Fractals in Trade Duration: Capturing Long-Range Dependence and Heavy Tailedness in Modeling Trade Duration

Wei Sun
Institute for Statistics and Mathematical Economics
University of Karlsruhe, Germany

Svetlozar Rachev*
Institute for Statistics and Mathematical Economics
University of Karlsruhe, Germany
and
Department of Statistics and Applied Probability
University of California, Santa Barbara, USA

Frank J. Fabozzi
Yale School of Management, USA

Petko S. Kalev
Department of Accounting and Finance
Monash University, Australia

*Corresponding author: Prof. Dr. S. Rachev, Institute for Statistics and Mathematical Economics, Uni-Karlsruhe, Postfach 6980, 76128 Karlsruhe, Germany. Email: rachev@statistik.uni-karlsruhe.de.

S. Rachev's research was supported by grants from Division of Mathematical, Life and Physical Science, College of Letters and Science, University of California, Santa Barbara, and the Deutschen Forschungsgemeinschaft. W. Sun's research was supported by grants from the Deutschen Forschungsgemeinschaft. P.S. Kalev's research was supported with a NCG grant from the Faculty of Business and Economics, Monash University. Data are supplied by Securities Industry Research Center of Asia-Pacific (SIRCA) on behalf of Reuters.

Fractals in Trade Duration: Capturing Long-Range Dependence and Heavy Tailedness in Modeling Trade Duration

Abstract

Several studies that have investigated a few stocks have found that the spacing between consecutive financial transactions (referred to as trade duration) tend to exhibit long-range dependence, heavy tailedness, and clustering. In this study, we empirically investigate whether a larger sample of stocks exhibit those characteristics. Based on the modeling mechanism of self-similar processes, this paper empirically compares stable distribution and fractional stable noise with several alternative distributional assumptions (lognormal distribution, fractional Gaussian noise, exponential distribution, and Weibull distribution) in modeling trade duration data. The empirical results suggest that fractional stable noise and stable distribution dominate these alternative distributional assumptions. Comparing goodness of fit in modeling trade duration data for stable distribution and fractional stable noise based on a procedure using bootstrap methods developed by the authors, this paper finds that empirically the autoregressive conditional duration model with stable distribution fits better than other combinations, while fractional stable noise itself fits better for the time series of trade duration.

Key Words: fractional stable noise, point processes, self-similarity, stable distribution, trade duration

JEL Classification: C41, G14

1. Introduction

There is considerable interest in the information content and implications of the spacing between consecutive financial transactions (referred to as trade duration) for trading strategies and intra-day risk management. Market microstructure theory, supported by empirical evidence, suggests that the spacing between trades be treated as a variable to be explained or predicted since time carries information and closely correlates with price volatility (see, Bauwens and Veredas (2004), Diamond and Varrecchia (1987), Engle (2000), Engle and Russell (1998), Hasbrouck (1996), and O'Hara (1995)). Manganelli (2005) finds that returns and volatility directly interact with trade duration and trade order size. Trade durations tend to exhibit long-range dependence, heavy tailedness, and clustering (see, Bauwens and Giot (2000), Dufour and Engle (2000), Engle and Russell (1998), and Jasiak (1998)).

These findings raise two questions that we address in this paper:

1. Can single stochastic processes which capture long-range dependence and heavy tailedness be used in modeling trade duration data ?

2. Can a relatively "powerful" distributional assumption in a relatively "simple" functional structure be used for efficient modeling of trade duration data?

It is necessary to treat long-range dependence, heavy tailedness, and clustering simultaneously in order to obtain more accurate predictions. Rachev and Mittnik (2000) note that for modeling financial data, not only does model structure play an important role, but distributional assumptions influence the modeling accuracy. The stable Paretian distribution[1] can be used to capture characteristics of trade duration since it is rich enough to encompass those stylized facts in such data, such as non-Guassian, heavy tails, long-range dependence, and clustering. Other researchers have shown the advantages of stable distributions in financial modeling (see, Fama (1963), Mittnik and Rachev (1993a, 1993b), Rachev (2003), and Rachev *et al.* (2005)). Meanwhile several studies have reported that long-range dependence, self-similar processes, and stable distribution are very closely related (see, Samorodnitsky and Taqqu (1994), Rachev and Mittnik (2000), Rachev and Samorodnitsky (2001), Doukhan *et al.* (2003), Racheva and Samorodnitsky (2003), and Rachev *et al.* (2005)).

The Hurst index is also used to model long-range dependence, see Hurst (1951,1955). It quantifies the degree of long-range dependence and measures the self-similarity scaling. Fortunately, one type of self-similar process can possess the Hurst index and stable distribution together (i.e., can capture both long-range dependence and heavy tailedness). This kind of stochastic process is a fractional stable noise generated from fractional stable motion (see, Samorodnitsky and Taqqu (1994)). Therefore, single stochastic processes can capture long-range dependence and heavy tailedness, answering the first question posed above. Based on estimating intensity of point processes, an autoregressive conditional duration (ACD) model is proposed by Engle and Russell (1998) for modeling trade duration with intertemporal correlation. The ACD model is a joint approach combining transition analysis and Engle's

[1] To distinguish between Gaussian and non-Gaussian stable distribution, the latter is usually named stable Paretian distribution or Lévy stable distribution. Referring to it as a stable Paretian distribution highlights the fact that the tails of the non-Gaussian stable density have Pareto power-type decay and Lévy stable is the recognition of pioneering works done by Paul Lévy to the characterization of non-Gaussian stable laws (see Rachev and Mittnik (2000)).

autoregressive conditional heteroscedasticity (ARCH) model. The motivation behind the ACD and the ARCH models is that in financial market events tend to occur in clusters. If fractional stable noise can be subordinated into the functional structure of the ACD model, then the second question can be answered.

In order to answer the two questions posed above, this paper introduces fractional stable noise as the single stochastic processes to model trade duration. In the empirical analysis of this paper, fractional stable noise is subordinated to the ACD model to model trade duration. As to self-similar processes, other single stochastic processes, such as fractional Gaussian noise which captures long-range dependence, are also is introduced as an alternative. Since the stable distribution itself can capture heavy tailedness and long-range dependence, we propose it as an alternative distribution that can better explain trade duration. In the empirical analysis, stable distribution is also subordinated to the ACD model. Some other distributions that are often used in modeling trade duration, such as lognormal distribution, exponential distribution and Weibull distribution, have been selected as alternative distributional assumptions in order to compare goodness of fit with the stable distribution and fractional stable noise. Utilizing two test statistics usually used to evaluate model performance under heavy-tailed assumptions, we examine trade duration for a sample of stocks to compare which distributional assumption fits better. By applying a newly developed test procedure that we formulate, based on a bootstrap method, we obtain empirical results that suggest the fractional stable noise and stable distribution dominate these alternative assumptions with high statistical significance. Comparing goodness of fit in the modeling of trade duration data for stable distribution and fractional stable noise, the empirical results indicate that the ACD model with stable distribution fits better than other combinations, while fractional stable noise itself fits better for the time series of trade duration.

The paper is organized as follows. A brief review of point processes and several trade duration models based on estimating intensity of such processes is provided in Section 2. In Section 3, we introduce two self-similar processes: fractional Gaussian noise and fractional stable noise. The method for estimating the parameters in the underlying process is introduced in Section 4. In Section 5, the methods of simulating fractional Gaussian noise and fractional stable noise are introduced. The empirical study based on trade duration data for 18 of the component stocks of the Dow Jones is reported in Section 6. In that section, we compare the goodness of fit of model under fractional stable noise and stable distribution together with distributions (the lognormal distribution, fractional Gaussian noise, exponential distribution, and Weibull distribution). We summarize our conclusions in Section 7.

2. Point processes in modeling durations

Given a probability space $(\Omega, \mathcal{A}, \mathrm{P})$, a family of random variables $(X_t)_{t\in\mathcal{T}}$ on Ω with values in some set M, (i.e., for all $t \in \mathcal{T}$ and $\mathcal{T}$ is some index set), $X_t : (\Omega, \mathcal{A}) \rightarrow (\mathrm{M}, \mathcal{B})$ is defined as a stochastic process with index set $\mathcal{T}$ and state space M. A sequence $(T_n)_{n\in N}$ of positive real random variables is a point process if $T_n(\omega) < T_{n+1}(\omega)$ for all $\omega \in \Omega$ and all $n \in N$; and $\lim_{n\rightarrow\infty} T_n(\omega) = \infty$, for all $\omega \in \Omega$. T_n is called the nth arrival time and $T_n = \sum_{i=1}^{n} \tau_i$; and $\tau_n = T_n - T_{n-1}$ (where $\tau_1 = T_1$) is called the nth *waiting time* (*duration*) for the *point process*.

A stochastic process $(N_t)_{t\in[0,\infty)}$ is a *counting process* if: $N_t : (\Omega, \mathcal{A}) \rightarrow (\mathrm{N}_0, \mathrm{P}(\mathrm{N}_0))$ for all $t \geq 0$,

$N_0 \equiv 0$; $N_s(\omega) \le N_t(\omega)$, for all $0 \le s < t$ and all $\omega \in \Omega$; $\lim_{s\to t, s>t} N_s(\omega) = N_t(\omega)$, for all $t \ge 0$ and all $\omega \in \Omega$; $N_t(\omega) - \lim_{s\to t, s>t} N_s(\omega) \in (0,1)$, for all $t > 0$ and all $\omega \in \Omega$; and $\lim_{t\to\infty} N_t(\omega) = \infty$, for all $\omega \in \Omega$. A point process $(T_n)_{n\in N}$ corresponds to a counting process $(N_t)_{t\in[0,\infty)}$ and vice versa, i.e.,

$$N_t(\omega) = |\{n \in N : T_n(\omega) \le t\}| \tag{1}$$

for all $\omega \in \Omega$ and all $t \ge 0$. For all $\omega \in \Omega$ and all $n \in N$,

$$T_n(\omega) = \min\{t \ge 0 : N_t(\omega) = n\} \tag{2}$$

The mean value function of the counting process is $m(t) = E(N_t)$, for $t \ge 0$, and $m : [0,\infty) \to [0,\infty), m(0) = 0$. m is an increasing and a right continuous function with $\lim_{t\to\infty} m(t) = \infty$. If the mean value function is differentiable at $t > 0$, then the first-order derivative is called *the intensity of the counting process*. Defining $\lambda(t) = \mathrm{d}m(t)/\mathrm{d}t$,

$$\lim_{\Delta t\to 0, \Delta t\ne 0} \frac{1}{|\Delta t|} P\left(|N_{t+\Delta t} - N_t| = 1\right) = \lambda(t) \tag{3}$$

and

$$\lim_{\Delta t\to 0, \Delta t\ne 0} \frac{1}{|\Delta t|} P\left(|N_{t+\Delta t} - N_t| \ge 2\right) = 0 \tag{4}$$

From the viewpoint of the point processes literature (for example, Cox and Isham (1980)), ultra-high frequency financial data can be described as marked point processes; that is, the state space M is a product space of $R^2 \otimes \mathcal{M}$ where $\mathcal{M}$ is the mark space. Engle (2000) pointed out that ultra-high frequency transaction data contain two types of processes: time of transactions and events observed at the time of the transaction. Those events can be identified or described by marks, such as trade prices, posted bid and ask price, and volume. The amount of time between events is the duration. The intensity is used to characterize the point processes and is defined as the expected number of events per time increment considered as a function of time. In survival analysis, the intensity equals the hazard rate. For n durations, $d_1, d_2, ..., d_n$, which are sampled from a population with density function f and corresponding cumulative distribution function F, the survival function $S(t)$ is:

$$S(t) = P[d_i > t] = 1 - F(t) \tag{5}$$

and the intensity or hazard rate $\lambda(t)$ is:

$$\lambda(t) = \lim_{\Delta t\to 0} \frac{P[t < d_i \le t + \Delta t \mid d_i > t]}{\Delta t} \tag{6}$$

The survival function and the density function can be obtained from the intensity,

$$\lambda(t) = \frac{f(t)}{S(t)} = \frac{-\mathrm{d}\log(S(t))}{\mathrm{d}t} \tag{7}$$

Several models have been proposed to model durations by estimating the intensity[2]. The favored models in the literature are the autoregressive conditional duration (ACD) model proposed by Engle and Russell (1998), the stochastic conditional duration (SCD) model by Bauwens and Veredas (2004),

[2]See recent reviews of Bauwens and Hautsch (2006) and Sun *et al.* (2006).

and the stochastic volatility duration (SVD) model by Ghysels *et al.* (2004). The ACD model expresses the conditional expectation of duration as a linear function of past durations and past conditional expectation. The disturbance is specified as an exponential distribution and as an extension of the Weibull distribution. The SCD model assumes that a latent variable drives the movement of durations. Then the expected durations in the SCD model are treated as the observed durations driven by a latent variable. The SVD model tries to capture the mean and variance of durations. The SCD and SVD models are mixed distributions models. The SCD model combines Weibull and gamma distributions while in the SVD model the durations are expressed as independently and exponentially distributed with a gamma heterogeneity. Extensions to these models have been suggested in the literature. Jasiak (1998) offers the fractional integrated ACD model, Gramming and Maurer (1999) replace the Weibull distribution by the Burr distribution, Bauwens and Giot (2000) propose the logarithmic ACD model, and Zhang *et al.* (2001) introduce the threshold ACD model. Bauwens and Giot (2003) propose an asymmetric ACD model; Feng *et al.* (2004) propose a linear non-Gaussian state-space version of the SCD model to capture the leverage effect of the expected durations.

The ACD(m,n) model specified in Engle and Russell (1998) is

$$d_i = \psi_i \, \varepsilon_i \tag{8}$$

$$\psi_i = \omega \, + \sum_{j=0}^{m} \alpha_j \, d_{i-j} \, + \sum_{j=0}^{n} \beta_j \, \psi_{i-j} \tag{9}$$

Bauwens and Giot (2000) give the logarithmic version of the ACD model as follows

$$d_i = e^{\psi_i} \, \varepsilon_i \tag{10}$$

Two possible specifications of conditional durations are

$$\psi_i = \omega \, + \sum_{j=0}^{m} \alpha_j \log d_{i-j} \, + \sum_{j=0}^{n} \beta_j \, \psi_{i-j} \tag{11}$$

and

$$\psi_i = \omega \, + \sum_{j=0}^{m} \alpha_j \log \varepsilon_{i-j} \, + \sum_{j=0}^{n} \beta_j \, \psi_{i-j} \tag{12}$$

Zhang *et al.* (2001) extend the conditional duration to a switching-regime version. Defining $L_q = [l_{q-1}, l_q)$, and $q = 1, 2, ..., Q$ for a positive integer Q, where $-\infty = l_0 < l_1 < ... < l_q = +\infty$ are the threshold values, d_i follows a q-regime threshold ACD (TACD(m,n)) model; that is:

$$\psi_i = \omega^{(q)} \, + \sum_{j=0}^{m} \alpha_j^{(q)} \, d_{i-j} \, + \sum_{j=0}^{n} \beta_j^{(q)} \, \psi_{i-j} \tag{13}$$

For example, if there is a threshold value l_h and $0 < h < q$, the TACD(1,1) model can be expressed as following:

$$\psi_i = \begin{cases} \omega_1 \, + \, \alpha_1 \, d_{i-1} \, + \, \beta_1 \, \psi_{i-1} & if \quad 0 < d_{i-1} \leq l_h \\ \omega_2 \, + \, \alpha_2 \, d_{i-1} \, + \, \beta_2 \, \psi_{i-1} & if \quad l_h < d_{i-1} < \infty \end{cases} \tag{14}$$

The threshold l_h determines the regime boundaries. Fernandes and Gramming (2005) propose the nonparametric tests for ACD models and suggested the practical application for estimation of intraday volatility patterns.

The SCD model given by Bauwens and Veredas (2004) takes the following form:

$$d_i = \Psi_i \, \varepsilon_i \tag{15}$$

where

$$\Psi_i = e^{\psi_i} \tag{16}$$

$$\psi_i = \omega \, + \, \beta \, \psi_{i-1} \, + \, u_i \tag{17}$$

in which $|\beta| < 1$, denotes I_{i-1} the information set before d_i, $u_i|I_{i-1} \sim N(0, \sigma^2)$, $\varepsilon_i|I_{i-1}$ follows some distribution with positive support, and u_i is independent of $\varepsilon_j|I_{i-1}$ for any i and j.

Ghysels *et al.* (2004) proposed a SVD model by assuming that durations are independently and exponentially distributed with gamma heterogeneity. More explicitly, the model can be expressed as:

$$d_i = \frac{U_i}{cV_i} \tag{18}$$

where U_i and V_i are two independent variables with exponential distribution and gamma(a, a) distribution. Then this expression can be transferred with suitable nonlinear transformations to the expression with Gaussian factors:

$$d_i = \frac{\Psi(1, \Phi(F_1))}{c\Psi(a, \Phi(F_2))} \; = \; \frac{H(1, F_1)}{cH(1, F_2)} \tag{19}$$

where F_1 and F_2 are i.i.d standard normal variables, Φ is the cdf of the standard normal distribution, and $\Psi(a, .)$ is the quantile function of the gamma(a, a) distribution.

3. Specification of the self-similar processes

The concept of semi-stable processes (which we today refer to as self-similar processes) was first introduced in Lamperti (1962). Let T be either R, $R_+ = \{t : t \geq 0\}$ or $\{t : t > 0\}$. The real-valued process $\{X(t), t \in T\}$ has stationary increments if $X(t+a) - X(a)$ has the same finite-dimensional distrituions for all $a \geq 0$ and $t \geq 0$. Then the real-valued process $\{X(t), t \in T\}$ is self-similar with exponent of self-similarity H for any $a > 0$, and $d \geq 1$, $t_1, t_2, ..., t_d \in T$, satisfying:

$$\Big(X(at_1), X(at_2), ..., X(at_d)\Big) \stackrel{d}{=} \Big(a^H X(t_1), a^H X(t_2), ...a^H X(t_d)\Big). \tag{20}$$

3.1 Fractional Gaussian noise

For a given $H \in (0, 1)$, there is basically a single Gaussian *H-sssi* [3] process, namely fractional Brownian motion (fBm), first introduced by Kolmogorov (1940). Mandelbrot and Wallis (1968) and Taqqu (2003) define fBm as a Gaussian *H-sssi* process $\{B_H(t)\}_{t \in R}$ with $0 < H < 1$. Mandelbrot and van Ness (1968) define the stochastic representation

$$B_H(t) := \frac{1}{\Gamma(H + \frac{1}{2})} \Big(\int_{-\infty}^{0} [(t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}}] dB(s) + \int_0^t (t-s)^{H-\frac{1}{2}} dB(s) \Big), \tag{21}$$

[3] The abbreviation of "*sssi*" means self-similar stationary increments. If the exponent of self-similarity H is to be emphasized, then "*H-sssi*" is adopted.

where $\Gamma(\cdot)$ represents the gamma function:

$$\Gamma(a) := \int_0^\infty x^{a-1} e^{-x} dx,$$

and $0 < H < 1$ is the Hurst parameter. The integrator B is ordinary Brownian motion. The principal difference between fractional Brownian motion and ordinary Brownian motion is that the increments in Brownian motion are independent while in fractional Brownian motion they are dependent. For fractional Brownian motion, Samorodnitsky and Taqqu (1994) define its increments $\{Y_j, j \in Z\}$ as fractional Gaussian noise (fGn), which is, for $j = 0, \pm 1, \pm 2, ...$, $Y_j = B_H(j-1) - B_H(j)$.

3.2 Fractional stable noise

While fractional Brownian motion can capture the effect of long-range dependence, it has less power to capture heavy tailedness. The existence of abrupt discontinuities in financial data, combined with the empirical observation of sample excess kurtosis and unstable variance, confirms the stable Paretian hypothesis identified by Mandelbrot (1963, 1983). It is natural to introduce the stable Paretian distribution in self-similar processes in order to capture both long-range dependence and heavy tailedness. Samorodinitsky and Taqqu (1994) introduce the α-stable *H-sssi* processes $\{X(t), t \in R\}$ with $0 < \alpha < 2$. If $0 < \alpha < 1$, the exponent of self-similarity are $H \in (0, 1/\alpha]$ and if $1 < \alpha < 2$, the exponent of self-similarity are $H \in (0, 1)$. In addition, Cohen and Samorodnitsky (2006) show that with exponent $H' = 1 + H(1/\alpha - 1)$, process $\{X(t), t \in R\}$ is a well-defined symmetric α-stable ($S\alpha S$) process. It has stationary increments and is self-similar. They show that (1) for $0 < \alpha < 1$, a family of H'*-sssi* $S\alpha S$ processes with $H' \in (1, 1/\alpha)$ is obtained, (2) for $1 < \alpha < 2$, a family of H'*-sssi* $S\alpha S$ processes with $H' \in (1/\alpha, 1)$ is obtained, and (3) for $\alpha = 1$, a family of *1-sssi* $S\alpha S$ processes is obtained.

There are many extensions of fractional Brownian motion to the stable distribution. The most commonly used is linear fractional stable motion (also called linear fractional Lévy motion), $\{L_{\alpha,H}(a, b; t), t \in (-\infty, \infty)\}$, which Samorodinitsky and Taqqu (1994) define as

$$L_{\alpha,H}(a, b; t) := \int_{-\infty}^{\infty} f_{\alpha,H}(a, b; t, x) M(dx), \tag{22}$$

where

$$f_{\alpha,H}(a, b; t, x) := a\Big((t-x)_+^{H-\frac{1}{\alpha}} - (-x)_+^{H-\frac{1}{\alpha}}\Big) + b\Big((t-x)_-^{H-\frac{1}{\alpha}} - (-x)_-^{H-\frac{1}{\alpha}}\Big), \tag{23}$$

and a, b are real constants. $|a| + |b| > 1$, $0 < \alpha < 2$, $0 < H < 1$, $H \neq 1/\alpha$, and M is an α-stable random measure on R with Lebesgue control measure and skewness intensity $\beta(x)$, $x \in (-\infty, \infty)$ satisfying: $\beta(\cdot) = 0$ if $\alpha = 1$. They define linear fractional stable noises expressed by $Y(t)$, and $Y(t) = X_t - X_{t-1}$,

$$\begin{aligned} Y(t) &= L_{\alpha,H}(a, b; t) - L_{\alpha,H}(a, b; t-1) \\ &= \int_R \Big(a\Big[(t-x)_+^{H-\frac{1}{\alpha}} - (t-1-x)_+^{H-\frac{1}{\alpha}}\Big] \\ &+ b\Big[(t-x)_-^{H-\frac{1}{\alpha}} - (t-1-x)_-^{H-\frac{1}{\alpha}}\Big]\Big) M(dx), \end{aligned} \tag{24}$$

where $L_{\alpha,H}(a, b; t)$ is a linear fractional stable motion defined by equations (22) and (23), and M is a stable random measure with Lebesgue control measure given $0 < \alpha < 2$. Samorodinitsky and Taqqu

(1994) show that the kernel $f_{\alpha,H}(a,b;t,x)$ is d-self-similar with $d = H - 1/\alpha$ when $L_{\alpha,H}(a,b;t)$ is $1/\alpha$-self-similar. This implies $H = d + 1/\alpha$ (see Taqqu and Teverovsky (1998) and Weron *et al.* (2005)).[4] In this paper, if there is no special indication, the fractional stable noise (fsn) is generated from a linear fractional stable motion.

4. Estimation in self-similar processes

4.1 Estimating the self-similarity parameter in fractional Gaussian noise

Beran (1994) discusses the Whittle estimator of self-similarity parameter. For fractional Gaussian noise, Y_t, let $f(\lambda; H)$ denote the power spectrum of Y after being normalized to have variance 1 and let $I(\lambda)$ the periodogram of Y_t; that is

$$I(\lambda) = \frac{1}{2\pi N}\left|\sum_{t=1}^{N} Y_t\, e^{i\,t\,\lambda}\right|^2. \tag{25}$$

The Whittle estimator of H is to find $\hat{H}$ that minimizes

$$g(\hat{H}) = \int_{-\pi}^{\pi} \frac{I(\lambda)}{f(\lambda;\hat{H})} d\lambda. \tag{26}$$

4.2 Estimating the self-similarity parameter in fractional stable noise

Stoev *et al.* (2002) proposed the least-squares (LS) estimator of the Hurst index based on the finite impulse response transformation (FIRT) and wavelet transform coefficients of the fractional stable motion. A FIRT is a filter $v = (v_0, v1, ..., v_p)$ of real numbers $v_t \in \Re, t = 1, ..., p$, and length is $p + 1$. It is defined for X_t by

$$T_{n,t} = \sum_{i=0}^{p} v_i\, X_{n(i+t)}, \tag{27}$$

where $n \geq 1$ and $t \in N$. The $T_{n,t}$ are the FIRT coefficients of X_t (i.e., the FIRT coefficients of the fractional stable motion). The indices n and t can be explained as "scale" and "location". If $\sum_{i=0}^{p} i^r v_i = 0$, for $r = 0, ..., q-1$, but $\sum_{i=0}^{p} i^q v_i \neq 0$, the filter v_i can be said to have $q \geq 1$ zero moments. If $\{T_{n,t}, n \geq 1, t \in N\}$ is the FIRT coefficients of fractional stable motion with the filter v_i that has at least one zero moment, Stoev *et al.* (2002) prove the following properties of $T_{n,t}$: (1), $T_{n,t+h} \stackrel{d}{=} T_{n,t}$, and (2), $T_{n,t} \stackrel{d}{=} n^H T_{1,t}$, where $h, t \in N$, $n \geq 1$. We assume that $T_{n,t}$ are available for the fixed scales n_j $j = 1, ..., m$ and locations $t = 0, ..., M_j - 1$ at the scale n_j, since only a finite number, say M_j, of the FIRT coefficients are available at the scale n_j. By using these properties, we have

$$E \log |T_{n_j,0}| = H \log n_j + E \log |T_{1,0}|. \tag{28}$$

The left-hand side of this equation can be approximated by

$$Y_{\log}(M_j) = \frac{1}{M_j} \sum_{t=0}^{M_j-1} \log |T_{n_j,t}|. \tag{29}$$

[4]Some properties of these processes have been discussed in Mandelbrot and Van Ness (1968), Maejima (1983), Maejima and Rachev (1987), Manfields *et al.* (2001), Rachev and Mittnik (2000), Rachev and Samorodnitsky (2001), Samorodnitsky (1994, 1998), Samorodinitsky and Taqqu (1994), and Cohen and Samorodnitsky (2006).

Then we get

$$\begin{pmatrix} Y_{\log}(M_1) \\ \vdots \\ Y_{\log}(M_m) \end{pmatrix} = \begin{pmatrix} \log n_1 & 1 \\ \vdots & \vdots \\ \log n_m & 1 \end{pmatrix} \begin{pmatrix} H \\ E \log |T_{1,0}| \end{pmatrix} + \begin{pmatrix} \sqrt{M}\Big(Y_{\log}(M_1) - E \log |T_{n_1,0}|\Big) \\ \vdots \\ \sqrt{M}\Big(Y_{\log}(M_m) - E \log |T_{n_m,0}|\Big) \end{pmatrix}. \tag{30}$$

We can express above equation as follows

$$Y = X\theta + \frac{1}{\sqrt{M}}\varepsilon, \tag{31}$$

where ε is the vector showing the difference between $\sqrt{M}Y_{\log}(M_m)$ and $\sqrt{M}E(\log |T_{n_m,0}|)$. Equation (31) shows that the self-similarity parameter H can be estimated by a standard linear regression of the vector Y against the matrix X. Stoev *et al.* (2002) show the details for implementing such a procedure.

4.3 Estimating the parameters of the stable Paretian distribution

The stable distribution requires four parameters for complete description: an index of stability $\alpha \in (0, 2]$ also called the tail index, a skewness parameter $\beta \in [-1, 1]$, a scale parameter $\gamma > 0$, and a location parameter $\zeta \in \Re$. There is unfortunately no closed-form expression for the density function and distribution function of a stable distribution. Rachev and Mittnik (2000) give the definition for the stable distribution: A random variable X is said to have a stable distribution if there are parameters $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, $\gamma \geq 0$ and ζ real such that its characteristic function has the following form:

$$E \exp(i\theta X) = \begin{cases} \exp\{-\gamma^{\alpha}|\theta|^{\alpha}(1 - i\beta(\sin\theta)\tan\frac{\pi\alpha}{2}) + i\zeta\theta\}, & if \quad \alpha \neq 1 \\ \exp\{-\gamma|\theta|(1 + i\beta\frac{2}{\pi}(\sin\theta)\ln|\theta|) + i\zeta\theta\}, & if \quad \alpha = 1 \end{cases} \tag{32}$$

and,

$$\text{sign}\theta = \begin{cases} 1, & if \quad \theta > 0 \\ 0, & if \quad \theta = 0 \\ -1, & if \quad \theta < 0 \end{cases} \tag{33}$$

Stable density is not only support for all of $(-\infty, +\infty)$, but also for a half line. For $0 < \alpha < 1$ and $\beta = 1$ or $\beta = -1$, the stable density is only for a half line.

In order to estimate the parameters of the stable distribution, the maximum likelihood estimation (MLE) method given in Rachev and Mittnik (2000) has been employed. Given N observations, $X = (X_1, X_2, \cdots, X_N)'$ for the positive half line the log-likelihood function is of the form

$$\ln(\alpha, \lambda; X) = N \ln \lambda + N \ln \alpha + (\alpha - 1)\sum_{i=1}^{N} \ln X_i - \lambda \sum_{i=1}^{N} X_i^{\alpha}, \tag{34}$$

which can be maximized using, for example, a Newton-Raphson algorithm. It follows from the first-order condition,

$$\lambda = N\left(\sum_{i=1}^{N} X_i^{\alpha}\right)^{-1} \tag{35}$$

that the optimization problem can be reduced to finding the value for α which maximizes the concentrated likelihood

$$\ln^*(\alpha; X) = \ln\alpha + \alpha\nu - \ln\left(\sum_{i=1}^{N} X_i^{\alpha}\right), \tag{36}$$

where $\nu = N^{-1}\Sigma_{i=1}^{N}\ln X_i$. The information matrix evaluated at the maximum likelihood estimates, denoted by $I(\hat{\alpha}, \hat{\lambda})$, is given by

$$I(\hat{\alpha}, \hat{\lambda}) = \begin{pmatrix} N\hat{\alpha}^{-2} & \sum_{i=1}^{N} X_i^{\hat{\alpha}} \ln X_i \\ \sum_{i=1}^{N} X_i^{\hat{\alpha}} \ln X_i & N\hat{\lambda}^{-2} \end{pmatrix}.$$

It can be shown that, under fairly mild condition, the maximum likelihood estimates $\hat{\alpha}$ and $\hat{\lambda}$ are consistent and have asymptotically a multivariate normal distribution with mean $(\alpha, \lambda)'$(see Rachev and Mittnik (2000)).

Other methods for estimating the parameters of a stable distribution (i.e., the method of moments based on the characteristic function, the regression-type method, and the fast Fourier transform method) are discussed in Stoyanov and Racheva-Iotova (2004a, 2004b).

5. Simulation of self-similar processes

5.1 Simulation of fractional Gaussian noise

Paxson (1997) gives a method to generate the fractional Gaussian noise by using the Discrete Fourier Transform of the spectral density. Bardet *et al.* (2003) give a concrete simulation procedure based on this method with respect to alleviating some of the problems faced in practice. The procedure is:

1. Choose an even integer M. Define the vector of the Fourier frequencies $\Omega = (\theta_1, ..., \theta_{M/2})$, where $\theta_t = 2\pi t/M$ and compute the vector $F = f_H(\theta_1), ..., f_H(\theta_{M/2})$, and

$$f_H(\theta) = \frac{1}{\pi}\sin(\pi H)\Gamma(2H+1)(1-\cos\theta)\sum_{t\in\aleph}|2\pi t + \theta|^{-2H-1},$$

   where $f_H(\theta)$ is the spectral density of FGN.

2. Generate $M/2$ i.i.d. exponential (exp(1)) random variables $E_1, ..., E_{M/2}$ and $M/2$ i.i.d. uniform ($U[0,1]$) random variables $U_1, ..., U_{M/2}$.

3. Compute $Z_t = \exp(2i\pi U_t)\sqrt{F_t E_t}$, for $t = 1, ..., M/2$.

4. Form the M-vector: $\tilde{Z} = (0, Z_1, ...Z_{(M/2)-1}, Z_{M/2}, \overline{Z}_{(M/2)-1}, ..., \overline{Z}_1)$.

5. Compute the inverse FFT of the complex Z to obtain the simulated sample path.

5.2 Simulation of fractional stable noise

Replacing the integral in equation (24) with a Riemann sum, Stoev and Taqqu (2004) generate the approximation of fractional stable noise. They introduce parameters $n, N \in \aleph$, and let the fractional stable noise $Y(t)$ expressed as

$$Y_{n,N}(t) := \sum_{j=1}^{nN} \Big((\frac{j}{n})_+^{H-1/\alpha} - (\frac{j}{n} - 1)_+^{H-1/\alpha} \Big) L_{\alpha,n}(nt - j), \tag{37}$$

where $L_{\alpha,n}(t) := M_\alpha((j+1)/n) - M_\alpha(j/n)$, $j \in \Re$. The parameter n is mesh size and the parameter M is the cut-off of the kernel function. Stoev and Taqqu (2004) describe an efficient approximation involving the Fast Fourier Transformation algorithm for $Y_{n,N}(t)$. Consider the moving average process $Z(m)$, $m \in \aleph$,

$$Z(m) := \sum_{j=1}^{nM} g_{H,n}(j) L_\alpha(m - j), \tag{38}$$

where

$$g_{H,n}(j) := \Big((\frac{j}{n})^{H-1/\alpha} - (\frac{j}{n} - 1)_+^{H-1/\alpha} \Big) n^{-1/\alpha}, \tag{39}$$

and where $L_\alpha(j)$ is the series of i.i.d standard stable Paretian random variables. Since $L_{\alpha,n}(j) \stackrel{d}{=} n^{-1/\alpha} L_\alpha(j), j \in \Re$, equations (37) and (38) imply $Y_{n,N}(t) \stackrel{d}{=} Z(nt)$, for $t = 1, ..., T$. Then, the computing of $Y_{n,N}(t)$ is moved to focus on the moving average series $Z(m)$, $m = 1, ..., nT$. Let $\tilde{L}_\alpha(j)$ be the $n(N+T)$-periodic with $\tilde{L}_\alpha(j) := L_\alpha(j)$, for $j = 1, ..., n(N+T)$ and let $\tilde{g}_{H,n}(j) := g_{H,n}(j)$, for $j = 1, ..., nN$; $\tilde{g}_{H,n}(j) := 0$, for $j = nN+1, ..., n(N+T)$. Then

$$\{Z(m)\}_{m=1}^{nT} \stackrel{d}{=} \Big\{ \sum_{j=1}^{n(N+T)} \tilde{g}_{H,n}(j) \tilde{L}_\alpha(n - j) \Big\}_{m=1}^{nT}, \tag{40}$$

because for all $m = 1, ..., nT$, the summation in equation (38) involves only $L_\alpha(j)$ with indices j in the range $-nN \leq j \leq nT - 1$. Using a circular convolution of the two $n(N+T)$-periodic series $\tilde{g}_{H,n}$ and $\tilde{L}_\alpha$ computed by using their Discrete Fourier Transforms (DFT), the variables $Z(n)$, $m = 1, ..., nT$ (i.e., the fractional stable noise), can be generated.

6. Empirical study

In this section, we report the results of empirical tests that investigate the goodness of fit of several candidate distributional assumptions.

6.1 The data

Ultra-high frequency data of 18 Dow Jones index component stocks based on NYSE trading for year 2003 are examined.[5] The companies in the sample are listed in Table 1. Our sample is considerably

[5] The data were provided by The Securities Industry Research Center of Asia-Pacific in Australia. The Dow Jones index consists of 30 stocks. The whole database we developed included data from 1996 to 2003 for stocks that remained in the index over the entire period. Only 18 stocks satisfied that requirement and we use the data of 2003 in this study.

larger than other studies that have investigated trade duration. Table 2 lists those studies and the stocks included in each one. Note that for the studies that include U.S. stocks, IBM is included in 7 of 11 studies and because the sample size is small, IBM constitutes a major part of those studies. IBM is included in our study also.

The trade durations were calculated for regular trading hours (i.e., overnight trading was not considered). Consistent with Engle and Russell (1998) and Ghysels *et al.* (2004), open trades are deleted in order to avoid effects induced by the opening auction. Therefore trade durations only from 10:00 to 16:00 are considered.

Figures 1 to 6 plot several sampled trade duration series. These figures show data characteristics that are consistent with the data patterns reported in the literature. For the trade durations of each stock in our study, sample period runs were performed from 4 January 2003 to 31 December 2003. We will let N denote the length of the sample, sub-sample series that have been randomly selected by a moving window with length T ($1 \leq T \leq N$). Replacement is allowed in the sampling. Stoev and Taqqu (2004) suggest that $2^{14} - 6,000 = 10,384$ is the optimal length for a fractional stable noise series to be simulated efficiently. Therefore, in the empirical analysis, sub-sample length (i.e., the window length) of $T = 10,384$ was chosen. A total of 684 sub-samples were randomly created.

The trade duration data are distributed asymetrically. All observations of duration are positive numbers. The stable distribution, fractional Gaussian noise, and fractional stable are all defined on both positive and negative supports. In our empirical study, we transfer the asymmetrically distributed series to the symmetrically distributed series before we estimate the corresponding parameters. Note that only the positive numbers of the generated series are considered in our simulation.

6.2 Preliminary tests

Table 1 shows the descriptive statistics of the trade duration data in our study. From the statistics reported in this table, it can be seen that excess kurtosis exists.

The Hurst index $H \in (0, 1)$ is the index of self-similarity. For Gaussian processes with stationary increments, when (1) $H \in (0, 0.5)$, the increments of a process tend to have opposite signs and thus are more zigzagging due to the negative covariance; (2) $H \in (0.5, 1)$, the covariance between these two increments is positive and less zigzagging of the process; (3) $H = 0.5$, the covariance between these two increments is zero. This can be restated as following: If the Hurst index is (1) less than 0.5, the process displays "anti-persistence" (i.e., positive excess return is more likely to be reversed and the performance in the next period is likely to be below the average, or in the contrary, negative excess return is more likely to be reversed and the performance in the next period is likely to be above the average); (2) greater than 0.5, the process displays "persistence" (i.e., positive excess return or negative excess return is more likely to be continued and the performance in the next period is likely to be the same as that in the current period); (3) equal to 0.5, the process displays no memory (i.e., the performance in the next period has equal probability to be below and above the performance in the current period).

For fractional stable processes, if the process has the index α ($0 < \alpha < 2$), then when $H = 1/\alpha$ which corresponds to a process with independent increments, we say this process has no memory. When $H > 1/\alpha$, the process displays long-range dependence and when $H < 1/\alpha$, the process displays negative dependence. In addition, long-range dependence is only possible when $\alpha > 1$, since $H \in (0, 1)$ (see Samorodnitsky and Taqqu (1994)).

In order to test for long-range dependence in stock returns, we use the methods introduced in Sections 4.1 and 4.2 to estimate the Hurst index under the Gaussian and stable assumptions. We employed the MLE method explained in Section 4.3 to estimate the stable parameter. The results, reported in Table 1, indicate that the Hurst index does not have an estimated value of 0.5 if fractional Gaussian noise is assumed. This suggests the occurrence of either long memory or short memory under the Gaussian assumption.[6] In Table 1, we can observe both fluctuation and long memory under the non-Gaussian stable assumption.

We use the Ljung-Box-Pierce Q-statistic based on autocorrelation function to test the serial correlation (i.e., the memory effect). The Q-statistic is given as follows:

$$Q :\sim \chi^2_m = N(N+2) \sum_{k=1}^{m} \frac{\rho_k^2}{N-k}, \tag{41}$$

where, N denotes the sample size, m the number of autocorrelation lags included in the statistic, and ρ_k the sample autocorrelation at lag order k which is

$$\rho_k = \frac{\sum_{t=1}^{N-k} y_t y_{t+k}}{\sum_{t=1}^{N} y_t^2}. \tag{42}$$

Ljung and Box (1978) show that the Q-statistic follows the asymptotic chi-square distribution with m degrees of freedom.

Table 3 shows that the null hypothesis that there is no serial correlation can be rejected at different lags. This table shows that the memory effect occurs in each duration series. In order to see when the memory effect vanishes, we compare the Q-statistic with its corresponding critical value. When the quotation of the Q-statistic and the corresponding critical value are less than 1, we cannot reject the null hypothesis that there is no serial correlation. Table 3 also shows such quotations. From this table, all the trade durations exhibit serial correlation. After 500 lags, the memory effect vanishes for 7 stocks and after 1500 lags, the memory effect vanishes for 13 stocks. From the ratios in Table 3, we find that the speed of autocorrelation decay is declining, which confirms the effect of long-range dependence.

6.3 The methodology of finding the best model

In our empirical study, we simulate a series for each distributional assumption with and without subordinating them into the ACD(1,1) structure. Then we compare the goodness of fit of the simulated together with originial trade duration series.

The ACD model can be defined as same as equation (8) and (9):

$$d_i = \psi_i \, u_i, \tag{43}$$

and

$$\psi_t^2 = \kappa + \sum_{i=1}^{p} \gamma_i \, d_{t-i} + \sum_{j=1}^{q} \theta_j \, \psi_{t-j}^2, \tag{44}$$

[6] There are various extensions of the self-similarity property for generalized random processes, see Dobrushin (1979).

u_t can be calculated from d_t/ψ_t. We define

$$\tilde{u_t} = \frac{d_t}{\hat{\psi}_t}, \tag{45}$$

where $\hat{\psi}_t$ is the estimation of ψ_t. In our empirical analysis, an ACD(1,1) model structure is adopted. The objective is to check the statistical characteristics exhibited by trade duration d_t and the error term $\tilde{u_t}$ in ACD(1,1) structure. We simulate d_t and $\tilde{u_t}$ with the ACD(1,1) structure based on the parameters estimated from the empirical series. Then we test the goodness of fit between the empirical series and the simulated series. Six candidate distributional assumptions — lognormal distribution, stable distribution, exponential distribution, Weibull distribution, fractional Gaussian noise, and fractional stable noise — are analyzed for estimation, simulation, and testing.

The Kolmogorov-Simirnov distance (KS) and Anderson-Darling distance (AD) proposed by Rachev and Mittnik (2000) are used as the criterion for the goodness of fit testing. They are defined as following:

$$KS = \sup_{x\in\Re}|F_s(x) - \tilde{F}(x)|, \tag{46}$$

and

$$AD = \sup_{x\in\Re}\frac{|F_s(x) - \tilde{F}(x)|}{\sqrt{\tilde{F}(x)(1-\tilde{F}(x))}}, \tag{47}$$

where $F_s(x)$ denotes the empirical sample distribution and $\tilde{F}(x)$ is the estimated distribution function. The major disadvantage of KS statistic researchers have argued is that it tends to be more sensitive near the center of the distribution than at the tails. But AD statistic can overcome this. The reliability of testing the empirical distribution will be increased with the help of these two statistics, with the KS distance focusing on the deviations around the median of the distribution and the AD distance on the discrepancies in the tails.

6.4 Results

The AD and KS statistics are calculated for the six candidate distributional assumptions. Table 4 reports the descriptive statistics of the computed AD and KS statistics. From Table 4, fractional stable noise and stable distribution exhibit a smaller mean value for the AD and KS statistics in comparison with the other four distributions. Figure 2 shows the boxplot of AD statistics of $\tilde{u_t}$ for the six alternative distributional assumptions investigated. Figure 3 shows the boxplot of AD statistics for d_t. Figure 4 shows the boxplot of KS statistics of $\tilde{u_t}$ for the six alternative distributional assumptions. Figure 5 shows the boxplot of KS statistics of d_t. These figures show that fractional stable noise and stable distribution have a small value of AD and KS statistics, confirming the results reported in Table 4. These results indicate that with or without an ACD(1,1) model structure, the fractional stable noise and stable distribution perform better than the other four tested distributional assumptions based on the criterion for goodness of fit testing.

From Figures 2 to 5, we can see that the fractional stable noise and the stable distribution fit $\tilde{u_t}$ and d_t better than other distributional assumptions. In order to empirically examine our conjecture, we formulate a statistical test procedure. Because we know that smaller AD and KS statistics mean better goodness of fit, in our test we are going to statistically test how significantly "smaller" AD and KS statistics are. The hypothesis test is:

$$H_0 : \mu_{criterion1} - \mu_{criterion2} \geq 0 \qquad (48)$$
$$H_1 : \mu_{criterion1} - \mu_{criterion2} < 0$$

where $\mu_{criterion}$ is the mean value of AD or KS statistics of the candidate distributional assumptions investigated. The distributions of AD and KS values are unknown. All AD or KS values are expressed as i.i.d. random variables $X_1, X_2, ..., X_n$, each with distribution function $F_X(\cdot|\theta)$. A $100(1-\alpha)\%$ upper confidence bound (UCB) is defined as $U(X_1, X_2, ..., X_n)$ for a function of $h(\theta)$ if for every θ,

$$P_\theta \left(h(\theta) \leq U(X_1, X_2, ..., X_n) \right) \geq 1 - \alpha \qquad (49)$$

and $(-\infty, U(X_1, X_2, ..., X_n)]$ is the $100(1-\alpha)\%$ upper confidence interval for $h(\theta)$. Similarly, $L(X_1, X_2, ..., X_n)$ is a $100(1-\alpha)\%$ lower confidence bound (LCB) for the function $h(\theta)$ for every θ

$$P_\theta \left(h(\theta) \geq L(X_1, X_2, ..., X_n) \right) \geq 1 - \alpha \qquad (50)$$

and $[L(X_1, X_2, ..., X_n), +\infty)$ is the $100(1-\alpha)\%$ lower confidence interval for $h(\theta)$.

As hypothesis testing and confidence intervals are dual concepts, the hypothesis testing in (50) is in fact evaluated by following test rules; that is, (1) if UCB is less than zero, H_0 can be rejected, (2) if LCB is greater than zero, H_0 cannot be rejected, and (3) if H_0 is greater than LCB but at the same time less than UCB, there is no statistically significant conclusion. Employing the bootstrap method introduced in DiCiccio and Efron (1996), the 99% bootstrap confidence intervals are reported in Table 5. From this table, at a high confidence level, fractional stable noise and stable distribution are more suitable to modeling trade duration data with or without support of an ACD(1,1) structure.

In comparing the fractional stable noise and stable distribution, it is unclear as to whether the fractional stable noise is better than the stable distribution or vice versa. Table 6 compares the supporting cases for the fractional stable noise and stable distribution. The fractional stable noise has a slightly better support than stable distribution. The stable distribution has a greater number of supporting cases in comparison to the fractional stable noise in modeling duration data with an ACD(1,1) structure.

7. Conclusions

The empirical research with very few stocks have demonstrated that trade duration data exhibit three characteristics: long-range dependence, heavy tailedness, and clustering. In this paper, we investigate the presence of these characteristics using a larger number of stocks and investigate whether for modeling trade duration data: (1) a single stochastic processes capturing long-range dependence and heavy tailedness and; (2) a relatively powerful distributional assumption in a relatively simple functional structure can be used.

To examine these issues, we introduce fractional stable noise and fractional Gaussian noise to capture long-range dependence and heavy tailedness in modeling the trade duration. In our empirical analysis, we investigate six distributional assumptions (fractional stable noise, fractional Gaussian noise, stable distribution, lognormal distribution, exponential distribution, and Weibull distribution) for modeling the trade duration for 18 Dow Jones index component stocks. By using parameters estimated from the empirical series, we simulate a series for each distributional assumption with and without subordinating them into an ACD(1,1) structure. Then we compare the goodness of fit for these generated series to the empirical series by adopting two test criteria for testing heavy-tailed distributions, the Kolmogorov-Simirnov and Anderson-Darling statistics. A test procedure is formulated based on a bootstrap method, and it is used in order to obtain empirical results.

The above test procedure yields empirical evidence which shows that the stable distribution and fractional stable noise are better in modeling trade duration than the exponential distribution, log-normal distribution, Weibull distribution, and fractional Gaussian noise. The results indicate that residuals from the ACD(1,1) model are more likely to be described by a stable distribution and trade durations exhibit the features of fractional stable noise. That is, stable distribution subordinated with an ACD(1,1) structure and fractional stable noise, demonstrate superior performance in the modeling of trade duration.

We argue that it is critical that the findings reported in this paper be taken into account in modeling trade duration. Many studies have found that stable distribution is a better description of financial data because it can capture heavy tailedness and has a close relationship with long-range dependence. As a self-similar process, fractional stable noise can capture almost all reported stylized facts in financial return data, such as heavy tailedness, long memory, non-Gaussian characters, and clustering. Therefore, if fractional stable noise and stable distribution can be properly employed in financial modeling, more accurate prediction might be realized by well-defined functional models.

References

[1] Bardet, J., Lang G., Oppenheim G., Philippe A., Taqqu M.: Generators of long-range dependent processes: a survey, In *Theory and Applications of Long-range Dependence*, P. Doulkhan, G. Oppenheim, and M. Taqqu (eds.), Boston: Birkhäuser 2003

[2] Bauwens, L.: Econometric analysis of intra-daily trading activity on Tokyo stock exchange, IMES Discuussion Paper Series, 2005-E-3 (2005)

[3] Bauwens, L., Veredas D.: The stochastic conditional duration model: a latent variable model for the analysis of financial durations, Journal of Econometrics **119**, 381-412 (2004)

[4] Bauwens, L., Giot P.: The logarithmic ACD model: an application to the bid-ask quote process of three NYSE stocks, Annals d'Economie et de Statistique **60**, 117-149 (2000)

[5] Bauwens, L., Giot P.: Asymmetric ACD models: introducing price information in ACD models, Empirical Economics **28**, 709-731 (2003)

[6] Bauwens, L., Gramming J., Veredas D.: A comparison of financial duration models via density forecasts, International Journal of Forecasting **20**, 589-609 (2004)

[7] Bauwens, L., Hautsch, N.: Modelling financial high frequency data using point processes, in Handbook of Financial Time Series, Springer-Verlag, forthcoming (2006)

[8] Beran, J.: Statistics for Long-Memory Processes. New York: Chapman & Hall 1994

[9] Cohen, S., Samorodnitsky G.: Random rewards, fractional Brownian local times and stable self-similar processes, The Annals of Applied Probability **16**, 1432-1461 (2006)

[10] Diamond, W., Verrecchia R.: Constraints on short-selling and asset price adjustment to private information, Journal of Financial Economics **18**, 277-311 (1987)

[11] DiCiccio, T. J., Efron B.: Bootstrap confidence intervals, Statistical Science **11**, 189-228 (1996)

[12] Dobrushin, R. L.: Gaussian and their subordinated self-similar random generalized fields, The Annals of Probability **7**, 1-28 (1979)

[13] Doukhan, P., Oppenheim G., Taqqu M.: (eds.), Theory and Applications of Long-Range Dependence. Boston: Birkhäuser 2003

[14] Dufour, A., Engle R.: Time and the price impact of a trade, Journal of Finance **55**, 2467-2498 (2000)

[15] Engle, R.: Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, Econometrica **50**, 987-1008 (1982)

[16] Engle, R.: The econometrics of ultra-high frequency data, Econometrica **68**, 1-22 (2000)

[17] Engle, R., Russell J. R.: Autoregressive conditional duration: a new model for irregularly spaced transaction data, Econometrica **66**, 1127-1162 (1998)

[18] Engle, R. Lunde A.: Trades and quotes: a bivariate point process, Journal of Financial Econometrics **1**, 159-188 (2003)

[19] Fama, E.: Mandelbrot and the stable Paretian hypothesis, Journal of Business **36**, 420-429 (1963)

[20] Feng, D., Jiang G. J., Song P.: Stochastic conditional duration models with "leverage effect" for financial transaction data, Journal of Financial Econometrics **2**, 390-421 (2004)

[21] Fernandes, M., Gramming J.: A family of autoregressive conditional duration models, Ensaios Econmicos, No.501 (2003)

[22] Fernandes, M., Gramming J.: Nonparametric specification tests for conditional duration models, Journal of Econometrics **127**, 35-68 (2005)

[23] Ghysels, E., Gouriéroux C., Jasiak J.: Stochastic volatility duration models, Journal of Econometrics **119**, 413-433 (2004)

[24] Hasbrouck, J.: Modeling microstructure time series, in Maddala G., and C. R Rao (eds.) Statistical Methods in Finance (Handbook of Statistics, Volume 14). Amsterdam: North-Holland 1996

[25] Hurst, H.: Long-term storage capacity of reservoirs, Transactions of the American Society of Civil Engineers **116**, 770-808 (1951)

[26] Hurst, H.: Methods of using long-term storage in reservoirs, Proceedings of the Institution of Civil Engineers, Part I, 519-577 (1955)

[27] Jasiak, J.: Persistence in intertrade durations, Finance **19**, 166-195 (1998)

[28] Kolmogoroff, A. N.: Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum, Comptes Rendus (Doklady) Acad. Sci. URSS (N.S.) **26**, 115-118 (1940)

[29] Lamperti, J.: Semi-stable stochastic Processes, Transactions of the American Mathematical Society **104**, 62-78 (1962)

[30] Mandelbrot, B. B.: New methods in statistical economics, Journal of Political Economy **71**, 421-440 (1963)

[31] Mandelbrot, B. B.: The Fractal Geometry of Nature. San Francisco: Freeman 1983

[32] Mandelbrot, B. B., Van Ness J.: Fractional brownian motions, fractional noises and applications, SIAM Review **10**, 422-437 (1968)

[33] Mandelbrot, B. B., Wallis J.: Noah, Joseph and operational hydrology, Water Resources Research **4**, 909-918 (1968)

[34] Maejima, M., Rachev S.: An ideal metric and the rate of convergence to a self-similar process, Annals of Probability **15**, 702-727 (1987)

[35] Manganelli, S.: Duration, volume and volatility impact of trades, Journal of Financial Markets **8**, 377-399 (2005)

[36] Mansfield, P., Rachev S., Samorodnitsky G.: Long strange segments of a stochastic process and long-range dependence, The Annals of Applied Probability **11**, 878-921 (2001)

[37] Mittnik, S., Rachev S.: Modeling asset returns with alternative stable models, Econometric Reviews **12**, 261-330 (1993a)

[38] Mittnik, S., Rachev S.: Reply to comments on "Modeling asset returns with alternative stable models" and some extensions, Econometric Reviews **12**, 347-389 (1993b)

[39] O'Hara, M.: Market Microstructure Theory. Cambridge: Blackwell 1995

[40] Paxson, V.: Fast, approximate synthesis of fractional Gaussian noise for generating self-similar network traffic, Computer Communication Review **27**, 5-18 (1997)

[41] Rachev, S.: (edt.) Handbook of Heavy Tailed Distributions in Finance. Amsterdam: Elsevier 2003

[42] Rachev, S., Mittnik S.: Stable Paretian Models in Finance. New York: Wiley 2000

[43] Rachev, S., Samorodnitsky G.: Long strange segments in a long range dependent moving average, Stochastic Processes and their Applications **93**, 119-148 (2001)

[44] Rachev, S., Menn C., Fabozzi F.: Fat-Tailed and Skewed Asset Return Distributions. Hoboken: Wiley 2005

[45] Samorodnitsky, G.: Possible sample paths of self-similar alpha-stable processes, Statistics and Probability Letters **19**, 233-237 (1994)

[46] Samorodnitsky, G.: Lower tails of self-similar stable processes, Bernoulli **4**, 127-142 (1998)

[47] Samorodnitsky, G., Taqqu M.: Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance. Boca Raton: Chapman & Hall/CRC 1994

[48] Stoev, S., Pipiras V., Taqqu M.: Estimation of the self-similarity parameter in liner fractional stable motion, Signal Processing **82**, 1873-1901 (2002)

[49] Stoev, S., and M. Taqqu, Simulation methods for linear fractional stable motion and FARIMA using the Fast Fourier Transform, Fractals **12**, 95-121 (2004)

[50] Stoyanov, S., Racheva-Iotova B.: Univariate stable laws in the field of finance—approximation of density and distribution functions, Journal of Concrete and Applicable Mathematics **2**, 38-57 (2004a)

[51] Stoyanov, S., Racheva-Iotova B.: Univariate stable laws in the field of finance—parameter estimation, Journal of Concrete and Applicable Mathematics **2**, 24-49 (2004b)

[52] Sun, W., Rachev, S., Fabozzi, F.: Long-range dependence, fractal processes, and intra-daily data, in Handbook of IT and Finance, Springer, forthcoming (2006)

[53] Zhang, M., Russell J., Tsay R.: A nonlinear autoregressive conditional duration model with applications to financial transaction data, Journal of Econometrics **104**, 179-207 (2001)

Table 1: Statistical characteristics of trade duration in 2003 for 18 stocks

Stock	$size$	$mean$	$variance$	$skewness$	$kurtosis$	$Hurst_{\tilde{u}_t}$	$Hurst_{d_t}$	$\alpha_{\tilde{u}_t}$	α_{d_t}
Alcoa Inc.	511817	10.8858	2214.9814	69.0315	5388.2732	0.6139	0.7202	1.3291	1.1022
American Express	819489	6.9582	2030.7830	77.7079	6318.8553	0.6379	0.5774	1.5673	1.3316
Caterpillar	571251	10.1333	3363.9955	58.7267	3679.0844	0.6932	0.6289	1.3475	1.0286
E.I.DuPont de Nemours	715515	7.8020	1629.7963	84.7517	7767.4205	0.6107	0.6379	1.3832	1.1930
Walt Disney	813090	6.8123	1354.3835	94.4269	9558.9300	0.6936	0.6943	1.3827	1.2792
Eastman Kodak Co.	475512	11.9756	3591.6955	57.6869	3600.8356	0.6445	0.7108	1.4295	1.1715
General Electric	1188851	4.8295	1470.4087	92.0469	8745.4964	0.4814	0.4725	1.4833	1.4413
General Motors	684082	8.4509	2821.2248	65.0970	4452.0523	0.4647	0.5679	1.2921	1.1606
IBM	986153	5.8425	1921.3140	81.4365	6837.2217	0.5186	0.4952	1.6492	1.2792
Int.Paper Company	606742	9.1403	1786.1335	79.9445	7072.3144	0.6195	0.7110	1.3677	1.0742
Coca-Cola Co.	695948	8.1239	2059.3277	75.4303	6113.4419	0.5676	0.5763	1.3470	1.1606
McDonalds	615302	9.0091	1801.3532	80.1190	7001.8869	0.5548	0.6555	1.3144	1.2437
3M Co.	739258	7.7422	2299.2715	71.6024	5407.2143	0.5819	0.5948	1.3705	1.2349
Altria Group	846140	6.7920	2078.7019	77.6406	6338.5890	0.5396	0.5842	1.3825	1.1282
Merck & Co.	875457	6.6009	2158.2266	76.6211	6145.4793	0.5861	0.5252	1.4886	1.1715
Procter & Gamble	780633	7.2657	1875.1364	80.2896	6796.9486	0.6853	0.5704	1.4437	1.2792
AT&T Inc.	553896	10.1818	2504.3505	67.7552	4979.9100	0.6280	0.6894	1.4133	1.1542
United Technologies	657286	8.5779	2240.6857	72.5846	5659.7341	0.6051	0.6544	1.3176	1.1282

Table 2: Studies of Trade Duration and Stocks Included in Sample

Study	Exchange	Stocks	Model distribution(s)
Bauwens (2005)	Tokyo Stock Exchange	Nippon Steel, Sony, Tokyo Electric, Toyota	Generalized gamma
Bauwens and Veredas (2004)	New York Stock Exchange	Boeing, Coca Cola, Walt Disney, Exxon	Lognormal, Weibull, Gamma
Bauwens and Giot (2000)	New York Stock Exchange	Boeing, Walt Disney, IBM	Exponential, Weibull
Bauwens and Giot (2003)	New York Stock Exchange	Walt Disney, IBM	Weibull
Bauwens *et al.* (2004)	New York Stock Exchange	Boeing, Coca Cola, Walt Disney, Exxon	Burr, Exponential, Generalized gamma, Weibull
Engle and Russell (1998)	New York Stock Exchange	IBM	Exponential, Weibull
Engle and Lunde (2003)	New York Stock Exchange	Bank-American, Walt Disney, General Motors	Exponential
		Federal National Mortagage, McDonald's,	
		Monsanto, Procter & Gamble, Schlumberger	
Feng *et al.* (2004)	New York Stock Exchange	Boeing, Coca Cola, IBM	Exponential, log-Weibull, log-Gamma
Fernandes and Gramming (2003)	New York Stock Exchange	Boeing, Coca Cola, Walt Disney, Exxon, IBM	Burr
Fernandes and Gramming (2005)	New York Stock Exchange	Exxon	Exponential, Weibull, Burr, Generalized gamma
Ghysels *et al.* (2004)	Paris Stock Exchange	Alcatel	Exponential, Gamma
Jasiak (1998)	Paris Stock Exchange	Alcatel	Weibull
	New York Stock Exchange	IBM	
Zhang *et al.* (2001)	New York Stock Exchange	IBM	Generalized gamma

Table 3: Ljung-Box-Pierce Q-test statistic for different lags at α=0.05. The italic numbers show the ratios of Q-test statistics compared with corresponding critical values.

stock	lag10	lag20	lag50	lag100	lag200	lag500	lag1000	lag1500
Alcoa Inc.	760.3	931.7	1379.5	2041.0	3224.5	5912.9	9897.8	13327.0
	41.5	*29.7*	*20.4*	*16.4*	*13.8*	*10.7*	*9.2*	*8.4*
American Express	330.5	336.2	344.5	352.8	372.8	405.6	596.7	691.3
	18.1	*10.7*	*5.1*	*2.8*	*1.6*	*0.7*	*0.6*	*0.4*
Caterpillar	413.9	446.1	516.6	623.5	814.5	1329.9	2093.9	2899.9
	22.6	*14.2*	*7.6*	*5.0*	*3.4*	*2.4*	*1.9*	*1.8*
E.I.DuPont de Nemours	371.6	401.4	455.6	527.1	644.4	840.9	1184.5	1376.5
	20.3	*12.7*	*6.7*	*4.2*	*2.7*	*1.5*	*1.1*	*0.8*
Walt Disney	294.5	319.2	374.8	452.1	606.5	936.2	1475.7	1620.3
	16.1	*10.1*	*5.5*	*3.6*	*2.5*	*1.6*	*1.3*	*1.0*
Eastman Kodak Co.	157.9	177.6	204.4	257.0	351.2	654.6	1078.8	1530.8
	8.6	*5.6*	*3.0*	*2.0*	*1.5*	*1.1*	*1.1*	*0.9*
General Electric	386.8	391.1	393.1	394.8	397.4	406.6	601.1	627.5
	21.1	*12.4*	*5.8*	*3.1*	*1.6*	*0.7*	*0.5*	*0.3*
General Motors	350.8	356.2	365.7	373.3	396.8	513.5	584.7	832.1
	19.1	*11.3*	*5.4*	*3.0*	*1.6*	*0.9*	*0.5*	*0.5*
IBM	159.9	165.7	169.7	174.0	181.2	195.7	357.1	418.0
	8.7	*5.2*	*2.5*	*1.3*	*0.7*	*0.3*	*0.3*	*0.2*
Int. Paper Company	344.3	400.2	519.7	698.0	986.6	1424.1	1846.4	2299.0
	18.8	*12.7*	*7.6*	*5.6*	*4.2*	*2.5*	*1.7*	*1.4*
Coca-Cola Co.	244.2	253.4	266.0	293.4	318.8	366.1	620.3	758.9
	13.3	*8.0*	*3.9*	*2.3*	*1.3*	*0.6*	*0.5*	*0.4*
McDonalds	233.2	256.1	292.2	346.9	442.6	899.3	1016.4	1343.8
	12.7	*8.1*	*4.3*	*2.7*	*1.8*	*1.6*	*0.9*	*0.8*
3M Co.	515.3	528.2	551.2	579.0	631.5	750.7	1044.4	1340.2
	28.1	*16.8*	*8.1*	*4.6*	*2.6*	*1.3*	*0.9*	*0.8*
Altria Group	279.1	284.2	292.6	299.0	319.1	357.7	556.0	617.1
	15.2	*9.0*	*4.3*	*2.4*	*1.3*	*0.6*	*0.5*	*0.3*
Merck & Co.	262.9	268.0	276.1	280.0	285.9	303.0	461.6	535.6
	14.3	*8.5*	*4.0*	*2.2*	*1.2*	*0.5*	*0.4*	*0.3*
Procter & Gamble	517.6	522.9	533.5	541.1	553.0	580.6	774.7	824.3
	28.2	*16.6*	*7.9*	*4.3*	*2.3*	*1.1*	*0.7*	*0.5*
AT&T Inc.	280.2	319.1	398.1	517.9	691.1	990.2	1370.2	1709.4
	15.3	*10.1*	*5.8*	*4.1*	*2.9*	*1.7*	*1.2*	*1.1*
United Technologies	327.1	356.2	417.2	491.2	621.1	834.9	1130.4	1389.9
	17.8	*11.3*	*6.1*	*3.9*	*2.6*	*1.5*	*1.1*	*0.8*
Critical Value	18.3	31.4	67.5	124.3	233.9	553.1	1074.7	1591.2

Table 4: Summary of KS statistics for alternative distributional assumptions, " * " indicates the test for d_t, otherwise for $\tilde{u}_t$. Mean, median, standard deviation ("std"), maximum value ("max"), minimum value ("min") and range of AD and KS statistics are presented in this table.

	AD_{mean}	AD_{median}	AD_{std}	AD_{max}	AD_{min}	AD_{range}	AD^*_{mean}	AD^*_{median}	AD^*_{std}	AD^*_{max}	AD^*_{min}	AD^*_{range}
FGN	2.7056	1.7220	2.7171	15.2900	0.1732	15.1170	4.4772	3.9688	3.0290	19.3230	0.2423	19.0810
fsn	0.4468	0.4207	0.1809	1.1384	0.1066	1.0318	0.4574	0.4307	0.1836	1.2071	0.1014	1.1057
lognormal	2.6100	1.6285	2.7004	15.4960	0.1652	15.3310	4.3202	3.8918	3.0646	19.6470	0.2407	19.4070
stable	0.4460	0.4233	0.1787	1.1384	0.1122	1.0262	0.4583	0.4372	0.1835	1.0795	0.1060	0.9736
exponential	1.0553	1.0516	0.0198	1.1521	0.9719	0.1802	1.0572	1.0526	0.0174	1.1246	1.0254	0.0993
weibull	1.0487	1.0472	0.0220	1.1103	0.9471	0.1632	1.0572	1.0526	0.0174	1.1246	1.0254	0.0993

c

	KS_{mean}	KS_{median}	KS_{std}	KS_{max}	KS_{min}	KS_{range}	KS^*_{mean}	KS^*_{median}	KS^*_{std}	KS^*_{max}	KS^*_{min}	KS^*_{range}
FGN	0.2615	0.2901	0.0866	0.3955	0.0644	0.3311	0.3157	0.3359	0.0715	0.4054	0.0902	0.3152
fsn	0.0381	0.0361	0.0101	0.0887	0.0197	0.0690	0.0493	0.0484	0.0097	0.0920	0.0288	0.0632
lognormal	0.2590	0.2816	0.0858	0.3984	0.0613	0.3371	0.3129	0.3334	0.0717	0.4030	0.0891	0.3139
stable	0.0378	0.0366	0.0099	0.0861	0.0195	0.0666	0.0495	0.0484	0.0101	0.0923	0.0281	0.0641
exponential	0.5265	0.5249	0.0091	0.5579	0.4856	0.0723	0.5278	0.5257	0.0081	0.5586	0.5126	0.0459
weibull	0.5236	0.5230	0.0105	0.5522	0.4729	0.0793	0.5278	0.5257	0.0081	0.5586	0.5126	0.0459

Table 5: Bootstrap 99% confidence intervals for mean of differences in AD and KS statistics, " * " indicates statistics for d_t, otherwise for $\tilde{u_t}$. "FGN" stands for fractional Gaussian noise, "fsn" stands for fractional stable noise, "exp" stands for exponential distribution, "wbl" stands for Weibull distribution.

T:	AD	AD^*	KS	KS^*
$E(T_{stable} - T_{FGN})$	(-2.5223, -1.9851)	(-4.3235 , -3.7124)	(-0.2318 , -0.2157)	(-0.2734 , -0.2594)
$E(T_{stable} - T_{lognormal})$	(-2.4281, -1.8893)	(-4.1642 , -3.5481)	(-0.2292 , -0.2132)	(-0.2706 , -0.2565)
$E(T_{stable} - T_{exp})$	(-0.6278 , -0.5913)	(-0.6175 , -0.5803)	(-0.4895 , -0.4878)	(-0.4793 , -0.4774)
$E(T_{stable} - T_{wbl})$	(-0.6211 , -0.5850)	(-0.6175 , -0.5804)	(-0.4868 , -0.4848)	(-0.4793 , -0.4773)
$E(T_{fsn} - T_{FGN})$	(-2.5236 , -1.9774)	(-4.3204 , -3.7077)	(-0.2315 , -0.2154)	(-0.2736 , -0.2597)
$E(T_{fsn} - T_{lognormal})$	(-2.4265 , -1.8888)	(-4.1638 , -3.5485)	(-0.2290 , -0.2130)	(-0.2708 , -0.2568)
$E(T_{fsn} - T_{exp})$	(-0.6274 , -0.5904)	(-0.6182 , -0.5815)	(-0.4893 , -0.4875)	(-0.4795 , -0.4776)
$E(T_{fsn} - T_{wbl})$	(-0.6203 , -0.5838)	(-0.6185 , -0.5815)	(-0.4866 , -0.4845)	(-0.4795 , -0.4776)
$E(T_{fsn} - T_{stable})$	(-0.0057 , 0.0074)	(-0.0077 , 0.0059)	(-0.0002 , 0.0007)	(-0.0007 , 0.0002)

Table 6: Supporting cases comparison of goodness of fit for fractional stable noise and stable distribution based on AD and KS statistics. Symbol " * " indicates the test for d_t, otherwise the test is for $\tilde{u}_t$. Symbol " $\succ$" means being preferred and "$\sim$" means indifference. Numbers shows the supporting cases to the statement in the first column and the number in parentheses give the proportion of supporting cases in the whole sample.

	AD	AD^*	KS	KS^*
$fsn \succ stable$	327	345	327	362
	(47.81%)	(50.44%)	(47.81%)	(52.93%)
$stable \succ fsn$	344	328	351	318
	(50.29%)	(47.95%)	(51.32 %)	(46.49%)
$fsn \sim stable$	13	11	6	4
	(1.90%)	(1.61 %)	(0.87%)	(0.58%)

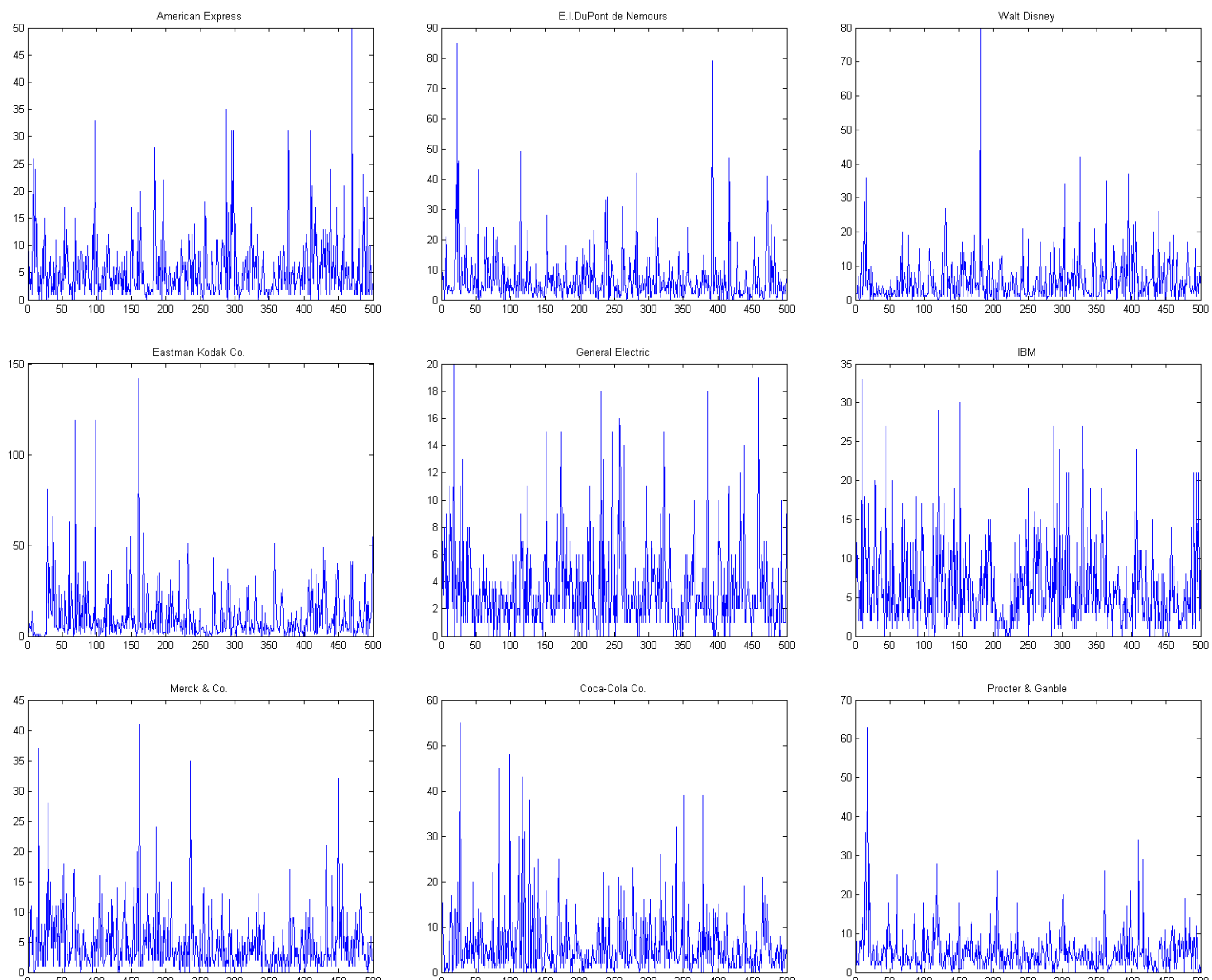

Figure 1: Plot of trade duration for several stocks.

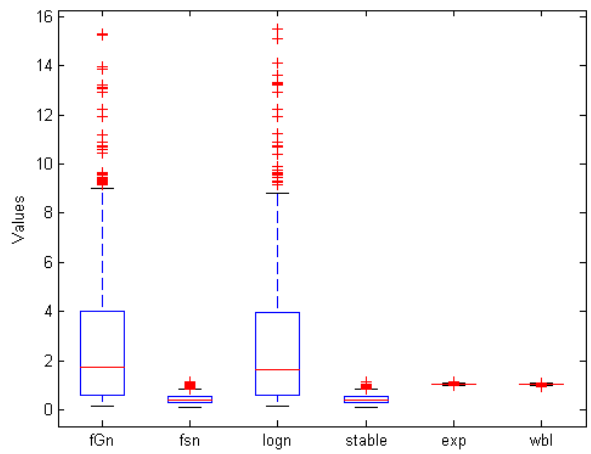

Figure 2: Boxplot of AD statistics for $\tilde{u}_t$ in alternative distributional assumptions.

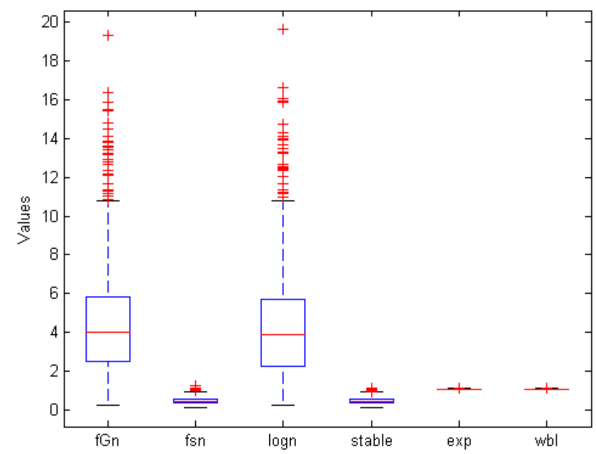

Figure 3: Boxplot of AD* statistics for d_t in alternative distributional assumptions.

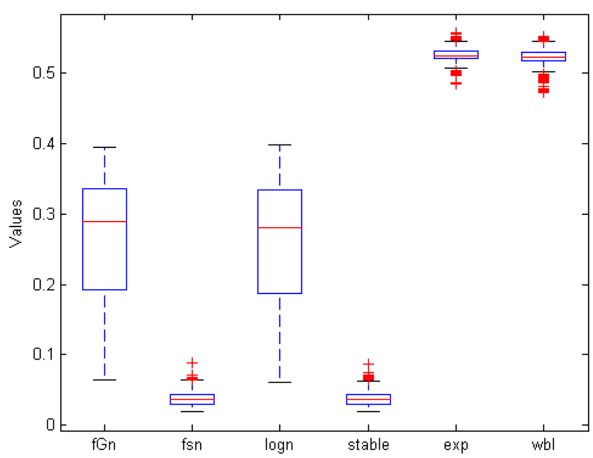

Figure 4: Boxplot of KS statistics for $\tilde{u}_t$ in alternative distributional assumptions.

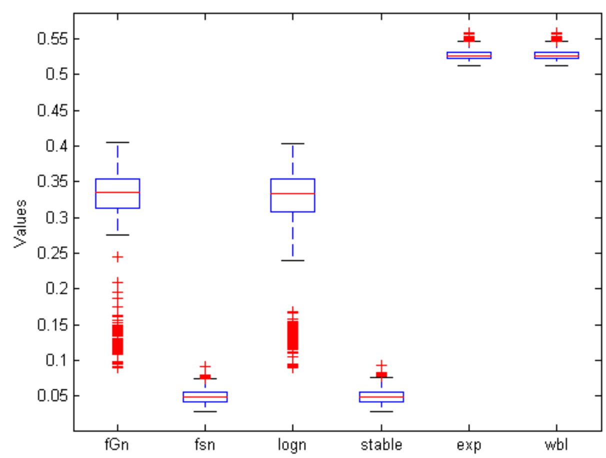

Figure 5: Boxplot of KS* statistics for d_t in alternative distributional assumptions.