Volatility is rough

Jim Gatheral

Baruch College WEISSMAN SCHOOL OF ARTS AND SCIENCES | Master of Financial Engineering

Baruch College

March 12, 2024

Outline of this talk

- The shape of the volatility surface
    - Scaling of implied volatility smiles
- Monofractal scaling of historical volatility
- Fractional Brownian motion (fBm)
- The Rough Fractional Stochastic Volatility (RFSV) model
- Estimating the roughness parameter H
- Forecasting realized variance

The SPX volatility surface as of 15-Feb-2023

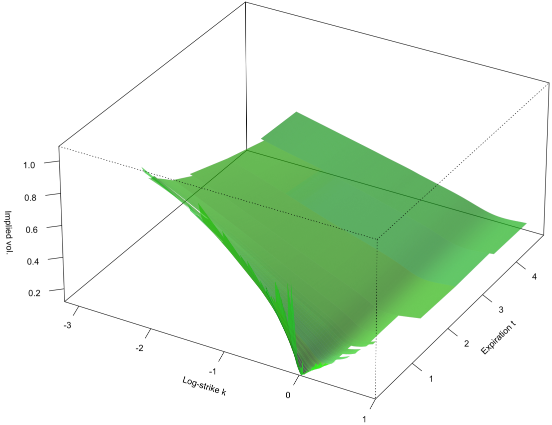


Figure 1: The SPX volatility surface as of 15-Feb-2023.[1]

[1] Data from OptionMetrics via WRDS.

Remarks on Figure 1

- Figure 1 is a slightly smoothed plot of estimated mid volatilities, not a fit!
    - There were 48 expirations and 6,749 put/call option pairs with non-zero bids as of the close on 15-Feb-2023.
- Notice how smooth this volatility surface is!
    - Bumps or dips would be tradable.

- Although the level and orientation of the volatility surface changes over time, it is a stylized fact that its rough shape stays very much the same.
    - The surface as of 15-Feb-2023 is typical.

SPX volatility smiles as of 15-Feb-2023

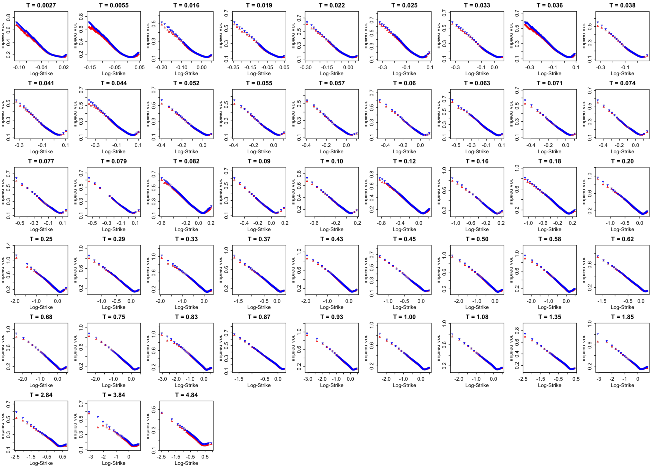


Figure 2: SPX volatility smiles as of 15-Feb-2023.

Interpreting a single smile

- We could say that the volatility smile (at least in equity markets) reflects two basic observations:
    - Volatility tends to increase when the underlying price falls,
        - hence the negative skew.
    - We don't know in advance what realized volatility will be,
        - hence implied volatility is increasing in the wings.
- It's implicit in the above that more or less any model that is consistent with these two observations will be able to fit one given smile.
    - Fitting two or more smiles simultaneously is much harder.

Term structure of at-the-money skew

- Given one smile for a fixed expiration, little can be said about the process generating it.
- In contrast, the dependence of the smile on time to expiration is intimately related to the underlying dynamics.
    - In particular model estimates of the term structure of ATM volatility skew defined as

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

are very sensitive to the choice of volatility dynamics in a stochastic volatility model.

Term structure of SPX ATM skew as of 15-Feb-2023

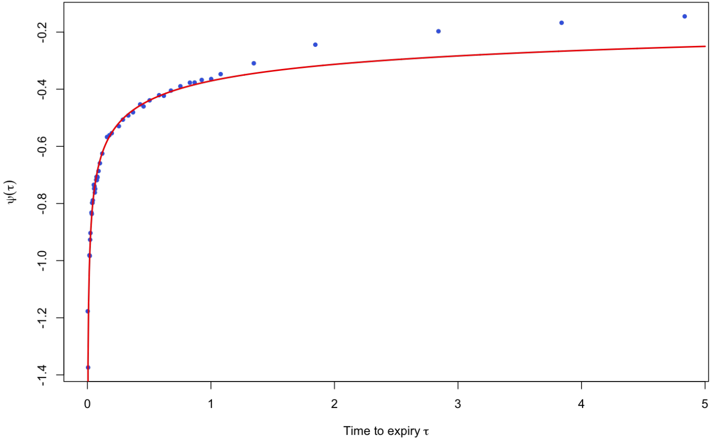


Figure 3: Term structure of ATM skew as of 15-Feb-2023, with power law fit $\tau^{-0.24}$ superimposed in red.

Stochastic volatility models

- A generic stochastic volatility model takes the form

$$\frac{dS_t}{S_t} = \sqrt{V_t}\, dZ_t$$
$$V_t = \int_{-\infty}^{t} F(\Omega_s)\, dW_s,$$

where $V_t\, dt = d\,\langle \log S\rangle_t$, F is some function, and Ω_t is the natural filtration generated by Z and W.

Fractional stochastic volatility models

- Non-Markovian models of the form

$$V_t = V_0 \exp\left\{\eta \int_0^t \frac{dW_s}{(t-s)^\gamma} + \text{ drift }\right\}$$

were shown by Alòs et al. in [ALV07] and then by Fukasawa in [Fuk11] to generate a short-dated ATM skew of the form

$$\psi(\tau) \sim \tau^{-\gamma}$$

with $\gamma = \frac{1}{2} - H$ and $0 < H < \frac{1}{2}$.

  - Such models, where the kernel decays as a power-law for small times, are called *rough volatility* models.

- The typical power-law behavior of the skew term structure for short times is one of the motivations for rough volatility models.

Skew term structure is not always power-law

- [GA23] show that a power-law fits the average skew term structure poorly for very short dates.
- [DDS23] show that skew term structure is typically a combination of two power-laws.

- We confirm in Figure 4, that the term structure of skew is not always power-law.
  - On 27-Dec-2022, the skew term structure is not even monotonic!
- We further confirm in Figure 5 that the skew term structure looks like a combination of two power-laws, at least on 15-Feb-2023, consistent with [DDS23].

Skew term structure is not always power-law

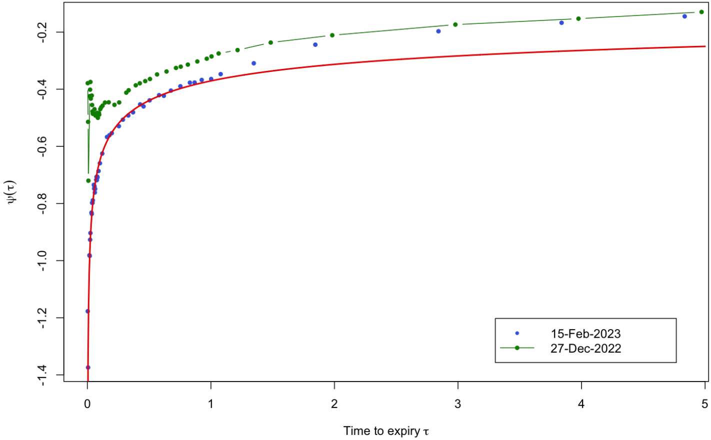


Figure 4: ATM skew term structure on two different dates. On 27-Dec-2022, the skew term structure is not even monotonic!

Log plot of skew term structure

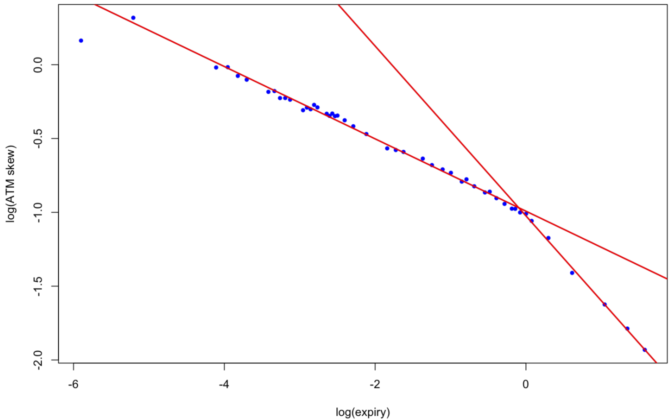


Figure 5: The skew term structure on 15-Feb-2023 looks like a superposition of two power-laws.

Total variance plots

- Define the implied total variance $w(k,\tau) := \sigma_{\rm BS}(k,\tau)^2\,\tau$.
- To avoid calendar spread arbitrage, we must have $w(k,\tau)$ non-decreasing in τ for fixed k.
    - If lines on a total variance plot cross, there is calendar spread arbitrage.
- The non-monotonic skew term structure on 27-Dec-2022 leads one to suspect calendar spread arbitrage.

- Let's check...

Total variance plot as of 15-Feb-2023

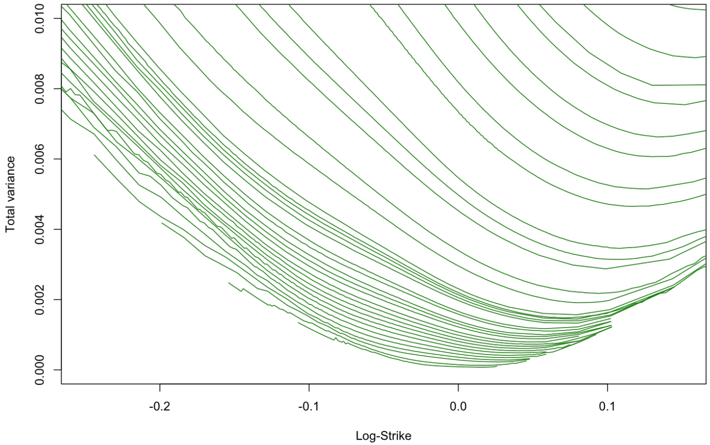


Figure 6: Beautiful total variance plot on 15-Feb-2023. No calendar spread arbitrage and no obvious trading opportunities.

Total variance plot as of 27-Dec-2022

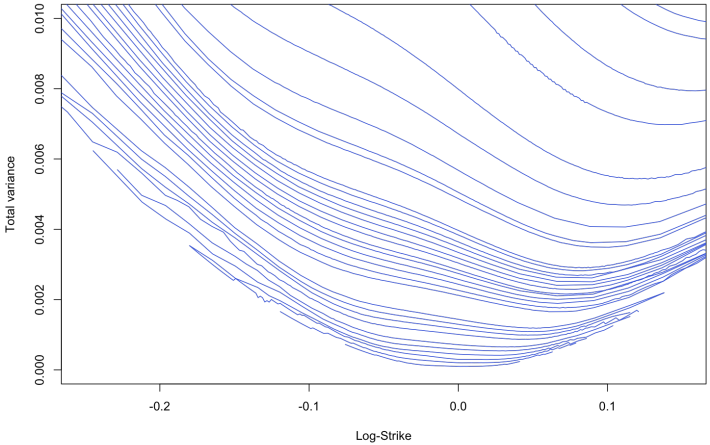


Figure 7: On 27-Dec-2022, no calendar spread arbitrage. However, some individual total variance curves are W-shaped.

Scaling of total variance

- The rough SABR formula of [FG22] suggests that we should have

$$\frac{w(k,\tau)}{w(0,\tau)} \approx f\left(\tau^{-\gamma}\,\frac{k}{\Sigma_{BS}(0)}\right).$$

  - Roughly speaking, total variance curves should scale as a power-law.
- Figure 8 does suggest close-to-power-law scaling, even in the 27-Dec-2022 case.

Scaling of total variance

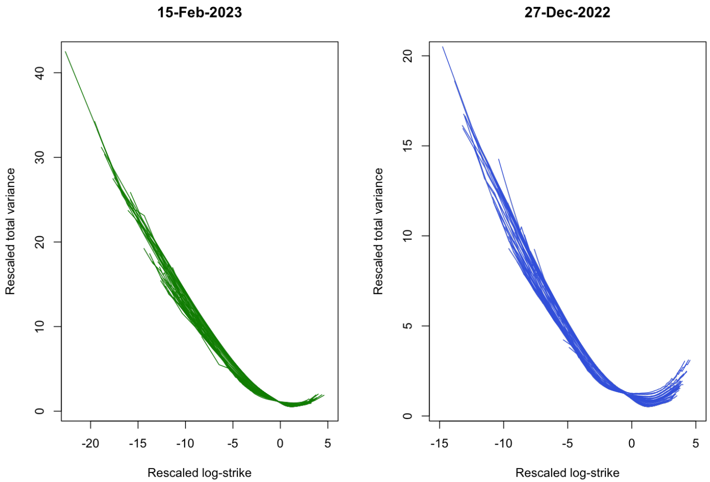


Figure 8: ATM skew term structure on two different dates. On 27-Dec-2022, the skew term structure is not even monotonic!

Fractional stochastic volatility models

- This simple scaling of volatility smiles suggests that rough volatility models should be consistent with option prices.
  - Despite that the term structure of skew is not always power-law.
- Were the instantaneous variance to follow something like

$$V_t = V_0 \exp\left\{ \eta \int_0^t \frac{dW_s}{(t-s)^\gamma} + \text{ drift } \right\},$$

  the time series of $\log V_t$ should also have simple scaling properties.

- Let's check ...

Power-law scaling of the historical volatility process

- The Oxford-Man Institute of Quantitative Finance used to make historical realized variance (RV) estimates freely available.
  - Unfortunately, no longer. The last date in my dataset is 06/28/2022.
- Using daily RV estimates as proxies for instantaneous variance, we may investigate the time series properties of V_t empirically.

History of SPX realized variance

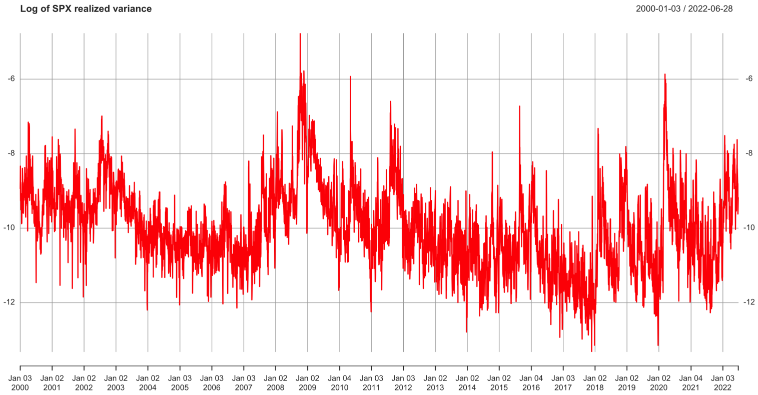


Figure 9: (Log) realized kernel estimates of SPX realized variance.

The smoothness of the volatility process

- For $q \geq 0$, we define the qth sample moment of differences of log-volatility at a given lag Δ[2]:

$$m(q,\Delta) = \langle |\log \sigma_{t+\Delta} - \log \sigma_t|^q \rangle$$

- For example

$$m(2,\Delta) = \left\langle (\log \sigma_{t+\Delta} - \log \sigma_t)^2 \right\rangle$$

is just the sample variance of differences in log-volatility at the lag Δ.

[2] $\langle \cdot \rangle$ denotes the sample average.

Scaling of $m(q, \Delta)$ with lag Δ

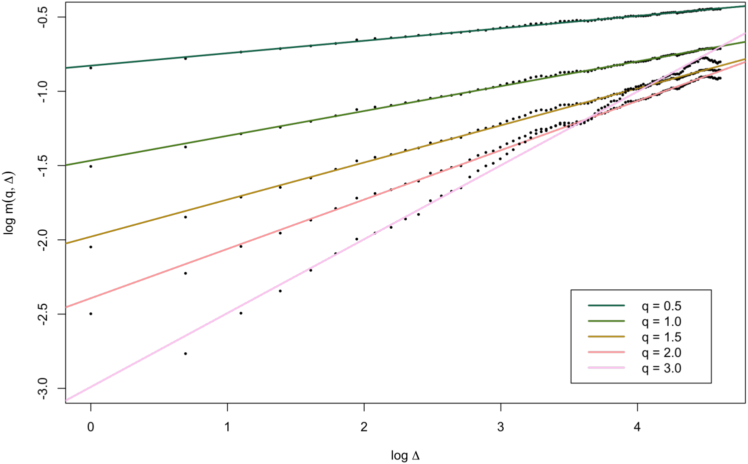


Figure 10: $\log m(q, \Delta)$ as a function of $\log \Delta$, SPX.

Scaling of ζ_q with q

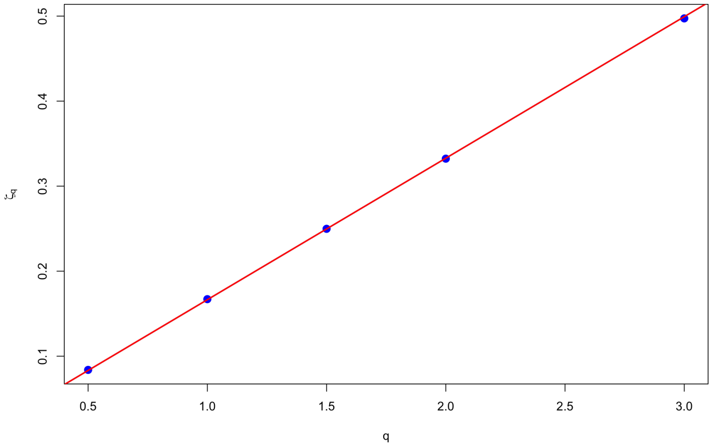


Figure 11: Scaling of ζ_q with q.

Monofractal scaling result

- From the log-log plot Figure 10, we see that for each q, $m(q,\Delta) \propto \Delta^{\zeta_q}$.
- And from Figure 11 the monofractal scaling relationship

$$\zeta_q = q\,H$$

  with $H \approx 0.17$.
    - Note also that our estimate of H is biased high because we proxied instantaneous variance V_t with its average over each day $\frac{1}{T}\int_0^T V_t\,dt$, where T is one day.
    - On the other hand, the time series of realized variance is noisy and this causes our estimate of H to be biased low.
- It is easily checked that H is not a constant, but varies with time.

Distributions of $(\log \sigma_{t+\Delta} - \log \sigma_t)$ for various lags Δ

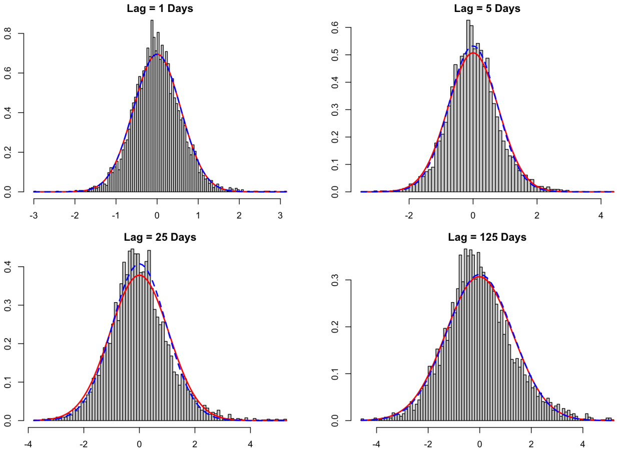


Figure 12: Histograms of $(\log \sigma_{t+\Delta} - \log \sigma_t)$ for various lags Δ; normal fit in red; $\Delta = 1$ normal fit scaled by $\Delta^{0.17}$ in blue.

Estimated H for all indices

Estimating the relationship

$$\left\langle (\log \sigma_{t+\Delta} - \log \sigma_t)^2 \right\rangle = \nu^2\, \Delta^{2\,H}$$

for all 31 indices in the Oxford-Man dataset yields:

Index	H	ν
.AEX	0.15	0.28
.AORD	0.11	0.30
.BFX	0.15	0.26
.BSESN	0.13	0.29
.BVLG	0.14	0.24
.BVSP	0.14	0.29
.DJI	0.17	0.28
.FCHI	0.14	0.29
.FTMIB	0.15	0.26
.FTSE	0.14	0.28
.GDAXI	0.15	0.27
.GSPTSE	0.15	0.30
.HSI	0.12	0.23
.IBEX	0.13	0.27
.IXIC	0.16	0.29
.KS11	0.12	0.28

Index	H	ν
.KSE	0.11	0.39
.MXX	0.09	0.29
.N225	0.13	0.30
.NSEI	0.13	0.32
.OMXC20	0.11	0.30
.OMXHPI	0.13	0.31
.OMXSPI	0.13	0.31
.OSEAX	0.14	0.26
.RUT	0.13	0.35
.SMSI	0.12	0.32
.SPX	0.17	0.30
.SSEC	0.13	0.32
.SSMI	0.19	0.19
.STI	0.07	0.24
.STOXX50E	0.11	0.35

Universality?

- In [GJR18], we compute daily realized variance estimates over one hour windows for DAX and Bund futures contracts, finding similar scaling relationships.
- We have also checked that Gold and Crude Oil futures scale similarly.
  - Although the increments $(\log \sigma_{t+\Delta} - \log \sigma_t)$ seem to be fatter tailed than Gaussian.
- In [BLP22] Bennedsen et al., estimate volatility time series for more than five thousand individual US equities, finding rough volatility in every case.

A natural model of realized volatility

- Distributions of differences in the log of realized volatility are close to Gaussian.
  - This motivates us to model σ_t as a lognormal random variable.
- Moreover, the scaling property of variance of RV differences suggests the model:

$$\log \sigma_{t+\Delta} - \log \sigma_t = \nu \left(W^H_{t+\Delta} - W^H_t \right) \tag{1}$$

where W^H is fractional Brownian motion.
  - If σ were continuous, distributions of $log \sigma$ were really Gaussian, and if H were constant, this model would be unique!
- In [GJR18], we refer to a stationary version of (1) as the RFSV (for Rough Fractional Stochastic Volatility) model.

Fractional Brownian motion (fBm)

- *Fractional Brownian motion* (fBm) $\{W_t^H; t \in \mathbb{R}\}$ is the unique Gaussian process with mean zero and autocovariance function

$$\mathbb{E}\left[W_t^H\, W_s^H\right] = \frac{1}{2}\left\{|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right\}$$

  where $H \in (0,1)$ is called the *Hurst index* or parameter.
- In particular, when $H = 1/2$, fBm is just Brownian motion.
  - If $H > 1/2$, increments are positively correlated.
  - If $H < 1/2$, increments are negatively correlated.

Volatility time series: data vs model

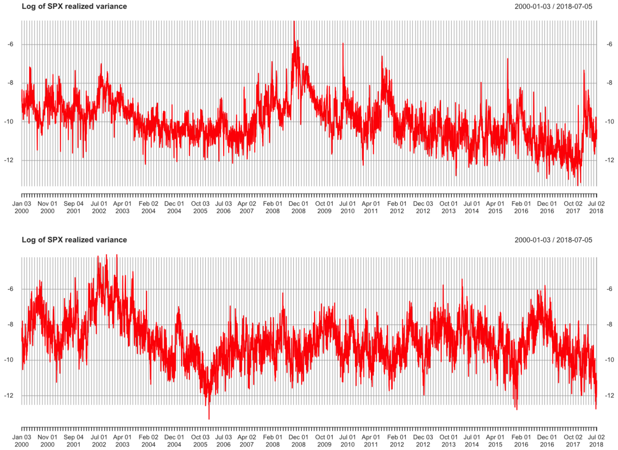


Figure 13: Log-volatility of SPX: True (above) and simulated (below).

Remarks on the comparison

- The qualitative features of simulated and actual graphs look very similar.
    - Persistent periods of high volatility alternate with low volatility periods.
- $H \sim 0.1$ generates very rough looking sample paths (compared with $H = 1/2$ for Brownian motion).
    - Hence *rough volatility*.
- On closer inspection, we observe fractal-type behavior.
    - The graph of volatility over a small time period looks like the same graph over a much longer time period.
- This feature of volatility has been investigated both empirically and theoretically in, for example, [BM03].
    - In particular, their Multifractal Random Walk (MRW) is related to a limiting case of the RSFV model as $H \to 0$.

More sophisticated estimators of H

- Numerous authors have pointed out that the estimates of H by linear regression in [GJR18] make sense only if the signal-to-noise ratio is sufficiently high.
  - A semimartingale volatility process with substantial estimation error would yield spuriously low estimates of H.
  - Some authors have *even* suggested that volatility may not be rough!
- More sophisticated estimators of H include
  - The ACF estimator of [BLP22]
  - The Whittle estimator of [FTW22]
  - The GMM estimator of [BCPV23]
- All of these authors conclude that volatility of SPX is indeed rough.

Heuristic derivation of the ACF estimator of [BLP22]

- Once again, the covariance structure of fBm is given by

$$\mathbb{E}\left[W_t^H\, W_s^H\right] = \frac{1}{2}\left\{t^{2H} + s^{2H} - |t-s|^{2H}\right\}.$$

- Up to a multiplicative factor, our model is

$$y_t = \log V_t = W_t^H.$$

- Then $\operatorname{var}[y_t] = t^{2H}$. and

$$\operatorname{cov}[y_t, y_{t+\Delta}] = \frac{1}{2}\left\{t^{2H} + (t+\Delta)^{2H} - \Delta^{2H}\right\}$$

- Dividing the covariance by the variance gives

$$\rho(\Delta) = \frac{1}{2}\left\{1 + \left(1 + \frac{\Delta}{t}\right)^{2H} - \left(\frac{\Delta}{t}\right)^{2H}\right\}$$

- Thus, for Δ/t sufficiently small,

$$1 - \rho(\Delta) = \frac{1}{2}\left(\frac{\Delta}{t}\right)^{2H} + O\left(\frac{\Delta}{t}\right).$$

  - Note in particular that we expect the ACF estimator to work best when $H \ll \frac{1}{2}$.

- Taking logs of each side, we obtain

$$\log(1 - \rho(\Delta)) = a + 2\,H\,\log\Delta.$$

  - H may be estimated efficiently by linear regression.

Graph of ACF-estimated H

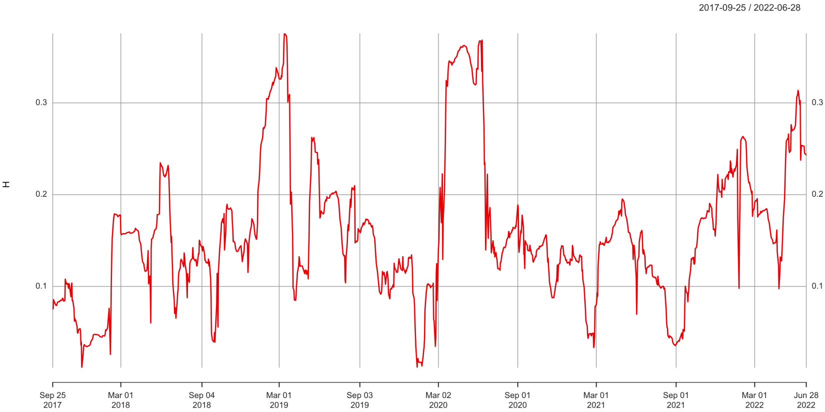


Figure 14: Plot of ACF estimates of H using 61 day windows. H is not constant!

The Chong-Todorov test

- Recently, in [CT23], Chong and Todorov devised a robust nonparametric test for rough volatility based on the observation that

  > *"volatility is rough only if changes in volatility are negatively correlated at high frequency".*

- They conclude

  > *"Implementing the test on SPY transaction data, we find evidence in support of rough volatility throughout the past eleven years".*

Fukasawa's robust regression

- Theorem 4 of [Fuk23] shows under lognormal assumptions,that if $H < \frac{1}{2}$, the covariance between realized variance and log-return over the period Δ should scale as $\Delta^{H+3/2}$.
- More precisely,

$$\operatorname{cov}\left(\sum_{i=1}^{k} \log R_{t+i\Delta}, X_{t+k\Delta} - X_t\right) \sim (k\Delta)^{H+\frac{3}{2}},$$

  where $X = \log S$ and R is realized variance.
- This should be very robust because the measurement errors in R and ΔX should be independent to a first approximation.
  - Unlike for example the errors in $\operatorname{var}\left(\sum_{i=1}^{k} \log R_{t+i\Delta}\right)$, which add.

Fukasawa's robust regression plot

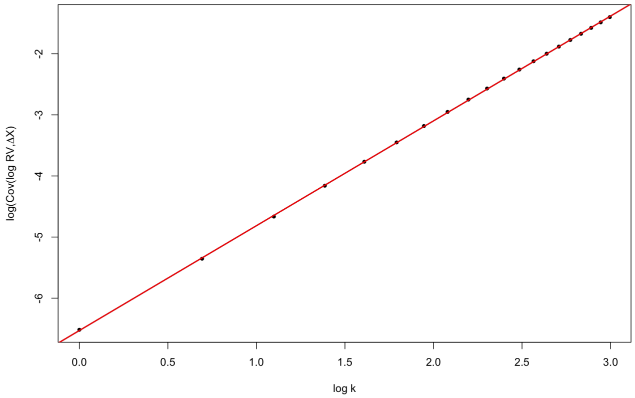


Figure 15: The covariance between log return and log realized variance vs $\log k$. The slope of the line is 1.7156 ± 0.0027 giving $H = 0.2156 \pm 0.0027$.

The forecast formula

- In the RFSV model (1), $\log V_t \approx 2\,\nu\, W^H_t + C$ for some constant C.
- [NP00] show that $W^H_{t+\Delta}$ is conditionally Gaussian with conditional expectation

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

and conditional variance

$$\mathrm{Var}[W^H_{t+\Delta}|\mathcal{F}_t] = c\,\Delta^{2H}.$$

where

$$c = \frac{\Gamma(3/2-H)}{\Gamma(H+1/2)\,\Gamma(2-2H)}.$$

The forecast formula

- Thus, we obtain

Variance forecast formula

$$\mathbb{E}^{\mathbb{P}}\left[V_{t+\Delta}\middle|\mathcal{F}_t\right] = \exp\left\{\mathbb{E}^{\mathbb{P}}\left[\log(V_{t+\Delta})\middle|\mathcal{F}_t\right] + 2\,c\,\nu^2\Delta^{2\,H}\right\} \tag{2}$$

where

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}}\left[\log V_{t+\Delta}\middle|\mathcal{F}_t\right]\\&= \frac{\cos(H\pi)}{\pi}\Delta^{H+1/2}\int_{-\infty}^{t}\frac{\log V_s}{(t-s+\Delta)(t-s)^{H+1/2}}ds.\end{aligned}$$

- [BLP22] confirm that this forecast outperforms the best performing existing alternatives such as HAR, at least at daily or higher timescales.

Pricing under rough volatility

Once again, the data suggests the following model for volatility under the real (or historical or physical) measure $\mathbb{P}$:

$$\log \sigma_t = \nu \, W^H_t.$$

Let $\gamma = \frac{1}{2} - H$. We choose the Mandelbrot-Van Ness representation of fractional Brownian motion W^H as follows:

$$W^H_t = C_H \left\{ \int_{-\infty}^{t} \frac{dW^{\mathbb{P}}_s}{(t-s)^\gamma} - \int_{-\infty}^{0} \frac{dW^{\mathbb{P}}_s}{(-s)^\gamma} \right\}$$

where the choice

$$C_H = \sqrt{\frac{2\,H\,\Gamma(3/2-H)}{\Gamma(H+1/2)\,\Gamma(2-2\,H)}}$$

ensures that

$$\mathbb{E}\left[W^H_t \, W^H_s\right] = \frac{1}{2}\left\{t^{2H} + s^{2H} - |t-s|^{2H}\right\}.$$

Pricing under rough volatility

Then

$$
\begin{aligned}
&\log V_u - \log V_t \\
&= 2\,\nu\, C_H \left\{ \int_t^u \frac{1}{(u-s)^\gamma}\, dW_s^{\mathbb{P}} + \int_{-\infty}^t \left[\frac{1}{(u-s)^\gamma} - \frac{1}{(t-s)^\gamma} \right] dW_s^{\mathbb{P}} \right\} \\
&=: 2\,\nu\, C_H\, \left[M_t(u) + Z_t(u) \right]. \qquad (3)
\end{aligned}
$$

- Note that $\mathbb{E}^{\mathbb{P}}\left[M_t(u) \middle| \mathcal{F}_t \right] = 0$ and $Z_t(u)$ is $\mathcal{F}_t$-measurable.
- To price options, it would seem that we would need to know $\mathcal{F}_t$, the entire history of the Brownian motion W_s for $s < t$!

The variance process under $\mathbb{P}$

Exponentiating (3), we get

$$
\begin{aligned}
V_u &= V_t \exp\left\{2\,\nu\, C_H\, M_t(u) + 2\,\nu\, C_H\, Z_t(u)\right\} \\
&= \mathbb{E}^{\mathbb{P}}\left[\, V_u \,|\, \mathcal{F}_t\right] \mathcal{E}\left(2\,\nu\, C_H \int_t^u \frac{dW_s^{\mathbb{P}}}{(u-s)^{\gamma}}\right).
\end{aligned}
$$

- The conditional distribution of V_u depends on $\mathcal{F}_t$ only through the variance forecasts $\mathbb{E}^{\mathbb{P}}\left[\, V_u \,|\, \mathcal{F}_t\right]$.
  - These variance forecasts depend explicitly on the history of the variance process.
  - Rough volatility models are (in principle) path-dependent!
- $\mathbb{E}^{\mathbb{Q}}\left[\, V_u \,|\, \mathcal{F}_t\right]$ can be obtained in a model-free manner from options prices.
  - For pricing, we don't need to know the history of the variance process.

Postscript

From [GJR18]:

> *It is of course plausible that other models are compatible with many of our observations. In fact, there are probably many ways to design a process so that most of our empirical results are reproduced (for example estimation errors when estimating volatility can be quite significant for some models, leading to downward biases in the measurement of the smoothness). However, what we show here is that we cannot find any evidence against the RFSV model. In statistical terms, the null hypothesis that the data generating process of the volatility is a RFSV model cannot be rejected based on our analysis.*

From the preface of [BFFGJR23]:

> *... any viable model describing the dynamics of volatility should exhibit a close resemblance to rough volatility.*

From [Fuk23]:

> *The pathwise roughness of volatility itself is, however, only secondary; volatility is only a hypothetical latent quantity in a diffusive scale, and its realization cannot be identified from finite data. Its roughness is a result of the distributional property of local self-similarity under a Brownian semimartingale framework. In this sense, the statement "volatility is rough" seems misleading and, indeed, has invoked some misunderstanding and confusion in the mathematical finance community.*

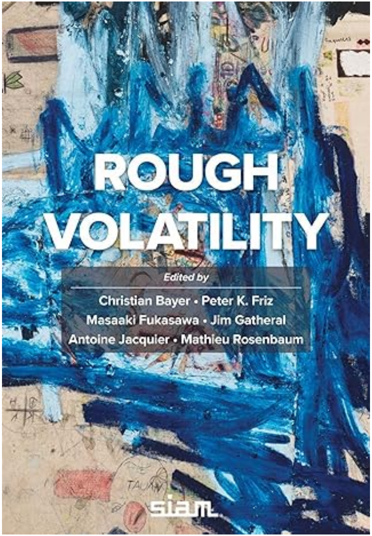
From [Fuk23] again:

> *A relevant scientific question is why volatility is rough (or more precisely, why the distributions of asset prices in the daily time scale have such universal properties that are well-explained by rough volatility). There are several studies that connect rough volatility and market microstructure dynamics. An economic reasoning is still absent. Research in this direction would require a deeper understanding of stochastic processes.*

Summary

- Approximate power-law scaling of volatility smiles suggests a scaling relationship for instantaneous variance.
- This leads us to incover a remarkable monofractal scaling relationship in historical volatility which now appears to be universal.
  - The rough volatility paradigm.
- A natural non-Markovian (path-dependent) stochastic volatility model under $\mathbb{P}$ then follows.
- The resulting volatility forecast beats existing alternatives.

For more inspiration ...



References

Elisa Alòs, Jorge A León, and Josep Vives.
On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility.
Finance and Stochastics, 11(4):571–589, 2007.

Christian Bayer, Peter K Friz, Masaaki Fukasawa, Jim Gatheral, Antoine Jacquier, and Mathieu Rosenbaum.
Rough Volatility.
SIAM, 2023.

Emmanuel Bacry and Jean François Muzy.
Log-infinitely divisible multifractal processes.
Communications in Mathematical Physics, 236(3):449–475, 2003.

Mikkel Bennedsen, Asger Lunde, and Mikko S Pakkanen.
Decoupling the short-and long-term behavior of stochastic volatility.
Journal of Financial Econometrics, 20(5), 961–1006, 2022.

Anine E. Bolko, Kim Christensen, Mikko S. Pakkanen and Bezirgen Veliyev.
A GMM approach to estimate the roughness of stochastic volatility.
Journal of Econometrics, 235(2) 745–778, 2023.

Carsten H. Chong and Victor Todorov.
The Fine Structure of Volatility Dynamics.
SSRN `https://papers.ssrn.com/abstract=4660138`, 2023.

Jules Delemotte, Stefano De Marco and Florent Segonne,
Yet Another Analysis of the SP500 At-The-Money Skew: Crossover of Different Power-Law Behaviours
SSRN, `https://papers.ssrn.com/abstract=4428407`, 2023.

Masaaki Fukasawa.
Asymptotic analysis for stochastic volatility: Martingale expansion.
Finance and Stochastics, 15(4):635–654, 2011.

Masaaki Fukasawa, Tetsuya Takabatake, and Rebecca Westphal,
Consistent estimation for fractional stochastic volatility model under high-frequency asymptotics.
Mathematical Finance, 32(4), 1086–1132, 2022.

Masaaki Fukasawa and Jim Gatheral.
A rough SABR formula.
Frontiers of Mathematical Finance, 1(1), 81–97, 2022.

Masaaki Fukasawa.
Wiener spiral for volatility modeling.
Theory of Probability & its Applications, 68(3), 481–500, 2023.

Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum.
Volatility is rough.
Quantitative Finance, 18(6):933–949, 2018.

Julien Guyon and Mehdi El Amrani
Does the Term-Structure of the At-the-Money Skew Really Follow a Power Law?
Risk.net, August 2023.

Carl J Nuzman and Vincent H Poor.
Linear estimation of self-similar processes via Lamperti's transformation.
Journal of Applied Probability, 37(2):429–452, 2000.