From no-arbitrage to rough volatility via market impact

Paul Jusselin and Mathieu Rosenbaum

École Polytechnique

2 February 2022

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 (Derman and Kani, 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.

Fractional Brownian motion (I)

- To allow for a wider range of smoothness, one can use the fractional Brownian motion in volatility modeling.
- Idea introduced by Comte and Renault in 1998 in the context of long memory modeling with $H > 1/2$.

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 satisfy

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

Mandelbrot-van Ness representation

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

The log-volatility

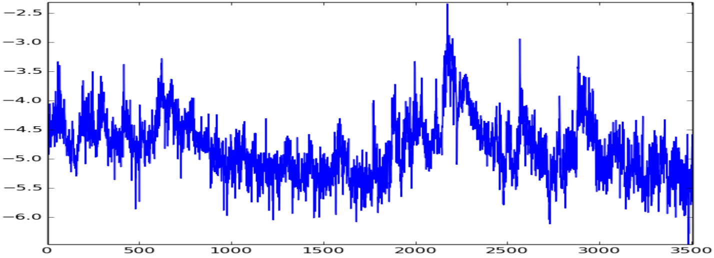

FIGURE – The log volatility $\log(\sigma_t)$ as a function of t, S&P.

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

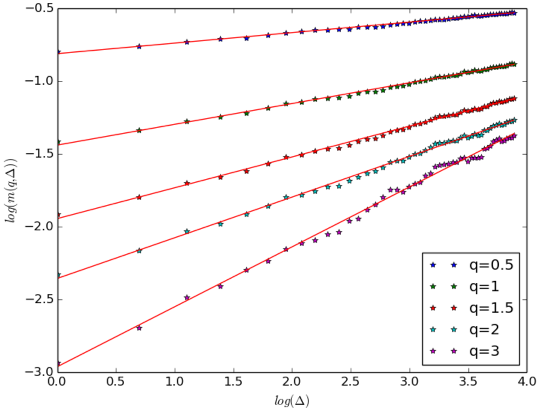

FIGURE – $\log(m(q,\Delta)) = \zeta_q \log(\Delta) + C_q$. The scaling is not only valid as Δ tends to zero, but holds on a wide range of time scales.

Monofractality of the log-volatility

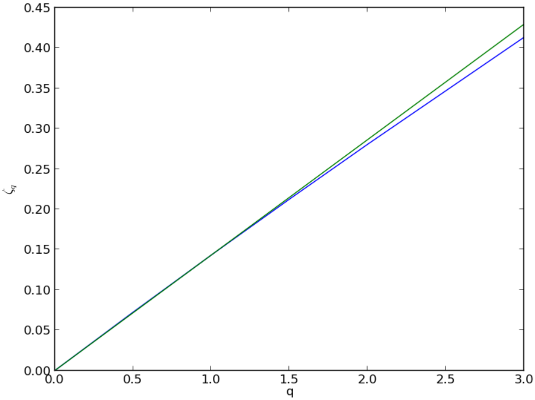

FIGURE – Empirical ζ_q and $q \to Hq$ with $H = 0.14$ (similar to a fBm with Hurst parameter H).

Distribution of the log-volatility increments

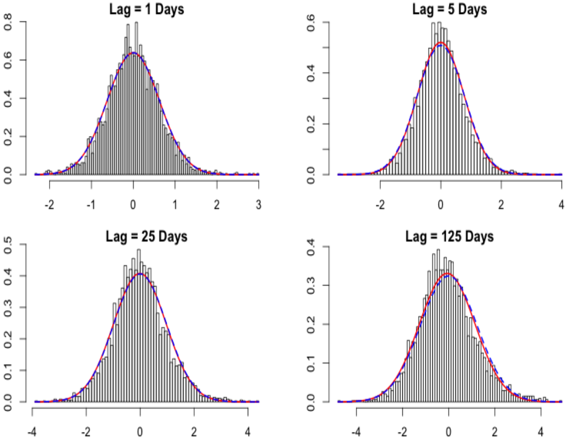

FIGURE – The distribution of the log-volatility increments is close to Gaussian.

Properties of the rough volatility models

Statistical analysis of rough volatility models

- The log-volatility behaves essentially as a fractional Brownian motion with Hurst parameter of order 0.1.
- More precisely, basically all the statistical stylized facts of volatility are retrieved when modeling it by a rough fractional Brownian motion.
- Such model also enables us to reproduce very well the behavior of the implied volatility surface, in particular the at-the-money skew (without jumps).
- Also very relevant for risk management of derivatives (closed form formulas, see for example the rough Heston model).
- The phenomenon is universal.

In this presentation

What we want to understand :

- Why is volatility rough ?
- Something universal in finance$\rightarrow$ should be related to some no arbitrage concept.
- Can we make this link ?

Table of contents

Market impact

Some definitions

- Market impact is the link between the volume of an order (either market order or metaorder) and the price moves during and after the execution of this order.
- We focus here on the impact function of metaorders, which is the expectation of the price move with respect to time during and after the execution of the metaorder.
- We call permanent market impact of a metaorder the limit in time of the impact function (that is the average price move between the start of the metaorder and a long time after its execution).

Market impact in practice, from Lillo et al.

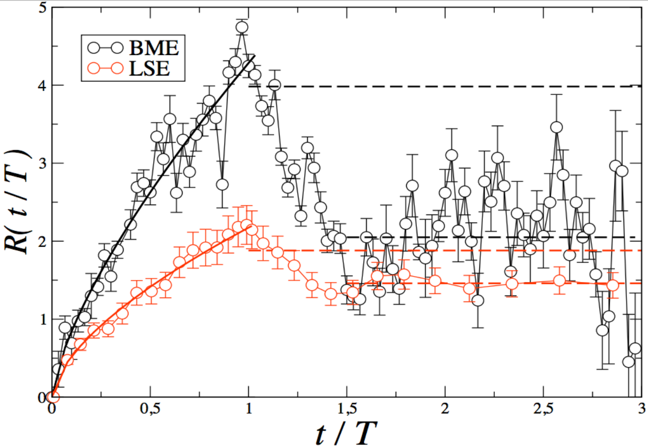

FIGURE – Market impact curves.

Market impact

Linear permanent impact

- Let P_t be the asset price at time t. Consider a metaorder with total volume V.
- $$PMI(V) = \lim_{s \to +\infty} \mathbb{E}[P_s - P_0 | V].$$
- Price manipulation is a roundtrip with negative average cost.
- From Huberman and Stanzl and Gatheral : Only linear permanent market impact can prevent price manipulation : $PMI(V) = kV$.

Market impact

CAPM like argument for linear permanent impact

- n investors in the market. Two dates : $t = 0$ and $t = 1$.
- N shares spread between the agents, price P for the asset.
- Every investor i estimates that the law of the price at time 1 has expectation E_i and variance Σ_i. He chooses his number of asset N_i such that

$$N_i = \text{argmax}_x[x(E_i - P) - \lambda_i x^2 \Sigma_i].$$

- We get

$$N_i = \frac{E_i - P}{2\lambda_i \Sigma_i}.$$

Market impact

CAPM like argument for linear permanent impact

- Since $\sum_{i=1}^n N_i = N$, we deduce

$$P = \frac{\sum_{i=1}^n \frac{E_i}{2\lambda_i \Sigma_i} - N}{\sum_{i=1}^n \frac{1}{2\lambda_i \Sigma_i}}.$$

- Let us now assume that the total number of shares becomes $N - N_0$ due to the action of some non-optimizing agent needing to buy some shares (for cash flow reasons for example). The new indifference price is

$$P^+ = P + \frac{N_0}{\sum_{i=1}^n \frac{1}{2\lambda_i \Sigma_i}} = P + kN_0.$$

Dynamics

Assumptions

- All market orders are part of metaorders.
- Let $[0, S]$ be the time during which metaorders are being executed (which can be thought of as the trading day). Let v_i^a (resp. v_i^b) be the volume of the i-th buy (resp. sell) metaorder and N_S^a (resp. N_S^b) be the number of buy (resp. sell) metaorders up to time S. Finally, write V_S^a and V_S^b for cumulated buy and sell order flows up to time S.
- We assume

$$P_S = P_0 + k\Big(\sum_{i=1}^{N_S^a} v_i^a - \sum_{i=1}^{N_S^b} v_i^b\Big) + Z_S = P_0 + k(V_S^a - V_S^b) + Z_S,$$

with Z a martingale term that we neglect.

Dynamics

Martingale assumption

- We furthermore assume that the price P_t is a martingale. We obtain

$$P_t = P_0 + \mathbb{E}\big[k(V^a_S - V^b_S)|F_t\big].$$

- We suppose that $\lim\limits_{S\to+\infty} \mathbb{E}\big[k(V^a_S - V^b_S)|F_t\big]$ is well defined. This means

$$\mathbb{E}\big[(V^a_{S+h} - V^b_{S+h}) - (V^a_S - V^b_S)|F_t\big] \to 0,$$

that is the order flow imbalance between S and $S+h$ is asymptotically (in S) not predictable at time t.

Dynamics

Price dynamics

- Under the preceding assumptions, we finally get

$$P_t = P_0 + k \lim_{S \to +\infty} \mathbb{E}\big[(V_S^a - V_S^b)|F_t\big].$$

- Martingale price.
- Linear permanent impact, independent of execution mode.
- The price process only depends on the global market order flow and not on the individual executions of metaorders. We thus do not need to assume that the market sees the execution of metaorders as it is usually done.
- Market orders move the price because they change the anticipation that market makers have about the future of the order flow.

Table of contents

Preliminary : Hawkes processes

Hawkes process

- 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. First introduction in finance : Chavez-Demoulin *et al.* (2005), Bowsher (2007).

Hawkes specification

Hawkes propagator

- We now assume that buy and sell order flows are modeled by independent Hawkes processes N^a and N^b with same parameters μ and ϕ. All orders have same unit volume.
- Later on we will consider an asymptotic setting so that the flows are defined on $[0, T]$ with $T \to +\infty$.
- To be very general, we allow the parameters to depend on T (but do not assume they do). So we write $N^{a,T}$, $N^{b,T}$, μ^T, $\phi^T = a^T\phi$ with $a^T < 1$ and $\int \phi = 1$ (stability condition).
- Note that the average intensity of our processes is essentially $\beta^T = \mu^T(1 - a^T)^{-1}$ (stationary case).

Price dynamic under Hawkes specification

Price equation

- In this case, the general equation above rewrites as the following propagator dynamic

$$P_t = P_0 + \int_0^t \zeta^T(t-s)(dN_s^{a,T} - dN_s^{b,T}),$$

with $\zeta^T(t) = \big(1 + \int_t^{+\infty} \psi^T(u) - \int_0^t \psi^T(u-s)\phi^T(s)dsdu\big)$.

- The propagator kernel compensates the correlation of the order flow implied by the Hawkes dynamics to recover a martingale price. Note that the kernel does not tend to 0 since there is permanent impact.

Adding our own transactions

Labeled order

- In the above framework, $N^{a,T}$ and $N^{b,T}$ are the flows of anonymous market orders.
- Now assume we arrive on the market, executing our own (buy) metaorder. Our flow is a Poisson process n on $[0, T]$ (can be generalized) with intensity $I^T = \gamma\beta^T$, $\gamma < 1$ (proportion γ of the total flow).
- According to the propagator approach, we get

$$P_t = P_0 + \int_0^t \zeta^T(t-s)(dN_s^{a,T} - dN_s^{b,T}) + \int_0^t \zeta^T(t-s)dn_s.$$

Impact function

Explicit market impact

- We get that the impact function of a metaorder executed between 0 and T is for $0 \leq t \leq T$

$$MI(t) := \mathbb{E}[P_t - P_0] = I^T \int_0^t \zeta^T(t-s)ds.$$

- We define

$$\overline{MI}^T(t) = \frac{MI_{tT}^T}{T\beta^T} = \int_0^t \chi^T(t-s)\mathrm{d}s,$$

with

$$\chi^T(s) = \gamma \frac{\zeta^T(Ts)}{1-a^T}.$$

Decomposing the impact

Transient and permanent market impact

- We have

$$\overline{MI}^T(t) = \int_0^t \chi^T(t-s)\mathrm{d}s,$$

$$\chi^T(s) = \gamma\big(1 + (1-a^T)^{-1}\int_{Ts}^{+\infty}\phi\big).$$

- The market impact kernel is the sum of a linear market impact representing the permanent component and of a transient term vanishing after the metaorder completion.
- Existence of transient part is equivalent (asymptotically) to the existence of a limit for $(1-a^T)^{-1}\int_{Ts}^{+\infty}\phi$.

Shape of the market impact

Power-law market impact

Assume the transient part of the market impact exists. Then for $t < 1$,

$$\lim_{T \to +\infty} \overline{MI}^T(t) - \gamma t = \gamma K t^{1-\alpha}$$

for some $K > 0$ and $\alpha \in (0, 1)$. Furthermore, we necessarily have $a^T \to 1$ (highly endogenous market) and the tail of the Hawkes kernel is power-law of order $x^{-(1+\alpha)}$.

Note that the celebrated square-root law (Bouchaud et al., Farmer et al., Pohl et al.) corresponds to $\alpha = 1/2$.

Limiting price process

Emergence of (hyper-)rough processes

Let $\bar{P}_t^T = \frac{1}{T\beta^T} P_t^T$ and assume $\mu^T(1-a^T)T$ tends to δ. As T goes to infinity, the limit P_t of $\bar{P}_t^T$ satisfies

$$P_t = B_{X_t},$$

$$X_t = \frac{2}{\delta}\int_0^t F^{\alpha,\lambda}(s)\mathrm{d}s + \frac{1}{\delta\sqrt{\lambda}}\int_0^t F^{\alpha,\lambda}(t-s)\mathrm{d}W_{X_s},$$

where B and W are Brownian motions, $\lambda = K\Gamma(1-\alpha)^{-1}$ and $F^{\alpha,\lambda}(t) = \int_0^t f^{\alpha,\lambda}(s)\mathrm{d}s$ with $f^{\alpha,\lambda}$ the density of the Mittag-Leffler distribution. Furthermore, X has Hölder regularity $\min(2\alpha, 1) - \varepsilon$.

Uniqueness in law of the limit

Characterization of the limit

Let X be the cumulated volatility process of the limiting price, $f \in C^0(\mathbb{R}^+, \mathbb{R}^-)$. The function $K(f,t) = \mathbb{E}[\exp\big(\int_0^t f(s)\mathrm{d}X_{t-s}\big)]$ satisfies

$$K(f,t) = \exp\Big(\int_0^t g(s)\mathrm{d}s\Big),$$

with g the (unique) solution of the Volterra Ricatti equation

$$g(t) = \int_0^t f^{\alpha,\lambda}(t-s)\Big(\frac{\delta}{4}g(s)^2 + \frac{2}{\delta}f(s)\Big)\mathrm{d}s.$$

The case $\alpha > 1/2$

Rough Heston limit

When $\alpha > \frac{1}{2}$, the rescaled price process variance is almost surely differentiable. Furthermore

$$P_t = \int_0^t \sqrt{Y_s} \mathrm{d}B_s,$$

$$Y_t = \frac{1}{\Gamma(\alpha)}\Big(\int_0^t (t-s)^{\alpha-1}\big(\frac{2}{\delta} - \lambda Y_s\big)\mathrm{d}s + \int_0^t (t-s)^{\alpha-1}\sqrt{Y_s}\mathrm{d}W_s\Big).$$

Therefore we have a rough Heston model with $H = \alpha - 1/2$.

Summary

From no-arbitrage to volatility

- We made two assumptions : Linear permanent impact and martingale price.
- Only modeling assumption : Hawkes dynamics for the order flow (reasonable...).
- This leads to rough volatility. In the square-root law case, $H \approx 0$.
- Going further : with the same type of arguments, using quadratic Hawkes processes, we are able to introduce a new class of log-normal type rough volatility processes with Zumbach effect, with many financial applications : the **Quadratic Rough Heston** models.