, D. & Gabrielli, A. Systemic Risk Analysis on Reconstructed Economic and Financial Networks. *Scientific reports* **5**, 1–12 (2015).
8. Bardoscia, M., Battiston, S., Caccioli, F. & Caldarelli, G. Pathways towards instability in financial networks. *Nature communications* **8**, 14416 (2017).
9. Novelli, L., Wollstadt, P., Mediano, P., Wibral, M. & Lizier, J. T. Large-scale directed network inference with multivariate transfer entropy and hierarchical statistical testing. *Network Neuroscience* **3**, 827–847 (2019).
10. Quax, R., Kandhai, D. & Sloot, P. Information dissipation as an early-warning signal for the Lehman Brothers collapse in financial time series. *Scientific reports* **3**, 1–7 (2013).
11. Gonzalo, J. & Granger, C. Estimation of common long-memory components in cointegrated systems. *Journal of Business and Economic Statistics* **13**, 27–35 (1995).
12. Hasbrouck, J. One Security, Many Markets: Determining the Contributions to Price Discovery. *The Journal of Finance* **50**, 1175–1199 (1995).
13. Harris, F. H. d., McInish, T. H., Shoesmith, G. L. & Wood, R. A. Cointegration, Error Correction, and Price Discovery on Informationally Linked Security Markets. *Journal of Financial and Quantitative Analysis* **30**, 563–579 (1995).
14. Putniņš, T. J. What Do Price Discovery Metrics Really Measure? *Journal of Empirical Finance* **23**, 68–83 (2013).
15. Baillie, R. T., Booth, G. G., Tse, Y. & Zabotina, T. Price discovery and common factor models. *Journal of Financial Markets* **5**, 309–321 (2002).
16. Stock, J. H. & Watson, M. W. Testing for Common Trends. *Journal of the American Statistical Association* **83**, 1097–1107 (1988).
17. Hagströmer, B. & Menkveld, A. J. Information Revelation in Decentralized Markets. *The Journal of Finance* **74**, 2751–2787 (2019).
18. Booth, G. G., So, R. W. & Tse, Y. Price discovery in the German equity index derivatives markets. *Journal of Futures Markets* **19**, 619–643 (1999).
19. Brandvold, M., Molnár, P., Vagstad, K. & Valstad, O. C. A. Price discovery on Bitcoin exchanges. *Journal of International Financial Markets, Institutions and Money* **36**, 18–35 (2015).
20. Zhang, Y.-J. & Wei, Y.-M. The crude oil market and the gold market: Evidence for cointegration, causality and price discovery. *Resources Policy* **35**, 168–177 (2010).
21. Figuerola-Ferretti, I. & Gonzalo, J. Modelling and Measuring Price Discovery in Commodity Markets. *Journal of Econometrics* **158**, 95–107 (2010).
22. Lehmann, B. N. Some Desiderata for the Measurement of Price Discovery Across Markets. *Journal of Financial Markets* **5**, 259–276 (2002).
23. Schreiber, T. Measuring Information Transfer. *Physical Review Letters* **85**, 461 (2000).
24. Shannon, C. E. A mathematical theory of communication. *The Bell System Technical Journal* **27**, 379–423 (1948).
25. Wibral, M. *et al.* Revisiting Wiener's principle of causality—interaction-delay reconstruction using transfer entropy and multivariate analysis on delay-weighted graphs. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 3676–3679 (IEEE, 2012).

26. Bossomaier, T., Barnett, L., Harré, M. & Lizier, J. *An Introduction to Transfer Entropy: Information Flow in Complex Systems* (Springer International Publishing, 2016).

27. Reserve Bank of Australia. Statement on Monetary Policy. https://www.rba.gov.au/publications/smp/2019/feb/pdf/statement-on-monetary-policy-2019-02.pdf (2019).

28. Wehrli, A. & Sornette, D. Classification of Flash Crashes Using the Hawkes(p,q) Framework. *Quantitative Finance* **22**, 213–240 (2022).

29. Han, M. F. & Westelius, M. N. J. *Anatomy of Sudden Yen Appreciations* (International Monetary Fund, 2019).

30. Bank, E. C. Asset purchase programmes. https://www.ecb.europa.eu/mopo/implement/app/html/index.en.html (2023).

31. Ricketts, L. R. *et al.* Quantitative easing explained. *Liber8 Econ. Inf. Newsl.* (2011).

32. of England, B. Quantitative easing. https://www.bankofengland.co.uk/monetary-policy/quantitative-easing.

33. Fischer, A. M. & Ranaldo, A. Does fomc news increase global fx trading? *Journal of Banking and Finance* **35**, 2965–2973 (2011).

34. Mueller, P., Tahbaz-Salehi, A. & Vedolin, A. Exchange rates and monetary policy uncertainty. *The Journal of Finance* **72**, 1213–1252 (2017).

35. Bundesbank, D. The Eurosystem's bond purchases and the exchange rate of the euro. *Monthly Report, January* (2017).

36. Bank, E. C. Monetary policy decisions. https://www.ecb.europa.eu/press/pr/date/2020/html/ecb.mp200312~8d3aec3ff2.en.html (2020).

37. Bank, E. C. Monetary policy in a pandemic: Ensuring favourable financing conditions. https://www.ecb.europa.eu/press/key/date/2020/html/ecb.sp201126~c5c1036327.en.html (2020).

38. Janczewski, A., Anagnostou, I. & Kandhai, D. Supplementary material for "Understanding Information Flow and Price Discovery in the Decentralized Foreign Exchange Market: An Information-Theoretic Approach". (2023).

39. Kraskov, A., Stögbauer, H. & Grassberger, P. Estimating Mutual Information. *Physical Review E* **69**, 066138 (2004).

40. Lizier, J. T. JIDT: An information-theoretic toolkit for studying the dynamics of complex systems. *Frontiers in Robotics and AI* **1**, 11 (2014).

41. Vejmelka, M. & Paluš, M. Inferring the directionality of coupling with conditional mutual information. *Physical Review E* **77**, 026214 (2008).

42. Lizier, J. Is KSG estimator deterministic? - [Java Information Dynamics Toolkit (JIDT) discussion] (2015).

43. Michalowicz, J. V., Nichols, J. M. & Bucholtz, F. *Handbook of Differential Entropy* (Crc Press, 2013).

44. Frenzel, S. & Pompe, B. Partial Mutual Information for Coupling Analysis of Multivariate Time Series. *Physical Review Letters* **99**, 204101 (2007).

45. Wibral, M., Vicente, R. & Lizier, J. T. *Directed Information Measures in Neuroscience* (Springer, 2014).

46. Runge, J. Detecting and Quantifying Causality from Time Series of Complex Systems. (2014).

47. Meyer, P. E., Kontos, K., Lafitte, F. & Bontempi, G. Information-Theoretic Inference of Large Transcriptional Regulatory Networks. *EURASIP J. on Bioinforma. Syst. Biol.* **2007**, 1–9 (2007).

48. Ragwitz, M. & Kantz, H. Markov models from data by simple nonlinear time series predictors in delay embedding spaces. *Physical Review E* **65**, 056201 (2002).

49. Lindner, M., Vicente, R., Priesemann, V. & Wibral, M. TRENTOOL: A Matlab open source toolbox to analyse information flow in time series data with transfer entropy. *BMC Neuroscience* **12**, 1–22 (2011).

50. Wibral, M. *et al.* Measuring Information-Transfer Delays. *PloS One* .

51. Virtanen, P. *et al.* SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **17**, 261–272, DOI: 10.1038/s41592-019-0686-2 (2020).

52. Lizier, J. T., Heinzle, J., Horstmann, A., Haynes, J.-D. & Prokopenko, M. Multivariate information-theoretic measures reveal directed information structure and task relevant changes in fMRI connectivity. *Journal of Computational Neuroscience* **30**, 85–107 (2011).

53. Kaiser, A. & Schreiber, T. Information transfer in continuous processes. *Physica D: Nonlinear Phenomena* **166**, 43–62 (2002).

Acknowledgements

We thank Dr. Rick Quax for his insightful feedback on the preliminary version of this manuscript, particularly regarding the information-theoretic inference algorithm.

Author contributions statement

A.J., I.A and D.K. conceived the project; A.J., I.A. gathered the data, A.J. analyzed the data; A.J., I.A and D.K. analyzed the results; A.J., I.A. and D.K. wrote the paper. The authors declare no competing interests. The opinions expressed in this work are solely those of the authors and do not represent in any way those of their current and past employers.

Additional information

Competing interests: The authors declare that they have no competing interests, financial or non-financial, that could influence or be perceived to influence the presented research. All authors have actively contributed to the work and have approved the final version of the manuscript.