Intraday Trading Invariance in the E-mini S&P 500 Futures Market

Torben G. Andersen, Oleg Bondarenko,
Albert S. Kyle and Anna A. Obizhaeva*

First Draft: July 2, 2014

This Draft: March 26, 2018

The trading activity in the E-mini S&P 500 futures contract between January 2008 and September 2011 is consistent with the following high-frequency invariance relationship: The return variation per transaction is log-linearly related to trade size, with a slope coefficient of -2. *This association applies both in the time series and across a pronounced intraday pattern. The documented factor of proportionality deviates sharply from prior hypotheses relating volatility to trading intensity. High-frequency trading invariance is motivated* a priori *by the intuition that market microstructure invariance, introduced by Kyle and Obizhaeva (2016*a*) to explain bets at low frequencies, also applies to transactions over short intraday intervals. It raises the prospect of identifying periods of market stress in real time and poses intriguing challenges for market microstructure research.*

Keywords: market microstructure, invariance, high-frequency trading, liquidity, volatility, volume, time series, intraday patterns.

* Andersen: Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, USA, NBER, and CREATES, t-andersen@kellogg.northwestern.edu. Bondarenko: Department of Finance (MC 168), University of Illinois at Chicago, 601 S. Morgan St., Chicago, IL 60607, USA, olegb@uic.edu. Kyle: Robert H. Smith School of Business, University of Maryland, College Park, MD 20742, USA, akyle@rhsmith.umd.edu. Obizhaeva: New Economic School, 100A Novaya Street, Skolkovo, Moscow, 143026, Russia, aobizhaeva@nes.ru. We are grateful to the CME Group for access to data from the CME DataMine system. Andersen acknowledges support from CREATES, Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation. Kyle works as a director for an asset management company which trades international stocks and has also worked as a consultant for various U.S. government agencies, including the SEC and CFTC. We appreciate comments from participants at various conferences and seminars.

I. Introduction

An extensive literature has documented persistent time series variation and strong correlation among daily market activity variables such as trading volume, transaction intensity, and return volatility. These dependencies remain substantial across weekly and monthly horizons. Numerous other studies demonstrate that these trading variables follow a pronounced periodic pattern across the trading day; there is a distinct intraday rhythm to market activity. Consequently, the features that dominate the daily series and intraday pattern operate at different time scales and display different stochastic properties. They have been explored in two largely separate strands of the literature, both empirically and theoretically.

The time series branch of the literature has its origins among the very earliest studies of daily return data. Motivated by initial contributions from Brada, Ernst and Van Tassel (1967), Mandelbrot and Taylor (1967) and Clark (1973), this literature derives a tight quantitative link among the market activity variables, couched within what we now label the mixture-of-distributions-hypothesis (MDH) framework. It stipulates that return volatility is related to concurrent trading activity either because the latter directly governs the former or because both are linked to an underlying economic quantity, such as news arrivals. Serial correlation in news arrivals may induce the persistent covariation we observe. Empirically, trading intensity is captured either through the transaction count or trading volume. There is an active debate regarding the relative merits of transaction count versus trade size in the MDH context, with both widely used in existing empirical studies. Of course, if trade size (volume per trade) is invariant or constant, these measures are proportional and possess identical explanatory power for return volatility. As we also document below, however, trade size varies systematically over time.

The intraday-oriented branch of the literature stresses the pronounced variation and correlation in trading intensity and volatility within the daily market cycle. For example, trading activity is often high at the open, moderate in the middle of the trading day, and then higher again towards the market close, generating a U-shaped pattern; see Wood, McInish and Ord (1985) for an early exploration. These findings inspired theories that rationalize the clustering of trading activity and return volatility in certain segments of the trading day, such as Admati and Pfleiderer (1988) and Hong and Wang (2000). Since such features are operative at a qualitative level across all major asset classes and market structures, they appear universal and likely arise endogenously from the interaction of trading strategies, the daily rhythm of news releases, and business activity. Despite their ubiquitous nature, there is little theory that develops predictions regarding the quantitative link among these series across the intraday cycle.

In this paper, we explore whether an MDH relationship applies simultaneously in the time series and across the intraday pattern. This is motivated by the fact that, in theory, the MDH relationship should remain valid for smaller discrete intraday intervals, provided the underlying market remains liquid. This suggests that we may undertake more powerful tests based on transaction level data. Moreover, while exploration of the distributional implications of the MDH for high-frequency data has a lengthy history (see Harris (1987) for an early study) there are no prior attempts at confronting the MDH with quantitative predictions regarding the nature of the interaction among the activity variables across the intraday pattern. As such, our tests are informative regarding the broader applicability of the theory and the reasons behind its shortcomings.

We find the qualitative correlation structure of return volatility and trading activity to be consistent with prior evidence. Nevertheless, our detailed tests of the MDH relationships imply extremely strong rejections. These systemic violations point towards a critical role for trade size in governing the joint distribution of the market activity variables. This inspires our thorough investigation of an alternative intraday trading invariance hypothesis, motivated by—yet not directly implied by—the invariance principles developed in Kyle and Obizhaeva (2016*a*).

The Intraday Trading Invariance (ITI) hypothesis stipulates that traders consider the exposure of the position they enter when submitting orders, so the transaction size tends to shrink, all else equal, in more volatile environments. As a result, return volatility will not be associated with a single trading intensity variable such as the number of transactions or volume, but rather a combination of such variables. Our exploration corroborates the main implications of the ITI hypothesis both for time series data and for the interdependencies across the intraday pattern.

To illustrate the diverse nature of the evidence we explore, Figures 1 and 2 depict a set of activity variables for the E-mini S&P 500 futures contract from January 2008 through September 2011. The figures are constructed from one-minute observations on trading volume V, return volatility σ, transaction count N, and trade size Q, across the time period of continuous trading in E-mini futures from -7:00 (17:00 of the prior day) to 15:15 Central Time (CT).[1]

Figure 1 displays daily values for the activity variables, obtained by averaging the one-minute observations across each trading day in our sample. As expected, we observe a strong degree of covariation between trading volume, transaction intensity, and return volatility. In particular, the persistent fluctuations in volatility are mirrored by broadly similar movements in trading activity, even if the latter series appear more irregular, or noisy. In addition, all major upward spikes in V and N are matched by corresponding peaks in volatility. Perhaps more unexpectedly, the lower right panel reveals a strong negative correlation between trade size and the other series. The dramatic decline in the average number of contracts exchanged per transaction, whenever volatility spikes, is particularly striking. This pronounced inverse time series association between trade size and volatility has hitherto not been explored in detail.[2] It defies the idea that trading activity "drives" return volatility, since this hypothesis implies a positive (or neutral) relationship between trade size and volatility. Instead, it suggests that agents actively

[1]Details on the construction of the measures are provided in Section III.A.

[2]The phenomenon is noted by Andersen and Bondarenko (2015) in their analysis of the forces driving the dynamics of the so-called VPIN metric.

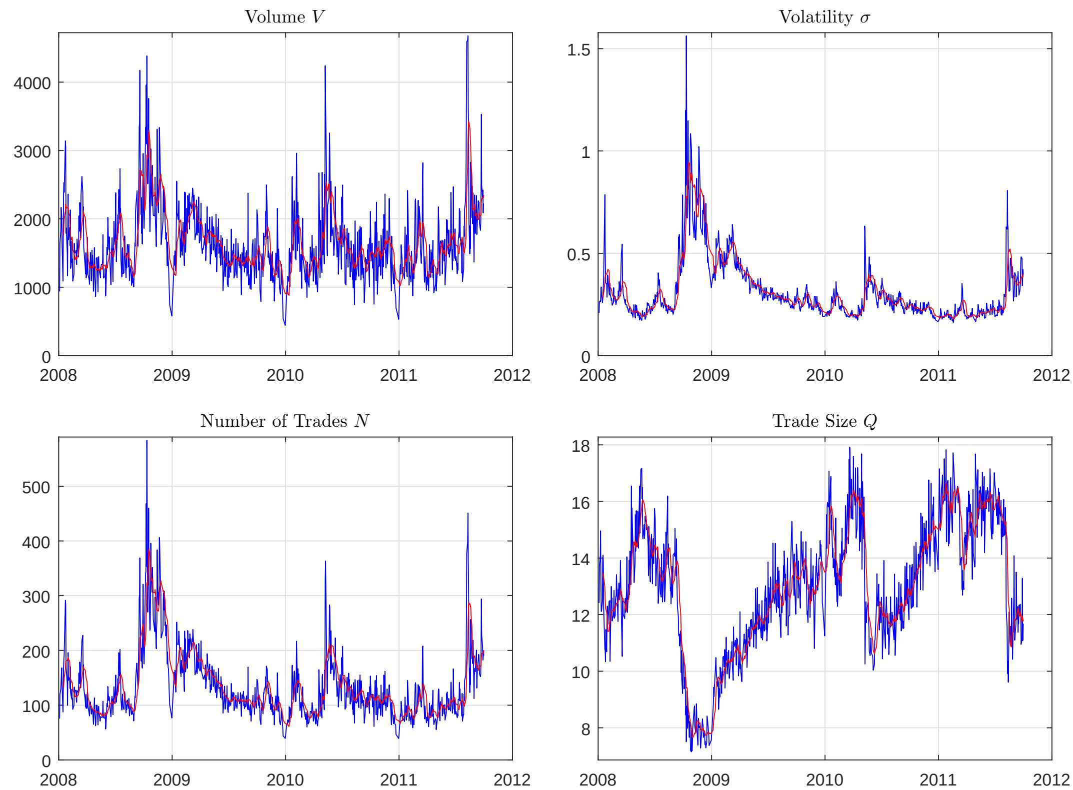

FIGURE 1. THIS FIGURE PLOTS THE TIME SERIES AVERAGES OF THE FOLLOWING ONE-MINUTE OBSERVATIONS FOR EACH TRADING DAY: VOLUME V_d, VOLATILITY σ_d (ANNUALIZED), NUMBER OF TRANSACTIONS N_d, AND TRADE SIZE Q_d. THE RED LINE INDICATES THE 2-WEEK MOVING AVERAGE. THE SAMPLE PERIOD IS JANUARY 4, 2008, TO SEPTEMBER 30, 2011.

limit their asset exposures as volatility, or risk, increases. Our task is to rationalize these features of the data jointly with the distinct and equally pronounced patterns observed across the intraday cycle below. Finally, we note that all series in Figure 1 appear stationary and, in particular, are void of any clear time trend.

Figure 2 depicts the systematic intraday variation in market activity variables, V, σ, N, and Q. The figure encompasses the full daily trading cycle in the futures market, with three regional trading zones separated by dashed vertical lines. The first regime spans -7:00 to 2:00 CT, covering the period before and during regular trading in Asia; the second regime contains observations in the interval 2:00 to 8:30 CT, roughly corresponding to standard European trading hours; and the third regime comprises 8:30 to 15:15 CT, matching active American trading hours. The series depicted in Figure 2 are averages, across all trading days in the sample, of the observations for each one-minute intraday interval.

There is extreme variation in the scaling of the diurnal pattern across the trading day. Volume increases more than fifty-fold and the transaction count grows twenty-five-fold as trading moves from Asian to American hours, while volatility and trade size only roughly double. Nonetheless, in line with prior findings, there is a strong

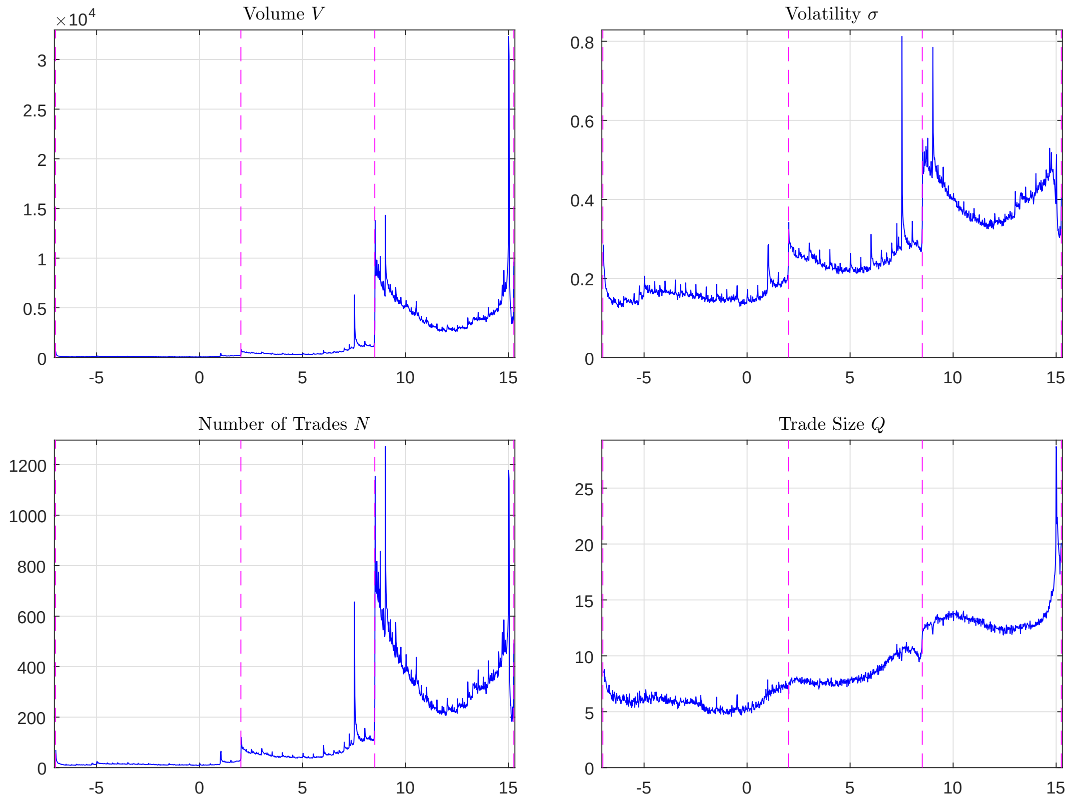

FIGURE 2. The figure depicts the averages, across all days, for contract volume V_t, annualized realized volatility σ_t, trade size Q_t, and number of trades N_t. The averages are computed at a granularity of one minute. The dashed vertical lines separate the three trading regimes corresponding to regular trading hours in Asia, Europe, and America. The sample period is January 4, 2008, to September 30, 2011.

commonality in the intraday variation of trading volume, transaction intensity, and return volatility. They all display a U-shaped pattern within the individual regions, and the pronounced spikes, typically associated with market openings or announcements, roughly coincide. In contrast, the fluctuations in trade size are more dissimilar across regions, with a distorted U-shape in the European zone and a distinctly different pattern during American trading. The latter feature represents a major discrepancy between the series in Figures 1 and 2. In contrast to patterns in Figures 1, the trade size correlates *positively* – not *negatively* – with return volatility in Figure 2. Rationalizing this distinct behavior across the time series and intraday pattern within the MDH framework is a tall order, but we shall see that the ITI offers a good resolution.[3] Finally, we have confirmed – consistent with prior evidence and the absence of arbitrage – that there is no systematic variation in the asset price level across the trading day.

[3]Alexander and Peterson (2007), Moulton (2005) and Gabaix et al. (2003) study time series variation in trade size, while Brennan and Subrahmanyam (1998) explore the intraday pattern.

A key variable in our empirical analysis is trade size. Ideally, we want to infer the number of contracts exchanged by the active party triggering a given transaction by submitting, say, a marketable limit order. Unfortunately, this quantity is often not discernible from the transaction data disseminated by exchanges. The problem arises, in part, from the practice of recording large trades, crossing several standing limit orders, as multiple transactions. For example, when a marketable order for 12 units is executed against four distinct limit orders involving, say, 2, 1, 5, and 4 units, it may be recorded as four distinct trades, reflecting execution of four (passive) limit orders, rather than as a single (active) order. Such issues can potentially be resolved if the full message record of the exchange is available. However, this is both expensive and cumbersome, requiring perfect alignment and sequencing of transactions. In fact, given the market fragmentation for many assets, with trading taking place simultaneously at several distinct venues, it is unclear whether proper sequencing and integration of the trading record across market segments is feasible.

These concerns explain our exclusive focus on the E-mini S&P 500 futures contract, traded on the CME Group Globex platform. This market has many desirable attributes. First, it is extremely liquid, with a turnover among the top two globally for exchange-traded futures, and it is a prime location for price discovery in the U.S. equity market. Second, it is a centralized market. The E-mini contract trades only through the Globex system; this ensures that we may identify the active trades and sequence them correctly. Third, the market operates almost continuously on weekdays, generating a huge degree of variation in the activity patterns across the Asian, European, and American segments; this provides us with excellent statistical power to identify the intraday interdependencies among the activity variables. Fourth, over our sample period, the CME Group data identifies the transaction size from the perspective of the active participant; hence, for the example above, we observe a single trade of 12 contracts. Fifth, the large tick size increases the quoted depth at best bid and offer; the market is sufficiently deep that orders for hundreds of contracts, worth tens of millions of dollars, regularly are consummated in one trade, rather than being broken up and executed over time.

Exploiting one-minute observations, we develop a log-linear regression specification relating the return variation per transaction and trade size. This representation encompasses the invariance relationship implied by a number of alternative theories, as they offer differing predictions for the slope coefficient on the trade size. In particular, MDH models stipulate coefficients of 0 or 1. Instead, we find the slope coefficient to be significantly negative, and often close to -2, both in the time series and intraday dimension. This implies that the return variation per transaction is inversely proportional to the square of the average trade size. As such, our findings align well with the prediction arising from the market microstructure invariance (MMI) principles, developed by Kyle and Obizhaeva (2016*a*). However, they develop the theory for bets or meta-orders, which are executed over long trading periods, not for transactions. The auxiliary assumptions required for their invariance principle to carry over to our high-frequency setting are very strict.

Hence, the hypotheses are distinct, and we refer to the invariance representation explored in this study as "Intraday Trading Invariance," or ITI.

Importantly, our tests for high-frequency trading invariance are unrelated to prior empirical work on MMI. We explore the nonlinear ITI relationship minute-by-minute, whereas MMI tests are couched in terms of the average trading intensity and return volatility over longer, typically monthly, horizons. Moreover, the latter focus on the variation of such monthly aggregates across distinct assets.[4] Consequently, existing tests of conformity with MMI rely on cross-stock comparisons. In contrast, we explore only a single market, but we check the invariance hypothesis across the pronounced variation over the daily trading cycle and, correspondingly, our time series tests reflect the accuracy of the nonlinear ITI relationship over one-minute observations. In fact, due to this inherent non-linearity, it is infeasible to test our specification of the ITI over longer intervals without imposing much stronger assumptions. In general, the ITI may not apply for time horizons that involve systematic variation in the underlying activity variables.

Since the ITI hypothesis invokes a notion of invariance, one may suspect kinship with the type of scaling laws documented elsewhere in the literature. For example, econo-physics studies invariance properties in the distribution of tail events, see, e.g., Gabaix et al. (2003) and Gillemot, Farmer and Lillo (2006). As we demonstrate below, the invariance properties of tail events do not imply that the ITI hypothesis holds, even if they can be compatible. Moreover, the scaling laws derived by Kyle and Obizhaeva (2017) using dimensional analysis, leverage neutrality, and invariance are consistent with the ITI relationships tested in this paper.

The remainder of this article proceeds as follows. Section II introduces the alternative MDH and ITI invariance hypotheses that we explore. Section III describes the sample, explains the construction of the variables, and verifies that our data display the usual features associated with tail invariance. Section IV provides an initial look at the empirical evidence by testing the invariance theories via daily time series observations. Section V develops the measurement and testing techniques for our exploration of the intraday invariance hypotheses. Section VI reports on the empirical results for both the time series and intraday pattern, documenting an overwhelming degree of support for ITI relative to the MDH theories. Section VII takes a detailed look at the performance of ITI during transitional periods, including the period around scheduled announcements, flash crash events, and shifts in trading from one regional market to another. Section VIII concludes. Section IX is an appendix providing supplementary information on data filtering, derivation of the estimation procedure, and additional empirical tests.

[4]Empirical studies of MMI hypotheses include Kyle and Obizhaeva (2016*a*), who study portfolio transitions with multiple-day durations across individual stocks, and Kyle and Obizhaeva (2016*b*), who rationalize the extent of market crashes through large (hidden) selling pressure, showing that the market impact is roughly consistent with MMI. In addition, Kyle, Obizhaeva and Tuzun (2016), Bae et al. (2016), and Kyle et al. (2017) explore monthly prints for individual stocks, the number of buy-sell switching points in retail trading accounts, and the number of news releases regarding firms, respectively, and relate these to monthly measures of market activity.

II. High-Frequency Trading Invariance

This section reviews a set of theoretical hypotheses that motivate the type of invariance propositions we explore in this paper. We cast them directly in a transaction level setting to accommodate our subsequent empirical work. Section II.A introduces basic notation for high-frequency observations, Section II.B presents some popular MDH specifications, and Section II.C develops a high-frequency parallel to the market microstructure invariance principle of Kyle and Obizhaeva (2016*a*), which we label Intraday Trading Invariance.

A. Basic Setting

The sample begins at time 0 and contains D trading days, each comprising T (small) intraday intervals of length $\Delta t = 1/T$. Thus, the full sample spans $[0, D]$ and contains the consecutive non-overlapping intervals $\tau = 1, \ldots, D \cdot T$. In order to identify the specific trading day and intraday time period associated with a specific interval, we also use double-index notation (d, t), with $d \in \mathcal{D} = \{1, \ldots, D\}$ denoting the trading day, and $t \in \mathcal{T} = \{1, \ldots, T\}$ the intraday interval. Hence, double-index interval (d, t) corresponds to interval $\tau = (d-1) \cdot T + t$ in our single-index notation.

For each interval τ, or equivalently (d, t), we indicate the random realization of any given quantity over the interval by including a tilde on top of the variable. Hence, the average (unsigned) number of contracts per transaction is denoted by $\widetilde{Q}_\tau$ or, alternatively, $\widetilde{Q}_{d,t}$. Similarly, using the single-index notation, we denote the average transaction price (dollars per contract) over the interval by $\widetilde{P}_\tau$, the average percentage return variance (per unit time) by $\widetilde{\sigma}^2_\tau$, the transaction rate (trades per unit time) by $\widetilde{N}_\tau$, and the cumulative trading volume (in contracts per unit time) by $\widetilde{V}_\tau$. Furthermore, we denote the logarithmic values of these quantities by corresponding lower case letters: $\widetilde{q}_\tau = \log(\widetilde{Q}_\tau)$, $\widetilde{p}_\tau = \log(\widetilde{P}_\tau)$, $\widetilde{s}_\tau = \log(\widetilde{\sigma}^2_\tau)$, $\widetilde{n}_\tau = \log(\widetilde{N}_\tau)$, and $\widetilde{v}_\tau = \log(\widetilde{V}_\tau)$. We note that, since $\widetilde{V}_\tau = \widetilde{N}_\tau \cdot \widetilde{Q}_\tau$, we have, $\widetilde{v}_\tau = \widetilde{n}_\tau + \widetilde{q}_\tau$. Finally, we refer to the expected value of a given variable in interval τ, conditional on information through time $\tau - 1$, by dropping the tilde. For example, the time $\tau - 1$ expected log trade size in interval τ is denoted, $q_\tau = E_{\tau-1}[\widetilde{q}_\tau]$.

B. The Mixture-of-Distributions Hypothesis

There is a long history of theories stipulating a relationship between trading intensity and return volatility, with the most celebrated contribution being Clark (1973). This literature dates back at least to Osborne (1962), who observes that if the returns associated with distinct trades are independent, then the return variance will be proportional to the number of transactions, $\sigma^2 \sim N$. Likewise, Brada, Ernst and Van Tassel (1967) show that the return defined over a fixed (large) number of transactions is approximately Gaussian, lending additional support to this proportionality relationship over intraday horizons. Osborne (1962) further notes that trading volume is proportional to the number of transactions, implying that return volatility is proportional to volume as well, $\sigma^2 \sim V$.

The above hypotheses imply that the return distribution is generated through a mixture of the distributions for trading activity and the i.i.d. return component associated with the individual trades, motivating the label "Mixture-of-Distributions Hypothesis" (MDH). Historically, the only official data regarding trading activity provided by the exchanges were the daily trading volume. Hence, subsequent studies largely tested, and confirmed, the qualitative implications of the MDH by documenting a strong positive association between the contemporaneous trading volume and return volatility, e.g., Granger and Morgenstern (1970), Morgan (1976) and Westerfield (1977).

To obtain an empirically tractable log-linear representation, we take logarithms in the proportional volatility-volume specification. Moreover, since the true return volatility and trading intensity are latent variables that we proxy by realized volatility measures and observed number of trades, we will only test trading invariance through estimates of the relations among the expected values of the market activity variables. Consequently, we take the expectations conditional on time $\tau - 1$ and, recognizing observation errors and frictions, we allow for an error term, leading to the following regression-style relationship:

The Mixture-of-Distributions-Hypothesis in Volume (MDH-V)

$$s_\tau \;=\; c + v_\tau + u_\tau, \qquad \textit{for } \tau = 1, \ldots, D \cdot T, \tag{1}$$

where c denotes a generic constant, and u_τ is a mean-zero residual.

While the distinction between the number of transactions and trading volume was largely ignored in the early literature, the identity of the relevant trading activity measure became a topic of separate interest later on. Jones, Kaul and Lipson (1994) find the number of trades better aligned with daily volatility, while Ané and Geman (2000) stipulate that intraday returns are i.i.d. Gaussian, when normalized by the (stochastic) transaction count. These studies suggest that the relevant variable is the number of transactions, generating the following alternative MDH specification:

The Mixture-of-Distributions-Hypothesis in Transactions (MDH-N)

$$s_\tau = c + n_\tau + u_\tau, \qquad \textit{for } \tau = 1, \ldots, D \cdot T, \tag{2}$$

where c denotes a constant and u_τ is a mean-zero residual.

Equations (1) and (2) provide special cases of a process evolving according to a stochastic business-time clock, as formally introduced into the modeling of financial returns by Mandelbrot and Taylor (1967). In the above setting, the driving process for return variability is directly observed, but it may generally be a latent economic variable such as economic news arrivals. Within this framework, and using transaction level data, Harris (1987) also finds empirical support for the association of volatility with both trading volume and the transaction count.[5]

The MDH specifications are intuitive, following directly from assumptions regarding the returns generated per transaction or per unit traded. Nonetheless, once we let market activity be governed by an underlying (latent) business-time clock, it becomes natural to contemplate more general relationships. In fact, Clark (1973) experiments with alternative functional representations and finds support for an increasing and convex impact of volume on volatility for daily data. However, as noted by Epps and Epps (1976), this comes at a cost, because this type of non-proportionality will not survive temporal aggregation. It renders the relationship inconsistent with the identical representation for transaction data. Instead, Epps and Epps (1976) argues that this type of specification is suitable at the intraday, or even the transaction-by-transaction, level. The point is that price impact may grow disproportionally with trade size. Thus, while the volatility per transaction is constant for fixed trade size, we should allow for a separate impact of the trade size, e.g., $\sigma^2/N \sim Q^\beta$, $\beta \geq 0$. The above reasoning leads to a direct generalization of equations (1) and (2):

The Generic Mixture-of-Distributions-Hypothesis (MDH)

$$s_\tau - n_\tau \;=\; c + \beta\, q_\tau + u_\tau, \qquad \textit{for } \tau = 1, \ldots, D \cdot T, \tag{3}$$

where c is a constant, $\beta \geq 0$, and u_τ is a mean-zero residual.

Equation (3) encompasses MDH-V and MDH-N for $\beta = 1$ and $\beta = 0$, respectively. At the same time, it allows for the trade size, q_τ, to be related to return volatility at the high-frequency level, beyond the effect of the sheer number of transactions. The findings of Clark (1973) and Epps and Epps (1976) suggest $\beta > 1$, while $0 < \beta < 1$ would correspond to a quantitative effect that falls between the predictions of MDH-V and MDH-N. Regardless of the actual value of β, if equation (3) is satisfied, the scenario describes specific invariant relationships between return volatility and trading activity variables.

Finally, we add a caveat regarding empirical exploration of the MDH relationship (3). It takes the form of a standard regression equation, suggesting that inference is feasible via standard techniques. As always, this is only correct under suitable regularity conditions. One primary requirement is that the specification is balanced: the relevant variables are stationary or at least of corresponding stochastic

[5]In addition, Andersen (1996), Bollerslev and Jubinski (1999), and Liesenfeld (2001) refine the volatility-trading intensity relationship in extended MDH specifications. Meanwhile, the striking results in favor of the MDH-N from transaction returns in Ané and Geman (2000) are not replicable according to Gillemot, Farmer and Lillo (2006), while Murphy and Izzeldin (2010) document a large degree of imprecision in the moment extraction of Ané and Geman (2000), and also conclude that normality of the standardized returns does not hold upon close scrutiny.

order. Upon inspection, this is not a trivial condition. While we may, reasonably, expect the volatility process to be stationary in the time dimension, this is more problematic for the transaction intensity, as documented by Tauchen and Pitts (1983). A variety of factors may induce an exogenous, secular trend in the trading activity, including an overall growing economy, shifts in asset allocation strategies of agents due to market innovations, and changes in the cost of trading. Thus, we may want to limit our MDH explorations to shorter time series for which we can ignore the impact of the secular trend, or we may seek to directly account for any time-varying trend in the empirical specification.

C. Intraday Trading Invariance

Another source of hypotheses for our investigation of alternative high-frequency trading invariance propositions comes from the recent work of Kyle and Obizhaeva (2016*a*). They develop a model in which return volatility is driven by the price impact of large speculative bets, representing portfolio reallocations among major investors. They obtain a market microstructure invariance (MMI) hypothesis, stating that the dollar risk transferred by a bet is invariant when measured in units of business time, with the latter reflecting the rate of bet arrivals.

While Kyle and Obizhaeva (2016*a*) develop the MMI theory by applying their invariance principles to large speculative bets, executed through sequences of smaller orders to reduce transaction costs, we obtain the Intraday Trading Invariance, or ITI, by stipulating that similar relationships apply to transactions conducted over short horizons. The sufficient conditions for the original invariance relationships to carry over into a high-frequency setting are strict, and we do not argue that they are valid in practice. Hence, we view the ITI as a purely empirical hypothesis, yet motivated by the invariance relationships that emanate from the MMI principle.

Translating the Kyle and Obizhaeva (2016*a*) prediction that the dollar risk transferred by a bet is invariant, when measured in units of business time, to an intraday setting, we obtain this assertion: the random variable, $\widetilde{I}_\tau = P_\tau \widetilde{Q}_\tau \widetilde{\sigma}_\tau / \widetilde{N}_\tau^{1/2}$, has an invariant distribution across all intervals τ. Assuming the agents, actively engaged in trading, are aware of the current volatility and transaction rates, σ_τ and N_τ, this implies that they adjust the scale of their trades to control the risk associated with the change in their asset portfolios. In effect, trade size is *random*, but "drawn" from a distribution that ensures the invariance relationship holds.

Our objective is not to test all implications of this invariance conjecture but rather to confront and compare its performance to that of the MDH. First, we note that, relative to the systematic variation in volatility and trading intensity across the daily cycle captured in Figure 2, the expected fluctuation in the price level is negligible. Hence, for exploration of the intraday pattern and, more generally, for periods covering short intervals of time, the expected variation in the price level is dwarfed by the variation in the other variables, and may effectively be treated as constant. This is the approach we adopt below, as it also has the advantage of rendering the discrepancy versus the MDH propositions transparent. Hence, our empirical hypothesis is not directly linked to the MMI theory.[6]

Consequently, over short intervals, we stipulate that the trading invariant is proportional to a nonlinear function of the market activity variables, $\widetilde{I}_\tau \sim \widetilde{Q}_\tau \cdot \widetilde{\sigma}_\tau / \widetilde{N}_\tau^{1/2}$. Taking the logarithms on both sides and then taking expectations, conditional on information at time $\tau-1$, we obtain an invariance relationship reminiscent of the general MDH representation (3),

$$E_{\tau-1}[\log \widetilde{I}_\tau] \;=\; E_{\tau-1}\left[\log\left(\widetilde{Q}_\tau \cdot \widetilde{\sigma}_\tau / \widetilde{N}_\tau^{1/2}\right)\right] \;=\; \text{c.}$$

Rearranging and acknowledging the presence of an error term, we obtain a convenient representation of the trading invariance relationship that we explore empirically:

Intraday Trading Invariance (ITI)

$$s_\tau - n_\tau \;=\; c \,-\, 2\, q_\tau \,+\, u_\tau, \qquad \textit{for } \tau = 1, \ldots, D \cdot T, \tag{4}$$

where c is a constant and u_τ is a mean-zero residual.

Compared to the MDH specification (3), the ITI hypothesis imposes the constraint $\beta = -2$. This contrasts sharply to the MDH implication that $\beta \geq 0$.

Intuitively, this stark discrepancy between the ITI hypothesis and MDH arises from the notion, in the ITI, that current market conditions modify the agents' desired trade sizes. All else equal, the ITI implies that the mean trade size drops if the return volatility rises or the trading intensity declines. No such endogenous response is present in the MDH. Empirically, this implies that the variation in trade size is critical for disentangling the alternative theories. Specifically, if the trade size does not vary systematically with market conditions, i.e., q_τ is constant, we have no ability to discriminate among the different theories, as return volatility will move in line with the trading intensity, whether captured through the transaction count or trading volume. Of course, Figures 1 and 2 reveal a large degree of systematic variation in the trade size, both within and across trading days, which greatly facilitates the empirical identification of the different effects.

III. Data Description

This section reviews the CME data available for our study, provides some summary statistics, and describes how we filter the observations to construct the variables ultimately used in the empirical analysis.

[6]For the intraday tests, the inclusion of the price level is literally immaterial. For completeness, Appendix IX.A reports time series results that include the price level in a manner inspired by the MMI. While the results do not change greatly, they are a bit weaker overall.

A. The E-mini S&P 500 Futures Data

We rely on the best bid-offer (BBO) files for the E-mini S&P 500 futures contract from CME DataMine. These "top-of-the-book" files provide tick-by-tick information regarding the best quotes, order book depth, trade price, and trade size, time-stamped to the second. The contract is traded exclusively on the CME Globex platform, so we observe all transactions executed over our sample, covering nearly four years, January 4, 2008 to September 30, 2011. The E-mini contract trades almost twenty-four hours a day, five days a week. The trading hours are governed by Central Time (CT), so all our references to time are expressed in CT.

The procedure behind the recording of trades and trade sizes is critical for our empirical tests. When an executable order arrives, it is often matched with more than one limit order resting at the top of the book at the time of execution. During our sample period, the exchange reported all contracts traded at an identical price against an incoming order as a single combined transaction quantity. Thus, the trade size is reported from the perspective of the active party placing the executable order. This convention is consistent with the notion of trade size associated with the intraday invariance principle. Since most markets do not record transactions in this manner, it will typically be impossible to test for invariance relationships at the high-frequency level, as we do here, unless one is able to identify the active orders initiating the recorded trades.[7]

The notional value of the E-mini S&P 500 futures contract is $50 times the value of the S&P 500 stock index, denominated in index points. The contract has a tick size of 0.25 index points ($12.50), equivalent to approximately 2 basis points of notional value during our sample. It has four expiration months per year. We use data for the front month contract until it reaches eight days to expiration, at which point we switch into the next contract. This ensures that we use the most actively traded contract throughout our analysis.

As indicated in Section I, our empirical analysis focuses on the uninterrupted trading session from 17:00 to 15:15 the following day, covering three largely separate regional zones. Thus, our trading week opens on Sunday at 17:00 and ends on Friday at 15:15. Daily trading in the Globex market is initiated by a batch auction at 17:00. The contracts traded during the opening auction are recorded during the first second of trading after 17:00. To generate a sample with a homogeneous trading protocol, we remove all transactions within the first second after 17:00. There is also trading on Monday through Thursday between 15:30 and 16:30. The market is often not particularly liquid in this trading window. Moreover, it involves another batch auction and it is not open on the first trading day of the week (Sunday afternoon) and post-holidays. In order to avoid issues of poor liquidity and non-homogeneity, we exclude this fairly inactive trading hour from our analysis.

[7]For example, if the exchange counts each limit order involved in trading as a distinct trade, the transaction count will reflect the limit order flow from the supply side that is intermediating trades. This will inflate the number of (active) transactions and shrink the trade size relative to the procedure used by the CME Group for the E-mini contract over our sample period.

Intraday trading invariance should hold for any exogenously selected subset of intervals within the trading day, so a priori exclusion of isolated or irregular trading segments does not affect the validity of our testing procedures. As detailed in Sections I and V.A, we focus on three intraday regimes corresponding, approximately, to the trading sessions in Asia, Europe, and America. They contain, respectively, 540, 390, and 405 one-minute intervals, constituting the three non-overlapping sets of intraday intervals, $\mathcal{T}_r$, $r = 1, 2, 3$. Furthermore, we exclude the entire day from our sample if there is an unusually low level of activity during one or more of the regimes. Such occurrences typically stem from a breakdown in data transmission or slow holiday-related trading. Since our invariance tests rely on averaging observations for given time-of-day intervals intertemporally as well as across intraday trading segments, it is critical to include regular trading data from the distinct periods in equal measure. The filtering procedure and the days eliminated are described in Appendix IX.B.[8]

After filtering, we have 899 complete trading days with, respectively, 236, 241, 237, and 185 days in the years 2008–2011. We let $\mathcal{D}_y$ represent the trading days in each calendar year, where $y = \{2008, 2009, 2010, 2011\}$. Hence, the observations for regime r in year y are contained in the set $\mathcal{D}_y \times \mathcal{T}_r$.

B. Descriptive Statistics

Figures 1 and 2 in Section I depict the time series and intraday evolution of the market activity variables across our sample period based on underlying one-minute observations. As detailed in the prior section, we have readily accessible data for the price, number of active trades, trade size, and aggregate trading volume. The return volatility is latent, however, and we use a (noisy) measure of realized volatility instead. The high liquidity of the E-mini contract renders it feasible to utilize six consecutive squared ten-second returns to compute a standard realized volatility measure without any major detrimental effects from microstructure noise. This provides a noisy, yet effectively unbiased, proxy for the true (average) spot variance for each one-minute interval. Nonetheless, the depicted time series and intraday estimates for volatility are precise and meaningful. They represent averages obtained over a vast number of ten-second intervals, either across intervals within a given trading day or across trading days. Consequently, these series benefit from the same error diversification principle that accounts for the accuracy of standard high-frequency realized-volatility estimators.[9]

The pronounced covariation among the activity variables, in both the time series and intraday dimension, is evident from Figures 1 and 2. At the same time, there

[8] We have confirmed that this approach improves the homogeneity of the sample, but the qualitative results are identical if we retain all trading periods across the sample.

[9] See Andersen, Bollerslev and Diebold (2010) or Ait-Sahalia and Jacod (2014) for a review of volatility estimation from high-frequency data. Realized volatility measures are often obtained from 78 five-minute or 390 one-minute squared returns. We obtain a substantial improvement by averaging across nearly 1000 trading days or thousands of ten-second intraday intervals.

are striking difference in the scaling of the variables across the regional trading zones and some distinct discrepancies in the interaction among the variables across the intraday pattern versus in the daily time series. Table 1 complements the figures with basic summary statistics.

Table 1—Descriptive Statistics for the E-mini S&P 500 Futures

	Asia	Europe	America	Combined			
				Mean	Min	Max	Ratio
Volatility	0.16	0.25	0.40	0.26	0.13	0.81	6.4
Volume	96	604	4788	1668	52	32328	624
# Trades	14	67	363	136	9	1272	144
Notional Value, $Mln	5	34	269	94	3	1804	614
Trade Size	6.0	8.5	13.4	9.0	4.6	28.7	6.3
Market Depth	56	274	999	406	36	3597	100
Bid-Ask Spread	26.4	25.6	25.1	25.8	25.1	27.9	1.1

The statistics are reported separately for the three regimes. The last four columns show the mean, minimum, maximum, and maximum-to-minimum ratio for the entire day. The volatility measure represents realized volatility, computed from six ten-second squared returns across each one-minute interval, and then averaged across all observations and reported in annualized terms. The volume, notional value, and number of trades are one-minute averages. Market depth is the average sum of the number of contracts at the best bid and ask. The bid-ask spread is measured in index points times 100.

It is striking to note from Table 1 that the average bid-ask spread for each regime barely exceeds the minimum value of 0.25 index points. Hence, the spread is binding and equals a single tick almost always, implying that the tick size is "large" for this liquid market. Similarly, a disproportionate number of trades in all three regimes involve a single contract, implying that the contract-size constraint is binding as well.[10]

In contrast, the market is less constrained in terms of consummating large transactions. This is evident from the extreme size variation across trades on display in Figure 3. While over half of the trades involve one or two contracts, there are, on average, a couple of transactions per minute during American trading hours exceeding 500 contracts. Table 1 shows there usually is enough depth at the best quotes to absorb trades of hundreds of contracts. Large trades are the primary determinant of the average trade size, suggesting this statistic is robust to distortions stemming from the minimum contract size. In this respect, the market provides a near ideal setting for exploring the interaction between trading intensity and volatility: it has an integrated electronic trading platform, a uniformly high degree of transaction activity, and a deep order book. The latter features allow us to observe pronounced variation in the trade size over time.[11]

[10]This feature is documented in Figure 17 of the Appendix IX.C. For early work on the impact of market frictions on the trading process, see, e.g., Harris (1994), Angel (1997), Goldstein and Kavajecz (2000), and Schultz (2000).

[11]Further study of the relationship between tick size and average trade size takes us beyond

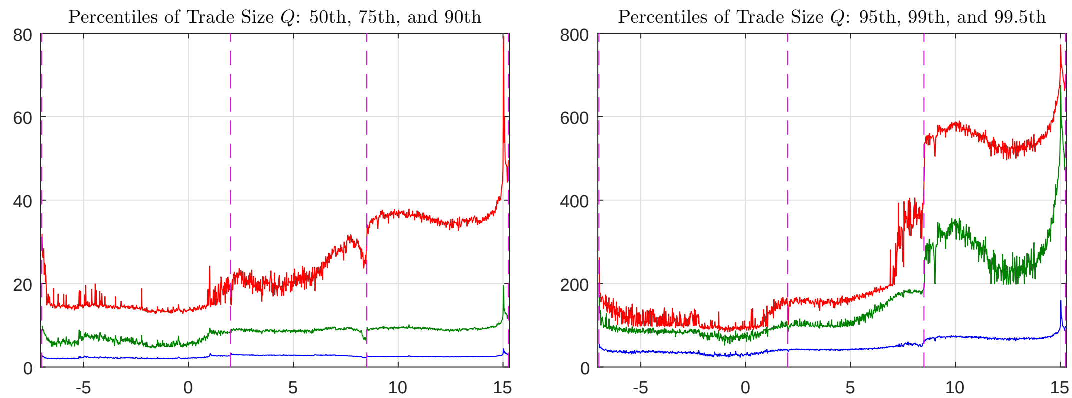

FIGURE 3. THE FIGURE PLOTS PERCENTILES OF THE TRADE SIZE Q. THE STATISTICS ARE COMPUTED AT THE GRANULARITY OF ONE-MINUTE. THE DASHED VERTICAL LINES SEPARATE THE THREE TRADING REGIMES.

We further note that, even during regular market conditions, we lose intermittent observations, as we test for invariance only when the market displays meaningful activity. The main constraint arises from our use of logarithmic values for each one-minute observation. To render this feasible, we exclude all intervals that feature zero volume (no trading activity) or no price variability (all six ten-second observations feature the identical price). The fraction of omitted observations in the three regimes are 28.36%, 8.01% and 1.20%. One may worry that the elimination of data *conditional* on zero intra-interval price changes may induce a bias to the estimated volatilities. As a robustness check, we also conduct our tests at the five-minute sampling frequency, which enables us to retain nearly all observations. We now filter only 1.26%, 0.12% and 0.11% of the observations across the three regimes. Our results are qualitatively identical and quantitatively very similar, so we only present results for the one-minute sampling. The latter provides a more challenging setup and allows us to enhance the granularity of the tests around key transition points, when the characteristics of the trading process shift abruptly.

C. Tail Invariance

A large literature documents the presence of a slow power law decay in the tails of the marginal distribution for many financial market variables, including volatility and trading intensity. Moreover, these features have been established at sampling frequencies ranging from transaction level to monthly. As such, the laws governing the tail behavior of volatility and trading intensity are often labelled scale-invariant, meaning their large fluctuations are highly non-Gaussian and obey specific regularities that elude traditional central limit theories. Instead, the behavior aligns well with stable distributions whose tail characteristics survive aggregation.

Importantly, these tail invariant features are entirely distinct from the trading invariance propositions explored in this paper, in the sense that neither implies

the scope of this paper.

the other. However, the tail characteristics do capture important aspects of the market dynamics. Furthermore, under certain conditions, they do shed additional light on the alternative MDH theories. Thus, in this section we provide a brief overview of the tail behavior of the activity variables across our sample.

The power laws pertain to the *marginal* distribution. Specifically, for observations on a variable y, the power law implies that the probability y exceeds a given tail threshold x is proportional to an inverse power of x,

$$P\,(\,y\,>\,x\,)\;\sim\;1\,/\,x^{\eta_y}\,. \tag{5}$$

The exponent η_y in equation (5) captures the rate of decay in the extreme observations as the tail threshold increases. A lower value for η_y implies a heavier tail and a more extreme deviation from Gaussian tail behavior.

In contrast, trading invariance stipulates that certain activity variables are governed by a *joint nonlinear* relationship, impacting their full joint distribution, as illustrated through our derivation of equations (1)-(4). These specifications have no direct implication for the shape of the tail of the variables, so the underlying concepts of invariance are, in principle, entirely unrelated.

Nevertheless, since tail power laws are near universal, it is natural to entertain the idea that both types of invariance apply simultaneously. In this case, under suitable assumptions, the MDH governs the relative size of the tail exponents, see, e.g., Gabaix et al. (2003, 2006).[12] This reasoning does not carry over to the ITI hypothesis, however, as it represents an equilibrium relationship among multiple variables. An extreme realization for the transaction count, say, may be accompanied either by an extreme value for volatility or the trade size, or by a combination of less extreme observations for both. As such, the ITI does not imply any simple relationship for the distinct tail exponents of the marginal distributions.

We now briefly summarize the empirical tail behavior of the activity variables in our E-mini S&P 500 futures sample. A few unusual features of the sample demand consideration. First, the series operate at distinctly different scales in the three trading segments. Hence, the vast majority of the outliers occur during the most active zone, namely American trading. Second, as a caveat, we note that the manner in which individual trades are recorded varies across market settings, so it is not obvious whether the transaction count series is comparable to those used in prior studies. Third, since we focus on high-frequency observations, there is a large degree of clustering among the outliers. The largest return volatility and trading volume episodes are associated with the financial crisis, the flash crash on May 6, 2010, and the aftermath of macroeconomic announcements. For these

[12]Nonetheless, the validity of an MDH relationship, along with the existence of power laws for the tails of volatility and trading activity, does not guarantee a specific relationship among the tail exponents. Such predictions also involve a restriction on the joint distribution of the variables and the residual in the MDH relationship. Gillemot, Farmer and Lillo (2006) argue that this condition is violated, even if the tests of Gabaix et al. (2003) are supportive of the relationship.

reasons, we conduct the analysis separately for each region, effectively treating them as distinct; we focus primarily on the volatility and volume series; and we analyze both raw and normalized one-minute observations. The normalization reflects division by the median value over the respective trading segments for the given day, so that we identify genuine outliers relative to the typical observations under prevailing market conditions.

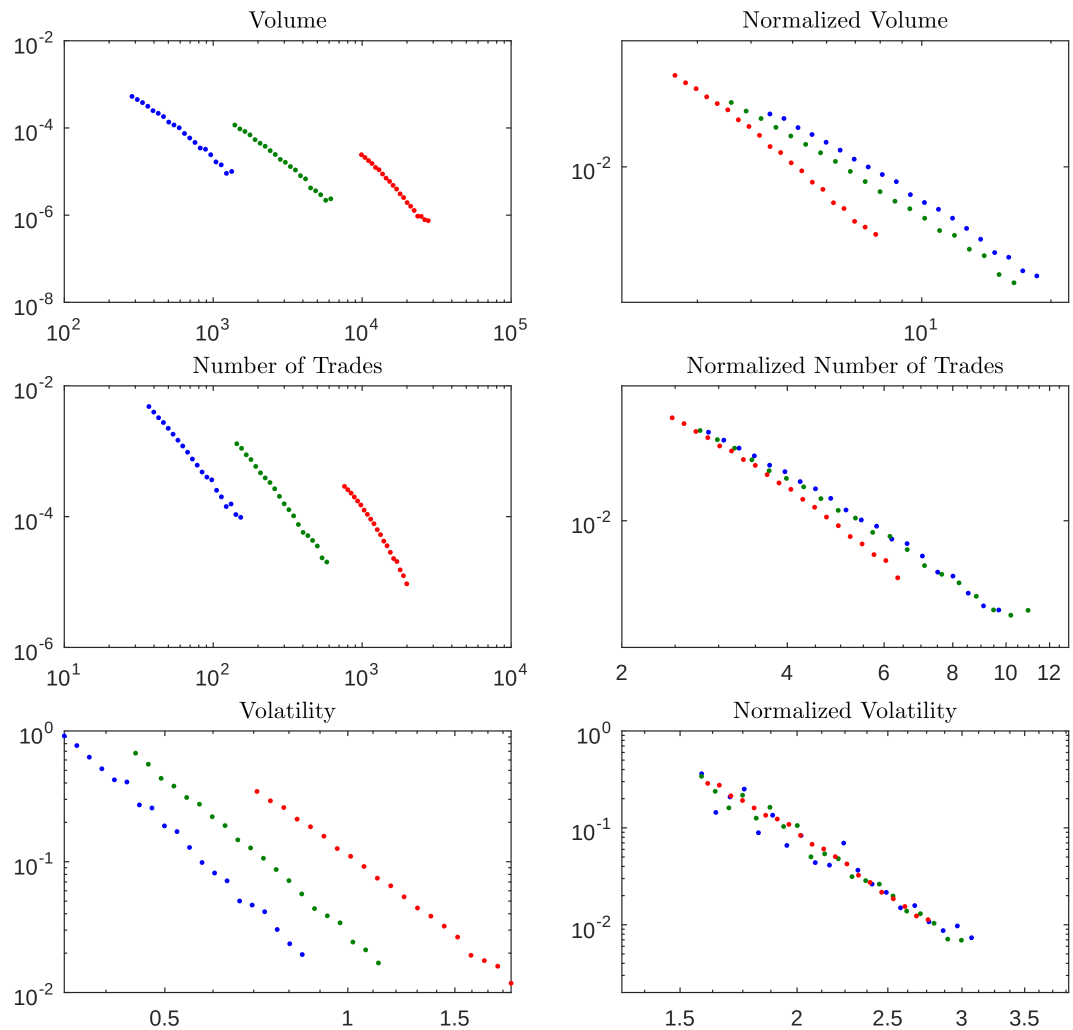

FIGURE 4. The tail distribution (top 10%) of the one-minute observations for volume V, number of trades N, and volatility σ in log-log scale. The regimes are depicted in blue (Asia), green (Europe) and red (America). The right panels show the tail distribution for the series normalized to have a median of unity each trading day.

Figure 4 depicts the top ten percentiles of the one-minute volume, trading intensity, and volatility observations. If a power law applies, these log-log plots should constitute straight lines for the largest observations, with the absolute value of the (negative) slope representing the tail exponent. The lines for the different trading zones, depicted in the left panels, are naturally separated, reflecting the higher activity levels—also in the tails—for the more active regions. However, once we normalize the observations, as in the panels on the right, they are quite similar, signaling that the size of the relative outliers and their decay rates across the three trading regions are comparable. This is an interesting finding, although tail invariance per se does not imply that these curves coincide, but merely that—sufficiently far in the tail—they are linear. We conclude that, to first order, the power law for the tail decay provides a good characterization for the activity variables in our study. This confirms that our activity variables display characteristics consistent with those observed for many other series.

As noted, under auxiliary assumptions, the relative values of the tail exponents will be determined by the MDH relationship. The exponents for volume, both across regimes and for raw as well as normalized data, in the top row of Figure 4 are close to 3, while the transaction count series in row two generate values between 3 and 4 and, finally, the volatility tail exponents associated with the bottom row fall between 4 and 6. These values are roughly consistent with the estimates obtained by Gabaix et al. (2003) for N and σ, but much higher for V, where they report a value of 3/2.[13] If we go along with the larger estimates of $\eta_\sigma = 6$, or equivalently, $\eta_{\sigma^2} = 3$, and accept $\eta_V = \eta_N = 3$, then we confirm the required balance between the tail exponents in the MDH-V and MDH-N relations.[14] However, since these estimates carry a great degree of uncertainty and the deviations from the above values are quite sizable in many cases, the (indirect) evidence based on the tails of the marginal distributions is at best mixed.

IV. Trading Invariance at the Daily Level

To set the stage, we initially explore standard MDH style regressions at the daily level. In addition, given the subsequent focus on ITI, we also check whether the average trade size adds explanatory power beyond the MDH specifications. Following equations (1) and (2), we aggregate the volume, transactions, and realized volatility measures across the trading cycle to obtain daily observations, then apply a logarithmic transformation and run the relevant regressions.

Since most prior work relies on data from markets with a single regional trading session, we report results corresponding to the U.S. trading hours only in the left panels of Figure 5, while the right panels reflect the full trading day. In the top row, we observe the expected positive MDH-V association between return volatility and trading volume. The second row documents an even closer MDH-N relationship between the volatility and transaction intensity, as the observations appear distributed quite tightly around the regression lines. In all cases, the regression slopes are highly significant, ranging from 1.66 to 1.90, and significantly larger than

[13] In line with our results, Gillemot, Farmer and Lillo (2006) also obtain values for η_V that exceed 3/2. As for Gabaix et al. (2003), their study is based on 15-minute observations.

[14] Recall the following rules for tail exponents of distributions of random variables x and y: $\eta_{x^\alpha} = \eta_x/\alpha$ and $\eta_{xy} = \min(\eta_x, \eta_y)$.

unity in all cases. Moreover, the degrees of explained variation for log volatility, R^2, are 53% and 48% in the first row, while it is 88% and 84% in the second, consistent with the conclusion by various authors that the transaction count is a better candidate for a business-time clock than the cumulative volume.

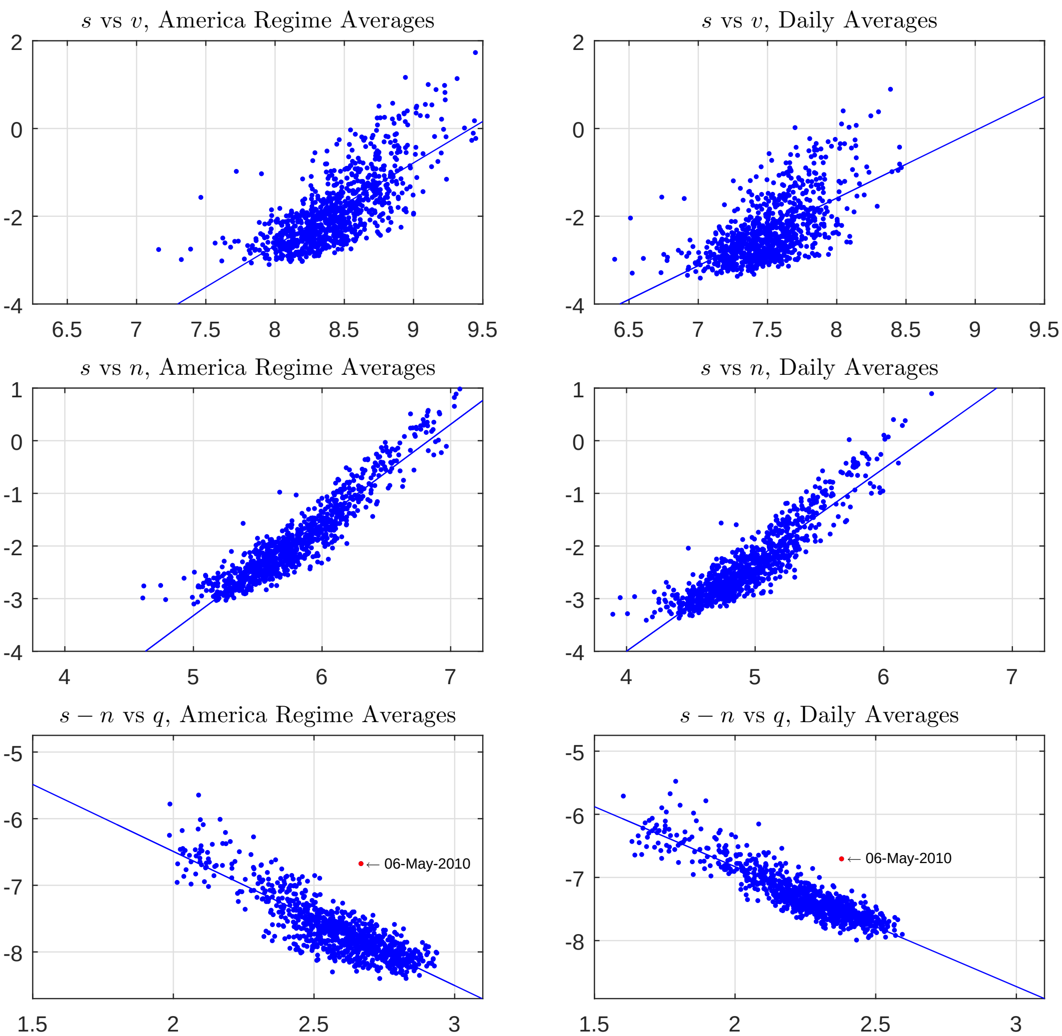

Figure 5. The left panels reflect the logarithms of daily observations across the U.S. trading hours. From top to bottom, they depict observations and regression lines for equations (1), (2), and (4), respectively. The right panels present the corresponding plots for the logarithms of the observations aggregated across the full trading day.

The regression slopes well in excess of unity in the two top rows indicate that the basic MDH variants do not apply for daily data. While this may reflect a genuine nonlinearity in the interplay among return volatility and trading intensity, it is problematic in the sense that this nonlinear relationship cannot be valid across the intraday sampling frequencies, as it is not preserved under temporal aggregation. We explore the evidence for the intraday MDH relationships in Section VI.

Besides the qualitative evidence regarding the MDH specifications, we also noted the apparent association between return volatility and trade size in Figure 1. This relationship is intrinsic to the ITI in equation (4). Consequently, the bottom row of Figure 5 displays the corresponding data points and fitted regression line. They are indeed negatively sloped with coefficients -2.01 and -1.94 which, notably, are not statistically different from the value of -2 implied by equation (4).

In terms of the implied explanatory power for log volatility, the ITI relationship displayed in the bottom row of Figure 5 reaches 92% and 94%, indicating a sizable incremental increase relative to the MDH-N relationship, especially for the full trading day regression.[15] We further note that one of the outliers in the bottom panels corresponds to the "flash crash" on May 6, 2010, when the volatility per transaction significantly exceeded the ITI prediction. A detailed exploration of the invariance relationships during this turbulent episode is provided in Section VII.B.

V. Testing Intraday Invariance

We now develop a formal framework for robust and powerful testing of the alternative intraday invariance hypotheses embedded within equation (3). Since they involve high-frequency expectations or real-time interactions among rapidly fluctuating variables, this requires careful measurement procedures.

A. Methodology for High-Frequency Measurement

In this section we develop practical measurement techniques for the latent log-transformed high-frequency variables, so that they are amenable for analysis through regression-based tests. For brevity, the more technical details behind our approach are deferred to Appendix IX.D.

The intraday activity measures consist of persistent, non-negative random variables that are subject to sizable idiosyncratic shocks. The multiplicative error models (MEM) provide a suitable framework for such settings.[16]

We let $\tilde{Y}_\tau$ denote a strictly positive random variable (representing V, N, or σ^2), and stipulate that the dynamics of $\tilde{Y}_\tau$ is given by,

$$\tilde{Y}_\tau \;=\; Y_\tau \cdot \tilde{U}_\tau, \tag{6}$$

[15]We obtain this measure by rewriting the regression as $s = c+n+\beta q+u$, where the coefficient for n is fixed at unity. This generates the identical slope coefficient for q as in Figure 5, but the R^2 now speaks to the explained variation for s alone. Since the coefficient on n is fixed, and not estimated to fit optimally for the given sample realizations, the playing field is leveled vis-a-vis the MDH relations in terms of the number of free parameters.

[16]The MEM family includes standard GARCH and stochastic volatility specifications. Early on, it was used to capture high-frequency return dynamics through the interaction of daily volatility factors and intraday diurnal features, see Andersen and Bollerslev (1997, 1998). Engle (2002) formulates the general modeling framework and applies it to both volatility and trade durations. Finally, Brownlees, Cipollini and Gallo (2011, 2012) provide successful applications to dynamic modeling of high-frequency trading volume.

where $\tilde{U}_\tau$ is strictly positive, i.i.d., with a mean of one and variance of σ_U^2. Equation (6) is a MEM model with time $\tau - 1$ conditional expectation $E_{\tau-1}[\tilde{Y}_\tau] = Y_\tau$ and $Var_{\tau-1}[\tilde{Y}_\tau] = Y_\tau^2 \cdot \sigma_U^2$. The conditional dependence structure is captured by the positive mean, Y_τ, which is unspecified and can be highly complex, incorporating long-run dependencies, diurnal patterns, and short-run serial correlation.

Our approach merely requires that we obtain unbiased estimates for Y_τ over each interval τ. A natural choice is to use the observed number of transactions and trading volume in the interval as noisy, yet unbiased, estimators for the corresponding ex-ante conditionally expected trading activity. Similarly, one may devise a high-frequency realized volatility estimator with similar attributes for the return variation, as detailed in Section III.B.

As documented in Appendix IX.D, one may now show that the corresponding logarithmic values of the activity variables, $\tilde{y}_\tau = \log \tilde{Y}_\tau$, take the form,

$$\tilde{y}_\tau \;=\; y_\tau \;+\; c \;+\; \tilde{\epsilon}_\tau, \tag{7}$$

where y_τ represents the conditionally expected logarithmic value of the activity variable, while $\tilde{\epsilon}_\tau$ are i.i.d. random variables with zero mean and finite variance.

In principle, one may now simply include these indicators of the interval-by-interval activity variables as proxies for the conditional expectations in the relevant regression. However, the estimator in equation (7) inevitably will be noisy as the variance of the error term, $\tilde{\epsilon}_\tau$, typically is large. If the regressors are subject to sizable measurement errors, we face severe errors-in-variables problems. Thus, we do not propose tests that exploit observations from individual intervals, but develop aggregation schemes that mitigate the errors, yet retain statistical power and allow for a natural economic interpretation.

Since we aim to test the invariance hypotheses along both the time series and intraday dimension, we construct the variables accordingly. The time series variation is obtained by aggregating a large set of log-transformed high-frequency observations each day, generating one data point per trading day. Similarly, the systematic diurnal variation is captured by averaging the observations for each intraday interval across many trading days, to isolate the time-of-day effect.

Formally, we obtain daily time series through the following aggregation scheme,

$$y_d \;=\; \frac{1}{T} \cdot \sum_{t=1}^{T} \tilde{y}_{dt} \;\approx\; c + \frac{1}{T} \cdot \sum_{t=1}^{T} y_{dt}\,, \qquad \text{for} \;\; d = 1, \ldots, D, \tag{8}$$

resulting in D daily observations.

Analogously, for the intraday pattern, time-of-day observations are generated as,

$$y_t \;=\; \frac{1}{D} \cdot \sum_{d=1}^{D} \tilde{y}_{dt} \;\approx\; c + \frac{1}{D} \cdot \sum_{d=1}^{D} y_{dt}, \qquad \text{for} \;\; t = 1, \ldots, T, \tag{9}$$

which produces T separate observations across the trading period.

Finally, given the sharp differences across the separate regions, we also routinely use time series that capture the evolution over time within a given region, rather than across all three trading zones. We define the subset of intervals for a given region on day d as $\{d\} \times \mathcal{T}_r$, where $\mathcal{T}_r$, $r = 1, 2, 3$ represents the set of intervals within each regime. Hence, the associated aggregation scheme becomes,

$$(10)\quad y_{d,r} \;=\; \frac{1}{|\mathcal{T}_r|} \cdot \sum_{t \in \mathcal{T}_r} \tilde{y}_{dt} \;\approx\; c + \frac{1}{|\mathcal{T}_r|} \cdot \sum_{t \in \mathcal{T}_r} y_{dt}\,, \qquad \text{for } d = 1, \ldots, D; \;\; r = 1, 2, 3,$$

where $|\mathcal{T}_r|$ denotes the number of elements in the set $\mathcal{T}_r$. This generates three distinct day-regime series, for a total of $D \cdot 3$ data points.

The aggregation schemes in equations (8)-(10) utilize about one thousand data points each. Thus, the law of large numbers ensures that the measurement errors are small, and the measures approximate the fluctuations in the underlying average of conditionally expected activity variables well, e.g., $y_d \approx \tilde{y}_d + c$. This sets the stage for meaningful inference exploiting standard regression techniques.

B. Regression-Based Tests for Intraday Invariance

We combine the log-linear representation of the intraday invariance hypotheses in Section II with the measurement and aggregation techniques outlined in Section V.A to obtain a robust regression-based framework for testing critical implications of the underlying theories. Specifically, for each of the activity variables, we construct the regression variables through the procedures in equations (8)-(10). For example, the time series evolution is represented through the fluctuations in the $\widetilde{n}_d$ series, or the corresponding region-specific daily developments, as given by $\widetilde{n}_{d,r}$, over $d = 1, \ldots, D$. Similarly, the trading intensity measure captures the intraday pattern through the variation of $\widetilde{n}_t$ across the intraday intervals $t = 1, \ldots, T$. Since each of these series should satisfy the invariance relationship under the null hypothesis, we will at times simplify notation by employing j as a generic index for the subscripts d, t, or (d, r) below, with $j = 1, \ldots, J$, running, as necessary, across the relevant trading days, time-of-day intervals, or day-and-region observations.

Accordingly, in terms of our aggregate activity variables, our log-linear regression specification (3) now takes the form,

$$(11)\qquad s_j \,-\, n_j \;=\; c \,+\, \beta\, q_j \,+\, u_j\,, \qquad j = 1, \ldots, J.$$

The intraday invariance hypotheses impose starkly different restrictions on equation (11), as the MDH-V implies $\beta = 1$, MDH-N predicts $\beta = 0$, and ITI stipulates $\beta = -2$. This enables us to test the theories against one another in a unified setting and compare their explanatory power through standard regression techniques.

The regression setting (11) has a number of noteworthy features. First, as already emphasized, the aggregation schemes greatly reduce the impact of measurement errors. Second, as is evident from the corresponding daily series depicted in Figures 1 and 2, the activity variables display a large degree of systematic variation across time and over the intraday cycle. The daily series fluctuate quite dramatically due to major shifts in the market environment, and the intraday pattern covers separate trading regions with huge discrepancies in the activity levels. These features help ensure a high signal-to-noise ratio for the relevant variables. Third, the separate tests for the region-specific time series ensure robustness against the impact of structural breaks occurring within a single, and potentially quite illiquid, trading zone. Fourth, trade size displays relatively little idiosyncratic variation in either dimension, alleviating lingering concerns regarding the impact of errors-in-variables for the regressor. Fifth, the regressand $s - n$ reflects a constraint implied by all three theories, so it improves the efficiency of our test under the null hypothesis. Equally importantly, it circumvents a severe multicollinearity problem, as all of the activity variables are highly correlated. For example, the time series correlation among n and q is 0.975. This renders alternative specifications featuring, say, n and q as separate regressors susceptible to imprecise and misleading inference.

VI. Empirical Evidence on Intraday Invariance

This section presents tests for intraday invariance, using the high-frequency measurement technique developed in Section V to mitigate the impact of noise and errors. Section VI.A explores the predictions in the time series dimension, consistent with the usual approach. In addition, using our comprehensive data, we check for robustness across trading regions, over shorter subsamples and, in Section VI.B, for more disaggregated series. Finally, Section VI.C provides empirical tests for invariance in the intraday dimension, which is new to the literature. Again, we explore robustness across subsamples and distinct trading regions.

A. Invariance in the Time Series Dimension

As an empirical matter, we observe large, persistent and correlated fluctuations in the activity variables across days, as illustrated in Figure 1. The log-linear invariance relations explored for daily data in Section IV point towards nonlinearities in the MDH specifications that render them incompatible with corresponding representations for intraday data, but that remains an empirical matter. In contrast, the ITI hypothesis held up well for the daily data, but this functional relationship is not consistent with temporal aggregation, if there is variability in the activity variables across the trading day. Figure 2 makes it clear that not only is such intraday variability present, but it is very pronounced. Thus, it is also an empirical question whether the ITI will apply at the high-frequency or transaction level.

As noted, the daily time series for each of the three trading regimes are distinct, with each clearly affected by its specific regional pattern of trading activity and economic news releases. In other words, we effectively have three alternative daily series available for our tests. Thus, beyond aggregating all one-minute observations for each activity variable across the full trading day, following equation (8), we also generate three (day-regime) data points per trading day by averaging separately each day across all intraday observations within the sets, $\mathcal{D} \times \mathcal{T}_r$, $r = 1, 2, 3$, according to equation (10). Furthermore, we consider the identical series, but for each calendar year separately, thus exploiting only the data $\mathcal{D}_y \times \mathcal{T}_r$, $r = 1, 2, 3$, for generating the three data points per day for year y. The caveat is that day-regime series (averaged over a singular regime) likely are noisier than the corresponding full trading day series because they contain fewer observations and, for some regimes, the liquidity is somewhat limited. This exacerbates potential measurement errors, and we may note a downward bias in the regression slope due to the errors-in-variable effect. To gauge the severity of such distortionary features, we also perform a joint test, where we include all three day-regime time series in one pooled regression, thus tripling the number of observations and generating additional time series variation in the regressor. Both effects should enhance the signal-to-noise ratio and improve inference.

To convey an initial sense of the findings, Figure 6 depicts the fit of the three separate invariance relationships through a parallel set of regressions to those provided in Figure 5, except we now include observations from each of the three regions in the left side panels. For each panel, the estimated and theoretical slope coefficient are indicated by the full drawn and dashed line, respectively.

The top panels refer to the MDH-V specification (1). The left panel illustrates a dramatic failure of this hypothesis in the pooled regression, exploiting all day-regime observations. The estimated slope is 0.31 and highly different from unity, which is the theoretical value. It is further evident that estimation for each individual region alone would generate a steeper slope. In fact, the point estimates for the Asian, European, and American trading zones are 0.52, 0.58, and 1.46, respectively, with the American zone having a sufficiently large standard error that it is insignificantly different from unity. Finally, in the right panel, the slope obtained from data averaged across the full trading day results in a slope estimate of 0.89, which is also not statistically different from unity at standard significance levels.

The middle row depicts outcomes for the MDH-N regression (2). The estimated slope coefficients are now all larger, with the Asian and European values at 0.78 and 1.00 both consistent with the theoretical slope of one, while the American zone produces a value significantly above unity at 1.54. As above, the pooled regression, including all data points across the three regions, generates an excessively flat slope of 0.44. Finally, if we average all one-minute observations across the trading day, we obtain an estimate well above unity at 1.45, which is also evident from the steep full drawn line in the right panel.

We conclude that the evidence for the basic versions of the MDH is, at best, mixed and contradictory. Ideally, all regressions would generate similar slope coefficients, but they vary sharply across trading zones. The lower values for the Asian and European regimes seem too extreme to reflect solely an errors-in-variable effect. Moreover, the observations deviate from the theoretical slopes in systematic ways, creating distinct clouds in certain regions off the estimated regression line. Yet, in all cases we do establish a significant positive slope coefficient.

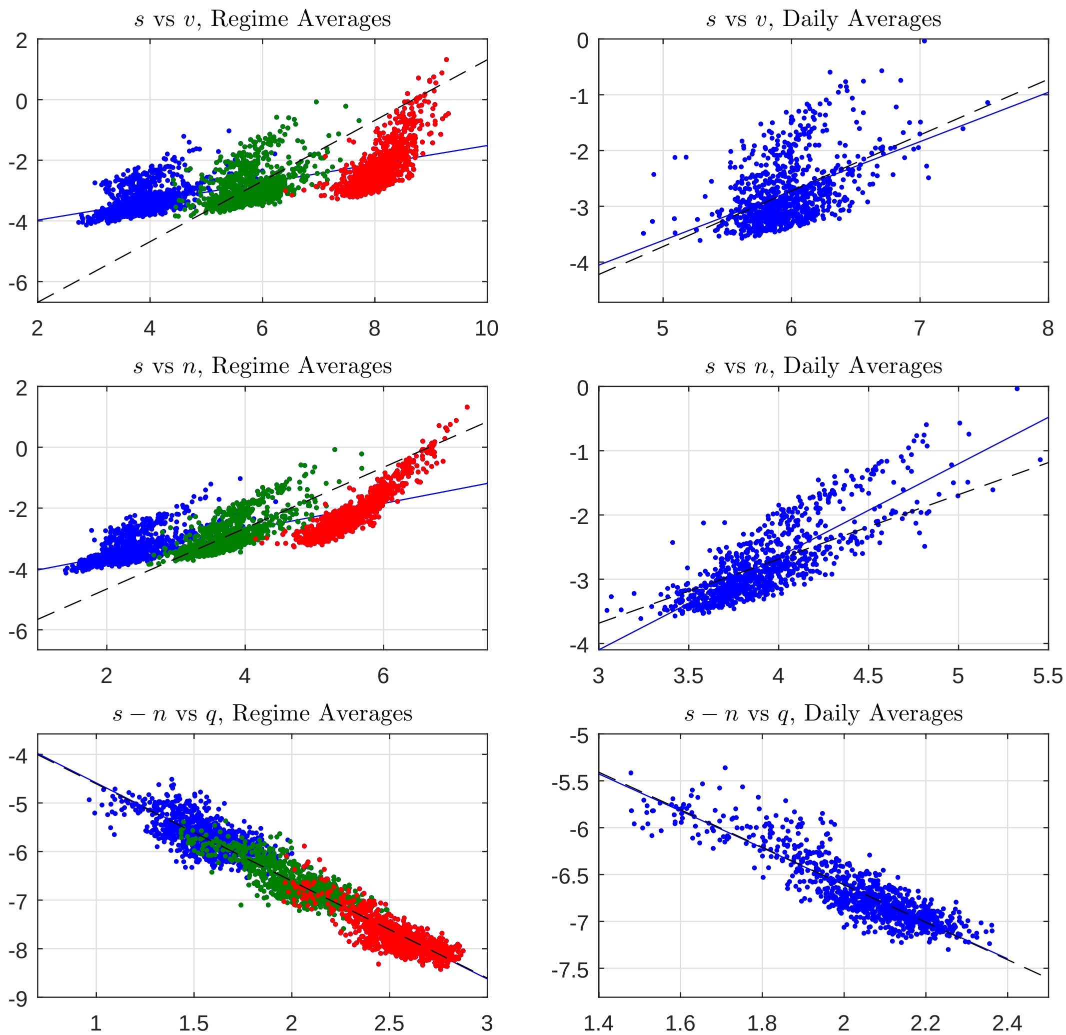

FIGURE 6. The panels reflect regressions for time series of high-frequency daily activity measures, obtained by first computing the one-minute measures, then taking the logarithm and averaging across the relevant segment of the trading day. The left panels concern measures obtained from the trading hours in the regional trading zones, while the right panels reflect the full trading day. From top to bottom, they depict observations along with fitted (full-drawn) and theoretical (dashed) regression lines for equations (1), (2), and (4).

Now, turning to the encompassing test for the invariance hypotheses, the bottom panels present evidence based on regression (3). The outcome is striking. The regression slopes are all negative and close to the ITI value of -2. In particular, contrary to the panels in the two top rows, the findings are near uniform across all regional trading zones. Moreover, for the daily averaged observations depicted in the bottom row of the right panel, the slope is estimated at -1.97 and is indistinguishable from -2. Details regarding the estimates for the encompassing regressions displayed in the left panels are provided in Panel A of Table 2.

TABLE 2—TIME-SERIES REGRESSION OF $s - n$ ONTO q

Panel A: 3 Regimes

	Nobs	c	β	se(c)	se(β)	$\bar{R}^2$
Asia	899	-3.53	-1.37	0.313	0.193	0.403
Europe	899	-3.26	-1.67	0.258	0.116	0.677
America	899	-3.15	-1.80	0.467	0.177	0.692
Pooled	2697	-2.57	-2.02	0.112	0.051	0.926

Panel B: 4 Years

	Nobs	c	β	se(c)	se(β)	$\bar{R}^2$
2008	708	-2.80	-1.93	0.098	0.043	0.939
2009	723	-2.22	-2.14	0.131	0.059	0.943
2010	711	-2.26	-2.17	0.154	0.084	0.927
2011	555	-3.21	-1.76	0.151	0.077	0.930
All	2697	-2.57	-2.02	0.112	0.051	0.926

This table reports on the OLS regressions $s_{d,r} - n_{d,r} = c + \beta \cdot q_{d,r} + u_{d,r}$. For each day, there are three observations corresponding to the regimes $r = 1, 2, 3$. In Panel A, the coefficients, standard errors, and $\bar{R}^2$ statistics are estimated separately for each regime. In Panel B, the regressions are estimated separately for each calendar year. The last year ends September 30, 2011. MDH-V, MDH-N and ITI predict $\beta = 1$, $\beta = 0$, and $\beta = -2$, respectively.

For the regime-specific regressions covering Asia and Europe, the slopes are less steep at -1.37 and -1.67, consistent with a lower signal-to-noise ratio stemming from errors-in-variables in the regressor, but for the American zone and the pooled regression, exploiting the day-regime observations across all regimes jointly, we obtain slopes that are indistinguishable from -2. The latter scenario, in particular, alleviates the errors-in-variable problem greatly and boosts the signal-to-noise ratio. Hence, the findings are very well aligned with the ITI hypothesis.

The greater variation of the regressors in the pooled regressions suggest that we may obtain sensible results also from shorter annual samples. Panel B of Table 2 reports on the findings from such annual subsample estimation. The slope coefficients are remarkably stable across time and all close to -2, except for the slightly lower (absolute) value obtained over the abbreviated final year of the sample, 2011. Thus, even for these moderate sample sizes, we find strong supportive evidence for the ITI hypothesis. In contrast, we emphatically reject the MDH variants which, if anything, would imply a positive slope coefficients for these regressions.

B. Invariance for Multi-Minute Bin Aggregation

Our empirical tests for intraday trading invariance depend critically on high-frequency measurement of trading intensity and return volatility. The minute-by-minute measures are invariably noisy. So far, we have performed aggregation across a large number of intraday observations to obtain time series comprised of reasonably precise daily measures. This approach involves a compromise where we sacrifice some systematic variation in the data along specific dimensions in order to improve the precision of the variables and, especially, the regressor.

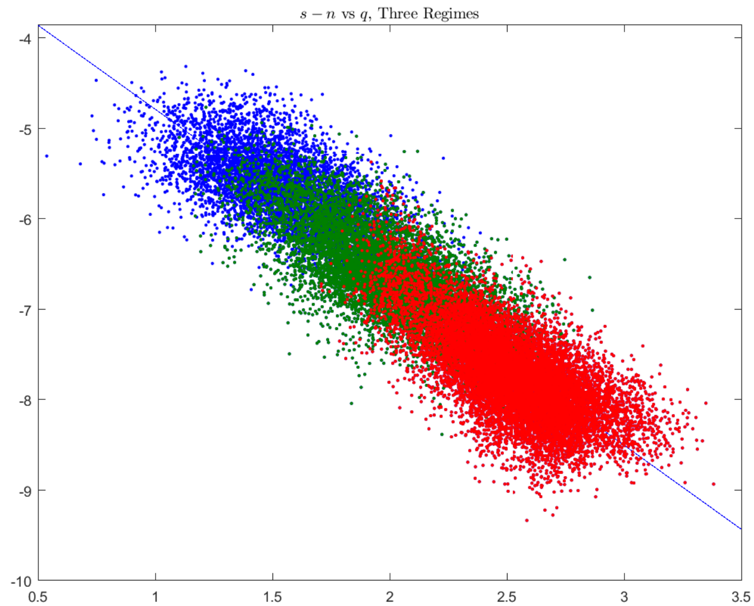

FIGURE 7. SCATTER PLOT OF $s_{d,b} - n_{d,b}$ VERSUS $q_{d,b}$ WITH DATA AGGREGATED ACROSS THE ONE-MINUTE OBSERVATIONS WITHIN BIN b FOR EACH TRADING DAY d. AS BEFORE, THE THREE REGIMES ARE REPRESENTED BY DISTINCT COLORS.

In this section, we seek to retain more of the high-frequency variation in the activity variables, while accepting that the individual measures become noisier. Specifically, we aggregate the one-minute observations into 31 separate bins per day. Each bin contains 90 consecutive minutes during the Asian regime, 39 minutes in the European regime, and 27 minutes in the American regime. This binning approach allows the measures to reflect both the intraday variation in trading activity across time bins and the time-series variation across days. The procedure generates 6, 10 and 15 day-bin observations for each day across the three regimes, respectively. Furthermore, it ensures that the regimes are relatively homogeneous in terms of the magnitude of the sampling and measurement errors. In total, we have nearly 30,000 day-bin ($D \times 31$) observations for the relevant activity variables, i.e., 31 observations for each of the variables over $D = 899$ days.[17]

Figure 7 depicts the scatter plot of $s_{d,b} - n_{d,b}$ versus $q_{d,b}$ for all day-bin observations along with the regression line obtained from the joint fit to all the data points. Consistent with the presence of non-trivial measurement errors, the points are now much more widely scattered around the regression line compared with Figure 6.

[17]A few observations are lost due to market closures or failures in the dissemination of data.

Table 3—Time-series Regression of $s - n$ onto q using binned Data

Panel A: 3 Regimes

	Nobs	c	β	se(c)	se(β)	$\bar{R}^2$
Asia	5394	-4.11	-0.99	0.191	0.098	0.348
Europe	8983	-3.87	-1.37	0.312	0.132	0.523
America	13473	-4.23	-1.38	0.515	0.187	0.445

Panel B: 4 Years

	Nobs	c	β	se(c)	se(β)	$\bar{R}^2$
2008	7303	-3.13	-1.78	0.089	0.032	0.858
2009	7465	-2.55	-1.98	0.106	0.040	0.847
2010	7347	-2.86	-1.91	0.111	0.063	0.782
2011	5735	-3.43	-1.67	0.123	0.056	0.806
All	27850	-2.93	-1.86	0.105	0.042	0.827

The table reports on OLS regressions $s_{d,b} - n_{d,b} = c + \beta \cdot q_{d,b} + u_{d,b}$ across days d and bins b. In Panel A, the coefficients, standard errors, and $\bar{R}^2$ statistics are estimated separately for each regime. In Panel B, the regressions are estimated separately for each calendar year. The last year ends on September 30, 2011. MDH-V, MDH-N and invariance predict $\beta = 1$, $\beta = 0$, and $\beta = -2$, respectively.

Nonetheless, as reported in Table 3, the slope estimate equals -1.84, which is close to the theoretical value of -2 implied by invariance. Likewise, for the individual years, the corresponding regression slopes attain a similar magnitude. In other words, the large degree of variation in the regressor across the bins may be sufficient to alleviate the severity of the errors-in-variable bias, both for the joint and the year-by-year regressions. Also, as expected with these more noisy measures, the regime-specific regressions have a less negative slope relative to earlier, consistent with a greater bias due to the lower signal-to-noise ratio.

C. Invariance Across the Intraday Activity Pattern

The systematic variation in the activity variables across the intraday pattern, documented in Figure 2 and Table 1, is an order of magnitude larger than the time series variation at the daily level depicted in Figure 1. As such, the daily activity cycle will provide a powerful test of the alternative trading invariance hypotheses. Moreover, it is a challenging test due to the heterogeneity in the market environment across the trading day. One, there are distinct spikes in activity at specific times associated with the release of macroeconomic news. Two, the opening and closing of related markets around the globe over the 24-hour cycle leads to a great diversity in the participation rate of different investor types and associated trading strategies over the trading day. A prime example is the arbitrage activities between the E-mini contract and the SPDR S&P 500 exchange traded fund (the "spider" or SPY), which is active only during regular trading in the U.S. equity markets. Three, and related, there are dramatic shifts in activity, when the majority of trading moves from one regional zone to another. This implies that a rotating cast of agents and institutions set the pace of the market over the course of the daily trading period. Four, these features naturally lead to huge discrepancies in the liquidity of the market among the different trading zones. All these factors render the proposition that the interaction of the activity variables, every minute of the day, satisfy the identical simple and uniform relationship quite heroic. Furthermore, to the best of our knowledge, this is the first time that the invariance hypotheses have been explored formally along this dimension.

The intraday test is based on regression (9) with, as before, the basic MDH-V implying $\beta = 1$, the MDH-N stipulating $\beta = 0$, and the ITI predicting a strong negative association, $\beta = -2$. Consequently, the systematic intraday pattern provides an equally clean test for the relative explanatory power of the different invariance hypotheses as those based on the time series behavior, and, given the huge variation across the daily activity cycle, it is arguably more powerful.

The regression results are presented in Figure 8 and Table 4.

TABLE 4—INDRADAY REGRESSION OF $s - n$ ONTO q

	Nobs	c	β	se(c)	se(β)	$\bar{R}^2$
Unfiltered	1335	-2.57	-2.02	0.139	0.074	0.966
Filtered	1273	-2.55	-2.03	0.105	0.051	0.979

This table reports on the Intraday OLS regression, $s_t - n_t = c + \beta \cdot q_t + u_t$. MDH-V, MDH-N and ITI predict $\beta = 1$, $\beta = 0$, and $\beta = -2$, respectively.

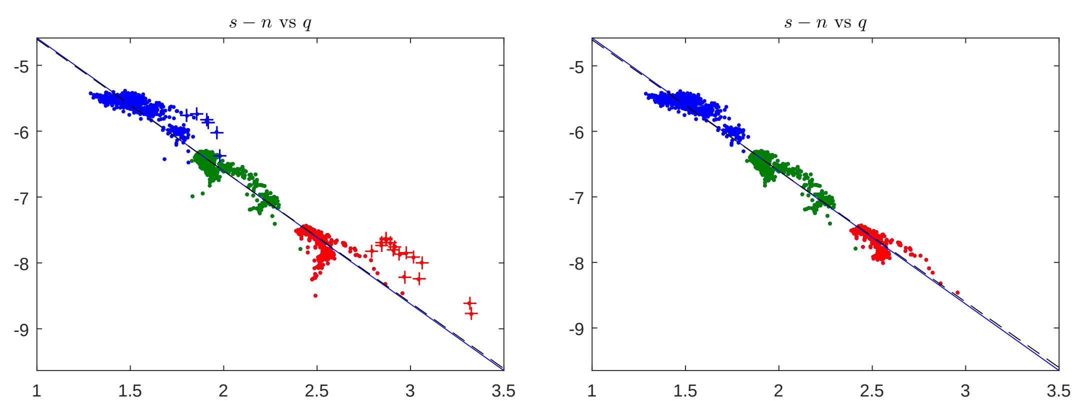

FIGURE 8. LEFT PANEL: INTRADAY SCATTER PLOT OF $s_t - n_t$ VERSUS q_t. THE OLS REGRESSION LINE IS SOLID AND THE PREDICTED INVARIANCE LINE IS DASHED. THE BLUE CROSSES INDICATE THE FIRST 6 MINUTES OF TRADING AND THE RED CROSSES THE LAST 16 MINUTES OF TRADING. THE RIGHT PANEL PRESENTS THE SAME SCATTER PLOT AS IN THE LEFT PANEL, EXCEPT THAT THE MINUTES AROUND THE REGIONAL MARKET OPENINGS AND CLOSES ARE REMOVED.

The top row of Table 4 offers strong support for the prediction of the ITI and a stinging rebuke of the MDH theories. The slope coefficient is estimated with good precision to be -2.02. It is not statistically different from -2, while the flat or positive slopes implied by the MDH variants are overwhelmingly rejected. The left panel of Figure 8 depicts the scatter plot of the 1,335 separate one-minute observations along with the fitted and theoretical regression lines. The clustering of the data points along the steeply sloped lines is evident overall, but also within each of the three trading regimes.

Most of the outliers in the left panel of Figure 8 stem from periods during which active trading is transitioning from one trading regime to another. If we eliminate observations corresponding to 3 minutes around 1:00 and 2:00 (the beginning of European hours), the 3+30 minutes around 8:30 (the beginning of American hours and the 9:00 news announcements) as well as the period 15:00-15:15 (the end of American hours, following the closure of the cash market), then we obtain the "filtered" results depicted in the right panel of Figure 8. The intraday invariance predictions are again validated in the sense that the regression slope still equals the predicted -2. At the same time, the vast majority of the outliers are eliminated by this procedure, suggesting that intraday trading invariance offers a particularly compelling account of the interaction among the activity variables in stable market settings. The consistent results obtained across the three distinct trading zones are especially noteworthy in this regard, as they differ dramatically in trading intensity and liquidity, so the ITI relationship holds up well in very different market conditions.

There are a variety of ways to assess the robustness of these quite remarkable findings by focusing on specific subsamples. The main complication is the presence of non-trivial measurement errors in our high-frequency observations. In particular, if the regression design reduces the genuine variation of the regressor relative to the size of the measurement errors, we will find, mechanically, that the regressions have less explanatory power, and we may observe a distinct downward bias in the (absolute) size of the estimated regression slope due to the errors-in-variables effect. With these caveats in mind, we first investigate the performance of the invariance hypothesis over annual subsamples. Due to the stability of the diurnal pattern for the activity variables, this approach retains the pronounced variation in the regressor, so the primary effect is simply a loss of observations.

Panel A of Table 5 reports the results for the year-by-year log-linear regressions. The individual regression slopes are now scattered more widely, but there is no evidence of any systematic deviation from the theoretical value of -2 implied by the ITI hypothesis.

An alternative robustness check involves regime-by-regime regressions. This approach does reduce the sample variation of the regressor substantially, raising concerns about the impact of measurement error. Panel B of Table 5 reports the regression results for all observations within the individual regimes, while Panel C reports the results when certain observations associated with the transition periods between regimes are excluded, as in Table 4.

In Panel B, we now observe a clear drop in the (negative) slope coefficients. Panel

Table 5—Intraday Regression of $s - n$ onto q

Panel A: 4 Years

	Nobs	c	β	se(c)	se(β)	$\bar{R}^2$
2008	1335	-2.87	-1.89	0.122	0.068	0.956
2009	1335	-2.21	-2.14	0.135	0.075	0.960
2010	1335	-2.25	-2.17	0.180	0.094	0.933
2011	1335	-3.23	-1.75	0.141	0.069	0.936
All	1335	-2.57	-2.02	0.139	0.074	0.966

Panel B: 3 Regimes, Unfiltered

	Nobs	c	β	se(c)	se(β)	$\bar{R}^2$
Asia	540	-3.87	-1.15	0.285	0.191	0.679
Europe	390	-3.18	-1.72	0.402	0.201	0.781
America	405	-5.06	-1.05	0.623	0.243	0.346
Pooled	1335	-2.57	-2.02	0.139	0.074	0.966

Panel C: 3 Regimes, Filtered

	Nobs	c	β	se(c)	se(β)	$\bar{R}^2$
Asia	526	-3.81	-1.19	0.286	0.190	0.710
Europe	388	-3.12	-1.74	0.387	0.195	0.806
America	359	-3.21	-1.78	0.917	0.374	0.607
Pooled	1273	-2.55	-2.03	0.105	0.051	0.979

This table reports on OLS regressions of the form $s_t - n_t = c + \beta \cdot q_t + u_t$. In Panel A, the coefficients, standard errors, and $\bar{R}^2$ statistics are estimated separately for each calendar year and then for the whole sample. The last year ends September 30, 2011. In Panels B and C, the regressions are estimated separately for each regime, except that in Panel C the minutes around the regional market openings and closures are removed. MDH-V, MDH-N and ITI predict $\beta = 1$, $\beta = 0$, and $\beta = -2$, respectively.

C shows that filtering improves matters considerably, but still leaves a non-trivial gap between the estimated slopes and -2. The discrepancy is especially striking for the Asian regime, where the diurnal pattern, and thus the systematic intraday variation of the regressor, is less pronounced compared to the other regions, and market activity generally is subdued.

As a final robustness check, we test the invariance relationship within each quarter of our sample. Figure 9 displays the slope coefficient for the q_t variable across quarters. It is evident that the slope coefficient is economically close to the theoretical value of -2 in all instances.[18]

In summary, our findings are broadly consistent with the ITI principle, while the MDH specifications are decidedly rejected. This implies, in particular, that there

[18]The confidence bands are obtained assuming we only face sampling error, but our construction of expectation proxies through the MEM approach inevitably introduces non-trivial measurement errors into the specification. While it is difficult to quantify this effect, it is clear that the confidence bands are too narrow and should be viewed with some skepticism.

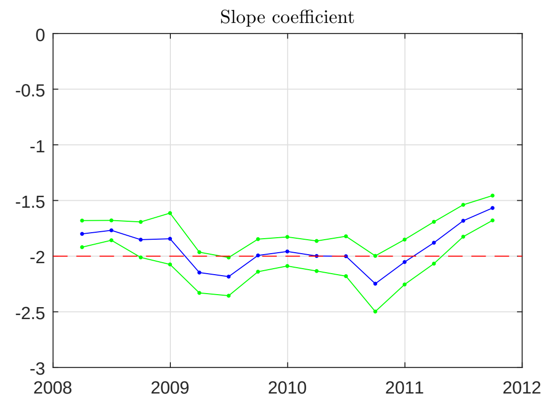

FIGURE 9. The figure plots the slope coefficient β for OLS regressions $s_t - n_t = c + \beta \cdot q_t + u_t$ estimated separately for each quarter. Also shown are confidence intervals computed as ± 2 standard errors. The red dashed line indicates the theoretical value predicted by invariance.

is strong support for the hypothesis that market participants actively adjust their trade size in response to varying market conditions in a manner consistent with the intraday trading invariance specification. Nonetheless, it is also evident that temporary deviations from the predicted relationships occur when the trading environment is changing rapidly. We present additional evidence on the performance of the ITI hypothesis during transition periods and turbulent market conditions in Section VII below.

VII. Invariance during Market Transitions

Certain events corresponding to particularly turbulent market conditions generate large outliers. In this section, we explore whether the intraday invariance principle can accommodate these dramatic episodes of elevated market activity. Specifically, we study the ITI relationship during macroeconomic announcements, during a so-called flash crash, and for the cash market open and close.

A. Macroeconomic Announcements

We now examine whether intraday invariance provides a good characterization of the market dynamics surrounding releases of macroeconomic announcements.

For the U.S. market, the most important announcements occur at 7:30. The Employment Report is usually released at 7:30 on the first Friday of the month and the Consumer Price Index at 7:30 on the second Friday of the month. Other releases at this time of day include the Producer Price Index, the Employment Cost Index, the U.S. Import/Export Price Indices, and Real Earnings. Market activity variables often exhibit distinct spikes immediately after such releases. Rationalizing the interaction among the trading variables and return volatility in the periods surrounding these events constitutes a challenging test for intraday invariance.

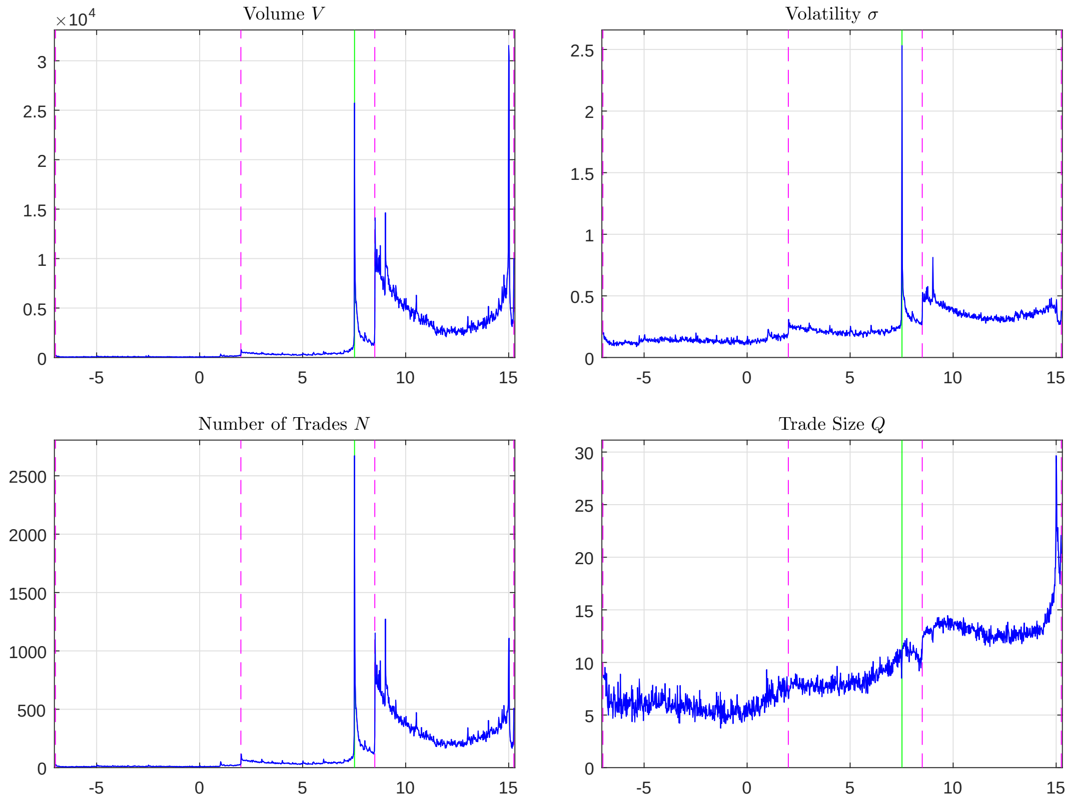

FIGURE 10. THE FIGURE DEPICTS AVERAGES ACROSS TRADING DAYS FOR WHICH AN ASSUMED 7:30 CT ANNOUNCEMENT INDUCED THE LARGEST SUBSEQUENT INCREASE IN TRADING ACTIVITY. THE STATISTICS INCLUDE CONTRACT VOLUME V_t (PER MINUTE), VOLATILITY σ_t, AVERAGE TRADE SIZE Q_t, AND THE NUMBER OF TRADES. THE AVERAGES ARE COMPUTED AT THE GRANULARITY OF ONE MINUTE. THE DASHED VERTICAL LINES SEPARATE THE TRADING REGIMES. THE SOLID VERTICAL LINE INDICATES THE TIMING OF THE ANNOUNCEMENT.

Since we seek to explore the robustness of the ITI during rapid changes in the market conditions, we focus on the main outliers induced by the 7:30 macroeconomic announcements. To identify the relevant events, we focus on the days with the largest increase in trading activity immediately after the news release. Specifically, we compute the ratio of the number of trades for the 10 minutes after 7:30 to the number of trades for the 60 minutes before 7:30. We then select the ten percent of days (90 out of 899) with the highest ratio. On average, this procedure identifies about two days each month with significant spikes in trading activity. In every case, we confirm that these episodes include a 7:30 announcement. In line with the terminology of Section III.A, we denote this subset of major announcement days by $\mathcal{D}_A$. We also zoom in further on these news releases and study the three minutes before and after 7:30; we denote this subset of minutes by $\mathcal{T}_A$.

Figure 10 depicts the intraday statistics averaged across the announcement days, $\mathcal{D}_A$. Compared to the full sample statistics in Figure 2, the 7:30 spike is obviously more pronounced for volume, number of trades, and volatility. Close inspection also reveals a small, yet distinctive, downward shift in the average trade size.

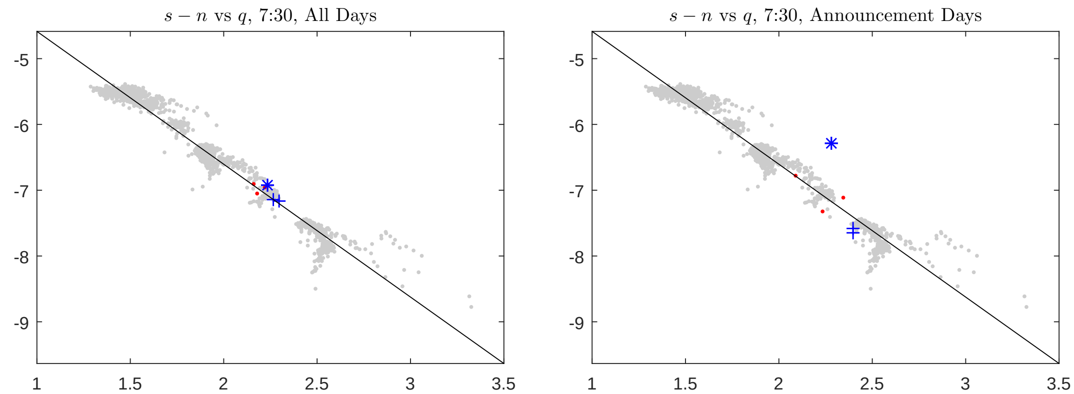

FIGURE 11. The figure provides scatter plots for $(s_t - n_t)$ versus q_t. The left and right panels plot averages across all days and 7:30 announcement days, respectively. The 3 minutes before 7:30 are indicated by red dots, the 3 minutes thereafter by a blue asterisk and two blue crosses, and all other minutes by light gray dots. The regression line from Figure 8 is also shown.

Figure 11 displays the fit for the ITI relationship, as captured by the plot of the averaged one-minute observations for volatility per transaction versus trade size. The left panel includes all trading days, while the right contains only major announcement days, $\mathcal{D}_A$. In both panels, the minutes within $\mathcal{T}_A$ are marked separately. The three minutes preceding 7:30 are represented by red dots, the 7:30-7:31 interval by the blue asterisk, and the following two minutes by blue crosses. The left panel reveals that none of the points surrounding the news release constitute unusual outliers. On announcement days on right panel, they remain in line with theoretical predictions, except for the exact one-minute interval covering the release (blue asterisk). The latter observation lies above the regression line, indicating excess volatility relative to trading volume. This is to be expected, if the price jumps with little or no trading at the point of release, consistent with the commonly observed dynamics around scheduled releases. The bid-ask spreads widen and trading stalls just prior to the release. After the number is reported, the quotes jump discretely in response to the news content, and only thereafter does an avalanche of trading hit the market. In short, the public release of news at predetermined times allows the headline number to be incorporated into quotes prior to any significant trading. Thus, modeling trading in the immediate aftermath of announcements may require an adjustment for the fraction of return volatility unrelated to order flow. Quantitatively, we find that the announcement-minute variance is about 0.7 above the fitted line. This implies that it is only about twice as large as during the same minute on normal days.

We conclude that, broadly speaking, the intraday invariance principle provides an accurate description of the interdependencies among the market activity variables, even around news releases, when the fluctuations are magnified greatly.[19]

[19] We have also confirmed that the results are qualitatively identical for the 9:00 macroeconomic news releases, which include New and Existing Home Sales, the Housing Market Index, Consumer

B. The Flash Crash

Figure 5 indicates that there is a major ITI outlier on the day of the flash crash. Hence, the relationship likely fails rather decisively at some point during that particular day. This section illustrates that ITI, indeed, may provide a useful lens through which to assess the stability of the market during the chaotic developments transpiring that day. Generating this type of insight should help us better understand the dynamics behind other episodes of extreme market behavior.[20]

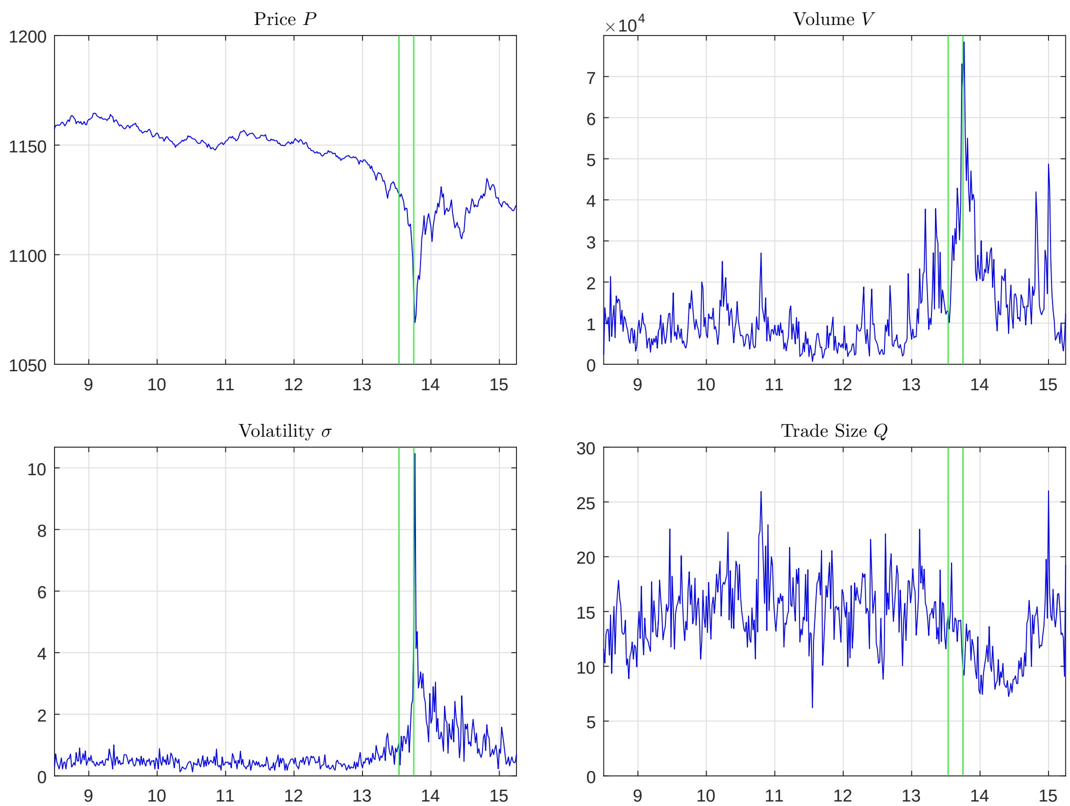

FIGURE 12. THE FIGURE SHOWS PRICE P, VOLUME V, VOLATILITY σ, AND TRADE SIZE Q ON MAY 6, 2010. THE STATISTICS ARE COMPUTED AT A GRANULARITY OF ONE MINUTE. THE SOLID VERTICAL LINES INDICATE THE TIMING OF THE FLASH CRASH.

The top left panel of Figure 12 depicts the price dynamics of the S&P 500 futures on May 6, 2010. The vertical lines indicate the timing of the crash from 13:32 to 13:45 CT, as identified in the joint report by the CFTC and SEC (2010a,b). In the morning, the market declined by about 3 percent amid rumors of a debt default by Greece, elections in the U.K., and an upcoming jobs report in the U.S. From

Sentiment, and Business Inventories, among others. These results are available upon request.

[20]Kyle and Obizhaeva (2016*b*) include the flash crash in their discussion of market crashes associated with the execution of large bets. Andersen and Bondarenko (2014, 2015) discuss alternative early warning signs for market turbulence.

13:40 to 13:45, prices first plummeted by 5.12% and then recovered by 5% over the next ten minutes, after a pre-programmed circuit breaker within the CME Globex platform halted trading for five seconds. Prices entered a free fall only during the last minute of the event window. The crash was accompanied by record trading volume and extreme realized volatility, as seen in Figure 12. This apparent breakdown in the provision of market liquidity is obviously of great interest.

We now examine whether an ITI-type relationship was in effect during the crash. Since the event lasted for less than an hour, the prior approach of averaging across a large number of observations is infeasible, and we focus instead on a stricter implication of the underlying hypothesis. We assume that the invariant, $\log I_{dt}$, is identically and independently distributed, as stipulated in Section II.C, either in terms of the raw statistic or after standardization for a time-of-day effect. We explore these propositions by computing the average $\log I_{dt}$ value over non-overlapping four-minute intervals.[21] Furthermore, we normalize them to obtain a zero mean and unit variance version of the sequence, using observations for $\log I_{dt}$ across all days at the same point in time. We then simply assess the timing and extent of ITI deviations by inspecting these intraday series for outliers.

The resulting realized values for the raw and standardized trading invariant $\log I_{dt}$ are displayed in Figure 13. For the standardized statistics on the right, we observe three consecutive large outliers—exceeding 3 standard deviations each—covering the minute including the five-second trading halt at the bottom of the flash crash and the following eleven minutes.[22] Thus, for the interval including the brief trading halt and the subsequent two intervals, we find that the realizations of $\log I_{dt}$, compared with the identical interval for the other days in the sample, represent 100th, 99.9th, and 99.4th percentile events. In comparison, for the four-minute interval ranging from five to one minute before prices bottomed out, the value of $\log I_{dt}$ corresponds to a 78th percentile event. In fact, prior to the three large outliers, there were no observations of $\log I_{dt}$ exceeding ± 2 standard deviations. In summary, the fluctuations in $\log I_{dt}$ were contained within the typical range both prior to the crash and during the time when prices dived, but it attained extreme positive values at the exact time when the market collapsed and then, after the five-second market break, as prices recovered swiftly.

How do we relate the breakdown in the ITI relationship, suggested by the consecutive outliers, to hypotheses regarding the cause of the collapse in liquidity provision?

Menkveld and Yueshen (2015) document that the co-integrating relationship between the E-mini futures and the SPY ETF broke down one minute prior to the trading halt and resumed about eight minutes afterwards. The three outliers in

[21] The four-minute block strikes a compromise between short blocks that adapt well to rapidly changing market dynamics and long blocks that facilitate averaging out measurement errors.

[22] At 1:45:28 CT, the CME Stop Logic Functionality was triggered and the trading was paused for five seconds to prevent a cascade of further price declines. When trading resumed at 1:45:33, prices stabilized and shorty thereafter the E-mini began to recover.

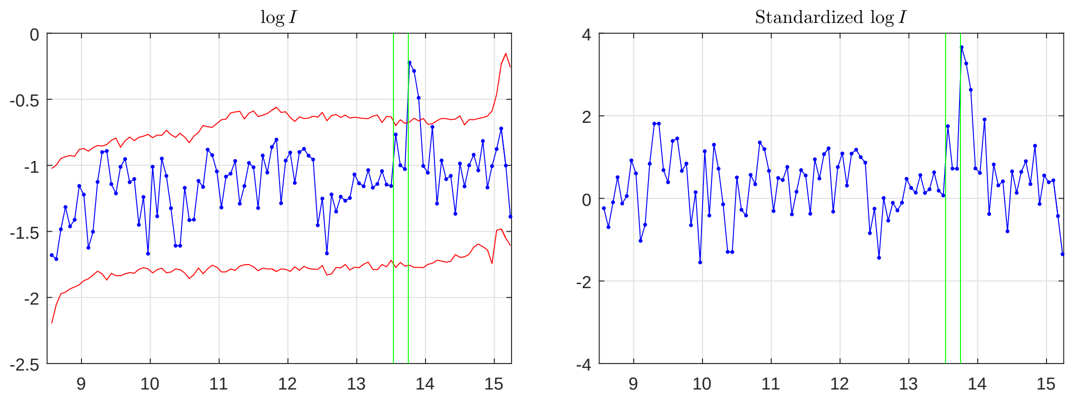

FIGURE 13. FOR MAY 6, 2010, THIS FIGURE DEPICTS THE LOG INVARIANT (LEFT PANEL) AND THE LOG INVARIANT STANDARDIZED TO HAVE ZERO MEAN AND UNIT VARIANCE (RIGHT PANEL) AT A 4-MINUTE FREQUENCY. THE RED LINES IN THE LEFT PANEL REPRESENT ± 2 STANDARD DEVIATION BOUNDS COMPUTED USING OBSERVATIONS ACROSS ALL DAYS. THE SOLID VERTICAL LINES INDICATE THE TIMING OF THE FLASH CRASH.

$\log I_{dt}$ coincide with this diagnosis of market disruption. Several factors hampered implementation of cross-market arbitrage. Margin requirements rose rapidly in response to the spike in volatility, there were problems with connectivity across trading platforms, and there was a general sense of confusion among market participants.

Kirilenko et al. (2014) report (their Table II) that traders representing institutional investors (Fundamental Buyers and Sellers) execute larger average trades than those who intermediate the trading process. In the last few minutes prior to the nadir of the crash, the participation rate of high-frequency traders increased, but then dropped immediately as prices bottomed out (their Figure 6).

Combined, these findings suggest that the breakdown in arbitrage relationships, occurring just as prices bottomed out, is associated with limited participation by market makers and arbitrageurs. Since these intermediaries tend to execute small orders, the average trade size rises, leading to unusually large realizations for the invariant, $\log I_{dt}$. Hence, we conjecture the breakdown in the ITI relationship reflects a sharp decline in intermediation relative to normal conditions. Notice that this does not imply a breakdown in the invariance relationship based on (large) bets, discussed by Kyle and Obizhaeva (2016*a*) and Kyle and Obizhaeva (2016*b*), illustrating the difference between their theory and our high-frequency ITI proposition.

C. The Cash Market Open and Close

We have argued that the ITI hypothesis provides a much better approximation to the interaction among the volatility and trading processes in the S&P 500 futures market than is feasible through the MDH theories. One primary reason is the recognition that the trade size responds endogenously to the changes in volatility and trading intensity across the daily trading cycle. Nonetheless, in Section VI.C, we also noted a systematic tendency for the ITI relationship to be violated when the market transitions from one regional trading zone to another. In this section, we assess the severity of this failure. During those transition phases, does it miss qualitatively by predicting shifts in the trade size, that are opposite to the direction of the actual changes? And quantitatively, how large are the actual deviation between the predicted and observed trade sizes?

Given our basic ITI specification, we may express the implied average trade size as a function of the expected volume and volatility, both of which we may estimate as above, using the techniques developed in Section V.A. Specifically, from the ITI relationship (4), and noting $v = n + q$, we obtain the theory-implied trade size q_t^*,

$$q_t^* \;=\; c \,+\, \frac{1}{3}\,[\,v_t \,-\, s_t\,]\,. \tag{12}$$

We identify the constant c via the (moment) condition, that the average implied log trade size q_t^* matches the sample average of the realized log trade size q_t.

Figure 14 depicts the realized average log trade size q_t and the corresponding implied log trade size q_t^*. The left panel shows that implied trade size generally tracks the actual trade size very closely, confirming that ITI provides an accurate account of how the average trade size responds to shifts in volume and return volatility on a minute-by-minute basis.

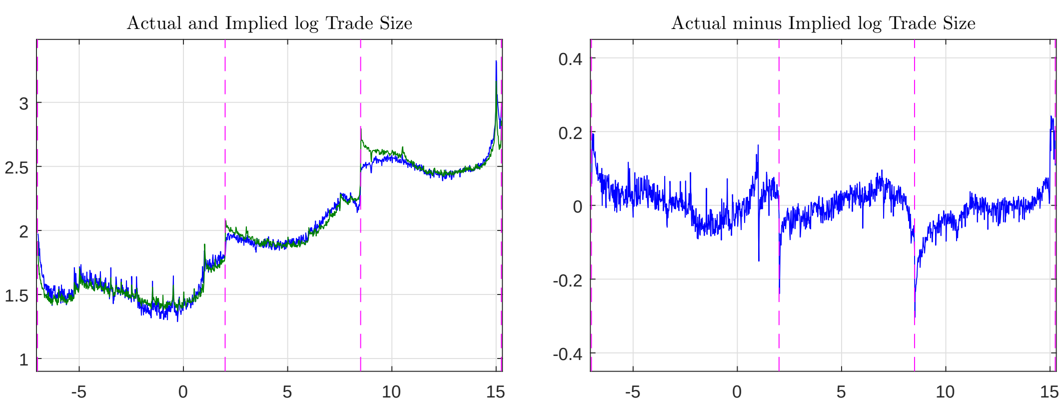

Figure 14. The left panel plots the empirical trade size q_t and the implied trade size q_t^* for each one-minute interval t, averaged across all trading days in the sample. The right panel depicts the prediction error $q_t - q_t^*$. The dashed vertical lines separate the three trading regimes.

A set of features of Figure 14 are noteworthy. First, the average log trade size follows a distinct U-shaped pattern within each regime. The most striking deviation from this pattern occurs from 1:00 to 2:00. Effectively, it represents a mixture of Asian and European trading hours, due to a couple of institutional features. One, there is no distinct opening time for European trading, as some exchanges in continental Europe open at 1:00, while the venues in London open one hour later. Two, the venues in Europe and America adhere to a summer time convention, whereas the major exchanges in Asia do not, thus further blurring the boundary between Asian and European trading hours.

Second, in spite of the dramatic spikes in the trading volume and return volatility around the release of scheduled macroeconomic announcements, there are no apparent outliers in average trade sizes at the typical times for those releases, such as 7:30 and 9:00. This is consistent with the evidence in Section VII.A, that intraday trading invariance provides an excellent approximation to the interaction among the market activity variables during announcement periods.

Third, the overall fit is excellent. The average trade size varies sixfold across the trading day. The vast majority of this variation is captured by the ITI-implied prediction involving a simple log-linear relationship based on the average trading volume and return volatility. We are unaware of any alternative theory providing comparable explanatory power for the systematic fluctuations in the transaction size across the daily trading cycle.

Fourth, we observe large discontinuities, when market activity shifts from one region to another. The trade size increases significantly when the European cash markets open, and it increases even more as trading shifts away from Europe and the U.S. cash markets open. In spite of the overall nice fit, there are clear indications that the ITI is less accurate in these instances.

The actual deviations are hard to discern from the left panel of Figure 14, as the close proximity of the predicted and observed trade sizes obscures the gap. Hence, the right panel depicts the log prediction error $q_t - q_t^*$. Most of the errors fall in the range $\pm 10\%$. Since the typical trade size is less than 10 contracts, as documented in Table 1, the average prediction errors amount to less than a single contract.

The right panel reveals the large discontinuity in trade size, when market activity shifts from one region to another. The downward spikes at 1:00, 2:00 and 8:30 represent clear outliers, revealing that the actual trade sizes are lower than the implied trade size, exactly at the time when an active regional cash market opens. Notice that ITI, correctly in qualitative terms, predicts a sharp increase in the trade size at these points but, systematically, the actual trade size jumps (quantitatively) somewhat less than predicted. Likewise, ITI predicts a sharp increase in trade size when trading transitions to a new regime, where the futures market is operating, but the cash market is closed. As before, the directional prediction of ITI is qualitatively correct but, in contrast to the case when the cash market is coming on line, the actual trade size now jumps more than predicted. These effects are visible at 15:00, as the U.S. cash market closes, and for the market opening at -7:00, i.e., 17:00 the preceding day. These abrupt shifts in trade size pose an interesting challenge for all theories of trade size.

Overall, we confirm our prior finding that the ITI provides a vastly superior approximation to the dynamics of the trade size than the MDH theories. As a striking illustration, we note that the jump in the number of transactions and volume in the minutes prior to the close of the cash market at 15:00, of the order of 300% and 200%, respectively, is accompanied by only a modest increase in return volatility; this is starkly at odds with either of the MDH specifications. In contrast, ITI accommodates the majority of this effect through a simultaneous sharp increase in the expected trade size.

In summary, at a qualitative level, the ITI hypothesis provides a truly impressive fit to the intraday dynamics of trade size At specific times, however, a systematic gap emerges between the predicted and actual trade sizes. The common trait, characterizing negative versus positive outliers, is that the cash markets are just opening for the former, while they are closed or closing for the latter. It suggests that part of the explanation stems from the different population of active traders at a given time. This reasoning is also consistent with the conjecture in Section VII.B that the malfunctions during the flash crash were accompanied by a shift in the proportion of intermediaries and arbitrageurs in the market. In other words, the actual trade size may exceed the predicted one following the U.S. close and the market opening at 17:00, because a relatively larger fraction of trades represent institutional investors. The latter execute larger trades compared to market makers and arbitrageurs, who trade frequently, but at small size.

We leave further investigation of why systematic deviations from the implied trade size occur at the open and close of active cash markets as well as, possibly, during certain longer stressful market episodes for future research.

VIII. Conclusion

The invariance properties of trading patterns in the E-mini S&P 500 are unexpected and powerful results, which raise interesting challenges for empirical and theoretical research in market microstructure. In particular, they highlight the importance of viewing the high-frequency market environment through the lens of an equilibrium involving the state of the primary activity variables. Since the discrepancy between ITI and traditional MDH representations is most readily captured through the notion of an endogenous trade size, we focused on the ability of ITI to mimic the average transaction size in Section VII.C. However, the prior literature emphasizes the correlation between market and return volatility, popularizing the notion of a stochastic clock for the price process. Thus, to most effectively contrast our findings to earlier work, we turn to this approach, exploring the ability of each hypothesis to generate a homogeneous time scale for the volatility process.

We transform the estimated return volatility according to the business time clock implied by the alternative high-frequency invariance theories. If a given representation is valid, it should produce a path for volatility bereft of systematic variation, i.e., the volatility should be evolving at a steady pace in business time. The suitable transformation may be obtained straightforwardly from the MDH-V, MDH-N, and ITI theories. Figure 15 depicts the time series for the actual (estimated) daily realized volatility in the upper left panel, while the business-time transformed and normalized log-volatility series according to the different hypotheses are provided in the three other panels. Clearly, the MDH-V and MDH-N representations merely generate a dampening of the underlying fluctuations in return volatility. They are

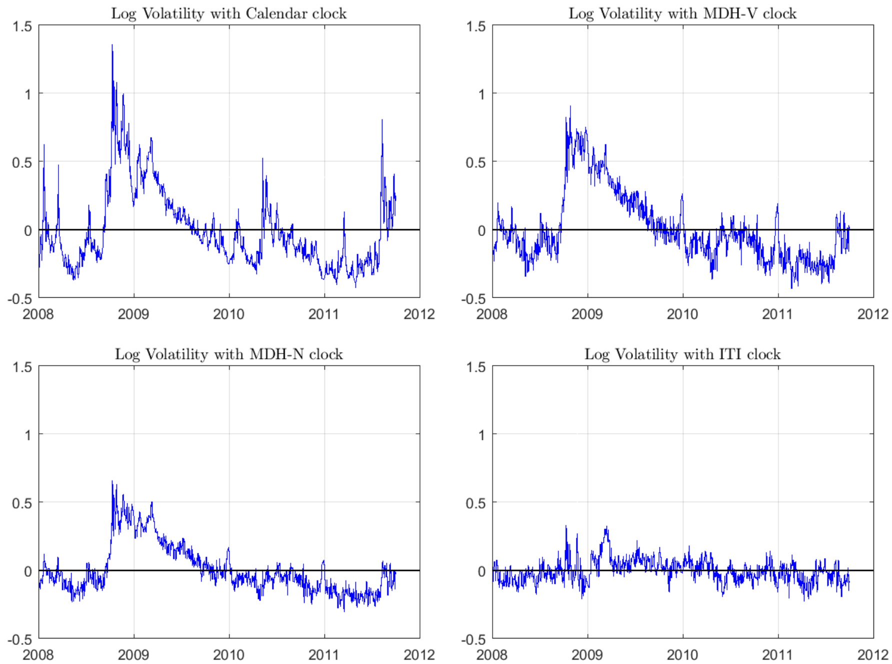

FIGURE 15. VOLATILITY COMPUTED USING DIFFERENT STOCHASTIC CLOCKS. VOLATILITY IS COMPUTED FOR EACH TRADING DAY CUMULATING 1-MIN FREQUENCY OBSERVATIONS. IT IS NORMALIZED TO HAVE MEAN 0.

not successful in eliminating the systematic variation in volatility over time, even if it is evident that the MDH-N transformation has the more meaningful impact. In contrast, the ITI panel in the bottom right panel has transformed volatility into a largely invariant series, subject solely to homoskedastic shocks. The main deviation from this pattern occurs around the financial crisis in 2008–2009, when the innovations are larger, and there is a tendency for the discrepancies to be positive, reflecting excess volatility relative to the ITI clock. Nonetheless, the improvement is dramatic relative to the implied MDH clocks.

Figure 16 applies the same transformations to the volatility across the intraday trading cycle. In this dimension, the MDH representations fail even more spectacularly. If anything, the systematic diurnal pattern in volatility are exacerbated by the MDH transformations. Clearly, the MDH failings are not linked to specific event, but reflect a glaring inability to accommodate the joint intraday variation of the activity variables in a consistent manner. In contrast, the ITI hypothesis again provides a good fit, generating a near homogenous series across the full trading cycle. Even so, minor systematic deviations are evident around the transition from one regime to another. The positive errors at time −7:00 and 15:00 CT indicate volatility exceeding the ITI prediction as the futures market is open, but related cash markets are closed. In contrast, the negative outliers at 1:00 and 2:00 CT as well as 8:30 CT correspond to times when the cash markets open in Europe and U.S., and volatility is lower than ITI would imply. These specific features were discussed at some length and motivate future research into the impact of a changing composition of active market participants during such transition points in the trading day.

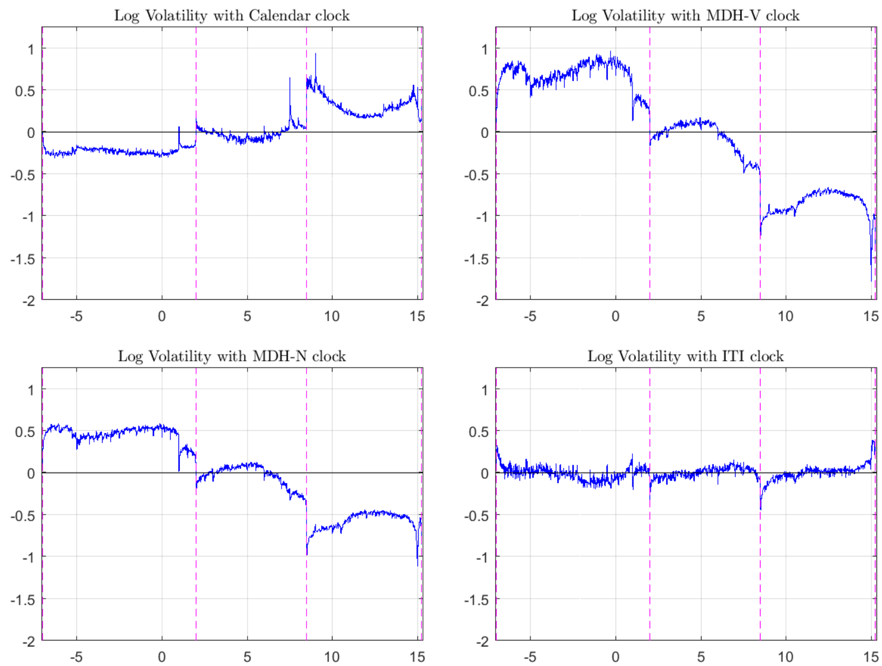

FIGURE 16. VOLATILITY COMPUTED USING DIFFERENT STOCHASTIC CLOCKS. INTRADAY VOLATILITY IS COMPUTED AT 1-MIN FREQUENCY AND AVERAGED ACROSS ALL TRADING DAYS. IT IS NORMALIZED TO HAVE MEAN 0.

Despite the obvious importance of the S&P 500 futures market—often characterized as the primary location for price discovery in U.S. equities–it is of interest to establish how universal our findings are. They could be an artifact of institutional arrangements or other unique features of the E-mini futures market. Possibilities include the large tick size, large contract size, tight integration with liquid cash markets, important role of high-frequency traders, presence of specialized automated trading algorithms, or specific features of the CME Globex matching engine. Any novel insights along these dimensions would be informative for empirical market microstructure research. On the contrary, if intraday trading invariance is a more universal phenomenon, then it speaks to deep structural issues related to how trading in financial markets operates. For example, intraday trading invariance may be a characteristic of electronic platforms on which traders shred big bets into many small orders, *and* it may also apply to non-electronic dealer markets in which bets are often executed as single trades. As such, intraday trading invariance may provide a fruitful framework for analyzing a host of issues surrounding market organization, liquidity, functional operation, general trading motives, and strategies. An empirical difficulty is the lack of a consistent recording of the transaction size for the active market participant in most trading venues due to reporting conventions and the increasing fragmentation of trading across competing venues.[23]

From a theoretical viewpoint, our findings pose a difficult conceptual question. Intraday trading invariance clearly reflects the spirit of the market microstructure invariance hypothesis, formulated for bet arrival rates and bet sizes. There are theoretical models such as Kyle and Obizhaeva (2016*a*) consistent with that hypothesis. Yet, trades are different from bets, traders usually shred bets into multiple small trades, optimizing their strategies to control transactions costs, as modeled in Kyle, Obizhaeva and Wang (2017).[24] We note, however, that Kyle and Obizhaeva (2017) suggest the combination of dimensional analysis, leverage neutrality, and invariance may help to explain theoretically how order shredding with restrictions on tick size and lot size might lead to intraday trading invariance.

In summary, the success of the invariance principle at the high-frequency level for the S&P 500 futures raises intriguing possibilities to shed new light on the intrinsic workings of modern financial markets. At a minimum, given the central position of the E-mini futures trading for the functioning of global financial markets, the issues raised pose a host of questions for future empirical and theoretical research.

REFERENCES

Admati, Anat R., and Paul Pfleiderer. 1988. "A Theory of Intraday Patterns: Volume and Price Variability." *Review of Financial Studies*, 1(1): 3–40.

Ait-Sahalia, Yacine, and Jean Jacod. 2014. *High-Frequency Financial Econometrics.* Princeton University Press.

Alexander, Gordon, and Mark Peterson. 2007. "An Analysis of Trade-Size Clustering and its Relation to Stealth Trading." *Journal of Financial Economics*, 84: 435–471.

Almgren, Robert, and Neil Chriss. 2000. "Optimal Execution of Portfolio Transactions." *Journal of Risk*, 3: 5–39.

Andersen, Torben G. 1996. "Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility." *Journal of Finance*, 1: 169–204.

Andersen, Torben G., and Oleg Bondarenko. 2014. "VPIN and the Flash Crash." *Journal of Financial Markets*, 17: 1–46.

Andersen, Torben G., and Oleg Bondarenko. 2015. "Assessing Measures of Order Flow Toxicity and Early Warning Signals for Market Turbulence." *Review of Finance*, 19(1): 1–54.

[23]Favorable evidence for invariance-style hypotheses for a wide set of futures contracts and individual stocks is reported recently in Benzaquen, Donier and Bouchaud (2016).

[24]For earlier contributions to the literature on optimal order shredding, see, e.g., Bertsimas and Lo (1998), Almgren and Chriss (2000), and Obizhaeva and Wang (2013).

Andersen, Torben G., and Tim Bollerslev. 1997. “Intraday Periodicity and Volatility Persistence in Financial Markets.” *Journal of Empirical Finance*, 4: 115–158.

Andersen, Torben G., and Tim Bollerslev. 1998. “DM-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer-Run Dependencies.” *The Journal of Finance*, 53: 219–265.

Andersen, Torben G., Tim Bollerslev, and Francis X. Diebold. 2010. “Parametric and Nonparametric Measurements of Volatility.” In *Handbook of Financial Econometrics, Volume 1 — Tools and Techniques.* , ed. Yacine Ait-Sahalia and Lars Peter Hansen, 67–137. North Holland.

Ané, Thierry, and Hélyette Geman. 2000. “Order Flow, Transaction Clock, and Normality of Asset Returns.” *Journal of Finance*, 55(5): 2259–2284.

Angel, James J. 1997. “Tick Size, Share Prices, and Stock Splits.” *Journal of Finance*, 52(2): 655–681.

Bae, Kyounghun, Albert S. Kyle, Eun Jung Lee, and Anna A. Obizhaeva. 2016. “Invariance of Buy-Sell Switching Points.” Working Paper, available at http://dx.doi.org/10.2139/ssrn.2730770.

Benzaquen, Michael, Jonathan Donier, and Jean-Philippe Bouchaud. 2016. “Unravelling the Trading Invariance Hypothesis.” Imperial College, London; Working Paper.

Bertsimas, Dimitris, and Andrew W. Lo. 1998. “Optimal Control of Liquidation Costs.” *Journal of Financial Markets*, 1: 1–50.

Bollerslev, Tim, and Dan Jubinski. 1999. “Equity Trading Volume and Volatility: Latent Information Arrivals and Common Long-Run Dependencies.” *Journal of Business & Economic Statistics*, 17: 9–21.

Brada, Josef, Harry Ernst, and John Van Tassel. 1967. “The Distribution of Stock Price Differences: Gaussian After All?” *Operations Research*, 14(2): 334–340.

Brennan, Michael, and Avanidhar Subrahmanyam. 1998. “The Determinants of Average Trade Size.” *The Journal of Business*, 71(1): 1–25.

Brownlees, Christian T., Fabrizio Cipollini, and Giampiero M. Gallo. 2011. “Intra-daily Volume Modeling and Prediction for Algorithmic Trading.” *Journal of Financial Econometrics*, 9: 489–518.

Brownlees, Christian T., Fabrizio Cipollini, and Giampiero M. Gallo. 2012. “Multiplicative Error Models.” In *Handbook of Financial Engineering and Econometrics: Volatility Models and Their Applications.* , ed. Luc Bauwens, Christian Hafner and Sebastian Laurent, 225–248. Wiley.

CFTC and SEC. 2010a. *Preliminary Findings Regarding the Market Events of May 6, 2010.* Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues. May 18, 2010. “Preliminary Report”.

CFTC and SEC. 2010b. *Findings Regarding the Market Events of May 6, 2010.* Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues. September 30, 2010. “Final Report”.

Clark, Peter K. 1973. “A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices.” *Econometrica*, 41(1): 135–155.

Engle, Robert F. 2002. “New Frontiers for ARCH Models.” *Journal of Applied Econometrics*, 17: 425–446.

Epps, Thomas W., and Mary Lee Epps. 1976. “The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture-of-Distributions Hypothesis.” *Econometrica*, 44(2): 305–321.

Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, and Eugene H. Stanley. 2003. “A Theory of Power-Law Distributions in Financial Market Fluctuations.” *Nature*, 423: 267–270.

Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, and Eugene H. Stanley. 2006. “Institutional Investors and Stock Market Volatility.” *The Quarterly Journal of Economics*, 121: 461–504.

Gillemot, László, J. Doyne Farmer, and Fabrizio Lillo. 2006. “There’s More to Volatility Than Volume.” *Quantitative Finance*, 6: 371–384.

Goldstein, Michael A., and Kenneth A. Kavajecz. 2000. “Eighths, Sixteenths, and Market Depth: Changes in Tick Size and Liquidity Provision on the NYSE.” *Journal of Financial Economics*, 56(1): 125–149.

Granger, Clive W.J., and Oscar Morgenstern. 1970. *Predictability of Stock Market Prices.* D.C. Heath.

Harris, Lawrence. 1987. “Transaction Data Tests of the Mixture of Distributions Hypothesis.” *Journal of Financial and Quantitative Analysis*, 22(02): 127–141.

Harris, Lawrence. 1994. “Minimum Price Variations, Discrete Bid-ask Spreads, and Quotation Sizes.” *Review of Financial Studies*, 7(1): 149–178.

Hong, Harrison, and Jiang Wang. 2000. “Trading and Returns under Periodic Market Closures.” *Journal of Finance*, 55(1): 297–354.

Jones, Charles M., Gautam Kaul, and Marc L. Lipson. 1994. “Transactions, Volume and Volatility.” *Review of Financial Studies*, 7(4): 631–651.

Kirilenko, Andrei, Albert S. Kyle, Tugkan Tuzun, and Mehrdad Samadi. 2014. "The Flash Crash: The Impact of High Frequency Trading on an Electronic Market." Working Paper, University of Maryland. Available at http://dx.doi.org/10.2139/ssrn.1686004.

Kyle, Albert S., and Anna A. Obizhaeva. 2016*a*. "Market Microstructure Invariance: Empirical Hypotheses." *Econometrica*, 84(4): 1345–1404.

Kyle, Albert S., and Anna A. Obizhaeva. 2016*b*. "Large Bets and Stock Market Crashes." Working Paper, available at http://dx.doi.org/10.2139/ssrn.2023776.

Kyle, Albert S., and Anna A. Obizhaeva. 2017. "Dimensional Analysis, Leverage Neutrality, and Market Microstructure Invariance." Working Paper, available at http://dx.doi.org/10.2139/ssrn.2785559.

Kyle, Albert S., Anna A. Obizhaeva, and Tugkan Tuzun. 2016. "Microstructure Invariance in U.S. Stock Market Trades." Working Paper, available at http://dx.doi.org/10.2139/ssrn.1107875.

Kyle, Albert S., Anna A. Obizhaeva, and Yajun Wang. 2017. "Smooth Trading with Overconfidence and Market Power."

Kyle, Albert S., Anna A. Obizhaeva, Nitish R. Sinha, and Tugkan Tuzun. 2017. "News Articles and the Equity Trading." Working Paper, available at http://dx.doi.org/10.2139/ssrn.1786124.

Liesenfeld, Roman. 2001. "A Generalized Bivariate Mixture Model for Stock Price Volatility and Trading Volume." *Journal of Econometrics*, 104: 141–178.

Mandelbrot, Benoît, and Howard M. Taylor. 1967. "On the Distribution of Stock Price Differences." *Operations Research*, 15(6): 1057–1062.

Menkveld, Albert J., and Bart Z. Yueshen. 2015. "The Flash Crash: a Cautionary Tale about Highly Fragmented Markets." Working Paper, available at http://dx.doi.org/10.2139/ssrn.2243520.

Morgan, I.G. 1976. "Stock Prices and Heteroscedasticity." *Journal of Business*, 49: 496–508.

Moulton, Pamela C. 2005. "You Can't Always Get What You Want: Trade-size Clustering and Quantity Choice in Liquidity." *Journal of Financial Economics*, 78(1): 89–119.

Murphy, Anthony, and Marwan Izzeldin. 2010. "Recovering the Moments of Information Flow and the Normality of Asset Returns." *Applied Financial Economics*, 20: 761–769.

Obizhaeva, Anna A., and Jiang Wang. 2013. "Optimal Trading Strategy and Supply/Demand Dynamics." *Journal of Financial Markets*, 16(1): 1–32.

Osborne, M.F.M. 1962. "Periodic Structure in the Brownian Motion of Stock Prices." *Operations Research*, 10: 345–379.

Schultz, Paul. 2000. "Stock Splits, Tick Size, and Sponsorship." *Journal of Finance*, 55(1): 429–450.

Tauchen, George E., and Mark Pitts. 1983. "The Price Variability-Volume Relationship on Speculative Markets." *Econometrica*, 51: 485–505.

Westerfield, Randolph. 1977. "The Distribution of Common Stock Price Changes: An Application of Transaction Time and Subordinated Stochastic Models." *Journal of Financial and Quantitative Analysis*, 12: 743–765.

Wood, Robert A., Thomas H. McInish, and J. Keith Ord. 1985. "An Investigation of Transaction Data for NYSE Stocks." *The Journal of Finance*, 40(3): 723–739.

IX. Appendix

A. Inclusion of Price in an ITI-Style Representation

This section discusses the findings from a set of representative regressions that introduce the price variation into the ITI hypothesis, inspired by the presence of P in the MMI specification.

As reported previously, variation in P is negligible for the regressions in Section VI.C, capturing intraday variation. These regressions produce identical estimates for the slope coefficient, and the overall fit is unchanged. The only difference arises for regressions exploring the time series variation, as in Section VI.A. Table 6 provides results mirroring those in Table 2.

B. High-Frequency Data Included in the Analysis

The starting point is all one-minute observations available for the trading of E-mini S&P 500 futures contracts on the CME Group Globex platform. As described in Section III.A, our sample covers January 4, 2008, to September 30, 2011. We use all observations within the continuous trading session, initiated each Sunday through Thursday, and ranging from 17:00 to 15:15 CT of the following day. Due to a few abbreviated trading sessions, light holiday-related trading intensity, and occasional loss of the data dissemination, there are a few gaps in our series.

To obtain a balanced dataset, we adopt a conservative filtering procedure: we drop from the sample an entire trading day if trading volume in any of the three regimes (Asian, European, or American) is very low, defined as less than 1/3 of the regime's average volume for the preceding month. Most commonly, the low trading activity happens due to holidays or abbreviated trading sessions. Overall, we discard 46 "irregular" trading days:

2008: 01/02, 02/04, 03/12-13, 03/24, 04/04, 06/12, 09/11, 11/03-04, 11/28, 12/11, 12/24, 12/26, 12/29, 12/31;

2009: 01/02, 03/12, 04/13, 06/11, 09/10, 09/14, 12/10, 12/24, 12/28-29, 12/31;

2010: 03/11, 04/05, 06/04, 06/10-11, 09/09, 11/26, 12/09, 12/10, 12/22-23, 12/27-29, 12/31;

2011: 03/10, 04/25, 06/09, 09/08.

The results are quantitatively very similar, if we adopt less aggressive filtering procedures.

C. Additional Details on the Trade Size

Figure 17 reports additional evidence on the distribution of the trade size in our sample. The heterogeneity is striking, with a large proportion of the transactions involving only a single or a couple of contracts, as well as the non-trivial occurrence of large trades involving the transfer of hundreds of contracts.

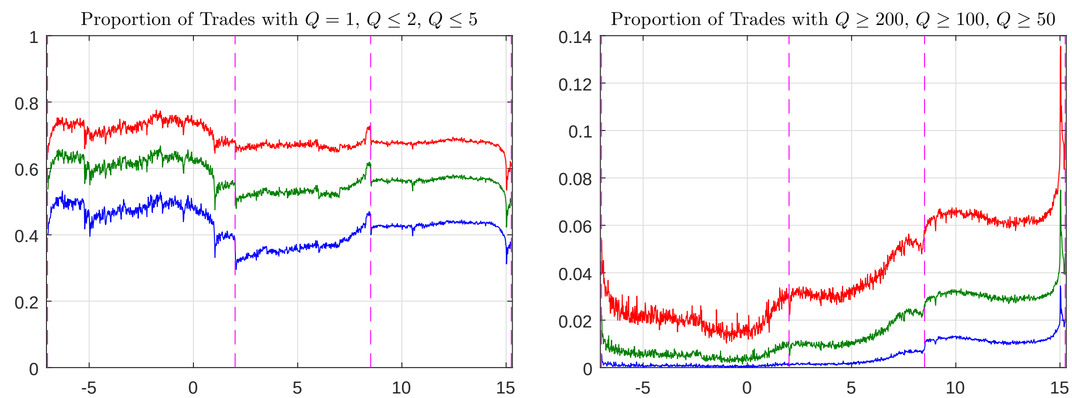

Figure 17. The figure plots the proportion of trades with sizes $Q = 1$, $Q \leq 2$, and $Q \leq 5$ contracts (left panel) and $Q \geq 200$, $Q \geq 100$, and $Q \geq 50$ contracts (right panel) in blue, green, and red, respectively. The proportions are computed at the granularity of one-minute. The dashed vertical lines separate the three trading regimes.

D. Details on the Multiplicative Error Models

As in equation 6, we let $\tilde{Y}_\tau$ denote the value of a strictly positive random variable, and stipulate that the dynamics of $\tilde{Y}_\tau$ is given by,

$$\tilde{Y}_\tau \;=\; Y_\tau \;\cdot\; \tilde{U}_\tau, \tag{13}$$

where $\tilde{U}_\tau$ is strictly positive and i.i.d. $(1, \sigma_U^2)$. Equation (13) is a MEM model with $E_{\tau-1}[\tilde{Y}_\tau] = Y_\tau$ and $Var_{\tau-1}[\tilde{Y}_\tau] = Y_\tau^2 \cdot \sigma_U^2$. The dependence structure is captured by the mean, Y_τ, which is unspecified and may incorporate complex temporal dependencies and seasonal patterns.

In this setting, we may construct simple unbiased estimators for σ_τ^2, V_τ and N_τ, subject to standard regularity conditions, as long as the interval $[\tau - \Delta t, \tau]$ is sufficiently short so that the intra-interval variation in the expected value is negligible. For volatility, we cumulate very high-frequency squared returns from within the interval τ while, for the transaction variables, we rely on the observed volume and transaction counts over the interval. Of course, sensible measures require a minimum of market activity, so we discard the interval if any of these estimators is zero. Hence, our tests involve only periods with a non-trivial degree of market activity.

Taking the logarithm of equation (13) and letting lower case letters denote the corresponding logarithmic values, so, e.g., $E_{\tau-1}[\tilde{y}_\tau] = y_\tau$, we have,

$$\tilde{y}_\tau \;=\; y_\tau \;+\; \tilde{u}_\tau, \tag{14}$$

where $\tilde{u}_\tau$ are i.i.d. random variables with mean $\mathbf{E}[\log \tilde{U}_\tau]$ and finite variance. The "innovation" term $\tilde{u}_\tau$ is not centered, but has a negative mean, due to Jensen's inequality, $E[\tilde{u}_\tau] < \log\left(E[\tilde{U}_\tau]\right) = 0$. Importantly, however, the mean of $\tilde{u}_\tau$ is constant across τ, $E[\tilde{u}_\tau] = c < 0$. Hence, equation (14) implies,

$$\tilde{y}_\tau \;=\; y_\tau \;+\; c \;+\; \tilde{\epsilon}_\tau, \tag{15}$$

where $\tilde{\epsilon}_\tau$ are i.i.d. random variables with zero mean and finite variance, while c denotes a generic constant which will differ across the activity variables.

Equation (15) provides guidance for the construction of estimators amenable for inclusion in log-linear regression-based tests. We construct unbiased estimators for the variables themselves, which typically may be obtained through suitable sample averages over a short interval. The logarithm of this sample average yields an estimator with an invariant bias which may be absorbed into the constant term in regression tests. Such estimators are, however, very noisy due to the presence of the sizable error term, $\tilde{\epsilon}_\tau$, in equation (15). Below, we pursue the strategy of mitigating the impact of measurement errors by averaging the estimator across a large number of separate intervals, thus diversifying these errors.

TABLE 6—TIME-SERIES REGRESSION OF $s - n + 2p$ ONTO q

Panel A: 3 Regimes

	Nobs	c	β	se(c)	se(β)	$\bar{R}^2$
Asia	899	9.46	-0.68	0.238	0.141	0.208
Europe	899	8.70	-0.64	0.234	0.114	0.261
America	899	7.75	-0.56	0.685	0.260	0.135
Pooled	2697	11.15	-1.85	0.171	0.084	0.861

Panel B: 4 Years

	Nobs	c	β	se(c)	se(β)	$\bar{R}^2$
2008	708	10.94	-1.69	0.307	0.134	0.839
2009	723	11.28	-2.04	0.192	0.107	0.917
2010	711	11.76	-2.15	0.164	0.092	0.910
2011	555	11.04	-1.73	0.179	0.084	0.919
All	2697	11.15	-1.85	0.171	0.084	0.861

This table reports on OLS regressions of the form, $s_{d,r} - n_{d,r} + 2p_{d,r} = c + \beta \cdot q_{d,r} + u_{d,r}$. For each day, there are three observations corresponding to the three regimes $r = 1, 2, 3$. In Panel A, the coefficients, standard errors, and $\bar{R}^2$ statistics are estimated separately for each regime. In Panel B, the data are pooled and the regressions are estimated separately for each calendar year, and then for the whole sample. The last year ends on September 30, 2011. MDH-V, MDH-N and ITI predict $\beta = 1$, $\beta = 0$, and $\beta = -2$, respectively.

First, we note that the results in Panel B, pooling the regime-day observations, generate qualitatively similar conclusions to those in Table 2. In contrast, the regime-specific regressions in the top three rows of Panel A lead to significantly lower slope coefficients than the ITI value of -2, and the explanatory power of the regression drops substantially. Hence, the inclusion of P does not appear warranted. Nonetheless, for the pooled regression, which greatly enhances the signal-to-noise ratio, the impact of the price variation is sufficiently blunted, that the ITI relationship cannot be rejected. This is, of course, consistent with the favorable results obtained for the pooled regressions in Panel B.