
The Long Memory of Order Flow in the Foreign Exchange Spot Market

Martin D. Gould*

*CFM–Imperial Institute of Quantitative Finance
Imperial College, London SW7 2AZ, UK
m.gould@imperial.ac.uk*

Mason A. Porter

*Oxford Centre for Industrial and Applied Mathematics
Mathematical Institute, University of Oxford
Oxford OX2 6GG, UK
CABDyN Complexity Centre
University of Oxford, Oxford OX1 1HP, UK*

Sam D. Howison

*Oxford Centre for Industrial and Applied Mathematics
Mathematical Institute, University of Oxford
Oxford OX2 6GG, UK
Oxford–Man Institute of Quantitative Finance
University of Oxford, Oxford OX2 6ED, UK*

Accepted 15 April 2016

Published 17 June 2016

We study the long memory of order flow for each of three liquid currency pairs on a large electronic trading platform in the foreign exchange (FX) spot market. Due to the extremely high levels of market activity on the platform, and in contrast to existing empirical studies of other markets, our data enables us to perform statistically stable estimation without needing to aggregate data from different trading days. We find strong evidence of long memory, with a Hurst exponent $H \approx 0.7$, for each of the three currency pairs and on each trading day in our sample. We repeat our calculations using data that spans different trading days, and we find no

*Corresponding author.

significant differences in our results. We test and reject the hypothesis that the apparent long memory of order flow is an artifact caused by structural breaks, in favor of the alternative hypothesis of true long memory. We therefore conclude that the long memory of order flow in the FX spot market is a robust empirical property that persists across daily boundaries.

Keywords: Long memory; autocorrelation; foreign exchange market; order flow; market microstructure.

1. Introduction

The autocorrelation properties of financial time series have been the subject of fierce debate for more than 50 years (Cont, 2005; Cont *et al.*, 1997; Gopikrishnan, 1999; Lo, 1991; Mandelbrot, 1963). Several important properties of financial markets have been reported to exhibit autocorrelations that decay slowly, often over periods of days or even months (Booth and Kaen, 1979; Booth *et al.*, 1982; Chakraborti *et al.*, 2011; Cont *et al.*, 1997; Greene and Fielitz, 1977; Mantegna and Stanley, 1999). Such observations have prompted some authors to conjecture that some financial time series exhibit a phenomenon known as *long memory* (Baillie, 1996; Beran, 1994; Cont, 2005), which means that the decay of autocorrelation is sufficiently slow that the sum of terms in their autocorrelation function (ACF) diverges to infinity.

In recent decades, the widespread uptake of electronic limit order books (LOBs; see Gould *et al.*, 2013) in many financial markets has facilitated the recording of order-flow data, which provides a detailed description of traders' actions and interactions at the microscopic scale. The availability of such data has ignited interest in the possibility that financial markets exhibit long memory at the level of individual order flow, and several empirical studies during the past decade have reported this to be the case in a wide variety of different markets (Bouchaud *et al.*, 2004; Lillo and Farmer, 2004; Mike and Farmer, 2008; Tóth *et al.*, 2015).

In a recent publication, Axioglou and Skouras (Axioglou and Skouras, 2011) challenged this view. They noted that, in order to construct sufficiently long time series to perform statistically stable estimation, existing studies of long memory in order flow have aggregated data from multiple trading days. Axioglou and Skouras argued that the apparent long memory reported by many studies is mostly an artifact caused by aggregating the data in this way. Specifically, they argued that the statistical properties of order flow change each day, and that concatenating order-flow series from different trading

days creates non-stationarities at the boundaries between daily series. Moreover, statistical tests are known to produce similar output for non-stationary series as they do for stationary series with long memory (Bhattacharya *et al.*, 1983; Giraitis *et al.*, 2001; Granger and Hyung, 2004). Therefore, distinguishing between these alternatives is a difficult task.

Assessing whether or not order flow really is a long-memory process is important for several reasons. From a practical perspective, the present values of a long-memory process are correlated with values in the distant future (Beran, 1994), so identifying and quantifying the strength of long memory is useful for forecasting. From a theoretical perspective, several recent publications suggest that long-range autocorrelations in order flow may hold the key to understanding the complex statistical properties of price formation in financial markets (e.g., price impact, volatility, and the heavy-tailed distribution of returns (Bouchaud *et al.*, 2009; Farmer *et al.*, 2006; Gerig, 2007; Tóth *et al.*, 2011; Wyart *et al.*, 2008)). Moreover, if order flow really is a long-memory process, then identifying the sources of long-range autocorrelations may provide insight into traders' strategic decision-making processes (Cont and Bouchaud, 2000; Tóth *et al.*, 2015).

In this paper, we perform an empirical study of a new, high-quality data set from a large electronic trading platform in the foreign exchange (FX) spot market to assess the long-memory properties of order flow for three liquid currency pairs. Due to the extremely high levels of market activity on the platform, and in contrast to existing empirical studies of other markets, our data enables us to perform statistically stable estimates of the long-memory properties of intra-day order flow without needing to aggregate data from different trading days. We are therefore able to exclude the possibility that our results are influenced by non-stationarities at the boundaries between different trading days, and we thereby avoid Axioglou and Skouras' (Axioglou and Skouras, 2011) criticism of previous studies. For each of the three currency pairs and on all trading days in our sample, we find strong, statistically significant evidence for long memory in order flow.

To investigate how aggregating data from different trading days impacts our results, we also concatenate pairs of adjacent intra-day order-flow series to create cross-day series, which cross daily boundaries. We repeat all of our calculations on these cross-day series, and we find that our results are very similar to those for the intra-day series. We test and reject the hypothesis that the apparent long memory that we observe is an artifact caused by structural breaks, in favor of the alternative hypothesis of true long memory.

We therefore conclude that the long memory of order flow in the FX spot market is a robust empirical property that persists across daily boundaries.

Several important differences separate our work from previous studies of long memory in order flow. First, our data originates from the FX market, whereas previous studies have analyzed data from equities markets (Axioglou and Skouras, 2011; Bouchaud *et al.*, 2004; Lillo and Farmer, 2004; Mike and Farmer, 2008; Tóth *et al.*, 2015). The FX market is the largest market in the world, so understanding its statistical properties is an important task. Second, the microstructural trade-matching rules in the FX market differ from those in equities markets. We provide a detailed comparison of these trading mechanisms in Sec. 4. By examining how these differences impact traders' actions, we are able to gain additional insight into the underlying causes of the long memory that we observe. Third, because of the extremely high levels of activity in our data, we are able to perform statistically stable estimation of the long-memory properties of intra-day order flow without needing to aggregate data from multiple trading days. We are therefore able to test and reject Axioglou and Skouras' conjecture that the apparent long memory of order flow is mostly an artifact caused by aggregating data in this way. Fourth, we employ a wide range of estimators and statistical techniques to ensure that our quantitative assessment of long memory is robust. By contrast, some previous empirical studies have based their conclusions on single estimators, whose output can be misleading (as we demonstrate in Sec. 5.1). Together, these differences enable us to perform a detailed analysis of the strength, origins, and nature of long memory in order flow in a highly liquid but hitherto unexplored market.

The paper proceeds as follows. In Sec. 2, we provide a detailed discussion of long memory in order flow. In Sec. 3, we review the findings of several empirical studies of long memory in order flow. We describe our data in Sec. 4. We present our main results in Sec. 5, and we discuss our findings in Sec. 6. We conclude in Sec. 7. In Appendix A, we present a technical overview of long memory. In Appendix B, we discuss the statistical techniques that we use to assess the long-memory properties of order flow. In Appendix C, we discuss the statistical tools that we use to distinguish between a long-memory series and a short-memory series with non-stationarities.

2. The Long Memory of Order Flow

In this paper, we perform an empirical analysis of the long-memory properties of order flow. First, we recall the formal definition of long memory and

provide a detailed discussion of the order-flow series that form the basis of our empirical study.

2.1. Long memory

Let

$$\{W_t\} = W_1, W_2, \dots \quad (1)$$

denote a real-valued, second-order stationary¹ time series with mean

$$\mathbb{E}(W_t) = \mu, \quad (2)$$

autocovariance function

$$\gamma(k) = \text{cov}(W_t, W_{t+|k|}), \quad (3)$$

and ACF

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)}. \quad (4)$$

The time series $\{W_t\}$ is said to exhibit *short memory* if

$$\lim_{N \rightarrow \infty} \sum_{k=-N}^N |\rho(k)| < \infty. \quad (5)$$

The time series $\{W_t\}$ is said to exhibit *long memory* if

$$\lim_{N \rightarrow \infty} \sum_{k=-N}^N |\rho(k)| = \infty. \quad (6)$$

For a technical introduction to long memory, see [Appendix A](#) and [Beran \(1994\)](#).

2.2. Limit order books

More than half of the world's financial markets use LOBs to facilitate trade ([Roşu, 2009](#)). In contrast to quote-driven systems, in which prices are set by designated market makers, trade in an LOB occurs via a continuous double-auction mechanism in which institutions submit orders. An *order* $x = (p_x, \omega_x, t_x)$ submitted at time t_x with price p_x and size $\omega_x > 0$ (respectively, $\omega_x < 0$) is a commitment by its owner to sell (respectively, buy) up to $|\omega_x|$ units of the asset at a price no less than (respectively, no greater than) p_x .

¹A time series $\{W_t\}$ is *second-order stationary* if its first and second moments are finite and do not vary with time ([Chatfield, 2000](#); [Taylor, 2008](#)).

Whenever an institution submits a buy (respectively, sell) order x , an LOB’s trade-matching algorithm checks whether it is possible for x to *match* to an active sell (respectively, buy) order y such that $p_y \leq p_x$ (respectively, $p_y \geq p_x$). If so, the matching occurs immediately and the owners of the relevant orders agree to trade the specified amount at the specified price. If not, then x becomes *active*, and it remains active until either it matches to an incoming sell (respectively, buy) order or it is *cancelled*.

Orders that result in an immediate matching upon arrival are called *market orders*. Orders that do not — instead becoming active orders — are called *limit orders*.² The *LOB* $\mathcal{L}(t)$ is the set of all active orders for a given asset on a given platform at a given time t . For a detailed introduction to LOBs, see [Gould *et al.* \(2013\)](#).

Many LOBs record comprehensive digital transcriptions of order flow on a given platform. These transcriptions provide an event-by-event account of the temporal evolution of $\mathcal{L}(t)$, and they thereby enable detailed empirical analysis of financial markets at the microscopic scale ([Cont, 2011](#)).

2.3. Order-sign series

Given a sequence of N consecutive arrivals of limit orders into $\mathcal{L}(t)$, the *order-arrival series*

$$\omega_{x_1}, \omega_{x_2}, \dots, \omega_{x_N} \tag{7}$$

is the time series of the arriving limit orders’ sizes. Similarly, given a sequence of M consecutive departures of active orders from $\mathcal{L}(t)$, the *order-departure series*

$$\omega_{x'_1}, \omega_{x'_2}, \dots, \omega_{x'_M} \tag{8}$$

is the time series of the departing active orders’ sizes. An entry in the order-arrival series always corresponds to the arrival of a new limit order, but an entry in the order-departure series can occur either because an active order is cancelled or because an incoming market order triggers a matching and thereby removes a limit order from the LOB.³ Together, the order-arrival and order-departure series completely determine the temporal evolution of $\mathcal{L}(t)$.

When studying the long-memory properties of order-flow series, it is customary to study the time series of *order signs*. For a given order x_i of size ω_{x_i} ,

²Some platforms allow other order types (such as fill-or-kill, stop-loss, or peg orders ([Knight Capital Group, 2015a](#)), but it is always possible to decompose the resulting order flow into limit and/or market orders. Therefore, we study LOBs in terms of these simple building blocks.

³Incoming market orders are not reported in the order-arrival series.

the order sign L_i is given by

$$L_i = \begin{cases} -1, & \text{if } \omega_{x_i} < 0, \\ +1, & \text{if } \omega_{x_i} > 0. \end{cases} \quad (9)$$

Recall from Sec. 2.2 that an order has negative size if and only if it is a buy order. Therefore, an order-sign series is simply a time series of ± 1 s, where -1 entries correspond to buy-order activity and $+1$ entries correspond to sell-order activity.

The reason for studying time series of order signs — instead of the corresponding time series of order sizes — is that empirical studies of a wide variety of different markets have reported that order sizes often vary over several orders of magnitude (see Gould *et al.* (2013) for a recent survey of empirical studies of LOBs). This brings into question the convergence properties of higher-order moments of time series of order sizes. By contrast, studying only the time series of order signs guarantees that all moments exist, while still providing insight into the long-range autocorrelation properties of buy and sell activity in order flow.

3. Literature Review

Early studies of the autocorrelation properties of order flow tended to focus on short-range (i.e., small-lag) autocorrelations in order-sign series. Hasbrouck (1988) studied the order-sign series for trades on the New York Stock Exchange (NYSE) during March–April 1985. He reported that lag-1 autocorrelations were strongly positive and that the mean sample ACF (see Appendix B.1) across all stocks in the sample was positive up to lags of at least 200. Biais *et al.* (1995) studied order-sign series for market orders, limit orders, and cancellations for 40 stocks traded on the Paris Bourse in 1991. For each type of order flow, they reported that any given event type (e.g., buy market order) was likely to be followed by another event of the same type. Ellul *et al.* (2003) and Yeo (2008) both reported similar findings for activity on the NYSE during 2001.

More recent work has focused on the long-memory properties of order-sign series. Lillo and Farmer (2004) studied order-sign series for limit order arrivals, market order arrivals, and cancellations for 20 stocks on the London Stock Exchange (LSE) during 1999–2002. They used a wide variety of statistical techniques and estimators to test and reject the hypothesis that these series were short-memory series, in favor of the alternative hypothesis of long memory. They also estimated the Hurst exponent (see Appendix A.3) for

each series and reported a mean value of $H \approx 0.7$. The cross-sectional variation in H across the stocks that they studied was small but significant. Because their sample ACFs contained no significant peaks or breaks corresponding to the length of a single trading day, Lillo and Farmer argued that long memory in order flow persists across daily boundaries. They also repeated their experiments on similar data from the NYSE and found similar results.

[Bouchaud *et al.* \(2004\)](#) studied long-range autocorrelations in the order-sign series for market orders on Euronext in 2001–2002. For all of the stocks that they studied, they reported that the sample ACFs decayed approximately according to a power law. They estimated each stock’s power-law exponent directly from its sample ACF and reported values that correspond to Hurst exponents (see Eq. (A.8)) ranging from $H \approx 0.65$ to $H \approx 0.9$. Similarly to Lillo and Farmer, [Bouchaud *et al.*](#) argued that long-range autocorrelations in order flow persist across daily boundaries.

[Mike and Farmer \(2008\)](#) studied order-sign series for both market orders and limit orders for 25 stocks traded on the LSE during 2000–2002. They used detrended fluctuation analysis (DFA) (see Appendix B.4) to estimate the Hurst exponent for each stock and reported values ranging from $H \approx 0.75$ to $H \approx 0.88$, with a mean of $H \approx 0.83$ across all stocks.

To date, two mechanisms have been proposed to explain the slow decay of autocorrelations in order-flow series. The first is that traders display “herding” behavior, either because they all respond similarly to common information or because they monitor each other’s actions and update their strategies by imitating those of their most successful competitors ([LeBaron and Yamamoto, 2007](#)). The second is that traders who wish to perform large trades decompose them into smaller chunks, which they submit over several days (or even months) to minimize their market impact ([Bouchaud *et al.*, 2009, 2004; Lillo *et al.*, 2005](#)). This strategy is commonly known as *order splitting*.

[Gerig \(2007\)](#) assessed the plausibility of these two explanations by studying order-flow series from the LSE. In contrast to most LOB data sets, Gerig’s data included information about the broker that submitted each order. This enabled him to compare the autocorrelation properties of order flow generated by individual brokers to those of the aggregate order flow generated by all brokers. He reported that correlations across different brokers decayed quickly to 0, but that autocorrelations in order flow from individual brokers exhibited long memory. He therefore argued that order splitting is a much more plausible explanation for long memory in order flow

than is herding. [Tóth *et al.* \(2012, 2015\)](#) also studied data containing brokerage identifiers from the LSE, and reached a similar conclusion.

Recently, however, [Axioglou and Skouras \(2011\)](#) challenged the notion that order flow exhibits long memory by arguing that this apparent effect was mainly an artifact caused by non-stationarities in the underlying order-flow series. Specifically, they noted that to construct sufficiently long time series to perform statistically stable estimation of long-range autocorrelations, existing studies have aggregated order-flow data from different trading days. They argued that such aggregation produces order-flow series with structural breaks at the daily boundaries. Because many statistical tests produce similar output for non-stationary series as they do for stationary series with long memory (see [Appendix C](#)), Axioglou and Skouras conjectured the apparent long memory in order flow is mainly an artifact caused by these structural breaks.

To test this hypothesis, Axioglou and Skouras studied the order-sign series for market orders on the LSE during 2005–2006. They first aggregated data from several different trading days, and they noted that standard statistical tests applied to this data concluded strongly in favor of long memory. They then constructed shorter time series by aggregating data across pairs of consecutive trading days. They applied the cumulative-sum change-point estimator (see [Eq. \(C.2\)](#)) to these series, and they were able to detect the daily boundaries with high accuracy. They then applied Berkes' change-point test (see [Appendix C](#)) to test the hypothesis that the apparent long memory in the cross-day series was actually due to a structural break. Working at the 5% significance level, they could not reject the null hypothesis of a piecewise stationary series with a structural break in favor of the alternative hypothesis of true long memory in about two thirds of the order-flow series that they studied. They concluded that although order flow exhibited significant autocorrelations within a single trading day, the strength of autocorrelations that persisted across daily boundaries was very weak.

4. Data

We have been granted access to a recent, high-quality data set from Hotspot FX ([Knight Capital Group, 2015a,b](#)), which is one of the largest multi-institution trading platforms in the FX spot market. According to the 2010 Triennial Central Bank Survey ([Bank for International Settlements, 2010](#)), the mean daily turnover of the global FX market was approximately 4.0 trillion US dollars. Approximately 37% of this volume was due to spot trades,

of which approximately 40% was conducted electronically. In total, the mean daily volume traded on all multi-institution electronic trading platforms was approximately US 0.6 trillion USD (Bank for International Settlements, 2010). The mean daily volume traded on Hotspot FX during the same period was approximately US 21.5 billion USD (Knight Capital Group, 2015c). Therefore, trade on Hotspot FX accounted for approximately 4% of all volume traded electronically in the FX spot market during this period.

Hotspot FX offers trade for more than 60 different currency pairs. A price for the currency pair XXX/YYYY denotes how many units of the *counter currency* YYY are exchanged per unit of the *base currency* XXX. Trade for each currency pair occurs in a separate LOB with price-time priority.

The Hotspot FX platform serves a broad range of trading professionals — including banks, financial institutions, hedge funds, high-frequency traders, corporations, and commodity trading advisers (Knight Capital Group, 2015b). As is customary on multi-institution trading platforms in the FX spot market, Hotspot FX enables institutions to specify *counterparty credit limits* (CCLs) for their trading counterparties. Each institution can only access the trading opportunities offered by counterparties with whom they possess sufficient bilateral credit. We call this market organization a *quasi-centralized limit order book* (QCLOB), because different institutions have access to different subsets of a centralized liquidity pool. For a detailed discussion of QCLOBs, see Gould *et al.* (2016). Examples of platforms that utilize QCLOBs include Reuters (Thomson–Reuters, 2011), EBS (2011), and Hotspot FX (Knight Capital Group, 2015b). For more details about trade on Hotspot FX, see Knight Capital Group (2015a).

The data that we study describes the full order-arrival and order-departure series (see Sec. 2.3) for the EUR/USD (Euro/US dollar), GBP/USD (Pounds sterling/US dollar), and EUR/GBP (Euro/Pounds sterling) currency pairs during the *peak trading hours* of 08:00:00–17:00:00 GMT on 30 trading days during May–June 2010. Although trade in the FX spot market continues to operate outside of these hours, more than 70% of the total traded volume for each of the three currency pairs occurs during these 9 hours each day.

For a given trading day D_i , we use the Hotspot FX data (see Sec. 4) to produce an ordered list of the limit order arrivals that occur during the peak trading hours of 08:00:00–17:00:00 GMT. We then use Eq. (9) to deduce the intra-day arrival sign series from this list. Similarly, we use the Hotspot FX data to produce an ordered list of the active order departures that occur during the same period, and we then use Eq. (9) to deduce the intra-day

Table 1. Minimum, maximum, and mean number of limit order (top panel) arrivals and (bottom panel) departures for EUR/USD, GBP/USD, and EUR/GBP, measured across the 30 trading days in our sample.

		EUR/USD	GBP/USD	EUR/GBP
Number of Arrivals	Minimum	3455561	2962688	2019826
	Maximum	6003406	5296372	3623053
	Mean	4533550.8	4340345.4	2932726.8
Number of Departures	Minimum	3449793	2961217	2019672
	Maximum	5992343	5293082	3622559
	Mean	4524175.4	4337320.6	2932171.5

departure sign series from this list. We repeat this process for each of the 30 trading days D_1, D_2, \dots, D_{30} in our sample.

In Table 1, we list the minimum, maximum, and mean number of limit order arrivals and departures across these 30 trading days. The number of arriving limit orders is largest for EUR/USD and smallest for EUR/GBP. For each of the three currency pairs and on most trading days that we study, the total number of limit order arrivals slightly exceeds the total number of departures, which implies that active orders accumulate throughout the trading day.

To quantify the imbalance between buying and selling activity, we also calculate the percentage of sell orders (i.e., percentage of +1 entries in the order-sign series) among limit order arrivals and departures each day (see Table 2). For each of the three currency pairs and for both arrivals and departures, sell orders account for close to 50% of order flow. Therefore, the level of buying activity is approximately equal to the level of selling activity on all trading days that we study.

Table 2. Minimum, maximum, and mean percentages of sell orders among limit order (top panel) arrivals and (bottom panel) departures for EUR/USD, GBP/USD, and EUR/GBP across the 30 trading days in our sample.

		EUR/USD (%)	GBP/USD (%)	EUR/GBP (%)
Percentage of Arriving Sell Orders	Minimum	48.664	49.389	48.817
	Maximum	50.870	51.255	50.252
	Mean	50.008	50.135	49.894
Percentage of Departing Sell Orders	Minimum	48.667	49.389	48.815
	Maximum	50.879	51.257	50.251
	Mean	50.009	50.134	49.895

5. Results

In this section, we present our empirical results for the arrival-sign series. The corresponding results for the departure-sign series are qualitatively similar.

5.1. Results for intra-day series

In Fig. 1, we plot the sample ACFs (see Appendix B.1) for each of the three currency pairs' intra-day arrival-sign series on 4th May 2010. The results for all other intra-day series on each day in our sample are qualitatively similar. Up to lags of about 25 events, the sample ACFs fluctuate between positive and negative values, which indicates that the order-flow series contain short-range negative autocorrelations. Although these autocorrelations have a magnitude below about 0.1 and are therefore relatively weak, this effect is present on each day in our sample, and we therefore deem it to be a robust statistical property of the data. We return to our discussion of these negative autocorrelations in Sec. 6.

In Fig. 2, we plot the intra-day sample ACFs in doubly logarithmic coordinates. To help reduce the noise at higher lags,⁴ we plot the mean sample ACFs, which we obtain by averaging the daily sample ACFs across all 30 days in our sample. After the short-term negative autocorrelations subside (which occurs before lag 50, the lower bound in our plots), the sample ACFs remain positive for lags of several thousands of events. This suggests that there are long-range, positive autocorrelations in the series.

In Fig. 3, we show rescaled-range plots (see Appendix B.2) for each of the three currency pairs' intra-day arrival-sign series. For each of the three currency pairs, the slope of the rescaled-range plot for each intra-day order-flow series is close to 0.5 for values of k below about 10,000. For larger values of k , the slope of each rescaled-range plot is above 0.5. This suggests that the intra-day order-flow sign series are long-memory processes.

To test the hypothesis of long memory more formally, we perform Lo's modified rescaled-range test (see Appendix B.3) on the intra-day arrival series (see Fig. 4). For each day in our sample, Lo's modified rescaled-range test causes us to reject the null hypothesis of short memory at the 5% significance level at all bandwidth choices that we study. Similarly, when using Andrews' plug-in estimator (see Eq. (B.10)) to estimate a suitable choice of bandwidth parameter from the data, Lo's test rejects the null hypothesis of short

⁴The statistical errors associated with estimating the sample ACF are approximately constant at all lags, but the signal strength is smaller at larger lags because there are fewer independent data points (Bouchaud *et al.*, 2009).

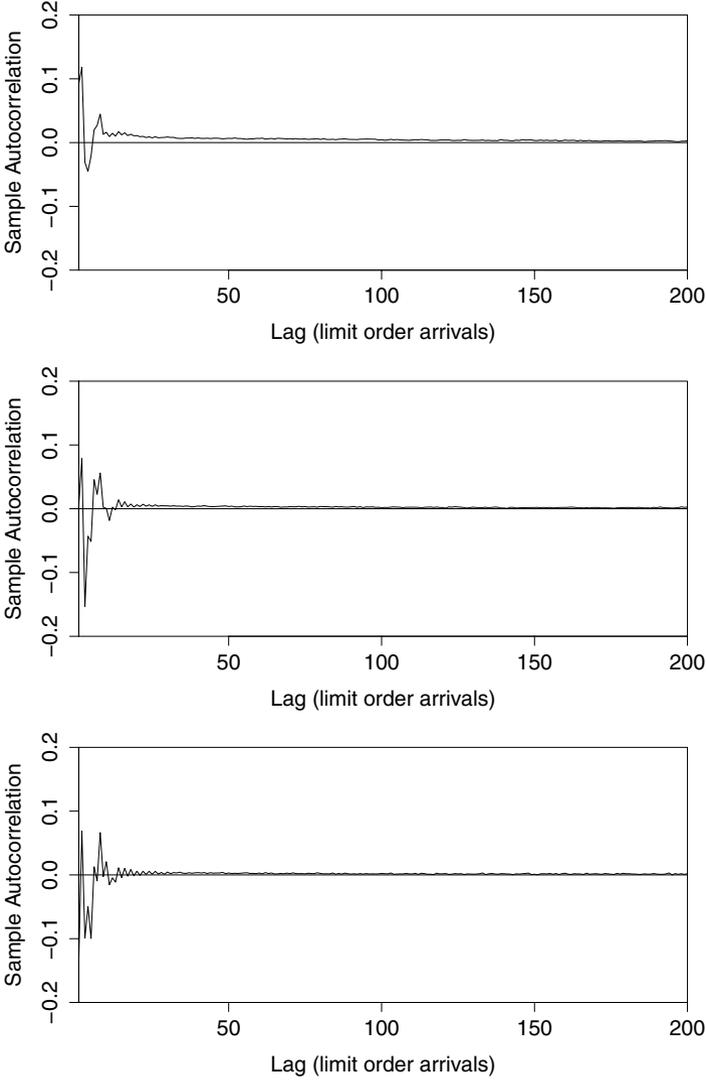


Fig. 1. Sample ACFs (see Appendix B.1) for (top row) EUR/USD, (middle row) GBP/USD, and (bottom row) EUR/GBP intra-day arrival-sign series, which we construct using Eq. (9). Each plot shows the sample ACF for 4th May 2010. The corresponding results for all other days in our sample and for the intra-day departure-sign series are qualitatively similar.

memory for each of the three currency pairs and on all 30 days in our sample. Therefore, Lo's test provides strong evidence that the intra-day order-sign series are long-memory processes.

Given the strong results of Lo's test, we now turn to assessing the strength of the long memory in intra-day order flow. To do so, we use two different methods to estimate the Hurst exponent H (see Appendix A.3): DFA

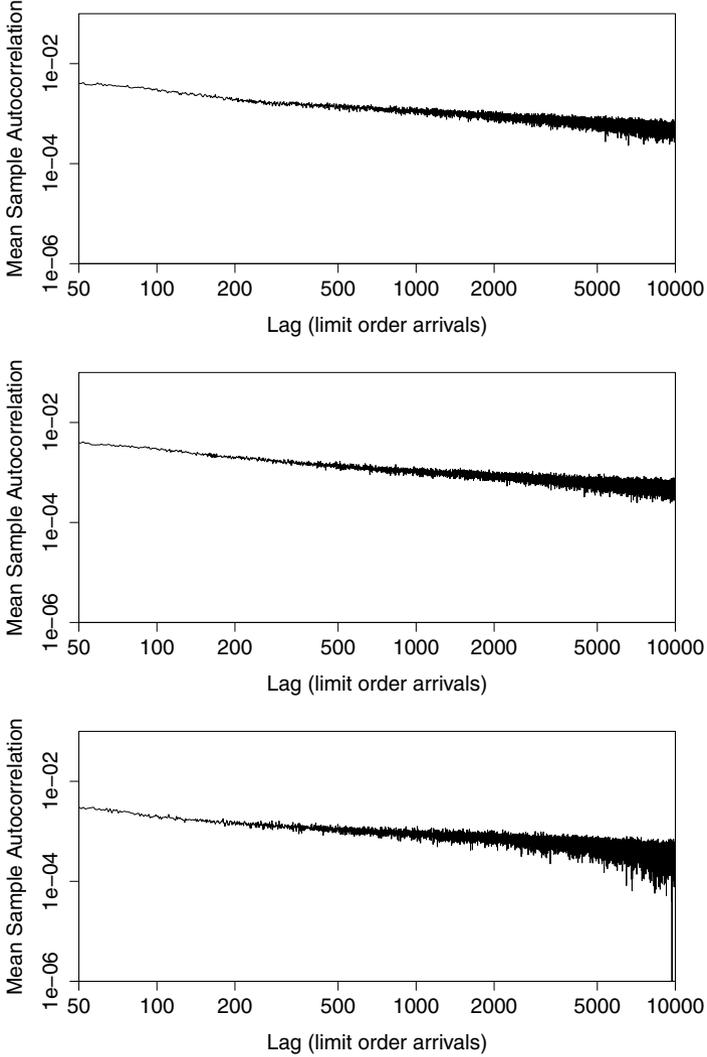


Fig. 2. Doubly logarithmic plots of mean sample ACFs (see Appendix B.1) for (top row) EUR/USD, (middle row) GBP/USD, and (bottom row) EUR/GBP intra-day arrival-sign series. We obtain each plot by averaging the daily sample ACFs across all 30 days in our sample. We omit lags smaller than 50 events because some values of the mean sample ACFs are negative in this range. The results for the intra-day departure-sign series are qualitatively similar.

(see Appendix B.4) and log-periodogram regression (see Appendix B.5). Due to the negative short-range autocorrelations that we observe in the sample ACFs (see Fig. 1), it is necessary to identify sensible choices of input parameters — namely, the minimal window length m_{\min} of a DFA and the number c of Fourier frequencies in a log-periodogram regression — when performing these estimation techniques.

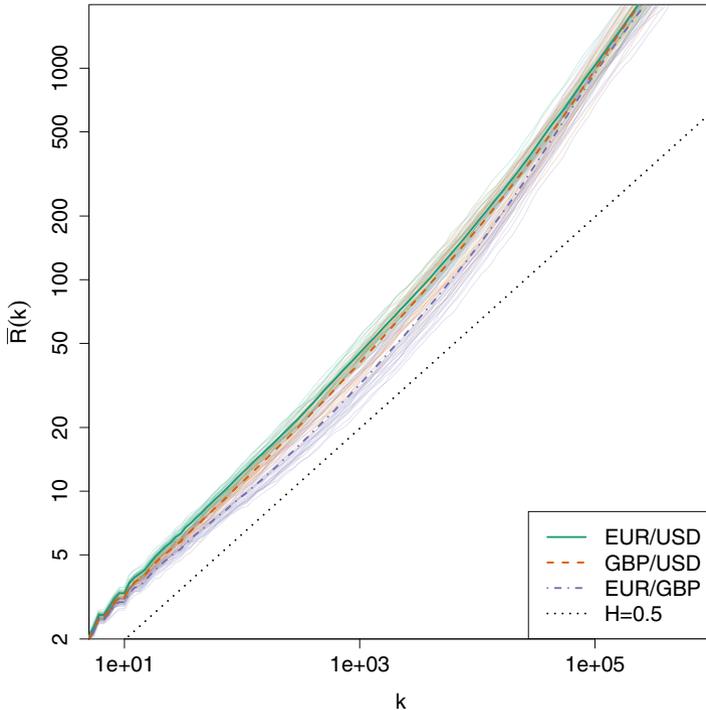


Fig. 3. Rescaled-range plots (see Appendix B.2) for (solid green curves) EUR/USD, (dashed orange curves) GBP/USD, and (dotted-dashed purple curves) EUR/GBP arrival-sign series. The pale curves indicate the rescaled-range statistics $\bar{R}(k)$ (see Eq. (B.5)) for a single intra-day series, and the darker curves indicate the mean across all 30 intra-day series. The dotted black line has a slope of 0.5. In these plots, we divide each intra-day series into $B = 100$ blocks; we also produced similar plots for several different values of $B \in [10, 1000]$ and obtained similar results. The results for the intra-day departure-sign series are qualitatively similar.

To identify a suitable choice of m_{\min} for our DFA estimates of H , we first plot the length- m mean detrended standard deviation $F(m)$ for several choices of m (see Fig. 5). For window lengths m that are smaller than about 25, the negative autocorrelations dominate the mean detrended standard deviations $F(m)$. These values of m are therefore unsuitable for calculating a DFA estimate of H . For values of m larger than about 100, the log-log plots of $F(m)$ follow an approximately straight line. We therefore perform our DFA estimates of H using $m_{\min} = 100$.

To identify a suitable choice of c for our log-periodogram regression estimates of H , we plot the log-periodogram regression estimates of H for several different values of c (see Fig. 6). In all cases, the estimates of H tend to decrease as c increases, and there is no clear plateau over which the estimates of H are stable. In the absence of an obvious choice for c , we use the popular

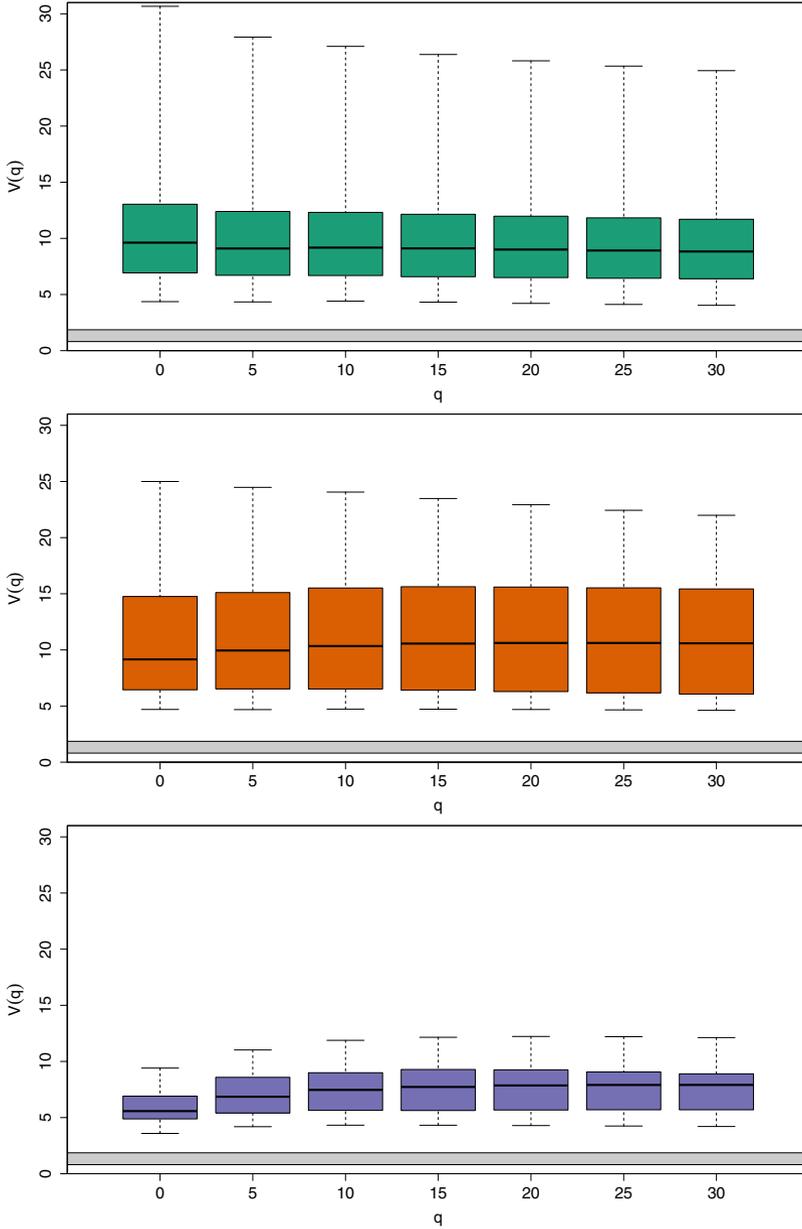


Fig. 4. Box plots of Lo’s modified rescaled-range test statistic $V(q)$ (see Appendix B.3) for the given choices of bandwidth parameter q for the (top) EUR/USD, (middle) GBP/USD, and (bottom) EUR/GBP intra-day arrival-sign series. For each choice of q , the boxes indicate the lower quartile, median, and upper quartile of $V(q)$ and the whiskers indicate the minimum and maximum of $V(q)$, across all 30 intra-day series. The light grey shading indicates the critical region for Lo’s modified rescaled-range test at the 5% significance level. The results for the intra-day departure-sign series are qualitatively similar.

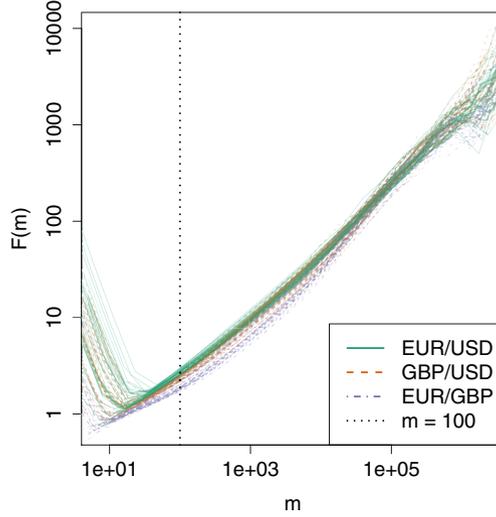


Fig. 5. Length- m mean detrended standard deviation $F(m)$ (see Appendix B.4) for the (solid green curves) EUR/USD, (dashed orange curves) GBP/USD, and (dotted-dashed purple curves) EUR/GBP arrival-sign series. Each curve corresponds to a single intra-day series. The dotted black line indicates $m = 100$. The results for the intra-day departure-sign series are qualitatively similar.

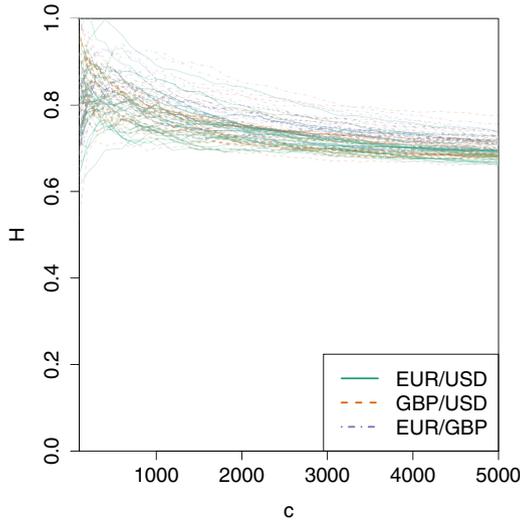


Fig. 6. Log-periodogram regression (see Appendix B.5) estimates of H for given number c of Fourier frequencies for the (solid green curves) EUR/USD, (dashed orange curves) GBP/USD, and (dotted-dashed purple curves) EUR/GBP arrival-sign series. Each curve corresponds to a single intra-day series. The results for the intra-day departure-sign series are qualitatively similar.

rule of thumb (Geweke and Porter-Hudak, 1983) $c = \sqrt{N}$, where N is the length of the given series. We stress, however, that the plots in Fig. 6 indicate that our log-periodogram regression estimates of H depend heavily on this choice, so using a different choice for c would produce quantitatively different results. For example, another popular rule of thumb (Taqqu *et al.*, 1995) is $c = 0.1 \times (N/2)$. Due to the extremely large size of the intra-day order-flow series, this choice produces estimates of $H \approx 0.5$, which we do not regard to be sensible given that the other statistical tests all suggest that the intra-day order-flow series exhibit long memory. The absence of a clear choice for c highlights a weakness of log-periodogram regression for the present application.

In Fig. 7, we plot the DFA and log-periodogram regression estimates of H (using our choices of $m_{\min} = 100$ and $c = \sqrt{N}$) for each intra-day arrival-sign series. For each of the three currency pairs, the DFA estimates of H cluster in the range from about 0.6 to about 0.8. The log-periodogram regression estimates of H tend to be slightly larger (they cluster in the range of about 0.65 to about 0.85). However, the latter results depend heavily on the choice of c (see Fig. 6), so we deem the DFA estimates to be more useful.

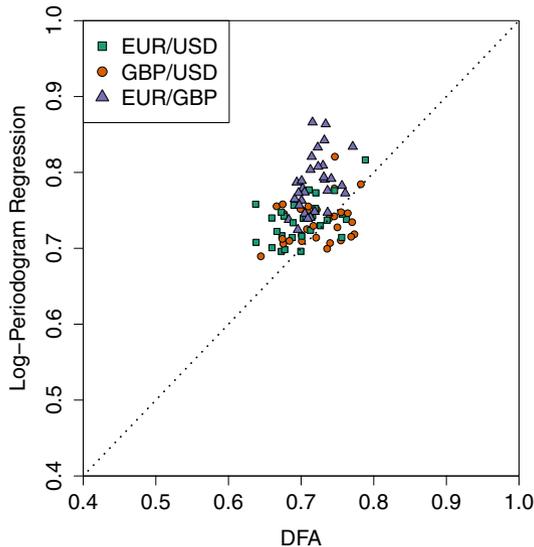


Fig. 7. DFA and log-periodogram regression estimates of the Hurst exponent H for the intra-day (green squares) EUR/USD, (orange circles) GBP/USD, and (purple triangles) EUR/GBP arrival-sign series. Each point corresponds to the estimates for a single trading day. The dotted black line indicates the diagonal. The results for the intra-day departure-sign series are qualitatively similar.

Table 3. DFA and log-periodogram regression estimates of the Hurst exponent H for the EUR/USD, GBP/USD, and EUR/GBP intra-day arrival-sign and departure-sign series. Each entry indicates the mean of the estimates across all intra-day series. The numbers in parentheses indicate 1 standard deviation of the estimates across all intra-day series.

		DFA	Log-periodogram regression
EUR/USD	Arrivals	0.70 (0.04)	0.74 (0.03)
	Departures	0.70 (0.04)	0.74 (0.03)
GBP/USD	Arrivals	0.72 (0.04)	0.74 (0.03)
	Departures	0.72 (0.03)	0.74 (0.03)
EUR/GBP	Arrivals	0.72 (0.02)	0.79 (0.04)
	Departures	0.71 (0.02)	0.78 (0.04)

In Table 3, we list the means and standard deviations of our DFA and log-periodogram regression estimates of H across all 30 days in our sample. As we noted above, our results for arrival-sign series and departure-sign series are very similar. In all cases, the mean estimates of H are many standard deviations larger than 0.5, which strongly supports the hypothesis that these series exhibit long memory. Based on our results in Figs. 5–7 and Table 3, we regard $H \approx 0.7$ to be a good estimate for the Hurst exponent of the arrival-sign series and departure-sign series for each of the three currency pairs.

5.2. Results for cross-day series

To assess how aggregating data from different trading days impacts our results, we now repeat all of our calculations using order-flow series that span daily boundaries. Specifically, for a pair of consecutive trading days D_i and D_{i+1} , we construct the *cross-day arrival-sign series* by concatenating the second half⁵ of the intra-day arrival-sign series from day D_i and the first half of the corresponding intra-day arrival-sign series from day D_{i+1} . We construct the *cross-day departure-sign series* similarly using the corresponding departure-sign series.

In all cases, we find that the sample ACFs, rescaled-range plots, and results from Lo’s modified rescaled-range test are qualitatively similar to those for the intra-day series (see Figs. 1–4). This provides strong evidence that the cross-day order-flow series exhibit long memory. To quantify the strength of this long memory, we calculate DFA and log-periodogram regression estimates of the Hurst exponent H using the same choices of input

⁵If an intra-day series has odd length, we round down to the previous integer.

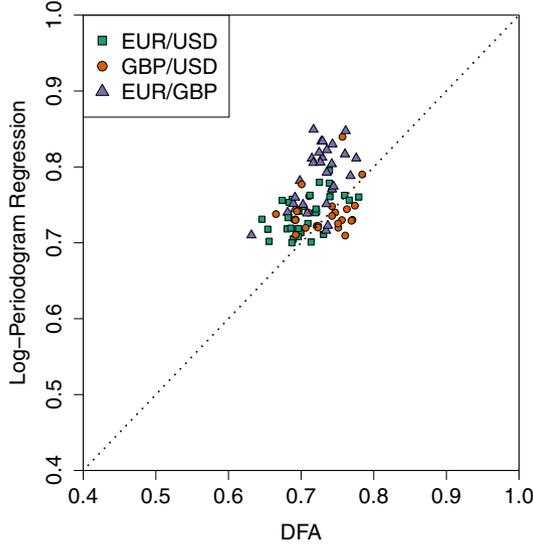


Fig. 8. DFA and log-periodogram regression estimates of the Hurst exponent H for the cross-day (green squares) EUR/USD, (orange circles) GBP/USD, and (purple triangles) EUR/GBP arrival-sign series. Each point corresponds to the estimates for a consecutive pair of trading days. The dotted black line indicates the diagonal. The results for the cross-day departure-sign series are qualitatively similar.

parameters ($m_{\min} = 100$ and $c = \sqrt{N}$) as we used for our corresponding estimates of H for the intra-day series. We plot our results in Fig. 8 and list the means and standard deviations of our estimates across all 29 cross-day periods in Table 4.

Our estimates of H for the cross-day series are very similar to the corresponding estimates for the intra-day series (see Fig. 8 and Table 3). As with

Table 4. DFA and log-periodogram regression estimates of the Hurst exponent H for the EUR/USD, GBP/USD, and EUR/GBP cross-day arrival-sign and departure-sign series. Each entry indicates the mean of the estimates across all cross-day series. The numbers in parentheses indicate 1 standard deviation of the estimates across all cross-day series.

		DFA	Log-periodogram regression
EUR/USD	Arrivals	0.71 (0.03)	0.74 (0.03)
	Departures	0.70 (0.03)	0.74 (0.03)
GBP/USD	Arrivals	0.73 (0.03)	0.74 (0.03)
	Departures	0.73 (0.03)	0.74 (0.03)
EUR/GBP	Arrivals	0.72 (0.03)	0.79 (0.04)
	Departures	0.72 (0.03)	0.78 (0.04)

the intra-day series, the DFA estimates of H cluster in the range of about 0.6 to about 0.8 and the log-periodogram regression estimates of H cluster in the range of about 0.65 to about 0.85. We therefore regard $H \approx 0.7$ to be a good estimate for the Hurst exponent of the cross-day order-sign series for each of the three currency pairs.

5.3. Long memory versus non-stationarity

In this section, we address the conjecture by [Axioglou and Skouras \(2011\)](#) that the apparent long memory in the cross-day series is mostly an artifact caused by structural breaks at the boundaries between different trading days (see Sec. 3).

To assess whether the estimator is able to identify the boundary between different trading days, we calculate the normalized cumulative-sum change-point estimate (see Eq. (C.2)) for each cross-day order flow series. Because the lengths of the intra-day series vary with the number of arrivals and departures each day, the locations of the boundaries between different trading days vary across different cross-day series. Consequently, the boundary between trading days D_i and D_{i+1} does not necessarily lie at the mid point in the relevant cross-day series. We therefore introduce a normalization to enable comparisons between different cross-day series. Given a cross-day series of length N with daily boundary $r^* \in \{2, 3, \dots, N-1\}$, and given a change-point estimator \hat{r}^* for r^* , the *normalized change-point estimator* is

$$\tilde{r} = \begin{cases} (\hat{r}^* - r^*)/r^*, & \text{if } \hat{r}^* \leq r^*, \\ (\hat{r}^* - r^*)/(r - r^*), & \text{otherwise.} \end{cases} \quad (10)$$

Observe that $\tilde{r} \in [-1, 1]$, and $\tilde{r} = 0$ if and only if $\hat{r}^* = r^*$.

In Fig. 9, we plot the empirical cumulative density functions (ECDFs) of \tilde{r} for each of the three currency pairs' cross-day arrival-sign series. The results for the cross-day departure-sign series are qualitatively similar. For the EUR/GBP cross-day series, there is no visible jump in the ECDF near $\tilde{r} = 0$. Therefore, the cumulative-sum change-point estimator performs very poorly at detecting the true locations of the daily boundaries in the cross-day series for this currency pair. For EUR/USD and GBP/USD, there is a small but discernible spike in the ECDFs near to $\tilde{r} = 0$, which indicates that the estimator performs somewhat better for these currency pairs. However, the distribution of the estimator's output across different trading days is still very broad, which suggests that its performance is still rather weak.

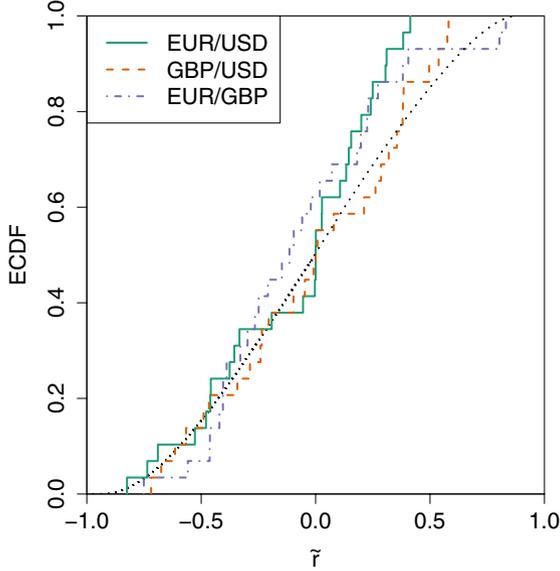


Fig. 9. ECDFs of the normalized cumulative-sum change-point estimator \tilde{r} (see Eq. (10) for the (solid green curve) EUR/USD, (dashed orange curve) GBP/USD, and (dotted-dashed purple curve) EUR/GBP cross-day arrival-sign series. The results for the cross-day departure-sign series are qualitatively similar. The dotted black curve illustrates the estimator’s null distribution for a second-order stationary series with no structural breaks, which we estimate by calculating the ECDF of the normalized cumulative-sum change-point estimates from 100,000 independent series that each consist of 1,000,000 random variables drawn independently and at random from the standard normal distribution.

Together, these results suggest that the apparent long-memory properties of the cross-day series are not strongly influenced by non-stationarities at the daily boundaries.

To test the hypothesis of non-stationarity more formally, we also perform Berkes’ change-point test (see Appendix C) on both the intra-day and cross-day order-flow series. We repeat the test for each of 0, 1, and 2 change points. We show box plots of our results in Fig. 10.

For each of the three currency pairs and for each of 0, 1, and 2 change-points, the test rejects the null hypothesis at the 1% significance level in all intra-day and cross-day arrival-sign and departure-sign series. The results for the intra-day series are again very similar to those for the cross-day series. This result provides strong evidence against the hypothesis that the apparent long-memory properties of the series are actually artifacts caused by non-stationarities. We therefore strongly reject Axioglou and Skouras’ (2011) hypothesis that the apparent long memory in the cross-day series is mostly an

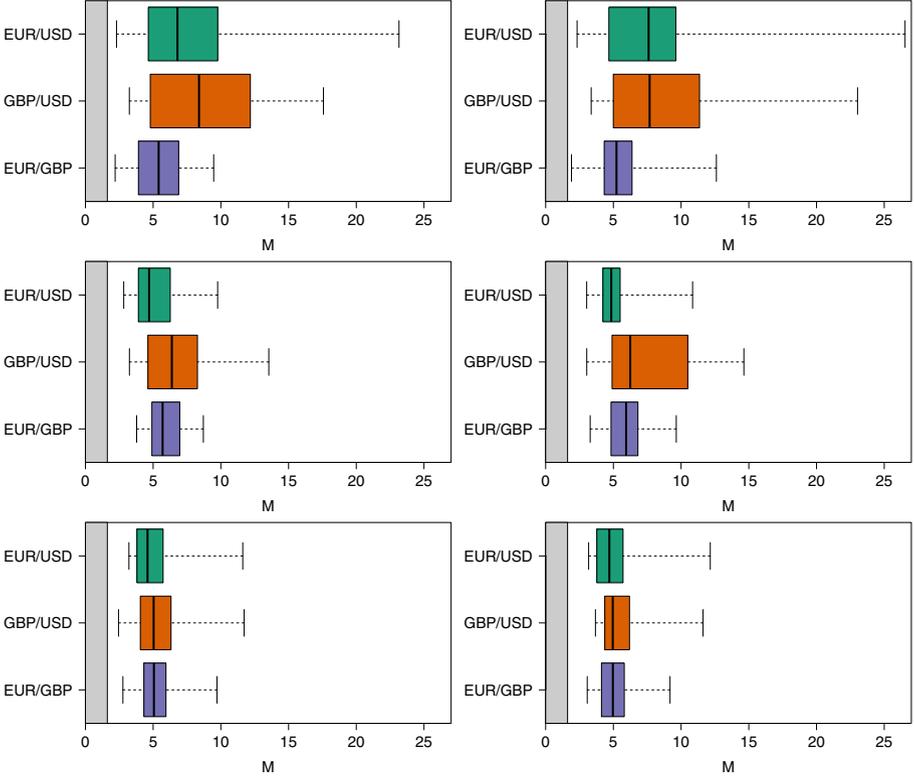


Fig. 10. Box plots of Berkes’ change-point test statistic M (see Appendix C) with (top row) 0 change points, (middle row) 1 change point, and (bottom row) 2 change points, for EUR/USD, GBP/USD, and EUR/GBP arrival-sign series. The left plots show the results for the intra-day series, and the right plots show the results for the cross-day series. The boxes indicate the lower quartile, median, and upper quartile of M , and the whiskers indicate the minimum and maximum of M , across all intra-day or cross-day series. In each case, we use Andrews’ (1991) data-driven plug-in estimator \hat{q} from Eq. (B.10) to calculate M . The light grey shading indicates the critical region for Berkes’ change-point test at the 1% significance level. The results for the cross-day departure-sign series are qualitatively similar.

artifact caused by structural breaks at the daily boundaries, in favor of the alternative hypothesis of true long memory that persists across daily boundaries. We return to this discussion in Sec. 6.

6. Discussion

The results of our statistical tests strongly support the hypothesis that both the arrival-sign and departure-sign series on Hotspot FX exhibit long memory that persists over several thousands of events. For each of the three currency pairs, our calculations suggest that this long memory in order flow

can be characterized by a Hurst exponent of $H \approx 0.7$. The variation in this result — both across the different currency pairs and across the different days in our sample — is small.

Our results raise several interesting points for discussion. As we noted in Sec. 3, several authors have argued that order splitting provides a more plausible explanation for long memory in order flow than does herding. Our results provide further evidence to support this argument. Other empirical studies of long memory in order flow have studied LOBs in which all institutions observe all order flow from all others. On Hotspot FX, by contrast, institutions can only see the order flow that originates from institutions with whom they possess bilateral credit (see Sec. 4). It is therefore reasonable to assume that herding effects are much weaker on Hotspot FX than they are on other platforms. Despite this important difference, our estimates of H on Hotspot FX are very similar to those reported for other platforms (see Sec. 3). We therefore argue that herding plays a minor role in the long-memory properties of order flow and that long-range autocorrelations are instead caused by order-splitting strategies (which are unaffected by the quasi-centralized nature of trade on Hotspot FX).

We turn next to the low-order negative autocorrelation that we observe in the sample ACFs (see Fig. 1). Although this negative autocorrelation is relatively weak, the effect is present for each day in our sample, and we therefore deem it to be a robust statistical property of the data. Interestingly, several other studies of long memory in order flow have not reported this behavior and have instead reported positive autocorrelations at small lags. This raises the question of why our results differ. We believe that the answer to this question lies in the extremely high activity levels in the FX spot market. In a study of order flow on the LSE, Tóth *et al.* (2012) used broker identifier codes to trace which order flow originated from which brokerage. For limit order arrivals, they reported that orders originating from the same broker were positively correlated with each other, but that orders originating from different brokers were negatively correlated with each other. Because of the extremely large number of participants in the FX spot market, an arriving limit order is likely to be followed by many other arriving limit orders from other participants. According to Tóth *et al.*'s findings, this should generate short-range, negative autocorrelations in order flow. At larger lags, however, the strength of this effect diminishes and the long-range autocorrelation effects become more apparent. We conjecture that the short-range negative autocorrelations were not reported in some other studies because this effect is weaker (and, therefore, not observable) in markets with lower activity levels.

It is also interesting to consider why the cumulative-sum change-point estimator (see Eq. C.2) performs much worse at detecting the daily boundaries between intra-day order-sign series in the FX spot market than was reported by Axioglou and Skouras for the LSE (Axioglou and Skouras, 2011). We conjecture that the answer lies in the structural differences between the trading days in these two markets. Trading on the LSE commences at 08:00:00 and ceases at 16:30:00 each day (The London Stock Exchange, 2015). By contrast, trading in the FX spot market occurs 24 h a day. Although we restrict our attention to the peak trading hours of 08:00:00–17:00:00 GMT, the absence of a market-wide closing time has several important consequences for the way that traders act.

First, many financial institutions require that traders unwind their positions (i.e., rebalance their net daily holdings to 0) before the end of each trading day (Lyons, 1995). The LSE market closure at 16:30:00 constitutes a hard deadline by which any traders who seek to unwind their positions must fulfill this goal, even if doing so requires them to trade at unfavorable prices. This may cause the statistical properties of order flow late in the trading day to differ substantially from those early in the trading day, and may therefore result in a structural break at the daily boundaries when concatenating data from different days. In the FX spot market, by contrast, there is no market-wide closing time.

Second, the absence of market opening and closing times in the FX spot market enables traders in different time zones to begin and end their trading days at different times. Hsieh and Kleidon (1996) noted that many traders spend the early part of their trading day assessing the state of the market, the middle part of their trading day performing the majority of their trades, and the late part of their trading day resetting their net inventory to 0. In markets with specified opening and closing times (such as the LSE), all traders progress through this cycle simultaneously. In the FX spot market, by contrast, traders in different time zones can choose the length and timing of their trading days as they wish. The flux of traders from different time zones into and out of the FX spot market may cause the statistical properties of order flow to differ from those in markets where all traders' trading days are aligned by the market opening and closing times.

Third, Axioglou and Skouras (2011) conjectured that most traders on the LSE reassess their trading strategy once per day, while markets are closed, then implement their chosen strategy throughout the next trading day. This is not a plausible description of the actions of traders in the FX spot market, which is always open.

Finally, we return to the question of whether the apparent long memory that we observe is actually an artifact caused by structural breaks. Although we test and strongly reject this hypothesis for zero, one, and two structural breaks, our results do not rule out the possibility of larger numbers of structural breaks. Indeed, it is plausible that intra-day order-sign series contain very large numbers of structural breaks, for reasons such as the arrivals or departures of institutions in the market or the release of macro-economic news. From a modelling perspective, however, models that rely on large numbers of structural breaks that occur at unknown points in time have several important drawbacks, such as being difficult to estimate and of little use for forecasting (see [Appendix C](#)). Moreover, due to their extremely large numbers of parameters, such models suffer a considerable risk of overfitting. By contrast, long-memory models are parsimonious, easy to simulate, and useful for forecasting. We therefore regard long-memory approaches to be more useful than alternative approaches that rely on fitting large numbers of structural breaks to explain the observed autocorrelation properties of the order-sign series.

It is important to note that our results do not rule out the possibility that order-sign series contain *both* long memory and structural breaks. Indeed, the co-existence of these effects could help to explain the differences between the behavior observed in different markets. According to our findings, intra-day order flow in the FX spot market exhibits long memory. Any possible structural breaks appear to have little impact. In equities markets, by contrast, the relatively low number of order arrivals each day makes precise empirical assessment of long memory much more difficult, and structural breaks appear to have a greater impact on the apparent long-memory properties of order flow. It is plausible that both structural breaks and true long memory could coexist in all financial markets and that the relative importance of these effects is different for different assets.

7. Conclusions

In this paper, we have investigated the long memory of order flow in the FX spot market. Due to the extremely high levels of activity on the platform that we study, and in contrast to other empirical studies on this topic, we were able to investigate the long-memory properties of intra-day series without needing to aggregate data from different trading days. For each of the three currency pairs and on each of the trading days that we studied, we found that both arrival-sign and departure-series exhibit long memory, with

a Hurst exponent of $H \approx 0.7$. We also uncovered a negative autocorrelation at shorter lags, which we conjectured was caused by the large number of participants in the FX spot market.

All of our results for data that crosses daily boundaries were similar to those for intra-day data, and we strongly rejected the hypothesis that the apparent long memory of order flow is an artifact caused by structural breaks. We therefore concluded that long memory is a robust statistical property of order flow on Hotspot FX that persists across daily boundaries. We also proposed several possible reasons why our findings differ from those reported by Axioglou and Skouras for the LSE (Axioglou and Skouras, 2011). Further empirical study of data from other markets will help to illuminate these issues further, and is therefore an important topic for future research.

Finally, we note that the existence of long memory in order flow raises an interesting question called the “efficiency paradox” (Farmer *et al.*, 2006): how can return series remain unpredictable given that order flow exhibits long memory? To date, there are two main hypotheses. Some authors (Bouchaud *et al.*, 2004, 2006) have argued that markets reside at a “self-organized critical point” in which liquidity takers cause long-range autocorrelations in order flow that exactly balance the long-range *negative* autocorrelations caused by liquidity providers. Others (Farmer *et al.*, 2006; Lillo and Farmer, 2004) have argued that predictability in order flow is offset by a negative correlation with available liquidity. At present, there is no clear consensus as to which approach best describes the temporal evolution of real markets, and further empirical and theoretical study of this question remains an important and exciting avenue for future research.

Appendix A. Long Memory

In this appendix, we provide a detailed description of long-memory processes. For further discussion of these topics, see Beran (1994).

A.1. Autocorrelation and long memory

Recall from Sec. 2.1 that a second-order stationary time series $\{W_t\} = W_1, W_2, \dots$ with ACF $\rho(k)$ (see Eq. (4)) is said to exhibit long memory if

$$\lim_{N \rightarrow \infty} \sum_{k=-N}^N |\rho(k)| = \infty. \quad (\text{A.1})$$

One way in which a time series can exhibit long memory is if there exists some constant $\alpha \in (0, 1)$ such that $\rho(k)$ decays asymptotically as a power of k :

$$\rho(k) \sim k^{-\alpha} L(k), \quad k \rightarrow \infty, \quad (\text{A.2})$$

where L is a slowly varying function⁶ (Cont, 2005; Lillo and Farmer, 2004; Lillo *et al.*, 2005). Smaller values of α correspond to slower decay of the long-range autocorrelations in $\{W_t\}$ (Bouchaud *et al.*, 2009; Lillo and Farmer, 2004).

A.2. The rescaled-range statistic

Let

$$\bar{W}_k = \frac{1}{k} \sum_{j=t+1}^{t+k} W_j. \quad (\text{A.3})$$

The *rescaled-range statistic* (Mandelbrot and Wallis, 1969b) is the ratio

$$Q(t, k) = \frac{R(t, k)}{S(t, k)}, \quad (\text{A.4})$$

where, for $t, k \in \mathbb{Z}_{>0}$ and for $i \in \{1, 2, \dots, k\}$,

$$R(t, k) = \max_{1 \leq i \leq k} \left[\sum_{j=t+1}^{t+i} (W_j - \bar{W}_k) \right] - \min_{1 \leq i \leq k} \left[\sum_{j=t+1}^{t+i} (W_j - \bar{W}_k) \right] \quad (\text{A.5})$$

and

$$S^2(t, k) = \frac{1}{k} \sum_{j=t+1}^{t+k} (W_j - \bar{W}_k)^2. \quad (\text{A.6})$$

The rescaled-range statistic Q measures the range of partial sums of deviations of the time series $\{W_1, W_2, \dots\}$ from its mean, rescaled by an estimate of its standard deviation (Lo, 1991).

A.3. The Hurst exponent

The following theorem by Mandelbrot (1975) provides a relationship between a time series' long-range autocorrelations and its rescaled-range statistic.

⁶ A function L is *slowly varying* if $\lim_{k \rightarrow \infty} L(zk)/L(k) = 1$ for all $z > 0$.

Theorem 1. *If $\{W_t\}$ is a second-order stationary process such that W_t^2 is ergodic and $t^{-H} \sum_{i=1}^t W_i$ converges weakly to a fractional Brownian motion⁷ with parameter H as $t \rightarrow \infty$, then*

$$k^{-H} Q(t, k) \xrightarrow{d} \eta \quad \text{as } k \rightarrow \infty, \quad (\text{A.7})$$

where η is a non-degenerate random variable and \xrightarrow{d} denotes convergence in distribution.

The constant H is called the *Hurst exponent* of $\{W_t\}$ (Beran, 1994; Hurst, 1951; Mandelbrot and Wallis, 1968, 1969a,b,c). A time series with a Hurst exponent of $H = 1/2$ is a short-memory process. For a long-memory process that satisfies the conditions of this theorem, H is related to α in Eq. (A.2) by

$$H = 1 - \frac{\alpha}{2}, \quad (\text{A.8})$$

and to β in Eq. (B.16) by

$$H = \frac{\beta + 1}{2}. \quad (\text{A.9})$$

Appendix B. Empirical Assessment of Long Memory

In many empirical situations, it is common to observe only a single, finite-length realization $\{w_1, w_2, \dots, w_N\}$ of $\{W_t\}$. If the statistical properties of $\{W_t\}$ are unknown, then estimating the long-memory properties of $\{W_t\}$ from $\{w_1, w_2, \dots, w_N\}$ entails considerable challenges (Beran, 1992, 1994; Mandelbrot and Wallis, 1969d). Most empirical studies employ heuristic methods for this task. The performance of such techniques on empirically observed series varies considerably, so it is common for empirical studies to evaluate the output of several heuristic methods rather than relying on a single estimator. In this appendix, we provide a detailed description of the techniques that we use throughout the paper. For further discussion and comparisons of these techniques, see Taquq *et al.* (1995).

B.1. Sample ACF

For an empirically observed time series $\{w_1, w_2, \dots, w_N\}$, let

$$\bar{w} = \frac{1}{N} \sum_{i=1}^N w_i \quad (\text{B.1})$$

⁷A *fractional Brownian motion* (Mandelbrot *et al.*, 1968) is a Gaussian process $B_H(t)$ with 0 drift that satisfies $B_H(0) = 0$ and $\mathbb{E}[B_H(t)B_H(s)] = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H})$ for some $H \in (0, 1)$.

denote the sample mean, let

$$\hat{\gamma}(k) = \frac{1}{N} \sum_{i=1}^{N-|k|} (w_{i+|k|} - \bar{w})(w_i - \bar{w}) \quad (\text{B.2})$$

denote the sample autocovariance function, and let

$$\hat{\rho}(k) = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} \quad (\text{B.3})$$

denote the sample ACF. It is very difficult to estimate the large- k decay of $\rho(k)$ from $\hat{\rho}$, so direct estimation of the long-memory properties of $\{W_t\}$ from $\hat{\gamma}$ often produces very poor results (Lillo and Farmer, 2004).

B.2. Rescaled-range plots

For a given *block number* $B \in \mathbb{N}$, let

$$G(k) = \left\{ t = \frac{N(i-1)}{B} + 1 \mid i = 1, \dots, B; t+k \leq N \right\}. \quad (\text{B.4})$$

A *rescaled-range plot* (also known as a *pox plot*) (Mandelbrot and Wallis, 1968, 1969b; Teverovsky *et al.*, 1999) is a plot of

$$\bar{R}(k) = \frac{1}{|G(k)|} \sum_{t \in G(k)} Q(t, k) \quad (\text{B.5})$$

versus k on doubly logarithmic axes, where $|G(k)|$ denotes the number of elements in $G(k)$. The slope of a rescaled-range plot for large values of k provides a rough estimate for the Hurst exponent H (Mandelbrot and Wallis, 1969b).

B.3. Lo's modified rescaled-range statistic

Lo (1991) noted that if a time series $\{W_t\}$ is subject to short-range autocorrelations, then the denominator $S(t, k)$ of the rescaled-range statistic $Q(t, k)$ is not a consistent estimator for the standard deviation of $\{W_t\}$. Therefore, an important difficulty in using the rescaled-range statistic to assess the long-memory properties of an empirically observed time series $\{w_1, w_2, \dots, w_N\}$ is that the finite-sample properties of Q are not invariant to short-range dependence. To address this problem, Lo proposed replacing the denominator of Q with the Newey–West⁸ estimator (Newey and West, 1987),

⁸For a suitable choice of q , Newey and West (1987) showed that $\hat{\sigma}(q)$ is a consistent estimator for the standard deviation of $\{W_t\}$, even if $\{W_t\}$ is subject to short-range autocorrelations.

which discounts short-range dependence in $\{W_t\}$ up to a specified lag $q < N$. The parameter q is called the *bandwidth parameter*. Lo's modified rescaled-range statistic (Lo, 1991) is

$$\tilde{Q}(q) = \frac{R(1, N)}{\hat{\sigma}(q)}, \quad (\text{B.6})$$

where

$$\hat{\sigma}^2(q) = \begin{cases} S^2(1, N), & \text{if } q = 0, \\ S^2(1, N) + 2 \sum_{i=1}^q \left(1 - \frac{i}{q+1}\right) \hat{\gamma}(i), & \text{otherwise.} \end{cases} \quad (\text{B.7})$$

Given $\tilde{Q}(q)$, the statistic

$$V(q) = \frac{\tilde{Q}(q)}{\sqrt{N}} \quad (\text{B.8})$$

can be used as a test statistic for the hypothesis test

$$H_0 : \{W_t\} \text{ is a short-memory process,}$$

$$H_1 : \{W_t\} \text{ is a long-memory process.}$$

This hypothesis test is called *Lo's modified rescaled-range test* (Lo, 1991). In the limit $N \rightarrow \infty$, the asymptotic critical region for the test at the 5% significance level is approximately $[0.809, 1.862]$.

Teverovsky *et al.* (1999) remarked that although Lo's modified rescaled-range statistic is a significant improvement over the original rescaled-range statistic, Lo's test can fail to reject H_0 for some time series with long memory. Moreover, they noted that both the size and the power of the test depend on q . The optimal choice of q for $\tilde{Q}(q)$ and $V(q)$ depends on the behavior of the spectral density f of $\{W_t\}$ (Andrews, 1991). If f is unknown (as is usually the case for empirically observed series), then there is no universal rule via which to choose q . In empirical applications, it is therefore customary to calculate $\tilde{Q}(q)$ and $V(q)$ using several different choices of q and/or to calculate a so-called *plug-in estimator* \hat{q} (Andrews, 1991; Axioglou and Skouras, 2011; Lo, 1991) by assuming that f is equal to the spectral density of a specified parametric process. Andrews (1991) derived plug-in estimators for several different parametric processes (including autoregressive, moving-average, and ARMA models).

For our calculations, we use the plug-in estimator for an AR(1) process

$$W_t = \phi W_{t-1} + \varepsilon_t, \quad (\text{B.9})$$

where $\phi \in \mathbb{R}$ is the autocorrelation parameter and ε_t is uncorrelated Gaussian noise. This estimator is given by (Andrews, 1991; Lo, 1991)

$$\hat{q} = \left[\left(\frac{3N}{2} \right)^{1/3} \left(\frac{2\hat{\phi}}{1 - \hat{\phi}^2} \right)^{2/3} \right], \quad (\text{B.10})$$

where $\hat{\phi}$ is the maximum-likelihood estimate of ϕ given $\{w_1, w_2, \dots, w_N\}$ and $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x .

B.4. Detrended fluctuation analysis (DFA)

DFA (Peng *et al.*, 1994) is a technique for estimating the Hurst exponent from an empirically observed series $\{w_1, w_2, \dots, w_N\}$. Let

$$w_i^* = \sum_{j=1}^i w_j, \quad i = 1, 2, \dots, N. \quad (\text{B.11})$$

For a given window length $m \in \mathbb{N}$ such that $m \leq N$, divide $\{w_1^*, w_2^*, \dots, w_N^*\}$ into non-overlapping windows of length m . For each $j \in \{1, 2, \dots, \lfloor N/m \rfloor\}$, label the data points in window j as $y_{1,j}, y_{2,j}, \dots, y_{m,j}$, perform an ordinary least-squares regression to fit a straight line to the m data points in the window, and let $\hat{y}_{i,j}$ denote the value of the regression line at the point $y_{i,j}$ for each $i \in \{1, 2, \dots, m\}$. For each window, calculate the *detrended standard deviation*

$$F_j(m) = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_{i,j} - \hat{y}_{i,j})^2}, \quad (\text{B.12})$$

and then calculate the *length- m mean detrended standard deviation*

$$F(m) = \frac{1}{\lfloor N/m \rfloor} \sum_{j=1}^{\lfloor N/m \rfloor} F_j(m). \quad (\text{B.13})$$

Repeat this process for several logarithmically spaced choices of window length m , and plot $F(m)$ versus m using doubly logarithmic axes. Identify the smallest value m_{\min} such that the plot is approximately straight for all $m \geq m_{\min}$. The DFA estimate of H is given by the slope of the best-fit line for $m \geq m_{\min}$.

B.5. Log-periodogram regression

The long-memory properties of $\{W_t\}$ can also be characterized by the behavior of its spectral density (Beran, 1994)

$$f(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma(k) e^{-ik\lambda} \quad (\text{B.14})$$

in the limit $\lambda \rightarrow 0$. If there exists a constant $l \in \mathbb{R}$ such that

$$f(\lambda) \rightarrow l \quad \text{as } \lambda \rightarrow 0, \quad (\text{B.15})$$

then $\{W_t\}$ is a short-memory process. If, by contrast, there exists a constant $\beta \in (0, 1)$ such that

$$f(\lambda) \sim \lambda^{-\beta} \quad \text{as } \lambda \rightarrow 0, \quad (\text{B.16})$$

then $\{W_t\}$ is a long-memory process. Larger values of β correspond to slower decay of the long-range autocorrelations in $\{W_t\}$ (Beran, 1994).

Estimating the behavior of $f(\lambda)$ close to 0 provides an alternative approach to estimating the Hurst exponent H . Let

$$\lambda_{j,N} = \frac{2\pi j}{N}, \quad j \in \left\{1, 2, \dots, \left\lfloor \frac{N-1}{2} \right\rfloor\right\} \quad (\text{B.17})$$

denote the *Fourier frequencies* of $\{w_1, w_2, \dots, w_N\}$, and let

$$I(\lambda_{j,N}) = \frac{1}{2\pi N} \left| \sum_{t=1}^N (w_t - \bar{w}) e^{-it\lambda_{j,N}} \right|^2 \quad (\text{B.18})$$

denote the corresponding *periodogram*. The slope of an ordinary least-squares regression of $\log(I(\lambda_{j,N}))$ onto $\log(\lambda_{j,N})$ for small $\lambda_{j,N}$ is an estimator for $-\beta$, and it is therefore an estimator for H (Beran, 1994; Lillo and Farmer, 2004; Taqqu *et al.*, 1995).

Despite the attractiveness of its computational simplicity, log-periodogram regression suffers from a substantial practical drawback (Beran, 1994): there is no universal rule for choosing the number c of Fourier frequencies with which to perform the regression. The scaling in Eqs. (B.15) and (B.16) only holds for $\lambda \rightarrow 0$, so choosing an overly large c leads to large bias, but choosing an overly small c leads to high variance. Robinson (1994) derived an expression for the optimal choice of c to minimize the mean squared error of the estimated cumulative spectral distribution function, but the optimal choice turns out to depend on the unknown value of H . Therefore, Robinson's expression does not provide a method for choosing c for an empirically

observed time series. Instead, most empirical studies use one of two rules of thumb: $c = \sqrt{N}$ (Geweke and Porter-Hudak, 1983) or $c = 0.1 \times (N/2)$ (Taqqu *et al.*, 1995). Despite their widespread use, neither of these rules are based on rigorous derivation or optimization.

Appendix C. Long Memory Versus Non-Stationarity

A key difficulty with assessing the long-memory properties of empirically observed time series is that many estimation techniques can produce similar output for a non-stationary series (e.g., a series with a monotonic trend (Bhattacharya *et al.*, 1983) or change in mean (Giraitis *et al.*, 2001; Granger and Hyung, 2004)) as they do for a stationary series with long memory (Rea *et al.*, 2009; Taqqu *et al.*, 1995; Xu *et al.*, 2005). Several authors have thus argued that the apparent long memory reported by empirical studies of financial time series is an artifact caused by non-stationarities (Axioglou and Skouras, 2011; Berkes *et al.*, 2006; Liu, 2000; Mikosch and Stărică, 2002).

Disentangling the statistical properties of a time series with long memory and a time series with non-stationarities is a difficult task, and the choice of whether to model such time series using a long-memory model or a non-stationary model often depends on the desired application. Long-memory models are parsimonious, straightforward to simulate, and there exist many techniques that require only mild assumptions to estimate their parameters from data (Beran, 1994). Non-stationary models can illuminate important features of an underlying series (such as the locations and frequency of structural breaks) that are not addressed by long-memory models, but they typically require the inclusion of either a large number of parameters (which can lead to over-fitting) or latent parameters (which can be difficult to estimate from data).

Many standard tests for non-stationarities in an empirically observed series have low power in the presence of long memory (Diebold and Rudebusch, 1991; Hassler and Wolters, 1994), and there is no universal test that is able to determine whether the apparent long-memory properties of an empirically observed time series are an artifact caused by some unknown form of non-stationarity (Granger and Hyung, 2004). However, Berkes *et al.* (2006) developed a hypothesis test that can help to distinguish between long memory and a specific type of non-stationarity that they called *structural breaks*.

Definition. The time series $\{Z_t\}$ is a *short-range dependent series with a structural break at time r^** if there exists a real-valued, second-order

stationary, short-memory process $\{\zeta_t\}$ and a constant $\mu^* \neq 0$ such that

$$Z_t = \begin{cases} \zeta_t, & t \leq r^*, \\ \mu^* + \zeta_t, & t > r^*. \end{cases} \quad (\text{C.1})$$

Given a finite-length empirical observation $\{z_1, z_2, \dots, z_N\}$ of a time series $\{Z_t\}$, *Berkes' change-point test* is the hypothesis test

H_0 : $\{Z_t\}$ is a short-memory process with one structural break,⁹

H_1 : $\{Z_t\}$ is a long-memory process.

Berkes' test uses the so-called *cumulative-sum change-point estimator* (Berkes *et al.*, 2006)

$$\hat{r}^* = \min \left\{ r : \max_{1 \leq j \leq N} \left| \sum_{i=1}^j z_i - \frac{j}{N} \sum_{i=1}^N z_i \right| = \left| \sum_{i=1}^r z_i - \frac{r}{N} \sum_{i=1}^N z_i \right| \right\}. \quad (\text{C.2})$$

For a given $q < N$, let

$$T^{(1)} = \frac{1}{\hat{\sigma}_{(1)}^2(q) \sqrt{\hat{r}^*}} \max_{1 \leq i \leq \hat{r}^*} \left| \sum_{j=1}^i z_j - \frac{i}{\hat{r}^*} \sum_{j=1}^{\hat{r}^*} z_j \right|,$$

$$T^{(2)} = \frac{1}{\hat{\sigma}_{(2)}^2(q) \sqrt{N - \hat{r}^*}} \max_{\hat{r}^* \leq i \leq N} \left| \sum_{j=\hat{r}^*}^i z_j - \frac{i - \hat{r}^*}{N - \hat{r}^*} \sum_{j=\hat{r}^*}^N z_j \right|,$$

where $\hat{\sigma}_{(1)}(q)$ and $\hat{\sigma}_{(2)}(q)$ are the values of $\hat{\sigma}(q)$ from Eq. (B.7) for the $\{z_1, z_2, \dots, z_{\hat{r}^*}\}$ and $\{z_{\hat{r}^*+1}, z_{\hat{r}^*+2}, \dots, z_N\}$ series, respectively. The test statistic for Berkes' test is

$$M = \max\{T^{(1)}, T^{(2)}\}. \quad (\text{C.3})$$

In the limit $N \rightarrow \infty$, the asymptotic critical value of M at the 1% significance level is about 1.72 (Berkes *et al.*, 2006).

Acknowledgements

We thank Julius Bonart, Jean-Philippe Bouchaud, Rama Cont, J. Doyne Farmer, Austin Gerig, Ben Hambly, and Gabriele La Spada for useful discussions. We thank Hotspot FX for providing the data for this project. We also thank Terry Lyons, Rich Plummer-Powell, and Justin Sharp for

⁹For brevity, we restrict our discussion to Berkes' change-point test for a single structural break. However, this framework can be extended to test the null hypothesis of any specified number of structural breaks against the alternative hypothesis of true long memory. For details, see Berkes *et al.* (2006).

technical support. M. D. G. and S. D. H. thank the Oxford-Man Institute of Quantitative Finance, and M. D. G. thank EPSRC and the James S. McDonnell Foundation for supporting this research. M. D. G. completed part of this work while at the University of Oxford.

References

- Andrews, D. W. K., 1991, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, *Econometrica* 59, 817–858.
- Axioglou, C., and S. Skouras, 2011, Markets Change Every Day: Evidence from the Memory of Trade Direction, *Journal of Empirical Finance* 18(3), 423–446.
- Baillie, R. T., 1996, Long Memory Processes and Fractional Integration in Econometrics, *Journal of Econometrics* 73(1), 5–59.
- Bank for International Settlements, 2010, Triennial Central Bank Survey: Report on Global Foreign Exchange Market Activity in 2010, Technical report, Bank for International Settlements, Available at: <http://www.bis.org/publ/rpfx10t.pdf>.
- Beran, J., 1992, Statistical Methods for Data with Long-Range Dependence, *Statistical Science* 7, 404–416.
- Beran, J., 1994, *Statistics for Long-Memory Processes*, Chapman and Hall, New York, NY, USA.
- Berkes, I., L. Horváth, P. Kokoszka, and Q. M. Shao, 2006, On Discriminating Between Long-Range Dependence and Changes in Mean, *The Annals of Statistics* 34, 1140–1165.
- Bhattacharya, R. N., V. K. Gupta, and E. Waymire, 1983, The Hurst Effect Under Trends, *Journal of Applied Probability* 20, 649–662.
- Biais, B., P. Hillion, and C. Spatt, 1995, An Empirical Analysis of the Limit Order Book and the Order Flow in the Paris Bourse, *The Journal of Finance* 50(5), 1655–1689.
- Booth, G. G., and F. R. Kaen, 1979, Gold and Silver Spot Prices and Market Information Efficiency, *Financial Review* 14(1), 21–26.
- Booth, G. G., F. R. Kaen, and P. Koveos, 1982, R/S Analysis of Foreign Exchange Rates Under Two International Monetary Regimes, *Journal of Monetary Economics* 10(3), 407–415.
- Bouchaud, J. P., J. D. Farmer, and F. Lillo, 2009, How markets slowly digest changes in supply and demand, In T. Hens, and K. R. Schenk-Hoppé, (eds.) *Handbook of Financial Markets: Dynamics and Evolution*, pp. 57–160, North-Holland, Amsterdam, The Netherlands.
- Bouchaud, J. P., Y. Gefen, M. Potters, and M. Wyart, 2004, Fluctuations and Response in Financial Markets: The Subtle Nature of ‘Random’ Price Changes, *Quantitative Finance* 4(2), 176–190.
- Bouchaud, J. P., J. Kockelkoren, and M. Potters, 2006, Random Walks, Liquidity Molasses and Critical Response in Financial Markets, *Quantitative Finance* 6(2), 115–123.
- Chakraborti, A., I. M. Toke, M. Patriarca, and F. Abergel, 2011, Econophysics Review I: Empirical Facts, *Quantitative Finance* 11(7), 991–1012.

- Chatfield, C., 2000, *Time-Series Forecasting*, Chapman and Hall, Boca Raton, FL, USA.
- Cont, R., 2005, Long-range dependence in financial markets, in J. Lévy-Véhel, and E. Lutton, (eds.), *Fractals in Engineering*, pp. 159–179, Springer, London, UK.
- Cont, R., 2011, Statistical Modeling of High-Frequency Financial Data, *IEEE Signal Processing Magazine* 28(5), 16–25.
- Cont, R., and J. P. Bouchaud, 2000, Herd Behavior and Aggregate Fluctuations in Financial Markets, *Macroeconomic Dynamics* 4(2), 170–196.
- Cont, R., M. Potters, and J. P. Bouchaud, 1997, Scaling in Stock Market Data: Stable Laws and Beyond, in B. Dubrulle, F., Graner, and D. Sornette, (eds.), *Les Houches Workshop*, pp. 75–85, Springer, New York, NY, USA.
- Diebold, F. X., and G. D. Rudebusch, 1991, On the Power of Dickey–Fuller Tests Against Fractional Alternatives, *Economics Letters* 35(2), 155–160.
- EBS, 2011, EBS Global Access User Guide. Retrieved 14 April 2015 from <https://emea.ebsspot.com/SHARED/HELP/userguide.pdf>.
- Ellul, A., C. W. Holden, P., Jain, and R. Jennings, 2003, A Comprehensive Test of Order Choice Theory: Recent Evidence from the NYSE, LSE Discussion paper 471, Available at <http://eprints.lse.ac.uk/24896/>.
- Farmer, J. D., A. N. Gerig, F., Lillo, and S. Mike, 2006, Market Efficiency and the Long-Memory of Supply and Demand: Is Price Impact Variable and Permanent or Fixed and Temporary? *Quantitative Finance* 6(2), 107–112.
- Gerig, A. N., 2007, A Theory for Market Impact: How Order Flow Affects Stock Price, Ph.D. thesis, University of Illinois at Urbana-Champaign, Champaign, IL, USA.
- Geweke, J., and S. Porter-Hudak, 1983, The Estimation and Application of Long-Memory Time-Series Models, *Journal of Time Series Analysis* 4(4), 221–238.
- Giraitis, L., P. S., Kokoszka, and R. Leipus, 2001, Testing for Long Memory in the Presence of a General Trend, *Journal of Applied Probability* 38(4), 1033–1054.
- Gopikrishnan, P., V. Plerou, L. A. N. Amaral, M. Meyer, and H. E. Stanley, 1999, Scaling of the Distribution of Fluctuations of Financial Market Indices, *Physical Review E* 60(5), 5305–5316.
- Gould, M. D., M. A. Porter, and S. D. Howison, 2016, Quasi-Centralized Limit Order Books, arXiv:1310.8387. Available at: <https://arxiv.org/abs/1502.00680>.
- Gould, M. D., M. A. Porter, S. Williams, M., McDonald, D. J., Fenn, and S. D. Howison, 2013, Limit Order Books, *Quantitative Finance* 13(11), 1709–1742.
- Granger, C. W. J., and N. Hyung, 2004, Occasional Structural Breaks and Long Memory with an Application to the S&P 500 Absolute Stock Returns, *Journal of Empirical Finance* 11(3), 399–421.
- Greene, M. T., and B. D. Fielitz, 1977, Long-Term Dependence in Common Stock Returns, *Journal of Financial Economics* 4(3), 339–349.
- Hasbrouck, J., 1988, Trades, Quotes, Inventories, and Information, *Journal of Financial Economics* 22(2), 229–252.
- Hassler, U., and J. Wolters, 1994, On the Power of Unit Root Tests Against Fractional Alternatives, *Economics Letters* 45(1), 1–5.
- Hsieh, D. A., and A. W. Kleidon, 1996, Bid-Ask Spreads in Foreign Exchange Markets: Implications for Models of Asymmetric Information, in J. Frankel, G.,

- Galli, and A. Giovannini, (eds.), *The Microstructure of Foreign Exchange Markets*, pp. 41–72, The University of Chicago Press, Chicago, IL, USA.
- Hurst, H., 1951, Long Term Storage Capacity of Reservoirs, *Transactions of the American Society of Civil Engineers* 116, 770–799.
- Knight Capital Group 2015a, Retrieved 14 April 2015 from http://www.hotspotfx.com/download/userguide/HSFX/HSFX_UserGuide-wrapper.html.
- Knight Capital Group 2015b, Retrieved 14 April 2015 from <http://www.hotspotfx.com/overview/index.jsp>.
- Knight Capital Group 2015c, Retrieved 14 April 2015 from http://www.hotspotfx.com/products/hotspot_volumes.jsp.
- LeBaron, B., and R. Yamamoto, 2007, Long-Memory in an Order-Driven Market, *Physica A* 383(1), 85–89.
- Lillo, F., and J. D. Farmer, 2004, The Long Memory of the Efficient Market, *Studies in Nonlinear Dynamics and Econometrics* 8(3), 1–33.
- Lillo, F., S. Mike, and J. D. Farmer, 2005, Theory for Long Memory in Supply and Demand. *Physical Review E* 71(6), 066122.
- Liu, M., 2000, Modeling Long Memory in Stock Market Volatility, *Journal of Econometrics* 99(1), 139–171.
- Lo, A. W., 1991, Long-Term Memory in Stock Market Prices, *Econometrica* 59(5), 1279–1313.
- Lyons, R. K., 1995, Tests of Microstructural Hypotheses in the Foreign Exchange Market, *Journal of Financial Economics* 39(2), 321–351.
- Mandelbrot, B. B., 1963, The Variation of Certain Speculative Prices, *The Journal of Business* 36(4), 394–419.
- Mandelbrot, B. B., 1975, Limit Theorems on the Self-Normalized Range for Weakly and Strongly Dependent Processes, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 31(4), 271–285.
- Mandelbrot, B. B., and J. W. Van Ness, 1968, Fractional Brownian Motions, Fractional Noises and Applications, *SIAM Review* 10(4), 422–437.
- Mandelbrot, B. B., and J. R. Wallis, 1968, Noah, Joseph, and Operational Hydrology, *Water Resources Research* 4(5), 909–918.
- Mandelbrot, B. B., and J. R. Wallis, 1969a, Computer Experiments with Fractional Gaussian Noises: Part 1, Averages and Variances, *Water Resources Research* 5(1), 228–241.
- Mandelbrot, B. B., and J. R. Wallis, 1969b, Computer Experiments with Fractional Gaussian Noises: Part 2, Rescaled Ranges and Spectra, *Water Resources Research* 5(1), 242–259.
- Mandelbrot, B. B., and J. R. Wallis, 1969c, Robustness of the Rescaled Range R/S in the Measurement of Noncyclic Long Run Statistical Dependence, *Water Resources Research* 5(5), 967–988.
- Mandelbrot, B. B., and J. R. Wallis, 1969d, Some Long-Run Properties of Geophysical Records, *Water Resources Research* 5(2), 321–340.
- Mantegna, R. N., and H. E. Stanley, 1999, *An Introduction to Econophysics: Correlations and Complexity in Finance*, Cambridge University Press, Cambridge, UK.

- Mike, S., and J. D. Farmer, 2008, An Empirical Behavioral Model of Liquidity and Volatility, *Journal of Economic Dynamics and Control* 32(1), 200–234.
- Mikosch, T., and C. Stárică, 2002, Long-Range Dependence Effects and ARCH Modeling, in P. Doukhan, G. Oppenheim, and M. Taqqu (eds.), *Theory and Applications of Long-Range Dependence*, pp. 439–459. Birkhäuser Boston, New York, NY, USA.
- Newey, W. K., and K. D. West, 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–708.
- Peng, C. K., S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, 1994, Mosaic Organization of DNA Nucleotides, *Physical Review E*, 49(2), 1685–1689.
- Rea, W., L. Oxley, M., Reale, and J. Brown, 2009, Estimators for Long-Range Dependence: An Empirical Study, arXiv:0901.0762. Available at: <http://arxiv.org/abs/0901.0762>.
- Robinson, P. M., 1994, Rates of Convergence and Optimal Spectral Bandwidth for Long Range Dependence, *Probability Theory and Related Fields* 99(3), 443–473.
- Roşu, I., 2009, A Dynamic Model of the Limit Order Book, *Review of Financial Studies* 22(11), 4601–4641.
- Taqqu, M. S., V. Teverovsky, and W. Willinger, 1995, Estimators for Long-Range Dependence: An Empirical Study, *Fractals* 3(4), 785–798.
- Taylor, S. J., 2008, *Modelling Financial Time Series*, World Scientific Publishing, Singapore.
- Teverovsky, V., M. S. Taqqu, and W. Willinger, 1999, A Critical Look at Lo's modified R/S Statistic, *Journal of Statistical Planning and Inference* 80(1), 211–227.
- The London Stock Exchange, 2015, SETS. Retrieved 14 April 2015 from <http://www.lseg.com/areas-expertise/our-markets/london-stock-exchange/equities-markets/trading-services/domestic-trading-services/sets>.
- Thomson–Reuters, 2016, Thomson Reuters FXAll. Available at: <http://financial.thomsonreuters.com/en/products/trading-capabilities/foreign-exchange-markets/fxall-electronic-trading-platform.html>.
- Tóth, B., Z. Eisler, F., Lillo, J., Kockelkoren, J. P., Bouchaud, and J. D. Farmer, 2012, How Does the Market React to Your Order Flow? *Quantitative Finance* 12(7), 1015–1024.
- Tóth, B., Y. Lempérière, C., Deremble, J., De Lataillade, J. Kockelkoren, and J. P. Bouchaud, 2011, Anomalous Price Impact and the Critical Nature of Liquidity in Financial Markets, *Physical Review X* 1(2), 021006.
- Tóth, B., I. Palit, F., Lillo, and J. D. Farmer, 2015, Why is Equity Order Flow so Persistent? *Journal of Economic Dynamics and Control* 51, 218–239.
- Wyart, M., J. P. Bouchaud, J. Kockelkoren, M., Potters, and M. Vettorazzo, 2008, Relation Between Bid-Ask Spread, Impact and Volatility in Order-Driven Markets, *Quantitative Finance* 8(1), 41–57.
- Xu, L., P. C. Ivanov, K. Hu, Z. Chen, A. Carbone, and H. E. Stanley, 2005, Quantifying Signals With Power-Law Correlations: A Comparative Study of

M. D. Gould, M. A. Porter & S. D. Howison

Detrended Fluctuation Analysis and Detrended Moving Average Techniques,
Physical Review E 71(5), 051101.

Yeo, W. Y., 2008, Serial Correlation in the Limit Order Flow: Causes and Impact,
Working Paper, SSRN eLibrary ID 971400. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=971400