Volatility is rough

Jim Gatheral[1], Thibault Jaisson[2] and Mathieu Rosenbaum[3]

[1]City University of New York, [2]École Polytechnique Paris, [3]University Paris 6

10 December 2014

Table of contents

Table of contents

Main classes of volatility models

Prices are often modeled as continuous semi-martingales of the form

$$dP_t = P_t(\mu_t dt + \sigma_t dW_t).$$

The volatility process σ_s is the most important ingredient of the model. Practitioners consider essentially three classes of volatility models :

- Deterministic volatility (Black and Scholes 1973),
- Local volatility (Dupire 1994),
- Stochastic volatility (Hull and White 1987, Heston 1993, Hagan et al. 2002,...).

In term of regularity, in these models, the volatility is either very smooth or with a smoothness similar to that of a Brownian motion.

Long memory in volatility

Definition

A stationary process is said to be long memory if its autocovariance function decays slowly, more precisely :

$$\sum_{t=1}^{+\infty} \mathrm{Cov}[\sigma_{t+x}, \sigma_x] = +\infty.$$

Power law long memory for the volatility :

$$\mathrm{Cov}[\sigma_{t+x}, \sigma_x] \sim C/t^{\gamma},$$

with $\gamma < 1$, is considered a stylized fact and has been notably reported in Ding and Granger 1993 (using extra day data) and Andersen *et al.*, 2001 and 2003 (using intra day data).

Fractional Brownian motion (I)

To take into account the long memory property and to allow for a wider range of smoothness, some authors have introduced the fractional Brownian motion in volatility modeling.

Definition

The fractional Brownian motion (fBm) with Hurst parameter H is the only process W^H to satisfy :

- Self-similarity : $(W^H_{at}) \overset{\mathcal{L}}{=} a^H (W^H_t)$.
- Stationary increments : $(W^H_{t+h} - W^H_t) \overset{\mathcal{L}}{=} (W^H_h)$.
- Gaussian process with $\mathbb{E}[W^H_1] = 0$ and $\mathbb{E}[(W^H_1)^2] = 1$.

Fractional Brownian motion (II)

Proposition

For all $\varepsilon > 0$, W^H is $(H - \varepsilon)$-Hölder a.s.

Proposition

The absolute moments of the increments of the fBm satisfy

$$\mathbb{E}[|W^H_{t+h} - W^H_t|^q] = K_q h^{Hq}.$$

Proposition

If $H > 1/2$, the fBm exhibits long memory in the sense that

$$\mathrm{Cov}[W^H_{t+1} - W^H_t, W^H_1] \sim \frac{C}{t^{2-2H}}.$$

Long memory volatility models

Some models have been built using fractional Brownian motion with Hurst parameter $H > 1/2$ to reproduce the long memory property of the volatility :

- Comte and Renault 1998 (FSV model) :

$$d \log(\sigma_t) = \nu dW_t^H + \alpha(m - \log(\sigma_t))dt.$$

- Comte, Coutin and Renault 2012, where they define a kind of fractional CIR process.

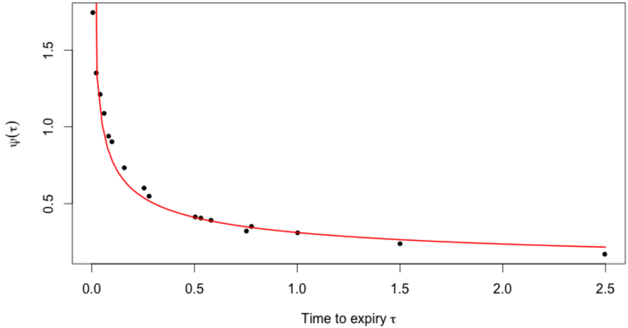
About option data

- Classical stochastic volatility models generate reasonable dynamics for the volatility surface.
- However they do not allow to fit the volatility surface, in particular the term structure of the ATM skew :

$$\psi(\tau) := \left| \frac{\partial}{\partial k} \sigma_{\mathrm{BS}}(k, \tau) \right|_{k=0},$$

where k is the log-moneyness and τ the maturity of the option.

About option data : the volatility skew



The black dots are non-parametric estimates of the S&P ATM volatility skews as of June 20, 2013 ; the red curve is the power-law fit $\psi(\tau) = A\,\tau^{-0.4}$.

About option data : fractional volatility

- The skew is well-approximated by a power-law function of time to expiry τ. In contrast, conventional stochastic volatility models generate a term structure of ATM skew that is constant for small τ.
- Models where the volatility is driven by a fBm generate an ATM volatility skew of the form $\psi(\tau) \sim \tau^{H-1/2}$, at least for small τ.

Table of contents

Intraday volatility estimation

We are interested in the dynamics of the (log)-volatility process. We use two proxies for the spot (squared) volatility of a day.

- A 5 minutes-sampling realized variance estimation taken over the whole trading day (8 hours).
- A one hour integrated variance estimator based on the model with uncertainty zones (Robert and R. 2012).

Note that we are not really considering a "spot" volatility but an "integrated" volatility. This might lead to some slight bias in our measurements (which can be controlled).

From now on, we consider realized variance estimations on the S&P over 3500 days, but the results are fairly "universal".

The log-volatility

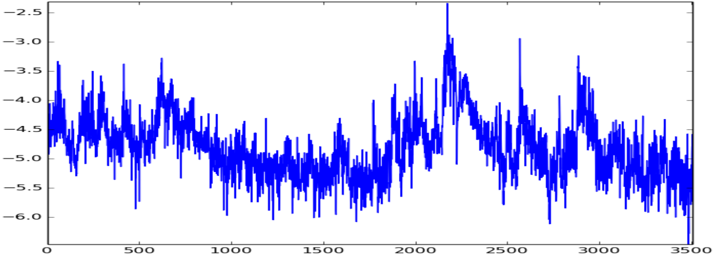


FIGURE : The log volatility $\log(\sigma_t)$ as a function of t, S&P.

Measure of the regularity of the log-volatility

The starting point of this work is to consider the scaling of the moments of the increments of the log-volatility. Thus we study the quantity

$$m(\Delta, q) = \mathbb{E}[|\log(\sigma_{t+\Delta}) - \log(\sigma_t)|^q],$$

or rather its empirical counterpart.

The behavior of $m(\Delta, q)$ when Δ is close to zero is related to the smoothness of the volatility (in the Hölder or even the Besov sense). Essentially, the regularity of the signal measured in l^q norm is s if $m(\Delta, q) \sim c\Delta^{qs}$ as Δ tends to zero.

Scaling of the moments

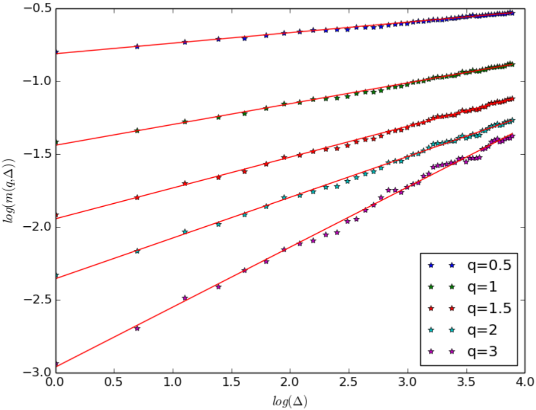


FIGURE : $\log(m(q,\Delta)) = \zeta_q \log(\Delta) + C_q$. The scaling is not only valid as Δ tends to zero, but holds on a wide range of time scales.

Monofractality of the log-volatility

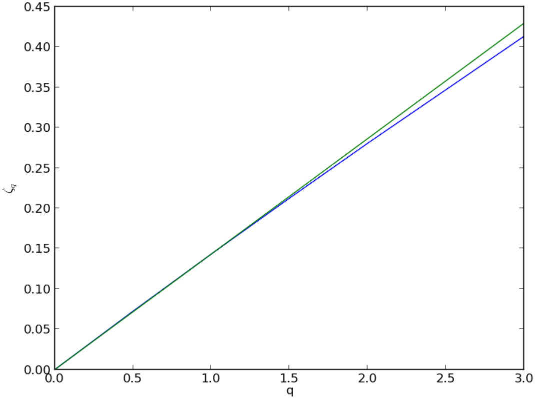


FIGURE : Empirical ζ_q and $q \to Hq$ with $H = 0.14$ (similar to a fBm with Hurst parameter H).

Distribution of the log-volatility increments

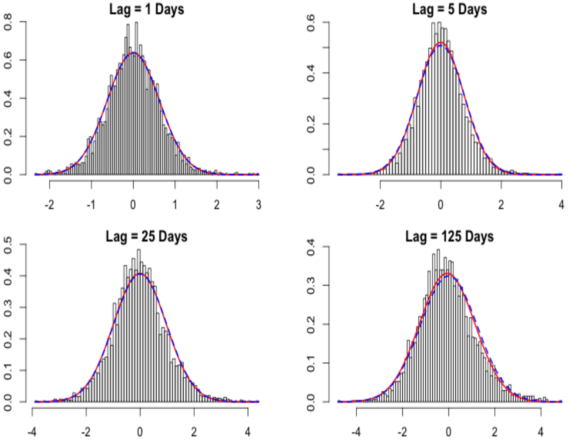


FIGURE : The distribution of the log-volatility increments is close to Gaussian.

A geometric fBm model ?

These empirical findings suggest we model the log-volatility as a fractional Brownian motion :

$$\sigma_t = \sigma e^{\nu W_t^H}.$$

However, this model is not stationary. In particular, the empirical autocovariance function of the (log-)volatility (which will be of interest) does not make much sense.

A geometric fOU model

We make it formally stationary by considering a fractional Ornstein-Uhlenbeck model for the log-volatility denoted by X_t

$$dX_t = \nu dW_t^H + \alpha(m - X_t)dt.$$

This process satisfies

$$X_t = \nu \int_{-\infty}^{t} e^{-\alpha(t-s)} dW_t^H + m.$$

We take the reversion time scale $1/\alpha$ very large compared to the observation time scale.

This model is a particular case of the FSV model. However, in strong contrast to FSV, we take H small and $1/\alpha$ large. Thus we call our model Rough FSV (RFSV).

$m(2, \Delta)$ and the parameters of the FSV

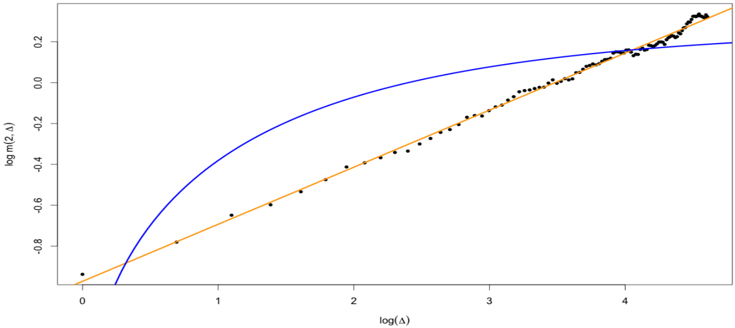


FIGURE : $\log(m(2, \Delta))$ as a function of $\log(\Delta)$ in our data (black) and in the FSV model with $H = 0.56$ (there is a closed formula) (blue). On real data, the scaling is not only valid as Δ tends to zero but holds on a wide range of Δ. In the FSV, the slope at the beginning of the graph is governed by the parameter H and then stationarity kicks in.

Behavior of our model at reasonable time scales

When the reversion time scale becomes large ($\alpha \to 0$), the mean reverting term is negligible and the log-volatility is almost a fBm.

Proposition

As α tends to zero,

$$\mathbb{E}\Big[\sup_{t \in [0,T]} |X_t^\alpha - X_0^\alpha - \nu W_t^H|\Big] \to 0.$$

Table of contents

Autocorrelogram of the (log-)volatility in our model

Proposition

Let $t > 0$, $\Delta > 0$. As α tends to zero,

$$\mathrm{Cov}[X_t^\alpha, X_{t+\Delta}^\alpha] = \mathrm{Var}[X_t^\alpha] - \frac{1}{2}\nu^2\,\Delta^{2H} + o(1).$$

Proposition

As α tends to zero,

$$\mathbb{E}[\sigma_{t+\Delta}\sigma_t] = e^{2\mathbb{E}[X_t^\alpha]+2\mathrm{Var}[X_t^\alpha]}e^{-\nu^2\frac{\Delta^{2H}}{2}} + o(1).$$

We now check these relations on the data.

Empirical autocorrelogram of the log-volatility

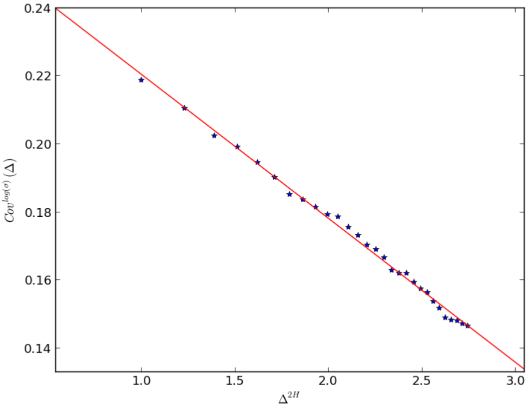


FIGURE : $\mathrm{Cov}[\log(\sigma_t), \log(\sigma_{t+\Delta})]$ as a function of Δ^{2H}. This fits the prediction of our model.

Empirical autocorrelogram of the volatility

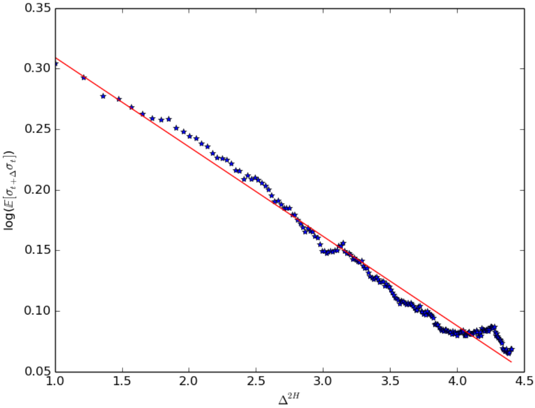


FIGURE : $\log(\mathbb{E}[\sigma_t\sigma_{t+\Delta}])$ as a function of Δ^{2H}. This fits again the prediction of our model.

Long memory in volatility

It is widely believed that the (log-)(squared-)volatility exhibits power law long memory

$$\mathrm{Cov}[\sigma_{x+t}, \sigma_x] \underset{t \to +\infty}{\sim} \frac{k}{t^\gamma},$$

with $\gamma < 1$.

We review two tests for this long memory property. We show that they can wrongly deduce a power law long memory on data generated from our model and are thus quite fragile.

Log-log autocovariance of the volatility

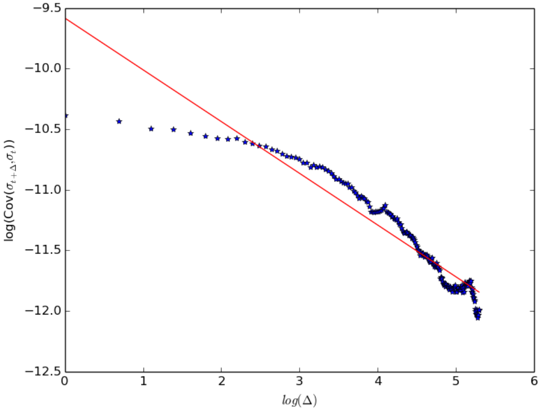


FIGURE : $\log(\mathrm{Cov}[\sigma_{x+\Delta}, \sigma_x])$ as a function of $\log(\Delta)$. The autocorrelation function does not behave as a power law function.

Scaling of the variance of the cumulated volatility

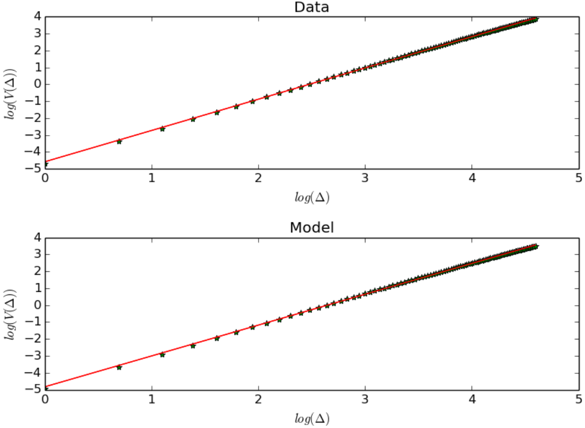


FIGURE : $V(\Delta) = \text{Var}[\sum_{t=1}^{\Delta} \sigma_t]$ as a function of $\log(\Delta)$ on empirical (above) and simulated (below) data. Power law long memory implies that it should behave as $\Delta^{2-\gamma}$, as we observe on the data and the model.

Fractional differentiation of the log-volatility

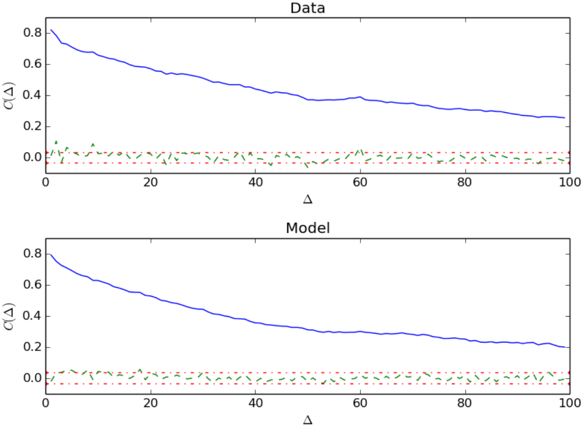


FIGURE : ACF of the log-volatility (blue) and of $\varepsilon = (1 - L)^d \log(\sigma)$, with $d = 0.4$ (green) on empirical (above) and simulated (below) data.

Multiscaling in finance

- An important property of volatility time series is their multiscaling behavior, see Mantegna and Stanley 2000 and Bouchaud and Potters 2003.
- This means one observes essentially the same law whatever the time scale.
- In particular, there are periods of high and low market activity at different time scales.
- Very few models reproduce this property, see multifractal models (Bacry *et al.*,...).

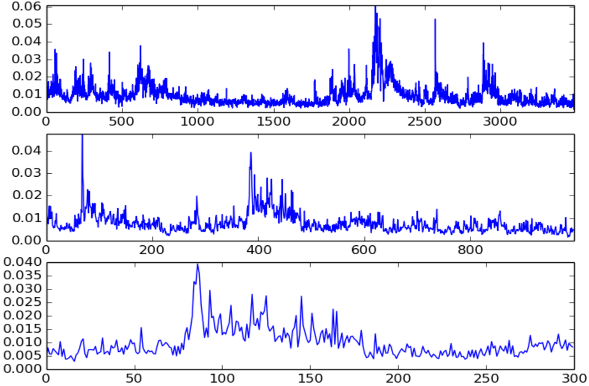


FIGURE : Empirical volatility over 10, 3 and 1 years.

Our model on different time intervals

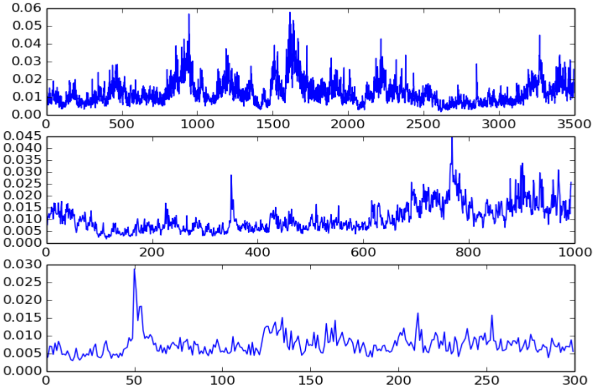


FIGURE : Simulated volatility over 10, 3 and 1 years. We observe the same alternations of periods of high market activity with periods of low market activity.

Apparent multiscaling in our model

- Let $L^{H,\nu}$ be the law on $[0,1]$ of the process $e^{\nu W_t^H}$.
- Then the law of the volatility process on $[0,T]$ renormalized on $[0,1]$: σ_{tT}/σ_0 is $L^{H,\nu T^H}$.
- If one observes the volatility on $T = 10$ years (2500 days) instead of $T = 1$ day, the parameter νT^H defining the law of the volatility is only multiplied by $2500^H \sim 3$.
- Therefore, one observes quite the same properties on a very wide range of time scales.
- The roughness of the volatility process ($H = 0.14$) implies a multiscaling behavior of the volatility.

Table of contents

Prediction of a fractional Brownian motion

There is a nice prediction formula for the fractional Brownian motion.

Proposition (Nuzman and Poor 2000)

For $H < 1/2$

$$\mathbb{E}[W^H_{t+\Delta}|\mathcal{F}_t] = \frac{\cos(H\pi)}{\pi}\Delta^{H+1/2}\int_{-\infty}^{t}\frac{W^H_s}{(t-s+\Delta)(t-s)^{H+1/2}}ds.$$

Our prediction formula

We apply the previous formula to the prediction of the log-volatility :

$$\mathbb{E}\left[\log\sigma^2_{t+\Delta}|\mathcal{F}_t\right] = \frac{\cos(H\pi)}{\pi}\Delta^{H+1/2}\int_{-\infty}^{t}\frac{\log\sigma^2_s}{(t-s+\Delta)(t-s)^{H+1/2}}ds$$

or more precisely its discrete version :

$$\mathbb{E}\left[\log\sigma^2_{t+\Delta}|\mathcal{F}_t\right] = \frac{\cos(H\pi)}{\pi}\Delta^{H+1/2}\sum_{k=0}^{N}\frac{\log\sigma^2_{t-k}}{(k+\Delta+1/2)(k+1/2)^{H+1/2}}.$$

We compare it to usual predictors using the criterion

$$P = \frac{\sum_{k=1}^{N-\Delta}(\widehat{\log(\sigma^2_{k+\Delta})} - \log(\sigma^2_{k+\Delta}))^2}{\sum_{k=1}^{N-\Delta}(\log(\sigma^2_{k+\Delta}) - \mathbb{E}[\log(\sigma^2_{t+\Delta})])^2}.$$

	AR(5)	AR(10)	HAR(3)	RFSV
SPX2.rv $\Delta = 1$	0.317	0.318	0.314	**0.313**
SPX2.rv $\Delta = 5$	0.459	0.449	0.437	**0.426**
SPX2.rv $\Delta = 20$	0.764	0.694	0.656	**0.606**
FTSE2.rv $\Delta = 1$	0.230	0.229	0.225	**0.223**
FTSE2.rv $\Delta = 5$	0.357	0.344	0.337	**0.320**
FTSE2.rv $\Delta = 20$	0.651	0.571	0.541	**0.472**
N2252.rv $\Delta = 1$	0.357	0.358	0.351	**0.345**
N2252.rv $\Delta = 5$	0.553	0.533	0.513	**0.504**
N2252.rv $\Delta = 20$	0.875	0.795	0.746	**0.714**
GDAXI2.rv $\Delta = 1$	0.237	0.238	0.234	**0.231**
GDAXI2.rv $\Delta = 5$	0.372	0.362	0.350	**0.339**
GDAXI2.rv $\Delta = 20$	0.661	0.590	0.550	**0.498**
FCHI2.rv $\Delta = 1$	0.244	0.244	0.241	**0.238**
FCHI2.rv $\Delta = 5$	0.378	0.373	0.366	**0.350**
FCHI2.rv $\Delta = 20$	0.669	0.613	0.598	**0.522**

Regression window and horizon

After a simple change of variable, the prediction of the log-volatility can be written :

$$\mathbb{E}[\log(\sigma_{t+\Delta}^2)|\mathcal{F}_t] \sim \frac{\cos(H\pi)}{\pi} \int_0^1 \frac{\log(\sigma_{t-\Delta u}^2)}{(u+1)\, u^{H+1/2}} du.$$

The only time scale that appears in the above regression is the horizon Δ.

As it is known by practitioners :

If trying to predict volatility one week ahead, one should essentially look at the volatility over the last week. If trying to predict the volatility one month ahead, one should essentially look at the volatility over the last month.

Conditional distribution of the fractional Brownian motion and prediction of the variance

Proposition (Nuzman and Poor 2000)

In law,

$$W_{t+\Delta}|\mathcal{F}_t = \mathcal{N}(\mathbb{E}[W_{t+\Delta}|\mathcal{F}_t], c\Delta^{2H})$$

with

$$c = \frac{\sin(\pi(1/2-H))\Gamma(3/2-H)^2}{\pi(1/2-H)\Gamma(2-2H)}.$$

Therefore, our predictor of the variance writes :

$$\mathbb{E}[\sigma^2_{t+\Delta}|\mathcal{F}_t] = e^{\mathbb{E}\left[\log(\sigma^2_{t+\Delta})|\mathcal{F}_t\right]+2\nu^2 c\Delta^{2H}}.$$

	AR(5)	AR(10)	HAR(3)	RFSV
SPX2.rv $\Delta = 1$	0.520	0.566	0.489	**0.475**
SPX2.rv $\Delta = 5$	0.750	0.745	0.723	**0.672**
SPX2.rv $\Delta = 20$	1.070	1.010	1.036	**0.903**
FTSE2.rv $\Delta = 1$	0.612	0.621	0.582	**0.567**
FTSE2.rv $\Delta = 5$	0.797	0.770	0.756	**0.707**
FTSE2.rv $\Delta = 20$	1.046	0.984	0.935	**0.874**
N2252.rv $\Delta = 1$	0.554	0.579	**0.504**	0.505
N2252.rv $\Delta = 5$	0.857	0.807	0.761	**0.729**
N2252.rv $\Delta = 20$	1.097	1.046	1.011	**0.964**
GDAXI2.rv $\Delta = 1$	0.439	0.448	0.399	**0.386**
GDAXI2.rv $\Delta = 5$	0.675	0.650	0.616	**0.566**
GDAXI2.rv $\Delta = 20$	0.931	0.850	0.816	**0.746**
FCHI2.rv $\Delta = 1$	0.533	0.542	0.470	**0.465**
FCHI2.rv $\Delta = 5$	0.705	0.707	0.691	**0.631**
FCHI2.rv $\Delta = 20$	0.982	0.952	0.912	**0.828**

Table of contents

Definition

Hawkes processes as models for the order flow

- The starting point of our microstructural analysis is the modeling of the order flow though Hawkes processes.
- A Hawkes process $(N_t)_{t\geq 0}$ is a self exciting point process, whose intensity at time t, denoted by λ_t, is of the form

$$\lambda_t = \mu + \sum_{0<J_i<t} \phi(t - J_i) = \mu + \int_{(0,t)} \phi(t-s)dN_s,$$

  where μ is a positive real number, ϕ a regression kernel and the J_i are the points of the process before time t.
- These processes have been introduced in 1971 by Hawkes in the purpose of modeling earthquakes and their aftershocks and are nowadays very popular in finance.

Hawkes processes in practice

Nearly unstable heavy-tailed Hawkes processes

When trying to calibrate such models on high frequency data, two main phenomena almost systematically occur :

- The L^1 norm of ϕ close to one$\rightarrow$ high degree of endogeneity of the market due to high frequency trading, see Bouchaud et al. 2013, Filimonov and Sornette 2013.
- The function ϕ has a power law tail$\rightarrow$ metaorders splitting.

Assumptions and asymptotic framework

Sequence of Hawkes processes

- We consider a sequence of point processes $(N_t^T)_{t\geq 0}$ indexed by T. We have $N_0^T = 0$ and the process is observed on the time interval $[0, T]$. Furthermore, our asymptotic setting is that the observation scale T goes to infinity.
- The intensity process (λ_t^T) is defined for $t \geq 0$ by

$$\lambda_t^T = \mu^T + \int_0^t \phi^T(t-s)dN_s^T,$$

where μ^T is a sequence of positive real number and ϕ^T a non negative measurable function on $\mathbb{R}^+$ which satisfies $\|\phi^T\|_1 < 1$.

Assumptions and asymptotic framework

Assumptions

There is some $\alpha \in (0,1)$ such that

$$\phi(x) \underset{x \to +\infty}{\sim} \frac{K}{x^{1+\alpha}}, \quad \lim_{T \to +\infty} T^{\alpha}(1 - \|\phi^T\|_1) = \delta\lambda > 0.$$

Normalized processes

We investigate the limit in law as T goes to infinity of the sequence of processes

$$a_T N^T_{tT}, \ t \in [0,1],$$

with a_T a suitable normalizing factor.

Agent based explanation for the behavior of the volatility

Limit theorem

For $\alpha > 1/2$, the sequence of renormalized Hawkes processes converges to some process which is differentiable on $[0,1]$. Moreover, the law of its derivative Y satisfies

$$Y_t = F^{\alpha,\lambda}(t) + \frac{1}{\sqrt{\mu^*\lambda}} \int_0^t f^{\alpha,\lambda}(t-s)\sqrt{Y_s}dB_s^1,$$

with B^1 a Brownian motion and

$$f^{\alpha,\lambda}(x) = \lambda x^{\alpha-1} E_{\alpha,\alpha}(-\lambda x^{\alpha}).$$

Therefore $H = \alpha - 1/2$. Furthermore, for any $\varepsilon > 0$, Y has Hölder regularity $\alpha - 1/2 - \varepsilon$.

Agent based explanation for the behavior of the volatility

Microstructural foundations for the RFSV model

- Then, it is clearly established that there is a linear relationship between cumulated order flow and integrated variance.
- Thus endogeneity of the market together with order splitting lead to a superposition effect which explains (at least partly) the rough nature of the observed volatility.