Journal of Empirical Finance 39 (2016) 105–128

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

A compound duration model for high-frequency asset returns[☆]

Eric M. Aldrich[a,*], Indra Heckenbach[b], Gregory Laughlin[c]

[a] *Department of Economics, University of California, Santa Cruz, United States*
[b] *Department of Physics, University of California, Santa Cruz, United States*
[c] *Department of Astronomy, Yale University, United States*

ARTICLE INFO

JEL classification:
C22
C41
C58
G12
G14
G17

Keywords:
High-frequency trading
US equities
News arrival

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 for 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.

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., 2014; 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

---

[☆] This work was supported by the Hellman Fellows Fund.

[*] Corresponding author.

E-mail addresses: ealdrich@ucsc.edu (E.M. Aldrich), iheckenb@ucsc.edu (I. Heckenbach), greg.laughlin@yale.edu (G. Laughlin).

http://dx.doi.org/10.1016/j.jempfin.2016.10.003
Received 9 April 2016; Received in revised form 1 September 2016; Accepted 9 October 2016

the context of high-frequency trading (over time scales ranging from milliseconds to minutes). We analyze a full year 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: during our data sample period, January–December 2014, 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 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 very weak volatility clustering when considered unconditionally throughout the day. However, when filtering price jumps by excluding the periods surrounding pre-scheduled news events, we confirm the existence of trade-time Gaussianity as well as a lack of volatility persistence. It is important to acknowledge that while removing jumps via news filtering is appropriate during our 2014 sample period, it may not be suitable for more volatile market periods outside of that time span.

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, 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 forward 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 clock-time 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, 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 ACD and 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) near-month 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 equities markets as a whole. This is attributed to its liquidity and the relationship of price formation