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: February 19, 2015

Preliminary Version: Not for Circulation

Intraday trading patterns in the E-mini S&P 500 futures contract between January 2008 and November 2011 are consistent with the following invariance relationship: The number of transactions is proportional to the $2/3$ power of the product of dollar volume and return volatility. Equivalently, the return variation per transaction is log-linearly related to trade size, with a slope coefficient of $-2$. This factor of proportionality deviates sharply from prior hypotheses relating volatility to transactions count or trading volume. Intraday trading invariance is motivated a priori by the intuition that market microstructure invariance, introduced by Kyle and Obizhaeva (2013) to explain bets at low frequencies, also applies to individual transactions at intraday frequencies.

Keywords: market microstructure, invariance, bets, high-frequency trading, liquidity, volatility, volume, business time, 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. 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
I. Introduction

An extensive literature has documented systematic intraday variation in market activity variables such as trading volume, number of transactions, and return volatility. These qualitative relationships are common across different time periods, asset classes, and market structures. It is natural to view these common variations as universal properties arising endogenously from the interaction of trading strategies. In this paper, we examine whether the number of trades, the size of individual trades, and return volatility have a quantitative relationship which remains constant both within and across days.

Our specification of a precise relationship describing variation in average trade size, transaction frequency, and return volatility is motivated by the market microstructure invariance hypothesis of Kyle and Obizhaeva (2013b). This theory is based on the intuition that markets for different financial securities operate similarly when viewed from the perspective of an appropriate security-specific “business-time” clock, which ticks at a rate proportional the arrival of low-frequency “bets” or meta-orders. Bets represent the sum of many uni-directional trades spread out across long time periods, which might be hours, days, or even weeks. The theory implies that the arrival rate of bets is proportional to the 2/3 power of the product of expected dollar volume and return volatility. Equivalently, the theory implies that the average size of bets is proportional to the cube root of the ratio of expected dollar volume to return variance. Using portfolio transition orders as proxies for bets, Kyle and Obizhaeva (2013b) show that variation in the average size of bets across stocks with different levels of trading activity conforms to the predictions of invariance.

In this paper, we examine empirically whether this invariance hypothesis can be generalized to explain the number of trades and the average size of the individual trades which are fragments of bets observed in tick data. Specifically, we hypothesize that when the market is examined over short time intervals like one minute, the number of transactions is proportional to the 2/3 power of the product of expected dollar volume and return volatility. This hypothesis is an empirically testable restriction on relating trading volume, return volatility, number of trades, and average trade size. This restriction is different from the accounting identity which states that trading volume is the product of the number of trades and average trade size. Since we measure variables at a one-minute frequency, we refer to our hypothesis as “intraday trading invariance.”

We test this hypothesis using tick-by-tick data on the E-mini S&P 500 futures contract, the most actively traded equity futures contract. The significant minute-
by-minute variation in trading activity over a 24-hour day provide statistical power for identifying patterns relating volume, volatility, and trade size. Using tick data from January 2008 to November 2011, we estimate that the number of trades varies with a minute-by-minute exponent of 0.6708, close to the predicted value of 2/3. Equivalently, we show that intraday trading invariance also predicts how the average size of individual trades increases with trading intensity but decreases with return volatility.

Intraday trading invariance is different from earlier theories describing how volume and volatility vary over time. We show that intraday trading invariance makes substantially more accurate predictions than the alternative theories of Mandelbrot and Taylor (1967) or Clark (1973), which use, respectively, the number of transactions or volume as the directing process for volatility.

We examine how intraday trading invariance holds up in times of extreme market activity. Intraday trading invariance holds up well following the release of scheduled macroeconomic announcements. During the flash crash of May 6, 2010, intraday trading invariance held up well as prices crashed but broke down as prices recovered. During the first three minutes after the market re-opened following a five-second trading halt which marked the bottom of the flash crash, average trade size was significantly larger than predicted by intraday trading invariance.

Intraday trading invariance also breaks down for a few minutes around the times when markets open and close at regularly scheduled times. Average trade size is smaller than predicted by intraday trading invariance when U.S. markets open. Average trade size is larger when U.S. markets close.

Taking together, these brief episodic breakdowns in intraday trading invariance suggest that intraday trading invariance is tied to the inherent features of properly functioning markets. Future modifications of the intraday trading invariance hypothesis might explain temporary breakdowns in intraday trading invariance during transition periods when market participants rapidly modify trading strategies. Unpredictable invariance breakdowns may serve as a contemporaneous indicator of breakdowns in the normal state of the market.

II. Invariance Hypotheses

Intraday trading invariance is based on making a careful distinction between a bet and a trade. Kyle and Obizhaeva (2013b) develop market microstructure invariance by applying an invariance principle to bets. Intraday invariance is based on applying an analogous invariance principle to trades. Intraday trading invariance is different from alternative theories relating volume, volatility, and number of trades.

In this section, we review the market microstructure invariance of Kyle and Obizhaeva (2013b) in subsection II.A, develop intraday trading invariance as an empirical hypothesis in subsection II.B, show how to test the theory during specific subperiods, such as bouts of extreme trading activity, in subsection II.D, and show
how intraday trading invariance differs from alternative theories relating volume, volatility, and number of trades inspired by previous literature in subsection II.C.

### A. Review of Market Microstructure Invariance

The market microstructure invariance hypothesis of Kyle and Obizhaeva (2013b) is based on the intuition that markets for different financial assets operate at different business-time scales. Business time passes at a rate proportional to the frequency with which bets arrive into the market. A bet represents a decision by an investor to change his risk exposure to the asset. In modern electronic markets, large bets are typically shredded into numerous smaller orders executed strategically over time to minimize price impact costs. For example, if a trader makes a decision to purchase 1,000 futures contracts and implements this decision by making 100 unidirectional purchases over a period of one hour, the first of the 100 purchases represents the arrival of the bet and the subsequent 99 purchases do not represent new bets.

Kyle and Obizhaeva (2013b) describe a theoretical model in which returns volatility results from the linear price impact of bets, noise traders bet on uninformative signals, and market makers take the other side of trades with informed and noise traders. When noise traders place more bets, trading activity increases as informed traders place more bets as well. Informed traders place bets at a rate reflecting a trade-off between the cost of acquiring private information and the expected trading profits net of market impact costs. In equilibrium, the dollar price impact cost and expected net trading profit from a bet are all proportional to the exogenous cost of private signals. This proportionality leads directly to the market microstructure invariance hypothesis, which states that the random dollar risk transferred by a bet is “invariant” when measured in units of business time.\(^1\)

To express this invariance relationship precisely, we need some notation. Let \(P\) denote the asset price (dollars per contract), let \(\tilde{Q}_B\) denote the random bet size (number of contracts), let \(\sigma^2\) denote the variance of returns expected by the market (unitless fraction per unit of time, e.g., with \(\sigma = 0.20^2 = 0.04\) per year corresponding to a standard deviation of returns of 20% per year), and let \(N_B\) denote the expected arrival rate of bets (number per unit of time). The expected bet arrival rate \(N_B\) measures the speed with which business time passes. Formally, define the “invariant” \(\bar{I}\) by

\[
\bar{I} := P \cdot \tilde{Q}_B \cdot \sigma \cdot N_B^{-1/2}.
\]

Invariance hypothesizes that randomness in bets size makes \(\bar{I}\) approximately independently and identically distributed across time and across assets.

\(^1\)Additional invariance relationships, relating trading activity to bid-ask spreads, price impact, and price efficiency also follow from the theory. We do not examine these additional invariance relationships in this paper.
Empirical analysis of invariance is complicated by the fact that business time — the bet arrival rate \(N_B\) — is not easily observable. Identifying bets requires assigning individual trades to traders and aggregating trades into bets. Kyle and Obizhaeva (2013b) circumvent this problem by invoking the auxiliary hypothesis that bet volume is proportional to the total trading volume. For expositional simplicity, we will assume that bet volume equals total volume, e.g., one side of every trade is a bet and the other side is an intermediation or market-maker trade. Thus, letting \(Q_B := E\{\tilde{Q}_B\}\) denote expected bet size and letting \(V\) denote expected trading volume (contracts per unit of time), we obtain
\[V = N_B \cdot Q_B.\]

Define expected “trading activity” \(W\) as the standard deviation of dollar mark-to-market gains or losses on all the expected trading volume in one unit of calendar time; we thus have \(W := P \cdot V \cdot \sigma\). Intuitively, the quantity \(W\) measures “total risk transfer” per unit of calendar time.

Invariance implies that the rate at which business time passes is proportional to the \(2/3\) power of trading activity \(W\). To see this, we can use \(W = P \cdot V \cdot \sigma\) and \(V = N_B \cdot Q_B\) to write
\[\tilde{I} := \frac{P \cdot \tilde{Q}_B \cdot \sigma}{N_B^{1/2}} = \frac{\tilde{Q}_B}{Q_B} \cdot \frac{P \cdot V \cdot \sigma}{\tilde{Q}_B} = \frac{Q_B}{Q_B} \cdot \frac{W}{N_B^{3/2}}.\]

Let \(I\) denote the mean of the invariant distribution, i.e., \(I := E\{\tilde{I}\}\). We obtain
\[I = \frac{P \cdot Q_B \cdot \sigma}{N_B^{1/2}} = \frac{W}{N_B^{3/2}}.\]

Since \(\tilde{I}\) has an invariant distribution, its mean \(I\) is a constant which does not depend on \(W\) or \(N_B\). It follows immediately that that \(N_B\) is proportional to \(W^{2/3}\).

The theory implies that \(I\), which is measured in dollars, is proportional to the cost of a bet.

In our empirical work, we focus on the corresponding linear representation that arises from taking the logarithm on both sides in equation (1). Define \(p := \log P\), \(\tilde{q}_B := \log \tilde{Q}_B\), \(s := \log \sigma^2\), \(n_B := \log N_B\); define \(q_B := E\{\log \tilde{Q}_B\}\). Then the invariant constant \(E\{\log \tilde{I}\}\) can be written
\[E\{\log \tilde{I}\} = p + q_B + \frac{1}{2}s - \frac{1}{2}n_B.\]

Note that the empirical content of equation (4) is slightly different from equation (3) because the means of the logs of \(\tilde{Q}\) and \(\tilde{I}\) are different from the logs of the means. Equations (3) and (4) are both implied by invariance because invariance implies that the distribution of \(\tilde{I}\) is invariant.

The hypothesis that the right-hand-side of equation (4) is constant can be expressed in different ways by placing different variables on the left-hand-side. Defining the log of trading activity as \(w := \log W = p + v + \frac{1}{2}s\), we can place \(n_B\) on the left-hand-side to write this proportionality relationship as
\[n_B = c + \frac{2}{3}w,\]
where $c$ denotes a generic invariant constant the value of which differs across equations. Since trading volume is the product of the arrival rate of bets and the expected size of bets, $v = n_B + q_B$, the same relationship can be equivalently written by placing log bet-size $q_B$ on the left-hand-side:

$$q_B = c + \frac{1}{3} \cdot w - \left[ p + \frac{1}{2} \cdot s \right].$$

Equations (5) and (6) formalize the intuition that as trading activity $w$ increases holding $s$ and $p$ fixed, $2/3$ of the increase results from increasing the arrival rate of bets (speed of business time) and $1/3$ from the magnitude of bets. This empirical relationship between the size and number of bets lies at the heart market microstructure invariance.

Several empirical hypotheses inspired by this type of invariance find support in the data. Kyle and Obizhaeva (2013) document invariance relationships for the size distributions of portfolio transition orders. Kyle, Obizhaeva and Tuzun (2012) document an invariance relationship for the size distribution of transactions in the Trade and Quote data set (TAQ). Kyle et al. (2014) study invariance relations for the number of monthly news articles. Bae et al. (2014) discuss an invariance relationship for the number of buy-sell switching points in the South Korean equity market. As noted, some of these relations require auxiliary identifying assumptions.

We now turn to an exploration of invariance for high-frequency trading patterns in the E-mini S&P 500 futures market. Our tests are not a direct implication of the underlying invariance theory. Instead, like Kyle, Obizhaeva and Tuzun (2012), they are based on extrapolating the theory from bets to individual trades.

**B. An Intraday Trading Invariance Hypothesis**

“Intraday trading invariance” is the hypothesis that the same invariance relationships which apply to bets also apply to the individual trades which add up to bets. Let $Q$ denote the average size of a trade and let $N$ denote the number of trades expected to arrive per unit of time. Since bets are the sum of many trades, intraday trading invariance can be obtained by adding to the invariance hypothesis for bets the additional hypothesis that the average number of trades per bet, which can be written as $Q_B/Q$ or $N/N_B$, is an also an invariant constant. Although we do not derive this empirical hypothesis from a theoretical model, we note that it is consistent with the intuition that the fixed cost of making a trade is constant.

Intraday trading invariance is much easier to test empirically than the corresponding invariance hypothesis for bets, for two reasons. First, while bets are difficult to observe, the size and number of trades are public information which can be observed in time-stamped tick data. Second, in today’s twenty-four-hour worldwide markets, huge fluctuations in trading volume and volatility are observed when Asian hours, European hours, and North American hours are compared. As we shall see, average trading volume in E-mini S&P 500 futures contracts increases more than fifty-fold and the transaction count increases by twenty five-fold.
when trading moves from Asian hours to U.S. hours; in contrast, volatility and trade size roughly double. The observability of trades and the huge variation in intra-day trading activity provides statistical power for detecting whether intraday fluctuations in minute-by-minute time series for average trade size $Q$ and expected number of trades $N$ bear a relationship to intraday expected returns volatility $\sigma$ which reflects the specific proportions implied by intraday trading invariance.

To test the intraday invariance hypothesis, we first define time series for average trade size, expected number of trades, price, and expected returns volatility over one-minute intraday time intervals during which expectations of these variables are not likely to be changing much. Let $d = 1, \ldots, D$ index the $D$ trading days in our sample and let $t = 1, \ldots, T$ index $T$ one-minute subintervals within each day. We extrapolate the invariance hypothesis for bets at low-frequencies in equation (1) or (4) to more readily observable intraday transaction variables at high-frequencies. The analogues to equations (1) and (4) become

$$
(7) \quad I_{dt} = P_{dt} \cdot Q_{dt} \cdot \sigma_{dt} \cdot N_{dt}^{-1/2} \quad \text{and} \quad E\{\log \tilde{I}_{dt}\} = p_{dt} + q_{dt} + \frac{1}{2} s_{dt} - \frac{1}{2} n_{dt}.
$$

This specification replaces expected bet size $Q_B$ and expected bet arrival rate $N_B$ from the invariance hypothesis for bets with the average trade size $Q_{dt}$ and expected number of transactions $N_{dt}$ for minute $t$ of date $d$. Likewise, $\sigma_{dt}$ refers to the expected volatility and $P_{dt}$ the average price over minute $t$ of date $d$. Analogously to the invariance hypothesis for bets, we conjecture that strategic interaction among market participants makes the distribution of $\tilde{I}$ identically and independently distributed across time. Thus, both $I_{dt}$ and $E\{\log \tilde{I}_{dt}\}$ are conjectured to be constants which do not vary over day $d$ or time-of-day $t$.

Equation (7) can be given a regression-style representation by writing

$$
(8) \quad n_{dt} = c + 2 \cdot q_{dt} + 2 \cdot p_{dt} + s_{dt} + u_{dt},
$$

where $u_{dt}$ indicates an uncorrelated error term. The intraday trading invariance hypothesis (7) involves only variables whose realizations are directly observable or readily approximated. Nonetheless, testing equation (8) requires dealing with significant measurement issues.

In equation (8), the variable $N_{dt}$ is not the realization of the number of trades in a particular minute and $\sigma_{dt}$ is not realized volatility. Instead, $N_{dt}$ is the market’s expectation of the speed of business time, and $\sigma_{dt}$ is the market’s expectation of volatility. Market participants choose expected trade size $Q_{dt}$ based on their trading needs and perception of the market conditions, as summarized by the their expectations of $N_{dt}$ and $\sigma_{dt}$. Volatility and transaction counts exhibit a large degree of idiosyncratic variation relative to market expectations, making their realizations differ from their a priori expectations.

To construct a test based on equation (8), we need estimators, or proxies, for the expected values. For the transaction count, trading volume, and volatility, we expect active trading participants to acquire real-time information on the state of
the market and therefore assume them to form corresponding unbiased expectations over the next minute. Hence, we treat the ex-post realizations of transaction count, trading volume, and volatility as unbiased, albeit noisy, estimators of ex-ante expectations.

To construct a test based on equation (8), we aggregate a large number of suitably transformed one-minute observations over numerous days. This alleviates the impact of sampling variation and measurement error. We can test the hypothesis of intraday trading invariance by making the identifying assumption that the large diurnal fluctuations which remain after such aggregation reflect fluctuations in the market’s expectations.

Another approach for constructing statistical tests is to aggregate across minutes of the day, allowing market expectations to show up in fluctuations in trading activity across days. Since intraday fluctuations are more pronounced than fluctuations across days, we reserve this approach for our robustness checks.

Let $\tilde{n}_{dt}$ denote the log transaction count for minute $t$ of day $d$ observed in transactions data. Define the average log transaction count variable $n_t$ for intraday interval $t$ by averaging the observations for this interval across all the days in the sample:

$$n_t = \frac{1}{D} \sum_{d=1}^{D} \tilde{n}_{dt}, \quad t = 1, \ldots, T. \tag{9}$$

Likewise, averaging across all intraday intervals on trading day $d$, we obtain

$$n_d = \frac{1}{T} \sum_{t=1}^{T} \tilde{n}_{dt}, \quad d = 1, \ldots, D. \tag{10}$$

Notice that $n_d$ is defined as the mean of the log of observed values, and not the log of the mean of observed variables. Minutes with no transactions are omitted. An alternative approach would be to take the mean of the logs. We construct $s_t, q_t, v_t, w_t,$ and $s_d, q_d, v_d, w_d$ similarly.

Using the relationships $v_{dt} = n_{dt} + q_{dt}$ and the definition $w_{dt} := p_{dt} + v_{dt} + s_{dt}/2$, equations (9) and (10) imply a regression relationship with the average of the log of the realized number of trades $n_j$ on the left side and realized log trading activity $w_j$ on the right:

$$n_j = c + \frac{2}{3} \cdot w_j + u^n_j; \tag{11}$$

here, $j$ refers to distinct intraday intervals, i.e., $j = t, \quad t = 1, \ldots, T,$ or different trading days, i.e., $j = d,$ with $d = 1, \ldots, D,$ while $u^n_j$ represents the regression residuals. Since our primary focus is to study intraday variations, we shall mostly be concerned with the relationship (11) for the former case, i.e., $j = t$.

Similarly, we can express the same invariance relationship by putting the average of log realized trade size $q_j$ on the left-side and the log of realized trading activity...
\[ q_j = c + \frac{1}{3} \cdot w_j - \left[ p + \frac{1}{2} \cdot s_j \right] + u_j^2. \]

Equations (11) and (12) represent our intraday trading invariance versions of equations (5) and (6). Holding returns variance \( s \) fixed as trading activity increases, these equations also imply that 2/3 of the increase comes from fluctuations in the transactions count and 1/3 from the size of trades. As returns variance \( s \) varies, these relationships are modified in specific ways.

The typical daily high-low range for price \( P \) is only 1%–2%, a thousand-fold less than for trading variables. In tests for systematic intraday fluctuations, fluctuation in the price level is immaterial. Furthermore, the average return is close to zero and displays no significant within-day variation. We therefore simplify the exposition by ignoring variation in the price level, folding \( p \) into the constant term \( c \).

### C. Alternative Empirical Hypotheses

There is a long history of theories describing the relationship between trading activity and return volatility, going back at least to Clark (1973). His work builds on the notion of subordination, or a stochastic business-time clock, introduced into the modeling of financial returns by Mandelbrot and Taylor (1967). Clark (1973) argues that increments to trading volume generate increments in business time, thus generating a direct proportionality between expected volume and return variation, i.e., \( \sigma_{dt}^2 \sim V_{dt} \), or equivalently, \( s_{dt} = c + v_{dt} = c + n_{dt} + q_{dt} \).

In contrast, Jones, Kaul and Lipson (1994) find the daily transactions count to be better aligned with daily volatility, and Ané and Geman (2000) assert that intraday returns become independently and identically distributed Gaussian when normalized by the (stochastic) transaction count. Tauchen and Pitts (1983), Andersen (1996), Bollerslev and Jubinski (1999), and Liesenfeld (2001) further refine and test the relationship between trading variables and return volatility. These hypotheses imply a proportionality between expected number of transactions and return variation, i.e., \( \sigma_{dt}^2 \sim N_{dt} \), or equivalently, \( s_{dt} = c + n_{dt} \).

If expected trade size is constant, intraday trading invariance (7) implies that the expected return variation is proportional to both the number of transactions and trading volume, \( \sigma_{dt}^2 \sim N_{dt} \) and \( \sigma_{dt}^2 \sim V_{dt} \). Assuming constant expected trade size, the hypothesis of Clark (1973), the hypothesis of Ané and Geman (2000), and our intraday trading invariance are all equivalent. In contrast, if changes in trade size are correlated with changes in return volatility and number of trades, equation (7) implies the number of transactions and volume are no longer proportional to return volatility. Since traders manage risk exposures actively in business time, we think it is likely that trade size varies systematically with volume and volatility. If so, the hypothesis of Clark (1973), the hypothesis of Ané and Geman (2000), and our intraday trading invariance all have quite different implications.

To illustrate these differences, we first modify equation (11) to ensure a similar 2/3 power representation different specifications. Using \( w := p + q + s/2 \), the Clark
(1973) specification takes the form

\begin{equation}
    n_j = c + \frac{2}{3} \left[ w_j - \frac{3}{2} q_j \right] + u_n^j.
\end{equation}

Likewise, the Ané and Geman (2000) formulation takes the form

\begin{equation}
    n_j = c + \frac{2}{3} \left[ w_j - q_j \right] + u_n^j.
\end{equation}

Both Clark (1973) and Ané and Geman (2000) view volatility as evolving at a uniform rate in business time governed by their respective notions of trading activity. It is thus natural to write an equivalent nested specifications for the alternative hypotheses by putting returns volatility per transaction on the left-side of a regression relationship. It is straightforward to show that all three hypotheses can be expressed in the following form:

\begin{equation}
    s_j - n_j = c + \beta \cdot q_j + u_q^j.
\end{equation}

The hypothesis of Clark (1973) implies \( \beta = 1 \), the hypothesis of Ané and Geman (2000) implies \( \beta = 0 \), and our intraday trading invariance hypothesis implies \( \beta = -2 \).

It is straightforward to explore the performance of the different specifications in this regression setting, and we do so in section IV.

D. Testing Intraday Trading Invariance During Specific Episodes

Particular market events create dramatic intraday fluctuations in trading activity and asset volatility. For example, the release of macroeconomic announcements typically induces an immediate price jump and a subsequent surge in trading volume and return volatility, see, e.g., Andersen and Bollerslev (1998). Similar extreme oscillations occur during stressful episodes associated with financial crises or market disruptions such as the “flash crash” of May 6, 2010. We can test whether the usual empirical regularities involving trading activity, return volatility, and trade size also hold during these extreme scenarios. If the market is functioning well, even during a crisis, apparently disruptive events may show up merely as a sharp elevation in the speed of the business clock.

Similarly, markets may function differently in the transition from one continent’s trading hours to another, specifically at the open and close of trading in the U.S., Europe, or Asia. Since the null hypothesis implies that the \( \tilde{I} \) variable is identically and independently distributed across all intraday intervals, we can test whether invariance relationships break down by focusing exclusively on the relevant opening or closing trading periods.

To develop notation for describing the observations involved in different scenarios, we denote the full set of trading days by \( D = \{1, \ldots, D\} \) and the set of all
intraday intervals by \( \mathcal{T} = \{1, \ldots, T\} \). Testing intraday trading invariance for the subset of days \( \mathcal{D}_k \), where a given macroeconomic announcement is released, involves observations in the set of intervals \( \{(d, t) : (d, t) \in \mathcal{O}_k = \mathcal{D}_k \times \mathcal{T}\} \). Similarly, if we seek to test intraday trading invariance only for a subset of intraday intervals \( \mathcal{T}_k \) around such announcements, then \( \mathcal{O}_k = \mathcal{D}_k \times \mathcal{T}_k \).

Before reporting results of statistical tests, we next provide a description of the data used in the empirical work.

### III. The E-mini S&P 500 Futures Data

The data we use are the best bid and 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 prices, and trade sizes, time stamped to the second. Since the contract is traded exclusively on the CME Globex electronic platform, our data contain all transactions executed during our sample, covering nearly four years from January 4, 2008, to November 4, 2011.

The procedure behind the recording of trades and trade sizes is critical for our empirical tests. When an executable order arrives into the limit order book, it is often executed against 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 of the contracts executed at the same price as a single combined transaction quantity. Thus, the trade size reflects the number of contracts traded when a marketable order crosses one or more standing limit orders at the top of the book. Classifying the transaction size from the perspective of the party placing an executable order is consistent with the motivation for intraday trading invariance.

The notional value of the E-mini S&P 500 futures contract is $50 times the value of the S&P 500 stock index 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. The E-mini contract 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 results in using the most actively traded contract throughout our analysis.

The E-mini contract trades essentially twenty-four hours a day, five days a week. From Monday to Thursday, trading is from 15:30 to 15:15 (Chicago Time) the following day, with a half-hour maintenance halt from 16:30 to 17:00. On Sunday, trading is from 17:00 to 15:15 the following Monday. We define three distinct trading regimes: 17:15 through 2:00, 2:00 through 8:30, and 8:30 through 15:15. These regimes roughly correspond to regular trading hours in Asia, Europe, and North America.

Changing the method of recording transactions can complicate testing for intraday trading invariance. For example, if the exchange instead counts each limit order involved in trading as a separate trade, the transaction count will reflect the order flow intermediated by the supply side. This would inflate the number of transactions and deflate trade size relative to the procedure used in the current paper.
America, respectively. Due to the unusually low trading volume and short trading hours around holidays, we discarded eleven days from our analysis, resulting in a total of \( D = 959 \) trading days. Each trading day is partitioned into \( T = 1,320 \) intervals of length \( \Delta t \) equal to one minute. The three trading regimes consist of \( T_1 = 525, T_2 = 390, \) and \( T_3 = 405 \) one-minute intervals, respectively, with \( T = T_1 + T_2 + T_3 \). For future convenience, we define the set of one-minute observations belonging to each trading regime using the notation established in Section II.D. Letting \( T_1 = \{1, \ldots, T_1\} \), \( T_2 = \{T_1 + 1, \ldots, T_2\} \) and \( T_3 = \{T_2 + 1, \ldots, T_3\} \), regime 1 involves the data within the set \( O_1 = D \times T_1 \), while regime 2 and 3 correspond to the sets \( O_2 = D \times T_2 \) and \( O_3 = D \times T_3 \), respectively.

Our analysis focuses on the contract volume \( V \), number of trades \( N \), average trade size \( Q \), and realized return volatility \( \sigma \) (computed from 10-second returns). The intraday variation of the four characteristics are depicted in Figure 1.

There is a striking diurnal pattern in the trading of E-mini S&P 500 futures, as illustrated by Figure 1. It depicts the share volume \( V_t \), volatility \( \sigma_t \), number of trades \( N_t \), and trade size \( Q_t \) for each minute across the trading day, where the observations are averaged across all days in the sample as in equation (9). Calendar time is displayed on the horizontal axis, with the three trading regimes separated by dashed vertical lines. The first regime spans from -6:45 (17:15 of the prior day) to 2:00 CT (Chicago Time), covering the period just before and during regular trading hours in Asia. The second regime spans from 2:00 to 8:30 CT, covering European trading hours. The third regime spans from 8:30 to 15:15 CT, covering active trading hours in North America.

The corresponding time series, obtained by averaging the intraday observations daily within each regime consistent with equation (10), are displayed in Figure 2. The series clearly display a fair degree of commonality in the dynamic features. Trading volume, volatility, and the number of trades increase substantially during the Société Générale scandal in January 2008 and the collapse of Bear Stearns in March 2008, rise once more during the financial crisis from September 2008 to February 2009, and spike during the flash crash in May 2010 as well as the second half of 2011. In contrast, the average trade size drops during the same periods. Finally, all these variables appear to be slightly subdued towards the end of each year.

Although the series exhibit common features, there are also systematic differences. In particular, as mentioned in section II.B, all variables have much lower values in regime 1 than regime 3. Table 1 provides detailed summary statistics.

Before engaging in the empirical analysis of intraday trading invariance, we provide a back-of-the-envelope assessment. Substituting the number of transactions \( N \) for the speed of business time \( N_B \) in equation (5), as discussed in section II.B, we find the ratio of the average number of trades to the \( 2/3 \) power of the product of the average dollar volume and volatility to be 15.6, 15.9 and 16.0 for the three regimes, respectively. The stability of this ratio suggests that further analysis along the lines of intraday trading invariance is warranted.
Figure 1. The figure shows averages across all days for share volume $V_t$ (per minute), annualized realized volatility $\sigma_t$, trade size $Q_t$, and number of trades, $N_t$. The averages are computed at a granularity of $\Delta t = 1$-minute. The dashed vertical lines separate the three trading regimes corresponding to trading hours in Asia, Europe, and North America.

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

<table>
<thead>
<tr>
<th></th>
<th>Regime 1</th>
<th>Regime 2</th>
<th>Regime 3</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volatility</td>
<td>0.16</td>
<td>0.25</td>
<td>0.40</td>
<td>0.26</td>
</tr>
<tr>
<td>Volume</td>
<td>92.01</td>
<td>600.73</td>
<td>4725.56</td>
<td>1663.97</td>
</tr>
<tr>
<td># Trades</td>
<td>13.87</td>
<td>66.74</td>
<td>360.04</td>
<td>135.70</td>
</tr>
<tr>
<td>Notional Value, $$\text{Mln}$</td>
<td>5.25</td>
<td>34.37</td>
<td>265.84</td>
<td>94.02</td>
</tr>
<tr>
<td>Market Depth</td>
<td>54.06</td>
<td>265.12</td>
<td>984.07</td>
<td>401.76</td>
</tr>
<tr>
<td>Bid-Ask Spread</td>
<td>26.54</td>
<td>25.69</td>
<td>25.13</td>
<td>25.86</td>
</tr>
<tr>
<td>Business Time</td>
<td>26.13</td>
<td>5.43</td>
<td>1.00</td>
<td>2.65</td>
</tr>
</tbody>
</table>

The statistics are reported separately for the three regimes and the entire day. Volatility is annualized. 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. Business time is proportional to $W^{-2/3}$ and, for ease of comparison, it is normalized to be unity in Regime 3.

The market microstructure invariance principle has additional implications re-
Figure 2. This figure plots the times series of the average statistics for each regime on each trading day: volume $V_d$ (per minute), volatility $\sigma_d$ (per minute, annualized), number of transactions $N_d$ (per minute), and average trade size $Q_d$ (during one minute). Asian hours of Regime 1 figures are plotted in blue, European hours of Regime 2 are in green, and North American hours of Regime 3 are in red. The sample period is January 4, 2008, to November 4, 2011.

Regarding the bid-ask spread and price efficiency. As an empirical matter, these implications depend on constraints that minimum tick size and minimum contract size impose on trading. For the E-mini futures, these frictions are significant. Table 1 shows that the average bid-ask spread in all three regimes is almost identical to the minimum value of 0.25 index points. In other words, the spread is binding and equals a single tick almost always. Similarly, figure 3 demonstrates that a disproportionate number of trades across all three regimes occur at the minimum contract size of one unit.

While the E-mini futures market provides a near ideal setting for exploring the interaction between trading intensity and volatility due to the integrated electronic trading platform and uniformly high degree of transaction activity, exploring invariance predictions regarding bid-ask spreads and trading costs requires dealing with difficult issues related to tick size and minimum contract size.

Finally, we reiterate that the summary statistics in this section involve the observed one-minute trade and volatility data, while the theoretical relations in equa-
Figure 3. The figure plots the proportion of trades with sizes $Q = 1$ and $Q \leq 2$ contracts (the left panel) and $Q \geq 200$ and $Q \geq 50$ contracts (the right panel) in blue and green, respectively. The proportions are computed at $\Delta t = 1$-minute granularity. The dashed vertical lines indicate the three trading regimes.

Evaluations (11) and (13)-(15) concern the market’s expected values for the logarithm of those variables. The one-minute sampling frequency represents a specific trade-off between bias and variance. If we sample at a lower frequency, we exploit more data in the computation of any given expectation and obtain a less noisy measure, reducing the variance of the estimator. On the other hand, as the time interval grows, the underlying variables vary more substantially across the interval, inducing an increasing bias in the estimator for the nonlinear mapping implied by the various hypotheses. By choosing a short time interval of one minute, we tilt in the direction of reducing the bias and thus mitigating the imprecision of our estimators stemming from sampling and measurement errors, by averaging the noisy log-transformed variables over a large set of one-minute intervals. Thus, we do not test whether the relations are satisfied for all distinct one-minute observations, but rather for specific subset of observations relative to others.

As discussed previously, for the transaction count $N$ and returns volatility $\sigma$, we assume that active traders are attuned to the market developments and form unbiased forecasts for the relevant combination of variables over short intervals. Hence, for the transaction count, we use the observed values directly in our analysis. For the latent volatility, due to the liquidity of the E-mini contract, we use six consecutive squared ten-second returns to compute a realized volatility measure without any major detrimental effects from microstructure noise. While this yields a very noisy estimator for the true local volatility, the subsequent aggregation across numerous one-minute intervals implies that we benefit from the same error diversification principle that accounts for the accuracy of standard high-frequency-based realized-volatility estimators.

The log transform presents a practical problem. During quiet market conditions, we may not observe a single trade over a given minute, or the latest trade
prices may be identical at the end of each of the six ten-second interval within a given minute, resulting in zero realized volatility measure. This renders the log transformation infeasible. Thus, we delete all one-minute intervals that feature zero realized volatility; the corresponding fractions of observations filtered out in the three regimes are 29.48%, 8.65% and 1.42%, respectively. Although intraday trading invariance is valid for any exogenously chosen subset of observations, the elimination of observations conditional on zero observed price changes will, in theory, induce an upward bias to the estimated volatilities of the remaining intervals. The filtered observations stem from periods where overall market activity also is extremely low, if not entirely devoid of trading, implying that various constraints may be binding and further complicate any meaningful test for invariance over these isolated intervals.

As a robustness check, we also test for invariance at the substantially lower five-minute sampling frequency, which enables us to retain almost all observations. We filter out only 1.63%, 0.32% and 0.29 % in the three regimes, respectively. Our results are quantitatively the same, and thus we chose to present only results for the one-minute sampling in our main analysis. The one-minute sampling provides a more challenging test and allows us to enhance the granularity of tests around key transition points, when the characteristics of the trading process shift abruptly.

IV. Empirical Results

This section presents tests for invariance relationships across the intraday patterns using specifications from section II.

A. Transactions Count and Trading Activity

We first discuss the regressions for transaction counts provided by equations (11), (13), and (14), which represent the invariance model, the model of Clark (1973), and the model of Ané and Geman (2000), respectively.

Figure 4 shows intraday scatter plots for each of the three models, where one-minute observations are obtained by averaging the across all days in the sample. The points in all three scatter plots lie approximately on a straight line, implying $R^2 > 0.99$ for the all three corresponding linear regressions. All three models predict the slope of the regression line to be $2/3$, represented by a dashed line in each panel. In the plot for the invariance hypothesis (panel 3), the dashed line with slope $2/3$ is visually indistinguishable from the fitted regression line. By contrast, the fitted slope for the two alternative theories are quite different from the predicted value of $2/3$. The closeness of the actual to the predicted slope is strong qualitative support for the intraday trading invariance hypothesis.

The points in sets $O_1, O_2$ and $O_3$ corresponding to the Asian, European and North American segments are depicted in blue, green and red colors, respectively. The patterns for each of the three regimes are consistent with the invariance predictions. There is, however, a cluster of 16 points marked by red crosses that lie below
the regression line for the invariance model. These observations correspond to the data from 15:00–15:15 CT, when the cash market for equities is closed but the futures market trading continues to take place until CME Globex trading stops at 15:15 CT. Remarkably, the slope generated by the crosses is also almost identical to 2/3. The parallel shift in observations suggests that our empirical model may be improved if it is adjusted appropriately for contemporaneous trading in the equity cash market and other related markets.

Additional details regarding the estimated regressions (11), (13), and (14) are provided in table 2. One reason that the $R^2$ of all three regressions are greater than 0.99 is that the number of transactions $n_t$ essentially appears on both sides of each regression equation, explicitly as the dependent variable and implicitly as an explanatory variables. Recalling the definition of trading activity $w_t := p_t + q_t + n_t + s_t/2$, we effectively regress $n_t$ on $p_t + n_t - q_t/2 + s_t/2$ in Clark’s model, $n_t$ on $p_t + n_t + s_t/2$ in Ané and Geman’s model, and $n_t$ on $p_t + n_t + q_t + s_t/2$ in the intraday invariance model.
Table 2—Intraday OLS Regression of $\log N$: The Three Models

<table>
<thead>
<tr>
<th></th>
<th>$c$</th>
<th>$\beta$</th>
<th>se($c$)</th>
<th>se($\beta$)</th>
<th>$\bar{R}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>2.4119</td>
<td>0.9757</td>
<td>0.0031</td>
<td>0.0016</td>
<td>0.9965</td>
</tr>
<tr>
<td>Model 2</td>
<td>1.7538</td>
<td>0.8490</td>
<td>0.0018</td>
<td>0.0006</td>
<td>0.9993</td>
</tr>
<tr>
<td>Model 3</td>
<td>0.8579</td>
<td>0.6708</td>
<td>0.0033</td>
<td>0.0007</td>
<td>0.9986</td>
</tr>
</tbody>
</table>

This table reports on intraday OLS regressions of $n_t$ onto $w_t - 3/2q_t$ (Clark), $w_t - q_t$ (Ané and Geman), and $w_t$ (Invariance). All models predict $\beta = 2/3$. For each regression there are $T = 1,320$ observations.

B. Return Volatility per Transaction and Trade Size

Equation (15) brings out clearly that the fundamental difference between intraday trading invariance and the two alternative theories can be expressed in terms of trade size. The theories offer dramatically different predictions regarding the relation between volatility per transaction and trade size, ranging from a positive relation (Clark, $\beta = 1$), no relation (Ané and Geman, $\beta = 0$), to a strongly negative relation (invariance, $\beta = -2$). This renders the regression equation (15) a particularly clean test of the relative explanatory power of the alternative hypotheses.

We estimate the regression lines across minutes of the day, where the log values of the relevant variables for each one-minute observation are averaged across all days in the sample. The results are presented in figure 5 and table 3.

![Figure 5](attachment:image.png)

**Figure 5.** The left panel shows the intraday scatter plot of $n_t - n_t$ versus $q_t$. The OLS regression line (solid) and the predicted invariance line (dashed) are shown. The right panel shows the same scatter plot with minutes around market opens and closes filtered out.

The left panel of figure 5 shows that the scatter plot of 1,320 intraday observations is now less tightly centered on the regression line than in figure 4, even though the relation again is again close to a log-linear with an overall adjusted $R^2$ of 0.967. Note that the estimated regression line predicted by intraday trading invariance
The table reports the results of the intraday OLS regression $s_t - n_t = c + \beta \cdot q_t + u_t$. Clark, Ané and Geman and invariance predict $\beta = 1$, $\beta = 0$, and $\beta = -2$, respectively. The regression has $T = 1,320$ observations.

(solid, with estimated slope $-2.0002$) is almost identical to the theoretical line (dashed, with slope $-2$). The Clark model and the Ané and Geman model are clearly rejected since the regression line is very far from being either flat or upward sloping.

Upon inspection, we find that the most of the outliers in the left panel of figure 5 stem from periods during which active hours shift from one continent to another. If we eliminate observations corresponding to 3 minutes around 1:00 CT and 2:00 CT (the beginning of European hours), the 3+30 minutes around 8:30 CT (the beginning of North American hours and the 9:00 announcements) as well as the last 16 minutes before 15:15 CT (the end of North American hours with the cash market closed), then we obtain the results depicted in the right panel of figure 5. The invariance predictions are again validated in the sense that the regression slope equals to predicted $-2$. Almost all outliers happen to be removed by this filtering procedure, thus suggesting that intraday invariance works particularly well when trading originates from a single stable trading zone.

These results are consistent with the interpretation that market participants actively adjust trade size in response to varying market conditions, in a manner consistent with intraday trading invariance. Temporary deviations from invariance occur around the times of day that regular trading hours shift from one continent to another and market participants are adjusting trade size most actively.

C. Macroeconomic Announcements

This section considers market dynamics during scheduled releases of macroeconomic announcements. We examine whether intraday trading invariance provides a good characterization of the market dynamics during those events.

In the U.S. market, the most important announcements occur at 7:30 CT. The Employment report is usually released at 7:30 CT on the first Friday of the month, and the Consumer Price Index is typically released at 7:30 CT 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 these releases. Explaining those spikes is a challenging test for intraday trading invariance.
To identify the relevant announcement days in our sample, we focus on the days with the largest increase in trading activity immediately after 7:30 CT. In particular, we compute the ratio of the number of trades for the 10 minutes after 7:30 CT to the number of trades for the 60 minutes before 7:30 CT. We then select the ten percent of trading days with the highest ratios. On average, this procedure identifies approximately two days each month with significant spikes in trading activity likely associated with 7:30 announcements. In the terminology of section II.D, we denote this subset by $\mathcal{D}_{A1}$, and refer to it as the “7:30 announcement days.” We also further zoom into the data and study the three minutes before and after 7:30 CT; we denote this subset by $\mathcal{T}_{A1}$.

Figure 6 depicts the intraday statistics averaged only across the 7:30 announcement days, $\mathcal{D}_{A1}$. Compared to similar statistics for the entire sample $\mathcal{D}$ in figure 1, the upward spikes at 7:30 CT are more pronounced for volume, number of trades, and volatility. Closer inspection also reveals a smaller, but clearly distinctive, downward shift in average trade size.

Figure 7 plots $n_t$ against $w_t$ for three minutes before and after 7:30 CT, the subset of observations $\mathcal{T}_{A1}$. In the left panel, we average the data across all days, i.e., the subset $\mathcal{D} \times \mathcal{T}_{A1}$. In the right panel, we average only across announcement days, i.e., the subset $\mathcal{D}_{A1} \times \mathcal{T}_{A1}$. As a benchmark, we also plot the regression line corresponding to equation (11), as in the bottom panel of figure 4. The three observations corresponding to the three minutes before the 7:30 CT announcements are depicted with points, whereas the three observations corresponding to the three minutes after the 7:30 CT announcements are depicted with crosses.

In the left panel of figure 7, all six points are very close to the regression line, implying that intraday invariance generally provides an accurate depiction of trading and volatility dynamics around 7:30 CT. In the right panel, all six points are still very close to the regression line, consistent with the interpretation that intraday invariance also holds during spikes in trading activity following macroeconomic announcements. Since trading activity tends to increase after the announcements, the observations for the three minutes after 7:30 CT move towards the upper right corner of the figure.

Figure 8 and figure 9 present similar evidence for the “9:00 announcement days.” The releases at 9:00 CT include announcements of New Home Sales, Existing Home Sales, the Housing Market Index, Consumer Sentiment, Consumer Confidence, and Business Inventories, among others. On average, the 9:00 announcements are less pronounced in terms of their effect on average trade size. Yet, the spikes for volume, number of trades, and volatility are still dramatic. As before, the results confirm that intraday trading invariance holds up well during those unusual market conditions.
**Figure 6.** The figure shows averages across the days for which an assumed 7:30 CT announcement resulted in biggest subsequent increase in trading activity. The statistics include share volume $V_t$ (per minute), volatility $\sigma_t$, average trade size $Q_t$, and number of trades. The averages are computed at $\Delta t = 1$-minute granularity. The dashed vertical lines indicate the three trading regimes.

**Figure 7.** This figure shows a scatter plot of $n_t$ versus $w_t$, but only for the 3 minutes before 7:30 (dots) and 3 minutes after 7:30 (crosses). The left panel plots the averages across all days, while the right panel plots the averages across the 7:30 announcement days. The solid line is the OLS regression line from the bottom panel of Figure 4.
Figure 8. The figure shows averages across the days for which the 9:00 CT announcement resulted in the largest subsequent increase in trading activity. The statistics include share volume $V_t$ (per minute), volatility $\sigma_t$, average trade size $Q_t$, and number of trades. The averages are computed at $\Delta t = 1$-minute granularity. The dashed vertical lines indicate the three trading regimes.

Figure 9. This figure shows a scatter plot of $n_t$ versus $w_t$, but only for the 3 minutes before 9:00 CT (dots) and 3 minutes after 9:00 CT (crosses). The left panel plots the averages across all days, while the right panel plots the averages across the 9:00 CT announcement days. The solid line is the OLS regression line from the bottom panel of Figure 4.
D. The Flash Crash

We next discuss how intraday trading invariance can help interpret the chaotic market conditions experienced during the flash crash of May 6, 2010. Are their general principles applicable to other episodes of extreme market conditions? Kyle and Obizhaeva (2013a) describe this episode as well as other stock market crashes associated with execution of large bets. Andersen and Bondarenko (2014, 2015) discuss other early warning signals for market turbulence.

Figure 10 shows (top left panel) the price dynamics of the E-mini S&P 500 futures contract on May 6, 2010. The vertical lines indicate the official timing of the crash event from 13:40 CT to 13:45 CT as identified in the joint report by Staffs of the CFTC and SEC (2010a,b). In the morning, the market declined by about 3 percent. There were rumors about a debt default by Greece, elections in the U.K., and upcoming jobs report in the U.S. From 13:40 CT to 13:45 CT, prices first plummeted 5.12% and then recovered 5% over the next ten minutes, after a pre-programmed circuit breaker built into the CME Globex electronic trading platform halted trading for five seconds. Prices entered a free fall only during the very last minute of the event window. The crash was accompanied by record trading volume and extreme realized volatility (top right and bottom left panels of figure 10). The apparent breakdown in the provision of market liquidity is obviously of great interest.

We examine whether intraday trading invariance relationships held during the market crash or the data displayed some unusual patterns. Since we study a single trading day and the crash event lasted for less than an hour, we focus on the prediction that the invariant log $I_{dt}$ is an identically and independently distributed random variable, as suggested by equation (7). We test this proposition informally by calculating the time series of average log $I_{dt}$ computed at a four-minute frequency, normalizing those variables to have a zero mean and a unit variance, and then checking for outliers in the intraday sequence on May 6, 2010.

The dynamics of the standardized trading invariant log $I_{dt}$ is shown in the lower right panel of figure 10. Kyle and Obizhaeva (2013b) find that the distribution of log $I_{dt}$ closely matches a normal distribution when portfolio transition orders are used as empirical proxies for bets. In figure 10, the time series of realized log $I_{dt}$ is clearly inconsistent with a normal distribution. There are three consecutive large outliers, exceeding 3 standard deviations each, which occurred during the first three four-minute intervals immediately following the five-second trading halt at the bottom of the flash crash. When compared with realizations of log $I_{dt}$ during the same time interval for other days in the sample, these three outlier points represent 100th, 99.9th, and 99.4th percentile events. Just before prices bottomed out, the value of the invariant log $I_{dt}$ corresponds to a 78th percentile event during the last four minutes before prices bottomed out. Prior to these outliers, there were no observations of log $I_{dt}$ exceeding ±2 standard deviations. Visual inspection suggests that fluctuations in log $I_{dt}$ were consistent with a normal distribution before the flash crash and during the time prices were crashing, but
the normal distribution broke down as prices began to recover.

\[ \text{Figure 10.} \text{ The figure shows price } P, \text{ volume } V, \text{ volatility } \sigma, \text{ and standardized log invariant } \log I \text{ on May 6, 2010. The solid vertical lines indicate the timing of the flash crash.} \]

What does the apparent breakdown in invariance, suggested by these three outlier points, tell us about how liquidity provision may have broken down during the flash crash?

Menkveld and Yueshen (2013) document that the co-integration relation between the E-mini contract and the SPDR S&P 500 exchange traded fund (the “spider” or SPY) broke down about one minute prior to the trading halt and resumed eight minutes after it. The three outlier points in \( \log I_{dt} \) coincide exactly with this period of breakdown in liquidity provision. Several factors hampered the 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 and uncertainty among market participants.

Kirilenko et al. (2012) report (see their table II) that the average trade size of groups of traders which represent institutional investors (Fundamental Buyers and Fundamental Sellers) is larger than the average trade size of groups which intermediate trading (High Frequency Traders, Market Makers, and Opportunistic
Traders). In the last few minutes before the bottom of the flash crash, the participation rate of high frequency traders increased, but it dropped immediately after prices bottomed out (see their figure 6).

Taken together, these results suggest that the breakdown in arbitrage relationships, which occurred immediately after prices hit bottom during the flash crash, is associated with a low participation rate by market makers and arbitragers. Since these intermediaries tend to execute orders with small trade size, we expect average trade size to rise when their participation rate falls. We conjecture that the invariant $\log I_{dt}$ had unusually large values because average trade size was larger than usual due to lack of participation by market makers and arbitragers whose trade size tends to be small. We therefore conjecture that apparent breakdown in invariance relationships, consisting of the three large outlier points in $\log I_{dt}$, indicates less participation by market makers and arbitragers than in normal market conditions.

E. The Cash Market Open and Close

Under the null hypothesis of intraday trading invariance, we can calculate an implied estimate for the average trade size as a function of volume and return volatility, both of which are either observable or readily estimated. Comparing these predicted trade sizes with actual minute-by-minute observations of average trade size, which may potentially exhibit highly idiosyncratic features at specific times during the daily cycle, represents a demanding test for intraday invariance.

From the basic intraday invariance relation (15) with $\beta = -2$ and $v = n + q$, we obtain the theory-implied trade size $q^*_t$ given by

$$q^*_t = c + \frac{1}{3} [v_t - s_t].$$

The constant $c$ is chosen to make the average implied log trade size $q^*_t$ match the sample average value of realized log trade size $q_t$. Figure 11 compares the actual average log trade size $q_t$ and the corresponding implied log trade size $q^*_t$. The left panel shows that implied trade size tracks very closely the actual trade size, thus confirming that intraday trading invariance provides an accurate explanation for how trade size responds to changes in volume and returns volatility on a minute-by-minute level.

The right panel shows the log prediction error $q_t - q^*_t$. Most of the prediction errors fall in the range $\pm 0.10$, i.e., approximately $\pm 10\%$. Since average trade size is less than 10 contracts (table 1), the prediction errors generally represent less than one contract. As discussed in sections IV.C and IV.E, most of extreme variations around public announcements and the cash market close are consistent with invariance relationship and therefore do not appear as outliers in figure 11.

The figure reveals the known U-shaped patterns for the average log trade size within each trading regime. Note that the most striking deviation from the U-shaped patterns occurs from 1:00 CT to 2:00 CT, blurring the time boundary
between Asian and European cash market trading hours.

Two institutional details explain this irregularity. First, there are two times at which European trading hours begin. Exchanges in Continental Europe open at 1:00 CT and trading venues in London open one hour later at 2:00 CT. Second, trading venues in Europe and North America adhere to a summer time convention, whereas the major trading venues in Asia do not. This makes the boundary between Asian and European trading hours shift back and forth depending on the time of the year.

There are striking distinct discontinuities in trade size when market activity shifts from one region to another due the opening and closing of regular trading hours. Trade size increases significantly as trading shifts from Asia to Europe when European cash markets open, and it increases even more as trading shifts from Europe to North America when U.S. cash markets open.

These shifts are also associated with non-trivial deviations between the actual and implied trade sizes. The downward spikes in the right panel of figure 11 at exactly 1:00, 2:00 and 8:30 represent clear outliers, revealing that actual trade sizes are somewhat lower than implied trade size at exactly the time when a more active cash market opens. The adjustments which intraday trading invariance makes to implied trade size based on changes in volume and volatility do not fully account for the difference between implied and actual trade size. In contrast, there is an upward spike at 15:00 CT at the U.S. cash market close.

The differences between actual and implied trade size after the European and U.S. open and after the U.S. close are quite persistent. Trade size is lower than predicted for at least one hour after the beginning of European trading hours and for two hours after beginning of North American trading hours.

The extreme increase in trade size after the 15:00 CT U.S. cash market close presents a challenge for theories of trade size. Figure 1 shows that the volume and the number of trades temporarily spike upward at 15:00 CT by a factor of
3 and a factor of 2, respectively. At the same time, returns volatility goes down by about 20 percentage point during the same minute. These patterns are clearly inconsistent with alternative models of Clark or Ané and Geman, both of which imply that returns variance must instead increase by a factor of 3 or 2, respectively.

In contrast, intraday trading invariance hypothesis predicts correctly that average trade size will increase, not decrease. The actual increase — by a factor of approximately 1.8 — is, however, more than 20% larger than the predicted increase.

Consistent with our conjecture for the flash crash, we conjecture that actual trade size is larger than predicted trade size after the U.S. close because a relatively larger fraction of trades represent institutional investors, which tend to make large trades, and a relative smaller fraction of trades represent market makers and arbitragers trades, who tend to make smaller trades. We leave for future research further investigation of why systematic deviations from implied trade size occur at the open and the close of active cash markets.

F. Robustness Checks

This section provides a set of complementary analyses to show the robustness of our findings. First, we consider whether the main conclusions hold over subperiods within the sample.

Intraday Invariance in Subsamples

Table 4 reports results for regression (11) for each of the individual calendar years in our sample. For conciseness, we focus only on the intraday trading invariance hypothesis. The alternative specifications continue to be decisively rejected.

<table>
<thead>
<tr>
<th>Year</th>
<th>$c$</th>
<th>$\beta$</th>
<th>$se(c)$</th>
<th>$se(\beta)$</th>
<th>$\bar{R}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>0.9110</td>
<td>0.6621</td>
<td>0.0037</td>
<td>0.0007</td>
<td>0.9983</td>
</tr>
<tr>
<td>2009</td>
<td>0.7674</td>
<td>0.6798</td>
<td>0.0036</td>
<td>0.0007</td>
<td>0.9984</td>
</tr>
<tr>
<td>2010</td>
<td>0.7901</td>
<td>0.6866</td>
<td>0.0046</td>
<td>0.0010</td>
<td>0.9972</td>
</tr>
<tr>
<td>2011</td>
<td>0.9769</td>
<td>0.6531</td>
<td>0.0051</td>
<td>0.0011</td>
<td>0.9965</td>
</tr>
<tr>
<td>All</td>
<td>0.8579</td>
<td>0.6708</td>
<td>0.0033</td>
<td>0.0007</td>
<td>0.9986</td>
</tr>
</tbody>
</table>

This table reports the results of OLS regressions, $n_t = c + \beta \cdot w_t + u_t$. Coefficients, standard errors, and $\bar{R}^2$ statistics are estimated both separately for each calendar year and then for the whole sample. The last year ends November 4, 2011. Each regression exploits $T = 1,320$ observations.

The four yearly estimates of 0.6621, 0.6798, 0.6866, and 0.6531 cluster around the predicted value of 2/3. As in the pooled sample, due to very small standard errors, a statistical test rejects the hypothesis that any of the coefficients equal 2/3. From an economic perspective, however, the differences between estimated and predicted values are almost immaterial.
Table 5 reports the results of the same regressions, but now estimated separately for each of the three trading regimes (Asian hours, European hours, and North American hours) across the full sample. Since the dispersion in levels of trading activity series is sharply reduced by not pooling observations across the trading regimes, the regressions may be relatively more distorted by statistical randomness and various market frictions such as periodic low trading activity, minimum tick size, and minimum contract size.

Table 5—OLS Regression: Intraday Patterns for Each Regime.

<table>
<thead>
<tr>
<th>Regime</th>
<th>$c$</th>
<th>$\beta$</th>
<th>se($c$)</th>
<th>se($\beta$)</th>
<th>$\bar{R}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regime 1</td>
<td>0.9722</td>
<td>0.6212</td>
<td>0.0109</td>
<td>0.0047</td>
<td>0.9707</td>
</tr>
<tr>
<td>Regime 2</td>
<td>0.8415</td>
<td>0.6740</td>
<td>0.0155</td>
<td>0.0035</td>
<td>0.9894</td>
</tr>
<tr>
<td>Regime 3</td>
<td>0.6436</td>
<td>0.7016</td>
<td>0.0389</td>
<td>0.0055</td>
<td>0.9754</td>
</tr>
</tbody>
</table>

This table reports the results of OLS regressions, $n_t = c + \beta \cdot w_t + u_t$. The coefficients, standard errors, and $\bar{R}^2$ statistics are estimated separately for each regime. The regressions exploit $T_1 = 525$, $T_2 = 390$, and $T_3 = 405$ observations, respectively.

The slope coefficients of 0.6212, 0.6740, and 0.7016 in table 5 center less tightly on the predicted value of 2/3.

**Intraday Invariance in Disaggregated Data**

Our tests thus-far have been based on creating bins by summing one-minute data across many days to minimize measurement error and obtain identification from significant fluctuations in levels of trading activity across a 24-hour day. In this section, instead of summing across days, we create bins by summing across consecutive minutes within each day. In principle, the bins should contain enough minutes so that our statistical tests are not distorted by low levels of trading activity generating a tiny number of trades and thus great measurement error within one bin. In particular, the bins should be large enough so that each bin contains a non-zero number of trades, making our log-transformation mathematically feasible.

Intraday trading invariance suggests that bins should be proportional to $W^{-2/3}$, where trading activity $W$ is measured within each bin. Thus, we choose the aggregation bins for the three regimes to contain 105, 26, and 5 consecutive minutes, respectively. This choice ensures that each regime is covered by an integer number of bins, the bin sizes are approximately proportional to the business clock $W^{-2/3}$ as shown in table 1. After aggregation, each day has 101 bins, and there are almost 100,000 ($101 \times D$) observations in the sample.

Using a slight abuse notation by letting the subscript $b$ indicate the bin within the trading day rather than the one-minute interval, we let $n_{db}$ and $w_{db}$ denote the log number of trades per minute and the log of trading activity per minute within the bins. Figure 12 shows a scatter plot of $n_{db}$ versus $w_{db}$. The observations are grouped tightly along the fitted line with a slope of 0.676, close to the predicted
value of $2/3$ implied by intraday trading invariance. Since less aggregation leads to more variation in measured trading activity than in figure 13, the cloud of observations from the different regimes overlap to a larger extent than observed previously. Despite large spikes in the intraday profiles for some of the underlying series in figure 1, none of the observations appear to be major outliers relative to the regression line.

**Figure 12.** Scatter plot of $n_{db}$ against $w_{db}$ with data aggregated across bins $b$ for each day $d$. As before, the three regimes are represented by distinct colors.

Table 6 reports a regression coefficient of pooled OLS regressions for the binned data. The estimated coefficient for the whole sample is 0.6692, and the estimates for each of the four calendar years are 0.6564, 0.6739, 0.6846, and 0.6706, respectively. Table 7 reports similar regression coefficients of 0.6500, 0.6503, and 0.6494, but now estimated separately for each of the regimes. All regressions produce a slope estimate close to the value of $2/3$ predicted by intraday trading invariance.

The binning method allows the regression coefficient to be influenced both by intraday variation in trading activity across bins and by time series variation across days. Taking both sources of variation into account, the regression coefficients remain close to the predicted value of $2/3$.

**Intraday Invariance in the Daily Time Series**

We next report regression results from further aggregating over minutes while not aggregating over days. Specifically, we generate three (regime-day) data points per day, each corresponding to an observation from the data within each of the trading regime sets $O_1$, $O_2$ and $O_3$. While this approach preserves intraday variation across
Table 6—OLS Regression: Binned Data

<table>
<thead>
<tr>
<th></th>
<th>c</th>
<th>β</th>
<th>se(c)</th>
<th>se(β)</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>0.8838</td>
<td>0.6564</td>
<td>0.0039</td>
<td>0.0006</td>
<td>0.9825</td>
</tr>
<tr>
<td>2009</td>
<td>0.7398</td>
<td>0.6739</td>
<td>0.0045</td>
<td>0.0007</td>
<td>0.9765</td>
</tr>
<tr>
<td>2010</td>
<td>0.7279</td>
<td>0.6846</td>
<td>0.0051</td>
<td>0.0008</td>
<td>0.9678</td>
</tr>
<tr>
<td>2011</td>
<td>0.7998</td>
<td>0.6706</td>
<td>0.0056</td>
<td>0.0008</td>
<td>0.9679</td>
</tr>
<tr>
<td>All</td>
<td>0.8004</td>
<td>0.6692</td>
<td>0.0024</td>
<td>0.0003</td>
<td>0.9743</td>
</tr>
</tbody>
</table>

This table reports the results of OLS regressions across bins, \( n_{db} = c + \beta \cdot w_{db} + u_{db} \). The coefficients, standard errors, and \( R^2 \) statistics are estimated separately for each calendar year and for the whole sample. The last year is incomplete and ends on November 4, 2011.

Table 7—OLS Regression: Binned Data, 3 Regimes

<table>
<thead>
<tr>
<th></th>
<th>c</th>
<th>β</th>
<th>se(c)</th>
<th>se(β)</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regime 1</td>
<td>0.7566</td>
<td>0.6500</td>
<td>0.0054</td>
<td>0.0020</td>
<td>0.9564</td>
</tr>
<tr>
<td>Regime 2</td>
<td>0.8821</td>
<td>0.6503</td>
<td>0.0055</td>
<td>0.0012</td>
<td>0.9563</td>
</tr>
<tr>
<td>Regime 3</td>
<td>0.9497</td>
<td>0.6494</td>
<td>0.0045</td>
<td>0.0006</td>
<td>0.9340</td>
</tr>
</tbody>
</table>

This table reports the results of OLS regressions \( n_{db} = c + \beta \cdot w_{db} + u_{db} \) across bins \( b \) and days \( d \). The coefficients, standard errors, and \( R^2 \) statistics are estimated separately for each regime.

regimes, it allows us to test whether variation across days is also consistent with invariance.

Figure 13 depicts the OLS fit for the regression of log number of trades on trading activity following equation (11) based on the pooled regime-day series. As before, the regression slope of 0.6682 is almost indistinguishable from the slope of 2/3 predicted by intraday trading invariance from equation (11). The alternative specifications associated with equations (13) and (14) again fail to generate the regression slopes predicted by the theories of Clark or Ané and Geman, respectively.

Figure 14 presents the results for a regression of log volatility per trade on log trade size, constructed along the lines of equation (15). It can be compared to the cross-sectional evidence in figure 5. Figure 5 shows that the points cluster along a fitted line with slope −1.98, close to the predicted slope of −2. We also overwhelmingly reject the alternative theories of Clark or Ané and Geman, which imply either a flat or positively-slope regression line.
V. Conclusion

The invariance properties of trading patterns in the E-mini S&P 500 are unexpected and powerful results, which raise interesting challenges for both empirical and theoretical research into market microstructure.

From an empirical viewpoint, it is interesting to ask how universal our current findings are. On the one hand, even if the results apply only to this very liquid futures market — often characterized as the primary location for price discovery in U.S. equities — they may still be an artifact of specific institutional arrangements
or other unique features of this specific marketplace. Possibilities include the large tick size, minimum contract size, tight integration with liquid cash markets, role of high frequency traders or other algorithms, or specific features of the CME Globex matching engine. On the other hand, if intraday trading invariance is a more universal phenomenon for financial markets, 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, or 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.

From a theoretical viewpoint, our findings pose a difficult conceptual question. Intraday trading invariance clearly captures the spirit of the market microstructure invariance hypothesis of Kyle and Obizhaeva (2013). Yet we have presented no formal theoretical model justifying the hypothesis in this paper. Since trades are different from bets, it is clearly something different. The theoretical models of Kyle and Obizhaeva (2013) may be interpreted as implying that invariance holds in a dealer market. The models of Kyle, Obizhaeva and Wang (2014, 2015) suggest that order shredding into infinitesimally small trades is optimizing strategic behavior aimed at controlling transactions costs. Yet we are aware of no theoretical research which which explains how order shredding with a minimum tick size and minimum contract size might lead to intraday trading invariance.

The issues raised in our paper pose a host of questions for future empirical and theoretical research in market microstructure.

REFERENCES


