Quantitative Methods in High-Frequency Financial Econometrics:Modeling Univariate and Multivariate Time Series

Wei Sun

Institute of Statistics and Mathematical Economics,

University of Karlsruhe, Germany

wei.sun@statistik.uni-karlsruhe.de

http://www.statistik.uni-karlsruhe.de/361.php

Introduction

1. Motivation
2. Fractal Processes
   - Fractional Gaussian noise
   - Fractional Stable noise
3. Unconditional Copulas and Tail Dependence
   - Gaussian copula
   - Student's t copula
   - Skewed Student's t copula
   - Tail dependence
4. Empirical Framework
   - Models for single stock returns
   - Models for single trade durations
   - Models for multivariate returns with symmetric correlation
   - Models for multivariate returns with asymmetric correlation
5. Conclusion

Stylized Facts of High-Frequency Stock Market Data

- Random durations (Dacorogna et al. (2001))
- Distributional properties
  - ▷ Fatter tails in the unconditional return distributions. (Bollerslev et al. (1992), Marinelli et al. (2000))
  - ▷ Stock returns are not independently and identically distributed. (Sun et al. (2007a))
- Autocorrelation (Bollwerslev et al. (2000), Wood et al. (1985))
- Seasonality (Gourieroux and Jasiak (2001))
- Clustering
  - ▷ Volatility clustering. (Engle (2000))
  - ▷ Trade duration clustering (Sun et al. (2006b))
- Long-range dependence. (Sun et al. (2007a))

Motivation

Modeling Irregularity and Roughness of Price Movement

- Capturing the stylized facts observed in high-frequency data
- Establishing a model for the study of price dynamics
- Simulating price movement based on the established model
- Testing the goodness of fit for the established model

Modeling Dependence Structure

- Dependence of price movement of a single asset
- Dependence of price movement between several assets

Why Fractal Processes?

- "The reasons are that the main feature of price records is roughness and that the proper language of the theory of roughness in nature and culture is fractal geometry" (Mandelbrot (2005)).
  - ▷ Mandelbrot (1982): The fractal geometry of nature. Freeman, New York.
  - ▷ Mandelbrot (1997): Fractals and scaling in finance. Springer, New York.
  - ▷ Mandelbrot (2002): Gaussian self-affinity and fractals. Springer, New York.
- Custom has made the increments' ratio be viewed as "normal" and thought the highly anomalous ratio has the limit $\alpha = 1/2$.
- Being the same at all instants in all financial data is a very important property. It has simplicity. But it also has a big flaw – a limit equal to $1/2$ is not available as parameter to be fitted to the data.
- The fractal processes allow $\alpha \neq 1/2$.
- A key feature of fractal processes is that it measures roughness by α and the value and/or the distribution of α is directly observable.
- The speed of volatility variation can be considered by the fractal processes.

What are Fractal Processes?

- Fractal processes (self-similar processes) are invariant in distribution with respect to changes of time and space scale. The scaling coefficient or self-similarity index is a non-negative number denoted by H, the Hurst parameter.
- Lamperti (1962) first introduced semi-stable processes (which we nowadays call self-similar processes).
- If $\{X(t+h) - X(h), t \in T\} \overset{d}{=} \{X(t) - X(0), t \in T\}$ for all $h \in T$, the real-valued process $\{X(t), t \in T\}$ has stationary increments. Samorodnisky and Taqqu (1994) provide a succinct expression of self-similarity: $\{X(at), t \in T\} \overset{d}{=} \{a^H X(t), t \in T\}$. The process $\{X(t), t \in T\}$ is called H-sssi if it is self-similar with index H and has stationary increments.
- In our study, two fractal processes are employed:
  - ▷ fractional Gaussian noise
  - ▷ fractional stable noise

Fractional Gaussian noise

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

$$B_H(t) := \frac{1}{\Gamma(H+\frac{1}{2})}\left(\int_{-\infty}^{0}\left[(t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}}\right]dB(s) + \int_{0}^{t}(t-s)^{H-\frac{1}{2}}dB(s)\right)$$

  where $\Gamma(\cdot)$ represents the Gamma function and $0 < H < 1$ is the Hurst parameter. The integrator B is the ordinary Brownian motion.
- As to the fractional Brownian motion, Samorodnitsky and Taqqu (1994) define its increments $\{Y_j, j \in Z\}$ as fractional Gaussian noise (fGn), which is, for $j = 0, \pm 1, \pm 2, \ldots$, $Y_j = B_H(j-1) - B_H(j)$.
- The main 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.

Fractional stable noise

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

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

where

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

and where 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 $\alpha=1$.

- Samorodinitsky and Taqqu (1994) define linear fractional stable noises expressed by $Y(t)$, and $Y(t) = X_t - X_{t-1}$, $Y(t) = L_{\alpha,H}(a,b;t) - L_{\alpha,H}(a,b;t-1)$, where $L_{\alpha,H}(a,b;t)$ is a linear fractional stable motion defined above, and M is a stable random measure with Lebesgue control measure given $0<\alpha<2$.

Stable Distribution

- 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 $\gamma > 0$,
  - 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 of 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 real ζ such that its characteristic function has the following form:

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

and,

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

Tail Dependence

- In financial data, we can observe that extreme events happen simultaneously for different assets. In a time interval, several assets might exhibit extreme values. Tail dependence reflects the dependence structure between extreme events. It turns out that tail dependence is a copula property.
- Letting $(Y_1, Y_2)^T$ be a vector of continuous random variables with marginal distribution functions F_1, F_2, then the coefficient of the upper tail dependence of $(Y_1, Y_2)^T$ is

$$\lambda_U = \lim_{u \to 1} P\Big(Y_2 > F_2^{-1}(u) | Y_1 > F_1^{-1}(u)\Big)$$

and the coefficient of the lower tail dependence of $(Y_1, Y_2)^T$ is

$$\lambda_L = \lim_{u \to 0} P\Big(Y_2 < F_2^{-1}(u) | Y_1 < F_1^{-1}(u)\Big)$$

If $\lambda_U > 0$, there exists upper tail dependence and the positive extreme values can be observed simultaneously. If $\lambda_L > 0$, there exists lower tail dependence and the negative extreme values can be observed simultaneously (Embrechts et al. (2003)).

Unconditional Copulas

- Sklar (1959) has shown:

$$\begin{aligned} F_Y(y_1, ..., y_n) &= P(Y_1 \le y_1, \ldots, Y_n \le y_n) \\ &= C(P(Y_1 \le y_1), \ldots, P(Y_n \le y_n)) \\ &= C(F_{Y_1}(y_1), \ldots, F_{Y_n}(y_n)) \end{aligned}$$

where F_{Y_i}, $i = 1, \ldots, n$ denote the marginal distribution functions of the random variables, Y_i, $i = 1, \ldots, n$.

- When the variables are continuous, the density c associated with the copula is given by:

$$c(F_{Y_1}(y_1), \ldots, F_{Y_n}(y_n)) = \frac{\partial^n C(F_{Y_1}(y_1), \ldots, F_{Y_n}(y_n))}{\partial F_{Y_1}(y_1), \ldots, \partial F_{Y_n}(y_n)}.$$

- The density function f_Y corresponding to the n-variate distribution function F_Y is

$$f_Y(y_1, ..., y_n) = c(F_{Y_1}(y_1), \ldots, F_{Y_n}(y_n)) \prod_{i=n}^{n} f_{Y_i}(y_i),$$

where f_{Y_i}, $i = 1, \ldots, n$ is the density function of F_{Y_i}, $i = 1, \ldots, n$ (see, Joe (1997), Cherubini et al. (2004), and Nelsen (2006)).

Gaussian copula

- Let ρ be the correlation matrix which is a symmetric, positive definite matrix with unit diagonal, and Φ_ρ the standardized multivariate normal distribution with correlation matrix ρ. The unconditional multivariate Gaussian copula is then

$$C(u_1, \ldots, u_n; \rho) = \Phi_\rho\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n)\right),$$

and the corresponding density is

$$c(u_1, \ldots, u_n; \rho) = \frac{1}{|\rho|^{1/2}} \exp\left(-\frac{1}{2}\lambda^T(\rho^{-1} - I)\lambda\right),$$

where $\lambda = (\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n))^T$ and u_n is the margins.

Student's t copula

- The unconditional (standardized) multivariate Student's copula $T_{\rho,\nu}$ can be expressed as

$$T_{\rho,\nu}(u_1,\ldots,u_n;\rho) = t_{\rho,\nu}\Big(t_\nu^{-1}(u_1),\ldots,t_\nu^{-1}(u_n)\Big),$$

where $t_{\rho,\nu}$ is the standardized multivariate Student's t distribution with correlation matrix ρ and ν degrees of freedom and t_ν^{-1} is the inverse of the univariate cumulative density function (c.d.f) of the Student's t with ν degrees of freedom. The density of the unconditional multivariate Student's t copula is

$$c_{\rho,\nu}(u_1,\ldots,u_n;\rho) = \frac{\Gamma(\frac{\nu+n}{2})}{\Gamma(\frac{\nu}{2})|\rho|^{1/2}}\left(\frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu+1}{2})}\right)^n \left(\frac{\Big(1+\frac{1}{\nu}\lambda^T\rho^{-1}\lambda\Big)^{-\frac{\nu+n}{2}}}{\prod_{j=1}^n\Big(1+\frac{\lambda_j^2}{\nu}\Big)^{-\frac{\nu+1}{2}}}\right),$$

where $\lambda_j = t_\nu^{-1}(u_j)$ and and u_n is the margins.

Skewed Student's t copula

- The skewed Student's t copula is defined as the copula of the multivariate distribution of X. Therefore, the copula function is

$$C(u_1, \ldots, u_n) = F_X(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n))$$

where F_X is the multivariate distribution function of X and $F_k^{-1}(u_k)$, $k = 1, n$ is the inverse c.d.f of the k-th marginal of X. That is, $F_X(x)$ has the density $f_X(x)$ defined above and the density function $f_k(x)$ of each marginal is

$$f_k(x) = \frac{aK_{(\nu+1)/2}\left(\sqrt{\left(\nu + \frac{(x-\mu_k)^2}{\sigma_{kk}}\right)\frac{\gamma_k^2}{\sigma_{kk}}}\right)\exp\left((x-\mu_k)\frac{\gamma_k}{\sigma_{kk}}\right)}{\left(\left(\nu + \frac{(x-\mu_k)^2}{\sigma_{kk}}\right)\frac{\gamma_k^2}{\sigma_{kk}}\right)^{-\frac{\nu+1}{4}}\left(1 + \frac{(x-\mu_k)^2}{\nu\sigma_{kk}}\right)^{\nu+1}}, \quad x \in R$$

where σ_{kk} is the k-th diagonal element in the matrix Σ.

Empirical Framework

Models for single stock returns

- Investigate the return distribution of German DAX stocks using intra-daily data under two separate assumptions regarding the return generation process (1) it does not follow a Gaussian distribution and (2) it does not follow a random walk.
- The high-frequency data at 1-minute frequency for 27 German DAX component stocks from January 7, 2002 to December 19, 2003 are investigated.
- The ARMA-GARCH Model is employed.

Empirical Framework I

The ARMA-GARCH Model

- ARMA model

$$y_t = \alpha_0 + \sum_{i=1}^{r} \alpha_i \, y_{t-i} + \varepsilon_t + \sum_{j=1}^{m} \beta_j \varepsilon_{t-j}.$$

- GARCH model

$$\sigma_t^2 = \kappa + \sum_{i=1}^{\mathrm{p}} \gamma_i \, \sigma_{t-i}^2 + \sum_{j=1}^{\mathrm{q}} \theta_j \, \varepsilon_{t-j}^2.$$

Since $\varepsilon_t = \sigma_t \, u_t$, u_t could be calculated from ε_t / σ_t. Defining

$$\tilde{u}_t = \frac{\varepsilon_t^s}{\hat{\sigma}_t},$$

where ε_t^s is estimated from the sample and $\hat{\sigma}_t$ is the estimation of σ_t. In our study, ARMA(1,1)-GARCH(1,1) are parameterized as marginal distributions with different kinds of u_t (i.e., normal distribution, fractional Gaussian noise, fractional stable noise, stable distribution, generalized Pareto distribution, and generalized extreme value distribution).

Empirical Framework I

The Goodness of Fit Tests

- Kolmogorov-Simirnov distance (KS)

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

- Anderson-Darling distance (AD)

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

- Cramer Von Mises distance (CVM)

$$CVM = \int_{-\infty}^{\infty} \left(F_s(x) - \tilde{F}(x) \right)^2 d\tilde{F}(x),$$

- Kuiper distance (K)

$$K = \sup_{x \in \Re} \left(F_s(x) - \tilde{F}(x) \right) + \sup_{x \in \Re} \left(\tilde{F}(x) - F_s(x) \right).$$

Empirical Framework I

Empirical Results

Statistical characteristics of stocks in the study.

	mean	std	skewness	kurtosis	max	min	$\tilde{\alpha}$	$\tilde{H}_{FGN}$	$\tilde{H}_{fsn}$
Adidas	1.83E-07	0.0012	0.4041	94.9171	0.0670	-0.0327	0.7903	0.4791	1.3168
Allianz	-3.01E-06	0.0015	-0.5477	212.8600	0.0949	-0.0905	1.4222	0.5203	0.5616
BASF	2.65E-07	0.0012	1.2349	121.8300	0.0863	-0.0341	1.2247	0.4548	0.4589
BAYER	-1.29E-06	0.0015	0.8794	85.5980	0.0877	-0.0521	1.3998	0.5137	0.4947
BMW	-1.27E-07	0.0012	0.1213	58.6840	0.0459	-0.0463	1.2010	0.4946	0.4873
Commerzbank	-4.40E-07	0.0018	0.2352	48.3320	0.0539	-0.0577	1.1694	0.5337	0.5186
Daimler	-7.83E-07	0.0013	-0.2877	56.4390	0.0469	-0.0597	1.4160	0.4628	0.4621
Dt.Bank	-5.72E-07	0.0012	-0.6959	68.6110	0.0339	-0.0571	1.3938	0.5332	0.5082
Dt.Post	2.16E-07	0.0015	-0.3476	41.8600	0.0393	-0.0512	0.8724	0.4258	1.0786
Dt. Telecom	-8.79E-07	0.0015	-0.8822	100.8200	0.0557	-0.0913	1.8337	0.5293	0.5190
Eon	-4.01E-07	0.0011	-0.2422	26.6700	0.0234	-0.0334	1.2434	0.4214	0.4381
Fressenius	-7.16E-07	0.0017	5.0585	646.8212	0.1982	-0.0447	0.8294	0.5411	1.2408
Henkel	3.64E-08	0.0011	1.9018	261.9911	0.0977	-0.0389	0.7562	0.4115	1.2339
Hypovereinsbank	-1.92E-06	0.0019	-2.3888	236.3100	0.0694	-0.1475	1.2688	0.5469	0.5304
Infineon	-2.36E-06	0.0020	-0.7092	145.4146	0.0707	-0.1016	1.6018	0.5351	0.5451
Linde	-1.01E-07	0.0013	-0.2952	44.1141	0.0340	-0.0392	0.7968	0.4716	1.2696
Lufthansa	-4.10E-07	0.0016	0.5289	79.2296	0.0821	-0.0541	1.2694	0.4647	0.4647
Man	3.80E-08	0.0016	0.3009	65.3610	0.0754	-0.0461	0.7645	0.5372	1.3134
Metro	2.66E-06	0.0015	0.1496	63.0310	0.0477	-0.0455	1.1293	0.4806	0.5084
Muechenerrueck	-3.52E-06	0.0015	-0.4968	105.6100	0.07250	-0.0748	1.3203	0.5842	0.5559
RWE	-9.71E-07	0.0013	-0.1861	27.6080	0.0353	-0.0355	1.2270	0.3960	0.4400
SAP	3.54E-06	0.0011	-0.1718	128.7802	0.0577	-0.0427	1.3693	0.5481	0.5485
Schering	-2.45E-07	0.0012	-0.3756	64.1430	0.0299	-0.0553	1.1411	0.4299	0.4271
Siemens	2.74E-06	0.0011	0.1990	49.7927	0.0379	-0.0350	1.4426	0.5251	0.5275
ThyssenKrupp	2.18E-06	0.0015	-0.1627	34.7571	0.0373	-0.0419	1.1064	0.5019	0.5126
Tui	-1.59E-06	0.0019	0.0812	145.4800	0.1253	-0.1032	0.8074	0.4965	1.1906
Volkswagen	-5.18E-07	0.0013	-0.5010	60.3170	0.0502	-0.0584	1.1807	0.4908	0.5042

Empirical Framework I

Empirical Results

ARCH-test for different lags at $\alpha = 0.05$.

$q =$	1min	2min	5min	10min	15min	20min	25min	30min	60min
Adidas	4403.1	6357.5	6530.4	6772.9	6821.2	6836.3	6894.1	6901.5	6986.5
Allianz	24.9	35.8	62.6	111.6	120.2	124.1	126.7	129.1	136.3
BASF	52.6	58.6	70.3	83.0	88.1	92.6	95.0	97.2	110.3
BAYER	2590.8	2595.1	2739.4	4511.9	4837.8	4840.4	4853.3	4881.6	5068.2
BMW	1211.3	1618.1	1780.2	1941.7	2030.9	2096.5	2129.6	2154.1	2262.1
Commerzbank	2595.7	2831.3	3693.9	4046.8	4109.8	4136.0	4181.2	4196.0	4264.1
Daimler	307.0	374.5	402.4	475.8	504.5	522.2	537.0	552.8	578.2
Dt.Bank	1172.6	1304.8	1504.4	1786.6	1911.8	1992.4	2073.5	2130.8	2325.2
Dt.Post	1552.0	3241.9	3390.5	3542.5	3601.2	3631.2	3661.7	3678.2	3702.1
Dt. Telecom	319.7	364.3	418.0	488.8	537.9	569.6	597.1	625.8	693.7
Eon	5050.6	5534.0	5896.9	6418.7	6668.7	6796.2	6940.1	7028	7251.6
Fressenius	9701.1	11503.0	12842.1	12981.0	13044.0	13083.0	13111.0	13124.0	13211.0
Henkel	131.2	139.6	148.1	185.4	194.2	196.0	197.4	198.9	211.5
Hypovereinsbank	75.9	90.2	100.9	121.6	127.9	154.2	154.6	163.1	179.7
Infineon	2477.1	2995.4	3056.6	3087.6	3101	3104.4	3113.1	3117.4	3128.6
Linde	3526.6	5312.5	5564.4	5797.4	5896.7	6048.9	6102.1	6146.3	6192.0
Lufthansa	1311.3	1605.0	3466.5	3519.5	3590.8	3626.7	3643.2	3649.3	3688.6
Man	3302.3	3890.1	4360.1	4530.6	4564.7	4572.9	4590.3	4600.9	4628.0
Metro	3654.2	4173.9	5454.1	5510.7	5551.1	5582.3	5597.6	5615.3	5652.3
Muechenerrueck	361.4	493.5	620.2	652.4	674.5	688.4	697.2	708.9	793.5
RWE	1747.6	2354.2	3049.4	3653.3	3875.2	3983.9	4042.6	4098.7	4284.3
SAP	351.0	377.9	410.2	448.5	466.4	483.8	496.7	502.7	515.4
Schering	2059.0	2335.1	2616.9	2763.7	2828.9	2865.3	2910.5	2937	2956.6
Siemens	732.8	896.3	1045.9	1178.5	1339.0	1398.2	1475.8	1513.6	1614.2
ThyssenKrupp	2237.2	3127.1	3399.1	3566.4	3671.6	3794	3824.2	3852.9	3892.2
Tui	154.4	251.8	299.5	331.4	346.1	358.9	366.9	374.6	395.7
Volkswagen	2682.0	3367.0	4020.8	4806.4	4906.4	4980.2	5020.3	5036.7	5114.0
Critical value	3.8415	5.9915	11.0705	18.307	24.996	31.4104	37.6525	43.773	67.5048

Empirical Framework I

Empirical Results

Ljung-Box-Pierce Q-test statistic for different lags at $\alpha = 0.05$.

$k =$	$10min$	$30min$	$1hr$	$2hr$	$4hr$	$1day$	$1wk$	$2wk$	$1mon$
Adidas	2016.7	2083.8	2131.5	2194.1	2321.3	2579.8	4910.9	7783.8	15872.0
Allianz	309.4	371.7	421.8	514.8	694.8	987.5	3923.6	7490.2	18300.0
BASF	624.7	665.5	708.2	800.2	961.6	1319.7	4014.8	7174.5	16150.0
BAYER	572.4	650.2	745.0	865.1	1136.3	1506.8	4453.2	7758.6	17859.0
BMW	653.2	698.7	741.5	849.2	983.3	1279.4	3855.4	6854.6	15903.0
Commerzbank	1990.2	2013.7	2067.1	2166.1	2289.2	2613.1	5127.9	8440.4	17903.0
Daimler	1324.3	1376.4	1428.9	1479.6	1618.1	1895.0	4461.3	7311.3	15993.0
Dt.Bank	710.6	760.1	830.1	933.4	1080.0	1453.6	3993.9	7076.4	16047.0
Dt.Post	2030.7	2096.6	2132.0	2187.1	2310.3	2579.8	5042.4	7870.4	16261.0
Dt. Telecom	2031.5	2084.9	2170.6	2289.2	2436.1	2794.6	5652.3	9147.6	19491.0
Eon	822.6	866.2	924.4	1020.4	1147.9	1515.8	4351.1	7554.7	17160.0
Fressenius	2565.9	2640.8	2680.1	2765.3	2885.1	3154.0	5705.7	8722.2	17796.0
Henkel	1742.1	1823.5	1878.5	1959.9	2130.6	2477.5	5168.2	8111.4	16844.0
Hypovereinsbank	910.2	959.8	1046.2	1138.5	1277.7	1586.7	3931.8	6924.0	16185.0
Infineon	1454.0	1518.1	1555.4	1619.4	1789.7	2071.6	4489.3	7482.8	16435.0
Linde	1048.7	1097.6	1136.9	1203.0	1371.1	1662.1	3986.9	6690.9	14994.0
Lufthansa	2491.6	2544.4	2590.2	2659.1	2823.4	3087.4	5568.6	8543.5	17758.0
Man	1817.7	1866.1	1913.5	1991.6	2134.2	2419.6	4778.6	7874.1	16652
Metro	1339.1	1400.2	1426.3	1506.9	1692.5	1954.3	4319.2	7230.2	16405.5
Muechenerrueck	220.2	280.9	332.8	409.6	555.2	880.1	3708.7	7020.0	16338.1
RWE	942.2	989.4	1051.8	1146.8	1303.7	1626.3	4683.1	8025.7	17937.3
SAP	484.2	511.3	554.4	629.7	748.9	1060.7	3254.8	6141.3	14647.2
Schering	1421.1	1451.2	1486.2	1579.2	1746.6	2006.5	4599.5	7511.9	15945.0
Siemens	776.5	850.9	929.6	1003.3	1158.3	1481.5	4221.6	7596.7	16957.0
ThyssenKrupp	2623.9	2693.6	2733.4	2813.3	2959.5	3206.6	5407.0	8185.1	16618.0
Tui	1561.9	1599.7	1633.4	1734.4	1873.2	2166.8	4732.4	7751.7	17175.0
Volkswagen	369.7	394.97	459.9	526.6	674.9	987.9	3598.3	6671.3	15722.0
Critical value	18.3	43.8	79.1	146.6	277.1	532.1	2515.1	4962.3	12256.0

Empirical Results

Summary of in-sample goodness of fit statistics for different models.

a. AD-statistic	AD_{mean}	AD_{std}	AD_{median}	AD_{max}	AD_{min}	AD_{range}
ARMA-GARCH-fGn	46.6768	54.3660	13.7335	55.8541	21.6282	34.2259
ARMA-GARCH-fsn	44.1625	53.2522	15.2382	64.6001	1.3917	63.2084
ARMA-GARCH-nor	46.7177	54.3751	13.7480	58.5747	21.0690	37.5057
ARMA-GARCH-sta	45.4108	53.5900	14.3638	95.2204	2.9886	92.2318
ARMA-GARCH-gev	46.6401	54.2656	13.7441	60.2271	21.0914	39.1357
ARMA-GARCH-gpd	51.2203	54.4755	20.2070	109.5363	3.5018	106.0244
b. KS-statistic	KS_{mean}	KS_{std}	KS_{median}	KS_{max}	KS_{min}	KS_{range}
ARMA-GARCH-fGn	0.4998	0.4992	0.0034	0.5285	0.4887	0.0398
ARMA-GARCH-fsn	0.4938	0.4965	0.0261	0.9725	0.2745	0.6980
ARMA-GARCH-nor	0.5003	0.4994	0.0043	0.5455	0.4893	0.0562
ARMA-GARCH-sta	0.5089	0.4974	0.0622	0.9725	0.3910	0.5815
ARMA-GARCH-gev	0.5000	0.4989	0.0057	0.5775	0.4825	0.0950
ARMA-GARCH-gpd	0.5698	0.5198	0.1335	1.0000	0.4165	0.5835
c. CVM-statistic	CVM_{mean}	CVM_{std}	CVM_{median}	CVM_{max}	CVM_{min}	CVM_{range}
ARMA-GARCH-fGn	449.5326	517.0503	203.7237	896.6917	82.9310	813.7607
ARMA-GARCH-fsn	445.2936	515.2956	202.2463	1473.6038	34.4169	1439.1868
ARMA-GARCH-nor	449.6701	517.3712	203.7739	889.0445	82.9532	806.0913
ARMA-GARCH-sta	454.7954	516.7259	230.4694	2985.6218	57.5066	2928.1152
ARMA-GARCH-gev	449.2438	517.0805	203.8510	886.1477	83.0288	803.1189
ARMA-GARCH-gpd	524.3399	521.2781	370.3320	1978.9581	52.5637	1926.3945
d. Kuiper-statistic	$Kuiper_{mean}$	$Kuiper_{std}$	$Kuiper_{median}$	$Kuiper_{max}$	$Kuiper_{min}$	$Kuiper_{range}$
ARMA-GARCH-fGn	0.9931	0.9937	0.0029	0.9985	0.9757	0.0227
ARMA-GARCH-fsn	0.9693	0.9862	0.0473	0.9985	0.5125	0.4860
ARMA-GARCH-nor	0.9931	0.9938	0.0029	0.9990	0.9750	0.0240
ARMA-GARCH-sta	0.9796	0.9877	0.0224	0.9990	0.6550	0.3440
ARMA-GARCH-gev	0.9913	0.9925	0.0048	0.9990	0.9570	0.0420
ARMA-GARCH-gpd	0.9696	0.9773	0.0287	1.0000	0.6505	0.3495

Empirical Results

Goodness of fit statistics for out-of-sample one week forecasting of different models.

a. AD-statistic	AD_{mean}	AD_{std}	AD_{median}	AD_{max}	AD_{min}	AD_{range}
ARMA-GARCH-fGn	30.1821	22.5228	13.6283	55.2241	21.5834	33.6407
ARMA-GARCH-fsn	27.6038	22.4110	12.2880	68.1149	1.2129	66.9021
ARMA-GARCH-nor	30.1927	22.5228	13.6421	59.2046	21.1361	38.0684
ARMA-GARCH-sta	28.8034	22.3886	13.0021	101.7023	2.6941	99.0082
ARMA-GARCH-gev	30.1205	22.5005	13.5541	59.8893	20.7111	39.1782
ARMA-GARCH-gpd	32.3273	23.9319	15.3931	108.9876	4.1084	104.8792
b. KS-statistic	KS_{mean}	KS_{std}	KS_{median}	KS_{max}	KS_{min}	KS_{range}
ARMA-GARCH-fGn	0.5018	0.5006	0.0049	0.5375	0.4905	0.0470
ARMA-GARCH-fsn	0.4965	0.4985	0.0278	0.9555	0.2820	0.6734
ARMA-GARCH-nor	0.5020	0.5010	0.0055	0.5615	0.4880	0.0735
ARMA-GARCH-sta	0.5105	0.4990	0.0617	0.9653	0.4049	0.5603
ARMA-GARCH-gev	0.5018	0.5005	0.0064	0.5705	0.4846	0.0858
ARMA-GARCH-gpd	0.5700	0.5210	0.1333	1.0000	0.4049	0.5951
c. CVM-statistic	CVM_{mean}	CVM_{std}	CVM_{median}	CVM_{max}	CVM_{min}	CVM_{range}
ARMA-GARCH-fGn	226.5630	92.4072	245.5856	950.0136	82.5743	867.4392
ARMA-GARCH-fsn	223.6907	91.6806	244.8920	1873.2304	34.0265	1839.2039
ARMA-GARCH-nor	226.6043	92.4969	245.5955	948.5973	82.7246	865.8726
ARMA-GARCH-sta	229.7204	92.0018	264.0086	2852.7912	77.7136	2775.0774
ARMA-GARCH-gev	225.5252	92.2819	243.0575	933.6036	82.4613	851.1423
ARMA-GARCH-gpd	248.8990	94.0770	282.0017	1929.6210	55.7156	1873.9054
d. Kuiper-statistic	$Kuiper_{mean}$	$Kuiper_{std}$	$Kuiper_{median}$	$Kuiper_{max}$	$Kuiper_{min}$	$Kuiper_{range}$
ARMA-GARCH-fGn	0.9935	0.9940	0.0032	1.0000	0.9715	0.0285
ARMA-GARCH-fsn	0.9698	0.9870	0.0481	0.9990	0.5362	0.4627
ARMA-GARCH-nor	0.9935	0.9940	0.0032	1.0000	0.9725	0.0275
ARMA-GARCH-sta	0.9801	0.9885	0.0231	0.9995	0.6876	0.3118
ARMA-GARCH-gev	0.9918	0.9930	0.0052	0.9995	0.9592	0.0402
ARMA-GARCH-gpd	0.9703	0.9780	0.0299	1.0000	0.6425	0.3575

Empirical Framework II

Models for single trade durations

- Ultra-high frequency data of 18 Dow Jones index component stocks based on NYSE trading for year 2003 are examined. 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. In the dataset, I observe many consecutive zero durations, which implies the existence of multiple transactions within a second. I aggregate these intra-second transactions within a second (see Engle and Russell (1998)).

- In the empirical analysis, an ACD(1,1) model structure is adopted. The objective is to check the statistical characteristics exhibited by trade duration $\boldsymbol{d_i}$ and the error term $\tilde{\boldsymbol{u}}_i$ in ACD(1,1) structure. I simulate $\boldsymbol{d_i}$ and $\tilde{\boldsymbol{u}}_i$ with the ACD(1,1) structure based on the parameters estimated from the empirical series. Then I 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. Trade durations are positive numbers, therefore the stable distribution, fractional Gaussian noise, and fractional stable noise are defined on positive supports correspondingly.

Empirical Framework II

The ACD model

-

$$d_i \;=\; \psi_i\, u_i,$$

-

$$\psi_i^2 \quad = \quad \kappa + \sum_{t=1}^{\mathrm{p}} \gamma_i\, d_{i-t} \;+\; \sum_{j=1}^{\mathrm{q}} \theta_j\, \psi_{i-j}^2,$$

- u_i can be calculated from d_i/ψ_i.

$$\tilde{u}_i \;=\; \frac{d_i}{\hat{\psi}_i},$$

where $\hat{\psi}_i$ is the estimation of ψ_i.

Empirical Framework II

Empirical Results

Statistical characteristics of trade duration in 2003 for 18 stocks

Stock	*size*	*mean*	*Min*	*Max*	$Hurst_{\tilde{u}_t}$	$Hurst_{d_t}$	$\alpha_{\tilde{u}_t}$	α_{d_t}
Alcoa Inc.	511817	10.8858	0	6194	0.6139	0.7202	1.3291	1.1022
American Express	819489	6.9582	0	6148	0.6379	0.5774	1.5673	1.3316
Caterpillar	571251	10.1333	0	6139	0.6932	0.6289	1.3475	1.0286
E.I.DuPont de Nemours	715515	7.8020	0	6114	0.6107	0.6379	1.3832	1.1930
Walt Disney	813090	6.8123	0	6134	0.6936	0.6943	1.3827	1.2792
Eastman Kodak Co.	475512	11.9756	0	6145	0.6445	0.7108	1.4295	1.1715
General Electric	1188851	4.8295	0	6196	0.4814	0.4725	1.4833	1.4413
General Motors	684082	8.4509	0	6147	0.4647	0.5679	1.2921	1.1606
IBM	986153	5.8425	0	6143	0.5186	0.4952	1.6492	1.2792
Int.Paper Company	606742	9.1403	0	6174	0.6195	0.7110	1.3677	1.0742
Coca-Cola Co.	695948	8.1239	0	6145	0.5676	0.5763	1.3470	1.1606
McDonalds	615302	9.0091	0	6165	0.5548	0.6555	1.3144	1.2437
3M Co.	739258	7.7422	0	6142	0.5819	0.5948	1.3705	1.2349
Altria Group	846140	6.7920	0	6149	0.5396	0.5842	1.3825	1.1282
Merck & Co.	875457	6.6009	0	6135	0.5861	0.5252	1.4886	1.1715
Procter & Gamble	780633	7.2657	0	6166	0.6853	0.5704	1.4437	1.2792
AT&T Inc.	553896	10.1818	0	6161	0.6280	0.6894	1.4133	1.1542
United Technologies	657286	8.5779	0	6161	0.6051	0.6544	1.3176	1.1282

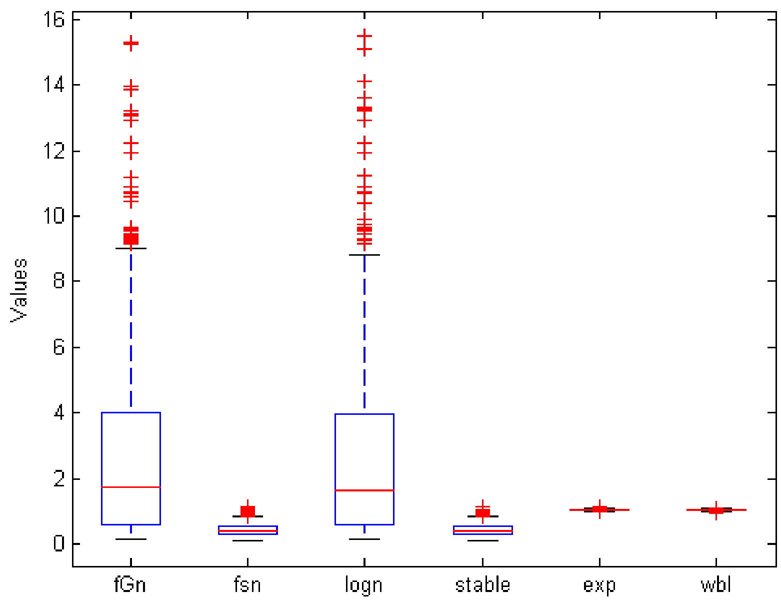
Empirical Results

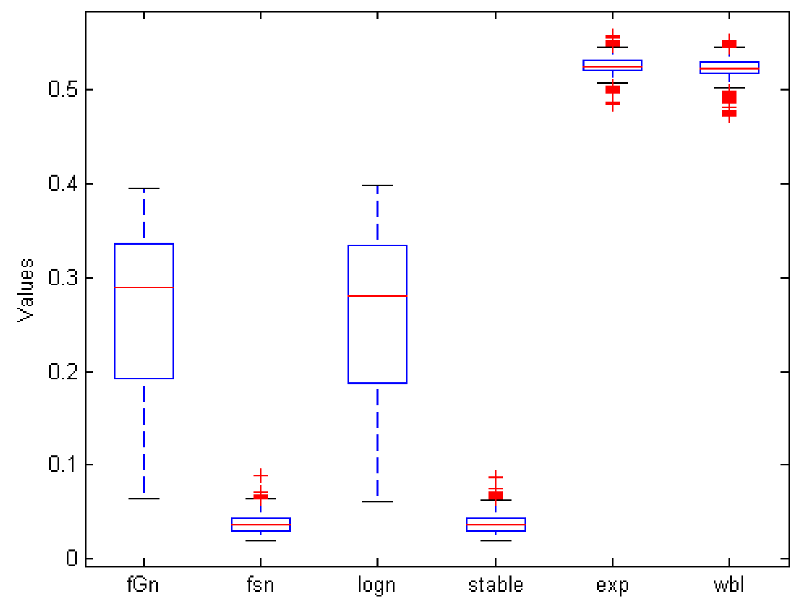
- Supporting cases comparison of goodness of fit for fractional stable noise and stable distribution based on AD and KS statistics. Symbol " * " indicates the test for d_t, otherwise the test is for $\tilde{u}_t$. Symbol " $\succ$" means being preferred and "$\sim$" means indifference. Numbers shows 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%)

Empirical Results

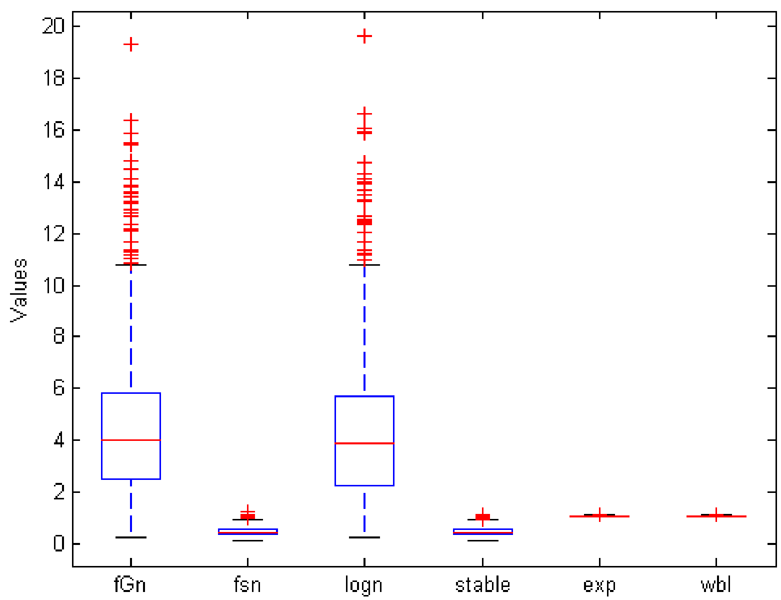
- Boxplot of AD and KSstatistics for $\tilde{u}_t$ in alternative distributional assumptions.

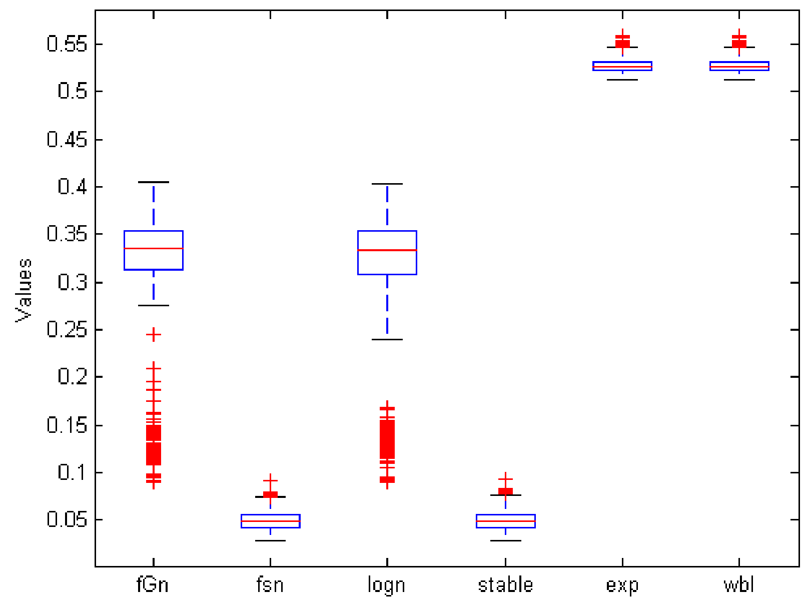




Empirical Results

- Boxplot of AD and KS statistics for d_t in alternative distributional assumptions.





Empirical Framework III

Models for multivariate returns with symmetric correlation

- The high-frequency data of the nine international stock indexes (i.e., AORD, DAX, FCHI, FTSE, HSI, KS200, N225, SPX, and STOXX) from January 8, 2002 to December 31, 2003 were aggregated to the 1-minute frequency level.
- The ARMA-GARCH Model as the Marginal Distribution.
- The Gaussian and Student's t copula for correlation.

Empirical Results

- J_ρ statistic of testing exceedence correlation at quantile=0.8. p values of rejecting the null hypothesis of symmetric correlation are reported in parentheses.

	DAX	FCHI	FTSE	HSI	KS200	N225	SPX	STOXX
AORD	1.0058	1.2757	2.3229	0.0894	0.7097	0.0610	0.5043	0.4164
	(0.3158)	(0.2586)	(0.1274)	(0.7649)	(0.3995)	(0.8048)	(0.4775)	(0.5187)
DAX		0.001	2.9011	2.161	0.7393	0.8153	0.0376	3.9866
		(0.9743)	(0.0885)	(0.1415)	(0.3898)	(0.3665)	(0.8461)	(0.0458)
FCHI			0.343	4.5819	3.1019	1.6173	0.1942	0.9737
			(0.5580)	(0.0323)	(0.0781)	(0.2034)	(0.6594)	(0.3237)
FTSE				0.1115	2.2976	0.2142	0.0644	0.1000
				(0.7384)	(0.1295)	(0.6434)	(0.7995)	(0.7517)
HSI					0.6317	0.6457	0.495	0.1771
					(0.4267)	(0.4216)	(0.4816)	(0.6738)
KS200						0.5464	0.0819	0.3178
						(0.4597)	(0.7746)	(0.5728)
N225							8.13E-05	0.242
							(0.9928)	(0.6226)
SPX								0.1221
								(0.7266)

Empirical Framework III

Empirical Results

- J_ρ statistic of testing exceedence correlation at quantile=0.95. p values of rejecting the null hypothesis of symmetric correlation are reported in parentheses.

	DAX	FCHI	FTSE	HSI	KS200	N225	SPX	STOXX
AORD	2.5350	1.2649	0.4065	0.9127	1.4148	0.3300	0.0256	0.7522
	(0.1113)	(0.2607)	(0.5237)	(0.3393)	(0.2342)	(0.5656)	(0.8726)	(0.3857)
DAX		0.5586	0.3197	0.3560	0.7390	0.0086	0.1135	1.5363
		(0.4548)	(0.5717)	(0.5506)	(0.3899)	(0.9258)	(0.7360)	(0.2151)
FCHI			1.8125	0.7246	0.0236	1.3454	0.6177	3.3126
			(0.1782)	(0.3946)	(0.8778)	(0.2460)	(0.4318)	(0.0680)
FTSE				0.9908	1.2734	1.1343	0.919	0.7235
				(0.3195)	(0.2591)	(0.2868)	(0.3377)	(0.3949)
HSI					1.2354	2.2382	0.4741	0.3233
					(0.2663)	(0.1346)	(0.4910)	(0.5695)
KS200						3.4007	1.6541	0.0108
						(0.0651)	(0.1983)	(0.9170)
N225							0.5622	0.0927
							(0.4533)	(0.7606)
SPX								0.5598
								(0.4543)

Empirical Framework III

Empirical Results

- Summary of the AD, KS and CVM statistics for alternative models for joint distribution. Mean, median, standard deviation ("std"), maximum value ("max"), minimum value ("min") and range of the AD, KS and CVM statistics are presented in this table.

	AD_{mean}	AD_{median}	AD_{std}	AD_{max}	AD_{min}	AD_{range}
Gaussian copula	0.9241	0.9374	0.0338	0.9718	0.8370	0.1348
Student's t copula	0.9237	0.9362	0.0340	0.9716	0.8382	0.1334
	KS_{mean}	KS_{median}	KS_{std}	KS_{max}	KS_{min}	KS_{range}
Gaussian copula	48.4519	55.5841	16.3230	67.9456	9.6306	58.3150
Student's t copula	48.4470	55.5060	16.3190	67.9740	9.9158	58.0580
	CVM_{mean}	CVM_{median}	CVM_{std}	CVM_{max}	CVM_{min}	CVM_{range}
Gaussian copula	785.6190	798.7101	24.5134	817.5083	729.5811	87.9272
Student's t copula	785.2964	798.1155	24.6323	817.9235	728.6673	89.2562

Empirical Framework IV

Models for multivariate returns with asymmetric correlation

- In this study, six indexes in German equity market (i.e., DAX, HDAX, MDAX, Midcaps, SDAX, and TecDAX) are considered.
- The high-frequency data of the six indexes in German equity market listed above from January 2 to September 30, 2006 were aggregated to the 1-minute frequency level.
- Employing high-frequency data has several advantages compared to the low-frequency data. First, with a very large amount of observations, high-frequency data offers a higher level of statistical significance. Second, high-frequency data are gathered at a low level of aggregation, thereby capturing the heterogeneity of players in financial markets. Third, using high-frequency data in analyzing the co-movement in an equity market can consider both the microstructure effects and macroeconomic factors.
- The ARMA-GARCH Model as the Marginal Distribution.
- The Skewed Student's t copula for correlation.

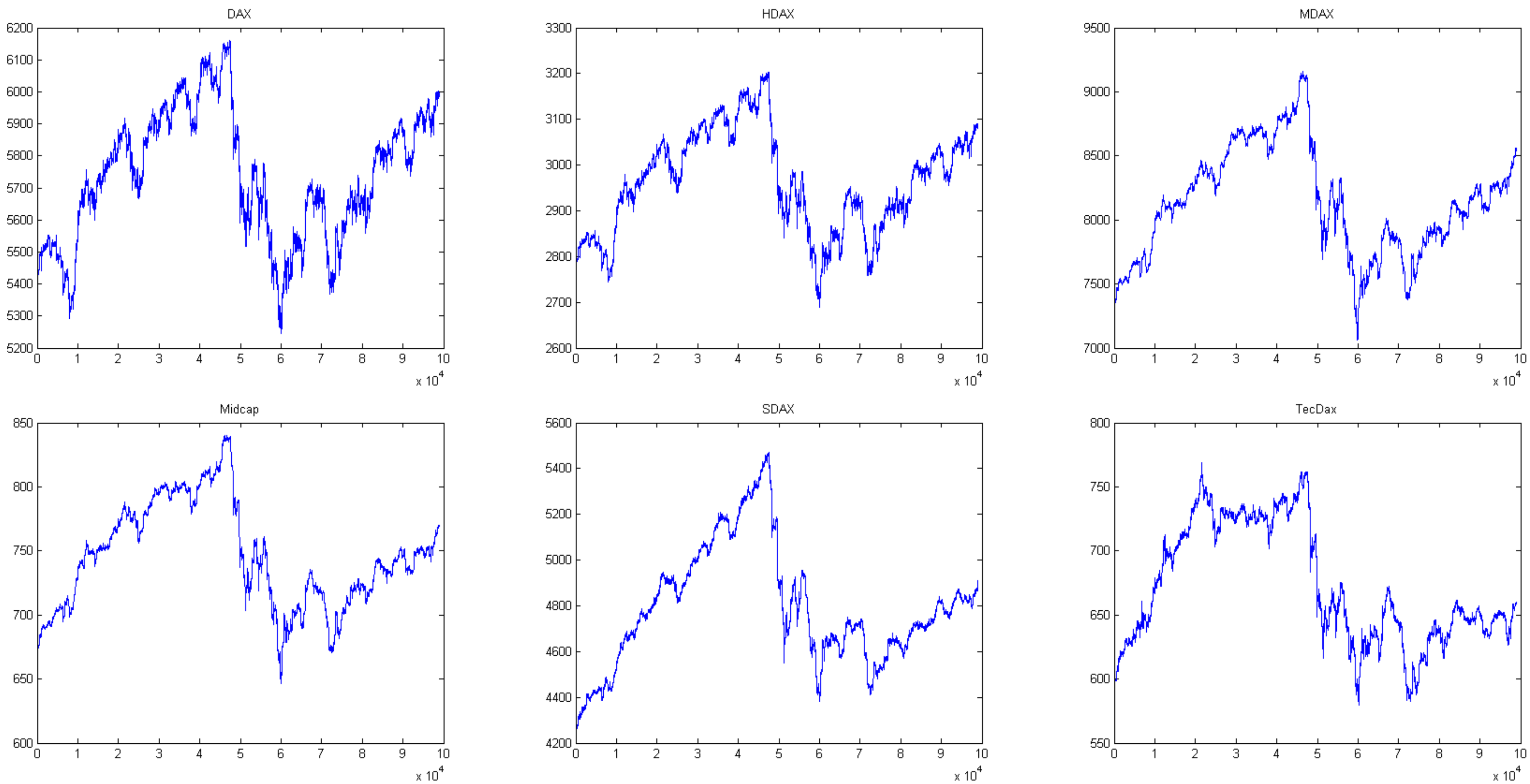
Empirical Framework

The Data

- DAX is a market index for Blue Chip stocks consisting of the 30 major German companies trading on the Frankfurt Stock Exchange.
- The HDAX includes the shares of all 110 companies from the DAX, MDAX, and TecDAX selection indices. Compared to the DAX, the HDAX represents a broader index covering all sectors and the shares of the largest companies listed in Prime Standard.
- The MDAX includes the 50 companies from classic sectors that rank immediately below the companies included in the DAX index. The company size is based on terms of order book volume and market capitalization.
- The SDAX is the small caps index for 50 smaller companies in Germany, which in terms of order book volume and market capitalization rank directly below the MDAX shares.
- Midcaps is the short name of Midcap Market Index which consisting 80 German companies trading on the Frankfurt Stock Exchange.
- The TecDAX stock index tracks the performance of 30 largest German companies from the technology sector.

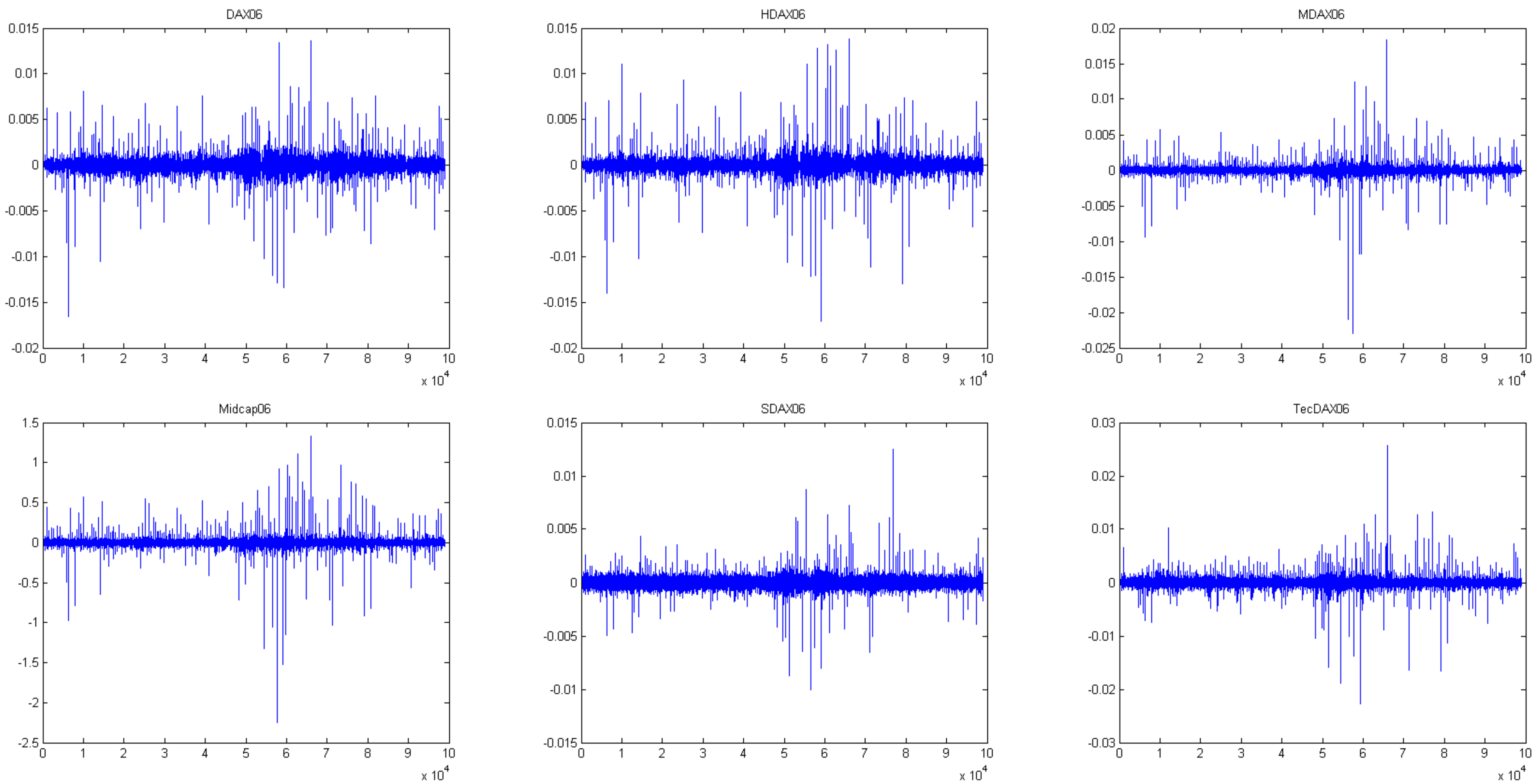
Empirical Results

- Plot of index dynamics.



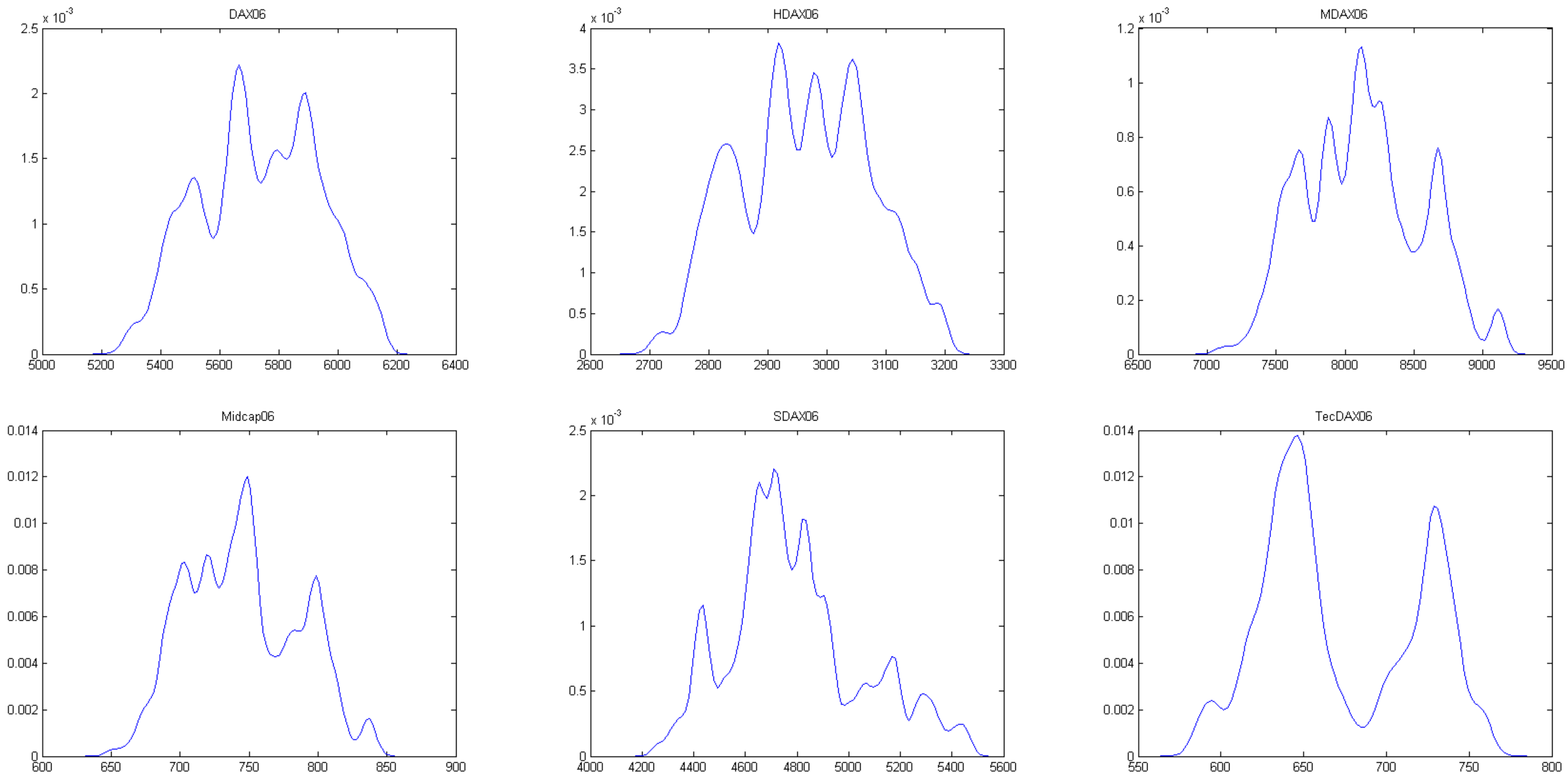
Empirical Results

- Plot of index return.



Empirical Results

- Estimated density using a kernel smoothing method.



Empirical Results

Mean of the in-sample and out-of-sample goodness of fit statistic for alternative models with joint distribution (Gaussian copula).

	In-sample						Out-of-sample					
	FGN	FSN	Normal	Stable	GEV	GPD	FGN	FSN	Normal	Stable	GEV	GPD
KS	0.5060	0.5048	0.5061	0.5120	0.5086	0.8831	0.5074	0.5073	0.5084	0.5139	0.5091	0.8831
AD	32.9751	32.8757	32.9751	33.3603	33.1234	57.4691	24.0592	24.0571	24.1089	24.3672	24.1423	41.8651
CVM	177.3689	177.3087	177.3719	177.8170	177.4900	561.5001	93.9545	93.9241	93.9594	94.2369	93.9836	297.2944
K	0.9982	0.9977	0.9983	0.9977	0.9979	0.9979	0.9983	0.9979	0.9986	0.9979	0.9981	0.9981

Mean of the in-sample and out-of-sample goodness of fit statistic for alternative models with joint distribution (t-copula).

	In-sample						Out-of-sample					
	FGN	FSN	Normal	Stable	GEV	GPD	FGN	FSN	Normal	Stable	GEV	GPD
KS	0.5049	0.5053	0.5062	0.5140	0.5085	0.8834	0.5080	0.5069	0.5075	0.5163	0.5108	0.8829
AD	32.9006	32.9230	32.9776	33.4938	33.1240	57.4796	24.0882	24.0333	24.0658	24.4880	24.2235	41.8567
CVM	177.3417	177.3131	177.3719	177.9565	177.5208	561.7679	93.9563	93.9064	93.9477	94.3381	94.0597	297.2955
K	0.9983	0.9976	0.9983	0.9975	0.9979	0.9981	0.9984	0.9977	0.9985	0.9978	0.9980	0.9982

Mean of the in-sample and out-of-sample goodness of fit statistic for alternative models with joint distribution (Skewed t-copula).

	In-sample						Out-of-sample					
	FGN	FSN	Normal	Stable	GEV	GPD	FGN	FSN	Normal	Stable	GEV	GPD
KS	0.5013	0.4993	0.5017	0.5075	0.5046	0.8570	0.5038	0.5018	0.5054	0.5088	0.5068	0.8576
AD	32.1926	31.7879	32.2002	32.3578	32.1734	54.0610	23.6406	23.4495	23.6880	23.8110	23.7328	40.0679
CVM	176.3376	175.6668	176.3228	176.2896	176.2375	537.6912	93.3896	93.0380	93.4694	93.4583	93.4071	284.7948
K	0.9902	0.9862	0.9901	0.9861	0.9883	0.9893	0.9909	0.9873	0.9911	0.9875	0.9896	0.9901

Empirical Framework IV

Empirical Results

- Summary statistics by groups of each creteria with respect to different models.

In-sample	FGN	FSN	Normal	Stable	GEV	GPD
Gaussian copula	52.9620	52.9217	52.9628	53.1717	53.0300	155.2125
Student t copula	52.9364	52.9347	52.9635	53.2405	53.0378	155.2822
Skewed t copula	52.5054	52.2350	52.5037	52.5352	52.4759	148.3996
Out-of-sample	**FGN**	**FSN**	**Normal**	**Stable**	**GEV**	**GPD**
Gaussian copula	29.8779	29.8716	29.8938	30.0290	29.9083	85.2602
Student t copula	29.8877	29.8611	29.8799	30.0851	29.9480	85.2583
Skewed t copula	29.6312	29.4941	29.6635	29.6914	29.6591	81.6776

Conclusion

- Based on a comparison of the goodness of fit criteria, the empirical evidence shows that the ARMA-GARCH model with fractional stable noise demonstrates better performance in modeling univariate high-frequency time series data.

- By using the same criteria of goodness of fit test in comparing marginal distributions, the multivariate Student's t copula with fractional stable ARMA-GARCH model has superior performance when modeling the co-movement of nine global equity market indexes.

- When the multivariate time series data exhibit asymmetric correlation, the multivariate skewed Student's t copula with fractional stable ARMA-GARCH model has superior performance when modeling the co-movement of six German equity market indexes.

- The advantage of the empirical study is threefold. First, using multi-dimensional copulas can reveal the tail dependence of in co-movement of several assets. Second, our model can capture long-range dependence, heavy tails, volatility clustering, and tail dependence simultaneously. Third, using high-frequency data, the impact of both macroeconomic factors and microstructure effects on asset return can be considered.