How Do the Lengths of the Lead Lag Time between Stocks Evolve? Tick-by-tick Level Measurements across Two Decades

Bing Anderson
California Polytechnic State University, USA
bianders@calpoly.edu
https://orcid.org/0000-0003-1870-1000

Received: 06 August 2022 / Revised: 28 October 2022 / Accepted: 4 November 2022 / Published online: 25 November 2022

ABSTRACT

There has been an extraordinary decrease in order execution time on stock exchanges in the past two decades. A related question is whether there has been a similar reduction in orders of magnitude for the lengths of the lead lag time between stocks. If the answer is affirmative, and the lengths of the lead lag time have long fallen below the human reaction time, algorithms have taken over information diffusion from one stock to another. Otherwise, humans continue to be in authority. In this study, the lengths of the lead lag time within pairs of stocks of large US companies are estimated using the Hayashi-Yoshida estimator, for each year from 2000 to 2022. We first construct stock pairs, with each pair containing two stocks from the same industrial sector. The median length of the lead lag time for each year shows a general trend of decline over time. From 2000 to 2005, the median lengths are a few seconds. By 2021 and 2022, they are less than 10 milliseconds. We also study a second construct in which stock pairs are randomly formed, but each pair contains stocks from two different sectors. The median length of the lead lag time for each year shows a decline over time, similar to the first construct. Overall, the lengths of the lead lag time in the second construct are not remarkably longer than those in the first construct. This shows that being in the same sector, at the tick-by-tick level, is not an important factor in determining the length of the lead lag time between stocks.

JEL Classification: G12; G14; G19

Keywords: Hayashi-Yoshida estimator, price discovery, cross-correlation, statistical arbitrage, high-frequency trading.

1. INTRODUCTION

The past two decades have seen a phenomenal increase in order execution speed, and a decrease in order execution time, on stock exchanges. In 2000, the average execution time was approximately 20 seconds on the New York Stock Exchange, whereas by 2010, it reduced to around 1 second (Haldane, 2011). It is now represented in microseconds on the fastest exchanges.

DOI: 10.7172/2353-6845.jbfe.2022.2.4

The question is whether there has been a similar reduction in orders of magnitude for the lengths of the lead lag time between stocks. For example, a short time after news on Chevron Corporation impacts its stock price, investors may realize that the news has repercussions on other stocks in the energy sector. Therefore, the length of the lead lag time between stock prices is a reflection of the time required to digest information. However, the duration of the lead lag time could potentially shed light on the nature of the information processing mechanism. If we indeed see orders of magnitude of lead lag time length reduction over the past two decades, the information is processed via automation by computers and algorithms. Alternatively, this process likely involves human decision making by traders and investors.

In this study, we use tick-by-tick data on eight pairs of significant US stocks from eight distinct industrial sectors and quantify the length of the lead lag time within each pair for each year between 2000 and 2022 in order to examine the evolution of lead lag periods between individual stocks over the past two decades. To study the impact of being in the same sector on the length of the lead lag time, we also randomly reassign the stocks into pairs from different sectors and remeasure the lengths of the lead lag time across the two decades.

This work provides substantial additions to the body of knowledge that already exists. The majority of lead lag literature focuses on the direction of the lead lag connection rather than the lengths of the lead lag period. How long is the lead lag duration on average for big liquid U.S. stocks? To the best of our knowledge, there has been no answer to this question in the current literature prior to this paper.

This research provided an answer for each year from 2000 until 2022. We are not aware of any prior investigation utilizing tick-by-tick data throughout such a wide time span. As it turns out, lead lag durations have a tendency to decrease with time. This finding has practical implications for academics who compare the outcomes of lead lag studies conducted over different time periods.

Eventually, this paper also answered the question posed above on whether the magnitude of the lead lag time points to information processing via humans versus via computers. Computers dominate, but occasionally humans may still play a role. This and other findings in this study have practical implications for practitioners. For instance, when a mutual fund wishes to liquidate a big position, does employing human traders optimize the net liquidation proceeds, or does this approach leave money on the table for high-frequency trading algorithms? When a hedge fund employs high-frequency lead lag statistical arbitrage algorithms, how frequently may it discover “anomalies” that represent potential profit opportunities? A peek at the three tables in this article and a count of the frequencies of “anomalies” in the tables may provide a general notion. Additionally, should these lead lag arbitrage algorithms focus on stock pairings from the same sector or from different sectors? In this paper, we have studied both the intra-sector and inter-sector cases.

The rest of the paper is organized as follows. Section 2 provides a review of the literature on lead lag studies. Section 3 introduces the data and the research methodology based on the Hayashi-Yoshida estimator. Section 4 presents and discusses the results. Section 5 concludes.

2. LITERATURE REVIEW

The lead lag relationship has been examined for different parts of the financial markets.

Hasbrouck (2003) analyzes the correlation between futures and exchange-traded funds (ETF) for stock indices and finds that the E-mini NASDAQ 100 and S&P 500 index futures have the highest price discovery. Price discovery is “the impounding of new information into the security price” (Hasbrouck, 1995). Most of the price discovery studies point out the direction of the lead lag relationship, but they rarely assess the exact lead lag duration. This is not surprising given that neither of the two primary approaches for price discovery research, the information share

DOI: 10.7172/2353-6845.jbfe.2022.2.4

methodology of Hasbrouck (1995) and the common factor component share methodology of Gonzalo and Granger (1995), is designed to measure the length of the lead lag time.

There is an extensive body of literature on lead lag research from the perspective of price discovery. Sapp (2002) studies price discovery in the spot foreign exchange market. Chakravarty, Gulen, and Mayhew (2004) investigate the role of option markets in stock price discovery. Mizrach and Neely (2008) examine how the futures markets contribute to the US Treasury price discoveries. There is also work on whether New York or London leads in the discovery of gold prices (Lucey, Larkin, & O'Connor, 2013) and on the stock price co-movements between Europe and the United States (Ben Ameur et al., 2018). More recently, Chen et al. (2021) find that regular index futures in Taiwan contribute more to price discovery than the mini index futures.

Recently, cryptocurrencies have become a focus of lead lag research. For example, Ji et al. (2019) examine six large cryptocurrencies, and conclude that return shocks from Bitcoin and Litecoin affect the rest the most. Using a wavelet approach, Mensi et al. (2019) find that Bitcoin leads Dash, Monero, and Ripple in the time frequency space. Corbet et al. (2018) conclude that "Bitcoin prices affect both Ripple (28.37%) and Lite (42.3%), but Ripple and Lite have limited influence on Bitcoin", and that Bitcoin clearly leads other cryptocurrencies in price movements. Ciaian, Rajcaniova, and Kancs (2018) investigate Bitcoin and 16 alternative cryptocurrencies, and find that Bitcoin price shocks impact the prices of 15 out of the 16 alternative cryptocurrencies in the short run. Using the VAR-based approach of Diebold and Yilmaz (2009), Koutmos (2018) analyzes 18 major cryptocurrencies and concludes that Bitcoin is the leader of return spillover to the rest. Most of the cryptocurrency lead lag studies use daily data.

Wang et al. (2022) investigate the lead lag link between the VIXs of individual stocks and the S&P 500 VIX. Xu and Yin (2017) probe stock market index volatility and how it relates to the index ETF volumes. Tolikas (2018) study the lead lag connection between the stock and the bond markets. Ballester and González-Urteaga (2020) examine the lead lag relationship between the sovereign credit default swap market and the stock market.

There are a few distinguishing factors between this study and the previous literature. In the bulk of existing studies, the lengths of the lead lag time are not measured. This paper is intended to quantify the lengths of the lead lag time. The majority of published works utilize daily or other low-frequency data. This study utilizes tick-by-tick data, sometimes known as ultra-high-frequency data, as there is no higher frequency than tick-by-tick data. The final distinctive feature of this study is that it analyzes the evolution of the lengths of the lead lag time over a period of 23 years.

3. DATA AND METHODOLOGY

Tick-by-tick data are purchased from Tick Data LLC, an authorized distributor of the New York Stock Exchange (NYSE) TAQ data. In comparable research, it is typical to utilize quotes rather than trade prices (see, for example, Huth & Abergel, 2014; Anderson, 2016). As explained in Anderson (2016), the possibility of the bid-ask bounce effect in trade prices is one argument for using quotes. According to Blume and Goldstein (1997), the NYSE dominates other U.S. exchanges in terms of initiating quote revisions and displaying the best quote prices. Only quotes from the NYSE during its regular trading hours are utilized in this study. The average of the bid and ask prices, or the mid-quote, is used as the observed price. When multiple quotes arrive exactly at the same time, their average is utilized.

Due to the vast amount of data involved, it is common for market microstructure research employing tick data to focus on a period of a few months. Hasbrouck (2003), for instance, examines the sample period between March 2000 and May 2000. Huth and Abergel (2014), for example, utilize data from March 2010 to May 2010 for their research. We use the month of May

DOI: 10.7172/2353-6845.jbfe.2022.2.4

in each year from 2000 to 2022. We pick 16 large liquid US stocks in eight different sectors to form eight pairs, with each pair being two stocks in the same sector: Coca-Cola (ticker symbol: KO) and Pepsi (PEP), Verizon (VZ) and AT&T (T), IBM (IBM) and HP (HPQ), Walmart (WMT) and Home Depot (HD), Raytheon (RTX) and Boeing (BA), JPMorgan (JPM) and Morgan Stanley (MS), Chevron (CVX) and Exxon Mobile (XOM), Pfizer (PFE) and Merck(MRK).

Using tick data for a study, a massive volume of data to work with is only a small portion of the difficulty. The fundamental issue is the asynchronous arrival of observations. The majority of time series econometrics tools deal with data that arrives at regularly spaced time intervals: each month, each day, each hour, etc. Casting high frequency asynchronous tick data into regularly spaced time intervals causes problems in measuring the lengths of the lead lag time (Finucane, 1999; Zhang, 2011).

Although it is possible to integrate the two asynchronous time series using a regression-based method (Finucane, 1999), we opt to employ a technique based on the cross-correlation estimator of Hayashi and Yoshida (2005). A significant advantage of the Hayashi-Yoshida estimator is that it reveals unambiguously the direction of the lead lag relationship as well as the length of the lead lag time.

In lead lag research, Huth and Abergel (2014), Dao, McGroarty, and Urquhart (2018), and Schei (2019) have utilized the Hayashi-Yoshida estimator. In 2018, bitcoin transactions on Bitfinex were 12 seconds ahead of those on Kraken, according to one of Schei's (2019) findings. To validate our implementation of the Hayashi-Yoshida estimator, we conduct the same analysis on the 2018 data and are able to duplicate Schei's findings to the second.

Below is a summary of how the length of the lead lag time is estimated. Consider stochastic differential equations to characterize the price dynamics of stocks P and Q:

$$dP_t = \mu_t^P dt + \sigma_t^P dB_t^P \tag{1}$$

$$dQ_t = \mu_t^Q dt + \sigma_t^Q dB_t^Q \tag{2}$$

where B_t^P and B_t^Q are standard Brownian motions with $d\langle B^P, B^Q\rangle_t = \rho_t dt$.

The stock price processes P_t and Q_t are observed at asynchronous sampling times, at $0 = t_0^P \le t_1^P \le \cdots \le t_{m-1}^P \le t_m^P = T$ for P and $0 = t_0^Q \le t_1^Q \le \cdots \le t_{n-1}^Q \le t_n^Q = T$ for Q. The sampling times should be independent of the prices P_t and Q_t.

Hayashi and Yoshida (2005) prove that a consistent estimator of the covariance between P and Q is

$$\sum\nolimits_{i,j} \delta_i^P \delta_j^Q 1_{\{C_{ij} \neq \varnothing\}} \tag{3}$$

with $C_{ij} = \left(t_{i-1}^P, t_i^P\right] \cap \left(t_{j-1}^Q, t_j^Q\right]$, $\delta_i^P = P_{t_i^P} - P_{t_{i-1}^P}$, $\delta_j^Q = Q_{t_j^Q} - Q_{t_{j-1}^Q}$. The indicator function $1_{\{C_{ij} \neq \varnothing\}}$ takes the value of 1 whenever the sampling intervals of P and Q, $\left(t_{i-1}^P, t_i^P\right]$ and $\left(t_{j-1}^Q, t_j^Q\right]$, have any overlap. Given the covariance estimator, the correlation ρ can be estimated as

$$\hat{\rho}(P,Q) = \frac{\sum_{i,j} \delta_i^P \delta_j^Q 1_{\{C_{ij} \neq \varnothing\}}}{\sqrt{\sum_i \left(\delta_i^P\right)^2 \sum_j \left(\delta_j^Q\right)^2}} \tag{4}.$$

Hoffmann, Rosenbaum, and Yoshida (2014) demonstrate that a lag time can be introduced to the time stamps of Q. The lagged Q can be correlated with the original P. Among all the different

DOI: 10.7172/2353-6845.jbfe.2022.2.4

lengths of the lag time, the one that maximizes the correlation is the actual length of the lead lag time between the two stocks.

Mathematically, let

$$Q_t^*(\nu) = Q_{t+\nu} \quad (5).$$

Then, P leads Q by

$$\text{argmax}_{-T<\nu<T}\hat{\rho}(P, Q^*(\nu)) \quad (6).$$

In our implementation, we test every second between -120 and 120 seconds for the lead lag duration. The time stamps of the original data are expressed in seconds from 2000 to 2005, milliseconds from 2006 to 2015 , microseconds from 2016 to 2018, and nanoseconds from 2019 to 2022. When the estimated length of the lead lag time is small, and the time stamp provides sufficient resolution, we repeat the estimation process with a step size of one order of magnitude finer. For example, when the time stamps are in units of milliseconds, and the first round of estimation with the [-120, 120] window yields 1 second for the length of the lead lag time, we run the estimation again from -12 seconds to 12 seconds with a smaller step size of 0.1 seconds. Similarly, if the situation warrants, we could run another round from -1.2 seconds to 1.2 seconds with a step size of 0.01 seconds, and so on.

4. RESULTS

Table 1 displays the estimation outcomes. For example, the cell in the very first row and the first column indicates that WMT leads HD by -8 seconds in May 2000, which means that HD actually leads WMT by 8 seconds.

There are significant variations in the lead lag duration both across different years and across different stock pairs. In the first column of Table 1, for instance, the lengths of the lead lag time between WMT and HD exhibit a general downward trend from 2000 to 2022. For both 2009 and 2011, the lengths of the lead lag time are measured in tens of milliseconds. Nonetheless, for the intervening year, 2010, the length of the lead lag time is a stunning 23 seconds.

Table 1
The lengths of the lead lag time between stocks within the same sector

Year	WMT leads HD	CVX leads XOM	MRK leads PFE	PEP leads KO	RTX leads BA	MS leads JPM	HPQ leads IBM	VZ leads T	Median of |Time|
2000	-8s	1s	8s	-2s	-7s	-3s	9s	2s	5s
2001	5s	-9s	6s	1s	-39s	6s	-4s	3s	5.5s
2002	-4s	-4s	3s	-2s	-1s	0s	-17s	0s	2.5s
2003	-2s	-4s	0s	1s	-3s	1s	4s	0s	1.5s
2004	1s	1s	0s	0s	0s	1s	1s	1s	1s
2005	1s	1s	-2s	0s	1s	2s	1s	0s	1s
2006	300ms	550ms	240ms	115ms	520ms	29ms	1.3s	570ms	410ms
2007	-90ms	70ms	-74ms	-100ms	-120ms	0ms	-70ms	-140ms	82ms
2008	-41ms	16ms	100ms	95ms	23ms	-6ms	-85ms	80ms	60.5ms

DOI: 10.7172/2353-6845.jbfe.2022.2.4

Table 1 – continued

Year	WMT leads HD	CVX leads XOM	MRK leads PFE	PEP leads KO	RTX leads BA	MS leads JPM	HPQ leads IBM	VZ leads T	Median of |Time|
2009	23ms	0ms	140ms	104ms	-22ms	56ms	38ms	140ms	47ms
2010	23s	-9ms	-10s	9ms	7ms	-16ms	-39ms	-9ms	12.5ms
2011	-49ms	1ms	85ms	37ms	-5ms	25ms	-70ms	8ms	31ms
2012	2ms	2ms	31ms	28ms	-12ms	33ms	-390ms	70ms	29.5ms
2013	-1ms	2ms	97ms	23ms	0ms	2ms	-58ms	-3ms	2.5ms
2014	5ms	0ms	11ms	4ms	1ms	1ms	-630ms	-1ms	2.5ms
2015	0ms	1ms	1ms	1ms	-5ms	-1ms	-22ms	0ms	1ms
2016	-3ms	1.3ms	2.8ms	16ms	-2.6ms	-3.2ms	-210ms	2.2ms	2.9ms
2017	-3ms	2ms	14ms	5ms	-2.3ms	-1.2ms	-1.05s	-2.8ms	2.55ms
2018	42ms	2.1ms	7ms	-680ms	8ms	-2.3ms	-120ms	5.3ms	7.5ms
2019	-0.6ms	8ms	9.8ms	300ms	-1.6ms	-1.3ms	-130ms	5.1ms	6.55ms
2020	2.3ms	7ms	29ms	130ms	-1.8ms	-1.3ms	-54ms	4ms	5.5ms
2021	1.2ms	0.4ms	1.9ms	13ms	0.1ms	-0.9ms	-43ms	-1.1ms	1.15ms
2022	-8.3ms	0.9ms	1.1ms	18ms	-0.7ms	-0.1ms	-9ms	-0.3ms	1ms

Note: The length of each lead lag time is measured in seconds(s) or milliseconds (ms).

Source: Author's own calculation.

Table 2
The lengths of the lead lag time between stocks in different sectors

Year	WMT leads IBM	CVX leads PFE	MRK leads JPM	KO leads XOM	RTX leads HD	MS leads T	HPQ leads PEP	VZ leads BA	Median of |Time|
2000	0s	17s	3s	-3s	18s	2s	-3s	11s	3s
2001	0s	3s	2s	6s	-3s	5s	-16s	-1s	3s
2002	9s	2s	-4s	6s	-3s	-2s	-5s	2s	3.5s
2003	2s	1s	1s	0s	0s	0s	1s	-1s	1s
2004	1s	1s	1s	1s	1s	0s	1s	1s	1s
2005	1s	2s	0s	2s	1s	1s	0s	0s	1s
2006	63ms	810ms	105ms	1.7s	720ms	1.3s	390ms	60ms	555ms
2007	-10ms	-230ms	-99ms	24ms	27ms	0ms	130ms	-18ms	25.5ms
2008	-170ms	27ms	-9ms	-54ms	180ms	52ms	-68ms	12ms	53ms
2009	0ms	380ms	80ms	-35ms	8ms	46ms	-59ms	-1ms	40.5ms
2010	-62ms	-49ms	-10ms	-3ms	44ms	17ms	13ms	6ms	15ms
2011	-60ms	93ms	9ms	13ms	11ms	69ms	-64ms	-55ms	57.5ms
2012	-80ms	220ms	52ms	-8ms	60ms	55ms	-41ms	-78ms	57.5ms
2013	-8ms	28ms	35ms	-32ms	1ms	12ms	-87ms	-19ms	3.5ms

DOI: 10.7172/2353-6845.jbfe.2022.2.4

Table 2 – continued

Year	WMT leads IBM	CVX leads PFE	MRK leads JPM	KO leads XOM	RTX leads HD	MS leads T	HPQ leads PEP	VZ leads BA	Median of |Time|
2014	-9ms	15ms	6ms	-11ms	14ms	-22ms	-47ms	-41ms	14.5ms
2015	-25ms	4ms	0ms	-1ms	1ms	-1ms	0ms	-16ms	1ms
2016	-31ms	29ms	2.4ms	-7.9ms	1.5ms	-1.9ms	-15ms	-130ms	11.45ms
2017	-18ms	58ms	0.52ms	1.2ms	0.12ms	62s	-0.8s	-24ms	21ms
2018	1.2ms	49ms	-3.3ms	-1.4ms	51ms	9.2ms	-64s	-5ms	7.1ms
2019	-1ms	29ms	-0.97ms	-11ms	1.6ms	12ms	-110s	-26ms	11.5ms
2020	-1.1ms	60ms	-0.76ms	0.8ms	-20ms	1.5ms	-0.5s	-5.4ms	3.45ms
2021	1.3ms	7.4ms	1.3ms	0.47ms	-1.2ms	0.39ms	-30s	-7.2ms	1.3ms
2022	0.028ms	5.1ms	-1.9ms	1.3ms	-14ms	28ms	-29ms	-17ms	9.55ms

Note: The length of each lead lag time is measured in seconds(s) or milliseconds (ms).

Source: Author's own calculation.

One scenario that could explain why this occurs is that when a large mutual fund or hedge fund wants to acquire or sell a substantial position in a stock, the process must often take at least a few months. If a large fund wants to acquire HD and employs human traders to do so, and if WMT is one of the stocks the human traders monitor, WMT can easily lead HD by 23 seconds for the month. The volume data may provide some evidence to support this claim. For May 2009, there are 1,629,219 WMT observations and 1,380,688 HD observations. For the month of May 2011, WMT has 2,331,075 and HD has 2,754,698. In May 2010, however, the amount of observations for WMT is 4,314,827 and for HD it is 5,400,357, significantly more than in either 2009 or 2011.

The median length of each year's lead lag times provides a more accurate depiction of the evolution of the length of the lead lag time between stocks throughout the years, considering the large variations in lead lag lengths between different years and different stock pairs. For each year, we calculate the median of the absolute values of the eight lead lag durations for the eight stock pairs. This median is displayed in the final column of Table 1.

It is evident by inspecting the median column that the lengths of the lead lag time reduce over time. From 2000 to 2005, they are a few seconds. In both 2004 and 2005, the median lead lag duration is just 1 second. Before that, in 2003, a mere 1.5 seconds. In 2006, it is a few hundred milliseconds. After that, from 2007 to 2012, tens of milliseconds. The median lengths of the lead lag time for all of the years after 2012 are less than 10 milliseconds.

To answer the question posed at the beginning of this paper, the decline of the median lengths of the lead lag time below the human reaction time over the years demonstrates that information is primarily processed by computers and that trades placed by algorithms are the most influential factors in determining the lead lag duration. Some may question the extent to which algorithms can interpret and process news. Scholtus, van Dijk, and Frijns (2014) examine high-frequency trading following the release of US macroeconomic news. They discover that a delay of just 0.3 seconds can already have a major impact. Humans are incapable of reading the news and executing a trade in 0.3 seconds. Clearly, computers are automatically processing and understanding the news and trading appropriately. What they investigate is not the length of the lead lag time between individual stocks per se, but the 300 milliseconds they found using data from 2009–2011 are not discordant with those provided in Table 1 in terms of magnitude.

DOI: 10.7172/2353-6845.jbfe.2022.2.4

Next, we analyze to what extent the lengths of the lead lag time depend on the two stocks in a pair belonging to the same industrial sector. According to Hou (2007), the lead lag effect is predominant within the same industry. Hou (2007) did not utilize tick level data, however. Will the length of the lead lag time grow considerably if two stocks in a pair are not from the same industry? To answer this question, we re-assign the stock pairs at random, ensuring that no two stocks in a given pair belong to the same industrial sector.

Table 3
The lengths of the lead lag time between stocks within the financial sector

Year	GS leads AIG	WFC leads PGR	MS leads TRV	BAC leads ALL	AXP leads BK	C leads PNC	JPM leads STT	COF leads USB	Median of |Time|
2000	-87s	4s	-3s	-6s	-1s	10s	39s	29s	8s
2001	10s	16s	38s	-11s	3s	10s	5s	-50s	10.5s
2002	-2s	58s	13s	10s	-3s	-5s	1s	-1s	4s
2003	1s	1s	25s	-1s	0s	1s	2s	-5s	1s
2004	-1s	0s	0s	1s	0s	0s	2s	5s	0.5s
2005	0s	-1s	-1s	0s	0s	1s	1s	2s	1s
2006	-1.02s	-240ms	-630ms	-2s	-18ms	-490ms	-1.6s	-390ms	560ms
2007	109ms	-110ms	-340ms	-320ms	110ms	22ms	2ms	800ms	110ms
2008	52ms	-160ms	-89ms	-200ms	74ms	-220ms	-77ms	24ms	83ms
2009	15s	-20ms	59ms	-270ms	-66ms	-24ms	-40ms	52ms	55.5ms
2010	-9ms	-19ms	2ms	-110ms	6ms	-27ms	-8ms	12ms	10.5ms
2011	13ms	-87ms	-104ms	-120ms	9ms	-73ms	-50ms	52ms	62.5ms
2012	8ms	-23ms	-74ms	-1.09s	53ms	-37ms	-34ms	46ms	41.5ms
2013	14ms	-33ms	0ms	-14ms	38ms	-20ms	-12ms	1ms	14ms
2014	61ms	3ms	-11ms	-19ms	12ms	-47ms	-24ms	-1ms	15.5ms
2015	1ms	-2ms	-1ms	-740ms	1ms	-1ms	-4ms	1ms	1ms
2016	15ms	-1.5ms	-49ms	88s	0.4ms	-15ms	-5.2ms	9.3ms	12.15ms
2017	-4.3ms	-10.5ms	-34ms	-220ms	59ms	-3.5ms	-20ms	2.1ms	15.25ms
2018	-2.7ms	-8.3ms	-1.5ms	-190ms	1.5ms	-2.7ms	-0.2ms	1.8ms	2.25ms
2019	-0.3ms	-18ms	-17ms	-130ms	3.6ms	-19ms	0ms	0.8ms	10.3ms
2020	9.7ms	-5.1ms	-33ms	-21ms	14ms	-14ms	-5.7ms	18ms	14ms
2021	10.7ms	8.5ms	-0.7ms	-11ms	-1.3ms	-13ms	3.4ms	0.9ms	5.95ms
2022	-0.4ms	-9ms	-2.1ms	-12ms	10.1ms	-7.1ms	-1ms	3.1ms	5.1ms

Note: Each length of the lead lag time is measured in seconds(s) or milliseconds (ms).

Source: Author's own calculation.

The results are displayed in Table 2. From 2000 to 2005, the median lengths of the lead lag time are a few seconds. In particular, in 2003, 2004, and 2005, the median lengths of the lead lag time are just 1 second. In 2006, it is a few hundred milliseconds, just like in Table 1. After that, there are a greater number of years in which the lead lag durations are tens of milliseconds compared to Table 1. Eventually, the median lengths of the lead lag time fall below 10 milliseconds. However,

DOI: 10.7172/2353-6845.jbfe.2022.2.4

the first instance of this occurs in 2015, as opposed to 2013 in Table 1. The overall declining trend is identical to that shown in Table 1.

In 3 of the 23 years examined, Tables 1 and 2 have the same median lengths of the lead lag time. In 8 years, Table 2's median length of the lead lag time is less than Table 1. In 12 years, Table 2's median length of the lead lag time is more than Table 1. In general, the length of the lead lag time between two stocks from different sectors is not significantly longer than that between two stocks from the same sector.

To further validate the results presented in Table 1 and Table 2, we next focus on a particular sector. Sixteen large, liquid stocks are chosen from the financial sector: American International Group (AIG), Allstate (ALL), American Express (AXP), Bank of America (BAC), Bank of New York Mellon (BK), Citigroup (C), Capital One (COF), Goldman Sachs (GS), JPMorgan (JPM), Morgan Stanley (MS), Progressive (PGR), PNC Financial Services (PNC), State Street (STT), Travelers (TRV), US Bancorp (USB), and Wells Fargo (WFC). The stocks are randomly divided into 8 pairs for lead lag analyses. The results are presented in Table 3. Table 3 confirms the findings from Tables 1 and 2. The median lengths of the lead lag time have a similar decline over the years. They are in seconds from 2000 to 2005, in hundreds of milliseconds for 2006 and 2007, in tens of milliseconds after that, till eventually in a few milliseconds for the most recent years.

5. CONCLUSION

For large liquid US stocks, the median length of the lead lag time within a pair, whether the pair is from the same sector or not, is at a few seconds even as back as the year 2000. As the years progress, the median lead lag duration drops to eventually a few milliseconds by 2022.

This demonstrates that the information diffusion from one stock to another occurs, mainly via computers and algorithms, rather than human insights and human analyses. As the trade execution times on stock exchanges decrease, computers become faster, and algorithms become more sophisticated, even a few hundred milliseconds are eventually too long to prevent statistical arbitrage from taking advantage of the lead lag connection. That is why the length of the lead lag time has to continue to drop, and eventually drops to a few milliseconds by 2022.

As seen by the evolution of the median length of the lead lag time throughout the years, the efficient market hypothesis is alive and largely accurate, notwithstanding the possibility of local or brief deviations. According to Easley, de Prado, and O'Hara (2012), by 2009, high-frequency trading accounts for nearly two-thirds of the US stock trading volume. The findings of this article undoubtedly support this conclusion. Any breach of the efficient market hypothesis that is not local nor transitory will be identified by computers, exploited for profit, and finally eradicated since all possible gains from the inefficiency have been harvested.

In investigating the length of the lead lag time between an equity index and its futures, a 1987 study revealed a lead lag duration of up to 45 minutes (Kawaller, Koch, & Koch, 1987), but a 1992 study discovered a length of the lead lag time of 15 minutes or less (Chan, 1992). We are convinced that the days of measuring the length of the lead lag time in minutes are forever gone, and we now understand that research on the length of the lead lag time undertaken for different eras cannot be directly compared, because lead lag duration tends to decrease as the years progress.

As stated in the literature review, cryptocurrencies have attracted the interest of scholars in recent years. The bulk of previous research on cryptocurrency lead lag relationship focuses on the direction of the connection as opposed to the length of the lead lag time. Most rely on daily or other low-frequency data. A direction for future research is to measure the lengths of cryptocurrency lead lag time and how they evolve over the years using tick-by-tick data.

DOI: 10.7172/2353-6845.jbfe.2022.2.4

References

Anderson, B. (2016). Stock price leads and lags before the golden age of high-frequency trading. *Applied Economics Letters*, *23*(3), 212–216. https://doi.org/10.1080/13504851.2015.1066481

Ballester, L., & González-Urteaga, A. (2020). Is there a connection between sovereign CDS spreads and the stock market? Evidence for European and US returns and volatilities. *Mathematics*, *8*(10), 1667. https://doi.org/10.3390/math8101667

Ben Ameur, H., Jawadi, F., Louhichi, W., & Idi Cheffou, A. (2018). Modeling international stock price comovements with high-frequency data. *Macroeconomic Dynamics*, *22*, 1875–1903. https://doi.org/10.1017/s1365100516000924

Blume, M. E., & Goldstein, M. A. (1997). Quotes, order flow, and price discovery. *The Journal of Finance*, *52*(1), 221–244. https://doi.org/10.1017/s1365100516000924

Chakravarty, S., Gulen, H., & Mayhew, S. (2004). Informed trading in stock and option markets. *The Journal of Finance*, *59*, 1235–1257. https://doi.org/10.1111/j.1540-6261.2004.00661.x

Chan, K. (1992). A further analysis of the lead-lag relationship between the cash market and stock index futures market. *Review of Financial Studies*, *5*, 123–152. https://doi.org/10.1093/rfs/5.1.123

Chen, Y. L., Lee, Y. H., Chou, R. K., & Chang, Y. K. (2021). Arbitrage trading and price discovery of the regular and mini Taiwan stock index futures. *Journal of Futures Markets*, *41*, 1–23. https://doi.org/10.1002/fut.22192

Ciaian, P., Rajcaniova, M., & Kancs, D. (2018). Virtual relationships: Short- and long-run evidence from Bitcoin and altcoin markets. *Journal of International Financial Markets Institutions & Money*, *52*, 173–195. https://doi.org/10.1016/j.intfin.2017.11.001

Corbet, S., Meegan, A., Larkin, C., Lucey, B., & Yarovaya, L. (2018). Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Economics Letters*, *165*, 28–34. https://doi.org/10.2139/ssrn.3070288

Dao, T. M., McGroarty, F., & Urquhart, A. (2018). Ultra-high-frequency lead–lag relationship and information arrival. *Quantitative Finance*, *18*(5), 725–735. https://doi.org/10.1080/14697688.2017.1414484

Diebold, F. X., & Yalmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *The Economic Journal*, *119*, 158–171. https://doi.org/10.3386/w13811

Easley, D., de Prado, M. M. L., & O'Hara, M. (2012). Flow toxicity and liquidity in a high-frequency world. *Review of Financial Studies*, *25*, 1457–1493. https://doi.org/10.1093/rfs/hhs053

Finucane, T. J. (1999). A new measure of the direction and timing of information flow between markets. *Journal of Financial Markets*, *2*, 135–151. https://doi.org/10.1016/s1386-4181(98)00010-x

Gonzalo, J., & Granger, C. (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business and Economic Statistics*, *13*(1), 27–35. https://doi.org/10.1017/ccol052179207x.013

Haldane, A. G. (2011). *The race to zero* [Speech at the International Economic Association Sixteenth World Congress]. Retrieved July 17, 2022, from https://www.bankofengland.co.uk/speech/2011/the-race-to-zero-speech-by-andy-haldane

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *Journal of Finance*, *50*, 1175–1199. https://doi.org/10.1111/j.1540-6261.1995.tb04054.x

Hasbrouck, J. (2003). Intraday price formation in U.S. equity index markets. *Journal of Finance*, *58*(6), 2375–2399. https://doi.org/10.2139/ssrn.252304

Hayashi, T., & Yoshida, N. (2005). On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli*, *11*(2), 359–379. https://doi.org/10.3150/bj/1116340299

Hoffmann, M., Rosenbaum, M., & Yoshida, N. (2013). Estimation of the lead-lag parameter from non-synchronous data. *Bernoulli*, *19*(2), 426–461. https://doi.org/10.3150/11-bej407

Hou, K. (2007). Industry information diffusion and the lead-lag effect in stock returns. *Review of Financial Studies*, *20*(4), 1113–1138. https://doi.org/10.2139/ssrn.463005

Huth, N., & Abergel, F. (2014). High frequency lead-lag relationships – Empirical facts. *Journal of Empirical Finance*, *26*, 41–58. https://doi.org/10.1016/j.jempfin.2014.01.003

Ji, Q., Bouri, E., Lau, C. K. M. & Roubaud, D. (2019). Dynamic connectedness and integration in cryptocurrency markets. *International Review of Financial Analysis*, *63*, 257–272. https://doi.org/10.1016/j.irfa.2018.12.002

Kawaller, I. G., Koch, P.D., & Koch, T. W. (1987). The temporal price relationship between S&P 500 futures and the S&P 500 index. *Journal of Finance*, *42*, 1309–1329. https://doi.org/10.2307/2328529

Koutmos, D. (2018). Return and volatility spillovers among cryptocurrencies. *Economics Letters*, *173*, 122–127. https://doi.org/10.1016/j.econlet.2018.10.004

Lucey, B. M., Larkin, C., & O'Connor, F. A. (2013). London or New York: Where and when does the gold price originate? *Applied Economics Letters*, *20*, 813–817. https://doi.org/10.2139/ssrn.2161905

Mensi, W., Rehman, M. U., Al-Yahyaee, K. H., Al-Jarrah, I. M. W., & Kang, S. H. (2019). Time frequency analysis of the commonalities between Bitcoin and major cryptocurrencies: Portfolio risk management implications. *North American Journal of Economics and Finance*, *48*, 283–294. https://doi.org/10.1016/j.najef.2019.02.013

DOI: 10.7172/2353-6845.jbfe.2022.2.4

Mizrach, B., & Neely, C. J. (2008). Information shares in the US Treasury market. *Journal of Banking and Finance*, *32*, 1221 – 1233. https://doi.org/10.20955/wp.2005.070

Sapp, S. G. (2002). Price leadership in the spot foreign exchange market. *Journal of Financial and Quantitative Analysis*, *37*, 425–448. https://doi.org/10.2307/3594987

Schei, B. N. (2019). *High-frequency lead-lag relationships in the Bitcoin market* [Copenhagen Business School Thesis]. Retrieved July 17, 2022, from https://kryptografen.no/wp-content/uploads/2019/06/High-Frequency-Lead-Lag-Relationships-in-The-Bitcoin-Market.pdf

Scholtus, M., van Dijk, D., & Frijns, B. (2014). Speed, algorithmic trading, and market quality around macroeconomic news announcements. *Journal of Banking and Finance*, *38*, 89–105. https://doi.org/10.2139/ssrn.2174901

Tolikas, K. (2018). The lead-lag relation between the stock and the bond markets. *The European Journal of Finance*, *24*(10), 849–866. *https://doi.org/10.1080/1351847x.2017.1340320*

Wang, Z., Bouri, E., Ferreira, P., Shahzad, S. J. H., & Ferrer, R. (2022). A grey-based correlation with multi-scale analysis: S&P 500 VIX and individual VIXs of large US company stocks. *Finance Research Letters*, *48*, 102872. https://doi.org/10.1016/j.frl.2022.102872

Xu, L., & Yin, X. K. (2017). Exchange traded funds and stock market volatility. *International Review of Finance*, *17*(4), 525–560. https://doi.org/10.1111/irfi.12121

Zhang, L. (2011). Estimating covariation: Epps effect, microstructure noise. *Journal of Econometrics*, *160*(1), 33–47. https://doi.org/10.2139/ssrn.885438

DOI: 10.7172/2353-6845.jbfe.2022.2.4