Rough volatility and leverage effect: From microscopic foundations to smile

Omar El Euch[1], Masaaki Fukasawa[2], Jim Gatheral[3], Thibault Jaisson[4], Mathieu Rosenbaum[1]

[1]University Paris 6, [2]City University of New York, [3]Osaka University, [4]Pictet

15 June 2016

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. The three most classical classes of volatility models are :

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

However, it has been recently shown that models where the volatility is driven by a fractional Brownian motion enable us to reproduce very well the behavior of historical data and of the volatility surface.

Fractional Brownian motion (I)

Definition

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

- Self-similarity : $(W^H_{at}) \stackrel{\mathcal{L}}{=} a^H (W^H_t)$.
- Stationary increments : $(W^H_{t+h} - W^H_t) \stackrel{\mathcal{L}}{=} (W^H_h)$.
- 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

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

Fractional models

FSV model

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

- Comte and Renault 1998 (FSV model) :

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

Here α is large to model a mean reversion effect.

Fractional models

RFSV model

However, statistical investigation of recent prices and options data rather suggests the use of rough versions of the preceding model, for example :

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

with H of order 0.1 and α very small (Rough FSV model).

Properties of RFSV-type models

Statistical analysis of the RFSV model

- Reproduces very well (almost) all the statistical stylized facts of volatility, with explicit formulas, see Gatheral *et al.*
- Very good fit of the volatility surface, in particular of the ATM skew.
- No power law long memory property.
- Applied to the RFSV model, statistical tests for long memory behave the same way as for real data and deduce, probably wrongly, the presence of long memory in the volatility.
- Multiscaling behaviour.
- Explicit prediction formulas for the future volatility, depending only on the parameter H, outperforming classical predictors. To forecast the volatility at time $t+\Delta$, one needs to consider the data in the past until $t-\Delta$.

Leverage effect and rough volatility

Leverage effect

- The leverage effect is a well studied phenomenon : negative correlation between price increments and volatility increments.
- Very easy to incorporate within a rough volatility framework : Use Mandelbrot-van Ness representation of the fractional Brownian motion :

$$W_t^H = \int_0^t \frac{dW_s}{(t-s)^{\frac{1}{2}-H}} + \int_{-\infty}^0 \left(\frac{1}{(t-s)^{\frac{1}{2}-H}} - \frac{1}{(-s)^{\frac{1}{2}-H}}\right) dW_s,$$

and correlate W with the Brownian motion driving the price.

Table of contents

Building the model

Necessary conditions for a good microscopic price model

We want :

- A tick-by-tick model.
- A model reproducing the stylized facts of modern electronic markets in the context of high frequency trading.
- A model helping us to understand the rough dynamics of the volatility from the high frequency behaviour of market participants.
- A model helping us to understand leverage effect.

How is leverage effect generated ?

Traditional macroscopic explanations for leverage effect

- Asset price declines→ company becomes automatically more leveraged since the ratio of its debt with respect to the equity value becomes larger→ risk of the asset (the volatility) should become more important.
- Forecast of an increase of the volatility should be compensated by a higher rate of return, which can only be obtained through a decrease in the asset value.

Microstructural component for leverage effect ?

- We want to address the following question : Can leverage effect be partly generated from high frequency features of the asset ?

Building the model

Stylized facts 1-2

- Markets are highly endogenous, meaning that most of the orders have no real economic motivations but are rather sent by algorithms in reaction to other orders, see Bouchaud *et al.*, Filimonov and Sornette.
- Mechanisms preventing statistical arbitrages take place on high frequency markets, meaning that at the high frequency scale, building strategies that are on average profitable is hardly possible.

Building the model

Stylized facts 3-4

- There is some asymmetry in the liquidity on the bid and ask sides of the order book. In particular, a market maker is likely to raise the price by less following a buy order than to lower the price following the same size sell order, see Brennan *et al.*, Brunnermeier and Pedersen, Hendershott and Seasholes.
- A large proportion of transactions is due to large orders, called metaorders, which are not executed at once but split in time.

Building the model

Hawkes processes

- Our tick-by-tick price model is based on Hawkes processes in dimension two, very much inspired by the approaches in Bacry *et al.* and Jaisson and R.
- A two-dimensional Hawkes process is a bivariate point process $(N_t^+, N_t^-)_{t\geq 0}$ taking values in $(\mathbb{R}^+)^2$ and with intensity $(\lambda_t^+, \lambda_t^-)$ of the form :

$$\begin{pmatrix} \lambda_t^+ \\ \lambda_t^- \end{pmatrix} = \begin{pmatrix} \mu^+ \\ \mu^- \end{pmatrix} + \int_0^t \begin{pmatrix} \varphi_1(t-s) & \varphi_3(t-s) \\ \varphi_2(t-s) & \varphi_4(t-s) \end{pmatrix} . \begin{pmatrix} dN_s^+ \\ dN_s^- \end{pmatrix} .$$

Building the model

The microscopic price model

- Our model is simply given by

$$P_t = N_t^+ - N_t^-.$$

- N_t^+ corresponds to the number of upward jumps of the asset in the time interval $[0, t]$ and N_t^- to the number of downward jumps. Hence, the instantaneous probability to get an upward (downward) jump depends on the location in time of the past upward and downward jumps.
- By construction, the price process lives on a discrete grid.
- Statistical properties of this model have been studied in details in Bacry *et al.*

Encoding the stylized facts

The right parametrization of the model

- Recall that

$$\begin{pmatrix} \lambda_t^+ \\ \lambda_t^- \end{pmatrix} = \begin{pmatrix} \mu^+ \\ \mu^- \end{pmatrix} + \int_0^t \begin{pmatrix} \varphi_1(t-s) & \varphi_3(t-s) \\ \varphi_2(t-s) & \varphi_4(t-s) \end{pmatrix} . \begin{pmatrix} dN_s^+ \\ dN_s^- \end{pmatrix}.$$

- High degree of endogeneity of the market→ L^1 norm of the largest eigenvalue of the kernel matrix close to one.
- No arbitrage→ $\varphi_1 + \varphi_3 = \varphi_2 + \varphi_4$.
- Liquidity asymmetry→ $\varphi_3 = \beta\varphi_2$, with $\beta > 1$.
- Metaorders splitting→ $\varphi_1(x),\ \varphi_2(x) \underset{x\to\infty}{\sim} K/x^{1+\alpha}$, $\alpha \approx 0.6$.

About the degree of endogeneity of the market

L^1 norm close to unity

- For simplicity, let us consider the case of Hawkes process in dimension 1 with kernel ϕ : N_t then represents the number of transactions between time 0 and time t.
- L^1 norm of the largest eigenvalue close to unity$\rightarrow L^1$ norm of ϕ close to unity. This is systematically observed in practice, see Hardiman, Bercot and Bouchaud ; Filimonov and Sornette.
- The parameter $\|\phi\|_1$ corresponds to the so-called degree of endogeneity of the market.

About the degree of endogeneity of the market

Population interpretation of Hawkes processes

- The parameter $\|\phi\|_1$ corresponds to the average number of children of an individual, $\|\phi\|_1^2$ to the average number of grandchildren of an individual,... Therefore, if we call cluster the descendants of a migrant, then the average size of a cluster is given by $\sum_{k\geq 1} \|\phi\|_1^k = \|\phi\|_1/(1 - \|\phi\|_1)$.
- Thus, the average proportion of endogenously triggered events is $\|\phi\|_1/(1 - \|\phi\|_1)$ divided by $1 + \|\phi\|_1/(1 - \|\phi\|_1)$, which is equal to $\|\phi\|_1$.

The scaling limit of the price model

Limit theorem

After suitable scaling in time and space, the long term limit of our price model satisfies the following Rough Heston dynamics :

$$P_t = \int_0^t \sqrt{V_s} dW_s - \frac{1}{2} \int_0^t V_s ds,$$

$$V_t = V_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda(\theta - V_s) ds + \frac{\lambda\nu}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \sqrt{V_s} dB_s,$$

with

$$d\langle W, B\rangle_t = \frac{1-\beta}{\sqrt{2(1+\beta^2)}} dt.$$

The scaling limit of the price model

Comments on the theorem

- The Hurst parameter $H = \alpha - 1/2$.
- Hence stylized facts of modern market microstructure naturally give rise to fractional dynamics and leverage effect.
- One of the only cases of scaling limit of a non ad hoc "micro model" where leverage effect appears in the limit. Compare with Nelson's limit of GARCH models for example.
- Uniqueness of the limiting solution is a difficult result. The proof requires the use of recent results in SPDEs theory by Mytnik and Salisbury.

Table of contents

Rough Heston model

Rough version of Heston model

- No clear definition of the Rough Heston model. Several candidates, relatively similar, see also Guennoun, Jacquier and Roome.
- We consider the following model :

$$dS_t = S_t \sqrt{V_t} dW_t,$$

$$V_t = V_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda(\theta - V_s) ds + \frac{\lambda\nu}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \sqrt{V_s} dB_s,$$

with

$$d\langle W, B \rangle_t = \rho dt.$$

Computing the characteristic functions

From microstructure to smile

- Deriving characteristic functions for our microscopic Hawkes-based price model and passing to the limit, we are able to compute characteristic functions in the Rough Heston model.

Characteristic functions

We write :

$$I^{1-\alpha}f(x) = \frac{1}{\Gamma(1-\alpha)}\int_0^x \frac{f(t)}{(x-t)^\alpha}dt, \;\; D^\alpha f(x) = \frac{d}{dx}I^{1-\alpha}f(x).$$

Theorem

The characteristic function at time t for the Rough Heston model is given by

$$L(a,t) = \exp\Big(\int_0^t g(a,s)ds + \frac{V_0}{\theta\lambda}I^{1-\alpha}g(a,t)\Big),$$

with $g(a,)$ the unique solution of the fractional Riccati equation :

$$D^\alpha g(a,s) = \frac{\lambda\theta}{2}(-a^2 - ia) + \lambda(ia\rho\nu - 1)g(a,s) + \frac{\lambda\nu^2}{2\theta}g^2(a,s).$$