Journal of Forecasting, *J. Forecast.* **33**, 391–408 (2014)
Published online 20 June 2014 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/for.2298

Ultra-High-Frequency Algorithmic Arbitrage Across International Index Futures

HAMAD ALSAYED AND FRANK MCGROARTY*
Southampton Management School, University of Southampton, UK

ABSTRACT

We show that persistent lead–lag relationships spanning mere fractions of a second exist in all three possible pairings of the S&P 500, FTSE 100 and DAX futures contracts. These relationships exhibit clear intraday patterns which help us to forecast mid-quote changes in lagging contracts with directional accuracy in excess of 85%. A simple algorithmic trading strategy exploiting these relations yields economically significant profits which are robust to market impact costs and the bid–ask spread. We find that price slippage and infrastructure costs are our most important limits to arbitrage. Our results support the impossibility of EMH view that informational inefficiencies incentivize arbitrageurs to eliminate mispricings. Copyright © 2014 John Wiley & Sons, Ltd.

KEY WORDS lead–lag relationships; futures markets; Hayashi–Yoshida cross-correlation estimator; statistical arbitrage

INTRODUCTION

Employing a recent advance in the statistical measurement of lead/lag relationships (Hayashi and Yoshida, 2005; Hoffmann *et al.*, 2013), this paper's principal contribution is to uncover the existence of sub-second price disequilibria which occur across international index futures, and which can be exploited via electronic trading algorithms.

Our work is motivated by evidence of increasing international financial integration in recent years (Ayuso and Blanco, 2001; Kearney and Lucey, 2004; Evans and Hnatkovska, 2014). If, on the one hand, markets were perfectly integrated, efficient and frictionless, then returns on closely related securities would exhibit perfect simultaneity and contemporaneous correlation, so as to preclude the possibility of cross-market arbitrage (De Jong and Nijman, 1998). On the other hand, there is wide evidence that frictions impede the flow of information across markets (Lo and MacKinlay, 1990) and give rise to lead–lag patterns.[1] Now, if the price adjustment process is not instantaneous, then we ask: Do lead–lag patterns across segmented markets induce cross-market predictability? Does information about future price adjustments admit the possibility of arbitrage?

Ours is the first paper to provide answers to these questions within the important but overlooked context of international stock index futures. Specifically, we focus on lead/lag relations in the three pairings between the S&P 500, FTSE 100 and DAX futures contracts. While previous works focus on the relation between futures contracts and their underlying stock indices, the vast majority highlight factors such as trading costs, price staleness and illiquidity in the index constituents which render spot–futures arbitrage infeasible (see, for example, Brennan and Schwartz, 1990). In contrast, futures contracts are highly liquid instruments with low upfront capital requirements and trading costs. Therefore, we suggest that pairs of futures contracts constitute an ideal setting under which to closely scrutinise the lead/lag relationship.

We complement the extant literature in four ways. First, we employ the Hayashi–Yoshida (Hayashi and Yoshida, 2005, hereafter HY) cross-correlation estimator to measure the speed of price adjustment. The HY estimator readily allows for irregular and non-synchronous quote arrival times without the need for coarse resampling. This feature allows us to examine microscopic lead–lag patterns, which is particularly useful as we find that the majority of price adjustments take place at sub-second intervals. Second, we employ Best Bid and Offer (BBO) quote data, time-stamped to the millisecond.[2] BBO quotes form the top level of the Central Limit Order Book, thus constituting a continuous price series free of stale quotes. The use of ultra-high-frequency data is crucially important, because the speed of price adjustment is likely to be high in electronically traded markets. Third, we place emphasis on both the forecasting accuracy and arbitrage returns in assessing the evidence against cross-market efficiency. Finally, we document the risks and costs which constitute the greatest limits to arbitrage in this market setting, and which differ from the limits to arbitrage that are usually acknowledged.

* Correspondence to: Frank McGroarty, Southampton Management School, University of Southampton, Highfield, Southampton, SO17 1BJ, UK. E-mail: f.j.mcgroarty@soton.ac.uk

[1] Examples include the S&P 500 futures versus spot in Kawaller *et al.* (1987), the Major Market Index futures versus spot in Stoll and Whaley (1990) and the FTSE 100 futures versus spot in Brooks *et al.* (2001).

[2] 1 millisecond is one-thousandth of a second.

Our results show evidence of highly significant lead/lag relationships between the three pairs. On average, the S&P 500 leads the FTSE 100 and DAX contracts by 98 and 349 milliseconds respectively, consistent with the notion that international price discovery originates in the USA. The FTSE 100 leads the DAX contract by 33 milliseconds on average.[3] We find some evidence of bidirectional causality in the FTSE–DAX pair, but very little between either transatlantic pair. Further, the FTSE–DAX pair is more strongly correlated than either transatlantic pair, in line with the notion of imperfect international integration (Jorion and Schwartz, 1986).

The lead/lag relations across all pairs exhibit strong intraday seasonality: correlations among all pairs peak around daily economic announcements at 13:30 GMT, and fall sharply when the FTSE 100 and DAX cash markets close at 16:30 GMT, rebounding in evening trade. The S&P 500 contract increases its lead over both FTSE 100 and DAX contracts at the announcement of data, at 14:30 GMT when the S&P 500 cash market opens and at the European close. It is important to note that none of these results is the product of statistical artefacts arising from differences in liquidity or quote arrival frequency.

Based on simple threshold triggers, mid-quote moves of leading contracts forecast subsequent moves of lagging contracts with a directional accuracy of over 85%. This accuracy relates directly to the sensitivity of the threshold: the more conservative the threshold, the 'cleaner' is the trigger signal and the more accurate the forecast.

Finally, a simple algorithmic trading exercise provides evidence against auto-efficiency across international markets. Leading mid-quote moves are used to generate buy/sell signals in lagging contracts, and the returns aggregated net of all costs. Realistic data transmission times are allowed for. Profits are economically and statistically significant—a strategy that trades only five contracts per trading signal generates aggregate profits of around GBP 100,000 per month. However, returns are sensitive to the risk of price slippage: around half the profits disappear if 20% of trades execute at the next best price. Further, the windows of arbitrage opportunity are narrow, with the most profitable trades rarely existing for more than 300 milliseconds. This suggests that technology costs are a significant consideration for arbitrageurs in this setting.

Overall, we conclude that information flows across international markets are not instantaneous, but rather exhibit brief delays. Importantly, these delays last long enough, and induce pricing anomalies large enough, to compensate arbitrageurs for appropriating pricing disequilibria. Our results accord with the Grossman and Stiglitz (1976, 1980) suggestion that temporary disequilibria incentivize arbitrageurs to correct pricing anomalies.

RELATED LITERATURE

Introduced in February 1982 by the Kansas City Board of Trade, the Value Line contract was the world's first index futures contract, followed shortly by the S&P 500 contract (Gulen and Mayhew, 2000). As futures and other derivative contracts become more pervasive, research exploring the link between these and their underlying markets, as well as across markets, becomes increasingly focal. Our work is related to two broad streams of literature which constitute a substantial body of work documenting inter-market pricing relationships.

The first stream explores price discovery and the temporal pricing relationship between futures contracts and the underlying cash index. The seminal research of Zeckhauser and Niederhoffer (1983) documents the possibility that futures prices can predict spot prices.[4] The early work involving intraday data of Kawaller *et al.* (1987) provides evidence that US futures prices lead spot prices with a time lag of around 45 minutes. The authors attribute this lag to the inertia of stock trading relative to futures trading, which implies that the futures market contributes most to price discovery. Similarly, Herbst *et al.* (1987) and Stoll and Whaley (1990) document that futures prices lead spot prices by around 8 minutes, whereas Wang and Wang (2001) show that price discovery is bidirectional when volatility is high.

The spot–futures relationship has also been explored for a number of different countries. Wahab and Lashgari (1993), Abhyankar (1998), and Brooks *et al.* (2001) and Brooks and Garrett (2002) analyse UK spot–futures data, Tse (1995) explores the Japanese spot–futures relationship, while the case of Greece is examined in Kenourgios (2004) and Andreou and Pierides (2008). These works largely confirm the notion that futures prices lead spot prices, suggesting that price discovery takes place in the futures market. But what is the significance of these pricing relationships?

This question is explored in the second stream of literature to which our work belongs. On the one hand, an important implication of a lead–lag relationship is the potential for cross-market return predictability and arbitrage. On the other hand, Brooks *et al.* (2001) point out that, while futures markets are capable of responding to new information quickly, the cash index can only fully respond once every constituent stock price updates. An important question then arises: Is a persistent lead–lag relationship evidence of market inefficiency?

While works such as Figlewski (1984, 1985), Brennan and Schwartz (1990), Thomas (2002) and Richie *et al.* (2008) and Cummings and Frino (2011) document evidence of significant disequilibria in the spot–futures relation,

[3] To illustrate just how short these time horizons are, Wikipedia lists the average time for a human eyeblink as between 100 and 400 milliseconds.
[4] The term *spot* refers to the underlying index cash market—e.g. the constituent stocks of an index.

there are two commonalities in these works: first, the nature of the pricing relation examined therein is not a temporal one based on lead–lag effects, but rather a mispricing approach based on a *cost of carry* model; second, disequilibria in these works are frequently attributed to transaction costs, market (im)maturity and liquidity effects that give rise to disequilibria while precluding arbitrage.

The literature exploring arbitrage based on lead–lag relationships is relatively sparse. Stoll and Whaley (1990) suggest that, although futures generally lead spot prices, the effect is not unidirectional, making spot–futures arbitrage difficult. Brooks *et al.* (2001) compare several forecasting models and demonstrate that the FTSE 100 futures contract can achieve over 65% accuracy when forecasting the FTSE 100 index. However, when a trading strategy based on this is tested, profits are not robust to trading costs.

Our paper relates to these works by placing emphasis on the predictability of returns and the profitability of arbitrage. However, unlike these works, we pursue this theme in an international setting. Indeed, one major commonality linking these works is a focus on inter-market relationships within the same country. Now, recent advances in communications technology have contributed to the integration of similar but otherwise fragmented markets. One direct consequence of this integration is a trend towards the equalization of expected returns across global markets—a single 'relevant' event is able to move global stock market indices jointly (Eiteman *et al.*, 1994; Medeiros *et al.*, 2009).

There are several works which explore temporal relationships across international markets (Eun and Shim, 1989; Hamao *et al.*, 1990; Abhyankar *et al.*, 1997; Antoniou *et al.*, 2003; Innocenti *et al.*, 2011). However, few focus explicitly on the link between lead–lag effects and arbitrage, none employs high-frequency data and none focuses on futures contracts. So why is doing so important? Employing high-frequency data is particularly important to uncover temporal relations which are invisible to the discrete-price observer. Goetzmann *et al.* (2005) document a dramatic increase in global market correlations over the last two decades, which increases international market integration and naturally pushes evidence of lead–lag relationships deeper into the sub-minute and sub-second space. Examining the link between lead–lag effects and arbitrage has important implications towards the theory of efficient markets. Finally, using futures data mitigates many of the limitations of spot–futures arbitrage, such as transaction costs and illiquidity. Therefore, exploring lead–lag relationships across different futures contracts provides an important platform on which to test for inter-market predictability and arbitrage.

In what follows, we employ BBO quotes time-stamped to the millisecond for the S&P 500, FTSE 100 and DAX futures. These contracts are arguably the most globally liquid financial securities, and by using them we mitigate to the greatest extent possible the transaction cost barrier and the spurious lead–lag relation generated by liquidity differences across the contracts. Anticipating that lead–lag patterns occur at very fine timescales, we employ the HY estimator, which is statistically robust to non-synchronous trading and irregular quote arrival times.

DATA AND METHODOLOGY

Here we provide details of our data and data refinement procedures, together with details of our implementation of the HY estimator. Also, we contrast the performance of HY against a common discretization procedure. Finally, we provide details of the forecasting and trading strategy we employ.

Futures contracts data

Our data consist of the three most liquid futures contracts globally, namely the DAX, FTSE 100 and S&P 500 E-mini futures contracts.[5] For each contract, we obtain BBO quotes containing price and volume information, time-stamped to the millisecond. Our data spans the period 9 January 2012 to 28 February 2012. Each contract is monitored for 12 hours per day, between 08:30 and 20:30 GMT. Because futures contracts operate on a quarterly expiration cycle, we focus this study on the the current 'active' contracts, namely those expiring in March 2012. In all, our dataset consists of a number of BBO quotes in excess of 100 million.

Our data are sourced from the CME, Eurex and NYSE Liffe exchanges. Since each of these exchanges is synchronized to an atomic clock (IOSCO, 2012), the quote time-stamps reported for each exchange are precise to within 1 microsecond, and contemporaneous to within 1 millisecond (Barua *et al.*, 2012). Table I provides an overview of the institutional features of the dataset.

We see from Table I that these contracts are highly liquid. Given that our aim is to ultimately establish the existence of arbitrage opportunities based on lead–lag relationships, our choice of dataset is deliberately conservative—it is well documented that liquidity and the availability of arbitrage are inversely related.

Inspired by Schultz and Shive (2010), a number of crucial data-filtering processes are applied. Specifically, data which qualifies any of the following are excluded:[6]

[5] The E-mini is the more liquid and electronically traded version of the S&P 500 contract.

[6] In all, this process removes around 0.2% of the data. However, the last two bullet-point filters do not actually capture any data but are included here to show that our approach is consistent with methods used by previous researchers.

Table I. **Futures Data.** Institutional features of the data set. The last two columns show, as of January 2012, the average number of contracts traded, and the average daily turnover

Contact Name	Symbol	Exchange	Multiplier	Minimum Increment	Average Daily Volume	Average Daily Trade Value
DAX	DA	Eurex	x €25	€ 0.50	137,217	€21.5 billion
FTSE 100	FT	NYSE-Liffe	x £10	£0.50	104,041	£7.7 billion
E-mini S&P 500	ES	CME	x $50	$0.25	1,971,484	$140 billion

Table II. **Exchange Costs.** This table shows the costs for order submission and clearing levied by the Eurex, NYSE-Liffe, and CME Group. These relate to trading the DAX, FTSE 100, and S&P 500 contracts, respectively

Contract Name	Exchange Fees	Clearing Fees
DAX	€ 0.50	- - -
FTSE	£0.25	£0.03
E-mini S&P 500	$0.75	$0.39

- bid price $\geq$ ask price;
- bid volume $= 0$ and/or ask volume $= 0$;
- ask price $>$ bid price by more than 25%;
- mid quote return $\geq 25\%$ or $\leq -25\%$.

Since we use BBO quotes, we are in effect continuously observing the current best quote within the central order book. New BBO quotes arrive when there is either a change in the current best price or a change in the available liquidity at the current best price. In the first scenario, if a new quote arrives that is better than the current best price, or an incoming trade consumes all the liquidity available at the best price, the BBO price and volume update. In the second scenario, if a new quote arrives *at* the current best price, or an incoming trade does not consume all the currently available liquidity, the BBO volume updates but the price does not.

In view of establishing lead–lag relationships, the difference between the above two scenarios is non-trivial. If one monitors changes in the mid-quote, the first scenario will always yield a non-zero return, while the second scenario will always yield a zero return. The main aim here is to establish whether mid-quote changes in one futures contract portend corresponding changes in another. This aim is best served by focusing purely on non-zero returns, and so all zero-returns from the data are removed. This process carries the benefit of allowing us to deem all non-zero returns as informative to the lead–lag relationship.[7] A similar procedure is performed in Huth and Abergel (2014).

Trading costs and the representative investor

Costs pertaining to futures trading consist of the bid–ask spread, which is accounted for directly in the BBO quote data. Further, each exchange charges a fee for order submission and clearing. These are outlined below.

The costs in Table II apply on a per-order basis, regardless of the size of the order. Given that the notional amount of a single futures contract is 10, 25 or 50 times its underlying value (see Table I), the costs in Table II are economically small. However, our approach warrants the submission of a vast number of executable orders aimed at exploiting minute pricing disequilibria across these markets. The costs in Table II are negligible for a long-term investor, but are a significant consideration for our purposes.

With that in mind, our representative investor is akin to a quantitative hedge fund utilizing a fully algorithmic strategy, with access to co-location services within the exchange buildings. We choose this type of investor for two reasons: first, all three contracts are 100% electronically traded. Hendershott *et al.* (2011) provide evidence that some 70% of NYSE trades are executed by investors of this type. Secondly, our approach aims to exploit temporal pricing disequilibria which exist predominantly in the sub-second horizon. Clearly, such speeds lie beyond the scope of human traders.

The Hayashi–Yoshida estimator

In Hayashi and Yoshida (2005),[8] the authors introduce a novel estimator of the covariance between two non-synchronous processes. Specifically, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let X_t and Y_t be two correlated

[7] Griffin and Oomen (2011) provide a thorough discussion of similar subsampling routines.
[8] For brevity, mathematical proofs in this section are omitted. However, we direct the interested reader to works in which these proofs are available.

processes such that

$$\begin{aligned} \mathrm{d}X_t &= \mu^X X_t \mathrm{d}t + \sigma^X X_t \mathrm{d}W_t^X \\ \mathrm{d}Y_t &= \mu^Y Y_t \mathrm{d}t + \sigma^Y Y_t \mathrm{d}W_t^Y \end{aligned} \tag{1}$$

where W_t^X and W_t^Y are Brownian motions with respect to $\mathbb{P}$. Assume that the correlation $\langle W_t^X, W_t^Y \rangle = \rho$. Assume further that X_t and Y_t are sampled at discrete observation times $0 = t_0^X \leq t_1^X \leq \cdots \leq t_n^X = T^X$ and $0 = t_0^Y \leq t_1^Y \leq \cdots \leq t_m^Y = T^Y$ respectively. Importantly, these observation times are assumed independent of each other and of X_t and Y_t—which in practical terms suggests two things. First, the quote arrival frequency of one asset does not influence that of the other; secondly, the quote arrival frequency of each asset does not depend on the value of that asset.

Let us define the following terms:

$$\begin{aligned} I_i^X &= \left(t_i^X, t_{i+1}^X\right] \\ I_j^Y &= \left(t_j^Y, t_{j+1}^Y\right] \end{aligned} \tag{2}$$

Here, I_i^X and I_j^Y denote time intervals between the arrival of quotes in assets X and Y respectively, and this concept is illustrated in Figure 1. The figure depicts a typical scenario encountered with high-frequency data. Quotes arrive irregularly for each asset, and asynchronously across assets. Robustly estimating the covariation between data which exhibit these phenomena is the subject of HY. Given equations (1) and (2), the HY estimator for the covariance $C^{X,Y}$ is

$$C^{X,Y} = \sum_{i,j} \Delta X\left(I_i^X\right) \Delta Y\left(I_j^Y\right) \mathbb{I}_{\left\{I_i^X \cap I_j^Y \neq \emptyset\right\}} \tag{3}$$

where Δ is the difference operator and $\mathbb{I}$ is an indicator function such that

$$\mathbb{I} = \begin{cases} 1, & \text{if } I_i^X \cap I_j^Y \neq \emptyset \\ 0, & \text{otherwise} \end{cases}$$

The covariance $C^{X,Y}$ in equation (3) yields the HY correlation coefficient ρ_{HY} as follows:

$$\begin{aligned} \rho_{\mathrm{HY}} &= \frac{C^{X,Y}}{\sigma^X \sigma^Y} \\ &= \frac{\sum_{i,j} \Delta X\left(I_i^X\right) \Delta Y\left(I_j^Y\right) \mathbb{I}_{\left\{I_i^X \cap I_j^Y \neq \emptyset\right\}}}{\sqrt{\sum_i \left[\Delta X\left(I_i^X\right)\right]^2 \sum_j \left[\Delta Y\left(I_j^Y\right)\right]^2}} \end{aligned} \tag{4}$$

In practice, equations (3) and (4) amount to summing the product of all returns between assets X and Y once they fully or partially share a time overlap. Because of this property, the HY estimator allows for the inclusion of all data points, without the need for regularizing (resampling) the data.

As it is currently stated, equation (4) measures the contemporaneous correlations between X and Y. Hoffmann *et al.* (2013) extend the estimator to allow for leads and lags. Following from equation (2), we define

$$\left(I_j^Y\right)_\ell = \left(t_j^Y + \ell, t_{j+1}^Y + \ell\right] \tag{5}$$

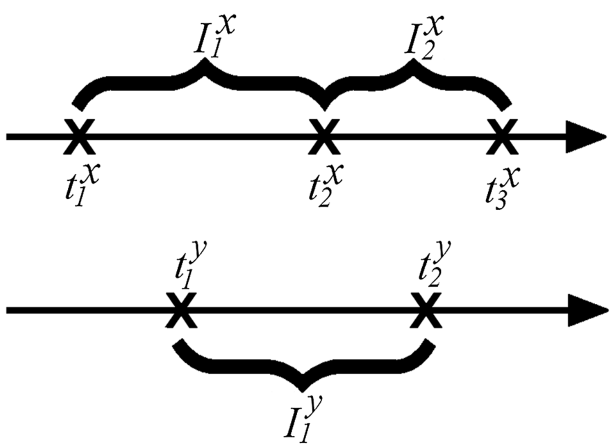

Figure 1. Time intervals. A time-line illustration of quote arrivals in X (top) and Y (bottom). Quotes are marked 'X'. Intervals between quotes are marked I^X and I^Y

where $\ell \in \mathbb{R}$ is the lag length (measured in units of time). Following the same procedure in equations (2)–(4), we arrive at a formula for the lagged version of the HY estimator:

$$\rho_{\mathrm{HY}}(\ell) = \frac{\sum_{i,j} \Delta X\left(I_i^X\right) \Delta Y\left(I_j^Y\right) \mathbb{I}_{\left\{I_i^X \cap \left(I_j^Y\right)_\ell \neq \emptyset\right\}}}{\sqrt{\sum_i \left[\Delta X\left(I_i^X\right)\right]^2 \sum_j \left[\Delta Y\left(I_j^Y\right)\right]^2}} \tag{6}$$

As it is now defined, equation (6) yields the entire cross-correlation curve (hereafter HY curve) between X and Y. In practice, evaluating (6) amounts to shifting the time-stamps of Y by amount ℓ and re-evaluating the HY correlation ρ_{HY}. Doing this for all ℓ yields the HY curve.

In practical applications, it is important to establish the lag length $\hat{\ell}$ that maximizes the correlation $\rho_{\mathrm{HY}}\left(\hat{\ell}\right)$. Doing so allows conclusions to be drawn about the temporal relationship between X and Y. For example, 'X leads Y by $\hat{\ell}$ seconds'. The lag $\hat{\ell}$ is defined as a solution to the following equation:

$$\left|\rho_{\mathrm{HY}}\left(\hat{\ell}\right)\right| = \max_{\ell \in \mathcal{G}} \left|\rho_{\mathrm{HY}}(\ell)\right| \tag{7}$$

over a time-grid $\mathcal{G}$. In practice, this amounts to finding the peak of the HY curve, equivalently, evaluating equation (6) for different values of ℓ until a maximum is obtained.

Importantly, if it is found that $\hat{\ell} \neq 0$, this would imply that the relationship between X and Y is not contemporaneous. Specifically, knowledge of one can be used to explicitly forecast future movements in the other. It is precisely this condition that we rely on to examine lead–lag relations across international futures contracts.

Finally, Huth and Abergel (2014) define the notion of a lead–lag ratio (henceforth LLR) which we employ in our analysis. It is defined as follows:

$$\mathrm{LLR} = \frac{\sum_u \rho_{\mathrm{HY}}^2(-\ell_u)}{\sum_u \rho_{\mathrm{HY}}^2(\ell_u)} \tag{8}$$

where u is chosen such that $\ell_u \geq 0$.

The quantity $\rho_{\mathrm{HY}}(\ell)$ in equation (6) provides the correlation coefficient when Y leads X by ℓ units of time and, similarly, the quantity $\rho_{\mathrm{HY}}(-\ell)$ provides the same measure when Y lags X (i.e. X leads Y) by ℓ units of time. With this in mind, the purpose of the LLR is simply to compare the evidence of these two phenomena. It is established in the literature that lead–lag relationships are often bidirectional (see Wang and Wang, 2001, among others). The LLR is informative in decoupling bidirectional relationships into relative strengths at different lags. Put simply, it is a metric which is useful in assessing the strength and direction of the lead–lag relationship.

Robustness to spurious lead–lag relations

This section evaluates the robustness of HY to differences in liquidity. Further, we contrast the performance of HY against a commonly employed method of measuring covariation.

It is well known that liquid[9] assets tend to lead less liquid ones, due to their ability to impound information faster (Lo and MacKinlay, 1990; Brooks *et al.*, 2001). However, liquidity differences could also give rise to spurious conclusions toward lead–lag relationships. For example, data that are known to be contemporaneously correlated but contain differences in trading activity must not exhibit any lead–lag effects (Voev and Lunde, 2007). Discovery of a non-zero lag in this case would, by definition, be spurious. It is therefore important that any employed estimator be robust to *artificial liquidity* effects, particularly when dealing with high-frequency data.

In this section, we examine the robustness of HY to spurious lead–lag relations by way of an exercise involving synthetic data. The benefit of using synthetic data is simple: we know what parameters we input; hence we know what output to expect.

Our exercise involves sampling from the stochastic processes X and Y in equation (1) along two independent Poisson time-grids with intensities λ_X and λ_Y respectively.[10] The data span 24 hours, and samples are observed to the nearest millisecond. Further, we impose that X leads Y by 400 milliseconds (0.4 seconds). At this lag length, we impose a correlation between X and Y of 90%.

We measure lead–lags between -10 and 10 seconds, with 10-millisecond increments. Specifically, we choose a grid $\mathcal{G}$ from (7) such that

$$\ell \in \{-10, -9.99, -9.98, \ldots, -0.01, 0, 0.01, \ldots, 9.98, 9.99, 10\}$$

[9] In this case, liquidity refers to trading/quoting activity.
[10] Here, λ_X and λ_Y can be thought of as proxies for illiquidity: the lower their value, the shorter the average interval between price updates.

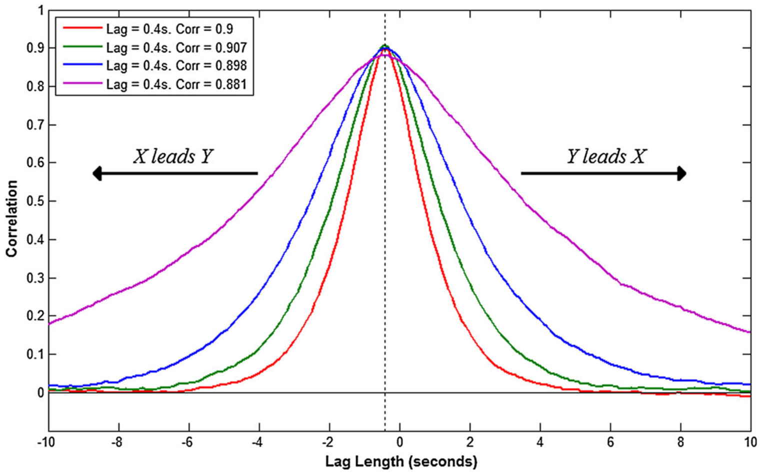

Figure 2. HY curve on synthetic data. This figure shows the HY-estimated correlation between X and Y at various lags $\ell \in [-10, 10]$. A dotted vertical line denotes the 'true' lead–lag relationship with X leading Y by 0.4 seconds $\left(\hat{\ell} = -0.4\right)$. The four curves represent repetitions of the HY exercise with various liquidity ratios: $\frac{\lambda_Y}{\lambda_X} = 0.5$ shown in red, $\frac{\lambda_Y}{\lambda_X} = 1$ in green, $\frac{\lambda_Y}{\lambda_X} = 2$ in blue and $\frac{\lambda_Y}{\lambda_X} = 5$ in purple

Further, we repeat the exercise for different liquidity combinations between assets X and Y. We achieve this by varying the ratio $\frac{\lambda_Y}{\lambda_X} \in \{0.5, 1, 2, 5\}$. This step is aimed at examining how the estimated lead–lag relation is affected by differences in liquidity. For example, the ratio $\frac{\lambda_Y}{\lambda_X} = 5$ suggests that the average interval between price updates for asset Y is five times that for asset X. Figure 2 shows the results of the HY exercise.

It is clear from Figure 2 that the HY estimator is robust to spurious temporal relations induced by liquidity differences between assets—it correctly recovers the true lead–lag relationship between X and Y under all combinations of $\frac{\lambda_Y}{\lambda_X}$. This property of HY is useful to the examination of temporal relationships using high-frequency data.

It is informative to contrast the performance of HY against that of a commonly employed methodology, namely *previous-tick interpolation* (henceforth PT).[11] For completeness, we present the results of the same exercise using the PT method.

The intuitive idea behind the PT method is to regularize irregularly spaced and non-synchronous prices onto a synchronous time-grid with fixed and constant intervals. Where there is no price update between two or more consecutive points on the regularized grid, the previous price is interpolated forward. Formally, and using the definitions in equation (2), let

$$\overline{X}_t = X\left(I_i^X\right) \text{ and } \overline{Y}_t = Y\left(I_j^Y\right) \tag{9}$$

denote the PT-interpolated versions of X and Y in equation (1), then for a fixed-interval grid of size M and interval length h, the PT correlation estimator is given by

$$\rho^{\text{PT}} = \frac{\sum_{i=1}^{M} \left(\overline{X}_{ih} - \overline{X}_{(i-1)h}\right)\left(\overline{Y}_{ih} - \overline{Y}_{(i-1)h}\right)}{\sigma^X \sigma^Y} \tag{10}$$

The extension of ρ^{PT} to incorporate leads and lags $\rho^{\text{PT}}(\ell)$ is similar to that of HY in equation (6).

We now repeat the estimation exercise using the PT method. We do this for various mesh sizes $h \in \{0.1, 0.5, 1\}$ seconds. The results are given in Figure 3.

There are two interesting phenomena evident in Figure 3. First, the PT estimation procedure is adversely affected by differences in liquidity. For example, in the bottom-left panel, where price updates in X occur on average five times as frequently as those in Y, the PT approach suggests that X leads Y by between 3 and 3.4 seconds, which is clearly not true. Second, although a smaller mesh size h yields a more accurate estimate of the true lag $\hat{\ell}$, it also results in a more biased estimate of the true correlation. This is not surprising, and is a well-known artefact of interpolation based on regularized grids (Epps, 1979).

Overall, the results of this section motivate the use of HY in dealing with high-frequency data.

[11] For a thorough discussion, see Voev and Lunde (2007), Hoffmann *et al.* (2013), and Griffin and Oomen (2011).

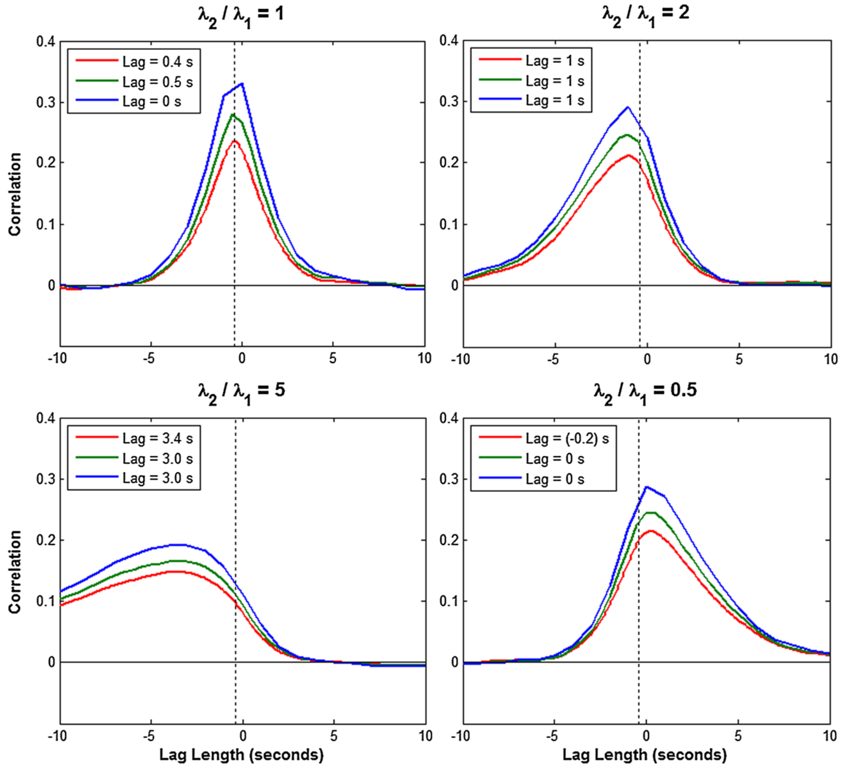

Figure 3. PT curve on synthetic data. This figure shows the PT-estimated correlation between X and Y at various lags $\ell \in [-10, 10]$. A dotted vertical line denotes the 'true' lead–lag relationship with X leading Y by 0.4 seconds $\left(\hat{\ell} = -0.4\right)$. The three curves in each panel represent repetitions of the PT exercise with various mesh sizes: $h = 0.1$ shown in red, $h = 0.5$ in green, and $h = 1$ in blue. The four panels represent different liquidity combinations $\frac{\lambda_Y}{\lambda_X}$ across X and Y. Lags which maximize the correlation are given by 'Lag'

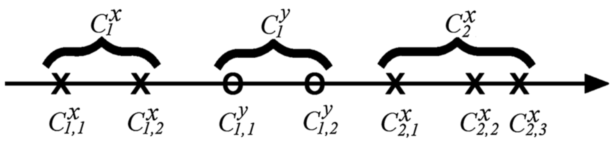

Figure 4. Quote clusters. A time-line illustration of contiguous sequences of mid-quote returns. Returns in asset X are marked **X** and returns in asset Y are marked **O**, while X-clusters and Y-clusters are marked with $C_{i,n}^X$ and $C_{i,n}^Y$ respectively

The trading strategy

Knowledge of the temporal relationship between a pair of contracts can be used to build statistical arbitrage strategies based on exploiting the temporal disequilibria. In this section, we show how we can apply our knowledge of the lead–lag relationship to forecast and exploit mid-quote changes in the lagging contract, based on mid-quote changes in the leading contract.

Our forecasting and trading devices are motivated by Huth and Abergel (2014) and Kozhan and Tham (2012). Given contracts X and Y, observing both time series simultaneously amounts to observing sets of contiguous quotes in X interspersed with sets of contiguous quotes in Y. Hence we define a *cluster* of mid-quote returns $\left\{C_{i,n}^X : i, n \in \mathbb{N}^+\right\}$ in contract X as a sequence of contiguous mid-quote returns in X uninterrupted by returns in Y. A similar definition of $C_{i,n}^Y$ holds for contract Y. Here, the subscript i refers to the cluster index (the number of clusters already observed), whereas the subscript n refers to the mid-quote return index within each cluster. This concept is illustrated graphically in Figure 4.

Once we distinguish the leading contract from the lagging contract within the HY framework, we use this information to build a statistical arbitrage trading strategy. Without loss of generality, let us assume contract X leads contract Y, then

$$N_i^Y = \begin{cases} +1, & \text{if } \max_n \left(C_{i,n}^X\right) \geq K \\ -1, & \text{if } \min_n \left(C_{i,n}^X\right) \leq -K \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

where N_i^Y is the position taken in contract Y, and where $+1$ and -1 denote long and short positions, respectively. The quantity $K > 0$ is a pre-set threshold. Intuitively, equation (11) describes a simple trading strategy: given that X leads Y, a signal to trade Y is generated whenever the price of X moves by an amount greater than K in the current X-cluster (i.e. before the price of Y moves).[12] Note that whether we choose to include or exclude zero returns from our data series makes no difference to the value of K. If Y moves in the interim, no trading signal is generated and the process restarts in the following X-cluster. The motivation for including a threshold in the trading strategy stems from the fact that exploiting microscopic price 'jumps' is fruitless, unless the magnitudes of these jumps exceed the costs associated with trading (Brooks *et al.*, 2001). It is important to note that the trading strategy (11) depends only on present and past price observations of the leading contract.

When a signal to trade Y is generated in the current X-cluster, a position in Y is immediately opened via an immediately executable limit order (i.e. a bid limit order at the prevailing ask price, or an ask limit order at the prevailing bid price) and held for the entire duration of the following Y-cluster. The position is then closed at the start of the subsequent X-cluster. Formally, we have

$$N_{i+1}^Y = \begin{cases} 0, & \text{if } \left\{C_{i+\epsilon,n}^X : n \geq 1, \epsilon \geq 1\right\} \\ N_i^Y, & \text{otherwise} \end{cases} \tag{12}$$

Taken together, equations (11) and (12) fully describe the algorithm by which we exploit temporal disequilibria between leading contract X and lagging contract Y. The calculation of profit P for each Y-cluster is as follows. We multiply the position taken in cluster i by the sum of the individual mid-quote returns in cluster i, and subtract trading costs:

$$P_i = N_i^Y \sum_n C_{i,n}^Y - c_i \tag{13}$$

where c_i is the all-inclusive cost of trading. Calculating total profits amounts to summing the individual P_i over all clusters in the dataset:

$$\text{total profit} = \sum_i P_i \tag{14}$$

Alongside the trading strategy, we measure the directional accuracy by which mid-quote changes in contract X portend mid-quote changes in contract Y. Specifically, we define

$$F_i = \begin{cases} 1, & \text{if } \operatorname{sgn}\left(\sum_n \left(C_{i,n}^Y\right)\right) = \operatorname{sgn}\left(\sum_n \left(C_{i,n}^X\right)\right) \\ 0, & \text{otherwise} \end{cases} \tag{15}$$

where $\operatorname{sgn}(\cdot)$ is the signum function. Equation (15) essentially measures whether the aggregate returns in the current X-cluster portend the sign of the aggregate returns in the subsequent Y-cluster. Arriving at a final measure of directional forecasting accuracy is intuitive: we sum all F_i and divide by the number of clusters in the dataset:

$$\text{directional forecasting accuracy} = \frac{\sum_i F_i}{\sup(i)} \times 100\% \tag{16}$$

where $\sup(\cdot)$ denotes the supremum.

An important question related to the forecasting accuracy is whether the accuracy is affected by the choice of signal threshold K in equation (11). One might expect that a stronger trading signal (i.e. higher K) would intuitively yield a higher directional forecasting accuracy of X over Y. To help investigate this, equations (15) and (16) are evaluated over all clusters *in which a trade has been generated.*

EMPIRICAL RESULTS

Here, we apply the analysis of the previous section to the three contract pairs within our dataset, namely the FTSE–DAX, S&P–FTSE and S&P–DAX March 2012 pairs. For the remainder of this paper, we define the following conventions: first, the leading asset, as in the above list, is named first; second, graphs pertaining to each pair will be coloured blue, green, and red, respectively. Third, the numerator of the lead–lag ratio (8) refers to the first-named contract; thus a highly asymmetrical left-heavy HY curve with LLR $\gg 1$ shows strong evidence that the first-named leads.

[12] Since all C denote mid-quote returns, it is easier to measure K in terms of ticks. Conversion between ticks and monetary value is done easily via Table I.

We start by exploring the three lead–lag relationships, and profile the intraday patterns of these relationships. Then, we forecast and trade lagging contracts based on signals generated by mid-quote changes in leading contracts. Finally, we discuss the limits to arbitrage across international futures contracts.

The lead–lag relationship between futures contracts

There are 35 full trading days (hereafter 'days') in our sample. For each day, we estimate the entire HY cross-correlation curve based on equation (6), then we average the curves for each pair over all days. We therefore ultimately obtain a single HY curve for each contract pair. This step helps quantify the relative strength and direction of the lead–lag relationships for each of our contract pairs.

For each contract pair, we measure leads and lags of mid-quote returns on a horizon of -30 to 30 seconds, with 5-millisecond increments. We justify this horizon by the fact that cross-correlations across all pairs diminish substantially within a few seconds. We choose a grid $\mathcal{G}$ from (7) such that

$$\ell \in \{-30, -29.995, -29.990, \ldots, -0.005, 0, 0.005, \ldots, 29.990, 29.995, 30\}.$$

The results of this exercise are presented in Figure 5. The figure displays evidence of strongly asymmetric lead/lag relationships, in which the S&P 500 contract leads both the FTSE 100 and DAX contracts. The relationship between the FTSE 100 and DAX contracts is less pronounced, though it is clear that the FTSE leads the DAX contract. These results are qualitatively consistent with other works documenting international price linkages (Innocenti *et al.*, 2011, among others), and the notion that international price discovery occurs in the USA. Further, by looking at the left-hand panels in Figure 5, what we find particularly striking is the speed at which the HY cross-correlation curves diminish towards zero across all three pairs. In other words, the panels in Figure 5 display evidence of high financial integration across the three contract pairs. This is confirmed by examining the right-hand panels, which reveal that point estimates of the lead–lag relationships, or equivalently the peaks of the HY curves, are of the order of milliseconds. This observation is striking because it has important implications for works examining international pricing relationships: by using non-granular data, it is easy to achieve the illusory effect of perfect contemporaneous correlation, as documented above ('Robustness to spurious lead–lag relations').

We quantify the strength and direction of the lead–lag relationships in terms of three parameters, namely the lag length $\hat{\ell}$ which maximizes the correlation between a given contract pair, the maximum correlation itself $\hat{\rho}^{HY}$ and the

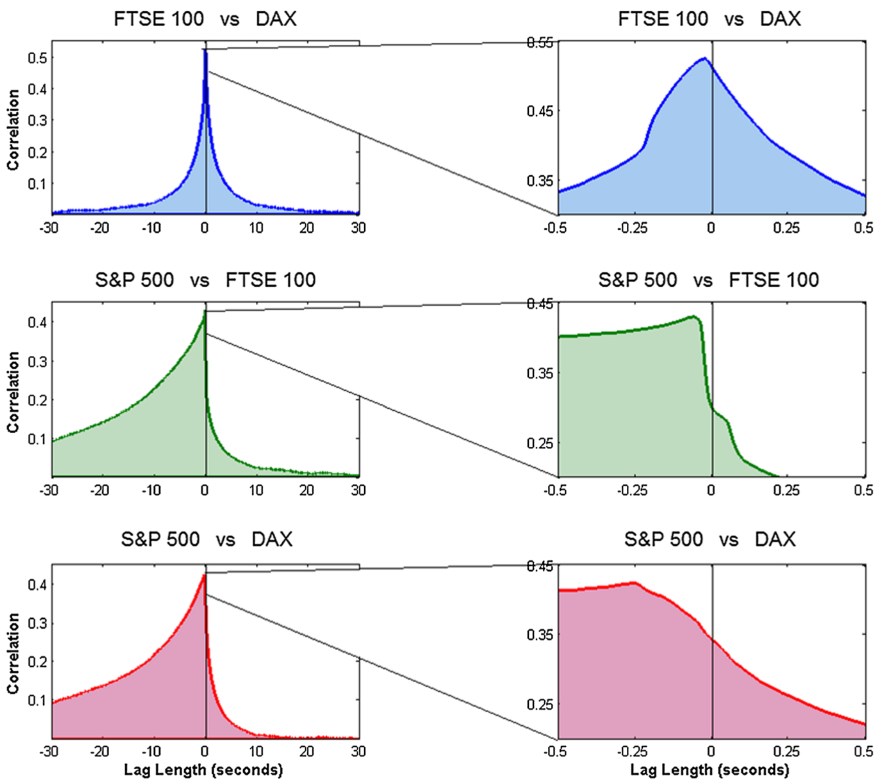

Figure 5. Lead–lag relationships. The HY curves plotted for each contract pair. Left panels show lag lengths $\ell \in [-30, 30]$ seconds. Right panels zoom in on lag lengths $\ell \in [-0.5, 0.5]$ seconds. In regions to the left of the vertical zero line in each panel, the contract whose name appears first in the title leads, and vice versa

Table III. **Lead/Lag Relationships.** Three contract pairs along with summary statistics. The third and fourth columns show, respectively, the mean and median of the lag length $\hat{\ell}$, by which the leader (second column) leads. The fifth and sixth column show, respectively, the maximum correlation $\hat{\rho}^{HY}$ and lead/lag ratios for each pair

Pair Name	Leading Asset	Mean Lag Length (milliseconds)	Median Lag Length (milliseconds)	Maximum Correlation	LLR
FTSE 100 / DAX	FTSE 100	25	33	52.36%	120
S&P 500 / FTSE 100	S&P 500	60	98	42.80%	19.64
S&P 500 / DAX	S&P 500	250	349	42.25%	15.96

lead–lag ratio LLR. These results are given in Table III. The table reveals intuitive insights into the nature of the lead–lag relationships under investigation. Lag lengths, maximum correlations and lead–lag ratios all appear to be directly related to the pairwise geographical distances between the markets.

Given that market information pertinent to any lead–lag relationship must traverse geographical distances, it is natural to wonder whether the lag lengths reported in Table III are not indicative of price discovery, but merely the result of delays in the physical transmission of data. The latter is unlikely. Built in 1999 by Global Crossing, the current transatlantic communication network (AC-1) spans 14,000 km, and links the US and major European markets via fibre-optic cables. One-way transmission between New York and London takes 32.4 milliseconds (see Johnson *et al.*, 2012, and references therein). By extrapolating this number to the distances between the CME, NYSE Liffe, and Eurex, it is unlikely that the lag lengths in Table III are the result of physical delays. Further, Hasbrouck and Saar (2012) suggest that the entire event–analysis—action cycle for co-located algorithmic trading systems is less than 2 milliseconds.

By examining the HY curves generated for each day in the sample, we reject at the 90% level the statistical hypothesis that the maximum correlation occurs at zero lag for the FTSE–DAX pair; and at the 99%-level for both transatlantic pairs. This result is not surprising, given the proximity of the NYSE Liffe to the Eurex, relative to either transatlantic distance.

The intraday profile of lead–lag relationships

It is a well-known stylized fact that financial markets produce intraday patterns (e.g. U-shaped volatility). Having established an overall lead–lag structure across the three contract pairs, we turn our attention to the intra–day profile of these relationships.

We split each day into 24 half-hourly intervals, spanning 08:30–20:30 GMT. For each interval, over each day, and for each contract pair, we measure the same three quantities as in Table III, namely the lead–lag ratio, the maximum correlation $\hat{\rho}^{\mathrm{HY}}$ and the lag length $\hat{\ell}$ which maximizes the correlation. We average these results over the number of days. Ultimately, we obtain three 24-point curves: one for each contract pair. The results of this exercise are shown in Figure 6. The figure reveals clear evidence of intraday seasonality exhibited for all three pairs. The daily events most interesting to us are the announcement of macroeconomic news at 13:30 GMT, the US cash market open at 14:30 GMT, and the close of both UK and German cash markets at 16:30 GMT. All three panels confirm that, for each pair, mid-quote returns in the first-named contracts portend mid-quote returns in the second-named contracts. Although this notion was established in the previous section, Figure 6 shows that this effect is consistent throughout the day.

The lead/lag ratio generally decreases (towards unity) throughout the day, indicating a diminishing asymmetry in the HY curves. This implies that the pairwise causal link between the leading and lagging contracts deteriorates throughout the day (Huth and Abergel, 2014). The maximum correlation remains largely range bound, and the lag length largely constant. Notable exceptions to the above are as follows. Around the announcement of macroeconomic data, the LLR and maximum correlation both increase, while in terms of lag length the US increases its leads over both UK and German contracts. This indicates that both UK and German traders react more attentively to cues from the US market, which impounds information first. The same effect occurs at the US cash market open. Following the close of the UK and German cash markets, the USA sharply increases its lead over each, but both the LLR and maximum correlation across all three pairs fall sharply. Taken together, these observations suggest that at the close of the UK and German markets the FTSE and DAX contracts are less sensitive to global factors, and more so to local factors (Jorion and Schwartz, 1986).

Trading results: exploiting temporal disequilibria

Here, we perform the algorithmic arbitrage exercise based on the methodology presented above (‘The trading strategy’). We reiterate that the trading strategy relies only on present and past mid-quote observations. Importantly, the trading strategy is simple in nature, and uses no *ex ante* knowledge of the dynamic lead–lag relationship presented above. This way, our entire dataset is taken as out-of-sample for forecasting and trading purposes. Although we discuss later how one may enhance the trading strategy by incorporating the information conveyed via the HY output, our results rely chiefly on the simple strategy (11)–(12), in the interests of robustness.

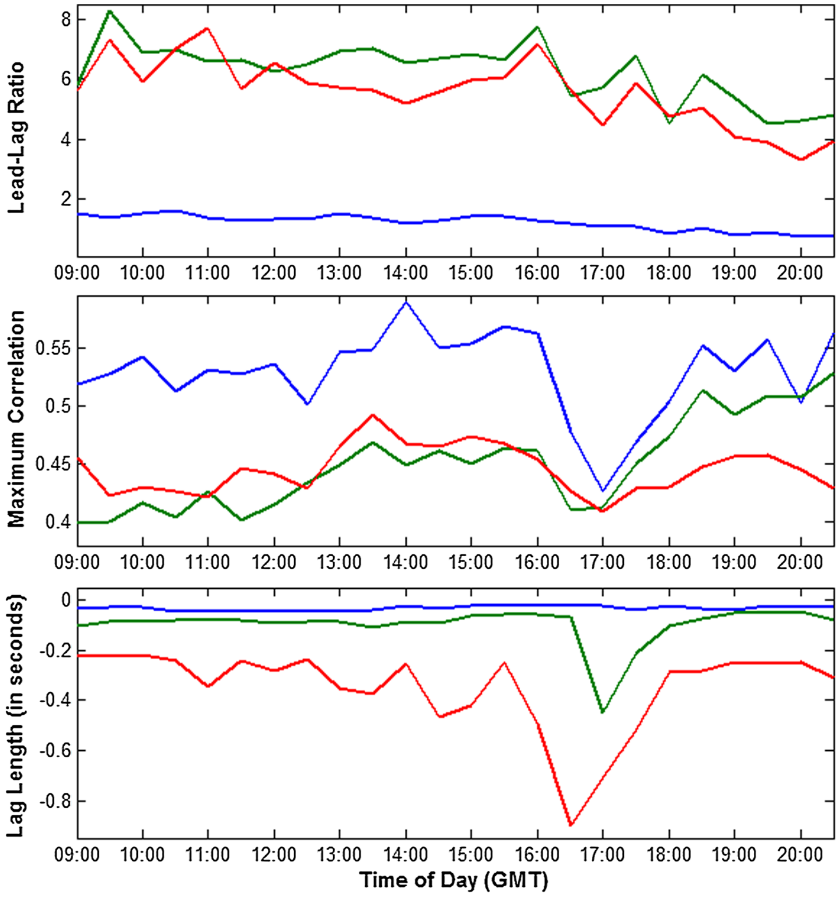

Figure 6. Lead–lag intraday profile. Intraday patterns in the lead–lag ratio (top), maximum correlation $\hat{\rho}^{HY}$ (middle) and lag length $\hat{\ell}$ (bottom). Times on the horizontal axes refer to the previous half-hour interval (e.g. 10:00 refers to the 09:30–10:00 interval)

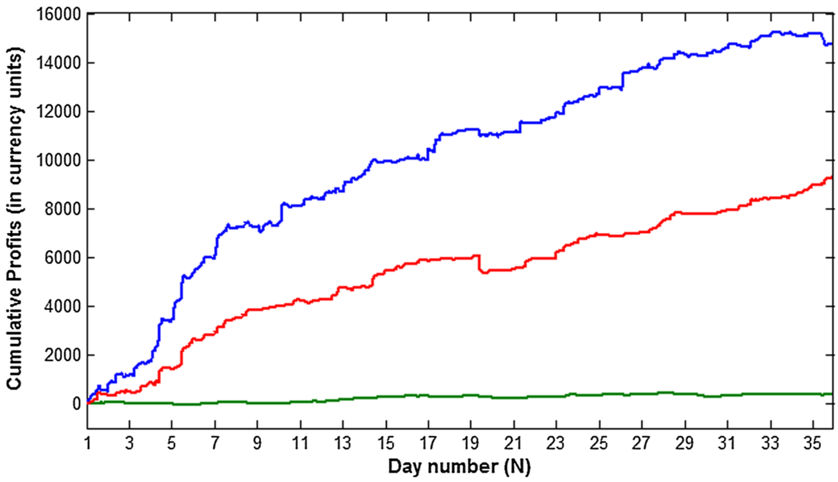

Figure 7. Cumulative profits. Profits gained by applying strategy (11)–(12) to the three contract pairs over the 35 days in our sample. The vertical axis is measured in euros for the FTSE–DAX and S&P–DAX pairs, and pounds sterling for the S&P–FTSE pair

We begin by presenting the accumulated profits gained by trading the FTSE–DAX, S&P–FTSE and S&P–DAX pairs over the number of days in our sample. Throughout, we maintain the convention that the first-named contract leads. Initially, we impose a signal threshold $K = 5$ ticks in equation (11). That is to say, allowing for necessary data transmission times, a trading signal is generated in the lagging contract whenever we encounter a cluster of leading quotes, wherein the mid-quote moves by five or more ticks in any aggregate direction. The position is then held open until the end of the lagging cluster. Later, we vary the signal threshold to examine its effect on the arbitrage profits. Furthermore, we initially impose that the arbitrageur trades a single contract per signal—a step that ensures no market impact costs, since we employ BBO quotes. Figure 7 shows the accumulated wealth gained by trading the three contract pairs, following the strategy (11)–(12). The curves denote profits over the 35-day sample period, net of the bid–ask spread and all order submission and clearing costs.

Following the lead of Suarez (2005), we present our cumulative profits in Figure 7 in absolute rather than percentage terms, grown from a small nominal initial investment and successively adding the profits from each individual trading opportunity without reinvestment of earlier gains. Futures contracts are inherently leveraged, since investors in them are not required to pay the full notional amount of the capital exposure that a futures contract affords to them, which is why the initial amount invested is so small. Typically, a clearing house will require the investor to post a *margin* equivalent to a small percentage of the contract's notional value (see Figlewski, 1984; Reilly and Brown, 1999; Hancock, 2005). The exact amount demanded by clearing houses to initiate a futures contract trade can vary depending on the recent historical volatility of the contract. There is a risk that, given an adverse market move, the investor may receive a margin call, or face having to close part of a trading position at an unfavourable point in time. However, in the historical trading exercise that we conducted, the effect of this proved negligible. The trading positions produced by our strategy were typically held for under 1 second and never for more than a few seconds. This means that we were not exposed to market risk in the conventional sense.

Although each profit time series trends consistently upwards, they exhibit substantially different magnitudes. The most profitable pair is the FTSE–DAX, while the least profitable is the S&P–FTSE. One reason for this may relate to the geographical proximity of London to Frankfurt: it is likely that the FTSE and DAX contracts respond to a larger set of overlapping idiosyncratic factors than either transatlantic pair, yielding a more consistent relationship. This suggestion is confirmed by a greater magnitude of correlation, shown in Figure 5. Further, the low profitability which characterizes the S&P–FTSE pair may be explained by Werner and Kleidon's (1996) suggestion that the US–UK market is among the most heavily arbitraged.

We now study the intraday profile of arbitrage. Specifically, we aggregate over our entire sample the directional forecasting accuracy, number of generated signals and durations of profitable disequilibria. The results are shown in Figure 8.

The top panel of Figure 8 shows a consistently high directional forecasting accuracy across all three contract pairs. Accuracy rates rise slightly following the US open, which accords with the idea that the UK and German markets follow the USA more closely during this period. Further, accuracy rates fall following the UK and German close, which is also an intuitive result exhibiting the opposite effect. Following Huth and Abergel (2014), we test the robustness of the directional forecasts against both a random forecast and a forecast generated via an autocorrelation in the lagging contract. The autocorrelation forecast slightly outperforms the random forecast, but is not statistically significantly better. The directional forecasts we present are significant at the 99% level for all contract pairs over all intervals in our sample, with the exception of the S&P–DAX pair between 20:00 and 20:30.

The middle panel of Figure 8 shows that the number of trades (or equivalently, trade signals) generated per interval increases sharply following the announcement of market data and again following the US open. This increase is then

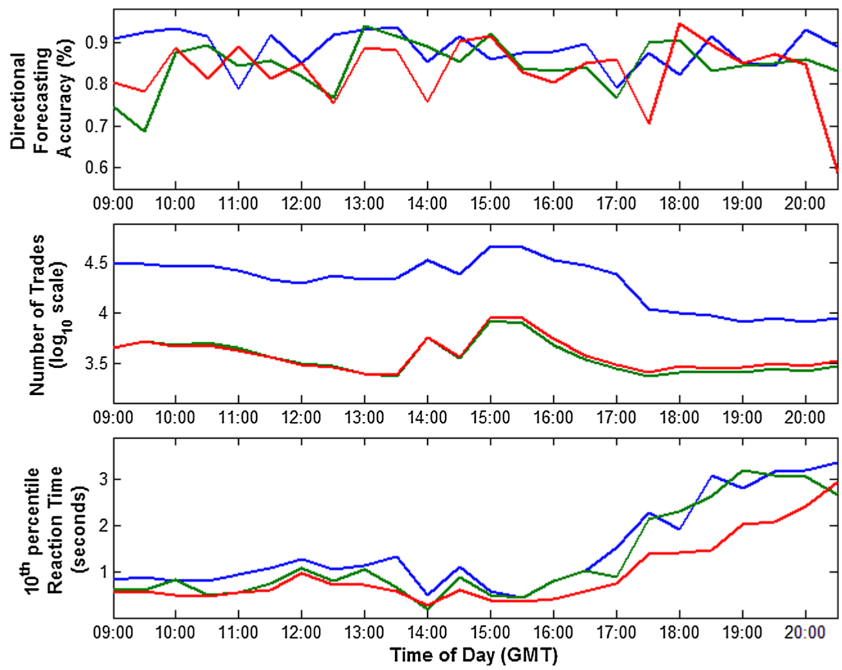

Figure 8. Intraday arbitrage profile. Intraday patterns of directional forecasting accuracy (top), the total number of individual trades on a $\log_{10}$ scale (middle) and the 10th percentile duration length of profitable disequilibria (bottom)

Table IV. **Varying the Signal Threshold.** This table shows, for each contract pair, the effect of varying the signal threshold K from 1 to 10 ticks. For each instance, the table shows the directional forecasting accuracy, total number of trades, profit per trade, and 10th percentile reaction time

Threshold tricks	1	2	3	4	5	6	7	8	9	10
					Directional Forecasting Accuracy (%)					
FTSE/DAX	70.87%	83.31%	90.02%	92.11%	92.28%	92.21%	92.61%	90.90%	89.64%	89.45%
S&P/FTSE	66.02%	66.55%	93.42%	93.72%	94.25%	93.69%	100.00%	100.00%	100.00%	100.00%
S&P/DAX	68.54%	69.26%	92.75%	93.32%	90.79%	91.62%	84.13%	84.13%	83.33%	83.33%
					Total Number of Trades					
FTSE/DAX	276,645	186,012	68,112	21,913	6,604	2,210	900	406	216	133
S&P/FTSE	47,844	44.135	595	539	67	55	14	13	5	5
S&P/DAX	49,559	46,461	799	766	70	66	15	14	4	4
					Profit per Trade (EUR or GBP)					
FTSE/DAX	-€, 14.62	-€ 9.84	-€ 4.60	-€ 0.36	-€ 3.41	-€ 6.35	-€ 11.43	-€ 12.45	-€ 15.20	-€ 13.29
S&P/FTSE	£7.49	£7.36	£0.44	£0.92	£3.77	£5.26	£19.44	£20.98	£40.44	£40.44
S&P/DAX	-€ 15.26	-€ 14.71	-€ 12.08	-€ 12.59	-€ 28.64	-€ 32.14	-€ 30.67	-€ 25.79	-€ 2.13	-€ 2.13
					10$^{\text{th}}$ Percentile Reaction Time (milliseconds)					
FTSE/DAX	1,166	1,277	1,432	1,668	1,628	1,130	769	549	275	162
S&P/FTSE	1,177	1,255	1,218	1,068	82	115	64	58	37	37
S&P/DAX	1,024	1,055	456	459	135	142	178	174	112	112

maintained for as long as all three cash markets remain open. The bottom panel plots the 10th-percentile *reaction times*, denoting interval durations between the point at which a trade signal is generated and the point at which the lagging contract begins to move following the signal. This panel shows that reaction times decrease following the announcement of market data, and remain low for as long as all three cash markets are open, before gradually rising following the close of the UK and German markets. Taken together, these results suggest that market participants who act to appropriate temporal disequilibria concentrate their activity during specific periods.

Varying the signal threshold

Up to now, we have assumed a pre-set signal threshold K. It is informative to study the effects of varying this threshold on the profitability of arbitrage.

Intuitively, a higher signal threshold yields better directional forecast profitability, at the expense of a lower trading frequency. Studying the effects of varying the signal threshold is informative to both academics and practitioners. To academics, it establishes a link between the frequency of temporal price disequilibria and risk. To practitioners, it permits an investigation into optimal trading rules.

We proceed by varying the signal threshold from $K \in \{1, 2, \ldots, 10\}$ ticks. For each instance and for each contract pair, we measure the aggregate directional forecasting accuracy, total number of trades, average profit per trade and 10th percentile reaction times. The results are shown in Table IV. Inspired by Schultz and Shive (2010), a number of crucial data-filtering processes are applied. Specifically, data which qualify any of the following are excluded:[13]

Table IV reveals several insights into the effects of varying the signal threshold on arbitrage returns. The effects are consistent across all contract pairs. First, increasing the signal threshold directly increases the directional forecasting accuracy. This result is consistent with the notion that increasing the signal threshold results in a *cleaner* signal with which to trade. Similarly, while the total number of trades falls sharply, the profit per trade increases. In particular, despite low signal thresholds yielding directionally accurate forecasts, trading based on low signal thresholds does not overcome costs. To practitioners, these results can be used to devise optimal trading strategies.

The final three rows of Table IV are interesting. The 10th percentile reaction times decline sharply as the signal threshold increases. In other words, arbitrageurs who wait for stronger trade signals do so at the cost of facing drastically shorter durations of disequilibria. This is consistent with the view that competition among arbitrageurs acts to enforce pricing efficiency (Grossman and Stiglitz, 1976, 1980), thereby perpetuating the need to invest in technological infrastructure. Clearly, this has a limit in the form of the speed of light. Over time, the pay-off to investing in infrastructure would yield diminishing marginal returns. Also, it is interesting to note that the directional forecasting accuracies for the FTSE–DAX and S&P–DAX contracts actually *decrease* for very high signal thresholds. However, this is the result of a small sample, given the very low number of trades at thresholds above eight ticks for both pairs.

[13] In all, this process removes around 0.2% of the data. However, it is interesting to note that the last two bullet-point filters do not capture any data.

It is informative to compare the reaction times stated in Table IV to the time it takes data to physically traverse the relevant geographical distances. Recall from above ('The lead–lag relationship between futures contracts') that one-way transmission between New York and London takes 32.4 milliseconds via Global Crossing's AC-1 fibre-optic cable. By extrapolating this time to the distances between the CME, NYSE Liffe and Eurex, it is highly unlikely that any of the FTSE–DAX or S&P–DAX trades lie inside the transmission speed boundary, even for high signal thresholds. This is particularly striking for the FTSE–DAX pair, which benefits from close geographical proximity, and which we have shown to be more profitable than either transatlantic pair. It is therefore puzzling why the 10th percentile reaction time for this pair should be significantly higher than either transatlantic pair. However, looking closely at the profits per trade is informative: the FTSE–DAX pair generates a large number of trades, each of which is relatively less profitable than either transatlantic pair. This suggests that the FTSE–DAX pair contains a relatively high idiosyncratic risk component which acts as a deterrent to arbitrage.

Looking at the S&P–FTSE pair, a signal threshold $K \geq 9$ suggests that for a number of arbitrage opportunities, namely those at 37 milliseconds, the duration of disequilibria lies at or just within the data transmission speed barrier. However, it is noteworthy that in our entire sample this phenomenon occurs for only one single trade. Overall, the S&P–FTSE result is consistent with Werner and Kleidon's (1996) suggestion that the US–UK market is among the most heavily arbitraged.

Finally, it is important to note that the computing processes behind generating a trading signal and managing a position require only the evaluation of one logical operation and one floating-point operation per observed mid-quote, as per equations (11) and (12). Modest desktop computing resources can perform these kinds of calculations within a small number of microseconds, akin to suggestions by Hasbrouck and Saar (2012). As for the millisecond environment in which this exercise operates, it is unlikely that the time required for calculation would significantly affect the visible duration of temporal disequilibria.

On the limits of arbitrage

Our dataset consists of the three most liquid and widely traded futures contracts globally, with no restrictions on trading or cross-border arbitrage.

In markets such as this, professional arbitrageurs often compete over a limited supply of available arbitrage opportunities. This competition creates an imbalance stemming from the excess demand by arbitrageurs for profitable trades. As a result, competition among arbitrageurs effectively reduces the magnitude and duration of profitable disequilibria. This in turn creates a natural selection mechanism by which arbitrageurs who possess the most powerful computing resources survive. Overall, this system poses two major limits to arbitrage, namely costs and risk.

Liquidity, price slippage, market impact and competition for scarce opportunities from other traders are all sources of risk for which arbitrageurs demand compensation. This fact may explain why even seemingly *riskless* arbitrage opportunities sometimes go unexploited. It is informative to explore the effects each of these risks has on arbitrageurs' capacity to appropriate pricing disequilibria, by examining their effects on profitability.

We parsimoniously capture the effects of liquidity risk, price slippage, market impact and competition from other arbitrageurs, following an approach by Kozhan and Tham (2012). Specifically, we impose that a certain proportion of the arbitrageur's trades are not executed at the observed (best) price, but at the next best limit order price instead, which is assumed to be available in infinite supply. The next best price in the context of this work means a one-tick increment down the limit order book, equivalent to GBP 5 and EUR 12.50 for the FTSE 100 and DAX contracts, respectively. The objective of this exercise is to see what effect this restriction has on the profitability of the individual arbitrageur. The results are shown in Figure 9.

For each contract pair, Figure 9 reveals the sensitivity of an arbitrageur's profits to the risk of executing part of a trade at the next best price. This sensitivity is particularly high for the FTSE–DAX pair, which is an intuitive result, given the generally low per-trade profitability of this pair as seen in Figure 8.

Up to this point, we have assumed that the trader executes a single contract per trade (which is the minimum quoted volume in all three futures contracts). With this in mind, it is interesting to ask: How is the profitability of the strategy affected by scaling up the number of contracts per trade? In the dataset, around 29% of FTSE 100 quotes on either the bid or ask side carry a size of five contracts or fewer, and around 46% carry 11 or more contracts. The corresponding figures for the DAX are 49% and 22% respectively. Clearly, when implementing a scaled-up version of this strategy, there is a risk of price slippage. For example, if a trader wishes to sell 15 contracts which she has previously bought, and there are only five contracts available at the current bid, she may be forced to sell 10 contracts at the next best (lower) bid price. Figure 9, along with the above analysis of the number of contracts available, is useful in understanding the effect of slippage on the profitability of arbitrage in this market setting.

Besides risk, the costs associated with trading in this market setting act as a significant deterrent to arbitrage. Figure 8 shows the directional forecasting accuracy to be high. However, in order to profit from a trade, an arbitrageur has to not only be directionally accurate but also overcome the bid–ask spread. Given that the majority of trades individually extract very small amounts of profit, the bid–ask spread is a major cost incurred by arbitrageurs. Furthermore, arbitrageurs face fixed rents and other costs incurred through co-location subscriptions and access to

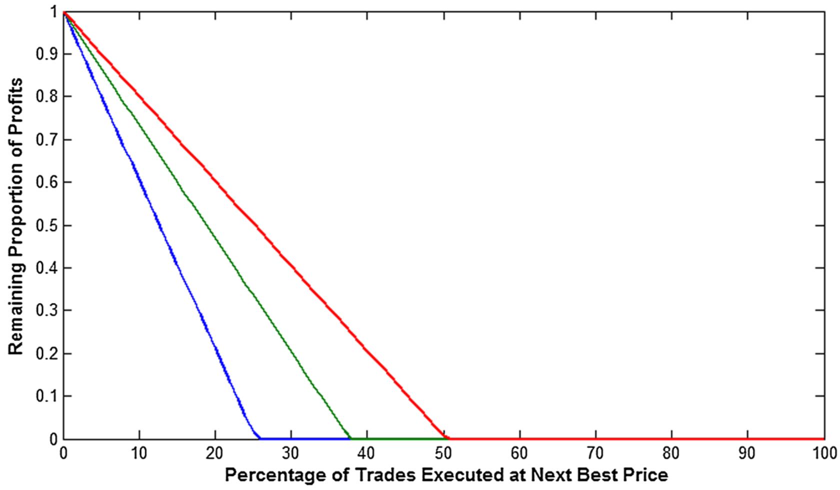

Figure 9. Slippage risk. The effect of having a percentage of trades fulfilled at the next best price (horizontal) on the proportion of otherwise available profit opportunities (vertical)

live market feeds. Hedge funds implementing these kinds of strategies must also hire analysts to create and maintain computer code. To the extent that these extra costs collectively diminish the profits reported in Table IV, they act as a deterrent to arbitrage. Due to the costs involved, one could suggest that arbitrage operations in this setting can most readily be undertaken by larger hedge funds and proprietary trading desks whose co-located infrastructure is not dedicated solely to this pursuit, but forms part of their overall business activity.

Overall, this section highlights that technological infrastructure costs and the risk of price slippage present significant limits to arbitrage in the market for international futures contracts.

CONCLUSION

This paper complements previous work by examining high-frequency pricing relationships in the important but overlooked market setting of international index futures. We document clear evidence of consistent sub-second lead–lag patterns across the S&P 500, FTSE 100 and DAX futures contracts (in order of leadership), which shows that the diffusion of information across these markets is not instantaneous.

Importantly, we show that these lead–lag patterns allow predictability of lagging assets, and give rise to profitable disequilibria. Taken together, these observations provide evidence against auto-efficiency across international futures markets, and suggest that arbitrage via algorithmic trading is an important component of the information flow across markets. However, these arbitrage profits are sensitive to the risk of price slippage, and profitable opportunities rarely exist for more than 300 milliseconds. Therefore, we highlight slippage risk and technological costs as the most significant limits to arbitrage in this market setting.

The data we employ consists of the three most liquid futures contracts, which are among the most liquid financial instruments in the world, with no capital controls and no restrictions to cross-border arbitrage. Based on this, we suggest that the results obtained in this paper are generalizable to other global markets, and we are keen to examine pricing relationships across other highly liquid instruments including other futures contracts. Furthermore, we highlight the importance of utilizing high-frequency data, based on compelling evidence that the price adjustment mechanism in these markets now mostly operates deep within the sub-second domain.

Future work could focus on enriching the perspective we presented above by extending our lead–lag analytical framework to multiple dimensions, e.g. by looking at the lead–lag relationships and cross-relationships between trading volumes, volatility, order flow and bid–ask spreads between various pairs of instruments. This would facilitate the analysis of global flows of information and transition effects across multiple geographical markets.

REFERENCES

Abhyankar A. 1998. Linear and nonlinear Granger causality: evidence from the UK stock index futures market. *Journal of Futures Markets* **18**: 519–540.

Abhyankar A, Copeland LS, Wong W. 1997. Uncovering nonlinear structure in real-time stock-market indexes: the S&P 500, the DAX, the Nikkei 225, and the FTSE-100. *Journal of Business and Economic Statistics* **15**(1): 1–14.

Andreou PC, Pierides YA. 2008. Empirical investigation of stock index futures market efficiency: the case of the Athens derivatives exchange. *European Journal of Finance* **14**: 211–23.

Antoniou A, Pescetto G, Violaris A. 2003. Modelling international price relations and interdependencies between the stock index and stock index futures markets of three EU countries: a multivariate analysis. *Journal of Business Finance and Accounting* **30**: 645–667.

Ayuso J, Blanco R. 2001. Has financial integration increased during the 1990s? *Journal of International Financial Markets, Institutions and Money* **11**: 265–287.

Barua A, Ranganayakulu J, Parakh N, Sharma AK. 2012. SEBI Bulletin. *Securities and Exchange Board of India* **10**(4): 359–365.

Brennan MJ, Schwartz ES. 1990. Arbitrage in stock index futures. *Journal of Business* **63**: 7–31.

Brooks C, Garrett I. 2002. Can we explain the dynamics of the UK FTSE 100 stock and stock index futures markets? *Applied Financial Economics* **12**: 25–31.

Brooks C, Rew AG, Ritson S. 2001. A trading strategy based on the lead–lag relationship between the spot index and futures contract for the FTSE 100. *International Journal of Forecasting* **17**: 31–44.

Cummings J, Frino A. 2011. Index arbitrage and the pricing relationship between Australian stock index futures and their underlying shares. *Accounting and Finance* **51**: 661–683.

De Jong F, Nijman T. 1998. Intraday lead–lag relationships between the futures, options, and stock market. *European Finance Review* **1**: 337–359.

Eiteman D, Stonehill A, Moffett MH, Moffet M. 1994. *Multinational Business Finance*. Addison-Wesley: Boston, MA.

Epps T. 1979. Co-movements in stock prices in the very short run. *Journal of the American Statistical Association* **74**: 291–298.

Eun C, Shim S. 1989. International transmission of stock market movements. *Journal of Financial and Quantitative Analysis* **24**(2): 241–256.

Evans MD, Hnatkovska V. 2014. International capital flows, returns and world financial integration. *Journal of International Economics* **92**(1): 14–33.

Figlewski S. 1984. Hedging performance and basic risk in stock index futures. *Journal of Finance* **39**: 657–670.

Figlewski S. 1985. Hedging with stock index futures: theory and application in a new market. *Journal of Futures Markets* **5**: 183–200.

Goetzmann W, Li L, Rouwenhorst KG. 2005. Long-term global market correlations. *Journal of Business* **78**(1): 1–38.

Griffin JE, Oomen RCA. 2011. Covariance measurement in the presence of non-synchronous trading and market microstructure noise. *Journal of Econometrics* **160**(1): 58–68.

Grossman S, Stiglitz J. 1976. Information and competitive price systems. *American Economic Review* **66**: 246–252.

Grossman S, Stiglitz J. 1980. On the impossibility of informationally efficient markets. *American Economic Review* **70**: 393–408.

Gulen H, Mayhew S. 2000. Stock index futures trading and volatility in international equity markets. *Journal of Futures Markets* **20**(7): 661–685.

Hamao Y, Masulis R, Ng V. 1990. Correlations in price changes and volatility across international stock markets. *Review of Financial Studies* **3**(2): 281–307.

Hancock GD. 2005. A Text Book Treatment of Calculating Returns on Non-Traditional Portfolios. *Review of Financial Economics* **14**(2): 173–186.

Hasbrouck J, Saar G. 2012. *Low-latency trading*, Working paper, New York University.

Hayashi T, Yoshida N. 2005. On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli* **11**(2): 359–379.

Hendershott T, Jones CM, Menkveld AJ. 2011. Does algorithmic trading improve liquidity? *Journal of Finance* **66**: 1–33.

Herbst AF, McCormack JP, West EN. 1987. Investigation of a lead–lag relationship between spot stock indices and their futures contracts. *Journal of Futures Markets* **7**: 373–381.

Hoffmann M, Rosenbaum M, Yoshida N. 2013. Estimation of the lead–lag parameter from non-synchronous data. *Bernoulli* **19**(2): 426–461.

Huth N, Abergel F. 2014. High frequency lead/lag relationships: empirical facts. *Journal of Empirical Finance* **26**: 41–58.

Innocenti A, Malpenga P, Menconi L, Santoni A. 2011. Intraday anomalies in the relationship between U.S. futures and European stock indexes. *Journal of Index Investing* **14**: 40–52.

IOSCO. 2012. Technological Challenges to Effective Market Surveillance Issues and Regulatory Tools. *Board of the International Organization of Securities Commissions.* August 2012.

Johnson N, Zhao G, Hunsader E, Meng J, Ravindar A, Carran S, Tivnan B. 2012. Financial black swans driven by ultrafast machine ecology. Available at SSRN.

Jorion P, Schwartz E. 1986. Integration vs. segmentation in the Canadian stock market. *Journal of Finance* **41**: 603–616.

Kawaller IG, Koch PD, Koch TW. 1987. The temporal price relationship between S&P 500 futures and the S&P 500 index. *Journal of Finance* **42**(5): 1309–1329.

Kenourgios DF. 2004. Price discovery in the Athens derivatives exchange: evidence for the FTSE/ASE-20 futures market. *Economic and Business Review* **6**: 229–243.

Kearney C, Lucey BM. 2004. International equity market integration: theory, evidence and implications. *International Review of Financial Analysis* **13**(5): 571–583.

Kozhan R, Tham WW. 2012. Execution risk in high-frequency arbitrage. *Management Science* **58**(11): 2131–2149.

Lo A, MacKinlay C. 1990. When are contrarian profits due to stock market overreaction? *Review of Financial Studies* **3**: 175–205.

Medeiros O, de Oliveira G, van Doornik B. 2009. Testing for lead–lag effects between the American and the Brazilian stock markets. *Working paper*. University of Brasilia.

Reilly FK, Brown KC. 1999. *Investment Analysis and Portfolio Management* (7th edn). South Western Educational Publishing.

Richie N, Daigler R, Gleason K. 2008. The limits to stock index arbitrage: examining S&P 500 futures and SPDRs. *Journal of Futures Markets* **28**12: 1182–1205.

Schultz P, Shive S. 2010. Mispricing of dual-class shares: profit opportunities, arbitrage, and trading. *Journal of Financial Economics* **98**: 524–549.

Stoll HR, Whaley RE. 1990. The dynamics of stock index and stock index futures returns. *Journal of Financial and Quantitative Analysis* **25**: 441–468.

Suarez E. 2005. Arbitrage opportunities in the depositary receipts market: myth or reality? *Journal of International Financial Markets, Institutions and Money* **15**: 469–480.

Thomas S. 2002. The saga of the first stock index futures contract: benchmarks, models, and learning. *Journal of Money, Credit, and Banking* **34**3: 767–808.

Tse YK. 1995. Lead–lag relationships between spot index and futures price of the Nikkei stock average. *Journal of Forecasting* **14**: 553–563.

Voev V, Lunde A. 2007. Integrated covariance estimation using high-frequency data in the presence of noise. *Journal of Financial Econometrics* **5**1: 68–104.

Wahab M, Lashgari M. 1993. Price dynamics and error correction in stock index and stock index futures markets: a cointegration approach. *Journal of Futures Markets* **13**: 711–742.
Wang P, Wang P. 2001. Equilibrium adjustment, basis risk and risk transmission in spot and forward foreign exchange markets. *Applied Financial Economics* **11**: 127–136.
Werner I, Kleidon A. 1996. U.K. and U.S. trading of British cross-listed stocks: an intraday analysis of market integration. *Review of Financial Studies* **9**: 619–664.
Zeckhauser R, Niederhoffer V. 1983. The performance of market index futures contracts. *Financial Analysts Journal* **39**1: 59–65.

Authors' biographies:
Hamad Alsayed is a PhD (Finance) candidate at the University of Southampton. He holds a BSc in Economics and Statistics, and a MSc in Mathematical Finance.

Frank McGroarty is Professor of Computational Finance and Investment Analytics at the Southampton Management School, University of Southampton, UK. He holds a PhD in Finance from the University of Southampton and has published numerous papers in leading academic journals. To date, his academic research has attracted over £3million in research funding. Prior to becoming an academic, he spent 13 years working as an investment analyst and fund manager in London for Barclays Global Investors, UBS and State Street.

Authors' address:
Hamad Alsayed and **Frank McGroarty**, Southampton Management School, University of Southampton, Southampton SO17 1BJ, UK.