

Dear Author,

Here are the proofs of your article.

- You can submit your corrections **online**, via **e-mail** or by **fax**.
- For **online** submission please insert your corrections in the online correction form. Always indicate the line number to which the correction refers.
- You can also insert your corrections in the proof PDF and **email** the annotated PDF.
- For fax submission, please ensure that your corrections are clearly legible. Use a fine black pen and write the correction in the margin, not too close to the edge of the page.
- Remember to note the **journal title**, **article number**, and **your name** when sending your response via e-mail or fax.
- **Check** the metadata sheet to make sure that the header information, especially author names and the corresponding affiliations are correctly shown.
- **Check** the questions that may have arisen during copy editing and insert your answers/ corrections.
- **Check** that the text is complete and that all figures, tables and their legends are included. Also check the accuracy of special characters, equations, and electronic supplementary material if applicable. If necessary refer to the *Edited manuscript*.
- The publication of inaccurate data such as dosages and units can have serious consequences. Please take particular care that all such details are correct.
- Please **do not** make changes that involve only matters of style. We have generally introduced forms that follow the journal's style. Substantial changes in content, e.g., new results, corrected values, title and authorship are not allowed without the approval of the responsible editor. In such a case, please contact the Editorial Office and return his/her consent together with the proof.
- If we do not receive your corrections **within 48 hours**, we will send you a reminder.
- Your article will be published **Online First** approximately one week after receipt of your corrected proofs. This is the **official first publication** citable with the DOI. **Further changes are, therefore, not possible.**
- The **printed version** will follow in a forthcoming issue.

Please note

After online publication, subscribers (personal/institutional) to this journal will have access to the complete article via the DOI using the URL: [http://dx.doi.org/\[DOI\]](http://dx.doi.org/[DOI]).

If you would like to know when your article has been published online, take advantage of our free alert service. For registration and further information go to: <http://www.link.springer.com>.

Due to the electronic nature of the procedure, the manuscript and the original figures will only be returned to you on special request. When you return your corrections, please inform us if you would like to have these documents returned.

Metadata of the article that will be visualized in OnlineFirst

ArticleTitle	Price response functions and spread impact in correlated financial markets	
Article Sub-Title		
Article CopyRight	The Author(s) (This will be the copyright line in the final PDF)	
Journal Name	The European Physical Journal B	
Corresponding Author	Family Name	Henao-Londono
	Particle	
	Given Name	Juan C.
	Suffix	
	Division	Fakultät für Physik
	Organization	Universität Duisburg-Essen
	Address	Lotharstraße 1, 47048, Duisburg, Germany
	Phone	
	Fax	
	Email	juan.henao-londono@uni-due.de
	URL	
	ORCID	
Author	Family Name	Krause
	Particle	
	Given Name	Sebastian M.
	Suffix	
	Division	Fakultät für Physik
	Organization	Universität Duisburg-Essen
	Address	Lotharstraße 1, 47048, Duisburg, Germany
	Phone	
	Fax	
	Email	
	URL	
	ORCID	
Author	Family Name	Guhr
	Particle	
	Given Name	Thomas
	Suffix	
	Division	Fakultät für Physik
	Organization	Universität Duisburg-Essen
	Address	Lotharstraße 1, 47048, Duisburg, Germany
	Phone	
	Fax	
	Email	
	URL	

ORCID

Schedule	Received	3 November 2020
	Revised	
	Accepted	4 March 2021
Abstract	Recent research on the response of stock prices to trading activity revealed long-lasting effects, even across stocks of different companies. These results imply non-Markovian effects in price formation and when trading many stocks at the same time, in particular trading costs and price correlations. How the price response is measured depends on data set and research focus. However, it is important to clarify how the details of the price response definition modify the results. Here, we evaluate different price response implementations for the Trades and Quotes (TAQ) data set from the NASDAQ stock market and find that the results are qualitatively the same for two different definitions of time scale, but the response can vary by up to a factor of two. Furthermore, we show the key importance of the order between trade signs and returns, displaying the changes in the signal strength. Moreover, we confirm the dominating contribution of immediate price response directly after a trade, as we find that delayed responses are suppressed. Finally, we test the impact of the spread in the price response, detecting that large spreads have stronger impact.	
Footnote Information		



Price response functions and spread impact in correlated financial markets

Juan C. Henao-Londono^a, Sebastian M. Krause, and Thomas Guhr

Fakultät für Physik, Universität Duisburg-Essen, Lotharstraße 1, 47048 Duisburg, Germany

Received 3 November 2020 / Accepted 4 March 2021
 © The Author(s) 2021

Abstract. Recent research on the response of stock prices to trading activity revealed long-lasting effects, even across stocks of different companies. These results imply non-Markovian effects in price formation and when trading many stocks at the same time, in particular trading costs and price correlations. How the price response is measured depends on data set and research focus. However, it is important to clarify how the details of the price response definition modify the results. Here, we evaluate different price response implementations for the Trades and Quotes (TAQ) data set from the NASDAQ stock market and find that the results are qualitatively the same for two different definitions of time scale, but the response can vary by up to a factor of two. Furthermore, we show the key importance of the order between trade signs and returns, displaying the changes in the signal strength. Moreover, we confirm the dominating contribution of immediate price response directly after a trade, as we find that delayed responses are suppressed. Finally, we test the impact of the spread in the price response, detecting that large spreads have stronger impact.

1 Introduction

While the definition of complexity varies, it is widely agreed upon that a system is referred to as complex if it, first, consists of a large number of interacting agents or constituents, respectively, second, is non-stationary, *i.e.*, cannot be described by standard equilibrium approaches, and, third, its interactions are typically not captured by microscopic governing equations, rather, by statistical rules. There are many examples, ranging from the traditional physical over biological to social and economic systems. The present interdisciplinary contribution studies financial markets, and it is put forward in the proven physics spirit: first, that every quantity which is used has to be a measurable observable; second, that the results given are quantitative and statistically sound; and third, that the methods are carefully and critically checked and verified. The emphasis of the present contribution is particularly on the third point: we evaluate the methods and show the impact of slight changes on the results. In view of an increasing interest in the analysis of response functions, we feel that this is a rewarding effort. It helps to answer the highly relevant question of the extent to which financial markets deviate from the largely Markovian behavior.

Financial markets use order books to list the number of shares bid or asked at each price. An order book is an electronic list of buy and sell orders for a specific secu-

rity or financial instrument organized by price levels, where agents can place different types of instructions (orders).

In general, the dynamics of the prices follow a random walk. There are two extreme models that can describe this behavior: the Efficient Market Hypothesis (EMH) and the Zero Intelligence Trading (ZIT). The EMH states that all available information is included in the price and price changes can only be the result of unanticipated news, which by definition are totally unpredictable [5, 7, 23, 35]. On the other hand, the ZIT assumes that agents instead of being fully rational, have “zero intelligence” and randomly buy or sell. It is supposed that their actions are interpreted by other agents as potentially containing some information [5, 7, 35, 44]. In both cases, the outcome is the same, and the prices follow a random walk. Reality is somewhere in-between [7, 35], and non-Markovian effects due to strategies or liquidity costs are not contained either.

There are diverse studies focused on the price response [3–5, 7, 13, 15, 16, 19, 22, 23, 29, 30, 32, 41–44]. In our opinion, a critical investigation of definitions and methods and how they affect the results is called for.

Regarding price self-response functions in Refs. [4, 5, 7], Bouchaud et al. found an increase to a maximum followed by a decrease as the time lag grows. In Ref. [16], Gerig found that larger sized transactions have a larger absolute impact than smaller sized transactions but a much smaller relative impact. In Ref. [29], it is found that the impact of small trades on the price is, in relative terms, much larger than that of large trades and the impact of trading on the price is quasi-permanent.

^ae-mail: juan.henao-londono@uni-due.de (corresponding author)

Table 1 Analyzed companies

Company	Symbol	Sector	Quotes ^a	Trades ^b	Spread ^c
Alphabet Inc.	GOOG	Information Technology (IT)	164,489	19,029	0.40\$
Mastercard Inc.	MA	Information Technology (IT)	98,909	6977	0.38\$
CME Group Inc.	CME	Financials (F)	98,188	3032	1.08\$
Goldman Sachs Group Inc.	GS	Financials (F)	160,470	26,227	0.11\$
Transocean Ltd.	RIG	Energy (E)	107,092	11,641	0.12\$
Apache Corp.	APA	Energy (E)	103,074	8889	0.13\$

^a Average number of quotes from 9:40:00 to 15:50:00 New York time during 2008.

^b Average number of trades from 9:40:00 to 15:50:00 New York time during 2008.

^c Average spread from 9:40:00 to 15:50:00 New York time during 2008

For price cross-responses functions, Refs. [3,44] revealed that the diagonal terms are on average larger than the off-diagonal ones by a factor ~ 5 . The response at positive times is roughly constant, what is consistent with the hypothesis of a statistically efficient price. Thus, the current sign does not predict future returns. In Ref. [44], the trends in the cross-responses were found not to depend on whether or not the stock pairs are in the same economic sector or extend over two sectors.

Here, we want to discuss, based on a series of detailed empirical results obtained on trade by trade data, that the variations in the details of the parameters used in the price response definition modify the characteristics of the results. Aspects like time scale, time shift, time lag, and spread used in the price response calculation have an influence on the outcomes. To facilitate the reproduction of our results, the source code for the data analysis is available in Ref. [24].

We delve into the key details needed to compute the price response functions, and explore their corresponding roles. We perform an empirical study in different time scales. We show that the order between the trade signs and the returns have a key importance in the price response signal. We split the time lag to understand the contribution of the immediate returns and the late returns. Finally, we shed light on the spread impact in the response functions for single stocks.

The paper is organized as follows: in Sect. 2, we present our data set of stocks. We then analyze the definition of the price response functions and describe the physical and trade time in Sect. 3. We implement different price responses for several stocks and pairs of stocks in Sect. 4. In Sect. 5, we show how the relative position between trade signs and returns has a huge influence in the results of the computation of the response functions. In Sect. 6 we explain in detail how the time lag τ behaves in the response functions. Finally, in Sect. 7, we analyze the spread impact in the price response functions. Our conclusions follow in Sect. 8.

2 Data set

Modern financial markets are organized as a double continuous auctions. Agents can place different types of orders to buy or to sell a given number of shares, roughly categorized as market orders and limit orders.

In this study, we analyzed trades and quotes (TAQ) data from the NASDAQ stock market. We selected NASDAQ, because it is an electronic exchange where stocks are traded through an automated network of computers instead of a trading floor, which makes trading more efficient, fast, and accurate. Furthermore, NASDAQ is the second largest stock exchange based on market capitalization in the world.

In the TAQ data set, there are two data files for each stock. One gives the list of all successive quotes. Thus, we have the best bid price, best ask price, available volume, and the time stamp accurate to the second. The other data file is the list of all successive trades, with the traded price, traded volume and time stamp accurate to the second. Despite the 1 s accuracy of the time stamps, in both files, more than one quote or trade may be recorded in the same second.

To analyze the response functions across different stocks in Sects. 4, 5, and 6, we select the six companies with the largest average market capitalization (AMC) in three economic sectors of the S&P index in 2008. Table 1 shows the companies analyzed with their corresponding symbol and sector, and three average values for a year.

To analyze the spread impact in response functions (Sect. 7), we select 524 stocks in the NASDAQ stock market for the year 2008. The selected stocks are listed in Appendix A.

This contribution specifically addresses methodical aspects related to the price response functions and spread impact in correlated financial markets. We carefully check and verify the methods used to evaluate response functions, spread impact, and related quantities. This is important as such observables quantify the deviation from the largely Markovian behavior of financial markets. We chose the year 2008 to clarify these methodical aspects. We plan to extend our results in a future study to different years for a comparison of price response functions and spread impact.

To avoid overnight effects and any artifact due to the opening and closing of the market, we systematically discard the first 10 and the last 10 min of trading in a given day [7,13,19,44]. Therefore, we only consider trades of the same day from 9:40:00 to 15:50:00 New York local time. We will refer to this interval of time as the “market time”. The year 2008 corresponds to 253 business days.

3 Price response function definitions

In Sect. 3.1, we introduce the fundamental quantities used in the price response definitions. In Sect. 3.2, we describe the physical time scale and the trade time scale. We introduce the price response functions used in literature in Sect. 3.3.

3.1 Key concepts

Market orders are executed at the best available buy or sell price, limit orders set a maximum purchase price for a buy order, or a minimum sale price for a sell order. If the limit price is not matched, the order will not be carried out [13, 14, 21, 35]. Limit orders often fail to result in an immediate transaction, and are stored in a queue called the limit order book [8, 14, 21, 29]. The order book also identifies the market participants behind the buy and sell orders, although some choose to remain anonymous. The order book is visible for all traders and its main purpose is to ensure that all traders have the same information on what is offered on the market. The order book is the ultimate microscopic level of description of financial markets.

Buy limit orders are called “bids”, and sell limit orders are called “asks”. At any given time, there is a best (lowest) offer to sell with price $a(t)$, and a best (highest) bid to buy with price $b(t)$ [5, 9, 11, 29, 35]. The price gap between them is called the spread $s(t) = a(t) - b(t)$ [5–7, 9, 13, 20, 35]. Spreads are significantly positively related to price and significantly negatively related to trading volume. Companies with more liquidity tend to have lower spreads [1, 2, 9, 17].

The average of the best ask and the best bid is the midpoint price, which is defined as: [5, 7, 13, 20, 29, 35]

$$m(t) = \frac{a(t) + b(t)}{2}. \quad (1)$$

As the midpoint price depends on the quotes, it changes if the quotes change. The midpoint price grows if the best ask or the best bid grow. On the other hand, the midpoint price decreases if the best ask or the best bid decrease.

Price changes are typically characterized as returns. If one denotes $S(t)$ the price of an asset at time t , the return $r^{(g)}(t, \tau)$, at time t and time lag τ is simply the relative variation of the price from t to $t + \tau$ [5, 10, 25–27, 34]:

$$r^{(g)}(t, \tau) = \frac{S(t + \tau) - S(t)}{S(t)}. \quad (2)$$

It is also common to define the returns as [3, 5, 10, 12, 13, 15, 16, 19, 31, 33]:

$$r^{(l)}(t, \tau) = \ln S(t + \tau) - \ln S(t) = \ln \frac{S(t + \tau)}{S(t)}. \quad (3)$$

Equations (2) and (3) agree if τ is small enough [5, 10].

At longer timescales, midpoint prices and transaction prices rarely differ by more than half the spread. The midpoint price is more convenient to study, because it avoids problems associated with the tendency of transaction prices to bounce back and forth between the best bid and ask [13].

We define the returns via the midpoint price as:

$$r(t, \tau) = \frac{m(t + \tau) - m(t)}{m(t)}. \quad (4)$$

The distribution of returns is strongly non-Gaussian and its shape continuously depends on the return period τ . Small τ values have fat tails return distributions [5]. The trade signs are defined for general cases as:

$$\varepsilon(t) = \text{sign}(S(t) - m(t - \delta)), \quad (5)$$

where δ is a positive time increment. Hence, we have:

$$\varepsilon(t) = \begin{cases} +1, & \text{If } S(t) \text{ is higher than the last } m(t) \\ -1, & \text{If } S(t) \text{ is lower than the last } m(t); \end{cases} \quad (6)$$

$\varepsilon(t) = +1$ indicates that the trade was triggered by a market order to buy and a trade triggered by a market order to sell yields $\varepsilon(t) = -1$ [5, 7, 19, 30, 38].

It is well known that the series of the trade signs on a given stock exhibit large autocorrelation. A very plausible explanation of this phenomenon relies on the execution strategies of some major brokers on given markets. These brokers have large transactions to execute on the account of some clients. To avoid market making movements because of an inconsiderable large order, they tend to split large orders into small ones [10].

3.2 Time definition

A direct comparison between the trade time scale and the physical time scale is not possible. To compare them directly, we need to assume whether the midpoint price or the trade signs are on the same scale. We assume the midpoint prices in the trade time scale to be the same as the midpoint prices in physical time scale. Therefore, we have the time lag for both computations in seconds. This approximation allows us to directly compare both scales to have an idea of the difference and similarities they have. In the other sections, as we are not directly comparing the time scales, the corresponding quantities of each time scale are not mixed. Thus, physical time scale is measured in seconds and trade time scale is measured in trades.

Due to the nature of the data, there are several options to define time for analyzing data.

In general, the time series are indexed in calendar time (hours, minutes, seconds, and milliseconds). Moreover, tick-by-tick data available on financial markets all over the world are time stamped up to the millisecond, but the order of magnitude of the guaranteed precision

is much larger, usually 1 s or a few hundreds of milliseconds [6, 10]. Several papers are used different time definitions (calendar time, physical time, event time, trade time, and tick time) [10, 18, 37]. The TAQ data used in the analysis have the characteristic that the trades and quotes cannot be directly related due to the time stamp resolution of only 1 s [44]. Hence, it is impossible to match each trade with the directly preceding quote. However, using a classification for the trade signs, we can compute trade signs in two scales: trade time scale and physical time scale.

The trade time scale is increased by one unit each time a transaction happens. The advantage of this count is that limit orders far away in the order book do not increase the time by one unit. The main outcome of trade time scale is its “smoothing” of data and the aggregational normality [10].

The physical time scale is increased by one unit each time a second passes. This means that computing the responses in this scale involves sampling [18, 44], which has to be done carefully when dealing for example with several stocks with different liquidity. This sampling is made in the trade signs and in the midpoint prices.

Facing the impossibility to relate midpoint prices and trade signs with the TAQ data in trade time scale, we will use the midpoint price of the previous second with all the trade signs of the current second. This will be our definition of trade time scale analysis for the response function analysis.

For physical time scale, as we can sampling, we relate the unique value of midpoint price of a previous second with the unique trade sign value of the current second.

Trade time scale

We use the trade sign classification in trade time scale proposed in Ref. [44] and used in Refs. [39, 40, 43] that reads:

$$\varepsilon^{(t)}(t, n) = \begin{cases} \text{sgn}(S(t, n) - S(t, n-1)), & \text{if } S(t, n) \neq S(t, n-1) \\ \varepsilon^{(t)}(t, n-1), & \text{otherwise} \end{cases} \quad (7)$$

$\varepsilon^{(t)}(t, n) = +1$ implies a trade triggered by a market order to buy, and a value $\varepsilon^{(t)}(t, n) = -1$ indicates a trade triggered by a market order to sell.

In the second case of Eq. (7), if two consecutive trades with the same trading direction did not exhaust all the available volume at the best quote, the trades would have the same price, and in consequence, they will have the same trade sign.

With this classification, we obtain trade signs for every single trade in the data set. According to Ref. [44], the average accuracy of the classification is 85% for the trade time scale.

The TAQ time step is 1 s, and as it is impossible to find the correspondences between trades and midpoint prices values inside a second step, We used the last midpoint price of every second as the representative

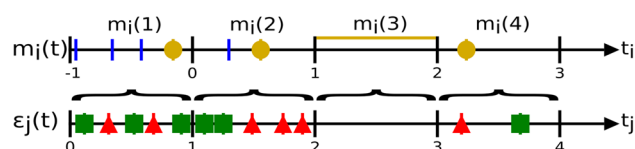


Fig. 1 Sketch of data processing for trade time scale. In the midpoint price time line, the vertical lines represent the change in price of the quotes and the circles represent the last price change in a quote in a second. In the trade signs time line, the squares represent the buy market orders and the triangles represent the sell market orders. The midpoint price time line and the trade sign time line are shifted in 1 s

value of each second. This introduce an apparent shift between trade signs and returns. In fact, we set the last midpoint price from the previous second as the first midpoint price of the current second [44].

As we know the second in which the trades were made, we can relate the trade signs and the midpoint prices, as shown in Fig. 1. For the trade time scale, there are, in general, several midpoint prices in a second. For each second, we select the last midpoint price value, and we relate it to the next second trades. In Fig. 1, the last midpoint price (circle) between the second -1 and 0 is related to all the trades (squares and triangles) in the second 0 to 1 , and so on. In the seconds when the quotes do not change, the value of the previous second (vertical line over the physical time interval) is used. Thus, all the seconds in the open market time have a midpoint price value, and in consequence return values. We assume that as long as no changes occurred in the quotes, the midpoint price remains the same as in the previous second.

The methodology described is an approximation to compute the response in the trade time scale. A drawback in the computation could come from the fact that the return of a given second is composed by the contribution of small returns corresponding to each change in the midpoint price during a second. As we are assuming only one value for the returns in each second, we consider all the returns in 1 s interval to be positive or negative with the same magnitude, which could not be the case. This could increase or decrease the response signal at the end of the computation.

Figure 2 illustrate this point. Suppose there are three different midpoint prices in 1 s interval, and thus, three different returns for these three midpoint price values. Furthermore, suppose that the volumes of limit orders with the corresponding midpoint prices are the same in the bid and in the ask (the returns have the same magnitude). In the case of the top left (top right) sketch, all the changes are due to the rise (decrease) of the midpoint price, that means, consumption of the best ask (bid), so all the contributions of the individual returns in the second are positive (negative), and in consequence, the net return is positive (negative). In the case of the bottom, the changes are due to a combination of increase and decrease of the midpoint price, so in the end, the individual returns sum up to a net return, which can be positive or negative, depending of the type

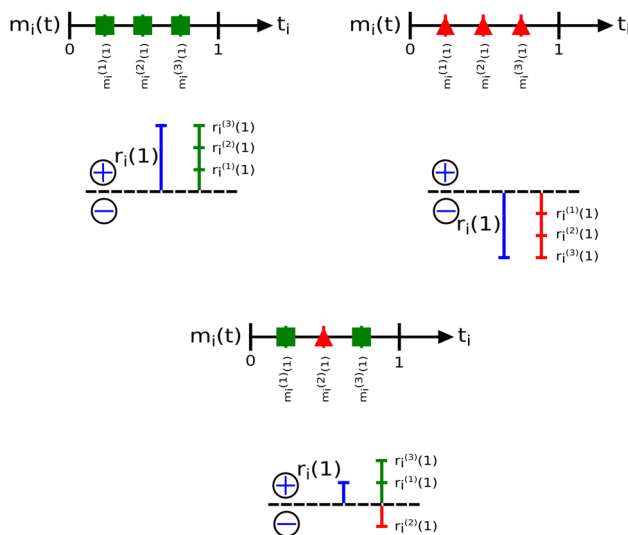


Fig. 2 Sketch of the return contributions from every midpoint price change in a second. The squares represent the rise of the price of the midpoint price and the triangles represent the decrease of the price of the midpoint price. We illustrate three cases: (top left) the changes of the midpoint prices and return are due to the rise of the prices, (top right) the changes of the midpoint prices and return are due to the decrease of the prices, and (bottom) the changes of the midpoint prices and return are due to a combination of rise and decrease of the prices. The blue vertical line represents the net return in each case

of midpoint price values in the interval. Thus, in this case, we are assuming in the end that all the returns were positive or negative, which probably was not the case, and in consequence, will increase or decrease the real value of the net return.

In all cases, we choose the last change in the midpoint price in a second interval as described before in Fig. 1. We use this method knowing that the variation in 1 s of the midpoint price is not large (in average, the last midpoint price of a second differ with the average midpoint of that second in 0.007%), and hence, it can give us representative information on the response functions.

Physical time scale

We use the trade sign definition in physical time scale proposed in Ref. [44] and used in Refs. [39,43], that depends on the classification in Eq. (7) and reads:

$$\varepsilon^{(p)}(t) = \begin{cases} \text{sgn} \left(\sum_{n=1}^{N(t)} \varepsilon^{(t)}(t, n) \right), & \text{If } N(t) > 0 \\ 0, & \text{If } N(t) = 0 \end{cases}, \quad (8)$$

where $N(t)$ is the number of trades in a second interval. $\varepsilon^{(p)}(t) = +1$ implies that the majority of trades in second t were triggered by a market order to buy, and a value $\varepsilon^{(p)}(t) = -1$ indicates a majority of sell market

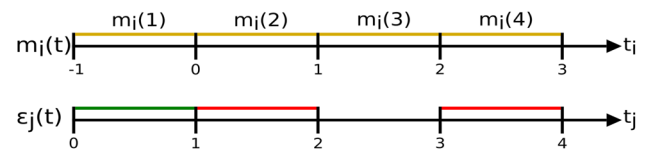


Fig. 3 Sketch of data processing for physical time scale. In the midpoint price time line, the horizontal lines between seconds represent the midpoint prices. In the trade sign time line, the horizontal lines between seconds represent the trade sign values. The midpoint price time line and the trade sign time line are shifted in 1 s

orders. In this definition, they are two ways to obtain $\varepsilon^{(p)}(t) = 0$. One way is that in a particular second, there are no trades, and then no trade sign. The other way is that the addition of the trade signs (+1 and -1) in a second be equal to zero. In this case, there is a balance of buy and sell market orders.

Market orders show opposite trade directions to limit order executed simultaneously. An executed sell limit order corresponds to a buyer-initiated market order. An executed buy limit order corresponds to a seller-initiated market order.

As the trade time scale, on the physical time scale, we use the same strategy to obtain the midpoint price for every second, so all the seconds in the open market time have a midpoint price value. Even if there is no change of quotes in a second, it still has a midpoint price value and return value.

In this case, we do not compare every single trade sign in a second, but the net trade sign obtained for every second with the definition, see Eq. (8). This can be seen in Fig. 3, where we related the midpoint price of the previous second with the trade sign of the current second. According to Ref. [44], this definition has an average accuracy up to 82% in the physical time scale.

3.3 Response function definitions

The response function measures price changes resulting from execution of market orders. In Refs. [4,5,7], Bouchaud et al. use a self-response function that only depends on the time lag τ . This function measures how much, on average, the price moves up (down) at time τ conditioned to a buy (sell) order at time zero. They found for France Telecom that the response function increases by a factor 2 between $\tau = 1$ and $\tau \approx 1000$ trades, before decreasing back. For larger τ , the response function decreases, and even becomes negative beyond $\tau \approx 5000$. However, in some cases, the maximum is not observed, and rather, the price response function keeps increasing mildly [5].

In Ref. [16], the price impact function is defined as the average price response due to a transaction as a function of the transaction's volume. Empirically, the function is highly concave [16]. The curvature of the price impact function is entirely due to the probability that a transaction causes a nonzero impact. The larger the size of the transaction, the larger the probability.

In Ref. [29], they found that the response function for three French stocks first increases from $\tau = 10s$ to a few hundred seconds, and then appears to decrease back to a finite value.

In Ref. [3], they defined a response function who measures the average price change of a contract i at time $t + \tau$, after experiencing a sign imbalance in contract j at time t . In this work, τ is used in units of five minutes.

In Ref. [28], local Gaussian correlations in financial and commodity markets were analyzed. These correlations distinguish between positive and negative local dependence. In this work, a breakpoint in August 2008 is found. The increased comovements between commodities and financial markets are more critical after the breakpoint, not only between commodities and the S&P500 and 10-year Treasuries, but also within commodities themselves.

In later works [19, 44], Grimm et al. and Wang et al. use the logarithmic return for stock i and time lag τ , defined via the midpoint price $m_i(t)$ to define a cross-response function. The response function measures how a buy or sell order at time t influences on average the price at a later time $t + \tau$. The physical time scale was chosen, since the trades in different stocks are not synchronous (TAQ data). They found that in all cases, an increase to a maximum is followed by a decrease. The trend is eventually reversed.

Finally, in Ref. [41], Wang et al. define the response function on a trade time scale (Totalview data), as the interest is to analyze the immediate responses. Here, the time lag τ is restricted to one trade, such that the price response quantifies the price impact of a single trade.

4 Price response function implementations

The main objective of this work is to analyze the price response functions. In general, we define the self- and cross-response functions in a correlated financial market as:

$$R_{ij}^{(\text{scale})}(\tau) = \left\langle r_i^{(\text{scale})}(t-1, \tau) \varepsilon_j^{(\text{scale})}(t) \right\rangle_{\text{average}}, \quad (9)$$

where the index i and j correspond to stocks in the market, $r_i^{(\text{scale})}$ is the return of the stock i in a time lag τ in the corresponding scale, and $\varepsilon_j^{(\text{scale})}$ is the trade sign of the stock j in the corresponding scale. The superscript *scale* refers to the time scale used, whether physical time scale (*scale* = p) or trade time scale (*scale* = t). Finally, The subscript *average* refers to the way to average the price response, whether relative to the physical time scale (*average* = P) or relative to the trade time scale (*average* = T).

We use the returns and the trade signs to define three response functions: trade time scale response, physical time scale response, and activity response.

To compare the three response functions, we define the following quantities:

$$E_{j,d}(t) = \sum_{n=1}^{N(t)} \varepsilon_{j,d}^{(t)}(t, n) = \text{sgn}(E_{j,d}(t)) |E_{j,d}(t)| \quad (10)$$

$$\varepsilon_{j,d}^{(p)}(t) = \text{sgn}(E_{j,d}(t)), \quad (11)$$

where the subscript d refers to the days used in the response computation. We use Eq. (10) to make easier the comparison between the results, as all the defined responses use the trade sign term.

In Sect. 4.1, we analyze the responses functions in trade time scale; in Sect. 4.2, we analyze the responses functions in physical time scale; and in Sect. 4.3, we define a response function to analyze the influence of the frequency of trades in a second.

4.1 Response functions on trade time scale

We define the self- and cross-response functions in trade time scale, using the trade signs in trade time scale. For the returns, we select the last midpoint price of every second and compute them. We use this strategy with the TAQ data set considering that the price response in trade time scale cannot be directly compared with the price response in physical time scale. In this case, we relate each trade sign in 1 s with the midpoint price of the previous second. Then, to compute the returns, instead of using trades as the time lag (it would make no sense as all the midpoint price are the same in 1 s), we use seconds. Thus, we force the response in trade time scale to have a physical time lag, and then, be able to compare with the physical time scale response. This approximation is feasible considering the discussion in Sect. 3. The price response function in trade time scale is defined as:

$$R_{ij}^{(t)}(\tau) = \left\langle r_i^{(p)}(t-1, \tau) \varepsilon_j^{(t)}(t, n) \right\rangle_T, \quad (12)$$

where the superscript t refers to the trade time scale. We explicitly calculate the average in Eq. (12):

$$\begin{aligned} R_{ij}^{(t)}(\tau) &= \frac{1}{\sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} N_{j,d}(t)} \\ &\quad \sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} \sum_{n=1}^{N(t)} r_{i,d}^{(p)}(t-1, \tau) \varepsilon_{j,d}^{(t)}(t, n) \\ &= \sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} r_{i,d}^{(p)}(t-1, \tau) \frac{\sum_{n=1}^{N(t)} \varepsilon_{j,d}^{(t)}(t, n)}{\sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} N_{j,d}(t)} \\ &= \sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} r_{i,d}^{(p)}(t-1, \tau) \text{sgn}(E_{j,d}(t)) w_{j,d}^{(t)}(t), \end{aligned} \quad (13)$$

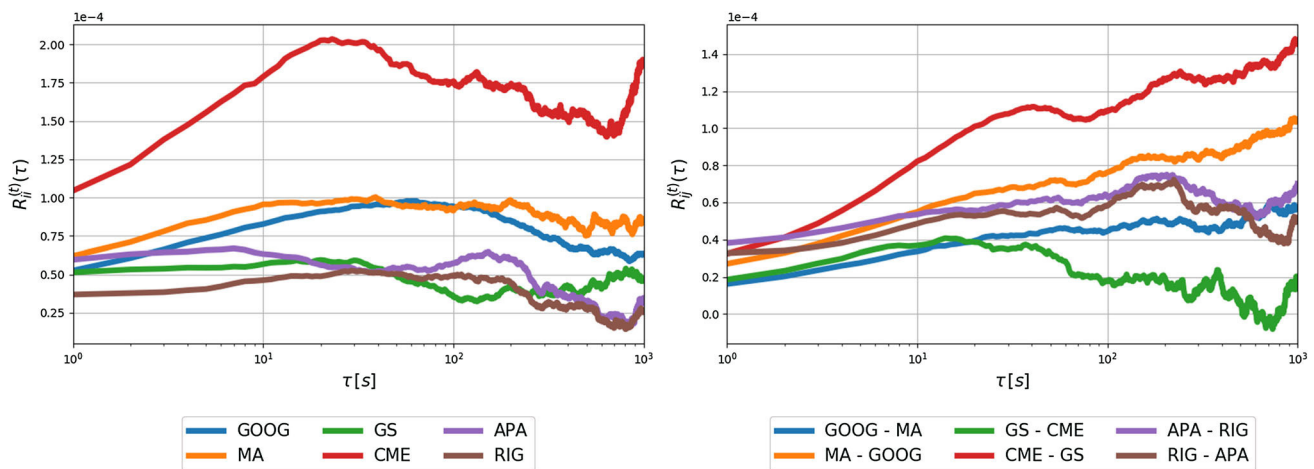


Fig. 4 Self- and cross-response functions $R_{ij}^{(t)}(\tau)$ in 2008 versus time lag τ on a logarithmic scale in trade time scale. Self-response functions (left) of individual stocks and cross-response functions (right) of stock pairs from the same economic sector

where:

$$w_{j,d}^{(t)}(t) = \frac{|E_{j,d}(t)|}{\sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} N_{j,d}(t)}, \quad (15)$$

is a weight function that depends on the normalization of the response.

To compute the response functions on trade time scale, we used all the trade signs during a day in market time. As we cannot associate an individual midpoint price with their corresponding trade signs, all the trade signs in 1 s are associated with the midpoint price of the previous second. As τ depends on the midpoint price, even if we are using trade signs in trade time scale, the value of τ is in seconds.

The results of Fig. 4 show the self-responses of the six stocks used in the analysis and the cross-responses for pairs of stocks representing three different economic sectors.

The self-response functions increase to a maximum and then slowly decrease. In some stocks, this behavior is more pronounced than in others. For our selected tickers, a time lag of $\tau = 10^3$ s is enough to see an increase to a maximum followed by a decrease. Thus, the trend in the self-response functions is eventually reversed. On the other hand, the cross-response functions have smaller signal strength than the self-response functions. For our cross-response functions of stocks in the same sectors, some couples exhibit the increase-decrease behavior inside a time lag of $\tau = 10^3$ s. Other couples seems to need a larger time lag to reach the decrease behavior.

4.2 Response functions on physical time scale

One important detail to compute the market response in physical time scale is to define how the averaging of the function will be made, because the response functions highly differ when we include or exclude $\varepsilon_j^{(p)}(t) =$

0 [44]. The price responses including $\varepsilon_j^{(p)}(t) = 0$ are weaker than the excluding ones due to the omission of direct influence of the lack of trades. However, either including or excluding $\varepsilon_j^{(p)}(t) = 0$ does not change the trend of price reversion versus the time lag, but it does affect the response function strength [43]. For a deeper analysis of the influence of the term $\varepsilon_j^{(p)}(t) = 0$, we suggest to check Refs. [43, 44]. We will only take in account the response functions excluding $\varepsilon_j^{(p)}(t) = 0$.

We define the self- and cross-response functions in physical time scale, using the trade signs and the returns in physical time scale. The price response function on physical time scale is defined as:

$$R_{ij}^{(p)}(\tau) = \left\langle r_i^{(p)}(t-1, \tau) \varepsilon_j^{(p)}(t) \right\rangle_P, \quad (16)$$

where the superscript p refers to the physical time scale. The explicit expression corresponding to Eq. (16) reads:

$$\begin{aligned} R_{ij}^{(p)}(\tau) &= \frac{1}{\sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} \eta[\varepsilon_{j,d}^{(p)}(t)]} \\ &\quad \sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} r_{i,d}^{(p)}(t-1, \tau) \varepsilon_{j,d}^{(p)}(t) \eta[\varepsilon_{j,d}^{(p)}(t)] \\ &= \sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} r_{i,d}^{(p)}(t-1, \tau) \frac{\varepsilon_{j,d}^{(p)}(t) \eta[\varepsilon_{j,d}^{(p)}(t)]}{\sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} \eta[\varepsilon_{j,d}^{(p)}(t)]} \\ &= \sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} r_{i,d}^{(p)}(t-1, \tau) \operatorname{sgn}(E_{j,d}(t)) w_{j,d}^{(p)}(t), \end{aligned} \quad (17)$$

where:

$$\eta(x) = \begin{cases} 1, & \text{If } x \neq 0 \\ 0, & \text{otherwise} \end{cases}, \quad (19)$$

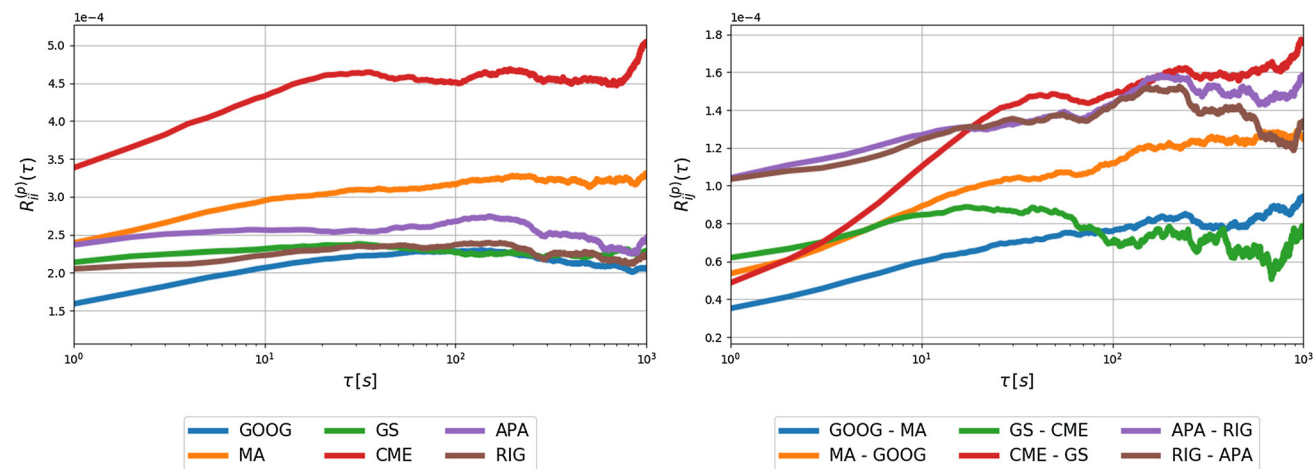


Fig. 5 Self- and cross-response functions $R_{ij}^{(p)}(\tau)$ excluding $\varepsilon_j^{(p)}(t) = 0$ in 2008 versus time lag τ on a logarithmic scale in physical time scale. Self-response functions (left) of individual stocks and cross-response functions (right) of stock pairs from the same economic sector

take only in account the seconds with trades and:

$$w_{j,d}^{(p)}(t) = \frac{\eta [\text{sgn}(E_{j,d}(t))]}{\sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} \eta [\text{sgn}(E_{j,d}(t))]}, \quad (20)$$

is a weight function that depends on the normalization of the response.

The results showed in Fig. 5 are the self- and cross-response functions in physical time scale. For the self-response functions, we can say again that in almost all the cases, an increase to a maximum is followed by a decrease. Thus, the trend in the self- and cross-response is eventually reversed. In the cross-response functions, we have a similar behavior with the previous subsection, where the time lag in some pairs was not enough to see the decrease of the response.

Compared with the response functions in trade time scale, the response functions in physical time scale are stronger.

4.3 Activity response functions on physical time scale

Finally, we define the activity self- and cross-response functions in physical time scale, using the trade signs and the returns in physical time scale. We add a factor $N_{j,d}(t)$ to check the influence of the frequency of trades in a second in the response functions. The activity price response function is defined as:

$$R_{ij}^{(p,a)}(\tau) = \left\langle r_i^{(p)}(t-1, \tau) \varepsilon_j^{(p)}(t) N(t) \right\rangle_P, \quad (21)$$

where the superscript a refers to the activity response function. The corresponding explicit expression reads:

$$R_{ij}^{(p,a)}(\tau) = \frac{1}{\sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} N_{j,d}(t)}$$

$$\begin{aligned} & \sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} r_{i,d}^{(p)}(t-1, \tau) \varepsilon_{j,d}^{(p)}(t) N_{j,d}(t) \quad (22) \\ &= \sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} r_{i,d}^{(p)}(t-1, \tau) \frac{\varepsilon_{j,d}^{(p)}(t) N_{j,d}(t)}{\sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} N_{j,d}(t)} \quad (23) \\ &= \sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} r_{i,d}^{(p)}(t-1, \tau) \text{sgn}(E_{j,d}(t)) w_{j,d}^{(a)}(t), \end{aligned}$$

where:

$$w_{j,d}^{(a)}(t) = \frac{N_{j,d}(t)}{\sum_{d=d_0}^{d_f} \sum_{t=t_0}^{t_f} N_{j,d}(t)}, \quad (24)$$

is a weight function that depends on the normalization of the response.

As $E_{j,d}(t)$ is the sum of +1 and -1 in 1 s and $N_{j,d}(t)$ is the number of trades in a second, $N_{j,d}(t) \geq E_{j,d}(t)$. $N_{j,d}(t) = E_{j,d}(t)$ only when all the trades in a second are buys (+1).

The trade weight $w_{j,d}^{(t)}(t)$ reduces noises. The physical weight $w_{j,d}^{(p)}(t)$ gives every step the same weight, and the activity weight $w_{j,d}^{(a)}(t)$ emphasizes seconds with large activity.

In Fig. 6, we can see how the three responses have approximately the same shape, but the strength of the signal varies depending on the definition. The frequency of trades have a large influence in the responses.

As predicted by the weights, the event response is weaker than the physical response, and the activity response is the strongest response.

We propose a methodology to directly compare price response functions in trade time scale and physical time scale. Additionally, we suggest a new definition to measure the impact of the number of trades in physical time

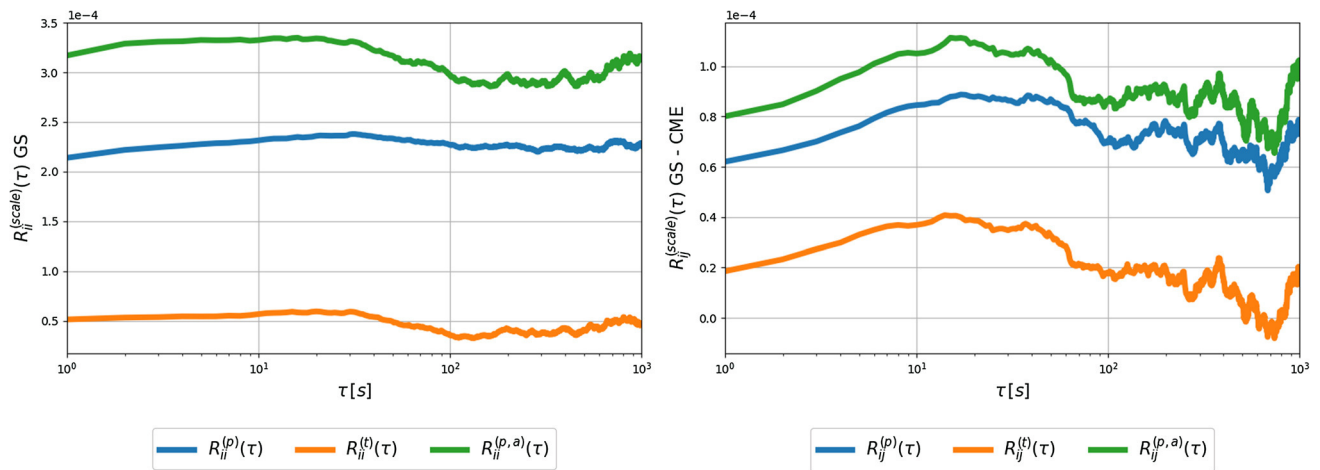


Fig. 6 Self- and cross-response functions $R_{ij}^{(\text{scale})}(\tau)$ excluding $\varepsilon_j^{(p)}(t) = 0$ in 2008 versus time lag τ on a logarithmic scale. Self-response functions (left) of Goldman Sachs Group Inc. stock and cross-response functions (right) of Goldman Sachs Group Inc.–CME Group Inc. stocks

scale. In the three curves in the figure can be seen the increase–decrease behavior of the response functions.

Our results are consistent with the current literature, where the results differ about a factor of two depending on the time scale. We note that the activity response function implementation is only a test and was never defined in previous works. That is why the difference of a factor of two can not be seen in Fig. 6 for $R_{ij}^{(p,a)}(\tau)$.

A special situation occurs when financial and commodity markets are analyzed to find correlations as shown in Ref. [28]. In this analysis, a breakpoint in August 2008 is found, where, before and after the point, the results are qualitatively different. However, our work is only focused in correlated financial markets. Our price response functions, on average, have the same behavior despite the time of the year.

5 Time shift response functions

The relative position between returns and trade signs directly impact the result of the response functions. Shifting the values to the right or to the left either in trade time scale or physical time scale has approximately the same effect.

To test this claim, we used the definition of the response function from Ref. [44] and add a parameter t_s that shifts the position between returns and trade signs. To see the impact of the time shift, we analyzed the stocks showed in Table 1 in the year 2008. We used different time shifts in the response function:

$$R_{ij}^{(\text{scale},s)}(\tau) = \left\langle r_i^{(\text{scale})}(t - t_s, \tau) \varepsilon_j^{(\text{scale})}(t) \right\rangle_{\text{average}}, \quad (25)$$

where the index i and j correspond to stocks in the market, $r_i^{(\text{scale})}$ is the return of the stock i in a time

lag τ with a time shift t_s in the corresponding scale and $\varepsilon_j^{(\text{scale})}$ is the trade sign of the stock j in the corresponding scale. The superscript scale refers to the time scale used, whether physical time scale (scale = p) or trade time scale (scale = t). $R_{ij}^{(\text{scale},s)}$ is the time shift price response function, where the superscript s refers to the time shift. Finally, the subscript *average* refers to the way to average the price response, whether relative to the physical time scale (average = P) or relative to the trade time scale (average = T).

We compute the response functions according to two cases. In one case, we set τ to a constant value and vary t_s , and in the other case, we set t_s to a constant value and vary τ .

In Sect. 5.1, we analyze the influence of the time shift between the trade signs and returns in trade time scale, and in Sect. 5.2, we analyze the influence of the time shift between the trade signs and returns in physical time scale.

5.1 Trade time scale shift response functions

On the trade time scale, we compute the response function:

$$R_{ij}^{(t,s)}(\tau) = \left\langle r_i^{(t)}(t - t_s, \tau) \varepsilon_j^{(t)}(t) \right\rangle_t. \quad (26)$$

In this case for $r_i^{(t)}$, we associate all the trade signs to a return value and create pseudo-midpoint price values in trade time scale. Then, we shift the trade signs and the returns by trades. Hence, the time lag and time shift are in trade time scale.

In Fig. 7, we show the response functions results for fixed τ values, while t_s is variable. In the different τ values of figures, the results are almost the same. The response functions are zero either if the time shift is larger than τ , or if the time shift is smaller than zero.

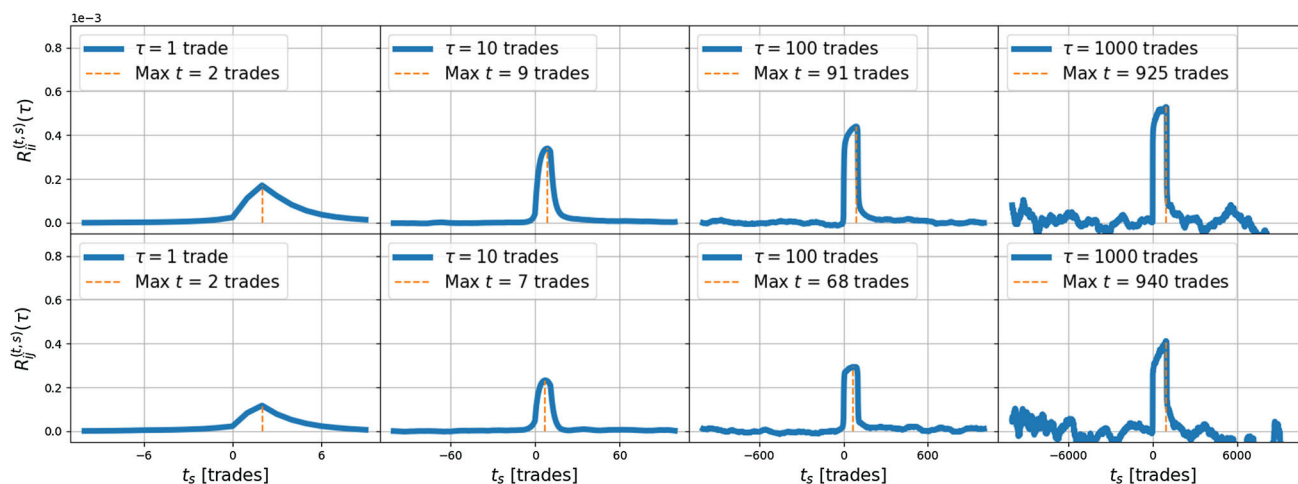


Fig. 7 Self-response functions $R_{ii}^{(t,s)}(\tau)$ in 2008 versus shift for the Transocean Ltd. stock (top) and cross-response functions $R_{ij}^{(t,s)}(\tau)$ in 2008 versus shift for the Transocean Ltd.-Apache Corp. stocks (bottom) in trade time scale

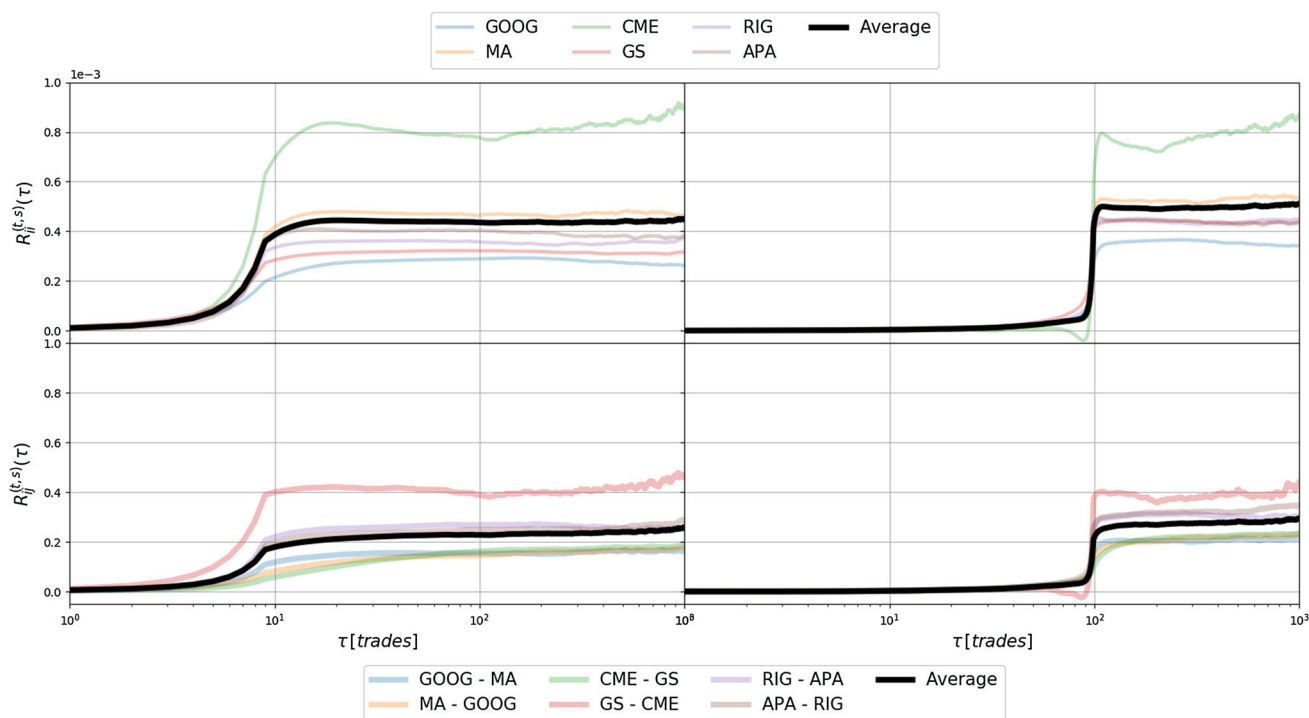


Fig. 8 Self- and cross-response functions $R_{ij}^{(t,s)}(\tau)$ in 2008 versus time lag τ on a logarithmic scale for different shifts in trade time scale. Self-response functions (top) of individual stocks and cross-response functions (bottom) of stocks pairs from the same economic sector. We use time lag values $\tau = 10$ trades (left) and $\tau = 100$ (right)

However, related to the time lag, there is a zone where the signal is different from zero. For values between zero and τ , there is a peak in a position related to τ . The response function grows and decreases relatively fast.

We tested the response function for fixed time shift values, while τ is variable. In Fig. 8, we use a time shift of 10 trades (left) and 100 trades (right). In both, self- and cross-response results are qualitatively the same. It can be seen that the response functions have a zero signal before the time shift. After the returns and trade

signals find their corresponding order, the signals grow. In comparison with the values obtained in Fig. 4, it looks like the response function values with large time shift are stronger. However, this is an effect of the averaging of the functions. As the returns and trade signs are shifted, there are less values to average, and then, the signals are stronger. Anyway, the figure shows the importance of the position order between the trade signs and returns to compute the response function.

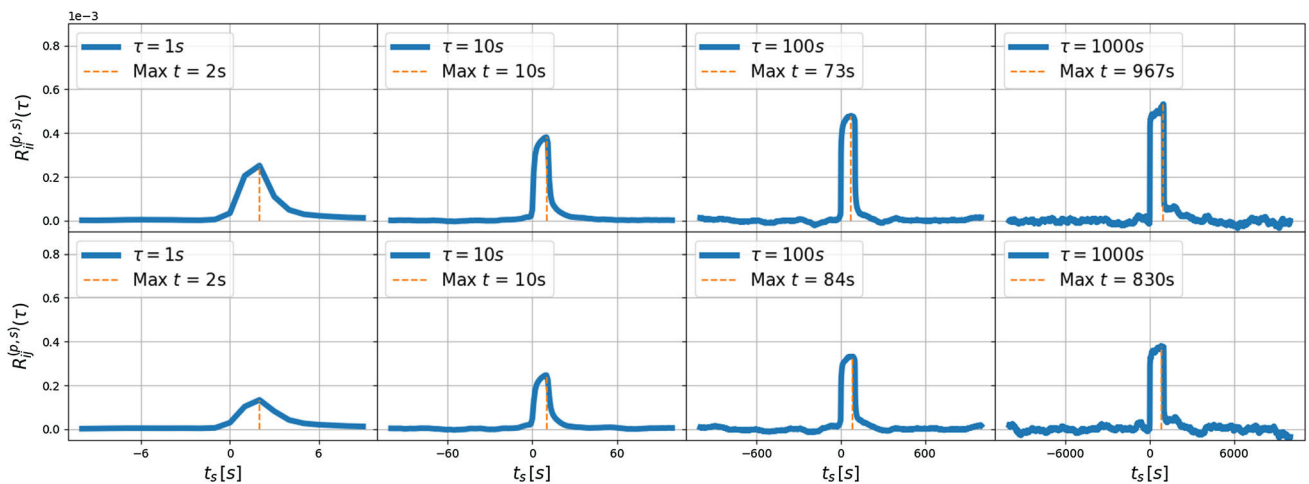


Fig. 9 Self-response functions $R_{ii}^{(p,s)}(\tau)$ excluding $\varepsilon_i^{(p)}(t) = 0$ in 2008 versus shift for the Transocean Ltd. stock (top) and cross-response functions $R_{ij}^{(p,s)}(\tau)$ excluding $\varepsilon_j^{(p)}(t) = 0$ in 2008 versus shift for the Transocean Ltd.-Apache Corp. stocks (bottom) in physical time scale

5.2 Physical time scale shift response functions

In the physical time scale, we compute the response function:

$$R_{ij}^{(p,s)}(\tau) = \left\langle r_i^{(p)}(t - t_s, \tau) \varepsilon_j^{(p)}(t) \right\rangle_t. \quad (27)$$

Similar to the results in Sect. 5.1, Fig. 9 shows the responses functions for fixed τ values, while t_s is variable. Again, the response functions are zero if the time shift is larger than the time lag, or if the time shift is smaller than zero. For every τ value, there is a peak. The peak grows and decays relatively fast. The response signal usually starts to grow in zero or a little bit earlier and grows to a value around to τ . In this zone, the response functions are different to zero.

The results for fixed time shift values and variable time lag ($t_s = 10$ s and $t_s = 100$ s) are shown in Fig. 10. The self- and cross-response results are qualitatively the same compared with the previous subsection. The response functions are zero before the time shift value. After the returns and the trade signs reach their order, the signals grow. The same effect of the apparent stronger signal can be seen here, and again, it is due to the averaging values.

The results in trade time scale and physical time scale can be explained understanding the dynamics of the market. A trade can or cannot change the price of a ticker. Therefore, when a change in price happens, a change in midpoint price, and consequently in returns happens. Thus, it is extremely important to keep the order of the events and the relation between them. When we shift the trade signs and returns, this order is temporarily lost and as outcome the signal does not have any meaningful information. When the order is recovered during the shift, the signal grows again, showing response function values different to zero. In this section, we were interested only in the order (shift)

and not in the responses values, which were analyzed in Sect. 4. A time shift smaller than zero does not have any useful information about the response. If the time shift is equal to zero, the signal is weak, due to the time needed by the market to react to the new information. On the other hand, a time shift larger than two steps shows the information is lost and the signal only grows when the original order is resumed.

Then, the question is what is the ideal time shift to compute the response functions. Our approach in Sect. 4 takes into account that the changes in the quotes are the ones that attract the agents to buy or sell their shares. Hence, they directly impact the trade signs. According to the results, the response can take up to two time steps in the corresponding scale to react to the change in quotes. Thus, a time shift larger than two time steps makes no sense. On the other hand, in the case of the physical time scale, where a sampling is used, to assure the selection of a midpoint price at the beginning of a second, it is a good strategy to use the last midpoint of the previous second as the first midpoint price of the current second. In this case, an apparent 1 s shift is used between returns and trade signs.

6 Time lag analysis

Regarding Eq. (2), we use a time lag τ in the returns to see the gains or loses in a future time. However, the strength of the return in the time lag should not be equal along its length. Then, we divide the full range time lag τ in an immediate time lag and in a late time lag as show in Fig. 11, where:

$$\tau = \tau' + (\tau - \tau'), \quad (28)$$

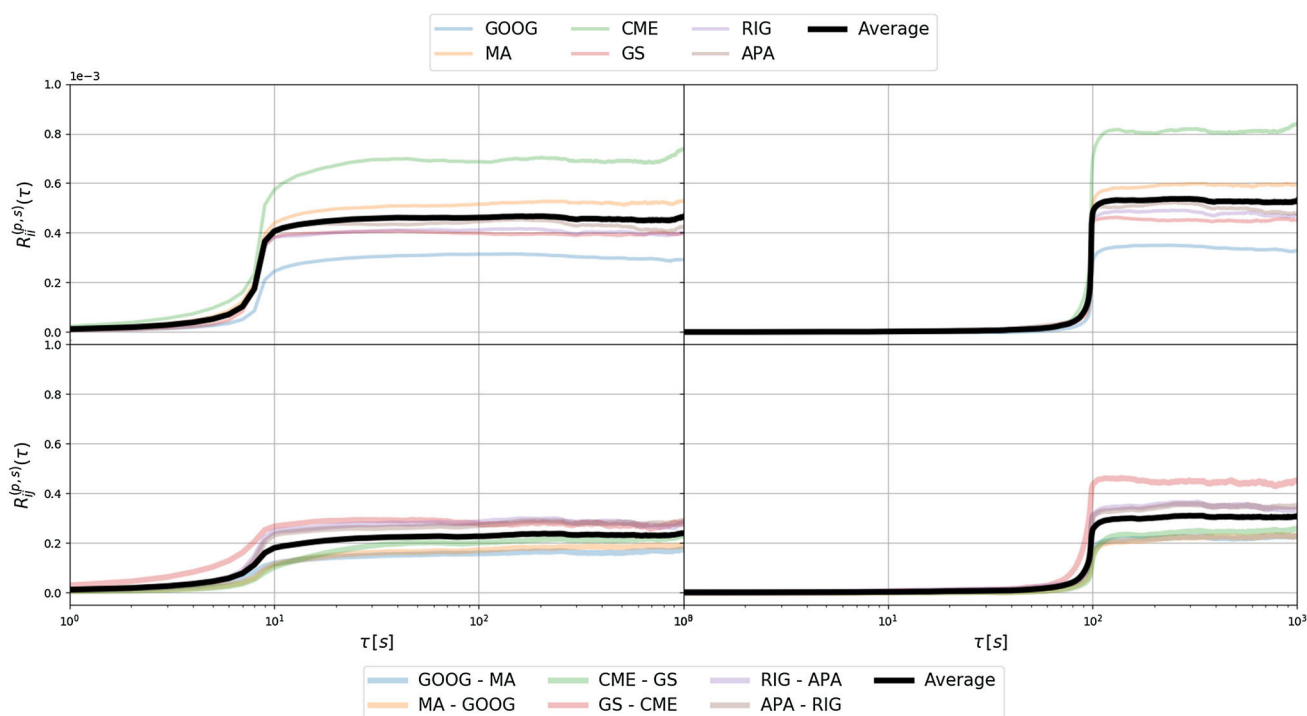


Fig. 10 Self- and cross-response functions $R_{ij}^{(p,s)}(\tau)$ excluding $\varepsilon_j^{(p)}(t) = 0$ in 2008 versus time lag τ on a logarithmic scale for different shifts in physical time scale. Self-responses functions (top) of individual stocks and cross-response functions (bottom) of stocks pairs from the same economic sector. We use time lag values $\tau = 10$ trades (left) and $\tau = 100$ (right)

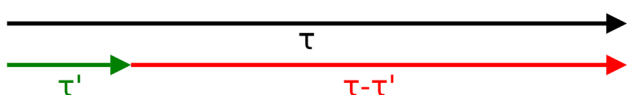


Fig. 11 τ value divided in short and long time lag

for $\tau' < \tau$. This distinguish the returns depending on the time lag as the short (immediate) return τ' with the long return $\tau - \tau'$. This approach is similar to the concept used in Ref. [41], where the price impact for a single trade is estimated by the immediate response on an event time scale. In our case, we check all over the range of the time lag in the price response function on a physical time scale.

To use the short and long time lag, we rewrite the returns in physical time scale as:

$$\begin{aligned} r_i^{(p,sl)}(t, \tau) &= \ln \left(\frac{m_i(t + \tau)}{m_i(t)} \right) \\ &= \ln \left[\left(\frac{m_i(t + \tau)}{m_i(t + \tau')} \right) \left(\frac{m_i(t + \tau')}{m_i(t)} \right) \right] \\ &= \ln \left(\frac{m_i(t + \tau)}{m_i(t + \tau')} \right) + \ln \left(\frac{m_i(t + \tau')}{m_i(t)} \right) \\ &\approx \frac{m_i(t + \tau) - m_i(t + \tau')}{m_i(t + \tau')} + \frac{m_i(t + \tau') - m_i(t)}{m_i(t)}, \end{aligned} \quad (29)$$

where the superscript sl refers to short-long and the second term of the right part is constant with respect

to τ . Replacing Eq. (29) in the price response function in physical time scale (Eq. 16), we have:

$$\begin{aligned} R_{ij}^{(p,sl)}(\tau) &= \left\langle r_i^{(p,sl)}(t - 1, \tau) \varepsilon_j^{(p)}(t) \right\rangle_P \\ &\approx \left\langle \frac{m_i(t - 1 + \tau) - m_i(t - 1 + \tau')}{m_i(t - 1 + \tau')} \varepsilon_j^{(p)}(t) \right\rangle_P \\ &\quad + \left\langle \frac{m_i(t - 1 + \tau') - m_i(t - 1)}{m_i(t - 1)} \varepsilon_j^{(p)}(t) \right\rangle_P, \end{aligned} \quad (30)$$

where the first term in the right side of Eq. (30) is the long response and the right term is the short response. Again, the right term of Eq. (30) is independent of τ .

The results in Fig. 12 show the short response, the long response, the addition of the short response and long response (Sum), the original response, a random response, and the value of τ' .

The main signal of the response function come from the short response. Depending on the stock and the value of τ' , the long response can increase or decrease the short response signal, but in general, the long response does not give a significant contribution to the complete response.

Before τ' , the short response and long response are the same, as the self and cross-response definition do not define values smaller than τ' , so it is computed as the original response. In the figure, the curves of the short and long response are under the curve of the original response. After τ' , the short response is a strong constant signal. On the other hand, the long response

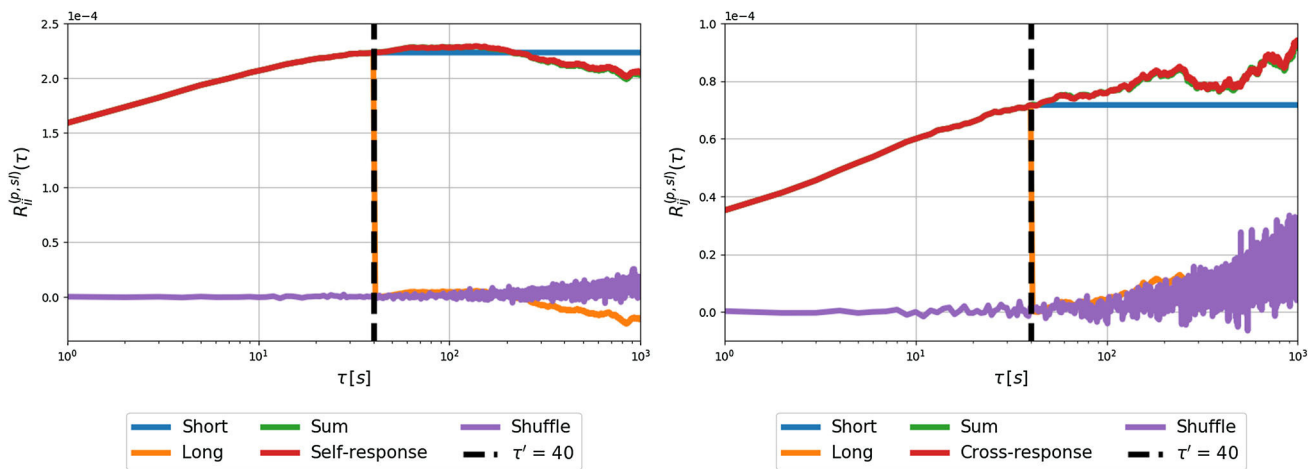


Fig. 12 Self- and cross-response functions $R_{ij}^{(p,sl)}(\tau)$ excluding $\varepsilon_j^{(p)}(t) = 0$ in 2008 versus time lag τ on a logarithmic scale using a $\tau' = 40$ in physical time scale. Self-response functions (left) of Alphabet Inc. stock and cross-response functions (right) of Alphabet Inc.-Mastercard Inc. stocks

immediately fades, showing the small contribution to the final response. To compare the significance of the long response, We added a random response made with the trade signs used to compute the response but with a shuffle order. The long response and the random response are comparable, and show how the long response is not that representative in the final response. If we add the short and long response, we obtain the original response. In Fig. 12, the original response (red line) has the same shape to the addition of the short and long responses (green line).

For the response functions that show the increase-decrease behavior in between the time lag $\tau = 10^3$, the peak is usually between $\tau = 10^1$ and $\tau = 10^2$. In these cases, the long response are always negative after the τ' value and are comparable in magnitude with the random signal. On the other hand, the response functions that require a bigger time lag to show the increase-decrease behavior have non-negative long responses, but, still, they are comparable in magnitude with the random signal.

According to our results, price response functions show a large impact on the first instants of the time lag. We proposed a new methodology to measure this effect and evaluate their consequences. The short response dominates the signal, and the long response vanishes.

7 Spread impact in price response functions

When we calculate the price response functions, the signal of the response depends directly on the analyzed stock. Thus, even if the responses functions are in the same scale, their values differ from one to another. We choose the spread [36] to group 524 stocks in the NASDAQ stock market for the year 2008 in physical time scale, and check how the average strength of the price self-response functions in physical time scale behaved

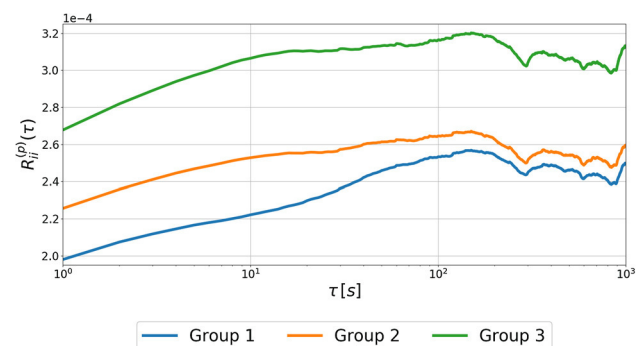


Fig. 13 Average price self-response functions $R_{ii}^{(p)}(\tau)$ excluding $\varepsilon_i^{(p)}(t) = 0$ in 2008 versus time lag τ on a logarithmic scale in physical time scale for 524 stocks divided in three representative groups

for this groups. For each stock, we compute the spread in every second along the market time. Then, we average the spread during the 253 business days in 2008. With this value, we group the stocks.

We used three intervals to select the stocks groups ($s < 0.05\%$, $0.05\% \leq s < 0.10\%$ and $0.10\% \leq s < 0.40\%$). The detailed information of stocks, the spread, and the groups can be seen in Appendix A. With the groups of the stocks defined, we averaged the price response functions of each group.

In Fig. 13, we show the average response functions for the three groups. The average price response function for the stocks with smaller spreads (more liquid) have in average the weakest signal in the figure. On the other hand, the average price response function for the stocks with larger spreads (less liquid) have in average the strongest signal. According to the results described in Sects. 3 and 4, the average price response functions for all the groups follow the increment to a maximum followed by a decrease in the signal intensity.

The strength of the price self-response function signals grouped by the spread can be explained knowing that the response functions directly depend on the trade signs. As long as the stock is liquid, the number of trade signs grow. Thus, at the moment of the averaging, the large amount of trades reduces the response function signal. Therefore, the response function decrease as long as the liquidity grows, and as stated in the introduction, the spread is negatively related to trading volume and, hence, firms with more liquidity tend to have lower spreads.

Finally, an interesting behavior can be seen in Fig. 13. Despite each stock has a particular price response according to the returns and trade signs, the average price response function for the different groups seems to be quite similar. As all the analyzed stocks come from the same market, we can infer that the general behavior of the market affects all the stocks, influencing in average a group response.

8 Conclusion

We went into detail about the response functions in correlated financial markets. We define the trade time scale and physical time scale to compute the self- and cross-response functions for six companies with the largest average market capitalization for three different economic sectors of the S&P index in 2008. Due to the characteristics of the data used, we had to classify and sampling values to obtain the corresponding quantities in different time scales. The classification and sampling of the data had impact on the results, making them smoother or stronger, but always keeping their shape and behavior.

The response functions were analyzed according to the time scales. We proposed a new approach to compare price response functions from different scales. We used the same midpoint prices in physical time scale with the corresponding trade signs in trade time scale or physical time scale. This assumption allowed us to compare both price response functions and get an idea of how representative was the behavior obtained in both cases. For trade time scale, the signal is weaker due to the large averaging values from all the trades in a year. In the physical time scale, the response functions had less noise and their signal were stronger. We proposed an activity response to measure how the number of trades in every second highly impact the responses. As the response functions can not grow indefinitely with the time lag, they increase to a peak, to then decrease. It can be seen that the market needs time to react and revert the growing. In both time scale cases depending on the stocks, two characteristics behavior were shown. In one, the time lag was large enough to show the complete increase-decrease behavior. In the other case, the time lag was not enough, so some stocks only showed the growing behavior.

We modify the response function to add a time shift parameter. With this parameter we wanted to analyze the importance in the order of the relation between

returns and trade signs. In trade time scale and physical time scale, we found similar results. When we shift the order between returns and trade signs, the information from the relation between them is temporarily lost, and as outcome, the signal does not have any meaningful information. When the order is recovered, the response function grows again, showing the expected shape. We showed that this is not an isolated conduct, and that all the shares used in our analysis exhibit the same behavior. Thus, even if they are values of time shift that can give a response function signal, empirically, we propose that this time shift should be a value between $t_s = (0, 2]$ time steps.

We analyzed the impact of the time lag in the response functions. We divided the time lag in a short and long time lag. With this division, we adapted the price response function in physical time scale. The response function that depended on the short time lag showed a stronger response. The long response function vanishes, and depending on the stock could take negative and non-negative values comparable to a random signal.

Finally, we checked the spread impact in price self-response functions. We divided 524 stocks from the NASDAQ stock market in three groups depending on the year average spread of every stock. The response functions signal were stronger for the group of stocks with the larger spreads and weaker for the group of stocks with the smaller spreads. A general average price response behavior was spotted for the three groups, suggesting a market effect on the stocks.

Acknowledgements. We thank S. Wang for fruitful discussions. One of us (JCHL) acknowledges financial support from the German Academic Exchange Service (DAAD) with the program “Research Grants - Doctoral Programmes in Germany” (Funding programme 57381412).

Author contributions

TG proposed the research. SMK and JCHL developed the method of analysis. The idea to look the time shift and to analyze the spread impact was due to JCHL, and the idea of the time lag analysis was due to SMK; JCHL carried out the analysis. All the authors contributed equally to analyze the results and write the paper.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors’ comment: ...].

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to

the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

A NASDAQ stocks used to analyze the spread impact

We analyzed the spread impact in the response functions for 524 stocks from the NASDAQ stock market for the year 2008. In Tables 2, 3, 4, and 5, we listed the stocks in their corresponding spread groups.

References

1. H.-J. Ahn, J. Cai, Y. Hamao, R.Y.K. Ho, The components of the bid-ask spread in a limit-order market: evidence from the Tokyo stock exchange. *J. Empir. Financ.* **9**(4), 399–430 (2002)
2. Y. Amihud, H. Mendelson, The effects of beta, bid-ask spread, residual risk, and size on stock returns. *J. Financ.* **44**(2), 479–486 (1989)
3. M. Benzaquen, I. Mastromatteo, Z. Eisler, J.-P. Bouchaud, Dissecting cross-impact on stock markets: an empirical analysis. *J. Stat. Mech. Theory Exp.* **2017**(2), 023406 (2017)
4. J.-P. Bouchaud, J. Kockelkoren, M. Potters, Random walks, liquidity molasses and critical response in financial markets. *Quant. Financ.* **6**, 115 (2004)
5. J.-P. Bouchaud, The subtle nature of financial random walks. *Chaos Interdiscip. J. Nonlinear Sci.* **15**(2), 026104 (2005)
6. J.-P. Bouchaud, D. Farmer, F. Lillo, *How Markets Slowly Digest Changes in Supply and Demand* (North-Holland, Amsterdam, 2008)
7. J.-P. Bouchaud, Y. Gefen, M. Potters, M. Wyart, Fluctuations and response in financial markets: the subtle nature of “random” price changes. *Quant. Financ.* **4**(2), 176–190 (2004)
8. J.-P. Bouchaud, M. Mézard, M. Potters, Statistical properties of stock order books: empirical results and models. *Quant. Financ.* **2**(4), 251–256 (2002)
9. C. Callahan, C. Lee, T. Yohn, Accounting information and bid-ask spread. *Account. Horiz.* **11**, 50–60 (1997)
10. A. Chakraborti, I. Toke, M. Patriarca, F. Abergel, Econophysics: empirical facts and agent-based models. [arXiv.org](https://arxiv.org/abs/0909.0001), Quantitative Finance Papers, 09 (2009)
11. K.H. Chung, B.F. Van Ness, R.A. Van Ness, Limit orders and the bid-ask spread. *J. Financ. Econ.* **53**(2), 255–287 (1999)
12. R. Cont, Empirical properties of asset returns: stylized facts and statistical issues. *Quant. Financ.* **1**, 223–236 (2001)
13. J.D. Farmer, L. Gillemot, F. Lillo, S. Mike, A. Sen, What really causes large price changes? *Quant. Financ.* **4**(4), 383–397 (2004)

Table 2 Information of the stocks in Group 1

Group 1		
Symbol	Company	Spread ^a
F	Ford Motor Company	0.01\$
Q	Qwest Communications Int.	0.01\$
ETFC	E-Trade Financial Corp.	0.01\$
PFE	Pfizer Inc.	0.01\$
MOT	Motus GI Holdings Inc.	0.01\$
AMD	Advanced Micro Devices	0.01\$
TLAB	Tellabs Inc.	0.01\$
INTC	Intel Corp.	0.01\$
TWX	Time Warner Inc.	0.01\$
CSCO	Cisco Systems Inc.	0.01\$
THC	Tenet Healthcare Corp.	0.01\$
LSI	Life Storage Inc.	0.01\$
MU	Micron Technology Inc.	0.01\$
EMC	EMC Corp.	0.01\$
MSFT	Microsoft Corp.	0.01\$
NOVL	Novell Inc.	0.01\$
JAVA	Sun Microsystems Inc.	0.01\$
ORCL	Oracle Corp.	0.01\$
S	Sprint Nextel Corp.	0.01\$
DELL	Dell Technologies Inc.	0.01\$
AMAT	Applied Material Inc.	0.01\$
SLE	Spark Energy Inc.	0.01\$
SBUX	Starbucks Corp.	0.01\$
DYN	Dynegy Inc.	0.01\$
DUK	Duke Energy Corp.	0.01\$
CMCSA	Comcast Corp.	0.01\$
IPG	Interpublic Group of Co.	0.01\$
BSX	Boston Scientific Corp.	0.01\$
GE	General Electric Company	0.01\$
SYMC	Symantec Corp.	0.01\$
C	Citigroup Inc.	0.01\$
CPWR	Ocean Thermal Energy	0.01\$
NVDA	Nvidia Corp.	0.01\$
YHOO	Yahoo Inc.	0.01\$
BMJ	Bristol-Myers Squibb Co.	0.01\$
ALTR	Altair Engineering Inc.	0.01\$
GLW	Corning Inc.	0.01\$
JDSU	JDS Uniphase Corp.	0.01\$
MYL	Mylan N.V.	0.01\$
HCBK	Hudson City Bancorp	0.01\$
SPLS	Staples Inc.	0.01\$
SGP	Siamgas and Petrochemicals	0.01\$
HST	Host Hotels & Resorts Inc.	0.01\$
AES	Aes Corp.	0.01\$
KFT	Kraft Foods Inc.	0.01\$
NTAP	NetApp Inc.	0.01\$
BAC	Bank of America Corp.	0.01\$
HD	Home Depot Inc.	0.01\$
SOV	Life Storage Inc.	0.01\$
QLGC	QLogic Corp.	0.01\$
T	AT&T Inc.	0.01\$
GPS	Gap Inc.	0.01\$

Table 2 continued

Group 1	Company	Spread ^a
DIS	Walt Disney Company	0.01\$
GM	General Motors Company	0.01\$
JNPR	Juniper Networks Inc.	0.01\$
LOW	Lowe's Companies Inc.	0.01\$
CBS	CBS Corp.	0.01\$
CAG	ConAgra Brands Inc.	0.01\$
LLTC	Linear Technology Corp.	0.01\$
DTV	DirecTV Group	0.01\$
HBAN	Huntington Bancshares Inc.	0.01\$
KG	Kinross Gold Corp.	0.01\$
CMS	CMS Energy Corp.	0.01\$
ODP	Office Depot Inc.	0.01\$
NVLS	Nivalis Therapeutics Inc.	0.01\$
WFC	Wells Fargo & Company	0.01\$
MO	Altria Group Inc.	0.01\$
VZ	Verizon Communications	0.01\$
DHI	D. R. Horton Inc.	0.01\$
SNDK	Sandisk Corp.	0.01\$
CCE	Coca-Cola Enterprises	0.01\$
NI	NiSource Inc.	0.01\$
EXPE	Expedia Group Inc.	0.01\$
AIG	American International	0.01\$
INTU	Intuit Inc.	0.01\$
LTD	Limited Brands Inc.	0.01\$
CNP	CenterPoint Energy Inc.	0.02\$
QCOM	Qualcomm Inc.	0.02\$
JBL	Jabil Inc.	0.02\$
TSN	Tyson Foods Inc.	0.02\$
CA	CA Inc.	0.02\$
XEL	Xcel Energy Inc.	0.02\$
AA	Alcoa Corp.	0.02\$
KR	Kroger Company	0.02\$
MRK	Merck & Company Inc.	0.02\$
NSM	Nationstar Mortgage Holdings	0.02\$
WMT	Walmart Inc.	0.02\$
FITB	Fifth Third Bancorp	0.02\$
EK	Eastman Kodak Company	0.02\$
PHM	PulteGroup Inc.	0.02\$
JPM	JP Morgan Chase & Co.	0.02\$
WAG	Walgreen Co.	0.02\$
SCHW	Charles Schwab Corp.	0.02\$
RF	Regions Financial Corp.	0.02\$
ADBE	Adobe Inc.	0.02\$
MAT	Mattel Inc.	0.02\$
PAYX	Paychex Inc.	0.02\$
PGR	Progressive Corp.	0.02\$
HPQ	HP Inc.	0.02\$
DOW	Dow Inc.	0.02\$
TE	TECO Energy Inc.	0.02\$
BBBY	Bed Bath & Beyond Inc.	0.02\$
JNJ	Johnson & Johnson	0.02\$

Table 2 continued

Group 1	Company	Spread ^a
IP	International Paper Company	0.02\$
RSH	Respiri Ltd.	0.02\$
MER	Mears Group PLC	0.02\$
HAL	Halliburton Company	0.02\$
KO	Coca-Cola Company	0.02\$
PBCT	People's United Financial Inc.	0.02\$
WU	Western Union Company	0.02\$
USB	U.S. Bancorp	0.02\$
MCHP	Microchip Technology Inc.	0.02\$
SO	Southern Company	0.02\$
BJS	BJ's Wholesale Club Holdings	0.02\$
MAS	Masco Corp.	0.02\$
NWL	Newell Brands Inc.	0.02\$
M	Macy's Inc.	0.02\$
CVS	CVS Health Corp.	0.02\$
CTSH	Cognizant Tech. Solutions	0.02\$
GNW	Genworth Financial Inc.	0.02\$
DFS	Discover Financial Services	0.02\$
KEY	KeyCorp	0.02\$
ADI	Analog Devices Inc.	0.02\$
SYI	Sysco Corp.	0.02\$

^a Average spread from 9:40:00 to 15:50:00 New York time during 2008

Table 3 Information of the stocks in Group 1

Group 1	Company	Spread ^a
WMB	Williams Companies Inc.	0.02\$
SE	Sea Ltd American Dep.	0.02\$
PG	Procter & Gamble Co.	0.02\$
CIEN	Ciena Corp.	0.02\$
LEN	Lennar Corp.	0.02\$
AKAM	Akamai Technologies	0.02\$
UNH	UnitedHealth Group Inc.	0.02\$
DD	DuPont de Nemours Inc.	0.02\$
KLAC	KLA Corp.	0.02\$
CIT	CIT Group Inc.	0.02\$
LEG	Leggett & Platt Inc.	0.02\$
WFMI	Whole Foods Market Inc.	0.02\$
MS	Morgan Stanley	0.02\$
VRSN	VeriSign Inc.	0.02\$
AN	AutoNation Inc.	0.02\$
RHT	Red Hat Inc.	0.02\$
CTXS	Citrix Systems Inc.	0.02\$
MDT	Medtronic plc.	0.02\$
NBR	Nabors Industries	0.02\$
MOLX	Molex Inc.	0.02\$
GILD	Gilead Sciences Inc.	0.02\$
CHK	Chesapeake Energy	0.02\$
TJX	TJX Companies Inc.	0.02\$
AMGN	Amgen Inc.	0.02\$
SWY	Safeway Inc.	0.02\$
XOM	Exxon Mobil Corp.	0.02\$
STZ	Constellation Brands	0.02\$
ADSK	Autodesk Inc.	0.02\$
LLY	Eli Lilly and Company	0.02\$

Table 3 continued

Group 1 Symbol	Company	Spread ^a
CTAS	Cintas Corp.	0.02\$
LIZ	Liz Claiborne Inc.	0.02\$
GCI	Gannett Co. Inc.	0.02\$
AXP	American Express Co.	0.02\$
TIE	Titanium Metals Corp.	0.02\$
SAI	SAIC Inc.	0.02\$
PDCO	Patterson Companies	0.02\$
WYE	Wyeth Inc.	0.02\$
COH	Cochlear Ltd.	0.02\$
CVG	Convergys Corp.	0.02\$
WDC	Western Digital Corp.	0.02\$
AVP	Avon Products Inc.	0.02\$
A	Agilent Technologies	0.02\$
JNY	Jones Apparel Group	0.02\$
SLM	SLM Corp.	0.02\$
TGT	Target Corp.	0.02\$
VLO	Valero Energy Corp.	0.02\$
MMC	Marsh & McLennan Co.	0.02\$
CPB	Campbell Soup Company	0.02\$
TYC	Tyco International PLC	0.02\$
MCD	McDonald's Corp.	0.02\$
HON	Honeywell International Inc.	0.02\$
DF	Dean Foods Company	0.02\$
NWSA	News Corp.	0.02\$
URBN	Urban Outfitters Inc.	0.02\$
RX	Recylex	0.02\$
BBY	Best Buy Inc.	0.02\$
CCL	Carnival Corp.	0.02\$
WMI	WMI Investment Corp.	0.02\$
POM	Polymet Mining Corp.	0.02\$
ERTS	Electronic Arts Inc.	0.03\$
MBI	MBIA Inc.	0.03\$
ADM	Archer-Daniels-Midland Co.	0.03\$
PEP	PepsiCo Inc.	0.03\$
CBG	CBRE Group Inc.	0.03\$
IR	Ingersoll Rand Inc.	0.03\$
HNZ	Heinz Company	0.03\$
CTX	Qwest Corp.	0.03\$
TSO	Tesoro Corp.	0.03\$
IGT	International Game Tech.	0.03\$
WYN	Wynnstay Group PLC	0.03\$
GT	The Goodyear Tire & Rubber	0.03\$
JCI	Johnson Controls Int.	0.03\$
JWN	Nordstrom Inc.	0.03\$
FRX	Fennec Pharmaceutical Inc.	0.03\$
FHN	First Horizon National Corp.	0.03\$
ABT	Abbott Laboratories	0.03\$
ADP	Automatic Data Processing	0.03\$
PBG	Pacific Brands Ltd.	0.03\$
ROST	Ross Stores Inc.	0.03\$
KBH	KB Home	0.03\$
YUM	Yum! Brands Inc.	0.03\$
MAR	Marriott International	0.03\$
STJ	St Jude Medical Inc.	0.03\$
FDO	Family Dollar Stores Inc.	0.03\$
ED	Consolidated Edison Inc.	0.03\$
UNM	Unum Group	0.03\$

Table 3 continued

Group 1 Symbol	Company	Spread ^a
ORLY	O'Reilly Automotive Inc.	0.03\$
SVU	SUPERVALU Inc.	0.03\$
EMR	Emerson Electric Company	0.03\$
MI	Marshall and Lisle Corp.	0.03\$
AEP	American Electric Power Co.	0.03\$
NEM	Newmont Corp.	0.03\$
MRO	Marathon Oil Corp.	0.03\$
ITW	Illionois Tool Works Inc.	0.03\$
FFIV	F5 Networks Inc.	0.03\$
CVX	Chevron Corp.	0.03\$
PCAR	PACCAR Inc.	0.03\$
OMC	Omnicom Group Inc.	0.03\$
XRAY	DENTSPLY SIRONA Inc.	0.03\$
COP	ConocoPhillips	0.03\$
PCS	MetroPCS Communications	0.03\$
RHI	Robert Half International Inc.	0.03\$
SEE	Sealed Air Corp.	0.03\$
COST	Costco Wholesale Corp.	0.03\$
FIS	Fidelity National Info. Services	0.03\$
PKI	PerkinElmer Inc.	0.03\$
BMC	BMC Software Inc.	0.03\$
RRD	R.R. Donnelley & Sons Co.	0.03\$
UTX	United Technologies Corp.	0.03\$
D	Dominion Energy Inc.	0.03\$
PBI	Pitney Bowes Inc.	0.03\$
ACAS	American Capital Ltd.	0.03\$
K	Kellogg Company	0.03\$
JCP	J. C. Penney Company	0.03\$
AMT	American Tower Corp.	0.03\$
ALL	Allstate Corp.	0.03\$
MWV	MeadWestvaco Corp.	0.03\$
HRB	H&R Block Inc.	0.03\$
NYT	New York times Company	0.04\$
RDC	Redcape Hotel Group	0.04\$
PTV	Pactiv Company	0.04\$
FISV	Fiserv Inc.	0.04\$
EXPD	Expeditors Int.of Washington	0.04\$
BBT	BB&T Corp.	0.04\$
PCG	Pacific Gas & Electric Co.	0.04\$
BIG	Big Lots Inc.	0.04\$
KMX	CarMax Inc.	0.04\$
TSS	Total System Services Inc.	0.04\$
BK	The Bank of N. Y. Mellon Corp.	0.04\$
TEL	Tellurian Inc.	0.04\$
KSS	Kohl's Corp.	0.04\$
CAT	Caterpillar Inc.	0.04\$
HSY	The Hershey Company	0.04\$
GIS	General Mills Inc.	0.04\$

^a Average spread from 9:40:00 to 15:50:00 New York time during 2008

Table 4 Information of the stocks in Groups 1 and 2

Group 1 Symbol	Company	Spread ^a
HAS	Hasbro Inc.	0.04\$
XTO	XTO Energy Inc.	0.04\$
PPL	PPL Corp.	0.04\$
HOG	Harley-Davidson Inc.	0.04\$
UPS	United Parcel Service	0.04\$
HSP	Hospira Inc.	0.04\$
CTL	CenturyLink Inc.	0.04\$
TDC	Teradata Corp.	0.04\$
BA	Boeing Company	0.04\$
BAX	Baxter International Inc.	0.04\$
PGN	Progress Energy Inc.	0.04\$
DRI	Darden Restaurants Inc.	0.04\$
JNS	Janus Capital Group Inc.	0.04\$
BMS	Bristol-Myers Squibb	0.04\$
RTN	Raytheon Company	0.04\$
CVC	Cablevision Systems	0.04\$
PWR	Quanta Services Inc.	0.04\$
LXK	Lexmark International	0.04\$
CELG	Celgene Corp.	0.04\$
MMM	3M Company	0.04\$
RSG	Republic Services Inc.	0.04\$
DNR	Denbury Resources Inc.	0.04\$
NE	Noble Corp.	0.04\$
NDAQ	Nasdaq Inc.	0.04\$
CINF	Cincinnati Financial	0.04\$
BIIB	Biogen Inc.	0.04\$
KMB	Kimberly-Clark Corp.	0.04\$
EIX	Edison International	0.04\$
NRG	NRG Energy Inc.	0.04\$
IVZ	Invesco Ltd.	0.04\$
AEE	Ameren Corp.	0.04\$
AAPL	Apple Inc.	0.04\$
CAH	Cardinal Health Inc.	0.04\$
MCO	Moody's Corp.	0.04\$
Group 2 Symbol	Company	Spread ¹
HOT	Hot Topic Inc.	0.05\$
GPC	Genuine Parts Co	0.05\$
PEG	Public Service Ent.	0.05\$
EFX	Equifax Inc.	0.05\$
COV	Covidien Ltd.	0.05\$
AET	Aetna Inc.	0.05\$
XL	XL Group Ltd.	0.05\$
DOV	Dover Corp.	0.05\$
VIAB	Viacom Inc.	0.05\$
ABC	AmerisourceBergen Corp.	0.05\$
APC	Anadarko Petroleum Corp.	0.05\$
CBE	Cooper Industries	0.05\$
FAST	Fastenal Company	0.05\$
MHP	McGraw-Hill Companies Inc.	0.05\$
AMZN	Amazon.com Inc.	0.05\$
EQR	Equity Residential	0.05\$
CL	Colgate-Palmolive Company	0.05\$
ECL	Ecolab Inc.	0.05\$
MKC	McCormick & Company Inc.	0.05\$
AOC	Aon Corp.	0.05\$

Table 4 continued

Group 1 Symbol	Company	Spread ^a
TXT	Textron Inc.	0.05\$
CSX	CSX Corp.	0.05\$
PNW	Pinnacle West Capital Corp.	0.05\$
KIM	Kimco Realty Corp.	0.05\$
WPI	Watson Pharmaceuticals	0.05\$
MET	Metlife Inc.	0.05\$
UST	ProShares Ultra	0.05\$
TMO	Thermo Fisher Scientific Inc.	0.05\$
HCP	HCP Inc.	0.05\$
SIAL	Sigma-Aldrich Corp.	0.05\$
MTW	Manitowoc Company	0.04\$
FLIR	FLIR Systems Inc.	0.05\$
EL	Estee Lauder Companies	0.05\$
AYE	Allegheny Energy Inc.	0.05\$
CI	Cigna Corp.	0.05\$
NFLX	Netflix Inc.	0.05\$
CHRW	C.H. Robinson Worlwide Inc.	0.05\$
GENZ	Genzyme Corp.	0.05\$
DDR	DDR Corp.	0.05\$
WFT	West Fraser Timber Co. Ltd.	0.05\$
MHS	Medco Health Solutions Inc.	0.05\$
DTE	DTE Energy Company	0.05\$
TRV	The Travelers Companies	0.05\$
NYX	NYSE Group IPO	0.05\$
COF	Capital One Financial Corp.	0.06\$
WLP	WellPoint Inc.	0.06\$
TWC	TWC Enterprises Ltd.	0.06\$
SWN	Southwestern Energy Co.	0.06\$
NSC	Norfolk Southern Corp.	0.06\$
CMA	Comerica Inc.	0.06\$
NKE	Nike Inc.	0.06\$
EQ	Equillium Inc.	0.06\$
CLX	Clorox Company	0.06\$
CAM	Corporate Actions Middleware	0.06\$
RAI	Reynolds American Inc.	0.06\$
NOC	Northrop Grumman Corp.	0.06\$
PLL	Piedmont Lithium Ltd.	0.06\$
SYK	Stryker Corp.	0.06\$
SRE	Sempra Energy	0.06\$
TIF	Tiffany & Co.	0.06\$
NUE	Nucor Corp.	0.06\$
OXY	Occidental Petroleum Corp.	0.06\$
FPL	First Trust New Opportunities	0.06\$
ESRX	Express Scripts Holding Co.	0.06\$
AIV	Apartment Investment Co.	0.06\$
BHI	Boulevard Holdings Inc.	0.06\$
FCX	Freeport-McMoRan Inc.	0.06\$
APOL	Apollo Group Inc.	0.06\$
SCG	Scentre Group Ltd.	0.06\$
CB	Chubb Limited	0.06\$
IBM	Int. Business Machines Corp.	0.06\$
GME	GameStop Corp.	0.06\$
DE	Deere & Company	0.06\$
LNC	Lincoln National Corp.	0.06\$
AVY	Avery Dennison Corp.	0.07\$
MCK	McKesson Corp.	0.07\$
PLD	Prologis Inc.	0.07\$

Table 4 continued

Group 1 Symbol	Company	Spread ^a
DGX	Quest Diagnostics Inc.	0.07\$
CERN	Cerner Corp.	0.07\$
IRM	Irom Mountain Inc.	0.07\$
COL	Rockwell Collins Inc.	0.07\$
NU	NeutriSci International Inc.	0.07\$
CVH	Coventry Health Care Inc.	0.07\$
WFR	MEMC Electronic Materials	0.07\$
AGN	Allergan PLC.	0.07\$
GAS	GAS	0.07\$
APH	Amphenol Corp.	0.07\$
EXC	Exelon Corp.	0.07\$
WEC	WEC Energy Group Inc.	0.07\$
AFL	AFLAC Inc.	0.07\$
NOV	National Oilwell Varco Inc.	0.07\$
SUN	Sunoco LP	0.07\$
BTU	Peabody Energy Corp.	0.07\$
ROK	Rockwell Automation Inc.	0.07\$
PCL	Plum Creek Timber Co. Inc.	0.07\$
JOYG	Joy Global Inc.	0.07\$
AMP	Ameriprise Financial Inc.	0.07\$
TAP	Molson Coors Beverage Co.	0.07\$

¹ Average spread from 9:40:00 to 15:50:00 New York time during 2008

Table 5 Information of the stocks in Groups 2 and 3

Group 2 Symbol	Company	Spread ^a
GD	General Dynamics Corp.	0.07\$
CSC	Computer Sciences Corp.	0.07\$
FII	Federated Investors Inc.	0.07\$
HIG	Hartford Fin. Services Group	0.07\$
ITT	ITT Inc.	0.07\$
TROW	T. Rowe Price Group Inc.	0.07\$
MDP	Meredith Corp.	0.07\$
GR	Goodrich Corp.	0.07\$
AKS	AK Steel Holding Corp.	0.07\$
PFG	Principal Financial Group	0.07\$
HUM	Humana Inc.	0.07\$
STI	SunTrust Banks Inc.	0.07\$
CMI	Cummins Inc.	0.08\$
STR	Questar Corp.	0.08\$
ZMH	Zimmer Holdings Inc.	0.08\$
FO	Fortune Brands Inc.	0.08\$
SII	Sprott Inc.	0.08\$
HRL	Hormel Foods Corp.	0.08\$
OI	O-I Glass Inc.	0.08\$
HCN	Welltower Inc.	0.08\$
ANF	Abercrombie & Fitch Co.	0.08\$
LM	Legg Mason Inc.	0.08\$
FDX	FedEx Corp.	0.08\$
PRU	Prudential Financial Inc.	0.08\$
DHR	Danaher Corp.	0.08\$

Table 5 continued

Group 2 Symbol	Company	Spread ^a
WY	Weyerhaeuser Company	0.08\$
OKE	ONEOK Inc.	0.08\$
SRCL	Stericycle Inc.	0.09\$
LUK	Leucadia National Corp.	0.09\$
BLL	Ball Corp.	0.09\$
FE	FirstEnergy Corp.	0.09\$
SWK	Stanley Black & Decker Inc.	0.09\$
ESV	Ensco PLC	0.09\$
SHW	Sherwin-Williams Company	0.09\$
BDX	Becton, Dickinson and Co.	0.09\$
VAR	Varian Medical Systems Inc.	0.09\$
TEG	Ten Ent. Group PLC	0.09\$
PH	Parker-Hannifin Corp.	0.09\$
ACE	ACE Ltd.	0.09\$
BNI	Burlington Northern Santa Fe	0.09\$
ACS	Affiliated Computer Services	0.09\$
PXD	Pioneer Natural Resources	0.09\$
AIZ	Assurant Inc.	0.09\$
WAT	Waters Corp.	0.09\$
ATI	Allegheny Technologies Inc.	0.09\$
ETN	Eaton Corp.	0.09\$
VTR	Ventas Inc.	0.09\$
Group 3 Symbol	Company	Spread ^a
HAR	Harman Int. Industries Inc.	0.10\$
NFX	Nasdaq Futures	0.10\$
EQT	EQT Corp.	0.10\$
HRS	Harris Corp.	0.10\$
COG	Cabot Oil & Gas Corp.	0.10\$
PPG	PPG Industries Inc.	0.10\$
MEE	Massey Energy Company	0.10\$
HP	Helmerich & Payne Inc.	0.10\$
DVN	Devon Energy Corp.	0.10\$
IFF	Int. Flavors & Fragrances	0.10\$
LH	Laboratory Corp.	0.10\$
UNP	Union Pacific Corp.	0.10\$
EW	Edwards Lifesciences Corp.	0.10\$
STT	State Street Corp.	0.10\$
RRC	Range Resources Corp.	0.10\$
SJM	J. M. Smucker Company	0.10\$
SNA	Snap-On Inc.	0.11\$
PNC	PNC Fin. Services Group	0.11\$
CEG	Constellation Energy Group	0.11\$
LMT	Lockheed Martin Corp.	0.11\$
FTI	TechnipFMC PLC	0.11\$
CNX	CNX Resources Corp.	0.11\$
ARG	Amerigo Resources Ltd.	0.11\$
DVA	DaVita Inc.	0.11\$
GS	Goldman Sachs Group Inc.	0.11\$
TMK	Torchmark Corp.	0.12\$
EMN	Eastman Chemical Co.	0.12\$
RIG	Transocean Ltd.	0.12\$
PX	Pelangio Exploration Inc.	0.12\$

Table 5 continued

Group 2 Symbol	Company	Spread ^a
RL	Ralph Lauren Corp.	0.12\$
R	Ryder System	0.12\$
MIL	Millipore Corp.	0.12\$
CRM	Salesforce.com Inc.	0.12\$
NTRS	Northern Trust Corp.	0.12\$
APD	Air Products and Chemicals	0.12\$
ANR	Alpha Natural Resources	0.13\$
NBL	Noble Energy Inc.	0.13\$
APA	Apache Corp.	0.13\$
FMC	FMC Corp.	0.13\$
BDK	Black and Decker Corp.	0.13\$
MUR	Murphy Oil Corp.	0.13\$
ROP	Roper Technologies Inc.	0.13\$
JEC	Jura Energy Corp.	0.13\$
HES	Hess Corp.	0.13\$
ETR	Entergy Corp.	0.13\$
MON	Monsanto Company	0.14\$
DV	Dolly Varden Silver Corp.	0.14\$
VFC	V.F. Corp.	0.14\$
GWV	W. W. Grainger Inc.	0.14\$
WHR	Whirlpool Corp.	0.15\$
MOS	Mosaic Company	0.15\$
SPG	Simon Property Group	0.15\$
EOG	EOG Resources Inc.	0.15\$
VMC	Vulcan Materials Co	0.16\$
X	United States Steel Corp.	0.16\$
SHLD	Sears Holding	0.16\$
LLL	L3 Technologies Inc.	0.17\$
WYNN	Wynn Resorts Ltd.	0.17\$
BXP	Boston Properties Inc.	0.17\$
PSA	Public Storage	0.17\$
PCP	Precision Castparts Corp.	0.18\$
VNO	Vornado Realty	0.18\$
DNB	Dun & Bradstreet Corp.	0.19\$
CLF	Cleveland-Cliffs Inc.	0.19\$
FLR	Fluor Corp.	0.20\$
PCLN	Priceline Group Inc.	0.20\$
MTB	M&T Banc Corp.	0.20\$
AVB	AvalonBay Comm. Inc.	0.21\$
BEN	Franklin Resources Inc.	0.21\$
DO	Diamond Offshore Drilling	0.23\$
CF	CF Industries Holdings	0.24\$
AZO	AutoZone Inc.	0.25\$
FLS	Flowserve Corp.	0.27\$
ICE	Intercontinental Exc. Inc.	0.28\$
MA	Mastercard Inc.	0.38\$
FSLR	First Solar Inc.	0.38\$
CMG	Chipotle Mexican Grill	0.38\$

^a Average spread from 9:40:00 to 15:50:00 New York time during 2008

14. J.D. Farmer, P. Patelli, I.I. Zovko, The predictive power of zero intelligence in financial markets. *Proc. Natl. Acad. Sci.* **102**, 2254 (2003)
15. X. Gabaix, P. Gopikrishnan, V. Plerou, H. Stanley, A theory of power-law distributions in financial market fluctuations. *Nature* **423**, 267–70 (2003)
16. A. Gerig, A theory for market impact: how order flow affects stock price. [arXiv.org](https://arxiv.org/), Quantitative Finance Papers, 04 (2008)
17. L.R. Glosten, L.E. Harris, Estimating the components of the bid/ask spread. *J. Financ. Econ.* **21**(1), 123–142 (1988)
18. J.E. Griffin, R.C.A. Oomen, Sampling returns for realized variance calculations: tick time or transaction time? *Econ. Rev.* **27**(1–3), 230–253 (2008)
19. S. Grimm, T. Guhr, How spread changes affect the order book: comparing the price responses of order deletions and placements to trades. *Eur. Phys. J. B* **92**, 1–11 (2018)
20. S. Krause, J. Fiegen, T. Guhr, Emergence of stylized facts during the opening of stock markets. *Eur. Phys. J. B* (2018)
21. F. Lillo, Introduction to market microstructure and heterogeneity of investors. 73–89 (2019)
22. J.F. Lillo, R.M. Farmer, Master curve for price-impact function. *Nature* **421**, 129–30 (2003)
23. J.F. Lillo, D. Farmer, The long memory of the efficient market. *Stud. Nonlinear Dyn. Econ.* **8**(3), 35 (2004)
24. J.C.H. Londono, First release—Price response functions and spread impact in correlated financial markets code. (2020)
25. M.C. Münnix, R. Schäfer, T. Guhr, Compensating asynchrony effects in the calculation of financial correlations. *Phys. A Stat. Mech. Appl.* **389**(4), 767–779 (2010)
26. M.C. Münnix, R. Schäfer, T. Guhr, Impact of the tick-size on financial returns and correlations. *Phys. A Stat. Mech. Appl.* **389**(21), 4828–4843 (2010)
27. M.C. Münnix, R. Schäfer, T. Guhr, Statistical causes for the Epps effect in microstructure noise. *Int. J. Theor. Appl. Financ.* **14**(08), 1231–1246 (2011)
28. Q.N. Nguyen, S. Aboura, J. Chevallier, L. Zhang, B. Zhu, Local Gaussian correlations in financial and commodity markets. *Eur. J. Oper. Res.* **285**(1), 306–323 (2020)
29. M. Potters, J.P. Bouchaud, More statistical properties of order books and price impact. *Phys. A Stat. Mech. Appl.* **324**(1), 133–140 (2003)
30. V. Plerou, P. Gopikrishnan, X. Gabaix, H.E. Stanley, Quantifying stock-price response to demand fluctuations. *Phys. Rev. E* **66**, 027104 (2002)
31. V. Plerou, P. Gopikrishnan, B. Rosenow, L.A.N. Amaral, H.E. Stanley, T. Guhr, Random matrix approach to cross correlations in financial data. *Phys. Rev. E* **65**, 066126 (2002)
32. V. Plerou, H.E. Stanley, X. Gabaix, P. Gopikrishnan, On the origin of power-law fluctuations in stock prices. *Quant. Financ.* **4**(1), 11–15 (2004)
33. B. Rosenow, Fluctuations and market friction in financial trading. *Int. J. Mod. Phys. C* **13**(03), 419–425 (2002)
34. T.A. Schmitt, D. Chetalova, R. Schäfer, T. Guhr, Non-stationarity in financial time series: generic features and tail behavior. *EPL (Europhys. Lett.)* **103**(5), 58003 (2013)
35. E. Smith, J.D. Farmer, L., (n.d.)

- 1105 36. M. Theissen, S.M. Krause, T. Guhr, Regularities and 1123
 1106 irregularities in order flow data. Eur. Phys. J. B **90**, 218 1124
 1107 (2017) 1125
 1108 37. I.M. Toke, *Market Making in an Order Book Model and 1126
 1109 Its Impact on the Spread* (Springer, Milan, 2011), pp. 1127
 1110 49–64 1128
 1111 38. B. Tóth, I. Palit, F. Lillo, J.D. Farmer, Why is equity 1129
 1112 order flow so persistent? J. Econ. Dyn. Control **51**(C), 1130
 1113 218–239 (2015) 1131
 1114 39. S. Wang, Trading strategies for stock pairs regarding to
 1115 the cross-impact cost. (2017)
 1116 40. S. Wang, T. Guhr, Local fluctuations of the signed
 1117 traded volumes and the dependencies of demands: a
 1118 copula analysis. J. Stat. Mech. Theory Exp. **2018**(3),
 1119 033407 (2018)
 1120 41. S. Wang, S. Neusüß, T. Guhr, Grasping asymmetric
 1121 information in price impacts. Eur. Phys. J. B **91**(11),
 1122 266 (2018)
42. S. Wang, S. Neusüß, T. Guhr, Statistical properties of
 market collective responses. Eur. Phys. J. B **91**(8), 191
 (2018)
 43. S. Wang, R. Schäfer, T. Guhr, Average cross-responses
 in correlated financial markets. Eur. Phys. J. B **89**(9),
 207 (2016)
 44. S. Wang, R. Schäfer, T. Guhr, Cross-response in corre-
 lated financial markets: individual stocks. Eur. Phys. J.
 B **89**(4), 1 (2016)

Author Query Form

**Please ensure you fill out your response to the queries raised below
and return this form along with your corrections**

Dear Author

During the process of typesetting your article, the following queries have arisen. Please check your typeset proof carefully against the queries listed below and mark the necessary changes either directly on the proof/online grid or in the 'Author's response' area provided below

Query	Details required	Author's response
1.	Please confirm if the corresponding author is correctly identified. Amend if necessary.	
2.	References are processed as alphabetical order. Kindly check and confirm.	
3.	Kindly provide the complete details for references [21, 22, 25, 35, 39] if possible.	
4.	Kindly check and confirm the complete details for references [5, 7, 15, 30].	
5.	Please provide an explanation as to why there is no data or why the data will not be deposited. Your explanation will be displayed as 'Authors' comment'.	