EPJ manuscript No.
(will be inserted by the editor)

Response functions and spread impact in foreign exchange markets

Juan C. Henao-Londono [a] and Thomas Guhr

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

Received: date / Revised version: date

Abstract Using data from the foreign exchange market, we analyzed the response functions of the majors pairs for three different years. We used two time definitions to compute the response functions: trade time scale and physical time scale. We computed the self-response functions for the seven pairs. Both responses increase to a maximum and then slowly decrease. Hence, the trend in the response functions is eventually reversed. To analyze the spread impact we used XX foreing pairs. We tested the impact in trade time scale and physical time scale. In both scales the less liquid pairs show a larger impact in the response due to the large spread.

PACS. 89.65.Gh Econophysics – 89.75.-k Complex systems – 05.10.Gg Statistical physics

Possible paper names (I will erase this when we choose a name)

The following are possible names for the paper:

- Details that influence the response functions results.
- Influence of the details in the response function measurement.
- Response function measurement in correlated financial markets.
- Response function calculation in correlated financial markets.
- Influence of the methodology in response functions results.

Or we can use another.

1 Introduction

The paper is organized as follows: in Sect. 2 we present our data set of foreign exchange pairs and briefly describe the physical and trade time. We then analyze the definition of the response functions in Sect. **??**, and compute them for the majors pairs. In Sect. **??** we show how the spread impact the values of the response functions. Our conclusions follows in Sect. 3.

2 Data set and time definition

In Sect. 2.1 we introduce the data set used in the paper. In Sect. 2.2 we describe the trade time scale and the physical time scale.

[a] *e-mail:* juan.henao-londono@uni-due.de

2.1 Data set

In this study, we analyze foreign exchange pairs from the foreign exchange market.

We selected the foreign exchange market because ...

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 one second accuracy of the time stamps, in both files more than one quote or trade may be recorded in the same second.

Due to the the time stamp accuracy, it is not possible to match each trade with the directly preceding quote. Hence, we cannot determine the trade sign by comparing the traded price and the preceding midpoint price [10]. In this case we need to do a preprocessing of the data to relate the midpoint prices with the trade signs in trade time scale and in physical time scale.

To analyze the response functions, we select the majors pairs (EUR/USD, GBP/USD, USD/JPY, AUD/USD, USD/CHF, USD/CAD and NZD/USD) for the years 2008, 2014 and 2019.

To analyze the spread impact in the response functions, we use XX foreign exchange pairs in the years 2010, 2014 and 2019. The list of the pairs can be seen in Appendix A.

In order to avoid overnight effects and any artifact due to the opening and closing of the foreign exchange market, we systematically discarded the first ten and the last ten minutes of trading in a given week [1,3,5,10]. Therefore, we only consider trades of the same week from Sunday

17:10:00 to Friday 16:50:00 New York local time. We will refer to this interval of time as the “market time”.

Table 1 shows the companies analyzed with their corresponding symbol and sector. The highest average number of quotes per day and the most liquid stock on average from our selection for the year 2008 is Alphabet Inc. The most traded stock on average from the group was Goldman Sachs Group Inc. On the other side, the stock with the less quotes, less traded and less liquidity on average for the analyzed year was CME Group Inc.

2.2 Time definition

A key concept in the analysis of the response functions is the time. Due to the nature of the data, they are several options to define it.

In general, the time series are indexed in calendar time (hours, minutes, seconds, milliseconds). Moreover, tick-by-tick data available on financial markets all over the world is time stamped up to the millisecond, but the order of magnitude of the guaranteed precision is much larger, usually one second or a few hundreds of milliseconds [2]. In several papers are used different time definitions (calendar time, physical time, event time, trade time, tick time) [2, 4, 6]. The TAQ data used in the analysis has the characteristic that the trades and quotes can not be directly related due to the time stamp resolution of only one second [10]. 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 [2].

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 [4, 10], 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.

Thus, trade sign values will be used in trade time scale and physical time scale and returns will be only used in physical time scale.

2.2.1 Trade time scale

We use the trade sign classification in trade time scale proposed by S. Wang et al. in [10] and used in [7, 8, 9] that reads

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

$\varepsilon^{t}(t,n) = +1$ 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 the classification, 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 [10], the average accuracy of the classification is 85% for the trade time scale.

TAQ time step is one second, 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 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, as explained in [10].

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. For the trade time scale, they 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. **??**, the last midpoint price (circle) between the second -1 and 0 is related with all the trades (squares and triangles) in the second 0 to 1, and so on. It is worth to note, in the seconds that there are no changes in the quotes, it is used the value of the previous second (vertical line over the physical time interval). Thus, all the seconds in the open market time have a midpoint price value, and in consequence returns values. We assume that as long as they were not changes in the quotes, the midpoint price remain the same as the one of the previous second.

We computed all the analysis for the trade time scale using Equations **??** and 1.

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 are supposing all the returns in one second 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.

Table 1. Analyzed companies.

Company	Symbol	Sector	Quotes[1]	Trades[2]	Spread[3]
Alphabet Inc.	GOOG	Information Technology (IT)	164489	19029	\$0.04
Mastercard Inc.	MA	Information Technology (IT)	98909	6977	\$0.38
CME Group Inc.	CME	Financials (F)	98188	3032	\$1.08
Goldman Sachs Group Inc.	GS	Financials (F)	160470	26227	\$0.11
Transocean Ltd.	RIG	Energy (E)	107092	11641	\$0.12
Apache Corp.	APA	Energy (E)	103074	8889	\$0.13

[1] Average number of quotes from 9:40:00 to 15:50:00 New York time.
[2] Average number of trades from 9:40:00 to 15:50:00 New York time.
[3] Average spread from 9:40:00 to 15:50:00 New York time.

illustrate with one example this point. Suppose one second interval, in which they are three different midpoint prices, and as result, three different returns for this three midpoint price values. Furthermore, consider that the volume of limit orders that have the corresponding midpoint prices are the same in the bid and in the ask (so 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 of midpoint price values in the interval. Thus, in this case, we are assuming at the end that all the returns were positive or negative, what probably was not the case, and in consequence will increase or decrease the real value of the net return.

In all the cases we choose the last change in the midpoint price in a second interval as described before (Fig. **??**). We use this method knowing that the variation in one second 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%), so it can give us valuable information about the response functions.

2.2.2 Physical time scale

We use the trade sign definition in physical time scale proposed by S. Wang et al. in [10] and used in [7,9], that depends on the classification in Eq. 1 and reads

$$\varepsilon^{p}\left(t\right)=\begin{cases}\mathrm{sgn}\left(\sum_{n=1}^{N(t)}\varepsilon^{t}\left(t,n\right)\right), & \text{If } N\left(t\right)>0\\ 0, & \text{If } N\left(t\right)=0\end{cases} \tag{2}$$

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 orders. In this definition, they are two ways to obtain $\varepsilon^p(t) = 0$. One way is that in a particular second there is not 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 in the trade time scale, in 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. It is worth to note again, that even if the second does not have a change in quotes, it will has still a midpoint price value and a 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. This can be seen in we related the midpoint price of the previous second with the trade sign of the current second.

According to [10], this definition has an average accuracy up to 82% in the physical time scale.

3 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 values 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. 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. The activity 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, from the theory this time shift should be a value between $t_s = [0, 2]$.

Finally, we analyzed the impact of the time lag in the response functions. We divided the time lag in a short and long time lag. The response function that depended on the short time lag, showed a stronger response. The long response function depending on the stock could take negative and non-negative values. However, in general the influence were not intense.

4 Author contribution statement

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

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)

Appendix A Foreign exchange pairs used to analyze the spread impact

References

1. Jean-Philippe Bouchaud, Yuval Gefen, Marc Potters, and Matthieu Wyart. Fluctuations and response in financial markets: the subtle nature of "random" price changes. *Quantitative Finance*, 4(2):176–190, Apr 2004.
2. Anirban Chakraborti, Ioane Toke, Marco Patriarca, and Frédéric Abergel. Econophysics: Empirical facts and agent-based models. *arXiv.org, Quantitative Finance Papers*, 09 2009.
3. J. Doyne Farmer, László Gillemot, Fabrizio Lillo, Szabolcs Mike, and Anindya Sen. What really causes large price changes? *Quantitative Finance*, 4(4):383–397, 2004.
4. Jim E. Griffin and Roel C. A. Oomen. Sampling returns for realized variance calculations: Tick time or transaction time? *Econometric Reviews*, 27(1-3):230–253, 2008.
5. Stephan Grimm and Thomas Guhr. How spread changes affect the order book: comparing the price responses of order deletions and placements to trades. *The European Physical Journal B*, 92:1–11, 2018.
6. Ioane Muni Toke. *"Market Making" in an Order Book Model and Its Impact on the Spread*, pages 49–64. Springer Milan, Milano, 2011.
7. Shanshan Wang. Trading strategies for stock pairs regarding to the cross-impact cost, 2017.
8. Shanshan Wang and Thomas Guhr. Local fluctuations of the signed traded volumes and the dependencies of demands: a copula analysis. *Journal of Statistical Mechanics: Theory and Experiment*, 2018(3):033407, mar 2018.
9. Shanshan Wang, Rudi Schäfer, and Thomas Guhr. Average cross-responses in correlated financial markets. *The European Physical Journal B*, 89(9):207, Sep 2016.
10. Shanshan Wang, Rudi Schäfer, and Thomas Guhr. Cross-response in correlated financial markets: individual stocks. *The European Physical Journal B*, 89(4), Apr 2016.