Rough Volatility: Fact or Artefact?

Quantitative Finance Seminars

Andrea Macrì

July 8, 2022

Scuola Normale Superiore

Introduction

Table of contents

$\rightarrow$ Investigating ”significancy“ of $H < 0.5$ when modelling stochastic volatility;

$\implies$ Non parametric method $\rightarrow$ Normalized p-th variation ;

$\rightarrow$ Investigating on ’artificial‘ paths’ rough behaviour;

$\implies$ Realized volatility vs. Actual Dynamics $\implies$ micro-structure noise?

$\rightarrow$ Investigating on ’real world‘ paths’ rough behaviour;

$\implies$ Application to S&P500;

$\implies$ Hausman test and application to MSFT;

$\rightarrow$ Conclusions.

$\implies$ Remarks;

$\implies$ Further directions.

When it comes of modelling assets' price dynamics, usually continuous semi-martingales are used:

$$dS_t = \mu^{\mathbb{Q}} S_t dt + \sigma S_t dW_t^{\mathbb{Q}} \tag{1}$$

$\mu \rightarrow$ "Usually" poses no threats;

$\sigma \rightarrow$ Needs to be modelled $\implies$ consistency with real world.

Rough rocesses

Starting from Comte and Renault 1998 and more recently Gatheral, Jaisson, and Rosenbaum 2018, usage of Fractional Brownian Motion (fBM) in volatility modelling has recently taken hold.

Definition 0. The fBM is BM on [0,T] whose increments need not to be independent. Unlike the BM its covariance function is[1]

$$\mathbb{E}[W_t^H W_s^H] = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H}) \;\; H \in [0,1] \;\; \forall s,t \in [0,T] \;, \;\; s \neq t \quad (2)$$

The H parameter controls for the raggedness of the resultant motion driven by such a process. The smaller the rougher, the bigger the smoother.

Definition 1.

$$W^H = \int_0^t \frac{dW_t}{(t-s)^\gamma} + \int_{-\infty}^0 \left[\frac{1}{(t-s)^\gamma} - \frac{1}{s^\gamma}\right] dW_s \;\; , \;\; \frac{1}{2} - H = \gamma \qquad (3)$$

[1] Representations by Mandelbrot and Ness 1968

Rough processes

Properties:

- self-similarity of the increments $W^H(at) \sim |a|^H W^H(t)$
- stationary increments: $W^H_t - W^H_s = W^H_{t-s}$;
- long range dependence: slow decay $\sim T^{2H-2}$ of auto-correlation of its increments, where H is the Hurst exponent;
- paths regularity: Ability to generate trajectories with varying levels of Hölder regularity (“ Roughness”).

Rough Volatility - Comte and Renault 1998 and Gatheral, Jaisson, and Rosenbaum 2018

An early contribution, that tries to model the 'volatility-clustering' phenomenon is Comte and Renault 1998, where the instantaneous volatility is modelled by:

$$\begin{cases} dS_t = \mu S_t dt + \sigma_t S_t dW_t \\ y_t = -\gamma y_y dt + \theta dW_t^H \quad ; \quad y_t = \ln(\sigma_t) \end{cases} \tag{4}$$

With W_t^H an FBM with $H \in (\frac{1}{2}, 1]$.

Gatheral, Jaisson, and Rosenbaum 2018 use a similar to (4) model but, relying on high-frequency data, infer fractal $\log(\sigma^2) \to m(q, \Delta) \sim \Delta^{\xi_q}$ that gives a regularity coefficient $\xi_q = q \times H$ for $H \in [0, \frac{1}{2})$.[2]

[2] Where $m(q, \Delta) = \frac{1}{n}[\log(RV)]^q_{\pi^n}$ is the sample moment of $\log(\sigma^2)$ differences.

Issues

Thus, specular effects to take into account:

> ‘Push’- effect: $H > 0.5 \impliedby$ long-range dependence in volatility$\leftarrow$ volatility *clustering*;
>
> ‘Pull’- effect: $H < 0.5 \impliedby$ very short intraday estimates of volatility $\leftarrow$ micro-foundational arguments.

$\implies$ However, as Gatheral, Jaisson, and Rosenbaum 2018 showed, volatility clustering is compatible with $H << 0.5$, the model with such H is able to identify spurious long memory[3].

Debate is still open on this front.

[3] not in the classical power law sense

Measuring the roughness of a path: the W estimator

Main contributions

The contribution to rough volatility literature is threefold in this work:

- $\rightarrow$ New non-parametric way to estimate the Hurst exponent of a path via *normalized* p-variation: $W(L, K, \pi, p, t, X)$;
- $\rightarrow$ Use of the $W(L, K, \pi, p, t, X)$ to show inconsistency between roughness in Realized Volatility (RV) and real volatility data;
- $\rightarrow$ Estimation of roughness from real data: $S\&P500$ and MSFT.

'Rough' discrepancy between σ_t and RV $\implies$ Cont and Das 2022b notice apparent roughness!

P-variation

Consider an adapted stochastic process $X(t,\omega) : [0,T] \times \Omega \to \mathbb{R}$ on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\in[0,T]})$. Fix a path $X(\cdot,\omega) = x$, then:

Definition 1. $x \in C^0([0,T],\mathbb{R})$ has finite p-variation along a sequence of partitions $\{\pi^n | n > 1\}$, if $\exists\ [x]^p_\pi :\ [0,T] \to \mathbb{R}_+$ such that[4]:

$$\forall t \in [0,T] \sum_{[t^n_j, t^n_{j+1}] \in \pi^n} |x(t^n_{j+1}) - x(t^n_j)|^p \xrightarrow[n\to\infty]{u.c.p.} [x]^p_\pi(t)$$

We denote $V^p_\pi([0,T],\mathbb{R})$ the set of all continuous paths with finite p-variation along π.

[4] $[x]^p_\pi$ is continuous and increasing function.

P-variation

Definition 2. The variation index of a path x along π^n is the *smallest* $p \geq 1$ for which x has finite p-variation.

$$p^{\pi}(x) = \inf\{p \geq 1 \ : \ x \in V_{\pi}^{p}([0,T],\mathbb{R})\}$$

Definition 3. The roughness index of a path x along π^n is defined as:

$$H^{\pi}(x) = \frac{1}{p^{\pi}(x)}$$

Notice: just an ω considered $\implies$ $H^{\pi}(x)$ may change path by path. For some classes of processes $\exists!$ *a.s.* one roughness index. E.g. BM has $P^{\pi}(B) = 2$ on any refining partition π (Dudley 2010).

Normalised p-variation

Definition 4. Consider a sequence of partitions π on [0,T] with vanishing mesh. $x \in V_\pi^p([0,T],\mathbb{R})$ has *npv* on π if exists $w(x,p,\pi) : [0,T] \to \mathbb{R}$ such that:

$$\forall t \in [0,T] \quad \sum_{\pi^n \cap [0,T]} \frac{|x(t_{i+1}^n) - x(t_i^n)|^p}{[x]_\pi^p(t_{i+1}^n) - [x]_\pi^p(t_i^n)} \times (t_{i+1}^n - t_i^n) \xrightarrow[n\to\infty]{} w(x,p,\pi)(t)$$

The class of continuous processes for which $w(\cdot,p,\pi)$ exists is called $N_\pi^p([0,T],\mathbb{R})$.

Notice: if $p = 2$ then $w(x,p,\pi)$ is the 'normalized quadratic variation'.

Normalised p-variation

Theorem

Let $x \in V^p_\pi([0,T],\mathbb{R})$, for some $p > 1$ where π be a sequence of partitions with vanishing mesh, if the p-variation exists and is continuous then $x \in N^p_\pi([0,T],\mathbb{R})$ and $\forall t \in [0,T]$, $w(x,p,\pi) = t$.

proof:(13)

Example

Let B be a Wiener process, on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\in[0,T]})$ and $(\pi^n)_{n\geq 1}$ is a sequence of partitions with vanishing mesh on $[0,t]$, then $\mathbb{P}(w(B,2,\pi)(t) = t) = 1$

proof:(14)

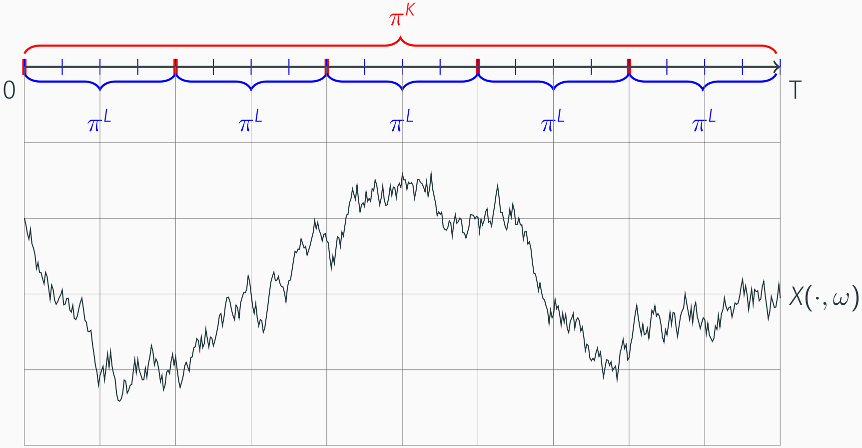
Example

Let same assumptions hold and consider B^H be an fBM with Hurst exponent H, then for a partition $\mathbb{T}$ we have: $\mathbb{P}(w(B^H, \frac{1}{H}, \mathbb{T})(t) = t) = 1$, with $p = \frac{1}{H}$.

Normalized p-variation estimator

Dealing with data $\Longrightarrow$ estimator for the normalized p-variation.

Consider a refining partition of the interval $[0, T]$ made by two frequencies: π^K and π^L, with π^K«π^L, such that:



Normalized p-variation estimator

π^K block frequency $\rightarrow$ as the partition is refining π^L is a sub-partition of π^K. Thus, the *discrete normalized p-variation* is:

$$W(L, K, \pi, p, t, X) = \sum_{\pi^K \cap [0,T]} \frac{\sum_{\pi^L \cap [t_j^K, t_{j+1}^K]} |X(t_{j+1}^L) - X(t_j^L)|^p}{|X(t_{j+1}^K) - X(t_j^K)|^p} \times (t_{j+1}^K - t_j^K) \quad (5)$$

The statistic (5) converges in probability to the normalized p-variation $w(x, p, \pi)(t)$ as $K, L \rightarrow \infty$, i.e.

$$W(L, K, \pi, p, t, X) \xrightarrow[K,L\rightarrow\infty]{P} w(x, p, \pi)(t)$$

$\Longrightarrow$ easy to deal with roughness in terms of (5)...

Normalized p-variation estimator

...in fact, the *variation index estimator* $\hat{p}_{L,K}(X)$ associated with the sampled process can be retrieved by solving:

$$W(L, K, \pi, \hat{p}_{L,K}(X), T, X) = T \tag{6}$$

$\implies$ we are able to find an estimator for the roughness index, defined as:

$$\hat{H}^{\pi}_{L,K}(X) = \frac{1}{\hat{p}_{L,K}(X)} \tag{7}$$

The asymptotic properties of these estimators under high frequency data are not available[5].

[5] They cite a work not yet available Cont and Das 2022a

Fact or Artefact? Numerical Simulations

Hurst Estimation

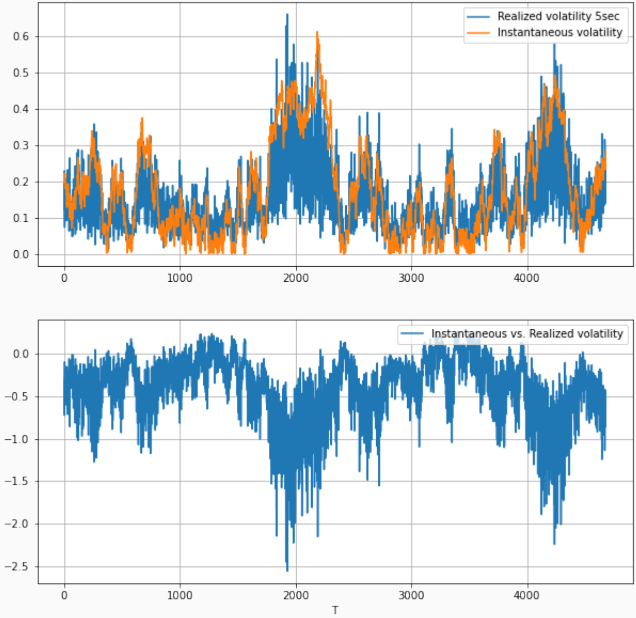
Being now equipped with the *discrete normalized p-variation* estimator, the Hurst exponent can thus be retrieved in an artificial environment.

Example (1)
Trivial (unreal) model:

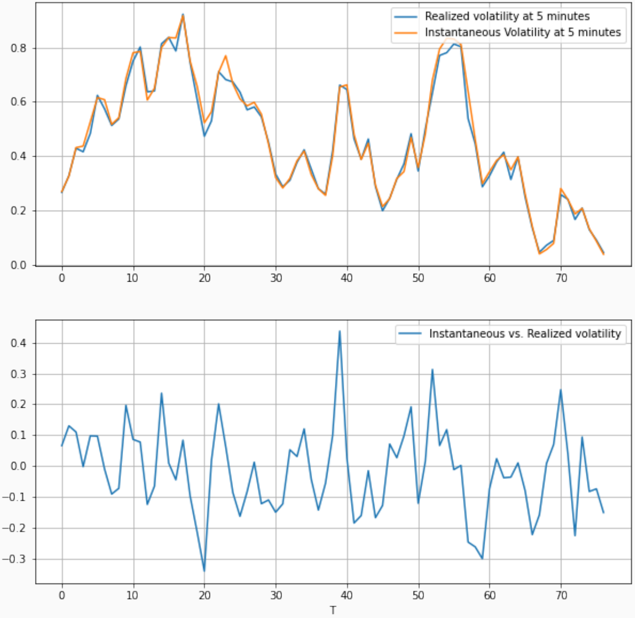
$$\begin{cases} dS_t = \sigma_t S_t dB_t \\ \sigma_t = |W_t| \\ B \perp\!\!\!\perp W \end{cases} \tag{8}$$

Simulation is carried on an interval $T = [0, 1]$, with $\Delta_t = \frac{1}{23400}$ observations.

Simulation experiments - Modulus of a Brownian Motion



Simulation experiments - Modulus of a Brownian Motion



Simulation experiments - Modulus of a Brownian Motion

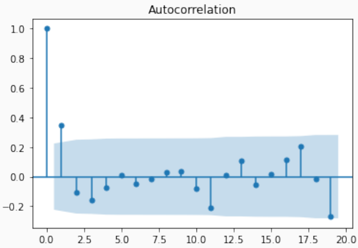


Figure 1: ACF for 5 minutes differences

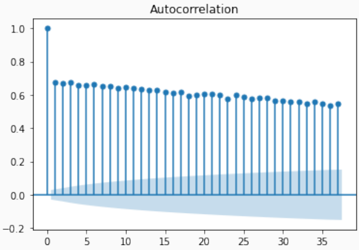
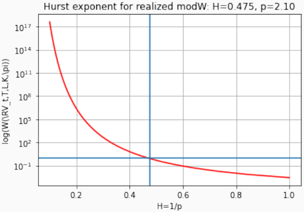


Figure 2: ACF for 5 seconds differences

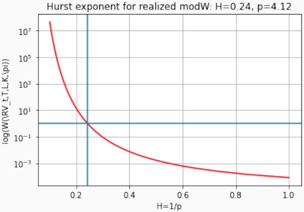
The $\sigma - RV_t$ is micro-structure noise, the 5 seconds sampling shows a non IID white noise behaviour, crucial for estimating H in many works such as in Fukasawa, Takabatake, and Westphal 2019. i.e.:

$$\sqrt{\Delta_n}(\ln(RV)^2 - \ln \int_{t_{j-1}}^{t_j} \sigma_s^2 dt) \to IID \sim \mathcal{N}(0,1)\sqrt{2/n}$$

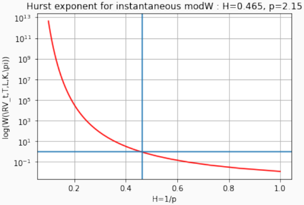
Simulation experiments - Modulus of a Brownian Motion



$\hat{H}_{L,K}$ for RV every 5 seconds,
$K = 500L = 500 \times 500$
(without noise)



$\hat{H}_{L,K}$ for RV every 5 seconds,
$K = 250L = 250 \times 250$
(with noise)



$\hat{H}_{L,K}$ for σ_t
$K = 250L = 250 \times 250$

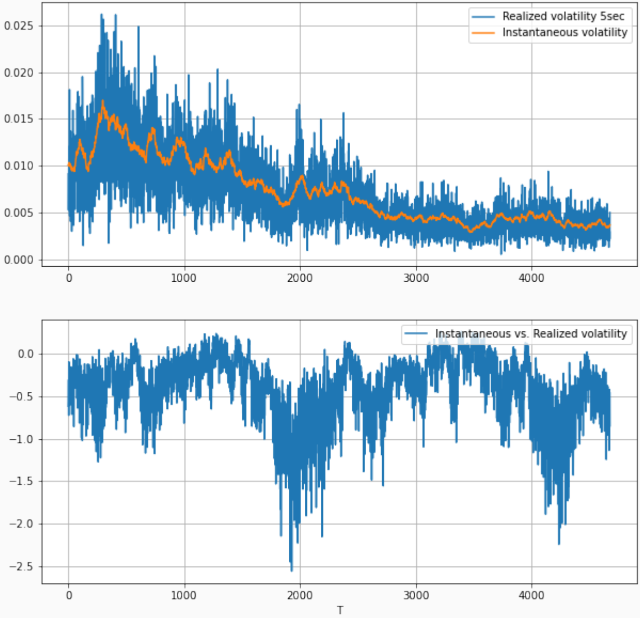
The estimated $\hat{H}_{L,K}$ via $W(\cdot)$ on $\ln(RV_t)$. Convergence is declared for $||W(\cdot) - t||_2^2 < 1e^{-3}$.

Simulation experiments - Ornstein-Uhlenbeck SV

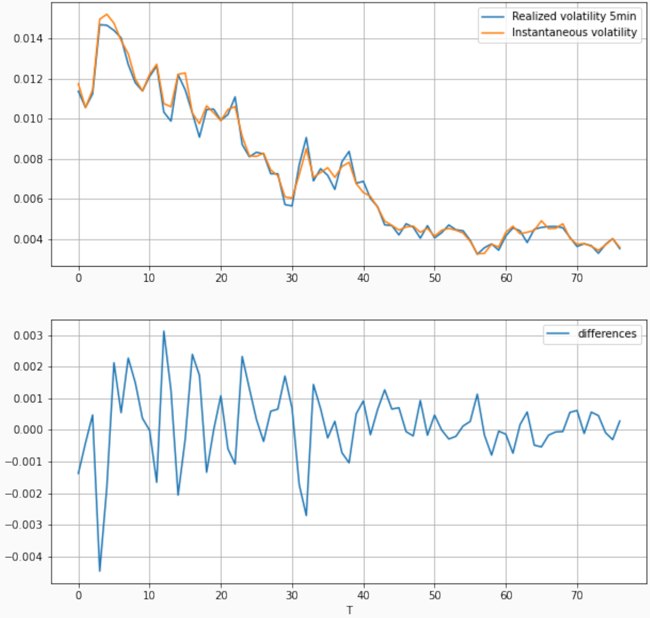
Example (2)
Coeteris paribus, a more realistic model:

$$\begin{cases} dS_t = \sigma_t S_t dB_t \\ \sigma_t = \sigma_0 \exp(Y_t) \\ dY_t = -\gamma Y_t dt + \theta dQ_t \\ \gamma = \theta = 1 \end{cases} \tag{9}$$

Simulation experiments - Ornstein-Uhlenbeck SV



Simulation experiments - Ornstein-Uhlenbeck SV



Simulation experiments - Ornstein-Uhlenbeck SV

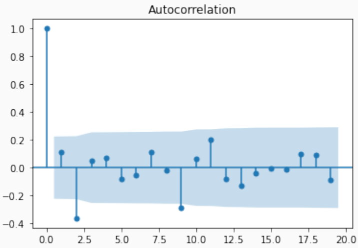


Figure 3: ACF for 5 minutes differences

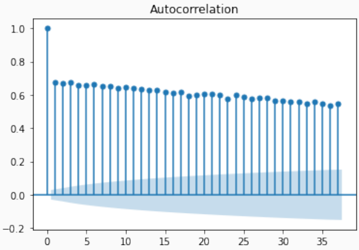
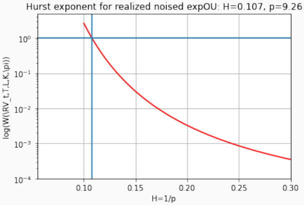


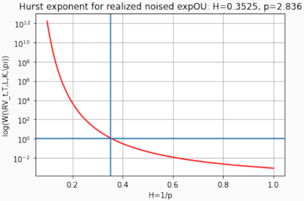
Figure 4: ACF for 5 seconds differences

The $\sigma - RV_t$ is micro-structure noise, the 5 seconds sampling shows a non IID white noise behaviour, crucial for estimating H in many works such as Fukasawa, Takabatake, and Westphal 2019. It can be concluded that the presence of noise jeopardises the framework based on IID noise.

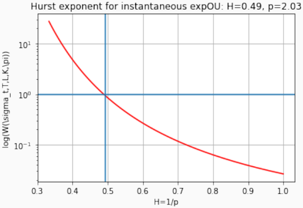
Simulation experiments - Ornstein-Uhlenbeck SV



$\hat{H}_{L,K}$ for RV every 5 seconds,
$K = 250L = 250 \times 250$
(with noise)



Average $\hat{H}_{L,K}$ for RV every 5 seconds,
$K = 500L = 500 \times 500$
(without noise)



Average $\hat{H}_{L,K}$ for σ_t
$K = 250L = 250 \times 250$

The estimated $\hat{H}_{L,K}$ via $W(\cdot)$ on $\ln(RV_t)$. Convergence is declared for $||W(\cdot) - t||_2^2 < 1e^{-3}$.

A fractional OU Model

Example (3)

Let us turn to the fractional processes instead:

$$\begin{cases} dS_t = \sigma_t S_t dB_t \\ \sigma_t = \sigma_0 \exp(Y_t) \\ dY_t = -\gamma Y_t dt + \theta dQ_t^H \\ B \perp\!\!\!\perp Q^H \end{cases} \tag{10}$$

With $H = (0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8)$, how does RV behaves when we know the underlying process?

The smoother nature of rough volatility

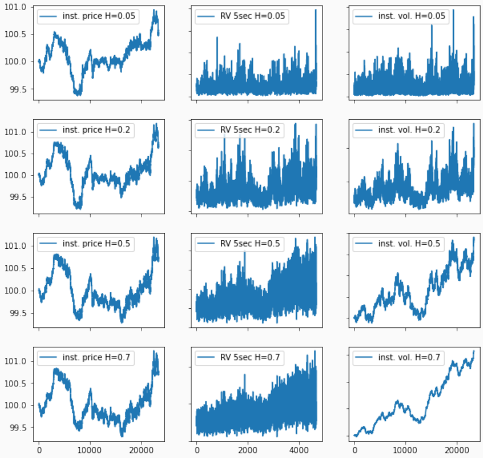


Figure 5: Bigger the H rougher RV, smoother sigma

Estimations for rough processes

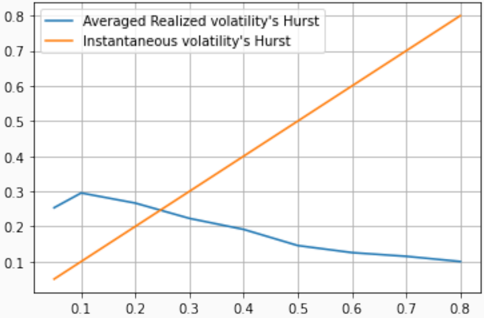


Figure 6: Plot of estimated Hurst parameter for Fractional driven models: Realised Volatility is not a good benchmark to estimate the Hurst parameter as it is always mis-estimated.

Roughening Around: S&P500 and Microsoft stocks

Following Cont and Das 2022b, a slightly similar analysis has been carried on the S&P500 index. Data have been collected from the Oxford-Man institute's data-set on realised volatility.

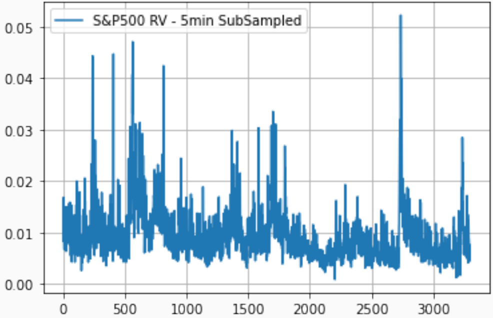


Figure 7: Oxford-Man realised volatility at 5 minutes sub sampled. On the x axis it is the number of observations. Observations range from 01/01/2016 to 01/06/2022.

Roughness analysis

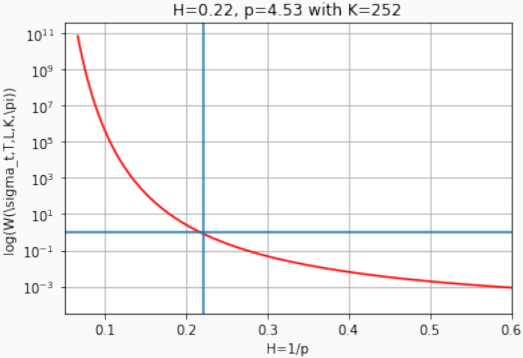


Figure 8: Estimate for the Hurst parameter for S&P500 with $K = 252$. Variance has been pre-sampled at 5 minutes.

It would be interesting to test on something dirtier..

A different proposal

Problems with Cont and Das 2022b in this phase:

- Are we sure we can call $\Delta(RV, \sigma_t)$ micro-structure noise on $S\&P500$? (notice: no σ_t and data already cleaned up)
- Are 5 minutes enough to rule out noise off of signals? Or is it too much?

A different proposal

Problems with Cont and Das 2022b in this phase:

- Are we sure we can call $\Delta(RV, \sigma_t)$ micro-structure noise on $S\&P500$? (notice: no σ_t and data already cleaned up)
- Are 5 minutes enough to rule out noise off of signals? Or is it too much?

$\Longrightarrow$ Building a new experiment!

MSFT x Hausmann

It would be better to carry on experimenting with a rationale:

- High frequency observations for simulated MSFT stock;
- Hausmann test on the mid-price $\rightarrow$ detecting noise frequency;
- Carry on same analysis with $W(K, L, p)$ estimator.

$\Longrightarrow$ Are their claims still correct?

Data

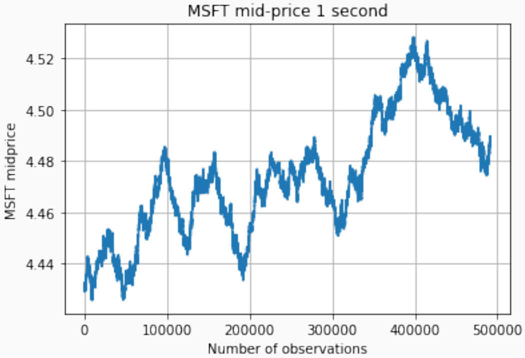


Figure 9: Series of the trade-price of the stock MSFT over the period April 1, 2018 - April 30, 2018. The sampling frequency adopted is one second. Data courtesy of Mariotti, Lillo, and Toscano 2022's authors.

Noise or not Noise?

According to Aït-Sahalia and Xiu 2019 presence of noise can be tested via an Hausman test.

- Assume $X_t = \sigma_0 dW_t$ that can be contaminated by noise $\rightarrow$ $\tilde{X}_t = X_t + a_0 U$ with $U \sim \mathcal{N}(0,1)$
- In absence of noise: $\Delta_n^{1/2}(\hat{\sigma}^2_{RV} - \sigma_0^2) \xrightarrow{d} \mathcal{N}(0, 2\sigma_0^4/T)$ however in presence of noise $\hat{\sigma}^2_{RV} = \sigma_0^2 + 2a_0^2\Delta_n^{-1} + o_p(1) \implies \hat{\sigma}^2_{RV} \rightarrow +\infty$ as $\Delta_n \rightarrow 0$.
- Consider the Maximum Likelihood Estimator for σ introduced by Aït-Sahalia, Mykland, and Zhang 2005[6]: $\Delta_n^{1/2}(\hat{\sigma}^2_{MLE} - \sigma_0^2) \xrightarrow{d} \mathcal{N}(0, 6\sigma_0^4/T)$[7]
- Leading to an Hausman test: $H_{1n} = \Delta_n^{-1} \frac{(\hat{\sigma}^2_{MLE} - \hat{\sigma}^2_{RV})}{\hat{V}_{1n}}$

$$\begin{cases} H_{1n} \xrightarrow{d} \chi^2_1 \text{ under } \mathbb{H}_0 : a^2 = 0 \\ H_{1n} \xrightarrow{p} \infty \text{ under } \mathbb{H}_1 : a^2 > 0 \end{cases} \tag{11}$$

[6] In this setting, a noise-robust parametric estimator is the MLE

[7] higher because we are controlling for noise when in fact there can be none

Results of Hausman test

Hausman test results at different frequencies for Microsoft stock (in April 2018). The symbol $\star$ ($\dagger$) indicates that the null of absence of noise is rejected (not rejected)[8] with a significance level of 5%.[9]

seconds	1	2	5	10	15	30	60
mid-price	$\star$	$\star$	$\star$	$\star$	$\star$	$\dagger$	$\dagger$

$\implies$ Use the $W(K, L, p)$ estimator should thus confirm the fact that micro-structure noise is responsible for roughness, ‘knowing’ the frequencies where noise can be found.[10]

[8] Rejection means that σ_{MLE} is inconsistent thus no noise. Notice: constant variance assumption.

[9] 9

[10] data: courtesy of Mariotti, Lillo, and Toscano 2022

Hurst parameter for MSFT

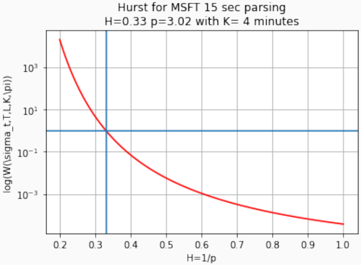


Figure 10: Hurst parameter estimation for MSFT with noise. 4 minute sampling

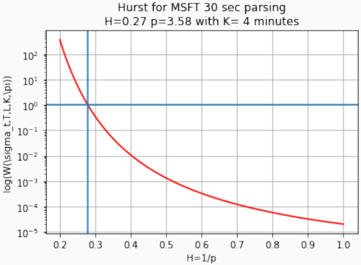


Figure 11: Hurst parameter estimation for MSFT without noise. 4 minute sampling

Hurst parameter for MSFT

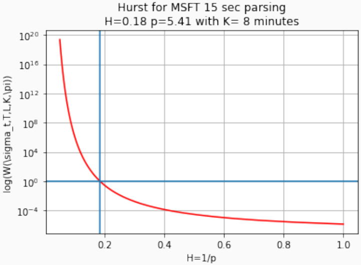


Figure 12: Hurst parameter estimation for MSFT with noise. 8 minute sampling

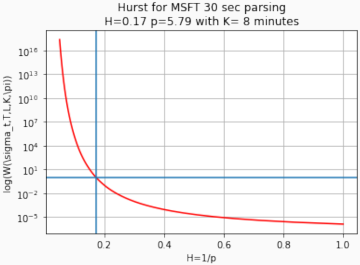


Figure 13: Hurst parameter estimation for MSFT without noise. 8 minute sampling

Hurst parameter for MSFT

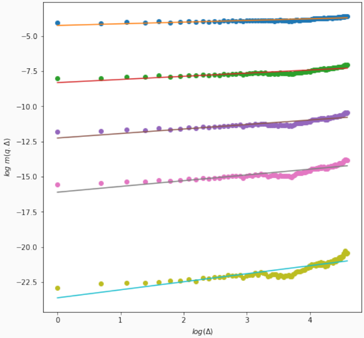


Figure 14: Hurst parameter estimation for MSFT for each q,$m(q, \Delta) \propto \Delta^{\zeta_q}$. 5 minute sampling

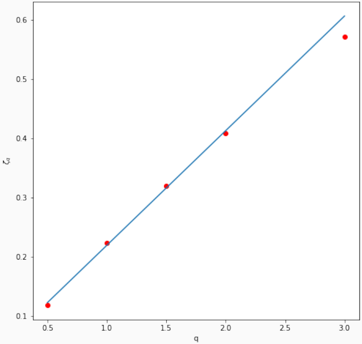


Figure 15: Hurst parameter estimation for MSFT *à la Gatheral* H=0.19. 5 minute sampling

Conclusions

Rough Volatility or Micro-structure Noise?

Finally, it can be concluded that:

- $H < 0.5$ in the fictional examples is showing an 'apparent roughness' $\Longleftarrow$ discretisation errors/micro-structure noise.
- What is observed in the market is consistent with an SV model driven by a Brownian motion $\Longleftarrow$ origins of roughness lie in micro-structure noise rather than the noise process driving spot volatility.
- Tests on S&P500 and on MSFT are not consistent with tests on fictional data spot driven by Brownian motions $\Longrightarrow$ different kind of behaviour.

$\Longrightarrow$ Observed roughness is compatible with $H = 0.5 \rightarrow$ Roughness is a characteristic of RV in lab experiments!

$\Longrightarrow$ On S&P500 we cannot conclude anything

$\Longrightarrow$ Experiments on MSFT contradict the thesis

Thank you for your attention!

References

- Aït-Sahalia, Yacine, Per A. Mykland, and Lan Zhang (Feb. 2005). "How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise". In: *The Review of Financial Studies* 18.2, pp. 351–416. ISSN: 0893-9454. DOI: *10.1093/rfs/hhi016*. eprint: *https://academic.oup.com/rfs/article-pdf/18/2/351/24421431/hhi016.pdf*. URL: *https://doi.org/10.1093/rfs/hhi016*.
- Aït-Sahalia, Yacine and Dacheng Xiu (2019). "A Hausman test for the presence of market microstructure noise in high frequency data". In: *Journal of Econometrics* 211.1, pp. 176–205.

Bennedsen, Mikkel, Asger Lunde, and Mikko S. Pakkanen (2017). “Hybrid scheme for Brownian semistationary processes”. In: *Finance and Stochastics* 21.4, pp. 931–965.

Comte, Fabienne and Eric Renault (1998). “Long memory in continuous-time stochastic volatility models”. In: *Mathematical finance* 8.4, pp. 291–323.

Cont, Rama and Purba Das (2022a). *Measuring the roughness of a signal.*

— (2022b). *Rough Volatility: Fact or Artefact?* URL: *arxiv.org/abs/2203.13820*.

Dudley, Richard M (2010). “Sample functions of the Gaussian process”. In: *Selected works of RM Dudley*. Springer, pp. 187–224.

El Euch, Omar, Masaaki Fukasawa, and Mathieu Rosenbaum (2016). *The microstructural foundations of leverage effect and rough volatility*. DOI: `10.48550/ARXIV.1609.05177`. URL: `https://arxiv.org/abs/1609.05177`.

Fukasawa, Masaaki, Tetsuya Takabatake, and Rebecca Westphal (2019). "Is volatility rough?" In: *arXiv preprint arXiv:1905.04852*.

Gatheral, Jim, Thibault Jaisson, and Mathieu Rosenbaum (2018). "Volatility is rough". In: *Quantitative finance* 18.6, pp. 933–949.

Gatheral, Jim and Roel C.A. Oomen (2010). "Zero-Intelligence Realized Variance Estimation". In: *Finance and Stochastics* 14.2, pp. 249–283.

Mandelbrot, B. and J.W. van Ness (1968). "Fractional Brownian motions, fractional noises and applications". In: *SIAM Review* 10.4, pp. 422–437.

- Mariotti, Tommaso, Fabrizio Lillo, and Giacomo Toscano (2022). *From Zero-Intelligence to Queue-Reactive: Limit Order Book modeling for high-frequency volatility estimation and optimal execution*. DOI: 10.48550/ARXIV.2202.12137. URL: https://arxiv.org/abs/2202.12137.
- Rogers, L. (2019). *Things we think we know*. URL: %5Cverb!https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjL3fW5gtX4AhUHQfEDHeDjC5gQFnoECAUQAQ&url=https%3A%2F%2Fwww.skokholm.co.uk%2Fwp-content%2Fuploads%2F2019%2F11%2FTWTWKpaper.pdf&usg=AOvVaw1ENrqp5ZuuhphEClc6VrUt!.

...However

- No properties for the estimator $W(\cdot)$ have been given,
- Estimates of $W(\cdot)$ really depend on the sampling chosen, which is the optimal sampling for the estimator?
- OM data-set is already corrected for many features called out by the authors $\rightarrow$ other features?
- RV may not be the ideal benchmark $\implies$ other estimators as in Gatheral and Oomen 2010 (σ_{MLE}, Realised kernel etc.)
- repeating the same procedure with other estimators of H (variance/covariance of increments cfr. Bennedsen, Lunde, and Pakkanen 2017)
- micro-structure noise is already accounted for in El Euch, Fukasawa, and Rosenbaum 2016 $\impliedby$ this is what generates leverage effect and roughness

$\implies$ Simpler models do not fit the ATM skew and volatility surface! $\rightarrow$ shall we model the non-IID noise? Will we replicate what is on the market?

Rogers model OU-OU

Rogers 2019 proposes a different perspective based on two arguments:

- Rough models are not really Markovian, so we need to know the entire history to make predictions
- what economic story could we tell that would result in a model where we need to know the entire history in order to predict the future?

$\implies$ Since the value of H varies from one index to another, there is clearly no universal law applicable to all assets

Rogers model OU-OU

Try an energetic OU process mean-reverting to a slower one for volatility:

$$\begin{cases} dY_t = \sigma_y dW'_t - \beta Y_t dt \\ dX_t = \sigma_x dW_t - \lambda(Y_t - X_t)dt \end{cases} \tag{12}$$

$\implies$ good convergence in values for long range, while for small range (high frequency) correct with additive IID noise $\implies$ almost perfect fit with data with noise modelled as MA(3)

Distribution of estimator for K=250

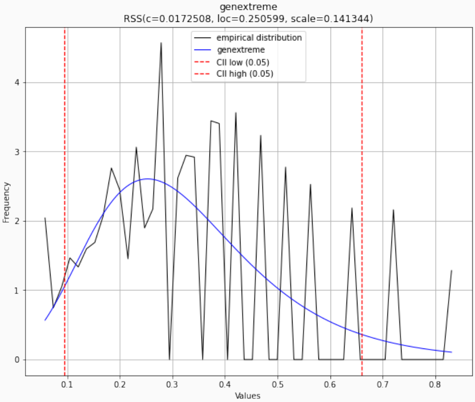


Figure 16: Fit of empirical distribution of $W(\cdot)$ estimator for OU SV (Gumbel distribution). $n = 5000, t = 23400, K = 250$

Distribution of estimator for K=500

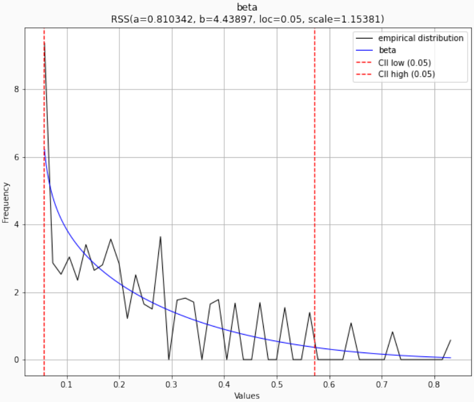


Figure 17: Fit of empirical distribution of $W(\cdot)$ estimator for OU SV (Beta distribution). $n = 5000, t = 23400, K = 500$

Distribution of estimator for K=10

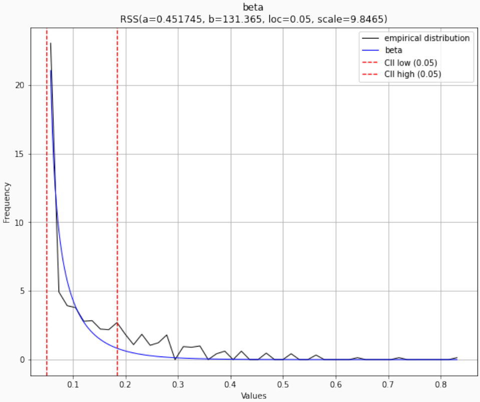


Figure 18: Empirical distribution of $W(\cdot)$ estimator for OU SV. $n = 5000, t = 23400, K = 10$

Proof theo page 11(1)

Assume $g(u) = d/du[x]^p_\pi(u)$. Since the p-variation is strictly increasing we have $\sup_{t\in[0,T]} \frac{1}{g(u)} \in (0,\infty)$ so by the mean value theorem:

$$\begin{aligned} w^n(t) &= \sum_{\pi^n\cap[0,t]} \frac{|x(t^n_{i+1}) - x(t^n_i)|^p}{[x]^p_\pi(t^n_{i+1}) - [x]^p_\pi(t^n)} \times (t^n_{i+1} - t^m i) \\ &= \sum_{\pi^n\cap[0,t]} \frac{|x(t^n_{i+1}) - x(t^n_i)|^p}{g(u^n_i)} \end{aligned} \tag{13}$$

By properties of riemman integral

$$\begin{aligned} &\sum_{\pi^n\cap[0,t]} \frac{|x(t^n_{i+1}) - x(t^n_i)|^p}{d(u^n_i)} \xrightarrow{n\to\infty} \int_0^t \frac{1}{g(u)} d[x]^p_\pi(u) \\ &= \int_0^t \frac{d[x]^p_\pi(u)/u}{g(u)du} = \int_0^t du = t \end{aligned}$$

The limit always exists and this concludes the proof

proof of (2)

proof for one path $\omega \in \Omega$ of a BM, $[\omega]_\pi(t) = t$, for a partition with vanishing mesh

$$\begin{aligned} w(\pi, 2, t) &= \sum_{\pi^n \cap [0,t]} \frac{|\omega(t^n_{i+1}) - \omega(t^n_i)|^2}{[\omega]^p_\pi(t^n_{i+1}) - [\omega]^2_\pi(t^n)} \times (t^n_{i+1} - t^m_i) \\ &= \sum_{\pi^n \cap [0,t]} \frac{|\omega(t^n_{i+1}) - \omega(t^n_i)|^p}{(t^n_{i+1} - t^m i)} \times (t^n_{i+1} - t^m_i) \\ &= \sum_{\pi^n \cap [0,t]} |\omega(t^n_{i+1}) - \omega(t^n_i)|^2 = t \quad a.s. \end{aligned} \tag{14}$$

As a consequence of the theorem:

$$\mathbb{P}(w(X, 2, \pi) = t) = 1 \qquad x_t = \int_0^t \sigma_u dB_t$$

QR limit order book 35

- correlation not only between order-arrival intensities and the corresponding queue size at each level, but also between intensities and the queue size at the corresponding level at the opposite side of the book
- dependence between the volume at the best level and order arrivals at the other levels is assumed
- QR model allows for exogenous dynamics by taking into account the flow of exogenous information that hits the market.

The test evaluates the consistency of an estimator when compared to an alternative, less efficient estimator which is already known to be consistent. It helps one evaluate if a statistical model corresponds to the data. 11

Volatility Modelling

Roughly speaking, volatility has been modelled in two major ways:

Local Volatility

$\sigma(t, S_t) \Longleftarrow$ continuum of quoted options:

$$\sigma_{\text{dup}}^2 = \frac{\frac{\partial C}{\partial T} + rK\frac{\partial C}{\partial K}}{\frac{1}{2}K^2\frac{\partial^2 C}{\partial K^2}}$$

Stochastic Volatility

$\sigma_t \Longleftarrow$ dynamics driven by a BM:

$$d\sigma_t = \mu(t, \sigma_t)dt + \Lambda(t, \sigma_t)dW_t$$