High-Frequency Cross-Market Trading: Model Free Measurement and Testable Implications

January 31, 2023

Dobrislav Dobrev[a], Ernst Schaumburg[b,*]

[a] *Dobrislav Dobrev: Federal Reserve Board of Governors, dobrislav.p.dobrev@frb.gov*
[b] *Ernst Schaumburg: ernst.schaumburg@gmail.com*

Abstract

We propose a simple model-free approach to measuring latency-driven interdependence in high speed cross-market trading activity. The resulting measures rely on high-precision timestamps alone, making them more generally applicable than return-based measures. They allow for sharp identification of prevailing low-latency lead-lag interlinkages in modern electronic markets and studying variations in the intensity of cross-market activity over time. As an empirical illustration to U.S. cash and futures markets, we show that the intensity of high-speed cross-market trading activity is directly associated with fleeting liquidity and volatility transmission both in normal times and episodes of market stress. This leads to new perspectives and testable implications for intrinsic times in finance such as endogeneity between volatility and cross-market activity and the cross-market interdependence in observed activity times.

Keywords: Low-latency trading, lead-lag relationships, cross-market interdependence, co-activity, volatility, liquidity, intrinsic time.

JEL classification: C58, C22, C14, G12, G13, G01.

[*] Corresponding author: Dobrislav Dobrev, Federal Reserve Board of Governors, 20th and C St NW, Washington, DC 20551; Phone: 202-549-5859; Email: dobrislav.dobrev@gmail.com
We are grateful for helpful discussions and comments from Robert Almgren, Torben Andersen, Federico Bandi, Tim Bollerslev, Francis Breedon, Alain Chaboud, Dan Cleaves, Michael Fleming, Ravi Jagannathan, Takaki Hayashi, Richard Haynes, Jorge Herrada, Peter Hoffmann, Frank Keane, Pete Kyle, Jia Li, Giang Nguyen, Andrew Patton, Clara Vega, Iryna Veryzhenko, Toshiaki Watanabe, Brian Weller, Nakahiro Yoshida, Filip Zikes, as well as seminar participants at Bank of England, Bank of Japan, Cass Business School, Commodity Futures and Trading Commission, Duke University, Federal Reserve Bank of Chicago, Financial Conduct Authority, Rutgers Business School, University of California Berkeley, University of Maryland, University of Padova, University of Tokyo, and participants at the Hitotsubashi Summer Institute Workshop "Frontiers in Financial Econometrics", Tokyo, August 4-5, 2015, the Ninth Annual SoFiE Conference, Hong Kong, June 15-17, 2016, the Financial Econometrics and Empirical Asset Pricing Conference, Lancaster University, June 30 - July 1, 2016, the 69th European Meeting of the Econometric Society, Geneva, August 22-26, 2016, the Alan Turing Institute Conference on "Algorithmic Trading: Perspectives from Mathematical Modelling", London, March 1, 2017, the Vienna-Copenhagen Conference on Financial Econometrics, Vienna, March 9-11, 2017, the Minisymposium on High Frequency Trading, University of Pittsburgh, March 25-26, 2017, the 3rd International Workshop on Financial Markets and Nonlinear Dynamics, Paris, June 1-2, 2017, the 2018 AFA Annual Meeting, Philadelphia, January 5-7, 2018, the 2018 Annual Meeting of CEBRA, Frankfurt, August 20-21, 2018, the Conference on Financial Stability Implications of New Technology, Atlanta, October 18-19, 2018, the 2019 MFA Annual Meeting, Chicago, March 7-9, 2019, the 7th Stevanovich Center Conference on Market Microstructure and High Frequency Data, University of Chicago, May 2-4, 2019, the 12th World Congress of the Econometric Society, Bocconi University, August 17-21, 2020, the International Conference on Intrinsic Time in Finance, Allensbach, May 6-7, 2022 and the Workshop on Volatility, Jumps and Bursts, Lancaster University, June 27-28, 2022.

1. Introduction

Market activity and volatility in the most liquid US fixed income and equity markets have long been recognized as important barometers of policy expectations and economic conditions that are scrutinized by market participants and policymakers alike. However, trading activity in many of today's leading financial markets is no longer dominated by long-term investors expressing their views about risk reward trade-offs. Instead, the majority of market activity has become associated with high-speed automated trading strategies aiming to optimally place and modify orders to take advantage of short lived trading opportunities often unrelated to views about fundamental values. An important aspect of this activity is that much of it takes place near contemporaneously across venues trading instruments with substantially similar risk characteristics such as the cash and futures markets studied in this paper.

Much of the debate of the impact of high-frequency trading (HFT) has centered on changes in liquidity provision and HFT's role for observed market fragilities and cross-market interlinkages during episodes of market stress. This is at least in part because of the difficulty in identifying the precise footprint of high-speed cross-market trading activity despite the unprecedented amount of market data available across many electronic markets today. Motivated by the need for improved data-driven inference, this paper develops a simple new approach for sharp identification of prevailing lead-lag relationships and intensities in low-latency cross-market trading activity based on a class of intuitive model-free measures that can be computed from publicly available trade and quote data with sufficient time stamp granularity. By virtue of capturing the time offset at which co-activity peaks, as well as the magnitude and dispersion of cross-market activity, the measures allow us to shed new light on a range of high-speed cross-market interlinkages in modern electronic markets, which we illustrate in applications to leading U.S. cash and futures markets.

First, while various types of cross-market trading, e.g. basis and spread trading, have historically always been an important component of market activity, the new market structure dominated by automated low-latency trading has given rise to high-speed cross-market inter-

The views expressed herein are those of the authors and should not be interpreted as reflecting the views of the Federal Reserve Board of Governors or any other person associated with the Federal Reserve System.

linkages subject to near deterministic technologically determined time lags usually measured in milliseconds or microseconds.[1] This allows the new cross-market activity measures we propose to reliably uncover key characteristics of low-latency lead-lag relationships between different markets such as the prevailing latency offset, the magnitude of peak cross-activity and the dispersion around it due to latency variations dictated by platform technology, timestamp granularity and market heterogeneity. In this regard, our measures provide a useful complement to popular approaches for studying lead-lag relationships through price cointegration and return correlation measures.[2] In contrast to such existing alternatives, though, our measures utilize only time stamps with sufficient granularity (but no prices or other quantities) and are exceedingly simple and cheap to compute, while also offering sharp identification. In particular, we provide new evidence contrasting U.S. Treasury and equity markets in terms of the relative importance of futures versus cash markets in the underlying low-latency lead-lag relationship. We also document the historical evolution of the prevalence of high frequency cross market trading in these markets that is broadly consistent with known timelines of advances in high speed communications, matching engine technology and the innovation in delivery of market data feeds.

Second, while many positive developments in market functioning due to HFT have been widely acknowledged,[3] we argue that the (price) efficiency gains associated with high-frequency cross-market trading come at the cost of making the real-time assessment of market liquidity across multiple venues more difficult. This is because order placement and execution in one market affects liquidity provision across related markets almost instantly. In particular, we demonstrate that, when applied to transaction times in one market versus quoted depth reduction times in another market, our new measures can help uncover the prevalence of fleeting liquidity in interlinked electronic markets. [4] Our findings point to prudent market making as the

[1]By adopting fiber optic and microwave tower technology, order placement and market information transmission speeds have rapidly been approaching the speed of light. Holden and Jacobsen (2014), among others, have noted some of the arising implications for reliable measurement of market liquidity in fast-paced electronic markets.

[2]See for example, Chan (1992), Huth and Abergel (2014), Laughlin, Aguirre and Grundfest (2014), Godfrey (2014), Benos, Brugler, Hjalmarsson and Zikes (2015), Buccheri, Corsi and Peluso (2018), Hayashi and Koike (2018), among many others.

[3]Notable studies include Brogaard, Hendershott and Riordan (2014),Hasbrouck and Saar (2013), Hendershott and Riordan (2013), Menkveld (2008), O'Hara (2015)

[4]Our measurement approach also allows for more formal data-driven analysis of anecdotal examples of fleeting liquidity such as the one famously described by Lewis (2014).

primary cause of so called liquidity mirages in U.S. Treasury markets, reflecting the challenges faced by large investors in accurately assessing available liquidity based on displayed market depth across different trading venues. Our findings in this regard support the notion that the modern market structure implicitly involves a trade-off between increased price efficiency and heightened uncertainty faced by many investors about overall liquidity available in the market.

Third, we show that, in view of its sensitivity to quoted price changes, low-latency cross-market activity tends to be positively associated with market volatility even after controlling for trading volume and number of transactions. In fact, the prevalence of high frequency cross market activity during episodes of high volatility can be seen as a key indicator of proper market functioning. In this regard, we find a clear distinction between the 2010 U.S. equity market flash crash and the 2014 U.S. Treasury market flash rally in terms of the intensity of high-speed cross-market activity and the associated tightness of the no-arbitrage price link between cash and futures markets during the market stresses. Both events were characterized by spikes in volatility and trading volume but while the latter exhibited well functioning markets and a spike in low-latency cross-market activity, the former saw a material breakdown in the cash-futures basis and a contemporaneous persistent drop in high-speed cross-market activity. These findings point to the important role played by high-speed cross-market trading in maintaining proper market functioning during market stress, consistent with related findings by Menkveld and Yueshen (2019). More broadly, we study the evolving link between cross-market activity and volatility in both U.S. Treasuries (between the ten-year Treasury note cash and futures markets) and equities (between the S&P 500 cash ETF and E-mini futures markets) over more than a decade from January 1, 2004 to September 30, 2015. Consistent with the rise in HFT, our model-free measure of cross-market activity expressed as the peak number of cross-active milliseconds (across all offsets) has become more strongly associated with volatility than trading volume and the number of trades in each market. This finding may simply reflect the fact that volatility can create brief dislocations in relative values spurring bursts of cross-market activity by high-frequency traders seeking to exploit short-lived trading opportunities. When liquidity is ample, cross-market activity can therefore capture incremental information about market volatility beyond traditional measures of overall market activity such as trading volume and the number of transactions. By contrast, the existing large body of empirical findings and alternative theories regarding the relationship between trading activity and volatility involve measures of overall market activity without explicitly accounting for feedback effects between volatility and the part of overall trading activity due to low-latency cross-market trading.[5] We further note that while studies such as Andersen et al. (2015), Ané and Geman (2000), Clark (1973), or Kyle and Obizhaeva (2016) aim to rectify a particular theory-implied form of the relationship between volatility and trading activity in a given single market, our main focus is to show that the cross-market component of high-frequency trading activity between closely related markets contains extra information about market volatility unspanned by trading volume and the number of trades in each market. Thus, we help establish cross-market activity as an increasingly more important driver of the evolving link between trading activity and volatility, thereby adding also to the long-standing literature on intrinsic time in finance. Overall, our findings show that accounting for high-speed cross-market trading is important for understanding liquidity provision and volatility transmission in modern electronic markets. Our measurement approach further yields testable implications for intrinsic times in finance such as endogeneity between volatility and cross-market activity and the cross-market interdependence in observed activity times.

The paper proceeds as follows. Section 2 introduces our model-free measures of cross-market activity based solely on timestamps, lays down the underlying theory and discusses key properties. Section 3 provides an empirical validation for studying lead-lag relationships and interdependence between transaction times. Section 4 shows an application to studying cross-market interdependence in quote update times and associated market liquidity responses to executed trades. Section 5 explores the link between low-latency cross-market activity and market volatility both in normal times and during notable episodes of market stress (flash events). Section 6 summarizes our main findings and testable implications for intrinsic times in finance along with associated avenues for future research.

[5] The vast literature on this subject listed in chronological order includes Ying (1966), Clark (1973), Epps and Epps (1976), Tauchen and Pitts (1983), Karpoff (1987), Schwert (1989), Harris (1987),Jones, Kaul and Lipson (1994), Andersen (1996), Bollerslev and Jubinski (1999), Ané and Geman (2000), Kyle and Obizhaeva (2016), Andersen, Bondarenko, Kyle and Obizhaeva (2015), among many others.

2. Measuring Low-Latency Market Interlinkages from Trading Activity Timestamps

We measure cross-market activity using transaction-level data with millisecond or higher precision for a pair of related markets such as the ones for benchmark U.S. Treasury notes traded on BrokerTec versus Treasury futures contracts traded on the CME or the ones for the SPDR S&P 500 ETF traded on NYSE versus E-Mini S&P 500 futures traded on the CME. This allows us to first identify "active" milliseconds in each market when activity (e.g. trades) occur and then compare the timing offset (millisecond lead or lag) of activity for two distinct instruments or markets. In particular, we define a measure of cross-market activity at each offset as the proportion of milliseconds with coincident trading at that offset relative to the total number of active milliseconds in the less active of the two markets. We further correct our measure for the expected proportion of such coincident activity that would occur if all trading were independent across markets. The measure can therefore be interpreted as the excess proportion of trading (relative to the least active market) accounted for by cross-market activity.

2.1. A Timing-Based Measure of Cross-Market Activity

To be specific, consider two markets, A and B, whose individual activity indicators given by $\mathbb{1}_{\{\text{market } A \text{ active in period } i\}}$ and $\mathbb{1}_{\{\text{market } B \text{ active in period } i\}}$ are observed over N time periods, $i = 1, \ldots, N$. Also observed at any integer offset $t \in [-N, N]$ is their cross-activity indicator $\mathbb{1}_{\{\text{market } A \text{ active in period } i\} \cap \{\text{market } B \text{ active in period } i+t\}}$. These indicators could reflect any particular market activity of interest such as the occurrence of trade executions, order submissions, quote updates, message cancellations, etc. whether it being for the market as a whole or only for some trading entities. The corresponding sequences of occurrence times in $\mathbb{R}^+$ for each of the two markets are given respectively by the two point processes $P^A = \{i_1^A, i_2^A, ..., i_{N_A}^A\}$ and $P^B = \{i_1^B, i_2^B, ..., i_{N_B}^B\}$, with corresponding sample sizes $\|P^A\| = N_A$ and $\|P^B\| = N_B$ getting large (almost surely) as N gets large.

This leads to the following two equivalent model-free ways of defining the raw cross-market activity measure at offset t:

Definition 1 (Raw cross-market activity at offset t: $\mathcal{X}_t^{\text{raw}}$).

$$\mathcal{X}_t^{\text{raw}} = \sum_{i=|t|}^{N-|t|} \mathbb{1}_{\{\text{market } A \text{ active in period } i\}\cap\{\text{market } B \text{ active in period } i+t\}} \tag{1}$$

$$\mathcal{X}_t^{\text{raw}} = \sum_{i_j^A \in P^A \cap [|t|, N-|t|]} \; \sum_{i_k^B \in P^B} \mathbb{1}_{\{i_k^B = i_j^A + t\}} \tag{2}$$

The diagram in Figure 1 illustrates the computation where six active time buckets, marked with “T”, are observed in market A along with nine active time buckets in market B. Applying the definition, at zero offset we find $\mathcal{X}_0^{\text{raw}} = 2$, while at offset $+1$ (using the convention that the offset is applied to the first market) we find $\mathcal{X}_1^{\text{raw}} = 5$ in this example.

For any one of the markets, the raw cross market activity measure defined above can be further decomposed into a product of the number of active time buckets times the fraction of time buckets that were cross active. The latter ratio represents a measure of relative cross-market activity which is interpretable as the fraction of activity in that market (potentially) related to cross market activity at any offset. For market A, the relative activity measure at offset t is thus given by

$$\mathcal{X}_t^{\text{rel,A}} = \frac{\sum_{i=|t|}^{N-|t|} \mathbb{1}_{\{\text{market } A \text{ active in period } i\}\cap\{\text{market } B \text{ active in period } i+t\}}}{\sum_{i=|t|}^{N-|t|} \mathbb{1}_{\{\text{market } A \text{ active in period } i\}}}$$

An analogous expression applies also for the other market B. When analyzing a pair of markets which may have very different levels of activity across time, it is advantageous to take the relative measure for the less active market so that the ratio can vary within the full range from 0 to 1. This leads to the following definition of relative cross-market activity at offset t expressed in two alternative ways:

Definition 2 (Relative cross-market activity at offset t: $\mathcal{X}_t^{\text{rel}}$).

$$\mathcal{X}_t^{\text{rel}} = \frac{\sum_{i=|t|}^{N-|t|} \mathbb{1}_{\{\text{market } A \text{ active in period } i\}\cap\{\text{market } B \text{ active in period } i+t\}}}{\min\left[\sum_{i=|t|}^{N-|t|} \mathbb{1}_{\{\text{market } A \text{ active in period } i\}}, \sum_{i=|t|}^{N-|t|} \mathbb{1}_{\{\text{market } B \text{ active in period } i+t\}}\right]} \tag{3}$$

$$\mathcal{X}_t^{\text{rel}} = \frac{\sum_{i_j^A \in P^A \cap [|t|, N-|t|]} \sum_{i_k^B \in P^B} \mathbb{1}_{\{i_k^B = i_j^A + t\}}}{\min\left[\, \| P^A \cap \, [|t|, N-|t|\,] \,\| \,,\, \| P^B \cap \, [|t|, N-|t|\,] \,\| \,\right]} \tag{4}$$

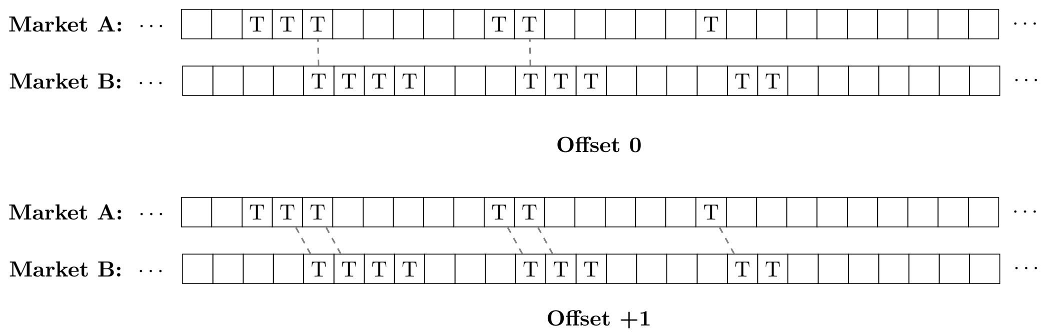

FIGURE 1: **Model-Free Measurement of Cross-Market Activity** The diagram represents the computation of the raw cross-market activity measure $\mathcal{X}_t^{\text{raw}} = \sum_{i=|t|}^{N-|t|} \mathbb{1}_{\{\text{market } A \text{ active in period } i\} \cap \{\text{market } B \text{ active in period } i+t\}}$ for two markets A and B at offset $t = 0$ time intervals (top panel) and offset $t = +1$ time interval (bottom panel). Empty squares in the tape of each market stand for time intervals without market activity in the corresponding market. Squares marked by "T" stand for time intervals when there is market activity in the corresponding market. Dashed lines between active time intervals in the tapes for market A and market B reflect intervals with coincidental activity in both market A and market B. The top panel depicts $\mathcal{X}_0^{\text{raw}} = 2$ given two instances of coincidental activity in both markets at offset $t = 0$ time intervals. The bottom panel depicts $\mathcal{X}_{+1}^{\text{raw}} = 5$ given five instances of coincidental activity in both markets at offset $t = +1$ time intervals. The corresponding values of the relative cross-market activity measure $\mathcal{X}_t^{\text{rel,A}} = \frac{\sum_{i=|t|}^{N-|t|} \mathbb{1}_{\{\text{market } A \text{ active in period } i\} \cap \{\text{market } B \text{ active in period } i+t\}}}{\sum_{i=|t|}^{N-|t|} \mathbb{1}_{\{\text{market } A \text{ active in period } i\}}}$ can then be obtained by way of dividing raw cross-market activity by the number of active time intervals for the less active market of the two, i.e. market A in this example that has six active intervals. This implies $\mathcal{X}_0^{\text{rel}} = \frac{1}{3}$ and $\mathcal{X}_{+1}^{\text{rel}} = \frac{5}{6}$. In practice, the time intervals are meant to be set in accordance with the prevailing latency of information transmission between the two markets of interest and the precision of recorded timestamps, e.g. milliseconds or tenths of milliseconds in the case of leading electronic trading platforms in U.S. Treasury or equity cash and futures markets (separated by a distance roughly equal to that between New York and Chicago, which is known to be covered at the speed of light in about 4-5 milliseconds, depending on the material the signal is traveling through, how close to a straight line is the transmission route, and sub-millisecond variations in latency due to signal interference effects).

The relative measure $\mathcal{X}_t^{\text{rel}}$ can be interpreted as the fraction of activity in the less active market attributable to activity in the other market with an offset t. Applied to the example given in Figure 1, $\mathcal{X}_0^{\text{rel}} = \frac{1}{3}$ while $\mathcal{X}_1^{\text{rel}} = \frac{5}{6}$. As further clarified by this example, the calculations required to obtain both the raw and relative measures of cross-market activity (1)-(2) and (3)-(4) given by definitions 1 and 2 above require only time stamps and are exceedingly simple and cheap (in CPU cycles) to implement. Importantly, the measures do not require a bandwidth choice beyond the time granularity for modelling activity occurrence times. Ideally this should be set to be slightly greater than the fastest transmission time between the analyzed markets to allow for meaningful differentiation between cross-market activity at non-zero lags. For example, one millisecond granularity would be an adequate choice for measuring cross-market activity between New York area cash markets and Chicago area futures markets subject to order transmission latency in the order of 4-5 milliseconds.[6]

For proper measurement and detection of abnormally high levels of cross-market activity it is further necessary to account for purely coincidental cross-market activity taking place between any two highly active markets, even if completely independent of each other. Assuming that the occurrence times i_j^A and i_k^B of the two point processes P^A and P^B become asymptotically independent as $|t| = |i_k^B - i_j^A| \to \infty$, we therefore adjust the relative cross market activity measure by subtracting off its average taken across very large positive and negative offsets. This leads to the following model-free definition of excess cross market activity at offset t:

Definition 3 (Excess cross-market activity at offset t: $\mathcal{X}_t$)**.**

$$\mathcal{X}_t = \mathcal{X}_t^{\text{rel}} - \mathcal{X}_\infty^{\text{rel}}\,, \text{ where } \mathcal{X}_\infty^{\text{rel}} = \frac{1}{2(T_2 - T_1)} \sum_{|t|=T_1+1}^{T_2} \mathcal{X}_t^{\text{rel}} \tag{5}$$

Under ergodicity and asymptotic independence between the lagged observation times for the two markets for large enough lags, this non-parametric adjustment is robust to clustering and varying local intensity rates of market activity. This ensures that $\mathcal{X}_t$ as defined by equation (5) represents a measure of excess cross-market activity at offset t relative to an independence null as formalized by the following two propositions.

Proposition 1 (Convergence to zero under the independence null)**.** *Let P^A and P^B be stationary and ergodic point processes on $\mathbb{R}^+$. Under the null of independence between P^A and P^B the excess cross-market activity measure $\mathcal{X}_t$ convergences almost surely to zero:*

$$\mathcal{X}_t \xrightarrow{a.s.} 0 \tag{6}$$

Proof. Ergodicity and stationarity of P^A and P^B under independence implies ergodicity and stationarity of $\mathcal{X}_t$. Consequently, $\mathbb{E}[\mathcal{X}_t^{\text{rel}}] = \mathbb{E}\left[\frac{1}{2(T_2-T_1)} \sum_{|t|=T_1+1}^{T_2} \mathcal{X}_t^{\text{rel}}\right]$. The ergodic theorem

[6] The higher is the chosen time granularity, the larger would be the dispersion of low-latency cross-market activity over a range of adjacent offsets due to the increasingly more dominant random latency variations at finer time scales.

then implies $\mathcal{X}_t^{\text{rel}} \xrightarrow{a.s.} \mathbb{E}[\mathcal{X}_t^{\text{rel}}]$. Combining, with $\mathbb{E}[\mathcal{X}_t^{\text{rel}}] = \mathbb{E}\left[\frac{1}{2(T_2-T_1)}\sum_{|t|=T_1+1}^{T_2} \mathcal{X}_t^{\text{rel}}\right]$ it follows that $\mathcal{X}_t \xrightarrow{a.s.} 0$. □

Proposition 2 (Asymptotic normality under the independence null)**.** *Let P^A and P^B be stationary and ergodic point processes on $\mathbb{R}^+$. Under the null of independence between P^A and P^B the excess cross-market activity measure $\mathcal{X}_t$ scaled by*

$$\sigma_{\mathcal{X}_\infty} = \frac{1}{\sqrt{2(T_2 - T_1)}} \left[\sum_{|t|=T_1+1}^{T_2} \left(\mathcal{X}_t^{rel} - \mathcal{X}_\infty^{rel} \right)^2 \right]^{1/2} \tag{7}$$

is asymptotically normally distributed:

$$\frac{\mathcal{X}_t}{\sigma_{\mathcal{X}_\infty}} \sim N(0,1) \tag{8}$$

Estimation uncertainty can thus be quantified using a suitably defined ergodic standard deviation estimate $\sigma_{\mathcal{X}_\infty}$. The independence assumption can be further relaxed to allow for finite (or asymptotically vanishing) dependence, thereby preserving the validity of the above two propositons for large enough offsets $|t|$ as $T1$ and hence $T2$ and $T_2 - T_1$ get large with N getting large. This enables inference of statistical significance of any peaks in cross-market activity obtained by computing $\mathcal{X}_t$ across different lags t.

However, gauging statistical significance in this way proves to be largely unnecessary in the empirical applications we consider in what follows. This is because the measured peak levels of $\mathcal{X}_t$ tend to be orders of magnitude larger than the corresponding estimation uncertainty given by $\sigma_{\mathcal{X}_\infty}$. Plotting $\mathcal{X}_t$ across a wide range of offsets t therefore makes it easy to implicitly confirm statistical significance by way of visually inspecting the peak of $\mathcal{X}_t$ vis-a-vis the typically much smaller variability of $\mathcal{X}_t$ in the tails. An important benefit of plotting directly $\mathcal{X}_t$ is its economic interpretability as a share of activity in the less active market attributable to cross-market activity at offset t.

The choice of the values of T_1 and T_2 is not very critical in practice as long as they are kept sufficiently large relative to the prevailing lead/lag offsets. In what follows, we analyze cash and futures markets data with millisecond timestamps and prevailing lead/lag offsets known to be in the range from zero to a few milliseconds. Therefore, we set $T_1 = 500$ milliseconds and $T_2 = 1000$ milliseconds without material change in the obtained results for any other values of $T_2 \gg T_1$ that are orders of magnitude larger than the prevailing lead/lag offset and are sufficiently apart from each other for averaging purposes. More formally, though, we recommend following a data-driven approach to validating these choice of $T1$ and $T2$ based on graphical assessment of the degree of lagged dependence between two point processes following Diggle's randomization testing procedure as in Dutilleul (2011).[7]

2.2. Key Features: Location, Dispersion and Magnitude of Peak Cross-Activity

The three main features of our measure are location, dispersion and magnitude of the peak cross-market activity across different offsets. We illustrate each feature below using a representative data sample over the period from Jul 1, 2014 to Dec 31, 2014 for the five-year and ten-year Treasury notes traded on the BrokerTec platform and the corresponding Treasury futures traded on the Chicago Mercantile Exchange.

2.2.1. Location

Figure 2 shows the relative cross-market activity, $\mathcal{X}_t$, for the ten-year and five-year Treasury notes. Based on our measure, cross-market trading accounts for around 8 percent of activity in the cash Treasury market on normal days (depicted by the gray shaded area) but almost double that on October 15, 2014, which was an exceptionally volatile day in Treasury markets. The location of the peak at zero is intuitive since the Treasury notes are traded on the same platform, and the sharpness of the spike indicates that random latency (i.e. the time it takes the platform to process trading instructions known to be around 0.2 milliseconds on the Brokertec cash platform) is small compared to the millisecond resolution at which the analysis is carried out.

2.2.2. Dispersion

The cross-market activity between five- and ten-year Treasury futures also exhibits a spike at zero offset (figure 3), indicating a similarly significant amount of nearly instantaneous trading

[7]This additional validation of the choice of T_1 and T_2 in our data is available upon request.

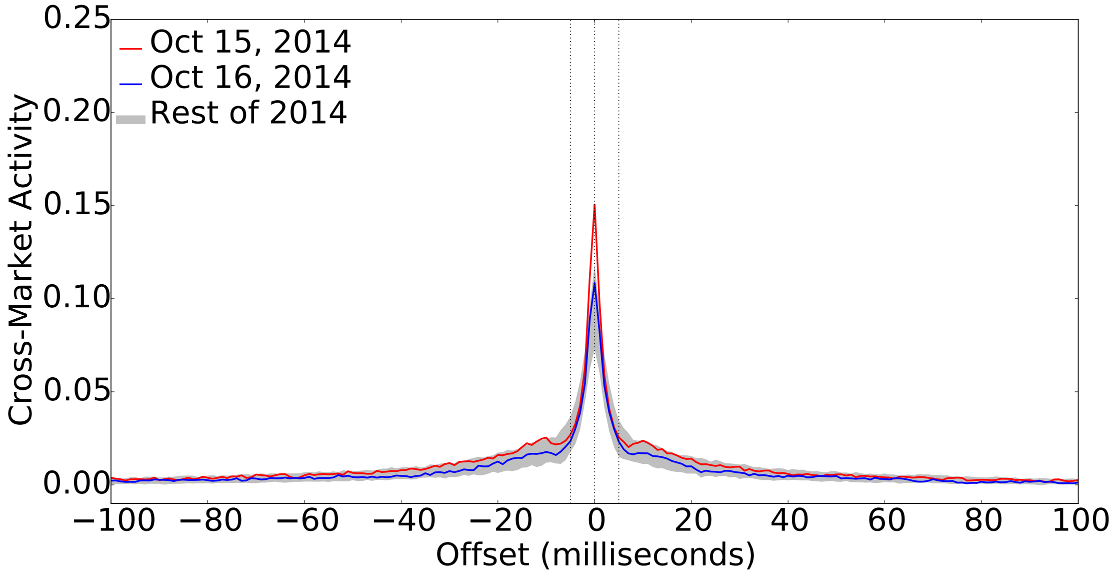

FIGURE 2: **Cross-Market Trading Activity Between 10-Year and 5-Year Treasury Cash Markets.** The intensity $\mathcal{X}_t$ of excess cross-market activity between the 10-Year and 5-Year Treasury cash markets is plotted at different offsets t applied to ten-year cash market time stamps. Vertical dashed lines are shown at 0 and +/-5 milliseconds. Excess cross-market activity is expressed as a fraction of ten-year cash market activity (the least active of the pairing). **Data Source:** BrokerTec (cash market data).

in these instruments. However, the spike is much more diffuse: it spreads over a much wider range of offsets. This may reflect a wider range of available connectivity options to CME as well as some occasional platform latency associated with Treasury futures trading.[8] Comparing figures 2 and 3 shows how the defined non-parametric measure of cross-market activity can be used to study the impact of market structure and technology differences on the observed trading behaviors.

2.2.3. Magnitude

To show how the magnitude of cross-market activity can be used to gauge the evolution of high-frequency cross-market trading along the yield curve, in figure 4 we plot the historical levels of cross-market trading at zero millisecond offset in the ten-year and five-year Treasury note on BrokerTec over the past decade. Early on, during 2004-2006, there was essentially no simultaneous trading in the two markets. Starting with a systems upgrade of the cash platform

[8]For more in-depth discussion of latency in futures trading, see Joint Staff Report (2015).

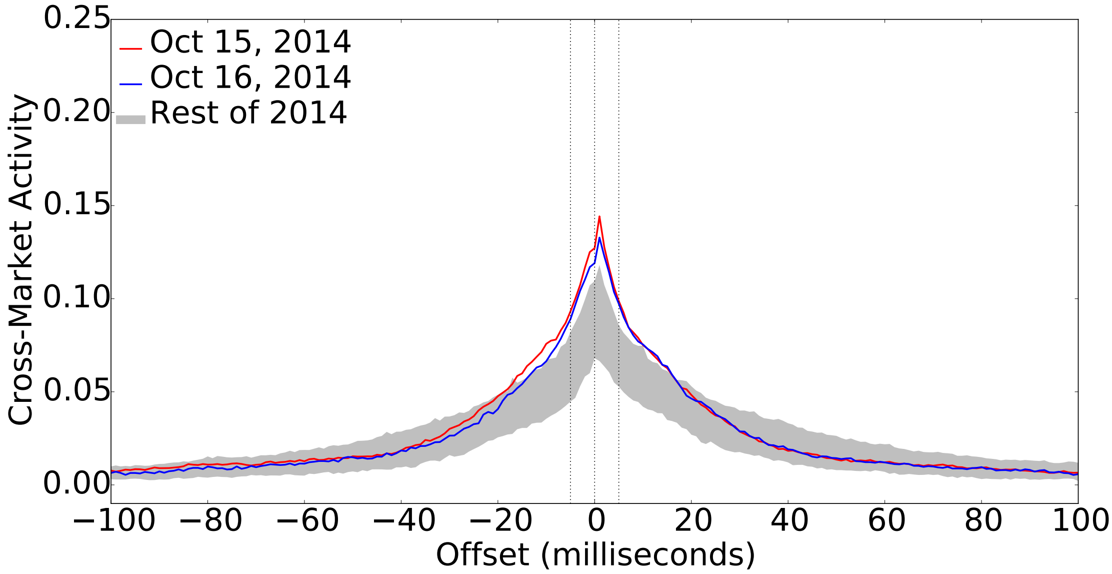

FIGURE 3: **Cross-Market Trading Activity Between 10-Year and 5-Year Treasury Futures Markets.** The intensity $\mathcal{X}_t$ of excess cross-market activity between the 10-Year and 5-Year Treasury futures markets is plotted at different offsets t applied to five-year futures market time stamps. Vertical dashed lines are shown at 0 and +/-5 milliseconds. Excess cross-market activity is expressed as a fraction of five-year futures market activity (the least active of the pairing). **Data Sources:** Nanotick (CME futures market data).

in March 2006, minor evidence of cross-market activity appears. In March 2012, a major upgrade to the BrokerTec market data feeds and matching engine significantly reduced platform latency, and synchronized trading activity more than doubled.[9] A more gradual but larger increase followed by mid to late 2013, as a number of further upgrades to order submission and market data protocols took place, with synchronized trading activity reaching an average of around 10 percent of total trading activity.[10] This example shows how the non-parametric measure of cross-market activity can be used to track changes in trends and patterns of trading behaviors resulting from discrete technological change.

[9]Specifically, BrokerTec introduced a new trading system for its U.S. based products using a modified version of the NASDAQ OMX Genium INET system, with changes taking effect on March 26, 2012. The upgraded platform enabled BrokerTec to increase order volume tenfold and decrease order input latency by up to 50 times. The platform's average latency dropped to less than 200 microseconds, with previous average latency approximately 10 milliseconds. Source: "BrokerTec Brings World Class Trading Platform To U.S. Fixed Income Market", NASDAQ Inc. press release, April 18, 2012.

[10]In particular these upgrades included the introduction of the ITCH (mid to late 2013) and OUCH (November 2012) protocols. OUCH is a low-level high performance native protocol that allows participants to enter, replace, and cancel orders and receive executions in an automated fashion on the Genium INET platform. ITCH is a high performance direct data feed protocol that displays all public orders and trades that occur on the platform.

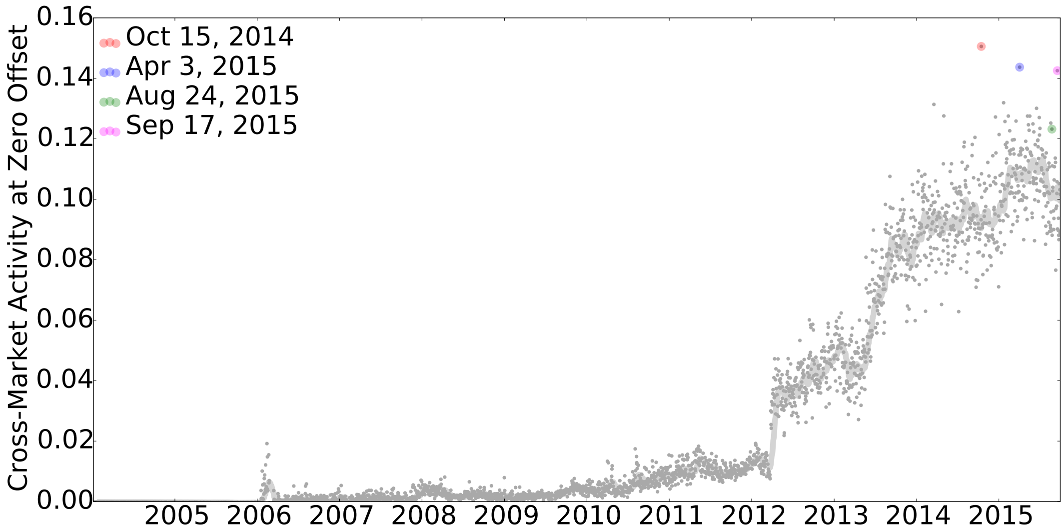

FIGURE 4: **Evolution of Cross-Market Trading Activity Between 10-Year and 5-Year Treasury Cash Markets.** The intensity $\mathcal{X}_t$ of excess cross-market activity between the 10-Year and 5-Year Treasury futures markets at zero millisecond offset of the time stamps is calculated daily over 2004-2015. The scatter plot of the obtained daily measures (gray dots) further singles out the following high-volatility days: October 15, 2014 (red dot), April 3, 2015 (blue dot), August 24, 2015 (green dot), September 17, 2015 (pink dot). Excess cross-market activity is expressed as a fraction of ten-year cash market activity (the least active of the pairing). **Data Sources:** BrokerTec (cash market data).

3. Low-Latency Lead-Lag Relations between Cash and Futures Markets

In this section, we apply the proposed model-free measures of cross-market activity to analyse low-latency lead-lag relationships between the cash markets and the associated derivatives markets for U.S. Treasuries and U.S. Equity indices. While lead-lag relationships have been studied extensively in the literature by comparing observed (often noisy) market returns, our measurement approach is able to shed new light by relying instead solely on the timing of activity based on millisecond or higher resolution time stamps.

The Chicago Mercantile Exchange matching engine is located in Aurora Illinois and is connected via fiber optics and microwave technology to the cash market exchanges. The Brokertec platform is located in Secaucus New Jersey, roughly 4.7 milliseconds from Aurora at the speed of light, which puts a natural lower bound on the offset at which cross-market activity may be expected to take place in more recent years.[11] The main US equity exchanges are primarily located in northern New Jersey with negligibly different distance to Aurora.[12] As we shall see throughout the analyses presented below, the roughly 5ms offset is clearly visible in our analysis and confirms huge impact of technology on trading behaviors observed in practice.

We analyze the link between cross-market activity and volatility in both US Treasuries (10-Year Treasury note cash and futures) and equities (S&P500 E-mini and SPY ETF) over the six-month period from July 1, 2014 to Dec 31, 2014. We measure cross-market activity on each trading day as the number of cross-active milliseconds at its peak across all offsets and restrict attention to the most active US portion of the electronic trading hours for each pair: 7:00-16:00 ET for the 10-Year T-Note and 9:30-16:00 ET for S&P500. While we carry out the analysis at millisecond frequency, it trivially generalizes to any other frequency with adequate time resolution for meaningful cross-activity measurements at different offsets.

3.1. The US Treasury Cash and Futures Markets

Figure 5 displays the cross-market activity measure for the cash and futures ten-year Treasury note markets. The measure is shown for October 15, 2014, a day with extraordinarily high volatility and trading; for October 16, 2014, another day with somewhat heightened volatility and trading; and for the rest of the other, more typical, trading days in October 2014.

The cash market platform (in Secaucus, New Jersey) and the futures market exchange (in Aurora, Illinois) exhibit very pronounced cross-market trading activity at an offset of +/-5 milliseconds, strikingly consistent with the time it takes to transmit information between the two trading venues using current microwave tower technology. Spikes occur at both negative and positive offsets, suggesting that sometimes Treasury futures lead cash Treasuries (+5 millisecond offset) and other times the cash market leads futures (-5 millisecond offset), a remarkable fact since it uncovers a more symmetric low-latency lead-lag relationship than the one known to be prevalent in the S&P 500 cash and futures markets.[13] The dip at zero is expected because coordinated trading within a millisecond is not possible (although the local leg of a trade could

[11] Since the refractive index of glass is around 1.5, communication via fiber-optic technology leads to a lower bound of roughly 7 ms.

[12] BATS in Weehawken, DirectEdge in Secaucus, Nasdaq in Carteret, and NYSE in Mahwah, NJ.

[13] See for example Laughlin et al. (2014), among others.

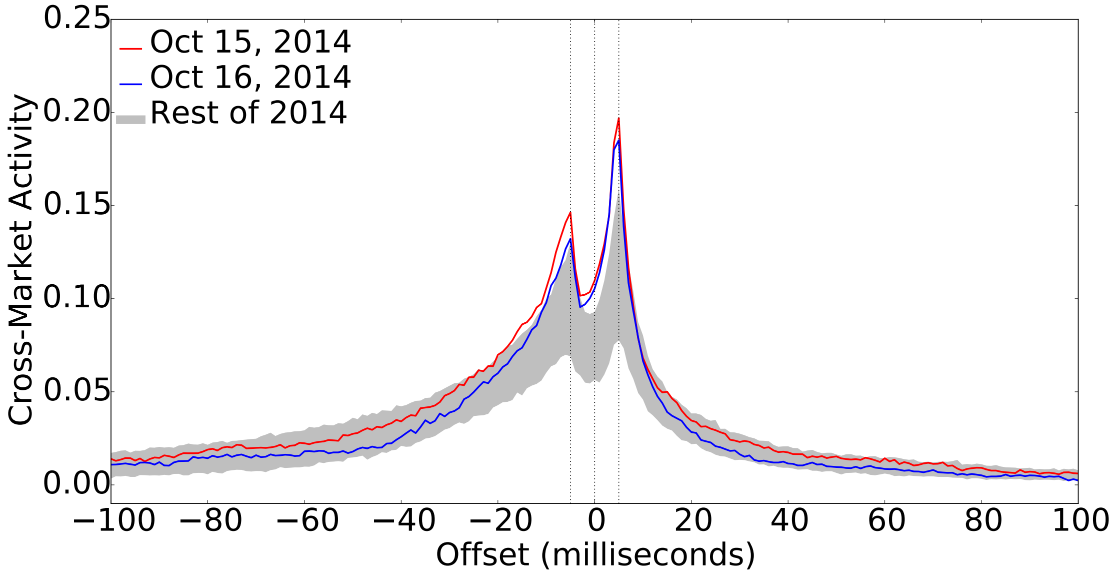

FIGURE 5: **Cross-Market Trading Activity Between 10-Year Treasury Cash and Futures Markets.** The intensity $\mathcal{X}_t$ of excess cross-market activity between the 10-Year Treasury cash and futures markets is plotted at different offsets t applied to cash market time stamps. Vertical dashed lines are shown at 0 and +/-5 milliseconds. Excess cross-market activity is expressed as a fraction of cash market activity (the least active of the pairing). **Data Sources:** BrokerTec (cash market data); Nanotick (CME futures market data).

purposely be delayed). However, the measure does not reach zero since trading activity tends to occur in short bursts and some amount of spurious cross-market activity may be picked up at very small offsets.

Overall, the large share of cross-market activity probably helps explain the extremely tight link between Treasury cash and futures markets. Consistent with this, the two highest volatility days in October 2014, the 15th and 16th, also had the most cross-market activity.

3.2. Measure Validation: Who Are the High-Speed Cross-Market Traders?

The ability of the cross-market activity measure to capture cross-market trading exclusively by modern high-speed traders has been further validated by its use in the Joint Staff Report (2015) on the October 15, 2014 U.S. Treasury markets flash rally. Figure 6 reproduces a comparison from the Joint Staff Report (2015) between the measured intensity of high-speed cross-market activity in Treasury cash and futures by the top 10 traditional bank/dealers that are not known to heavily use high-speed trading technology (left panel) and the top 10 principal

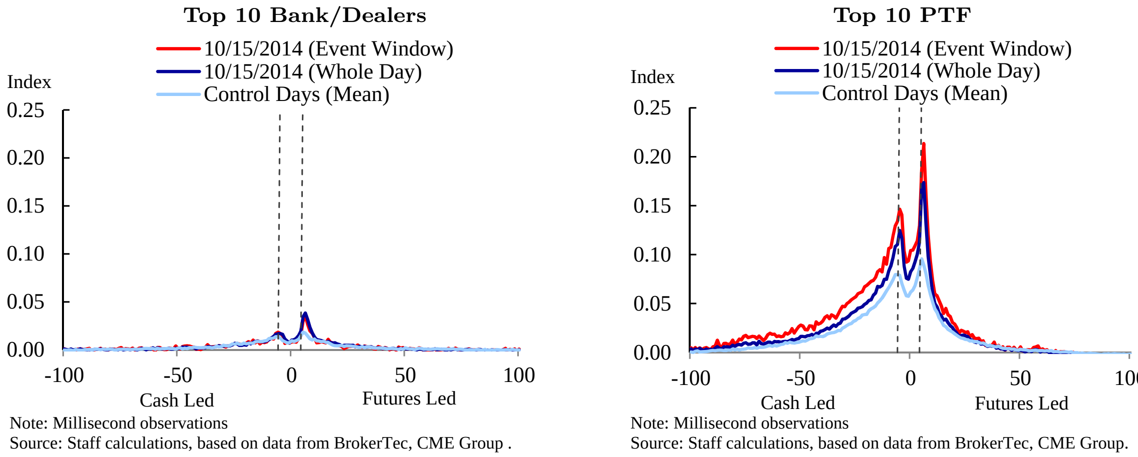

FIGURE 6: **Validation of Measured Cross-Market Trading Activity Between 10-Year Treasury Cash and Futures Markets.** The figure reproduces Figures B.1 and B.2 from the Joint Staff Report (2015) on the October 15, 2014 U.S. Treasury markets flash rally. **Left Panel:** Cross-market activity by the top 10 bank/dealers. **Right Panel:** Cross-market activity by the top 10 principal trading firms (PTF).

trading firms (PTF) known for their heavy use of such technology (right panel). As can be seen from this comparison, only the PTF cross-activity pattern shown in the right panel of Figure 6 is consistent with the market-wide pattern from the middle panel of Figure 5. This confirms that the cross-market activity measure successfully identifies exclusively the low-latency cross-market activity of modern high-speed traders. Importantly, it accomplishes this by utilizing solely high-precision time stamps from the anonymous record of all trading activity, even in the absence of any entity identifying information.

3.3. The S&P 500 E-Mini Futures and Cash Markets

Figure 7 displays the average measure of cross-market activity for the S&P 500 cash and futures markets. Of note is the pronounced asymmetry of the spike in the measure at +5 milliseconds for the S&P 500 compared to the 10-Year US Treasury. The much higher spike for the positive 5ms offset is consistent with the widely known leading role played by the S&P futures. Unlike prior studies based on returns and correlation measures, though, our measure successfully uncovers this relationship solely based on trading activity time stamps and pins down the precise offset with very high precision.[14]

[14]For comparison, see Laughlin et al. (2014), among others.

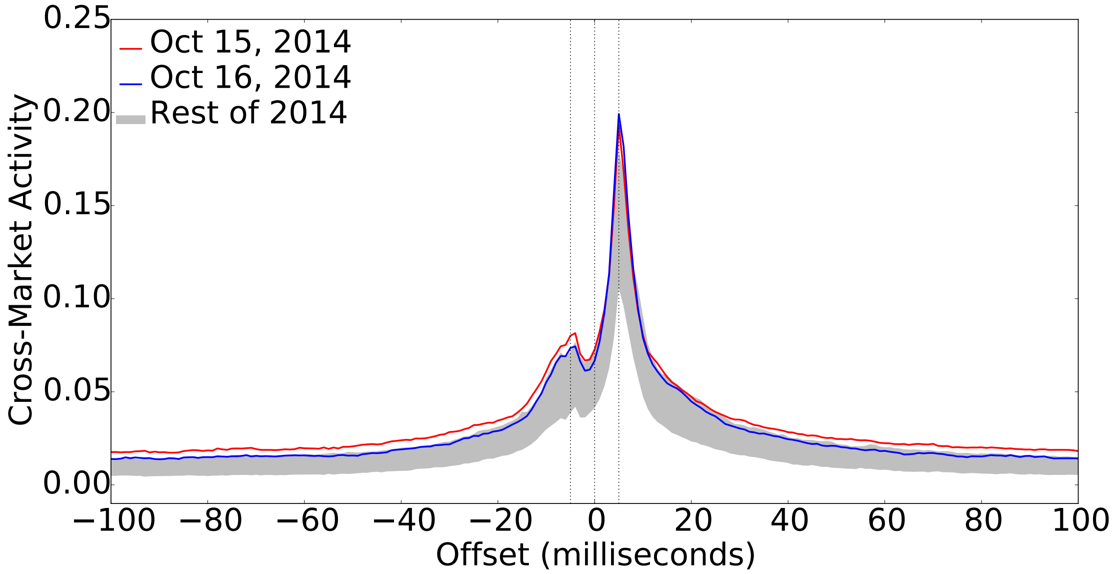

FIGURE 7: **Cross-Market Trading Activity Between S&P 500 Cash and Futures Markets.** The intensity $\mathcal{X}_t$ of excess cross-market activity between the S&P 500 cash and futures markets is plotted at different offsets t applied to cash market time stamps. The cash market is represented by the SPDR S&P 500 ETF (SPY). The futures market is represented by the E-Mini S&P 500 futures contract. Vertical dashed lines are shown at 0 and +/-5 milliseconds. Excess cross-market activity is expressed as a fraction of cash market activity (the least active of the pairing). **Data Sources:** Thomson Reuters Tick History (cash market data); Nanotick (CME futures market data).

3.4. The US Treasury Cash and S&P 500 E-Mini Futures Markets

A different pattern emerges when we compare trading activity across substantially different highly active markets, underscoring the fact that the cross-market activity measure is not simply picking up spurious coactivity. For example, overall cross-market activity in the ten-year Treasury notes and E-mini S&P 500 futures is much lower (see figure 8) and indicates that, to the extent that cross-market activity does occur, the E-mini market leads the cash Treasury market (+5 millisecond offset) but not vice versa.

3.5. Tracing the Evolution in High-Speed Cross-Market Trading Activity between Cash and Futures Markets

The ability of the cross-market activity measure to quantify the intensity of low-latency cross-market activity by high-speed traders opens up the possibility to gauge in a purely data-driven way the otherwise largely anecdotal evolution of high-speed cross trading activity over time between different cash and futures markets. As a first step in this direction, Figure 9

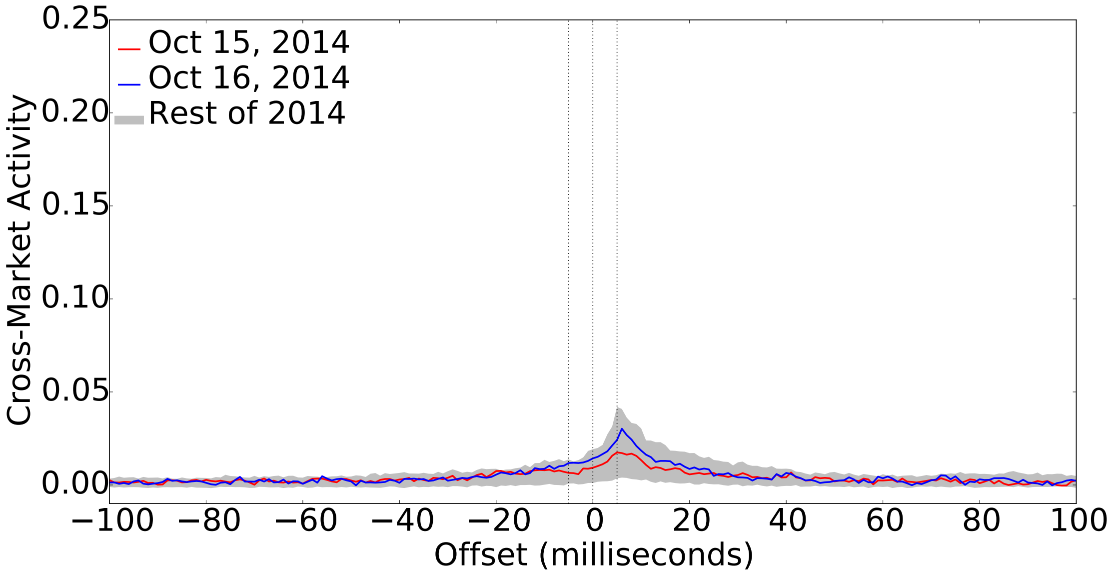

FIGURE 8: **Cross-Market Trading Activity Between 10-Year Treasury Cash and E-Mini S&P 500 Futures Markets.** The intensity $\mathcal{X}_t$ of excess cross-market activity between the 10-year Treasury cash and E-Mini S&P 500 futures markets is plotted at different offsets t applied to cash market time stamps. Vertical dashed lines are shown at 0 and +/-5 milliseconds. Excess cross-market activity is expressed as a fraction of cash market activity (the least active of the pairing). **Data Sources:** BrokerTec (cash market data); Nanotick (CME futures market data).

depicts the obtained data-driven reconstruction of the growth in high-speed cross-trading over the eight year period from 2010 to 2017 for ten-year Treasury cash and futures vis-à-vis S&P 500 cash and futures.

The left panel of Figure 9 reveals that high-speed cross-market trading has consistently been more intense in S&P 500 cash and futures (the red line) than in Treasury cash and futures markets (the blue line). The respective prevailing levels appear to have largely stabilized at around 15 percent of overall activity for the S&P 500 (since 2013) and nearly 10 percent of overall activity for Treasuries (since 2014). High-speed cross-market trading for S&P 500 stood around 10 percent of overall activity even as early as 2010, while in the case of U.S. Treasuries the level of high-speed cross-market trading was pretty negligible until 2012. The rapid rise in high-speed cross-market trading after 2012 is consistent with the major upgrade to the BrokerTec trading engine discussed in section 2.2.3.

The right panel of Figure 9 further illustrates that the prevailing millisecond lead-lag of high-speed cross-market activity has been steadily declining since 2010. The uncovered pattern

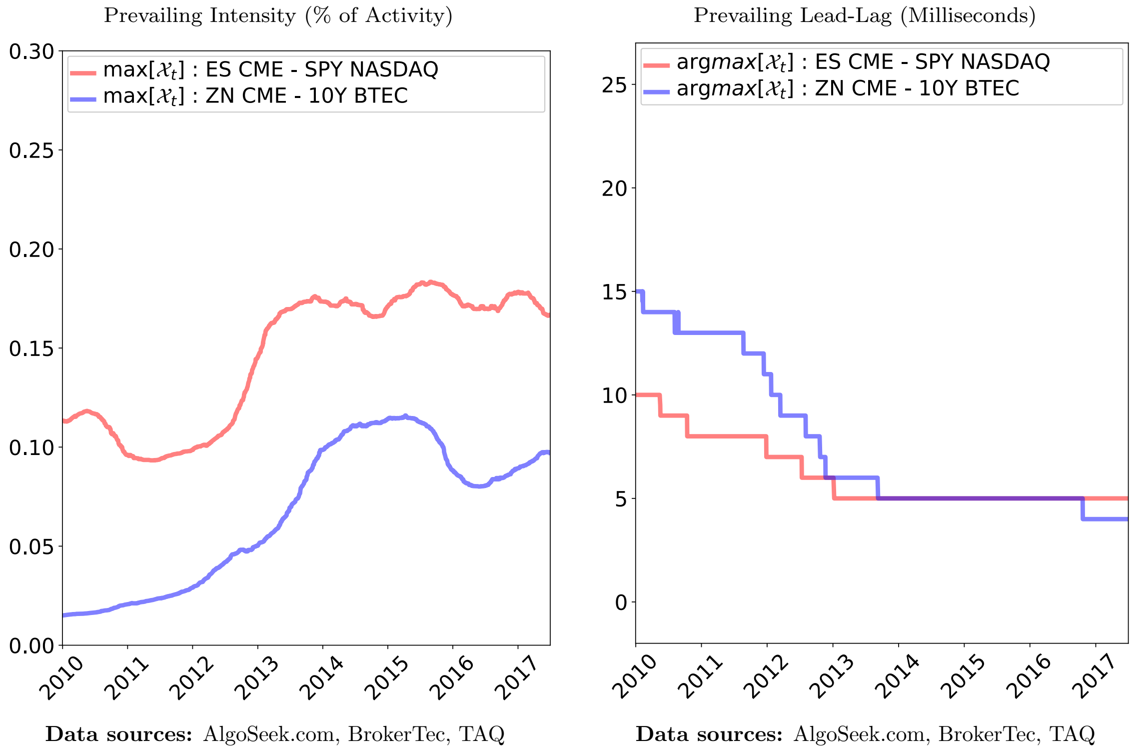

FIGURE 9: **Evolution of the Intensity and Prevailing Lead-Lag of High-Speed Cross-Market Trading between Cash and Futures Markets.** We plot the prevailing intensity (left panel) and millisecond lead-lag (right panel) of high-speed cross-market trading activity between S&P 500 cash and futures markets (in red) and 10-Year Treasury Note cash and futures markets (in blue). **Left Panel:** Prevailing Intensity of Cross-Market Activity between Cash and Futures Markets. **Right Panel:** Prevailing Lead-Lag of Cross-Market Activity between Cash and Futures Markets.

of declines is strikingly consistent with the fiber-optic and microwave communication technology improvements over this period and the BrokerTec trading engine in 2012 that were discussed in section 2.2.3. In 2013 for the S&P 500 (red line) and soon after also for Treasuries (blue line) the prevailing lead-lag stabilized close to the fastest connectivity speed between these two markets near its theoretical minimum for the speed of light. Overall, the uncovered patterns are in agreement also with anecdotal reporting that the rise of high-speed trading in U.S. Treasury markets has been more recent compared to similar developments in equity markets. As such, the obtained quantitative measurements should provide a good basis for comparing and controlling for the varying intensity of high-speed cross-market activity over time.

4. Cross-Market Activity and Fleeting Liquidity

In this section we aim to demonstrate that by relying solely on high-precision timestamps the proposed measurement approach is more generally applicable than return-based measures of lead-lag relationships in cross-market activity. Market efficiency is often pointed to as a main benefit of automated and high-frequency trading (HFT) in U.S. Treasury markets. Fresh information arriving in the market place is reflected in prices almost instantaneously, ensuring that market makers can maintain tight spreads and that consistent pricing of closely related assets generally prevails. While the positive developments in market functioning due to HFT have been widely acknowledged, we argue that the (price) efficiency gain comes at the cost of making the real-time assessment of market liquidity across multiple venues more difficult. This situation, which we term the liquidity mirage, arises because market participants respond not only to news about fundamentals but also market activity itself. This can lead to order placement and execution in one market affecting liquidity provision across related markets almost instantly. The modern market structure therefore implicitly involves a trade-off between increased price efficiency and heightened uncertainty about the overall available liquidity in the market.

Our measurement approach of cross-market activity makes it possible to assess this trade-off in a purely data-driven fashion. To this end, we consider jointly the trading and quote update activity in the ten-year U.S. Treasury note and futures in the month of October 2014 during which the October 15 "flash rally" occurred. Trading in the cash market takes place on the BrokerTec and eSpeed interdealer platforms (both located in Secaucus, New Jersey), while the corresponding front-month U.S. Treasury futures contract trades on the Chicago Mercantile Exchange (CME). All three venues feature anonymous electronic central limit order books with trading largely dominated by principal trading firms and bank dealers that often employ automated and low-latency trading techniques, as documented in Joint Staff Report (2015).

To an investor in the U.S. Treasury market, the eSpeed, BrokerTec, and CME platforms represent distinct liquidity pools which typically exhibit tight spreads and significant depth at the top of the order book. Thanks to low-latency cross-market trading activity, the prices on all three platforms are likely to be competitive (market efficiency at work!). However, the liquidity effectively available to the investor at any point in time is highly unlikely to be the sum of top-of-book depths across the platforms. To see why, consider a New York-based investor submitting a buy order to all three venues. Given the short distance to the interdealer platforms, the investor's orders will reach one of them first. Suppose the BrokerTec order is matched first and the trader gets the requested quantity at the best offer. As soon as the BrokerTec transaction is observed in the market data feed, colocated low-latency market participants may immediately seek to cancel top-of-book offers on eSpeed and CME or submit competing buy orders to eSpeed and CME.

The former would be consistent with prudent risk management by market makers while the latter would be an example of opportunistic trading ahead of anticipated order flow, also known as front running. Due to random fluctuations in network latency and the close proximity between eSpeed and BrokerTec, the investor's order may or may not arrive in time to get filled at the expected best offer on eSpeed. However, low-latency traders will almost certainly be able to preempt the order's arrival at CME, leaving it in a position where it may not be filled at all. Hence, the displayed market depth across distinct liquidity pools can convey a misleading impression of the aggregate available liquidity.

4.1. Causes of the Liquidity Mirage: Prudent Market Making

To investigate how low-latency liquidity providers respond to incoming market data, we study the order book reactions to trades across platforms based on cash market data from BrokerTec and CME futures market data from Nanotick. We focus our analysis on the extent to which CME trades may cause a reduction in depth on BrokerTec, as captured by the cross-market trading activity measure $\mathcal{X}_t$. In this case, the measure can be interpreted as the proportion of CME trades associated with top-of-book depth reduction on BrokerTec in excess of what might be expected by pure chance.

Figure 10 shows that on average as much as 20 percent of Treasury futures trades at the bid (offer) are associated with depth reduction at the bid (offer) on the BrokerTec platform. Moreover, this BrokerTec order book reaction to CME trades peaks at a roughly 5 millisecond delay, which matches the current shortest possible transmission time between the two venues using cutting-edge microwave transmission technology. The evidence therefore supports the hypothesis that rapid depth reduction by low-latency liquidity providers contributes to the

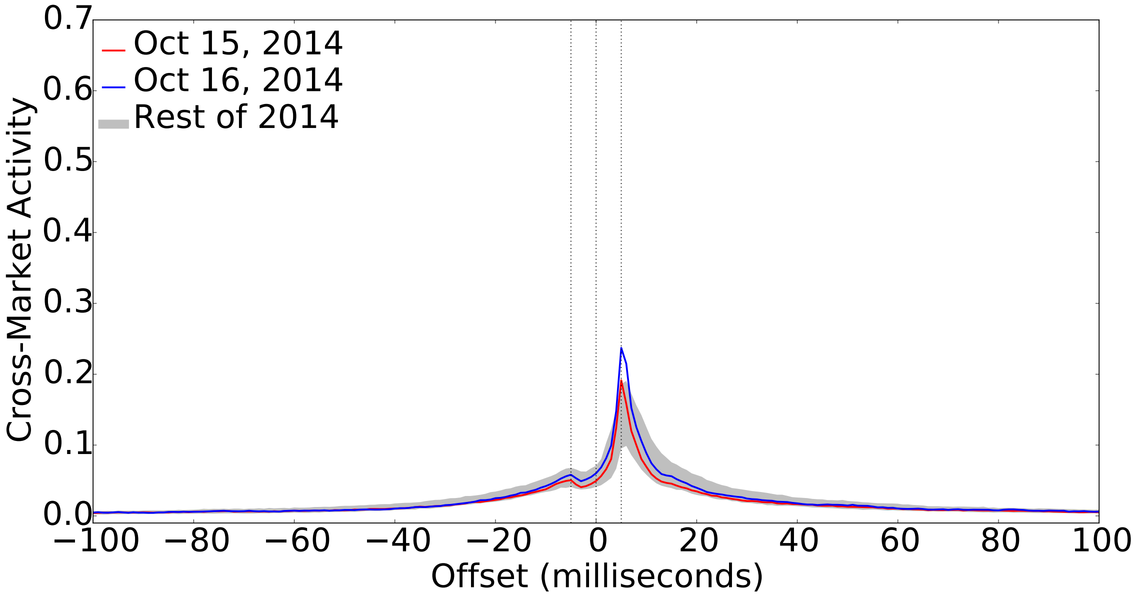

FIGURE 10: **CME Trades versus Depth Reductions on BrokerTec.** Cross-market activity is expressed as a share of futures trades and reflects the incidence rate of a trade at the bid (offer) on CME coinciding with a reduction in depth at the best bid (offer) on BrokerTec. A positive offset indicates that a CME trade happens prior to a depth reduction on BrokerTec. The gray area represents the 2.5th to 97.5th interpercentile range of the daily cross-market activity measures over the second half of 2014. **Data Sources:** BrokerTec (cash market data); Nanotick (CME futures market data).

liquidity mirage. It is also worth highlighting that October 15, while exhibiting extreme price volatility, is not an outlier in this regard. As such, we did not find any evidence that the liquidity mirage was more pronounced on October 15 compared with our control days.

An even stronger link exists between large CME trades (those with a trade size greater than fifty contracts) and top-of-book depth changes on BrokerTec. Figure 11 displays a remarkable spike indicating that nearly 60 percent of large CME trades are followed by BrokerTec depth reductions with an offset of 5 milliseconds. These findings underscore the fact that the total liquidity available to an investor at a given point in time in practice may tend to be closer to the depth of the market that gets accessed first rather than the sum of the depths in each market.

We stress that the documented fleeting liquidity patterns are entirely consistent with prudent market making in an anonymous central limit order book environment where the informational advantage of market makers does not lie in having proprietary access to customer flows (as New York Stock Exchange specialists once did), but rather in their speed and ability to process

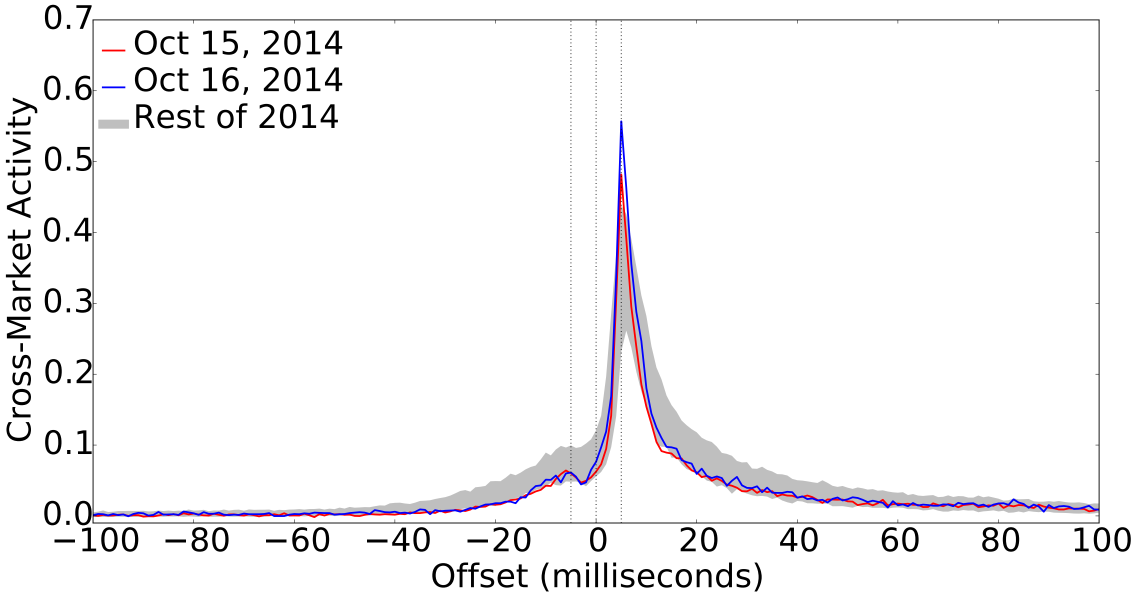

FIGURE 11: **Large CME Trades versus Depth Reductions on BrokerTec.** Cross-market activity is expressed as a share of futures trades and reflects the incidence rate of a trade at the bid (offer) on CME coinciding with a reduction in depth at the best bid (offer) on BrokerTec. A positive offset indicates that a CME trade happens prior to a depth reduction on BrokerTec. The gray area represents the 2.5th to 97.5th interpercentile range of the daily cross-market activity measures over the second half of 2014. **Data Sources:** BrokerTec (cash market data); Nanotick (CME futures market data).

complex market data.

4.2. Causes of the Liquidity Mirage: Opportunistic Directional Trading or Front Running

A second potential cause for the liquidity mirage is that traders with a speed advantage may attempt to trade ahead of (expected) incoming order flow. In this scenario, they would react to buy (sell) orders at CME by immediately submitting buy (sell) orders on BrokerTec in anticipation that prices may tick up (down) due to incoming order flow. However, the evidence for this behavior turns out to be quite weak in the data, based on our cross-market activity measure.

Figure 12 shows the cross-market activity measure for buy and sell trades on CME and BrokerTec. There is modest evidence that buy (sell) trades take place in a coordinated fashion at an offset of +/-5 milliseconds, but this accounts for very little of the overall trading activity on the BrokerTec platform (the less active of the two markets). For large orders (not shown here), the evidence is even weaker even though the incentives to front run would be even greater.

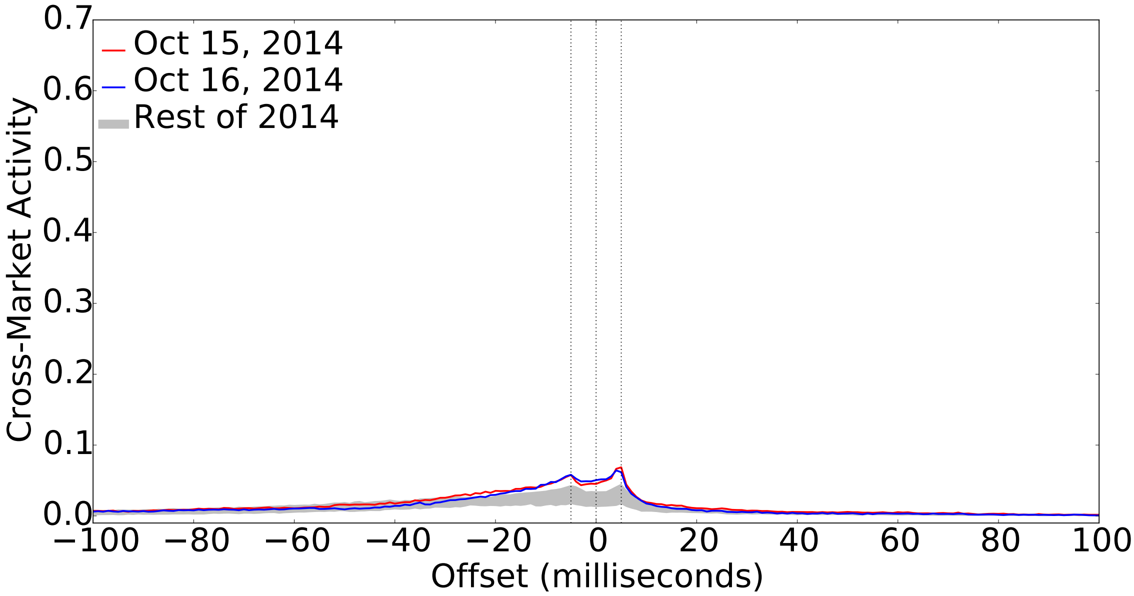

FIGURE 12: **CME Trades versus BrokerTec Trades.** Cross-market activity is expressed as a share of cash trades and reflects the incidence rate of a trade at the offer (bid) on CME coinciding with a trade at the offer (bid) on BrokerTec. A positive offset indicates that a CME trade happens prior to a BrokerTec trade. The gray area represents the 2.5th to 97.5th interpercentile range of the daily cross-market activity measures over the second half of 2014. **Data Sources:** BrokerTec (cash market data); Nanotick (CME futures market data).

Overall, the simplicity of the above empirical analysis of alternative drivers of fleeting liquidity further demonstrates the ability of our model-free measurement approach to uncover low-latency cross-market activity absent return series. In particular, this shows also that our measures of low-latency cross-market activity relying solely on recorded timestamps can be applied more generally than return-based measures.

5. Cross-Market Activity and Volatility Transmission

The close relationship between market volatility and trading activity is a long-established fact in financial markets.[15] On the one hand, a large body of empirical findings and compelling alternative theories advocate that trading volume comoves with measures of within-day price variability. On the other hand, in recent years much of the trading in U.S. Treasury and equity markets has been associated with nearly simultaneous trading between the leading cash and

[15] A survey of the early literature on the subject can be found in Karpoff (1987), while notable later examples include Schwert (1989), Ané and Geman (2000), Kyle and Obizhaeva (2016) and Andersen et al. (2015), among many others.

futures platforms. Thus, the striking cross-activity patterns we uncover in both high-frequency cross-market trading and related cross-market order book changes naturally lead to the question of how the cross-market component of overall trading activity is related to volatility.

Leaving aside the pronounced asymmetry in figure 7 versus figure 5, the spikes in cross-market activity on October 15 and 16, 2014, stand out as being well-aligned with the heightened volatility and trading observed on those days.[16] Cross-market trading and quoting activity thus appears to be related to variations in market volatility, which can create (short-lived) dislocations in relative valuations as market participants respond to news about fundamentals or market activity itself. We further demonstrate empirically that the peak number of cross-active milliseconds (across all offsets) comoves more strongly with market volatility than generic market-activity proxies such as trading volume and the number of transactions. This pattern is consistent with a positive feedback effect by which an increase in volatility can spur additional trading activity by creating cross-market trading opportunities. This observation stands in contrast to the more conventional view in the finance and economics literature which holds that trading activity predominantly causes volatility but not vice versa.

We study the evolution of the link between cross-market activity and volatility in both U.S. Treasuries (ten-year Treasury note cash and futures) and equities (S&P 500 E-mini and SPY ETF) over more than a decade from January 1, 2004 to September 30, 2015. Our analysis starts with a closer look at intraday patterns and the degree of alignment between volatility and cross-market activity over a representative sample period from July 1, 2014 to December 31, 2014 before proceeding with the full decade-long sample. We measure cross-market activity on each trading day as the number of cross-active milliseconds at its peak across all offsets and restrict attention to the most active electronic U.S. trading hours for each pair: from 7:00 to 16:00 ET for the ten-year Treasury note and from 9:30 to 16:00 ET for the S&P 500. While we carry out the analysis at millisecond frequency, it trivially generalizes to any other frequency with adequate time resolution for meaningful cross-activity measurements at different offsets.

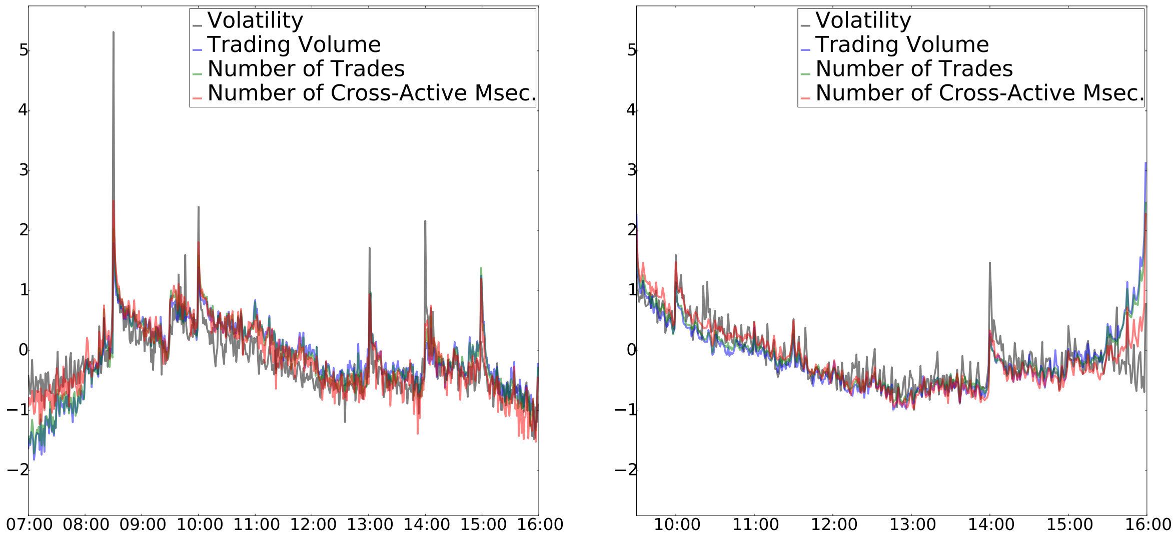

FIGURE 13: **Intraday Patterns in Different Trading Activity Measures and Volatility.** We plot the minute-average peak number of cross-active milliseconds between cash and futures markets (red line), number of trades (blue line) and trading volume (green line) against the minute-average realized volatility (gray line) in the cash market. The averages are taken over the period from July 1, 2014 to December 31, 2014 and the values are brought to the same scale by representing them as the binary logarithm of the ratio to their intraday mean. Zero indicates equality to the intraday mean; positive/negative one indicates twice/half the intraday mean. **Left Panel:** Ten-Year U.S. Treasury Note. **Right Panel:** S&P 500. **Data Sources:** BrokerTec; Nanotick; Thomson Reuters Tick History.

5.1. Intraday Patterns in Cross-Market Trading Activity and Volatility

Figure 13 shows that for both the ten-year Treasury note and S&P 500, the prevailing intraday volatility pattern is matched very closely by the diurnal pattern in cross-market activity as measured by the number of cross-active milliseconds between the cash and futures markets. The biggest volatility spikes for U.S. Treasuries (left panel of figure 13) occur at 8:30, 10:00, 13:00, and 14:00 ET around known times of news announcements, Treasury auctions, and the release of Federal Open Market Committee announcements and meeting minutes. For the S&P 500, only the spikes at 10:00 and 14:00 ET stand out to a lesser degree (right panel of figure 13). For U.S. Treasuries, there is also a notable peak around 15:00 ET corresponding to the CME market close for all pit-traded interest rate options.

Furthermore, in terms of correlation, the number of cross-active milliseconds is tracking the

[16]See "Joint Staff Report: The U.S. Treasury Market on October 15, 2014", `http://www.treasury.gov/press-center/press-releases/Documents/Joint_Staff_Report_Treasury_10-15-2015.pdf`.

intraday volatility pattern somewhat more closely than either trading volume or the number of trades. However, the tight range of most observed values within negative and positive one (excluding the extremes) suggests that interday as opposed to intraday variation may provide a better measure of the degree to which the different activity series relate to volatility.

5.2. Day-to-Day Variations in Volatility and Cross-Market Trading Activity

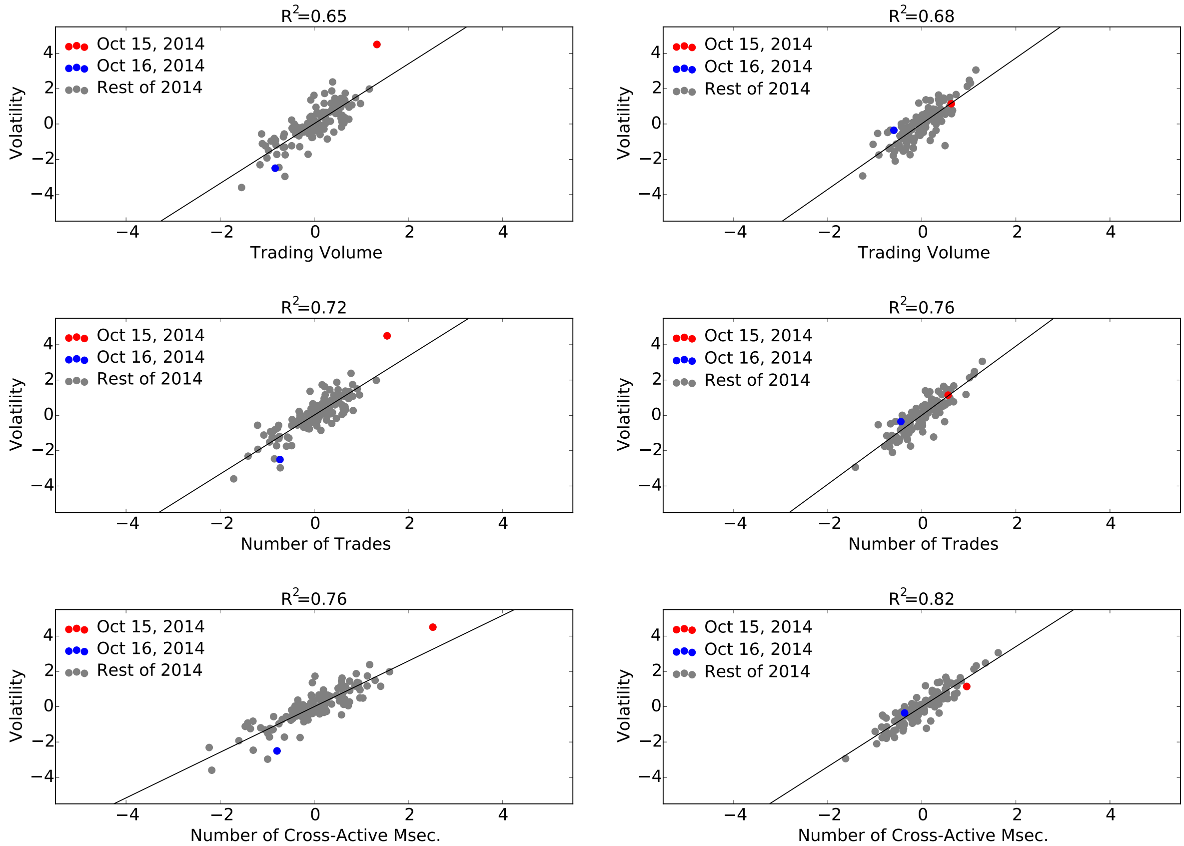

FIGURE 14: **Degree of Association between Different Trading Activity Measures and Volatility.** Daily changes in logarithmic trading volume (top row), number of trades (middle row) and peak number of cross-active milliseconds between cash and futures markets (bottom row) are plotted and regressed against daily changes in logarithmic realized volatility over the period from July 1, 2014 to December 31, 2014. The R squared measure of goodness of fit is reported above each panel. **Left Column:** Ten-Year U.S. Treasury Note. **Right Column:** S&P 500. **Data Sources:** BrokerTec; Nanotick; Thomson Reuters Tick History.

Figure 14 shows changes in daily logarithmic realized volatility plotted against changes in each daily logarithmic activity measure over the July 1 to December 31, 2014 sample period. For both the ten-year Treasury (left column) and the S&P 500 (right column) markets, the day-to-day changes in volatility appear to be more closely associated with the day-to-day changes in the peak number of cross-active milliseconds (bottom row) than with the changes in the number of trades (middle row) or trading volume (top row). In particular, the R squared for the number of cross-active milliseconds exceeds by more than 0.1 the R squared for trading volume and about half that amount the R squared for the number of trades.

Moreover, the number of cross-active milliseconds often appears to crowd out both trading volume and the number of trades if included jointly as regressors for volatility. This result is quite remarkable since it establishes that the number of cross-active milliseconds subsumes both trading volume and the number of trades in terms of information content that is available about volatility. It is also worth highlighting that while October 15 (in red) and October 16 (in blue) are known to have exhibited extreme volatility, they are not large outliers in terms of the strong linear relationship observed between changes in volatility and changes in the number of cross-active milliseconds (bottom row).

5.3. Evolution of the Relationship between Cross-Market Trading Activity and Volatility

Historical Degree of Association between Different Trading Activity Measures and Volatility

	Ten-Year US Treasury Note			**S&P 500**		
Year	**Trading Volume**	**Number of Trades**	**Number of Cross-Active Msec.**	**Trading Volume**	**Number of Trades**	**Number of Cross-Active Msec.**
2004	0.51	0.52	**0.55**	0.55	**0.65**	0.63
2005	0.64	0.66	**0.67**	0.60	**0.65**	0.61
2006	0.61	**0.65**	0.61	0.59	**0.71**	0.67
2007	0.50	0.57	**0.64**	0.63	**0.76**	0.76
2008	0.41	**0.44**	0.43	0.54	0.69	**0.69**
2009	0.45	**0.47**	0.43	0.44	0.56	**0.63**
2010	0.48	0.52	**0.62**	0.58	0.62	**0.67**
2011	0.52	0.58	**0.64**	0.51	0.68	**0.72**
2012	0.56	0.60	**0.65**	0.52	**0.60**	0.59
2013	0.64	0.72	**0.75**	0.61	0.71	**0.73**
2014	0.63	0.69	**0.75**	0.65	0.71	**0.76**
2015	0.56	0.67	**0.73**	0.57	0.63	**0.67**

TABLE 1: **Historical Degree of Association between Different Trading Activity Measures and Volatility.** The table reports yearly R squared measures of goodness of fit obtained by regressing year by year the daily changes in logarithmic realized volatility on the daily changes in logarithmic trading volume, number of trades, and peak number of cross-active milliseconds over the period from Jan 1, 2004 to Sep 30, 2015. The numbers in bold represent the largest values obtained separately for each year for the Ten-Year U.S. Treasury Note and S&P 500. **Data Sources:** BrokerTec; Nanotick; Thomson Reuters Tick History.

To better assess the extent to which market volatility has become more closely associated with high-frequency cross-market trading activity, table 1 reports historical R squared measures of goodness of fit for each year from Jan 1, 2004 to Sep 30, 2015 obtained by regressing year by year the daily changes in logarithmic realized volatility on the daily changes in logarithmic trading volume, number of trades, or number of cross-active milliseconds for the 10-Year US Treasury (left part of the table) and S&P 500 (right part of the table). In particular, measuring the peak number of cross-active milliseconds day by day and taking logarithmic differences of the daily measures during each twelve-month period limits the impact of secular trends in latency and trading practices over the past decade resulting from technological improvements and the related evolution in high-frequency trading. Table 1 thus strongly indicates that with the rise in high-frequency trading in recent years, cross-market activity as measured by the peak number of cross-active milliseconds between the cash and futures markets has typically been more tightly linked to volatility than standard activity measures such as overall trading volume or the number of trades in either the cash or futures markets.

Even more strikingly, the proposed model-free measure of cross-market activity appears to contain extra information about volatility beyond what is implied by naively combining the trading volumes and counts for both the cash and futures markets. Table 2 shows the high statistical significance of the loading on the peak number of cross-active milliseconds in yearly OLS regressions including also the trading volumes and number of trades for both the cash and futures markets. The regression specification takes the form:

$$\Delta rv_t = const + \beta[n_1] \cdot \Delta n_{1,t} + \beta[n_2] \cdot \Delta n_{2,t} + \beta[v_1] \cdot \Delta v_{1,t} + \beta[v_2] \cdot \Delta v_{2,t} + \beta[x_{12}] \cdot \Delta x_{12,t} + \epsilon_t \ ,$$

where daily changes in logarithmic realized volatility (rv_t) are regressed on daily changes in the logarithmic number of trades and trading volume in both the futures market ($n_{1,t}$ and $v_{1,t}$) and the cash market ($n_{2,t}$ and $v_{2,t}$) as well as daily changes in the logarithmic measure of cross-market activity between the two markets ($x_{12,t} = \log(\max_s \mathcal{X}^{\text{raw}}_{s,t})$) given by the peak number of cross-active milliseconds across all different offsets. In addition to the high statistical significance of the model-free measure of cross-market activity, the resulting improvements in R squared are most visible in the years when high-frequency trading has become prevalent for each market pair

Regression of Volatility on Cash and Futures Market Trading Activity Measures

	OLS Coefficients						Student's t-statistics						Adjusted R^2			
Year	**const**	**n_1**	**n_2**	**v_1**	**v_2**	**x_{12}**	**const**	**n_1**	**n_2**	**v_1**	**v_2**	**x_{12}**	**Without x_{12}**	**With x_{12}**	**Absolute Increase**	**Relative Increase**
Panel A: Ten-Year U.S. Treasury Note																
2004	0.0	3.1	-1.0	-0.6	0.5	-0.3	0.8	**8.2**	-1.4	-1.4	0.9	-1.4	0.74	0.74	0.00	0.01
2005	0.0	1.6	0.3	-0.5	-0.4	0.6	0.1	**4.2**	0.8	-1.8	-1.5	**5.9**	0.74	0.75	0.01	0.05
2006	0.0	1.5	1.0	-0.6	-0.5	0.1	0.4	**4.8**	**2.5**	**-2.5**	-1.5	0.8	0.74	0.74	0.00	0.01
2007	0.0	1.8	0.8	-1.0	-0.8	0.7	-0.2	**6.4**	1.8	**-3.2**	**-2.6**	**3.4**	0.76	0.78	0.02	0.10
2008	0.0	1.6	0.0	-0.2	-0.1	0.1	0.0	**5.0**	-0.1	-0.9	-0.2	1.4	0.68	0.68	0.00	0.01
2009	0.0	2.9	-0.4	-1.3	0.4	0.0	-0.3	**5.3**	-0.3	**-4.2**	0.4	-0.3	0.60	0.60	0.00	0.00
2010	0.0	2.4	-0.3	-0.6	-0.3	0.3	-0.7	**7.8**	-0.6	**-2.5**	-0.6	**2.1**	0.74	0.74	0.01	0.03
2011	0.0	1.0	0.5	0.1	-0.9	0.7	-0.8	1.5	1.0	0.1	**-2.2**	**2.9**	0.69	0.70	0.02	0.06
2012	0.0	0.5	0.1	-0.1	0.0	0.8	-0.9	1.9	0.2	-0.2	0.1	**4.0**	0.67	0.70	0.03	0.10
2013	0.0	0.8	1.4	-0.7	-0.9	0.8	0.3	1.8	**3.3**	-1.8	-1.9	**4.1**	0.77	0.78	0.01	0.06
2014	0.0	1.1	0.7	-0.1	-1.2	0.8	-0.4	**4.3**	**2.3**	-0.6	**-4.0**	**5.4**	0.79	0.81	0.01	0.07
2015	0.0	0.5	1.6	0.1	-1.6	0.6	0.5	1.3	**3.4**	0.3	**-5.2**	**3.8**	0.78	0.79	0.01	0.06
Panel B: S&P 500																
2004	0.0	1.1	0.5	-0.7	0.0	0.3	-0.7	**2.9**	**3.3**	**-3.0**	-0.1	**2.1**	0.65	0.66	0.01	0.02
2005	0.0	0.7	0.0	-0.2	0.1	0.4	-0.1	**3.1**	0.3	-1.8	**2.0**	**2.2**	0.65	0.66	0.01	0.03
2006	0.0	1.6	0.2	-0.7	0.1	0.1	-0.2	**5.2**	1.6	**-3.8**	0.5	0.7	0.75	0.75	0.00	0.00
2007	0.0	1.8	-0.6	-1.2	0.2	1.1	-1.0	**4.8**	**-3.0**	**-9.1**	1.5	**3.8**	0.80	0.82	0.02	0.11
2008	0.0	2.7	0.4	-1.9	0.0	0.3	-0.2	**5.3**	**2.0**	**-4.9**	-0.2	**2.2**	0.81	0.81	0.00	0.03
2009	0.0	1.3	0.4	-1.3	0.0	0.8	-1.7	**3.1**	1.8	**-6.4**	-0.1	**2.0**	0.68	0.72	0.04	0.13
2010	0.0	0.1	0.7	-0.6	-0.1	1.0	-0.4	0.3	**2.2**	**-2.8**	-0.4	**2.1**	0.71	0.74	0.03	0.10
2011	0.0	1.0	0.6	-1.3	-0.6	1.3	-0.7	**3.6**	**2.1**	**-6.3**	**-2.4**	**4.6**	0.74	0.77	0.03	0.12
2012	0.0	1.6	0.1	-1.3	-0.1	0.7	0.7	**3.2**	0.3	**-3.6**	-0.4	**2.2**	0.64	0.65	0.01	0.04
2013	0.0	1.2	0.3	-1.2	-0.3	1.2	-0.8	**2.9**	0.9	**-3.9**	-1.7	**4.4**	0.72	0.76	0.03	0.12
2014	0.0	1.0	-0.1	-1.1	-0.2	1.4	-0.2	**2.2**	-0.2	**-3.1**	-1.5	**4.4**	0.76	0.81	0.05	0.20
2015	0.0	0.1	0.7	-0.1	-1.1	1.6	0.1	0.4	1.6	-0.4	**-3.8**	**4.9**	0.69	0.75	0.05	0.17

TABLE 2: **Regression of Volatility on Cash and Futures Markets Trading Activity Measures.** The table reports results for yearly OLS regressions $\Delta rv_t = const + \beta[n_1] \cdot \Delta n_{1,t} + \beta[n_2] \cdot \Delta n_{2,t} + \beta[v_1] \cdot \Delta v_{1,t} + \beta[v_2] \cdot \Delta v_{2,t} + \beta[x_{12}] \cdot \Delta x_{12,t} + \epsilon_t$ of daily changes in logarithmic realized volatility (rv_t) on daily changes in logarithmic number of trades and trading volume in both the futures market ($n_{1,t}$ and $v_{1,t}$) and the cash market ($n_{2,t}$ and $v_{2,t}$) as well as daily changes in the model-free logarithmic measure of cross-market activity between the two markets ($x_{12,t}$) given by the peak number of cross-active milliseconds across all different offsets. Student's t-statistics in bold are significant at 0.05 level after Newey-West adjustment with 22 lags. The relative increase in adjusted R squared is computed by regressing volatility residuals on cross-market activity residuals after partialling out the number of trades and trading volume for each market from both volatility and cross-market activity. The sample period is from Jan 1, 2004 to Sep 30, 2015 and gray boxes indicate when high-frequency cross-market trading is known to have become prevalent in each of the considered market pairs. **Panel A:** Ten-Year U.S. Treasury Note. **Panel B:** S&P 500. **Data Sources:** BrokerTec; Nanotick; Thomson Reuters Tick History.

(starting from 2011 for the ten-year US Treasury Note and from 2009 for S&P 500). We further note that while other studies such as Andersen et al. (2015), Ané and Geman (2000), Clark (1973), or Kyle and Obizhaeva (2016) aim to rectify a particular theory-implied form of the relationship between volatility and trading activity in a given single market, our main goal is to show that the cross-market component of high-frequency trading activity between closely related markets contains extra information about market volatility unspanned by trading volume and the number of trades in each market.

To address possible concerns stemming from the high degree of multi-collinearity between the trading activity measures, we further consider an alternative specification that brings all cross-correlations below 0.5 by way of suitable transformations. Table 3 shows the OLS regression results for one such rearrangement in which both the cross-market activity and the number of trades in the futures market are taken relative to the number of trades in the cash market, while the trading volume in each market is transformed to average trade size (daily trading volume taken relative to trade count). By design this yields the same R squared measures of goodness of fit while better revealing the statistical significance of cross-market activity in comparison to the rest of the activity variables. In particular, it becomes clear that while most of the explanatory power stems from the chosen proxy for the baseline trading activity level (the number of trades in the cash markets $n_{2,t}$), the proposed measure of cross-market activity relative to the baseline activity level ($x_{12,t} - n_{2,t}$) has consistently been highly statistically significant, especially after high-frequency cross-market trading activity is known to have grown up in the considered cash and futures markets.

The resulting increase in terms of the adjusted R squared measure of goodness of fit is more evident when regressing volatility residuals on cross-market activity residuals after partialling out the number of trades and trading volume for each market from both volatility and cross-market activity (denoted as relative increase in adjusted R squared in the last column of tables 2 and 3). Thus, for the ten-year U.S. Treasury Note cash and futures markets at least 5% of the variability in volatility unexplained by the number of trades and trading volume in both cash and futures markets appears to be consistently associated with cross-market activity from 2011 onwards when high-frequency cross-market trading is known to have become prevalent in these markets. Even more strikingly, with the rise in high-frequency cross-market trading between S&P 500 cash and futures markets from 2009 onwards, the proposed measure of cross-market activity consistently can explain as much as 10% of the variability in volatility that remains

Regression of Volatility on Transformed Cash and Futures Market Trading Activity Measures

	OLS Coefficients						Student's t-statistics						Adjusted R^2			
Year	const	n_1-n_2	n_2	v_1-n_1	v_2-n_2	x_{12}-n_2	const	n_1-n_2	n_2	v_1-n_1	v_2-n_2	x_{12}-n_2	Without x_{12}-n_2	With x_{12}-n_2	Absolute Increase	Relative Increase
Panel A: Ten-Year U.S. Treasury Note																
2004	0.0	2.5	1.7	-0.6	0.5	-0.3	0.8	**9.5**	**9.3**	-1.4	0.9	-1.4	0.74	0.74	0.00	0.01
2005	0.0	1.2	1.6	-0.5	-0.4	0.6	0.1	**3.0**	**23.0**	-1.8	-1.5	**5.9**	0.74	0.75	0.01	0.05
2006	0.0	0.9	1.4	-0.6	-0.5	0.1	0.4	**3.9**	**14.5**	**-2.5**	-1.5	0.8	0.74	0.74	0.00	0.01
2007	0.0	0.8	1.4	-1.0	-0.8	0.7	-0.2	**4.2**	**14.4**	**-3.2**	**-2.6**	**3.4**	0.76	0.78	0.02	0.10
2008	0.0	1.3	1.4	-0.2	-0.1	0.1	0.0	**7.3**	**27.1**	-0.9	-0.2	1.4	0.68	0.68	0.00	0.01
2009	0.0	1.6	1.6	-1.3	0.4	0.0	-0.3	**5.0**	**12.6**	**-4.2**	0.4	-0.3	0.60	0.60	0.00	0.00
2010	0.0	1.8	1.5	-0.6	-0.3	0.3	-0.7	**12.2**	**18.8**	**-2.5**	-0.6	**2.1**	0.74	0.74	0.01	0.03
2011	0.0	1.1	1.4	0.1	-0.9	0.7	-0.8	**4.6**	**17.3**	0.1	**-2.2**	**2.9**	0.69	0.70	0.02	0.06
2012	0.0	0.4	1.3	-0.1	0.0	0.8	-0.9	1.9	**19.4**	-0.2	0.1	**4.0**	0.67	0.70	0.03	0.10
2013	0.0	0.2	1.4	-0.7	-0.9	0.8	0.3	0.5	**14.2**	-1.8	-1.9	**4.1**	0.77	0.78	0.01	0.06
2014	0.0	1.0	1.3	-0.1	-1.2	0.8	-0.4	**4.9**	**20.1**	-0.6	**-4.0**	**5.4**	0.79	0.81	0.01	0.07
2015	0.0	0.6	1.2	0.1	-1.6	0.6	0.5	**2.2**	**10.2**	0.3	**-5.2**	**3.8**	0.78	0.79	0.01	0.06
Panel B: S&P 500																
2004	0.0	0.4	1.3	-0.7	0.0	0.3	-0.7	**2.0**	**11.2**	**-3.0**	-0.1	**2.1**	0.65	0.66	0.01	0.02
2005	0.0	0.5	1.1	-0.2	0.1	0.4	-0.1	**2.5**	**21.7**	-1.8	**2.0**	**2.2**	0.65	0.66	0.01	0.03
2006	0.0	0.9	1.3	-0.7	0.1	0.1	-0.2	**5.1**	**18.6**	**-3.8**	0.5	0.7	0.75	0.75	0.00	0.00
2007	0.0	0.5	1.3	-1.2	0.2	1.1	-1.0	1.8	**14.5**	**-9.1**	1.5	**3.8**	0.80	0.82	0.02	0.11
2008	0.0	0.8	1.5	-1.9	0.0	0.3	-0.2	**3.4**	**14.3**	**-4.9**	-0.2	**2.2**	0.81	0.81	0.00	0.03
2009	0.0	0.0	1.2	-1.3	0.0	0.8	-1.7	0.1	**9.0**	**-6.4**	-0.1	**2.0**	0.68	0.72	0.04	0.13
2010	0.0	-0.5	1.1	-0.6	-0.1	1.0	-0.4	-1.6	**9.4**	**-2.8**	-0.4	**2.1**	0.71	0.74	0.03	0.10
2011	0.0	-0.2	1.1	-1.3	-0.6	1.3	-0.7	-0.9	**11.0**	**-6.3**	**-2.4**	**4.6**	0.74	0.77	0.03	0.12
2012	0.0	0.3	1.0	-1.3	-0.1	0.7	0.7	0.8	**7.3**	**-3.6**	-0.4	**2.2**	0.64	0.65	0.01	0.04
2013	0.0	0.0	1.1	-1.2	-0.3	1.2	-0.8	-0.2	**13.9**	**-3.9**	-1.7	**4.4**	0.72	0.76	0.03	0.12
2014	0.0	-0.2	1.0	-1.1	-0.2	1.4	-0.2	-0.4	**11.3**	**-3.1**	-1.5	**4.4**	0.76	0.81	0.05	0.20
2015	0.0	0.0	1.2	-0.1	-1.1	1.6	0.1	-0.1	**10.9**	-0.4	**-3.8**	**4.9**	0.69	0.75	0.05	0.17

TABLE 3: **Regression of Volatility on Transformed Cash and Futures Markets Trading Activity Measures.** The table reports results for yearly OLS regressions $\Delta rv_t = const + \beta[n_1] \cdot \Delta(n_{1,t} - n_{2,t}) + \beta[n_2] \cdot \Delta n_{2,t} + \beta[v_1] \cdot \Delta(v_{1,t} - n_{1,t}) + \beta[v_2] \cdot \Delta(v_{2,t} - n_{2,t}) + \beta[x_{12}] \cdot \Delta(x_{12,t} - n_{2,t}) + \epsilon_t$ of daily changes in logarithmic realized volatility (rv_t) on daily changes in the logarithmic number of trades in futures markets relative to cash markets ($n_{1,t} - n_{2,t}$), logarithmic number of trades in cash markets ($n_{2,t}$), logarithmic average trade size in futures markets ($v_{1,t} - n_{1,t}$) and cash markets ($v_{2,t} - n_{2,t}$), as well as daily changes in the model-free logarithmic measure of cross-market activity between the two markets given by the peak number of cross-active milliseconds across all different offsets relative to the number of trades in the cash market ($x_{12,t} - n_{2,t}$). Student's t-statistics in bold are significant at 0.05 level after Newey-West adjustment with 22 lags. The relative increase in adjusted R squared is computed by regressing volatility residuals on cross-market activity residuals after partialling out the number of trades and trading volume for each market from both volatility and cross-market activity. The sample period is from Jan 1, 2004 to Sep 30, 2015 and gray boxes indicate when high-frequency cross-market trading is known to have become prevalent in each of the considered market pairs. **Panel A:** Ten-Year U.S. Treasury Note. **Panel B:** S&P 500. **Data Sources:** BrokerTec; Nanotick; Thomson Reuters Tick History.

unexplained by the number of trades and trading volume in both markets.[17]

[17] As an additional robustness check available upon request we further document that the cross-market compo-

5.4. High-Speed Cross-Market Trading Activity during Flash Events

Notable flash events such as the May 2010 U.S. equity market flash crash and the October 2014 U.S. Treasury market flash rally further underscore the importance of better understanding any potential links between the intensity of high-speed cross-trading activity and observed differences in market functioning under stress. It is therefore useful to revisit and contrast such events through the lens of our measure of high-speed cross-market activity. Its ability to reliably quantify the intensity of high-speed cross-trading over relatively narrow intraday event windows is particularly valuable for identifying any potentially short-lived changes in cross-market activity under stress vis-à-vis historical baselines.

To this end, Figure 15 compares the equity market flash crash on May 6, 2010 (left panel) with the Treasury market flash rally on October 15, 2014 (right panel) in terms of observed cash-futures price synchronicity and measured intensity of high-speed cross-trading activity. One notable distinction between the two flash events is that during the equity flash crash depicted in the left panel, price dynamics for the futures market and those for the cash markets lost their no-arbitrage price link and markedly diverged from each other. In contrast, during the Treasury flash rally, shown in the right panel, futures and cash prices remained remarkably tightly linked.

Each panel in Figure 15 further shows the intraday evolution of the locally measured intensity of high-speed cross-trading activity (blue line) and its respective 95% confidence bounds (blue range) computed from pre-crash data. This comparison reveals that high-speed cross-market activity was abnormally low during the May 6, 2010 equity market flash crash when cash and futures markets lost their no-arbitrage price link. By contrast, cross-market activity was abnormally high during the October 15, 2014 Treasury market flash rally when cash and futures markets remained tightly linked, even under stress. This is consistent with recent findings by Menkveld and Yueshen (2019) that a breakdown of cross-market arbitrage activity renders markets more fragile and further suggests a potentially beneficial role played by high-speed cross-trading activity for maintaining the no-arbitrage price link between closely related markets also during periods of significant market stress.

nent of high-frequency trading activity remains statistically significant in the above regression also after expanding the regression specification with often used market liquidity measures.

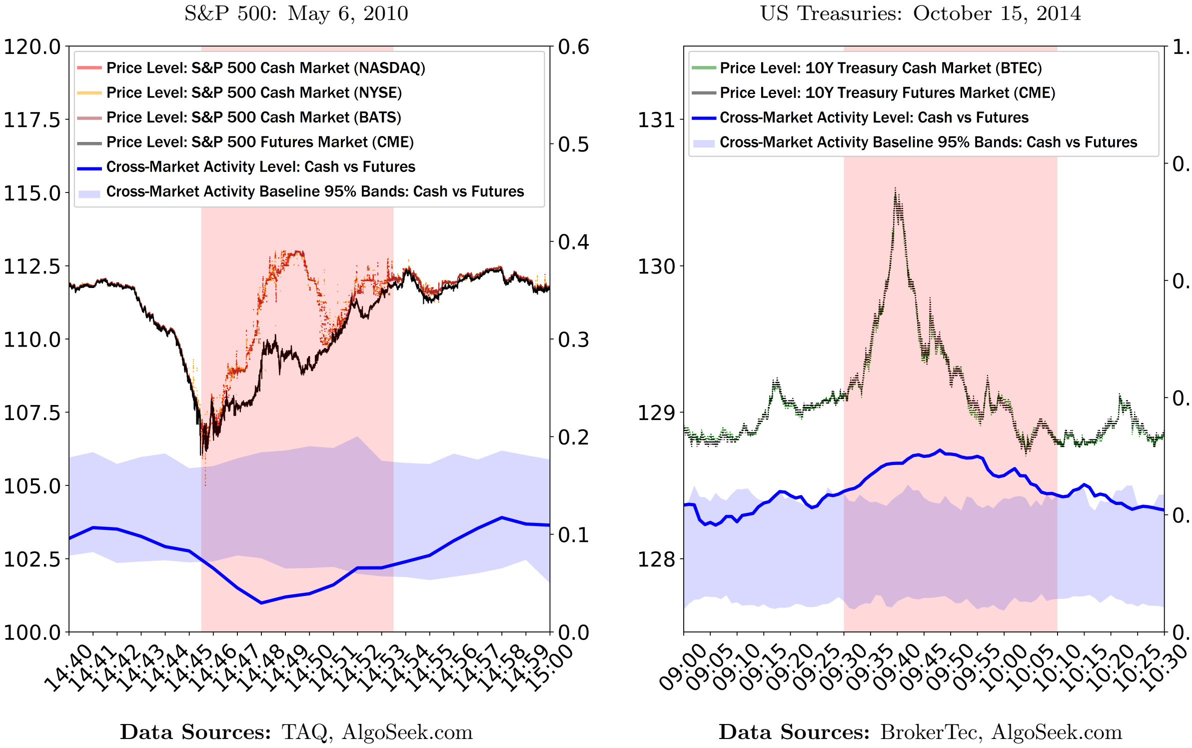

FIGURE 15: **High-Speed Cross-Market Trading Activity and Price Synchronicity During Flash Events.** We plot flash event price dynamics for the futures market (the dark gray line) and those for cash markets (the red, yellow, brown, and green lines) vis-à-vis locally measured intensity of high-speed cross-trading activity (blue line) and its respective 95% confidence bounds (blue range) computed from pre-event data. **Left Panel:** S&P 500 Flash Crash on May 6, 2010. **Right Panel:** U.S. Treasuries Flash Rally on October 15, 2014.

On the flip side, modern electronic markets may have become more heavily dependent on the sufficiently active presence of high-speed cross-market traders and liquidity providers. The incremental explanatory power of cross-market activity for market volatility beyond single-market activity measures documented in the previous subsection can be interpreted as supporting the intuition that high-speed cross-market arbitrageurs tend to become more active during more volatile market periods when relative mispricings are naturally more likely to occur. As such, there appears to be rationale for a potential amplification mechanism by which a spike in volatility during flash events naturally leads to elevated high-speed cross-trading activity. This might put further upward pressure on volatility as a result of elevating the overall level of trading activity.

Nonetheless, as seen from the comparison between the two flash events in Figure 15, abnormally low levels of high-speed cross-trading activity may be associated with far more severe disruptions in the functioning of modern interconnected markets. This is evidenced by the loss of the futures-cash arbitrage price link during the 2010 U.S. equity market flash crash. Any potential vulnerabilities arising from abnormally low cross-market activity may have further implications also for assessing the impact of uncoordinated trading halts across multiple related trading venues. In particular, a trading halt by a single trading venue would induce a breakdown in cross-market trading activity with related markets where trading continues. For example, during the October 2016 sterling flash crash a breakdown in high-speed cross-market trading activity occurred due to the trading halt in futures markets. This has been cited as a potential contributing factor to the elevated price disruptions in the sterling cash market. In this regard, more generally, ability to measure variations in low-latency cross-market trading activity can help better distinguish between different episodes of market turbulence and the degree of systemic importance of high-speed liquidity provision across different markets.

6. Conclusion

We propose a class of intuitive model-free measures of high-speed cross-market interlinkages in modern electronic markets capturing the prevailing lead-lag at which low-latency cross-market trading activity peaks, as well as its magnitude and dispersion. The measures avoid reliance on noisy high-frequency return series and directly employ time stamp information for sharp identification of the prevailing lead-lag relationships in cross-market trading activity based on publicly available trade and quote data with sufficient time stamp granularity. The measures can thus be used to help uncover and quantify low-latency cross-market interlinkages such as so called "fleeting liquidity" in interlinked electronic markets. As an empirical illustration, we employ the measures to contrast U.S. Treasury and equity markets in terms of the underlying low-latency lead-lag relationship between futures and cash markets and provide new evidence pointing to prudent market making as the primary cause of liquidity mirages in U.S. Treasury markets. We further show that low-latency cross-market activity between cash and futures markets tends to be positively associated with market volatility even after controlling for trading activity in each market. We argue that a positive relationship between high-speed cross-market activity and market volatility is a sign of proper market functioning as we find a clear distinction between the 2010 U.S. equity market flash crash and the 2014 U.S. Treasury market flash rally in terms of observed intensity of low-latency cross-market activity and tightness of the no-arbitrage price link between cash and futures markets.

The simplicity of our cross-market activity measures makes them particularly well suited for large scale market wide analysis, and monitoring of market functioning. Knowing the prevalence and variations in low-latency market inter-linkages sheds new light on the topology of cross-market liquidity provision and volatility transmission. Overall, our findings show that accounting for high-speed cross-market trading activity opens up new perspectives and testable implications for intrinsic times in finance such as endogeneity between volatility and cross-market activity and the cross-market interdependence in observed activity times. Thus, we see further scope for improving the available modelling toolkit in these areas.

References

Andersen, T.G., 1996. Return volatility and trading volume: An information flow interpretation of stochastic volatility. The Journal of Finance 51, 169–204.

Andersen, T.G., Bondarenko, O., Kyle, A.S., Obizhaeva, A.A., 2015. Intraday trading invariance in the E-mini S&P 500 futures market. Working Paper .

Ané, T., Geman, H., 2000. Order flow, transaction clock, and normality of asset returns. The Journal of Finance 55, 2259–2284.

Benos, E., Brugler, J.A., Hjalmarsson, E., Zikes, F., 2015. Interactions among high-frequency traders. Bank of England Working Paper No. 523 .

Biais, B., Foucault, T., Moinas, S., 2015. Equilibrium fast trading. Journal of Financial Economics 116, 292 – 313.

Bollerslev, T., Jubinski, D., 1999. Equity trading volume and volatility: Latent information arrivals and common long-run dependencies. Journal of Business and Economic Statistics 17, 9–21.

Brogaard, J., Hendershott, T., Riordan, R., 2014. High-frequency trading and price discovery. Review of Financial Studies .

Buccheri, G., Corsi, F., Peluso, S., 2018. High-frequency lead-lag effects and cross-asset linkages: a multi-asset lagged adjustment model. Working Paper .

Budish, E., Cramton, P., Shim, J., 2015. The high-frequency trading arms race: Frequent batch auctions as a market design response. The Quarterly Journal of Economics .

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.

Clark, P.K., 1973. A subordinated stochastic process model with finite variance for speculative prices. Econometrica 41, 135–155.

Dutilleul, P., 2011. Spatio-Temporal Heterogeneity: Concepts and Analyses. Cambridge University Press.

Epps, T.W., Epps, M.L., 1976. The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distributions hypothesis. Econometrica 44, 305–321.

Godfrey, K., 2014. Detecting pairs trading (or any related trading) in two financial securities. Working Paper, University of Western Australia. .

Harris, L., 1987. Transaction data tests of the mixture of distributions hypothesis. The Journal of Financial and Quantitative Analysis 22, 127–141.

Hasbrouck, J., Saar, G., 2013. Low-latency trading. Journal of Financial Markets 16, 646–679.

Hayashi, T., Koike, Y., 2018. Multi-scale analysis of lead-lag relationships in high-frequency financial markets. Working Paper .

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

Haynes, R., Roberts, J., 2015. Automated trading in futures markets. CFTC White Paper .

Hendershott, T., Riordan, R., 2013. Algorithmic trading and the market for liquidity. Journal of Financial and Quantitative Analysis 48, 1001–1024.

Holden, C.W., Jacobsen, S., 2014. Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. The Journal of Finance 69, 1747–1785.

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

Joint Staff Report, 2015. The U.S. Treasury Market on October 15, 2014. U.S. Department of the Treasury, Board of Governors of the Federal Reserve System, Federal Reserve Bank of New York, U.S. Securities and Exchange Commission, U.S. Commodity Futures Trading Commission.

Jones, C.M., Kaul, G., Lipson, M.L., 1994. Transactions, volume, and volatility. Review of Financial Studies 7, 631–651.

Karpoff, J.M., 1987. The relation between price changes and trading volume: A survey. The Journal of Financial and Quantitative Analysis 22, 109–126.

Kyle, A.S., Obizhaeva, A.A., 2016. Market microstructure invariance: Empirical hypotheses. Econometrica 84, 1345–1404.

Laughlin, G., Aguirre, A., Grundfest, J., 2014. Information transmission between financial markets in Chicago and New York. Financial Review 49, 283–312.

Lee, C.M.C., Ready, M.J., 1991. Inferring trade direction from intraday data. The Journal of Finance 46, 733–746.

Lewis, M., 2014. Flash Boys: A Wall Street Revolt. W. W. Norton.

Menkveld, A.J., 2008. Splitting orders in overlapping markets: A study of cross-listed stocks. Journal of Financial Intermediation 17, 145 – 174.

Menkveld, A.J., Yueshen, B.Z., 2019. The flash crash: A cautionary tale about highly fragmented markets. Management Science 65, 4470–4488.

O'Hara, M., 2015. High frequency market microstructure. Journal of Financial Economics 116, 257 – 270.

Schwert, G.W., 1989. Why does stock market volatility change over time? The Journal of Finance 44, 1115–1153.

Tauchen, G.E., Pitts, M., 1983. The price variability-volume relationship on speculative markets. Econometrica 51, 485–505.

Yao, C., Ye, M., 2018. Why Trading Speed Matters: A Tale of Queue Rationing under Price Controls. The Review of Financial Studies 31, 2157–2183.

Ying, C.C., 1966. Stock market prices and volumes of sales. Econometrica 34, 676–685.

Appendix A. Description of High Frequency Data Sources and Timestamp Precision

We obtain high frequency market data for both U.S. Treasuries (ten-year Treasury note and futures) and equities (S&P 500 E-mini futures and SPDR SPY ETF) with millisecond or higher precision of the timestamps over the period 2004-2017. In particular, for the ten-year Treasury cash market we rely on Brokertec data with platform-provided microsecond precision of the timestamps. For the S&P 500 cash market represented by the SPDR ETF we use data available through Thomson Reuters Tick History (TRTH) and the Wharton Research Data Services (WRDS) NYSE TAQ database, both of which contain exchange-provided timestamps recorded with millisecond precision. Finally, for the CME Treasury and S&P 500 futures market we rely on high-frequency market data captured on-site by AlgoSeek.com with millisecond resolution of the timestamps over the period 2010-2017 as well as data by Nanotick with higher sub microsecond resolution of the timestamps over the six-month sample period from July 1, 2014 to December 31, 2014. As such, the timestamp precision of these data sets is more than adequate enough for constructing reliable cross-market activity measures at millisecond frequency.

We further note that while measuring the prevailing offset and dispersion of cross-market activity can get affected by timestamp distortions due to extra latency incurred at data capture, measuring relative changes in the magnitude of cross-market activity tends to be a lot more robust across alternative data sources. This inherent robustness greatly expands the ability to extract meaningful information from the proposed cross-market activity measures across a wider range of data sources and time periods. In particular, it allows us to study changes in the magnitude of cross-market activity over the past decade from January 1, 2004 to September 30, 2015 by using the longer history of futures market data available through TRTH despite the somewhat coarser than millisecond accuracy of most TRTH-provided timestamps in view of the extra time lag incurred due to centralized off-site data processing in London.