

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Aldrich, Eric M.; Heckenbach, Indra; Laughlin, Gregory

# Working Paper The random walk of high frequency trading

Working Paper, No. 722

**Provided in Cooperation with:** University of California Santa Cruz, Economics Department

*Suggested Citation:* Aldrich, Eric M.; Heckenbach, Indra; Laughlin, Gregory (2015) : The random walk of high frequency trading, Working Paper, No. 722, University of California, Economics Department, Santa Cruz, CA, https://doi.org/10.2139/ssrn.2481201

This Version is available at: https://hdl.handle.net/10419/125552

#### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

#### Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



# WWW.ECONSTOR.EU

# The Random Walk of High Frequency Trading<sup>\*</sup>

Eric M. Aldrich<sup>†</sup> Department of Economics University of California, Santa Cruz Indra Heckenbach<sup>‡</sup> Department of Physics University of California, Santa Cruz

Gregory Laughlin<sup>§</sup> Department of Astronomy and Astrophysics University of California, Santa Cruz

> First Draft: August 15, 2014 This Draft: June 1, 2015

#### Abstract

This paper builds a model of high-frequency equity returns by separately modeling the dynamics of trade-time returns and trade arrivals. Our main contributions are threefold. First, we characterize the distributional behavior of high-frequency asset returns both in ordinary clock time and in trade time. We show that when controlling for pre-scheduled market news events, trade-time returns of the near-month E-mini S&P 500 futures contract are well characterized by a Gaussian distribution at very fine time scales. Second, we develop a structured and parsimonious model of clock-time returns using a time-changed Brownian motion composed with a general, non-Lévy directing process. Particular cases of this model allow for leptokurtosis and volatility clustering in clock-time returns, even when trade-time returns are Gaussian. Finally, we highlight conditions on the directing process which are required in order to generate proper volatility dynamics while simultaneously matching the unconditional distribution of returns. In-sample fitting and out-of-sample realized volatility forecasting demonstrate the strength of our model relative to leading candidates.

Keywords: High-frequency trading, US Equities, News arrival.

**JEL Classification**: C22, C41, C58, G12, G14, G17.

<sup>\*</sup>This work was supported by the Hellman Fellows Fund.

<sup>&</sup>lt;sup>†</sup>Email: ealdrich@ucsc.edu.

<sup>&</sup>lt;sup>‡</sup>Email: iheckenb@ucsc.edu

<sup>&</sup>lt;sup>§</sup>Email: glaughli@ucsc.edu

# 1 Introduction

Modern electronic exchanges function in a manner that outwardly display the properties that are expected of an efficient, liquid market. Bid-offer spreads are narrow in comparison to the price of the underlying instrument being traded, volumes are high, high-frequency traders compete to make markets, and information regarding price discovery is disseminated at nearly the speed of light (see, e.g. Brogaard et al. (2013) and Hasbrouck and Saar (2013)). Within such an environment, the disparate, shifting spectrum of intentions of a wide range of market participants is continuously being aggregated, and so a naive, but nonetheless reasonable, expectation is that the Central Limit Theorem should play a fundamental role, and that short-period returns should adhere to a Gaussian distribution.

Indeed, from the pioneering work of Bachelier (1900) through the development of the Black-Scholes options pricing model (Black and Scholes (1973)), modern finance has traditionally held that market price movements can be approximated to a somewhat useful degree by a Gaussian random walk. In reality, observed distributions of market returns are markedly non-Gaussian. Regardless of venue and asset class, returns distributions invariably have fat tails and display the phenomenon of volatility clustering. A rich literature exists which describes both the characterization and the modeling of the observed departures from normality (for a review, see Bouchaud (2005)).

In this contribution, we carry out a ground-level re-examination of the process that generates short-period market returns within the context of high-frequency trading (over time scales ranging from milliseconds to minutes). We analyze six complete months of recent, millisecond-resolution tick data from the extremely liquid near-month E-mini S&P 500 futures contract traded at the Chicago Mercantile Exchange.

Our analysis begins with a significant empirical insight: outside of pre-scheduled news announcement periods, high-frequency asset returns are well described by a Gaussian distribution when trade time is employed. Brada et al. (1966) introduced the notion of trade time to show that asset returns distributions are nearly Gaussian if the returns process is subordinated with successive transactions (trades) acting as the subordinator. Mandelbrot and Taylor (1967) showed that a Gaussian random walk composed with a subordinating trade-time process is fully consistent

2

with a fat-tailed, Lévy-stable distribution, as suggested in Mandelbrot (1963). Clark (1973) used an alternative subordinator, time measured by volume of transactions, to obtain similar results. More recently, Ane and Geman (2000) shows that coarsely sampled intra-day returns also conform to a Gaussian distribution when measured in trade (transaction) time.

Our analysis demonstrates that the Gaussianity of trade-time returns does not immediately extend to high-frequency intra-day returns. That is, high-frequency, trade-time returns exhibit heavy tails and volatility clustering when considered unconditionally throughout the day. However, by excluding the periods surrounding pre-scheduled news events, we confirm the existence of tradetime Gaussianity as well as a lack of volatility persistence.

Building on the empirical observations above, our paper makes two theoretical contributions. First, we develop a general time-changed Brownian motion model for high-frequency asset returns. similar to those of Press (1967), Mandelbrot and Taylor (1967) and Clark (1973), but allowing for more general (non-Lévy) directing processes. In particular, we allow the directing process to be characterized by inter-trade durations which follow the autoregressive conditional duration (ACD) model of Engle and Russell (1998) as well as the Markov-Switching Multifractal Duration (MSMD) model of Chen et al. (2013) and Zikes et al. (2014), which builds on the work of Mandelbrot et al. (1997), Calvet et al. (1997) and Fisher et al. (1997), as well as subsequent work by Calvet and Fisher (2001), Calvet and Fisher (2002) and Calvet (2004). In this dimension our work differs substantially from that of Ane and Geman (2000): where they begin with a nonparametric estimate of the distribution of clock-time asset returns and work backwards to implicitly define the nonparametric density of trades that would be consistent with trade-time Gaussianity, we work forwards by first compounding a parametric distribution of trade-time returns with a parametric model of duration times (and hence, an associated trade arrival process) to characterize the distribution of clocktime returns. Our contribution is significant because it promotes a structured and parsimonious approach to approximating the observed evolution of asset returns.

Second, we develop conditions under which the directing process can generate realistic dynamics for asset returns. The early work on subordinated Brownian motion, cited above, focused entirely on the unconditional distribution of returns. Our work, however, shows that this same class of models is flexible enough to capture volatility persistence when the proper directing process is used. In a similar vein, Carr and Wu (2004) show how this class of models can account for the leverage effect.

The financial econometrics literature has utilized stochastic time changes to explain volatility dynamics, but this has typically been done by working with volatility directly. Madan and Seneta (1990) is an early example of this. More recently, Andersen et al. (2007), Andersen et al. (2010) and Todorov and Tauchen (2014) "devolatize" intra-day clock-time returns using short-term volatility measures in order to achieve conditional Gaussianity. Such devolatization devices are akin to stochastic time changes in that they control for the random flow of information and assume that the underlying returns process is Gaussian. The upshot is that they use latent volatility as a surrogate for information content while we use the observed transaction record. Our theoretical choice to focus on transactions is a result of the strong empirical evidence suggesting that this is the proper device for time deformation.

Using Monte Carlo simulations, we show that our compound duration model is a good characterization of observed clock-time returns and that the stochastic transformation between clock time and trade time for our data is most effectively explained using MSMD durations. In particular, we highlight the in-sample strengths of both versions of our model (ACD and MSMD durations) relative to a benchmark compound Poisson as well as a more traditional GARCH model that has been adapted to high-frequency data (Engle (2000)). We also conduct an out-of-sample realized volatility forecasting exercise which demonstrates that despite its high degree of parameterization, the compound MSMD model significantly outperforms competing models.

Our paper proceeds as follows. We begin by describing our data in Section 2 and provide an analysis of the distributional characteristics of the data during news-affected and non-news-affected subperiods in Section 3. In Section 4, we describe the model and determine conditions under which it can produce volatility dynamics. Section 5 estimates the model and compares Monte Carlo simulations with observed data while Section 6 reports out-of-sample realized volatility forecasting results. Section 7 concludes.

# 2 Data

In this paper we focus our analysis exclusively on the Chicago Mercantile Exchange (CME) nearmonth E-mini S&P 500 Futures contract (commodity ticker symbol ES). Although the CME provides a variety of E-mini products, the E-mini S&P 500 futures contract is the most heavily traded, and for this reason it is commonly referred to as *the* E-mini. As indicated by its name, the E-mini (ES) is a futures contract that trades at 1/5th the size of the standard S&P 500 futures contract. It has a notional value of 50 times the index. We obtained the full record of trades for the period 1 January 2014 to 31 December 2014 by parsing the CME historical files, encoded in FIX format, which we use to estimate and evaluate our model in Section 5.

Despite the fact that the E-mini is a futures contract that does not trade on equities exchanges, its statistical behavior characterizes the dynamics of the equity markets as a whole. This is attributed to its liquidity and the relationship of price formation and information transmission between the futures and equities exchanges in Illinois and New Jersey, as studied in Laughlin et al. (2014).

E-mini futures trade Monday through Friday, starting at 5:00 p.m. Central Time on the previous day and ending at 4:15 p.m., with an additional daily maintenance trading halt from 3:15 p.m. to 3:30 p.m., Central Time. We aggregate multiple trades occurring within single milliseconds as unit transactions and assign to them the final, in-force price of the millisecond as the price of the trade. While not a perfect approximation, this assumption exploits the fact that multiple transactions with the same time stamp are nearly always attributable to a single aggressor order filling several resting orders at the same price level and also allows us to circumvent singularities associated with zero durations in our subsequent models. Our resulting data for the sample period contains a total of 11,875,293 such transactions.

The quoted price of the E-mini corresponds to the index value of the S&P 500. The minimum quoted price increment is  $\Delta P = 0.25$  index points, which corresponds to an actual increment of \$50  $\times 0.25 = $12.50$  for a single contract. For the remainder of the paper we will use *quoted* E-mini prices, measured in points, which correspond directly to the S&P 500 index value. Between January and June of 2014, the S&P 500 Index traded between  $P_{min} = 1732$  and  $P_{max} = 1960$ , indicating a

typical minimum percentage increment in the index price of  $100 \times \frac{\Delta P}{(P_{min}+P_{max})/2} \sim 0.013\%$ .

Because we are interested in investigating the distributions of intra-day asset returns during news event periods and non-event periods, we consider subsamples of the data that sort according to news events. Since the E-mini is a futures contract on a market aggregate, it is almost exclusively affected by major macroeconomic announcements, and not by smaller scale, industryor firm-specific news. For this reason, we classify news event periods with the EconoDay calendar (econoday.com), which lists major pre-scheduled news announcements in the U.S. and which powers calendars for outlets such as Bloomberg and the Wall Street Journal. In particular, we focus attention on the EconoDay 'Market Moving Indicator' and 'Merit Extra Attention' series, as past experience dictates that these are the events most relevant to markets. Since 10:00 a.m. is the most common news announcement time during regular market hours, we form an event-driven dataset of all E-mini trades that occurred during a 1000-second (roughly 16-minute) window following a news announcement at that time. That is, the event-driven dataset is comprised of trades from approximately 10:00-10:16 a.m. that follow any pre-scheduled news announcement during the sample period. We form a corresponding non-event-driven subsample that constitutes all trades during the same time window on days when announcements were not scheduled. A 1000-second window encompasses participants ranging from the fastest algorithmic traders to human traders who manually read the news, consider its implications, and trade on their resulting conclusions.

For the remainder of the paper, we will refer to the event-driven subsample as the *active* data and the non-event subsample as the *passive* data. Our initial separation of daily data by news resulted in 61 active data periods and 67 passive data periods, but due to complications with timing event windows<sup>1</sup>, we discarded the 12 days corresponding to the University of Michigan Consumer Sentiment reports, which reduced the sample to 57 active periods and 58 passive periods. Despite the fact that trading is heavier following news events, the former sample is only about 20% larger: 364,264 records in the active data and 302,840 records in the passive data. Hence, our results below

<sup>&</sup>lt;sup>1</sup>The University of Michigan Consumer Sentiment report is released twice monthly at 9:55 a.m., Eastern Time. Since some of these announcements also coincided with other 10:00 a.m. announcements during our sample period, a 1000-second time window would not have provided an appropriate filter relative to events on other days. For uniformity, we simply exclude the Consumer Sentiment days.

are not an artifact of clock time (both data sets are drawn from the same intra-day time periods) or sample size.

# 3 Empirical Distributions of Intra-Day Asset Returns

In this section we emphasize some of the key features of observed intra-day asset returns distributions. We define clock-time returns as

$$r_{\delta}(t) = p(t) - p(t - \delta) \tag{1}$$

where p(t) is the price of an asset at time t and  $\delta$  is the clock-time duration under consideration (such as 1000 milliseconds). An alternative definition of returns can be provided in trade time, where time increments are measured by a fixed number of trades:

$$r_m(n) = p(n) - p(n-m),$$
 (2)

where n represents the n-th trade and m represents the number of trades in a unit of time. Note that trade time is distinct from volume time: the former defines time by number of transactions (regardless of the size of the transactions) while the latter defines time by number of contracts traded. Although the two are closely related (volume is much less variable in trade time than clock time) they are distinct. As mentioned in the introduction, Clark (1973) subordinates returns with volume as a time increment. However, within the context of the historically well documented volume/volatility relationship (Karpoff (1987), Gallant et al. (1992), Tauchen et al. (1996), Aldrich (2013)), Jones et al. (1994) shows that there is little additional information content in volume beyond the number of transactions. In each case, the objective is to use a measure of time that controls for latent information arrival. We choose to follow Jones et al. (1994) and adopt number of transactions as the best surrogate.

Empirical asset returns (measured in clock time) typically exhibit notable features such as leptokurtosis and conditional heteroskedasticity, regardless of time scale. Much effort has been expended over the course of decades to model the heavy tails of returns distributions (Mandelbrot (1963)) as well as the strong autocorrelation of volatility (Engle (1982) and Bollerslev (1986)). In the subsequent analysis we show that the observations of Brada et al. (1966), Mandelbrot and Taylor (1967), Clark (1973) and Ane and Geman (2000), that trade-time returns are nearly Gaussian, extend to high-frequency, intra-day returns when controlling for pre-scheduled news announcements.

#### 3.1 Intra-day Clock-Time and Trade-Time Distributions

As reported in Section 2, the 57 active 1000-second time intervals contain 364,264 transactions, while the 58 passive 1000-second time intervals contain a total of 302,840 transactions. This corresponds to an average of 6.39 transactions per second in the active subsample and 5.22 transactions per second in the passive subsample. In a prior data subsample, for 2013, we found the respective averages to be closer to 3 and 4 transactions per second. To strike a balance, we assume in the remaining analysis that 1 transaction corresponds to 0.25 seconds, or 250 milliseconds. We consider time aggregated returns for  $\delta = \{250, 500, 1000, 4000, 10000, 30000\}$  milliseconds (i.e. our largest time scale is 30 seconds) and  $m = \{1, 2, 4, 40, 400, 4000\}$  trades, which according to our approximation are roughly corresponding time intervals.

Figure 1 isolates empirical returns distributions for a single time scale during both active and passive news regimes:  $\delta = 10000$  ms (10 seconds) and m = 40 trades. The upper row of plots depict the empirical density functions of discretely observed returns over the entire sample period, superimposed upon Gaussian distributions that are estimated by maximum likelihood. These are shown with the vertical axis on a log scale in order to highlight discrepancies in the tails of the distributions. The second row of plots in Figure 1 show corresponding quantile-quantile (Q-Q) plots of the distributions in the upper panels. Departures from Gaussianity are highlighted via nonlinearities in the Q-Q plots.

Figure 1 shows that intra-day clock-time returns have heavy tails for both active and passive time periods, suggesting that the data generating process has higher probabilities of extreme events than a Gaussian distribution, even during passive, non-event periods. The same is true of activeperiod trade-time returns. The characteristic heavy tails of the empirical distributions are visible in the Q-Q plots as a convex-concave departure from linearity. Remarkably, the empirical density



Figure 1: Empirical density and Q-Q plots for clock-time and trade-time returns for both active and passive subsamples. The clock-time interval is  $\delta = 10000$  milliseconds (10 seconds) and the trade-time interval is m = 40 (chosen to roughly match the clock-time interval).

and corresponding Q-Q plot in the last column of Figure 1 show that passive-period, trade-time returns conform closely to a Gaussian distribution – they do not have the same propensity for extreme events as the other datasets. This feature forms the backbone of our model in Section 4.

Figure 2 shows Q-Q plots of clock-time and trade-time returns during active and passive periods for all time scales that we consider. The panels in the first and second rows of the figure compare the empirical quantiles of the clock-time returns with the theoretical quantiles of the best-fit Gaussian distributions for active and passive periods, respectively. The panels in the third and fourth rows are the same for trade-time returns.

Not surprisingly, the upper rows of Figure 2 show that clock-time returns distributions are markedly different from a Gaussian density over a variety of intra-day time scales, and exhibit patterns of leptokurtosis, especially at fine time scales. In addition to the diminishing tail weight with increasing  $\delta$ , comparison of active-period and passive-period clock-time returns shows that active, news-event periods have much heavier tails. This is exactly as we would anticipate, since periods following news announcements should have a higher frequency of large price movements. The third



Figure 2: Sample Q-Q plots for returns in both clock time and trade time and in both active and passive subsamples. Each panel corresponds to a clock-time or trade-time interval during either the active or passive subsample of data. The vertical axes depict sample quantiles of the data and the horizontal axes depict theoretical quantiles of best-fit Gaussian distributions.

row of Figure 2 shows that trade-time returns during the active subsample are quite similar to those of clock-time returns during active and passive subsamples: leptokurtosis diminishing with m. The final row of the figure, however, highlights the surprising result noted above: during passive, non-event time periods, trade-time returns conform quite closely to a Gaussian distribution.

Gaussianity of high-frequency, trade-time returns is not immediately apparent when studying intra-day data because trading during news event periods has a large impact on the unconditional distribution of returns. For considerations of space, we have not included Q-Q plots for the full sample of data (without sorting on news events), however the corresponding empirical distributions look quite similar to those we have shown for the active subsample. That is, trading during limited periods of pre-scheduled news announcements has a very large impact on the tail probabilities of the unconditional returns distributions. In contrast, as the bottom panels of Figure 2 indicate, if one only considers returns during non-event times, a Gaussian distribution provides an excellent fit. This is true for our very limited subsample of passive period returns (1000 seconds following 10:00 am on each day without a news announcement at that time) as well as for the full sample of intra-day returns that excludes the 1000 second periods following each news announcement.

Figures 3 and 4 depict day-by-day sample autocorrelation functions (ACFs) of returns and squared returns (respectively) for the clock-time and trade-time intervals considered in the Q-Q plots of Figure 2. For each lag value on the x-axis, the 0.05, 0.5 and 0.95 quantiles of the ACFs are depicted in dark red or blue. The corresponding ACF for the entire sample, treated contiguously, is shown as a solid black line. We emphasize, however, that given the large breaks between the daily subsamples of our data, it is not correct to compute the ACF for the entire data sample. The black dotted lines depict 95% Bartlett bounds for the null hypothesis of no serial correlation.

It is generally accepted that asset returns exhibit no significant autocorrelation except at very fine time scales, where bid/offer bounce and mean reversion induce negative correlation among adjacent trades. Figure 3 corroborates these stylized facts: for all series and at all time scales, returns show no autocorrelation except at the first lag, which demonstrates significant evidence of reversion. The strength of the reversion is greatest at the finest time scales, but is still present even for  $\delta = 30000$  ms and m = 120, where the ACF bundles for each dataset are slightly shifted towards negative values and eclipse the negative Bartlett bounds.

ACFs of squared returns, shown in Figure 4, are typically used to diagnose the persistence of volatility, which is a well documented property of financial asset returns. In particular, the panels in the second row of Figure 4, which correspond to passive-period, clock-time returns, are exemplary of the persistent nature of volatility: a high degree of autocorrelation at low lags and a subsequent low-level of persistence over a long horizon, which is exhibited by the vertical displacement of autocorrelation values for  $\delta \leq 1000$  ms. A similar vertical asymmetry is also apparent for the active-period clock-time returns. The remaining panels show that trade-time returns exhibit volatility persistence at low lags, but that the persistence quickly disappears with no apparent long-horizon autocorrelation (i.e. no vertical displacement of the distributions of autocorrelations).



Figure 3: Sample autocorrelation functions of returns in both clock time and trade time and in both active and passive subsamples. Each panel corresponds to a clock-time or trade-time interval during either the active or passive subsample of data.

# 4 Model

Following Mandelbrot and Taylor (1967), we model asset prices as time-changed Brownian motion:

$$p(t) = \sigma B(\tau(t)), \tag{3}$$

where  $\{B(\tau)\}$  is standard Brownian motion and  $\{\tau(t)\}$  is an increasing process known as the directing process. We restrict attention to pure-jump directing processes, which is a characteristic inherited by the price process. Intuitively,  $\{\tau(t)\}$  is a stochastic time change which represents business or transaction time. If  $\{\tau(t)\}$  is restricted to be a Lévy process, then its increments are, by definition, independent and it is referred to as a *subordinator*. Notable examples of such Lévy directing processes are the Gamma process, which results in the Variance-Gamma model of Madan



Figure 4: Sample autocorrelation functions of squared returns in both clock time and trade time and in both active and passive subsamples. Each panel corresponds to a clock-time or trade-time interval during either the active or passive subsample of data.

and Seneta (1990) and Madan et al. (1998), and the Exponential process, which is related to the compound Poisson process. Our model, on the other hand, does not enforce the Lévy restriction of independent increments, but allows for more general directing processes. In fact, the dynamics of returns will depend importantly on the dynamics of the directing process.

For a given increasing, pure-jump directing process,  $\{\tau(t)\}$ , we define two auxiliary processes:

$$t_i = \min_t \tau(t) = i, \quad i = 1, 2, \dots$$
 (4)

$$N(t) = \sum_{i=1}^{\infty} \mathbb{I}_{\{t_i \le t\}}, \quad t > 0.$$
(5)

The point process  $\{t_i\}$  represents the jump or transaction times and  $\{N(t)\}$  is the counting process

induced by the directing process. The point process increments,

$$d_i = t_i - t_{i-1}, (6)$$

are the durations between transactions. The assumption of a Lévy directing process implies independence of the point process and inter-trade durations. On the other hand, non-independent increments to the directing process allows for time dependence in the processes defined in Equations (4) and (5) as well as the inter-trade durations.

The increments to the subordinated Brownian motion in Equation (3) can be expressed as

$$r_{\delta}(t) = \sum_{i=1}^{N_{\delta}(t)} r_i \tag{7}$$

where

$$r_{\delta}(t) = p(t) - p(t - \delta) \tag{8}$$

$$N_{\delta}(t) = N(t) - N(t - \delta) \tag{9}$$

$$r_i \stackrel{i.i.d.}{\sim} \mathcal{N}(\mu, \sigma), \forall i.$$
 (10)

The interpretation is that increments to the subordinated Brownian motion are the summation of a random number of standard Brownian increments. The probability density function of  $r_{\delta}(t)$  is,

$$f(r_{\delta}(t)|\mu,\sigma) = \sum_{k=1}^{\infty} f\left(\sum_{i=1}^{k} r_i \middle| N_{\delta}(t) = k, \mu, \sigma\right) f(N_{\delta}(t) = k).$$
(11)

Equation (11) is a finite Gaussian mixture model with mixture weights that vary according to the probability distribution of  $N_{\delta}(t)$ . The model can also be characterized as a two-stage hierarchical model in which a number of trades is drawn from the distribution of  $N_{\delta}(t)$  in the first stage and a single  $\delta$ -period return is drawn from the Gaussian distribution of  $r_{\delta}(t) = \sum_{i=1}^{N_{\delta}(t)} r_i \sim$  $\mathcal{N}\left(N_{\delta}(t)\mu, \sqrt{N_{\delta}(t)\sigma}\right)$  in the second stage.

Equation (11) is a form of the mixture of distributions hypothesis (MDH, Epps and Epps (1976)) for asset returns and has been utilized in various forms in the finance literature since the work of Press (1967). While certain cases of MDH models mix directly over time varying volatility, and hence allow for temporal dependence in the directing process, much of the focus of stochastic time change MDH models has been on Lévy directing processes, with the objective to match empirical moments of unconditional returns distributions. It is, in fact, the temporal dependence of the directing process which governs the dynamics of return volatility, a result that we develop at the end of this section. First, we highlight several specializations of the model in Equation (7), which we compare in Section 5. It is important to note that these specializations can either be characterized by their directing processes, associated point processes, counting processes or inter-trade durations. With the exception of the compound Poisson process, we will typically find it convenient to focus on their characterization in terms of inter-trade durations.

#### 4.1 Compound Poisson Process

A starting point for modeling trade arrivals would be to assume they follow a Poisson process:

 $N_{\delta}(t) \sim \text{Poisson}(\gamma \delta)$ 

where  $\gamma$  is the trade arrival intensity parameter. In this case Equation (11) becomes

$$f(r_{\delta}(t)|\mu,\sigma) = \sum_{k=1}^{\infty} \frac{1}{\sigma\sqrt{2\pi k}} \exp\left\{-\frac{1}{2} \frac{\left(\sum_{i=1}^{k} r_{i} - k\mu\right)^{2}}{k\sigma^{2}}\right\} \times \exp\{-\gamma\delta\} \frac{(\gamma\delta)^{k}}{k!}.$$
 (12)

This is the compound Poisson process developed by Press (1967) and Press (1968). Although, its density function cannot be obtained in closed form, it can be approximated by Monte Carlo simulation: first making independent draws from the Gaussian distribution and then accumulating random numbers of those Gaussians according to integer deviates drawn from the Poisson density. When  $\delta$  is large relative to  $\gamma$ , the  $\mathcal{N}(\gamma\delta,\sqrt{\gamma\delta})$  density serves as good approximation to the Poisson( $\gamma\delta$ ) density, which results in an approximate Gaussian density for  $r_{\delta}(t)$ . However, for small values of  $\delta$  (very short calendar time intervals) the compound Poisson process does exhibit leptokurtosis.

In this single case, we focus on the counting process as the representation of time deformation and that the assumption of Poisson trade arrivals implies that inter-trade durations are distributed as an Exponential random variable, with rate  $\gamma$ . This is also the only time deformation model that we will consider which has Lévy increments. In the remainder of this section, we will focus on models that represent time deformation with dependent, inter-trade durations.

#### 4.2 Compound ACD Process

Engle and Russell (1998) introduce a dynamic model for inter-trade durations known as the autoregressive conditional duration (ACD) process. Widely considered the benchmark for dynamic models of inter-trade durations, the ACD process exploits a GARCH-style structure for expected waiting times:

$$d_i = \psi_i \varepsilon_i \tag{13}$$

$$\psi_{i} = \omega + \sum_{j=0}^{m} \alpha_{j} d_{i-j} + \sum_{j=0}^{q} \beta_{j} \psi_{i-j}, \qquad (14)$$

where  $d_i$  represents inter-trade durations,  $\psi_i = \mathbb{E}[d_i]$  and  $\varepsilon_i$  represents an innovation at time *i*. Engle and Russell (1998) refer to Equations (13) and (14) as the ACD(m,q) model and separately consider both Exponential and Weibull innovations. As with the GARCH model, Engle and Russell (1998) demonstrate that a simple Exponential-ACD(1,1) model is very effective at capturing the dynamics of inter-trade durations.

While ACD models are typically combined with ARCH/GARCH models to generate highfrequency asset price dynamics (Engle (2000)), we propose a compound-duration model of the form in Equation (7), utilizing counting processes that are induced by ACD durations. We subsequently refer to this model as the compound ACD processes.

#### 4.3 Compound Multifractal Process

Chen et al. (2013) and Zikes et al. (2014) develop the Markov-switching multifractal duration (MSMD) model, which builds on the multifractal volatility model of Calvet and Fisher (2001), Calvet and Fisher (2002) and Calvet (2004). While Zikes et al. (2014) develop a general characterization of the MSMD model, we focus on the specific case considered by Chen et al. (2013). The core components of the model are a set of  $\bar{k}$  latent state variables,  $M_{k,i}$ , that obey a two-state Markov-switching process with varying degrees of persistence,  $\gamma_k$ , for  $k = 1, 2, \ldots, \bar{k}$ . That is, the distribution of trade durations,  $d_i$ , is governed by the equations,

$$d_i = \frac{\varepsilon_i}{\lambda_i} \tag{15a}$$

$$\varepsilon_i \sim Exp(1)$$
 (15b)

$$\lambda_i = \lambda \prod_{k=1}^k M_{k,i} \tag{15c}$$

$$M_{k,i} = \begin{cases} M & \text{with probability } \gamma_k \\ M_{k,i} & \text{otherwise} \end{cases}$$
(15d)

 $M_{k,i-1}$  otherwise

$$\gamma_k = 1 - \left(1 - \gamma_{\bar{k}}\right)^{b^{k-\bar{k}}} \tag{15e}$$

$$M = \begin{cases} m_0 & \text{with probability } 1/2 \\ 2 - m_0 & \text{otherwise.} \end{cases}$$
(15f)

Hence, the MSMD can be succinctly characterized by five parameters:  $\bar{k} \in \mathbb{N}$ ,  $\lambda > 0$ ,  $\gamma_{\bar{k}} \in (0, 1)$ ,  $b \in (1, \infty)$  and  $m_0 \in (0, 2]$ . The intuition is that conditional on knowing the values of the latent state variables, inter-trade durations are Exponentially distributed with intensity parameter  $\lambda_i$ . However, as time evolves, the latent states,  $M_{k,i}$ , switch values with varying degrees of persistence,  $\gamma_k$ . This causes the unconditional distribution of trade durations to be a mixture of Exponentials, which is consistent with the over-dispersion property of the observed data, described in Chen et al. (2013). The latent states can be interpreted as shocks that have varying impacts over diverse timescales, some having short-horizon and others have long-horizon effects. The value *b* governs a tight relationship between the persistence parameters,  $\gamma_k$ , and is responsible for the parsimony of the model: even with a large number of latent states,  $\bar{k}$ , the model is always characterized by a total of five parameters. The choice of *b* dictates the degree of heterogeneity in values of persistence parameters. For more insight regarding the MSMD model, see Chen et al. (2013) and Calvet and Fisher (2008).

The distribution of  $r_{\delta}(t)$  retains its hierarchical structure under the MSMD model, with the number of trades per unit of time being drawn from a mixture of Poisson distributions in the first stage. We refer to  $r_{\delta}(t)$  as a *compound multifractal process* when the Gaussian mixture weights correspond to count probabilities associated with MSMD durations. The variability of the MSMD stochastic intensity parameter,  $\lambda_i$ , causes MSMD durations to exhibit far greater heterogeneity than those of the constant-intensity Exponential. Chen et al. (2013) liken stochastic intensity in duration models to stochastic volatility in returns models: "Just as stochastic volatility 'fattens' Gaussian conditional returns distributions, so too does MSMD 'over-disperse' exponential conditional duration distributions." (p. 9). In fact, we find that stochastic intensity plays the dual role of duration over-dispersion and returns tail fattening: the dispersion of probability over a greater variety of counts, relative to the simple Poisson model, induces greater heterogeneity in the Gaussian mixture, which generates fatter tails for  $r_{\delta}(t)$ . Intuitively, the random variable  $r_{\delta}(t)$  switches between a greater variety of differing sums of Gaussians with higher probability. In addition, the compound multifractal model explicitly generates volatility persistence by producing autocorrelation in the inter-trade duration distribution, via Markov-switching latent states. Similar mechanisms in the ACD process result in leptokurtosis and volatility persistence for clock-time returns.

#### 4.4 Volatility Persistence

The compound returns model of Equations (7) - (10) is able to simultaneously generate fat tails and volatility persistence. The latter feature, however, is dependent on the dynamics of the counting process, which we now show.

**THEOREM 1** Given a counting process N(t) and a Gaussian mixture process  $r_{\delta}(t)$  as specified in Equations (7) – (10), then

$$\gamma_{r_{s,k}^2} = \sigma^4 \gamma_{N_{\delta},k}, \quad k = 0, 1, 2, \dots,$$
(16)

where  $\gamma_{X_{\delta},k} = \operatorname{Cov}(X_{\delta}(t), X_{\delta}(t-\delta k))$  is the *k*th  $\delta$ -period autocorrelation for  $X \in \{r^2, N\}$ .

#### Proof

$$\gamma_{r_{\delta}^2,k} = \operatorname{Cov}\left(r_{\delta}^2(t), r_{\delta}^2(t-k\delta)\right) \tag{17}$$

$$= \operatorname{Cov}\left(\left[\sum_{i=1}^{N_{\delta}(t)} r_{i}\right]^{*}, \left[\sum_{j=1}^{N_{\delta}(t-k\delta)} r_{j}\right]^{*}\right)$$
(18)

$$= \operatorname{Cov}\left(\underbrace{\sum_{i=1}^{N_{\delta}(t)} r_{i}^{2} + 2\sum_{i=1}^{N_{\delta}(t)-1} \sum_{n=i+1}^{N_{\delta}(t)} r_{i}r_{n}}_{A}, \underbrace{\sum_{j=1}^{N_{\delta}(t-k\delta)} r_{j}^{2} + 2\sum_{j=1}^{N_{\delta}(t-k\delta)-1} \sum_{m=j+1}^{N_{\delta}(t-k\delta)} r_{j}r_{m}}_{B}}_{B}\right)$$
(19)

$$=\underbrace{\mathbb{E}\left[\operatorname{Cov}\left(A,B|N_{\delta}(t),N_{\delta}(t-k\delta)\right)\right]}_{=0} + \operatorname{Cov}\left(E[A|N_{\delta}(t)],E[B|N_{\delta}(t-k\delta)]\right)$$
(20)

$$= \operatorname{Cov}\left(N_{\delta}(t)\sigma^{2}, N_{\delta}(t-k\delta)\sigma^{2}\right)$$
(21)

$$= \sigma^4 \text{Cov} \left( N_{\delta}(t), N_{\delta}(t-k\delta) \right).$$
(22)

$$=\sigma^4\gamma_{N_{\delta},k}.$$
(23)

Theorem 1 shows that the persistence of squared returns is directly related to the persistence of the counting process. An immediate implication follows.

**LEMMA 1** Given a Lévy counting process, N(t), i.e. a renewal process, a Gaussian mixture process  $r_{\delta}(t)$  as specified in Equations (7) – (10) does not exhibit serial correlation among its squared values.

It follows from Lemma 1 that the compound Poisson process will not exhibit volatility persistence. Further, any subordinated Brownian motion with a Lévy directing process, and its corresponding Gaussian mixture model, will not generate volatility dynamics. On the other hand, the ability of a non-Lévy directing processes to generate volatility dynamics will depend on the dynamics of the counting process. This suggests that the ability of the compound ACD and MSMD models to match volatility dynamics, will depend on the ability of their associated counting processes to match the dynamics of observed transaction counts. The ACD and MSMD models accomplish this by directly modeling serial dependence among inter-trade durations, which propagates to the associated counting processes. While the result stated above is an exact, finite-sample result, Deo et al. (2009) establish conditions under which long memory in durations asymptotically propagates to the associated counting process and clock-time returns within a pure-jump model. However, it is important to note that in our model, serial correlation of durations is sufficient but not necessary to generate the appropriate serial correlation in volatility via the counting process; it is possible for trade-count persistence to exist in the presence of independent durations.

### 5 Estimation and Results

Using the non-event E-mini returns data described in Sections 2 and 3, we estimate the parameters of the component distributions in Equation (11) and simulate from the mixture model. In particular, we first estimate the duration models described in Section 4, which we then compound with an estimated Gaussian distribution for trade-time returns to synthesize a distribution for clock-time returns. Using Monte Carlo approximations for the distribution of clock-time returns, we evaluate the candidate models using several measures of goodness-of-fit and distributional distance.

#### 5.1 Estimation

Since the observed inter-trade durations are not a contiguous sequence of data, we estimate each of the models on a daily basis, using the 1000 seconds of post 10:00 am data on the days without news announcements at that time. The result is 58 sets of estimates, whose medians and standard deviations are reported in Table 1. As a matter of comparison, we also report the single point estimates for each model using the full sample of data as a single contiguous sequence of observations. The following subsections provide details on estimation for each of the model components.

#### 5.1.1 Poisson/Exponential

The assumption of Poisson-distributed trade arrivals with mean  $\gamma$  corresponds to duration times that are Exponentially distributed with mean  $\nu = \frac{1}{\gamma}$ . It is trivial to show that the maximum likelihood estimate of the Exponential mean is

$$\hat{\nu} = \frac{1}{\hat{\gamma}} = \frac{1}{n} \sum_{i=1}^{n} d_i,$$
(24)

Model	Parameter	Median	Std. Dev.	Full Sample
Exponential	$\gamma$	5.13e-03	1.95e-03	5.25e-03
MSMD	λ	9.84e-02	1.49e-02	9.44e-02
MSMD	$\gamma_{ar k}$	6.28e-01	1.03e-01	6.40e-01
MSMD	b	5.46e00	7.21e00	5.22e00
MSMD	$m_0$	1.24e-01	2.33e-02	1.30e-01
ACD	ω	2.04e00	5.29e-01	1.95e00
ACD	$\alpha$	2.25e-01	3.00e-02	2.30e-01
ACD	eta	4.37e-01	8.93e-02	4.57e-01
Gaussian	$\mu$	0.00e00	5.58e-04	-3.88e-05
Gaussian	σ	1.18e-01	6.27 e-03	1.18e-01

Table 1: Parameter estimates for duration models and Gaussian trade-time returns distributions.

where  $d_i$ , i = 1, 2, ..., n are the observed inter-trade durations. The median and standard deviation of daily estimates of  $\gamma$  are reported in Table 1.

#### 5.1.2 MSMD

Following Chen et al. (2013) and Zikes et al. (2014), we evaluate the likelihood of the MSMD model, associated with Equations (15a) – (15f), using the nonlinear filtering method of Hamilton (1989) and maximize the likelihood with a standard hill-climbing algorithm. To estimate all parameters of the MSMD model, we iterate over candidate values of  $\bar{k}$  and estimate the remaining four parameters,  $\lambda$ ,  $\gamma_{\bar{k}}$ , b, and  $m_0$ . Although we do not report the full set of results here, we find that the log likelihood (both the median of daily log likelihood values and the single value for the full sample) is maximized for  $\bar{k} = 8$ , with a plateau that begins at  $\bar{k} = 7$ . For purposes of concision, we adopt  $\bar{k} = 7$  for the remainder of the paper.

#### 5.1.3 ACD

As outlined in Engle and Russell (1998) the likelihood function for the ACD model can be obtained in closed form. We obtain daily and full-sample estimates of the model via maximum likelihood.

#### 5.1.4 Gaussian

Building on the empirical observations of Section 3 and the assumptions of the compound duration model of Equation (7), we treat non-news trade-time returns as i.i.d. Gaussian random variables and estimate their moments via maximum likelihood. In this case, the maximum likelihood estimates of the parameters of the distribution are simply the sample average and standard deviation of trade-time returns. Daily and full sample estimates are reported in Table 1.

#### 5.2 Simulation and Evaluation

We obtain Monte Carlo approximations of clock-time returns distributions using the Gaussian mixture model, expressed in Equation (11). We do this in a hierarchical fashion, first simulating inter-trade durations from the Exponential, MSMD and ACD models, pairing the durations with independent draws of trade-time returns from the estimated Gaussian density, and finally aggregating individual returns within a fixed clock time interval. For each model, we generate a dataset equivalent to the observed data: 58 days, each comprising 1e+06 milliseconds worth of data.

Following the procedure outlined above, we aggregate returns for clock-time intervals  $\delta = \{250, 500, 1000, 5000, 10000, 30000\}$  milliseconds until we obtain  $n = \{232000, 116000, 58000, 11600, 58000, 1914\}$  clock-time returns, respectively, which correspond to the number of observations in the data for those time intervals. The individual simulations under each of the duration models use the same trade-time returns; they only differ in the elapsed time between observations. It is important to mention two adjustments that we make in order to simulate clock time returns. First, since E-mini returns are discrete and only observed at increments of 0.25 points, we simulate trade-time returns from the continuous Gaussian distribution described above and then discretize to the nearest 0.25 increment. For example, a simulated trade-time return of 0.13 would be discretized to 0.25, while a simulated trade-time return of 0.12 would be discretized to zero. Second, we perform

a similar discretization of simulated durations (under all models) by rounding values to the nearest millisecond. Since zero durations are not allowed in our framework, all simulated durations below one millisecond are rounded upward.

As a matter of comparison, we also estimate and simulate clock-time returns under the ACD-GARCH(1,1), suggested by Engle (2000), using ACD durations in the GARCH volatility equation. In particular, following Engle (2000),

$$\frac{r_i}{d_i} = \sigma_i \varepsilon_i \tag{25}$$

$$\sigma_i^2 = \omega + \alpha \varepsilon_{i-1}^2 + \beta \sigma_{i-1}^2 + \gamma d_i^{-1}$$
<sup>(26)</sup>

$$\varepsilon_i \stackrel{i.i.d.}{\sim} WN(0,1).$$
 (27)

Since the compound Poisson process has no ability to generate serial correlation in clock-time return volatility, the ACD-GARCH model is a more reasonable benchmark against which to compare the compound duration model.

Figure 5 depicts daily sample autocorrelation functions for observed inter-trade durations, as well as the daily simulations under each of the duration models. The faint lines represent ACFs for each day and the darker lines represent 0.05, 0.5 and 0.95 quantiles of the distributions of autocorrelations at each lag. By construction, the Exponential durations are independent of each other, and so do not exhibit any serial correlation. The ACD and MSMD models, on the other hand, do a much better job of capturing observed dynamics: the median ACD autocorrelation remains elevated until lag 7 or 8, while that of the MSMD does not fall to zero until roughly lag 15. The empirical counterpart falls rapidly until lag 10-15, at which point it remains elevated at a low level and with very slow decay. Further, the quantiles of the empirical autocorrelations are heavily skewed toward positive values for the full range of lags that are depicted in the plot. This low-level persistence of observed durations is not captured by any of the models, although the MSMD performs slightly better than the ACD in this dimension: its quantiles are skewed towards positive autocorrelation until about lag 40. It is important to note that while the MSMD model appears to have some directional advantage in matching the autocorrelation structure of observed durations, it is certainly much noisier than the ACD model.



Figure 5: Daily sample autocorrelation functions for observed durations (blue) and durations simulated under the estimated Exponential (red), ACD (yellow) and MSMD (green) models. Faint lines represent daily ACFs and darker lines represent 0.05, 0.5 and 0.95 quantiles of the distributions of autocorrelations at each lag. The dotted black lines are 95% Bartlett bounds for the null hypothesis of no serial correlation.

Figure 6 shows Q-Q plots of the clock-time returns simulations for each value of  $\delta$  that we consider. The panels in the first (red) row correspond to the compound Poisson (Exponential) model, the second (yellow) row corresponds to the compound ACD model, the third (green) row corresponds to the compound MSMD model and the fourth (purple) row corresponds to the ACD-GARCH(1,1) model. The panels in the final (blue) row of the figure are a reproduction of the E-mini passive-period clock-time Q-Q plots shown in Figure 2. We note that while the limits of the y-axes are constant within row, they differ across rows: this was necessary in order to capture the full range of quantiles across models and should be considered when comparing the models. It is immediately apparent from the figure that the compound Poisson model does a very poor job of capturing leptokurtosis, except at the lowest values of  $\delta$ , where the tails of the distributions of the

simulated returns are still lighter than those of the data. The compound ACD model only provides a slight improvement. On the other hand, the compound MSMD and ACD-GARCH models exhibit much heavier tails for all time scales – tails that are heavier than those of the data. Interestingly, as  $\delta$  increases, the central quantiles of the ACD-GARCH conform more closely to a Gaussian, with only a very small proportion of extreme quantiles. These extreme tails are most likely the additive result of persistence in both the volatility and duration equations: the simple GARCH(1,1) is known to exhibit leptokurtosis, which now appears to be accentuated in the presence of duration autocorrelation. The compound MSMD, however, exhibits more appropriate (although slightly extreme) tail heaviness, which can be attributed to the fact that the MSMD model is interpreted as a mixture of Exponential distributions. Compared to simple Exponential durations, the persistence of the latent states in the MSMD model generates more variation in inter-trade durations, which leads to a more heterogeneous mixture of Gaussian densities in Equation (11), resulting in a greater degree of leptokurtosis.

Panels (a) and (b) of Table 2 report summary statistics that formalize the degree of leptokurtosis in the simulated returns depicted in Figure 6. Under the null hypothesis of Gaussian clock-time returns, the variance of the sample excess kurtosis,  $\hat{\kappa}$ , is:

$$\sigma_{\hat{\kappa}}^2 = \frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)},\tag{28}$$

where *n* is the number of clock-time returns during a single day (i.e. n = 4000 for  $\delta = 250$  ms). An implication of this result is that  $\sqrt{n}\hat{\kappa} \stackrel{d}{\rightarrow} \mathcal{N}(0, 24)$  under the null hypothesis of Gaussianity. Panel (a) of Table 2 reports the fraction of times  $\hat{\kappa} \in (-z_{1-\alpha/2}\sigma_{\hat{\kappa}}, z_{1-\alpha/2}\sigma_{\hat{\kappa}})$ , where  $z_{1-\alpha/2}$  is the  $(1 - \alpha/2)$  quantile of the standard Normal density. We set  $\alpha = 0.05$ . The values in the table demonstrate that for low values of  $\delta$ , all models and the data uniformly reject the hypothesis of Gaussianity on all days. However, as  $\delta$  increases, the simulated compound Poisson/Exponential returns suddenly switch from near 100% rejection to very low rejection rates, while the compound MSMD and ACD-GARCH(1,1) models attenuate their rejection rates by only a small amount. The only model that reasonably tracks the data in this case is the compound ACD, which moderately attenuates rejection rates for increasing  $\delta$ .

Panel (b) of Table 2 reports the Kullback-Leibler divergences of the empirical distributions



Figure 6: Sample Q-Q plots for simulated clock-time returns under each of the compound duration models considered in this paper as well as the ACD-GARCH(1,1). The bottom row of panels is a reproduction of the E-mini passive-period clock-time Q-Q plots shown in Figure 2. Q-Q plots are shown for returns computed across clock-time intervals  $\delta = \{250, 500, 1000, 5000, 10000, 30000\}$  milliseconds.

of daily Kurtosis values for each model and each clock-time interval  $\delta$ , relative to the theoretical limiting Gaussian density under the null hypothesis for clock-time returns. In order to compute the distance measures, we used a Gaussian kernel smoother to approximate the empirical density of kurtoses. Once again, the compound Poisson/Exponential performs worst, with distances from the limiting distribution that are much too low relative to those observed in the data. In a less extreme manner, the same is true of the compound ACD model. This latter result is interesting in

		δ					
Model	250	500	1000	5000	10000	30000	
	(a) Fraction of Accepted Hypotheses (Daily Kurtosis)						
Compound Exp.	0.00e+00	0.00e+00	5.17e-02	9.48e-01	9.66e-01	9.31e-01	
Compound ACD	0.00e+00	0.00e+00	0.00e+00	6.90e-01	9.48e-01	9.83e-01	
Compound MSMD	0.00e+00	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1.38e-01	
ACD-GARCH	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1.90e-02	7.24e-02	
Data	0.00e+00	0.00e+00	2.24e-01	6.90e-01	7.93e-01	9.48e-01	
(b) Kullback-Leibler Divergence (Daily Kurtosis)							
Compound Exp.	$1.15e{+}02$	$2.33e{+}01$	$3.93e{+}00$	1.06e-01	5.68e-02	4.44e-02	
Compound ACD	$4.63e{+}01$	$1.60e{+}01$	$6.55e{+}00$	4.31e-01	8.85e-02	8.41e-02	
Compound MSMD	7.91e+00	7.20e + 00	5.96e + 00	$4.25e{+}00$	$3.35e{+}00$	1.84e + 00	
ACD-GARCH	6.36e + 00	5.30e + 00	4.78e + 00	2.92e+00	1.89e+00	5.15e-01	
Data	$4.45e{+}00$	3.69e + 00	3.02e + 00	$1.73e{+}00$	$1.10e{+}00$	1.70e-01	
(c) Fraction of Accepted Hypotheses (Ljung-Box)							
Compound Exp.	9.66e-01	9.83e-01	9.83e-01	9.31e-01	9.66e-01	1.00e+00	
Compound ACD	9.31e-01	9.14e-01	9.48e-01	9.66e-01	9.31e-01	9.83e-01	
Compound MSMD	6.90e-02	1.38e-01	2.93e-01	7.41e-01	9.48e-01	9.83e-01	
ACD-GARCH	6.90e-02	2.59e-01	8.10e-01	9.66e-01	$1.00e{+}00$	1.00e+00	
Data	1.72e-01	6.21e-01	8.45e-01	9.48e-01	9.66e-01	9.83e-01	
(d) Kullback-Leibler Divergence (Ljung-Box)							
Compound Exp.	8.66e-02	5.74e-02	1.20e-01	1.52e-01	1.22e-01	4.39e-01	
Compound ACD	4.84e-02	3.51e-02	1.53e-01	1.76e-01	2.38e-01	1.92 e- 01	
Compound MSMD	$3.34e{+}00$	$2.66\mathrm{e}{+00}$	2.03e+00	7.86e-01	8.32e-01	$1.31e{+}00$	
ACD-GARCH	3.38e+00	$1.96\mathrm{e}{+00}$	7.92e-01	4.92e-01	6.43e-01	4.75e-01	
Data	1.78e+00	6.00e-01	3.04e-01	2.05e-01	3.73e-01	3.90e-01	

**Table 2:** Sample statistics under null hypotheses related to the distributions of daily kurtosis and autocorrelation estimates of squared returns. Panels (a) and (c) report fractions of accepted null hypotheses and panels (b) and (d) report Kullback-Leibler distances of empirical distributions of daily estimates from limiting distributions under the null hypotheses. All statistics are computed for each of the compound duration models considered in this paper as well as the ACD-GARCH(1,1) and the passive-period clock-time Emini data, using returns computed across clock-time intervals  $\delta = \{250, 500, 1000, 5000, 10000, 30000\}$ milliseconds. light of the fact that the compound ACD has the best rejection rates in panel (a). The remaining models do a better job of tracking the data, with the ACD-GARCH model exhibiting distances that are slightly more congruent with the data than those of the compound MSMD.

Although we do not show them here, the daily sample autocorrelation functions of returns under each of the models exhibit no autocorrelation and look almost identical to those shown in the bottom panels of Figure 3. One exception is the sharp negative autocorrelation at the first lag: while none of the models captures the negative autocorrelation attributed to bid/offer bounce and mean reversion in finely sampled data (low  $\delta$ ), this dynamic is not explicitly modeled in our framework and is not expected to be present.

Figure 7 shows daily sample autocorrelation functions for squared returns simulated under each of the duration models. The panels of Figure 7 correspond to those of Figure 6. Persistence among the autocorrelations is present in all but the compound Poisson/Exponential model. The variability of the compound MSMD, and ACD-GARCH(1,1) models appears to be much larger than that of the data, and the shape of the autocorrelations for the ACD-GARCH models does not appear to conform as closely to the data for low values of  $\delta$ : in the case of the ACD-GARCH(1,1), the low-lag autocorrelations begin at values that are too extreme and subsequently drop too quickly. The compound ACD model, on the other hand, appears to exhibit both the right variability and shape, relative to what is observed in the data, but for  $\delta \leq 1000$  ms the distribution of autocorrelations is symmetrically displaced around zero by lag 10, while that of the data remains positively displaced for many more lags. This latter feature is captured by the compound MSMD model which, despite excessive variation, has a distribution of autocorrelations with an appropriate displacement for low values of  $\delta$ .

To formalize the measures of daily autocorrelation among models, panels (c) and (d) of Table 2 report hypothesis rejection rates and Kullback-Leibler divergences for daily Ljung-Box statistics computed using squared clock-time returns. The Ljung-Box statistic is defined as

$$Q_l = n(n+2) \sum_{i=1}^{l} \frac{\hat{\rho}_i^2}{n-i},$$

where  $\hat{\rho}_i$  is the sample autocorrelation for lag *i*, *n* is the number of observations in the data (for each day), and *l* is the number of lags over which the statistic is computed. Under the null hypothesis



Figure 7: Sample autocorrelation functions for simulated clock-time squared returns under each of the compound duration models considered in this paper as well as the ACD-GARCH(1,1). The bottom row of panels is a reproduction of the E-mini passive-period clock-time ACFs shown in Figure 4. ACFs are shown for squared returns computed across clock-time intervals  $\delta = \{250, 500, 1000, 5000, 10000, 30000\}$  milliseconds.

that all autocorrelations are jointly zero,  $Q_l \sim \chi^2(l)$ . For the summary statistics reported in Table 2, we set l = 30, but the results are robust to a variety of other choices. Similar to panel (a), panel (c) reports the fraction of times the daily Ljung-Box statistic fall within a one-sided  $(1 - \alpha) \chi^2$  confidence interval:  $\hat{Q}_l \in (0, \chi^2_{1-\alpha}(l))$ , where  $\alpha = 0.05$ . The values in the table show that the compound Poisson/Exponential and compound ACD models very rarely reject the null hypothesis, unlike the data which frequently rejects for low values of  $\delta$ . The remaining models display rejections rates that are similar to the data, with the ACD-GARCH(1,1) performing slightly better for intermediate values of  $\delta$  but the compound MSMD model performing best for both low and high values of  $\delta$ .

Panel (d) of Table 2 reports Kullback-Leibler divergences for the empirical densities of daily Ljung-Box statistics relative to the theoretical  $\chi^2$  distribution under the null hypothesis of no serial correlation in squared clock-time returns. As with the rejection rates, the compound Poisson/Exponential and compound ACD models exhibit divergences that most different from those in the data – almost uniformly too low across values of  $\delta$ . The remaining models conform more closely to the divergences observed in the data, with the ACD-GARCH model performing better than the compound MSMD.

#### 5.3 Discussion

The foregoing results highlight the strengths of compound duration models in the presence of a time deformation process that exhibits serial correlation. While no model performs uniformly better than the rest, the compound MSMD durations achieve a better balance in terms of matching both the tail heaviness and volatility clustering that is observed in the data. Several observations are in order.

First, the empirical results corroborate the finding of Theorem 1, which shows that the lack of dependence in the compound Poisson trade arrival process renders it unable to deliver serial correlation in volatility. On the other hand, the ability of the remaining models to capture volatility persistence is directly related to the persistence of its inter-trade durations: the autocorrelation structure observed in Figure 5 immediately propagates to Figure 7. Notably, the manner in which the compound MSMD model exhibits a low level of volatility persistence for many lags when  $\delta \leq 1000$  ms is directly related to the low-level persistence of MSMD durations shown in Figure 5.

Second, we see that the compound Poisson model does not exhibit an appropriate degree of leptokurtosis. While a more flexible distributional assumption for the trade arrival process might rectify this deficiency, volatility dynamics will not be present when the directing process of Equation (3) is Lévy. In particular, although the variance gamma models of Madan and Seneta (1990), Madan et al. (1998) and Carr et al. (2002) are much better at capturing extreme events, the nature of their directing processes precludes serial correlation in volatility.

Third, although the ACD-GARCH model does an acceptable job of capturing volatility dynamics, it exhibits excessive leptokurtosis relative to the data for all horizons. We attribute this to the fact that the ACD-GARCH compounds two mechanisms for generating tail-heaviness: first, it models trade-time returns as a GARCH process and second, it allows the GARCH to depend on autoregressive durations. In contrast, our empirical results in Section 3 suggest that trade-time returns conform much more closely to a Gaussian random walk at fine time scales.

Fourth, despite empirical advances in measuring the leverage effect in high-frequency data (Aït-Sahalia et al. (2013)) and theoretical advances in explaining the leverage effect via time deformation (Carr and Wu (2004)), we do not attempt to explain the leverage effect in this paper. Our model treats the distributions of trade-time returns and trade arrival as independent, which, as suggested by Renault and Werker (2011), is an oversimplification. This is a simplification we adopt in order to focus on other influences that the trading process has on the distribution of asset returns and their dynamics. Future research could generalize our model to jointly model trade-time returns and durations and to incorporate the leverage effect. One method would be to require the baseline MSMD intensity parameter  $\lambda$  to depend on past or concurrent returns, or other joint factors.

Fifth, a significant portion of the financial econometrics literature works directly with volatility as a time deformation device. Madan and Seneta (1990), Andersen et al. (2007), Andersen et al. (2010) and Todorov and Tauchen (2014) are examples. We view both time-varying inter-trade duration and time-varying volatility as outgrowths of a latent time deforming information process. However, we work directly with durations because they are more primitive to the trading process and because our empirical work suggests that this might be the most direct way to model time deformation. We thus allow volatility to be a direct result of inter-trade duration, which is our surrogate for a time deformation device.

# 6 Realized Volatility Forecast Performance

Each of the models proposed in Section 4 and evaluated in Section 5 has a differing number of parameters: the compound Poisson has 2, the compound ACD as 5, the compound MSMD has 7 and the ACD-GARCH(1,1) has 4. To deal with the potential issue of overfitting, and as an alternate method of comparison, we evaluate each model's performance in out-of-sample realized volatility (RV) forecasting.

We begin with a new data set – all E-mini transactions during the period 1 July 2014 - 31 December 2014, which comprises all transactions in the six months subsequent to the data considered in the previous sections of this paper. Using the new data, we propose the following forecasting procedure: (1) for each day that the E-mini traded and for each model under consideration, use all transactions from 2:15 - 3:00 p.m. to estimate the parameters of the model and to simulate one hour of transactions/returns; (2) compute 1-minute and 5-minute realized volatilities for E-mini returns between 3:00 and 4:00 p.m. The overall objective is to utilize a contiguous segment of transactions each day that are not affected by news announcements in order to estimate and forecast. Although a number of news releases are regularly scheduled for 2:00 p.m. Eastern Time, the time span we use for fitting and forecasting is free of macroeconomic news announcements during the test period.

Panel (a) of Table 6 reports summary quantiles for the empirical distributions of the realized volatilities. There are a total of 110 days in the test period.  $\delta$ -minute realized volatility is defined as

$$RV(\delta) = \sum_{j=1}^{T/\delta} r_{\delta}^2(t+j\delta), \qquad (29)$$

where T is the total amount of time in the sample. In our case, T = 60 minutes and, as a result, all of the estimates in the table correspond to hourly volatility computed under two values of  $\delta$ . As expected, panel (a) shows that hourly volatility is a very noisy quantity. This is both a result of inherent noisiness, as well as estimation error. Although we have a wealth of data (thousands of transactions for each day, during the time window under consideration), we cannot simply improve our RV estimates by reducing  $\delta$  and sampling the data more finely: microstructure effects introduce a bias that render this infeasible (Andersen et al. (1999)). Much of the empirical RV literature has found 5-minute RV to be a suitable, bias-free quantity (Patton and Sheppard (2009)); we also report 1-minute RV since it allows the use of more, albeit potentially bias affected, data.

(a) Realized Volatility Empirical Quantiles (%)								
Time Interval	Min	0.05	Median 0.95		Max			
$1 \min$	0.06556	0.1081	0.1869	0.5174	0.7187			
5 min	0.04886	0.07867	0.1665	0.5170	0.9176			
(b) RMSE (%)								
Time Interval	Compound Exp.	Compound ACD	Compound MSMD		ACD-GARCH			
$1 \min$	0.4075	0.2556	0.1519		0.2097			
$5 \min$	0.4230	0.2799	0.1553		0.2316			
(c) Diebold-Mariano Statistics								
	Comp. Exp.	Comp. ACD	Comp. MSI	MD	ACD-GARCH			
Comp. Exp.	NA	7.35	8.88		7.01			
Comp. ACD	7.12	NA	5.53		2.38			
Comp. MSMD	9.75	14.41	NA		-4.89			
ACD-GARCH	20.54	2.20	-2.42		NA			

Table 3: Out of sample realized volatility (RV) forecasting results. Panel (a) reports quantiles of the empirical distributions of 1-minute and 5-minute RV. Panel (b) reports root-mean-square errors of 1-minute and 5-minute RV under each model. Panel (c) reports two sets of pairwise Diebold-Mariano statistics under the null hypothesis of equal prediction accuracy among models – the values above the diagonal assume no serial correlation in forecast errors and those below the diagonal truncate the error ACFs at lag k = 30.

Panel (b) of the table reports root-mean-square errors (RMSE) for both 1-minute and 5-minute RV forecasts under each model. The results show that the compound MSMD model readily beats each of the others in terms of out-of-sample RV forecasting, despite the fact that it has more parameters. Not surprisingly, the compound Poisson/Exponential performs worst. It is interesting, however, that among the two remaining models the ACD-GARCH(1,1) exhibits much better forecast performance than the compound ACD. While the in-sample comparison was not decisively better for either model, the out-of-sample performance suggests the relatively good in-sample results for the compound ACD might be a simple artifact of overfitting.

Panel (c) of the table reports two sets of Diebold and Mariano (1995) statistics. The statistics proposed by Diebold and Mariano (1995) test the null hypothesis of equal predictive accuracy for competing forecasts, even in the presence of contemporaneous and serial correlation. Since our forecasts are separated by periods of time that are large relative to the forecast time scale, it is not expected that serial correlation is an issue. We corroborate this fact with the empirical autocorrelation functions, although we do not report them here for considerations for space. As such, we compute Diebold-Mariano statistics both for the case of no serial correlation among forecasts, as well as a rather long truncation lag of k = 30 for the ACF. The former statistics are reported above the diagonal in panel (c) and the latter are reported below the diagonal. In all cases, the test statistics are significant at the 5% level or better and demonstrate that the differences in forecast performance reported in panel (b) are not a result of sampling error.

Our conclusion is that the compound MSMD model is significantly better at explaining realized volatility out of sample. As with the in-sample fitting, this result suggests that the compound MSMD model provides a fairly good representation of the stochastic process that drives asset price dynamics.

# 7 Conclusion

The empirical and theoretical work of this paper is intimately linked to the work of Mandelbrot (1963), Clark (1973), Brada et al. (1966), Mandelbrot and Taylor (1967) and Ane and Geman (2000), which show that fat-tailed returns distributions are consistent with a Gaussian random walk directed by an appropriate stochastic process. Our empirical insight is that after controlling for pre-scheduled, market-wide news announcements, the subordinating process for a highly liquid market aggregate (the E-Mini S&P 500 near-month futures contract) is simply characterized by a model of high-frequency trade arrival. Our theoretical contribution is to compose a parsimonious inter-trade duration model that serves as an appropriate directing process with a Gaussian random

walk to arrive at a hierarchical model of returns in clock time and to show when this model can simultaneously explain the dynamics and unconditional nature of observed returns. Returns in the tail of the distribution arise as a consequence of faster random walks generated by periods where the trading rate is high and volatility persistence is generated by serial correlation within the trade arrival process. The upshot is that outside of pre-scheduled news-affected periods, the observed non-Gaussianity in the E-mini returns distribution can be fully attributed to the temporal clustering of trades.

Further work appears warranted. While the compound Markov Switching Multifractal Duration model provides the best in-sample and out-of-sample fit compared to other models we consider, it falls far short of explaining the low-level of persistence exhibited in observed inter-trade durations. To the extent that trade-time returns abide by a random walk (as suggested by our empirical work), improvements in modeling the directing process would yield excellent dividends in explaining returns dynamics. Further, while we have treated durations and returns as mutually independent processes, modeling their joint distribution could potentially enhance the model.

Finally, it is possible that the Gaussian spectrum of trade-time returns is a feature that is largely specific to the heavily traded E-mini. Expanding the work of this paper to a broader set of assets could lead to important innovations to the model. Additionally, it would be useful to adapt the model to explain the evolution of returns under conditions of market stress.

# References

- Aït-Sahalia, Y., Fan, J., and Li, Y. (2013), "The leverage effect puzzle: Disentangling sources of bias at high frequency," *Journal of Financial Economics*, 109, 224–249.
- Aldrich, E. M. (2013), "Trading Volume in General Equilibrium with Complete Markets," Working Paper.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (1999), "Realized Volatility and Correlation," *Working Paper*.
- Andersen, T. G., Bollerslev, T., and Dobrev, D. (2007), "No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications," *Journal of Econometrics*, 138, 125–180.
- Andersen, T. G., Bollerslev, T. I. M., and Frederiksen, P. E. R. (2010), "Continuous-time Models, Realized Volatilities, and Testable Distributional Implications for Daily Stock Returns," *Journal* of Applied Econometrics, 261, 233–261.
- Ane, T. and Geman, H. (2000), "Order Flow, Transaction Clock, and Normality of Asset Returns," *The Journal of Finance*, LV, 2259–2284.
- Bachelier, L. (1900), Théorie de la Spéculation, Paris: Gauthier-Villars.
- Black, F. and Scholes, M. (1973), "The Pricing of Options and Corporate Liabilities," Journal of Political Economy, 81, 637–654.
- Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327.
- Bouchaud, J.-P. (2005), "The subtle nature of financial random walks." Chaos, 15, 26104.
- Brada, J., Ernst, H., and van Tassel, J. (1966), "The Distribution of Stock Price Differences: Gaussian after All?" *Operations Research*, 14, 334–340.
- Brogaard, J., Hendershott, T., and Riordan, R. (2013), "High Frequency Trading and Price Discovery," Working Paper.
- Calvet, L. and Fisher, A. (2001), "Forecasting multifractal volatility," Journal of Econometrics, 105, 27–58.
- (2002), "Multifracticality in Asset Returns: Theory and Evidence," The Review of Economics and Statistics, 84, 381–406.
- Calvet, L., Fisher, A., and Mandelbrot, B. (1997), "Large Deviations and the Distribution of Price Changes," Cowles Foundation Discussion Paper No. 1165.
- Calvet, L. E. (2004), "How to Forecast Long-Run Volatility: Regime Switching and the Estimation of Multifractal Processes," *Journal of Financial Econometrics*, 2, 49–83.
- Calvet, L. E. and Fisher, A. J. (2008), Multifractal Volatility: Theory, Forecasting, and Pricing, Burlington, MA: Academic Press.
- (2012), "Extreme Risk and Fractal Regularity in Finance," SSRN Electronic Journal.

- Carr, P., Geman, H., Madan, D. B., and Yor, M. (2002), "The Fine Structure of Asset Returns: An Empirical Investigation," *The Journal of Business*, 75, 305–333.
- Carr, P. and Wu, L. (2004), "Time-changed Lévy processes and option pricing," Journal of Financial Economics, 71, 113–141.
- Chen, F., Diebold, F. X., and Schorfheide, F. (2013), "A Markov-switching multifractal inter-trade duration model, with application to US equities," *Journal of Econometrics*, 177, 320–342.
- Clark, P. K. (1973), "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41, 135–155.
- Deo, R., Hurvich, C. M., Soulier, P., and Wang, Y. (2009), "Conditions for the Propagation of Memory Parameter From Durations To Counts and Realized Volatility," *Econometric Theory*, 25, 764.
- Diebold, F. X. and Mariano, R. S. (1995), "Comparing Predictive Accuracy," Journal of Business & Economic Statistics, 13, 134–144.
- Easley, D., de Prado, M. M. L., and O'Hara, M. (2012), "The Volume Clock: Insights into the High Frequency Paradigm," *Journal of Portfolio Management*.
- Engle, R. F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987–1008.
- (2000), "The Econometrics of Ultra-High-Frequency Data," *Econometrica*, 68.
- Engle, R. F. and Russell, J. R. (1998), "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data," *Econometrica*, 66, 1127–1162.
- Epps, T. W. and Epps, M. L. (1976), "The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture-of-Distributions Hypothes," *Econometrica*, 44, 305–321.
- Fama, E. F. (1965), "The Behavior of Stock-Market Prices," The Journal of Business, 38, 34–105.
- Fisher, A., Calvet, L., and Mandelbrot, B. (1997), "Multifractality of Deutschemark / US Dollar Exchange Rates," Cowles Foundation Discussion Paper No. 1165.
- Gallant, A. R., Rossi, P. E., and Tauchen, G. (1992), "Stock Prices and Volume," The Review of Financial Studies, 5, 199–242.
- Ghysels, E., Gouriéroux, C., and Jasiak, J. (1995), "Market Time and Asset Price Movements: Theory and Estimation," *CIRANO Wokring Paper No. 95s-32*.
- Grammig, J. and Maurer, K.-O. (2000), "Non-monotonic hazard functions and the autoregressive conditional duration model," *Econometrics Journal*, 3, 16–38.
- Hamilton, J. D. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57, 357–384.
- Hasbrouck, J. (2003), "Intraday Price Formation in U.S. Equity Index Markets," The Journal of Finance, LVIII, 2375–2399.

- Hasbrouck, J. and Saar, G. (2013), "Low-latency trading," *Journal of Financial Markets*, 16, 646–679.
- Jones, C. M., Kaul, G., and Lipson, M. L. (1994), "Transactions, Volume, and Volatility," The Review of Financial Studies, 7, 631–651.
- Karpoff, J. M. (1987), "The Relation Between Price Changes and Trading Volume: A Survey," *The Journal of Financial and Quantitative Analysis*, 22, 109–126.
- Kullback, S. and Leibler, R. A. (1951), "On Information and Sufficiency," The Annals of Mathematical Statistics, 22, 79–86.
- Laughlin, G., Aguirre, A., and Grundfest, J. (2014), "Information Transmission between Financial Markets in Chicago and New York," *The Financial Review*, 49, 283–312.
- Lunde, A. (1999), "A Generalized Gamma Autoregressive Conditional Duration Model," Working Paper, 1–30.
- Madan, D. B., Carr, P. P., and Chang, E. C. (1998), "The Variance Gamma Process and Option Pricing," *European Finance Review*, 2, 79–105.
- Madan, D. B. and Seneta, E. (1990), "The Variance Gamma (V.G.) Model for Share Market Returns," *The Journal of Business*, 63, 511–524.
- Mandelbrot, B. (1963), "The Variation of Certain Speculative Prices," The Journal of Business, 36, 394–419.
- Mandelbrot, B., Fisher, A., and Calvet, L. (1997), "A Multifractal Model of Asset Returns," Cowles Foundation Discussion Paper No. 1164.
- Mandelbrot, B. and Taylor, H. M. (1967), "On the Distribution of Stock Price Differences," Operations Research, 15, 1057–1062.
- Mandelbrot, B. B. (1973), "Comments on: "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," by Peter K. Clark," *Econometrica*, 41, 157–159.
- Patton, A. J. and Sheppard, K. (2009), "Optimal combinations of realised volatility estimators," International Journal of Forecasting, 25, 218–238.
- Press, S. J. (1967), "A Compound Events Model for Security Prices," *The Journal of Business*, 40, 317–335.
- (1968), "A Modified Compound Poisson Process with Normal Compounding," Journal of the American Statistical Association, 63, 607–613.
- Renault, E. and Werker, B. J. (2011), "Causality effects in return volatility measures with random times," *Journal of Econometrics*, 160, 272–279.
- Tauchen, G., Zhang, H., and Liu, M. (1996), "Volume, volatility, and leverage: A dynamic analysis," Journal of Econometrics, 74, 177–208.
- Tauchen, G. E. and Pitts, M. (1983), "The Price Variability-Volume Relationship on Speculative Markets," *Econometrica*, 51, 485.

- Todorov, V. and Tauchen, G. (2014), "Limit Theorems for the Empirical Distribution Function of Scaled Increments of Ito Semimartingales at High Frequencies," *The Annals of Applied Probability*, 24, 1850–1888.
- Zikes, F., Barunik, J., and Shenai, N. (2014), "Modeling and Forecasting Persistent Financial Durations," *Working Paper*.