Rough volatility

Mathieu Rosenbaum

École Polytechnique

August 2022

Table of contents

1. Introduction
2. Microstructural foundations for rough volatility
3. An important application : The rough Heston formula
4. Rough volatility and Zumbach effect
5. Quadratic rough Heston model and the VIX market
6. On the universality of the volatility formation process

Table of contents

1 Introduction

2 Microstructural foundations for rough volatility

3 An important application : The rough Heston formula

4 Rough volatility and Zumbach effect

5 Quadratic rough Heston model and the VIX market

6 On the universality of the volatility formation process

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

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 $\mathbb{E}[(W^H_1)^2] = 1$.

Proposition

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

Mandelbrot-van Ness representation

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

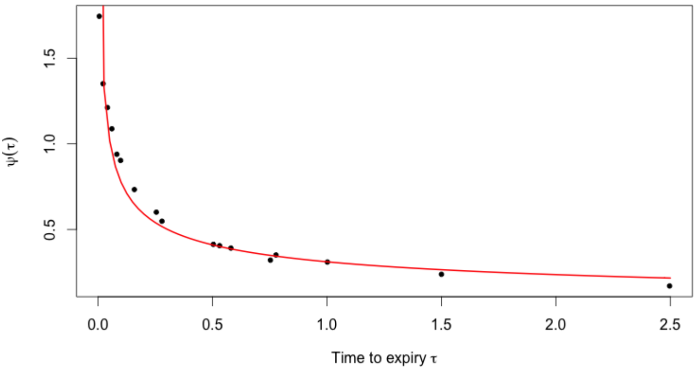
About option data

- 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 ; 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.

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).

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

The log-volatility

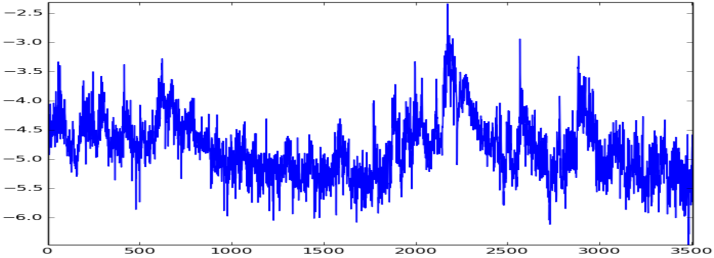


FIGURE – The log volatility $\log(\sigma_t)$ as a function of t, S&P.

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

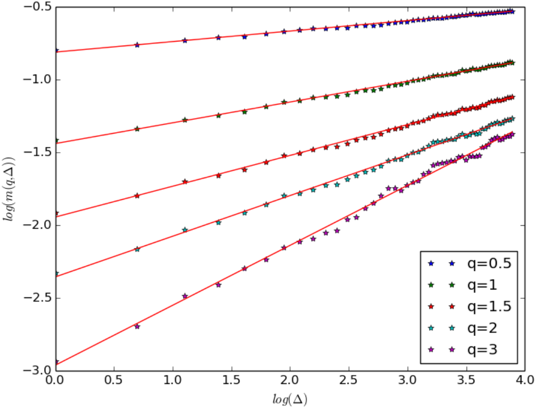


FIGURE – $\log(m(q, \Delta)) = \zeta_q \log(\Delta) + C_q$. The scaling is not only valid as Δ tends to zero, but holds on a wide range of time scales.

Monofractality of the log-volatility

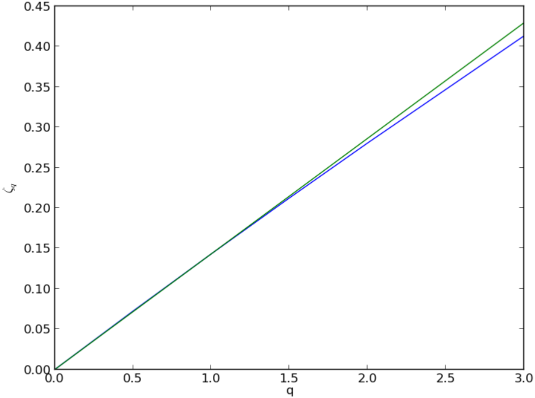


FIGURE – Empirical ζ_q and $q \to Hq$ with $H = 0.14$ (similar to a fBm with Hurst parameter H).

Distribution of the log-volatility increments

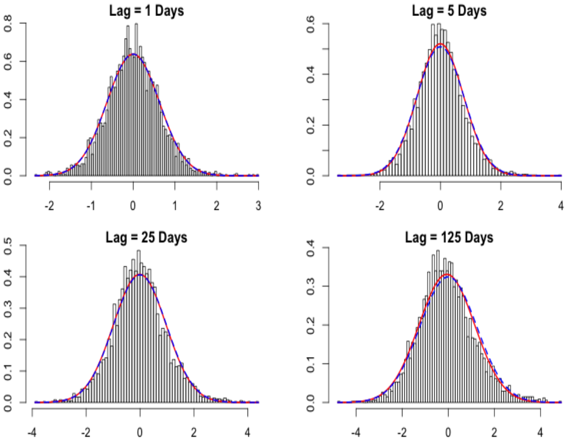


FIGURE – The distribution of the log-volatility increments is close to Gaussian.

A geometric fBm model

The RFSV model

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

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

A particularly intriguing property : volatility multiscaling

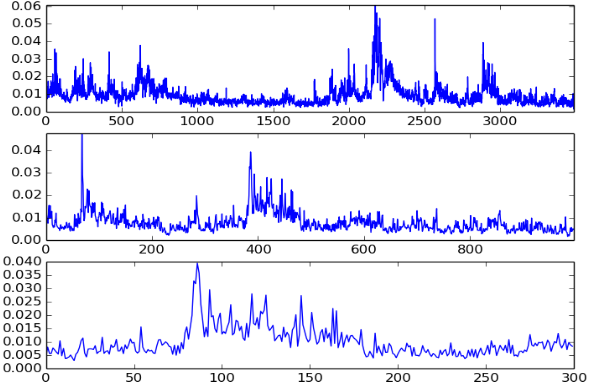


FIGURE – Empirical volatility over 10, 3 and 1 years.

Rough volatility on different time intervals

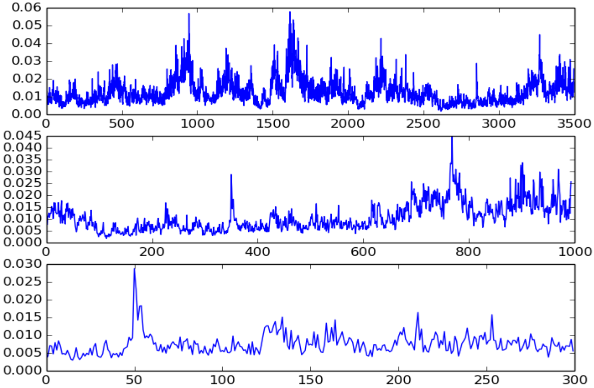


FIGURE – Simulated volatility over 10, 3 and 1 years. We observe the same alternations of periods of high market activity with periods of low market activity.

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 ($H=0.1$) implies a multiscaling behavior of the volatility.

On the universality of rough volatility

Towards universality

- **Rough volatility appears to be a universal phenomenon** : Similar values for H on more than 10.000 assets.
- We want to understand why : Microstructural foundations for rough volatility. (*Hawkes processes*)
- Can we also deal with complex derivatives such as VIX options in the rough volatility framework ? (*Zumbach effect and Quadratic rough Heston*)
- We obtain very accurate volatility forecasts thanks to rough volatility models. Can we deduce from this paradigm some fundamental universal patterns in the endogenous part of the volatility formation process ? (*Deep learning*)

Literature

Rough volatility network

- https ://sites.google.com/site/roughvol/
- Forthcoming book

Table of contents

1. Introduction
2. Microstructural foundations for rough volatility
3. An important application : The rough Heston formula
4. Rough volatility and Zumbach effect
5. Quadratic rough Heston model and the VIX market
6. On the universality of the volatility formation process

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 dynamic of volatility from the high frequency behavior of market participants.

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.
- 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.
- 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.

Encoding the stylized facts

The right parametrization of the model

- Recall that

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

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

Rough Heston 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 Hurst parameter H satisfies $H = \alpha - 1/2$.

An even deeper result

No-arbitrage implies rough volatility and power law market impact

- We have shown that combining typical behaviours of market participants at the high frequency scale automatically generates rough volatility.
- We can actually prove that only assuming no-statistical arbitrage implies rough volatility.
- The key phenomenon to obtain this result is the **market impact**.
- In a perfect market from a statistical arbitrage viewpoint, $H = 0$.
- There is a one to one connection between the value of H and the shape of the market impact curve.
- $H = 0$ corresponds to square-root market impact.

Table of contents

1 Introduction

2 Microstructural foundations for rough volatility

3 An important application : The rough Heston formula

4 Rough volatility and Zumbach effect

5 Quadratic rough Heston model and the VIX market

6 On the universality of the volatility formation process

Deriving the characteristic function of the rough Heston model

Strategy

- From our last theorem, we are able to derive the characteristic function of our high frequency Hawkes-based price model.
- We then pass to the limit.

Characteristic function of rough Heston models

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

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

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

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

Comments on the theorem

The rough Heston formula

- The formula is the very same as the celebrated Heston formula, up to the replacement of a classical time derivative by a fractional derivative.
- This formula allows for fast derivatives pricing and risk management.
- Thanks to this approach, we can derive the infinite dimensional Markovian structure underlying rough Heston models, leading to explicit hedging formulas.

Table of contents

1. Introduction
2. Microstructural foundations for rough volatility
3. An important application : The rough Heston formula
4. Rough volatility and Zumbach effect
5. Quadratic rough Heston model and the VIX market
6. On the universality of the volatility formation process

Further aspects of volatility

Super-Heston rough volatility and Zumbach effect

- All the works on microstructural foundations of rough volatility have produced rough Heston type models.
- In the context of rough models, there are other aspects of volatility that one could wish to understand from a microstructural perspective.
- Going beyond the square root associated to the dynamic of the volatility in the rough Heston model$\rightarrow$ additional additive or multiplicative factor leading to fatter volatility tails : *Super-Heston rough volatility*.
- *Zumbach effect*.

Further aspects of volatility

Zumbach effect (Zumbach *et al.*) : description

- Feedback of price returns on volatility.
- Price trends induce an increase of volatility.
- In the literature (notably works by J.P. Bouchaud and co-authors), a way to reinterpret the Zumbach effect is to consider that the predictive power of past squared returns on future volatility is stronger than that of past volatility on future squared returns.
- To check this on data, one typically shows that the covariance between past squared price returns and future realized volatility (over a given duration) is larger than that between past realized volatility and future squared price returns.
- We refer to this version of Zumbach effect as *weak Zumbach effect*.

Further aspects of volatility

Weak and strong Zumbach effect

- It is shown in Gatheral *et al.* that the rough Heston model reproduces the weak form of Zumbach effect.
- However, it is not obtained through feedback effect, which is the motivating phenomenon in the original paper by Zumbach. It is only due to the dependence between price and volatility induced by the correlation of the Brownian motions driving their dynamics.
- In particular in the rough Heston model, the conditional law of the volatility depends on the past dynamic of the price only through the past volatility.
- We speak about *strong Zumbach effect* when the conditional law of future volatility depends not only on past volatility trajectory but also on past returns.

A convenient microscopic model encoding Zumbach effect

Quadratic Hawkes processes

- Inspired by Blanc *et al.*, we model high frequency prices using quadratic Hawkes processes.
- Jump sizes of the price P_t are i.i.d taking values -1 and 1 with probability $1/2$ and jump times are those of a point process N_t with intensity

$$\lambda_t = \mu + \int_0^t \phi(t-s)\mathrm{d}N_s + Z_t^2, \text{ with } Z_t = \int_0^t k(t-s)\mathrm{d}P_s.$$

- The component Z_t is a moving average of past returns.
- If the price has been trending in the past, Z_t is large leading to high intensity. On the contrary if it has been oscillating, Z_t is close to zero and there is no feedback from the returns on the volatility. So Z_t is a (strong) Zumbach term.

One particular scaling limit

Quadratic rough Heston model

$$\mathrm{d}S_t = S_t\sqrt{V_t}dW_t, \;\; V_t = a(Z_t - b)^2 + c,$$

where a, b and c some positive constants and Z_t follows

$$Z_t = \int_0^t f^{\alpha,\lambda}(t-s)\theta_0(s)\mathrm{d}s + \int_0^t f^{\alpha,\lambda}(t-s)\sqrt{V_s}\mathrm{d}W_s,$$

with $\alpha \in (1/2, 1)$, $\lambda > 0$ and θ_0 a deterministic function.

- Z_t is *path-dependent* : a weighted average of past returns.
- c : minimal instantaneous variance.
- $b > 0$: asymmetry of the feedback effect.
- a : sensitivity of the volatility feedback.
- **A log-normal rough volatility model with strong Zumbach effect.**

Table of contents

1 Introduction

2 Microstructural foundations for rough volatility

3 An important application : The rough Heston formula

4 Rough volatility and Zumbach effect

5 Quadratic rough Heston model and the VIX market

6 On the universality of the volatility formation process

The Volatility Index

Definition of the VIX

- Introduced in 1993 by the CBOE.
- VIX is the square root of the price of a specific basket of options on the S&P 500 Index (SPX) with maturity $\Delta = 30$ days such that

$$\mathrm{VIX}_t = -\frac{2}{\Delta}\sqrt{\mathbb{E}[\log(S_{t+\Delta}/S_t)|\mathcal{F}_t]} \times 100,$$

  with S the SPX index.
- VIX futures and VIX options exist.

The joint calibration problem

VIX options

- More than 500,000 VIX options traded each day.
- Quite wide spreads for VIX options : non-mature market.
- VIX is by definition a derivative of the SPX, any reasonable methodology must necessarily be consistent with the pricing of SPX options.
- Designing a model that jointly calibrates SPX and VIX options prices is known to be extremely challenging.
- This problem is sometimes considered to be *the holy grail of volatility modeling*.
- We simply refer to it as the *joint calibration problem*.

The joint calibration problem

Attempts to solve the joint calibration problem

- Theoretical approch by J. Guyon : the joint calibration problem is interpreted as a model-free constrained martingale transport problem. Perfect calibration of VIX options smile at time T_1 and SPX options smiles at T_1 and $T_2 = T_1 + 30$ days. Hard to be extended to any set of maturities and high computational cost.
- Models with jumps : most of them fail to reproduce VIX smiles for maturities shorter than one month.
- Continuous models : Unsuccessful so far. Interpretation : the very large negative skew of short-term SPX options, which in continuous models implies a very large volatility of volatility, seems inconsistent with the comparatively low levels of VIX implied volatilities

The VIX conjecture

The joint calibration problem and continuous models

- “So far all the attempts at solving the joint SPX/VIX smile calibration problem [using a continuous time model] only produced imperfect, approximate fits”.
- “Joint calibration seems out of the reach of continuous-time models with continuous SPX paths”.
- Investigating Guyon’s work one can realise the following : a necessary condition for a continuous model to fit simultaneously SPX and VIX smiles is the inversion of convex ordering between volatility and the local volatility implied by option prices.
- The intuition behind this condition could be reinterpreted as some kind of strong Zumbach effect.
- Natural for us to investigate the ability of super-Heston rough volatility models to solve the joint calibration problem.

Calibration for one day in history 19 May 2017

Parameters calibration with Deep Learning

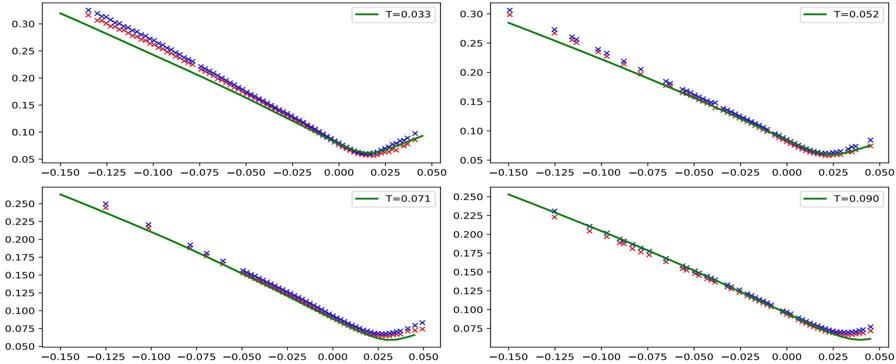


FIGURE – Implied volatility on SPX options for 19 May 2017. Blue and red points are bid and ask of market implied volatilities. Model implied volatility smiles from the model are in green. Strikes are in log-moneyness, maturity in year.

Calibration for one day in history 19 May 2017

Parameters calibration with Deep Learning

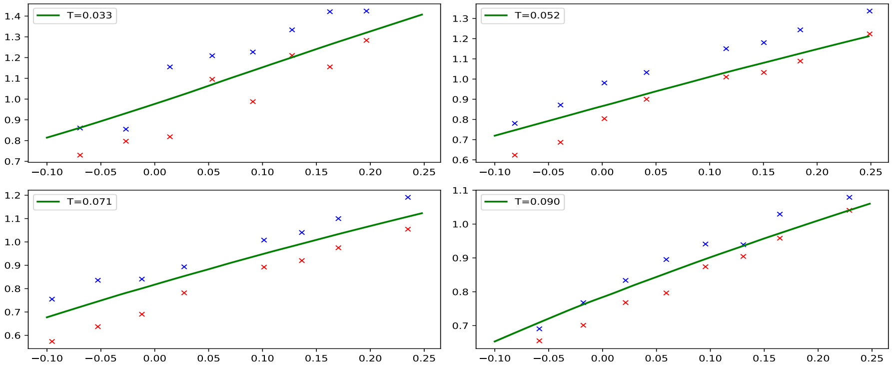


FIGURE – Implied volatility on VIX options for 19 May 2017. Blue and red points are bid and ask of market implied volatilities. Model implied volatility smiles from the model are in green. Strikes are in log-moneyness, maturity in year.

Take home message for the joint calibration problem

Thanks to the quadratic rough Heston model

- 6 parameters.
- VIX smiles in the bid-ask spread.
- Global shape of the implied volatility surface of the SPX very well reproduced
- Very accurate SPX skews of orders -1.5 (shortest maturites), -1 (longer maturities), as for market data.

Table of contents

1. Introduction
2. Microstructural foundations for rough volatility
3. An important application : The rough Heston formula
4. Rough volatility and Zumbach effect
5. Quadratic rough Heston model and the VIX market
6. On the universality of the volatility formation process

Forecasting devices for next day realized volatility

Parametric methods

- AR(p) :

$$\hat{\sigma}_t = \alpha_0 + \sum_{j=1}^{p} \beta_j \sigma_{t-j} \,,$$

- HAR :

$$\hat{\sigma}_t = \alpha_0 + \beta_1 \sigma_{t-1} + \beta_2 \frac{1}{5} \sum_{j=1}^{5} \sigma_{t-j} + \beta_3 \frac{1}{22} \sum_{j=1}^{22} \sigma_{t-j} \,,$$

- RFSV ($d \log \sigma_t = \nu dW_t^H$) :

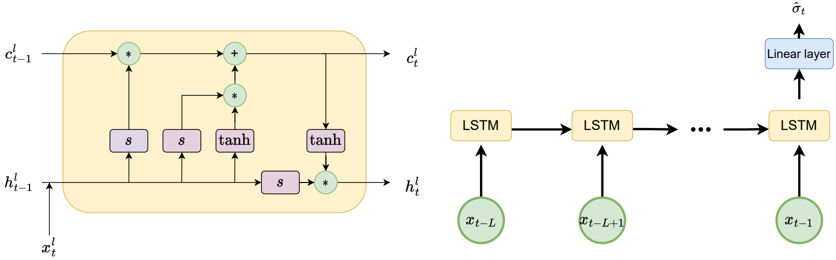
$$\widehat{\log \sigma_t} = \frac{\cos(H\pi)}{\pi} \int_{-\infty}^{t-1} \frac{\log \sigma_s}{(t-s+1)(t-s)^{H+1/2}} ds \,,$$

$$\hat{\sigma}_t = c \exp(\widehat{\log \sigma_t}).$$

Forecasting devices for daily realized volatility

A universal non-parametric method

- LSTM recurrent neural network, with similar weights for each asset, trained on a pooled dataset.
- Inputs are $x_t = (\sigma_t^2)$ or $x_t = (\sigma_t^2, r_t)$, where r_t is the daily return at time t, with variable length for history.



Dataset

Description

- 5-minutes intraday prices of Russell 1000 and STOXX Europe 600 constituents, for years between 2010 and 2020.
- 862 names from the US market and 503 names from the European market.
- Scaling for each stock :

$$\sigma_t = \frac{\sigma_t}{\sqrt{\langle \sigma_t^2 \rangle}}, \qquad r_t = \frac{r_t - \langle r_t \rangle}{\sqrt{\langle (r_t - \langle r_t \rangle)^2 \rangle}}.$$

- We focus mostly on the US market. The data of the European market is used for an out-of-sample double-check.
- We use the pooled dataset of 862 stocks over years 2010 - 2015 to train the LSTM network. The period 2016 - 2020 is used for out-of-sample evaluation.

Evaluation metric

Relative mean square error

- $$\text{MSE}(\sigma, \hat{\sigma}) = \frac{1}{T} \sum_{t=1}^{T} (\hat{\sigma}_t - \sigma_t)^2,$$

  where T is the number of trading days of the out-of-sample period.
- We focus on each model's relative performance compared to that of the HAR model so that we compute instead $(\text{MSE}_\text{m}/\text{MSE}_\text{HAR})$, for $m \in \{\text{AR}(22), \text{RFSV}, \text{LSTM}\}$.

Capturing universality with LSTM

Hurst parameter

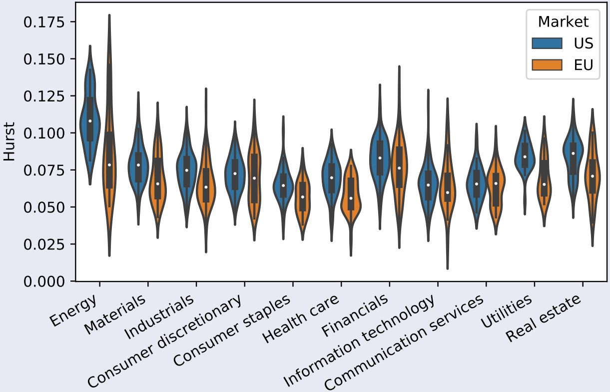


FIGURE – Empirical distribution of the estimated Hurst parameters inside each sector.

Capturing universality with LSTM

Parametric vs non-parametric

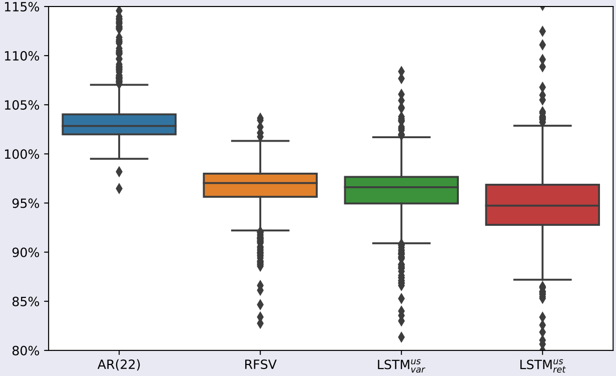


FIGURE – Boxplot showing each model's out-of-sample MSE relative to the HAR model for the stocks of the US market.

Capturing universality with LSTM

Parametric vs non-parametric

- AR(22) underperforms the HAR model (overfitting).
- RFSV outperforms the HAR model. It is remarkable as it involves essentially no parameters ($H = 0.055$, $c = 1.03$).
- LSTM^{us}_{var} and LSTM^{us}_{ret} outperform the other parametric models, especially when we incorporate past returns data. This indicates that the potential universal volatility formation mechanism across assets, relating past volatilities and returns to current volatilities, allows us to calibrate a universal model based on all assets, where the risk of overfitting is reduced due to enriched realized scenarios.
- We check for potential sector/stock (transfer learning)/market specific or time dependent component in the volatility formation process but consistently found that our universal network could not be significantly improved.

Uncovering the universal volatility formation process

A quadratic rough Heston inspired forecast

- Following the idea on Zumbach effect in the QRH model, we propose the following forecasting device :

$$\hat{\sigma}_t^2 = a(Z_{t-1} - b)^2 + c$$

with $Z_t = \int_{-\infty}^{t} \frac{(t-s)^{H-\frac{1}{2}}}{\Gamma(H+\frac{1}{2})} \sigma_s \mathrm{d}W_s$.

- We finally consider the following forecast

$$(1-\lambda)\hat{\sigma}^{RFSV} + \lambda\hat{\sigma}^{QRH},$$

with $H = 0.055, c = 1.03, a = 0.043, b = 0.74, c = 0.55$.

Uncovering the universal volatility formation process

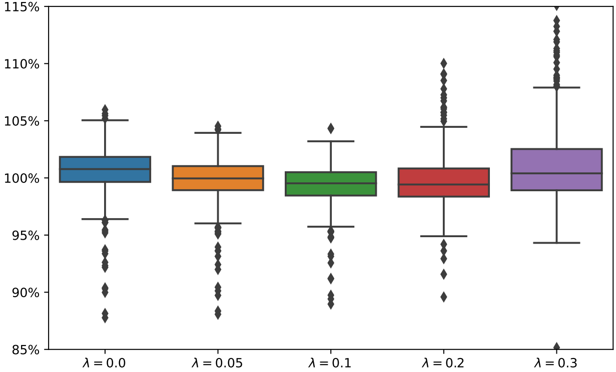


FIGURE – Out-of-sample performance of the forecast $(1-\lambda)\hat{\sigma}^{RFSV} + \lambda\hat{\sigma}^{QRH}$ relative to LSTM^{eu}_{ret} in the EU market.

Conclusion

Universality of the volatility formation process

- The universal LSTM network, trained on a pooled dataset of hundreds of stocks, outperforms consistently the asset-specific parametric models based on past volatilities.
- Similar superior performances hold on assets that are not part of the training set, even on those of a different market. Fine-tuning the universal model with the data of each stock does not help improve the performance.
- These observations suggest the existence of a universal volatility formation mechanism from a nonparametric perspective.
- A simple combination of the RFSV and QRH forecasts with fixed parameters perform similarly to our LSTM network.
- From a parametric perspective, this shows that the main features of this universal volatility formation process can be well described by the rough volatility paradigm boosted with Zumbach effect.