Rough volatility

Jim Gatheral
(joint work with Christian Bayer, Peter Friz,
Thibault Jaisson, Andrew Lesniewski, and Mathieu Rosenbaum)

Baruch COLLEGE
The City University of New York

National School of Development, Peking University,
Tuesday November 4, 2014

Outline of this talk

- The volatility surface: Stylized facts
- A remarkable monofractal scaling property of historical volatility
- Fractional Brownian motion (fBm)
- The Rough Fractional Stochastic Volatility (RFSV) model
- The Rough Bergomi (rBergomi) model
- Tentative numerical results

The implied volatility smile

- The implied volatility $\sigma_{\text{BS}}(k,\tau)$ of an option (with log-moneyness k and time to expiration τ) is the value of the volatility parameter in the Black-Scholes formula required to match the market price of that option.
- Plotting implied volatility as a function of log-moneyness k generates the *volatility smile*.
- Plotting implied volatility as a function of both k and τ generates the *volatility surface*, explored in detail in, for example, [The Volatility Surface].

The SPX volatility surface as of 15-Sep-2005

- We begin by studying the SPX volatility surface as of the close on September 15, 2005.
  - Next morning is *triple witching* when options and futures set.
- We will plot the volatility smiles, superimposing an SVI fit.
  - SVI stands for "stochastic volatility inspired", a well-known parameterization of the volatility surface.
  - We show in [Gatheral and Jacquier] how to fit SVI to the volatility surface in such a way as to guarantee the absence of static arbitrage.
- We then interpolate the resulting SVI smiles to obtain and plot the whole volatility surface.

The March expiry smile as of 15-Sep-2005

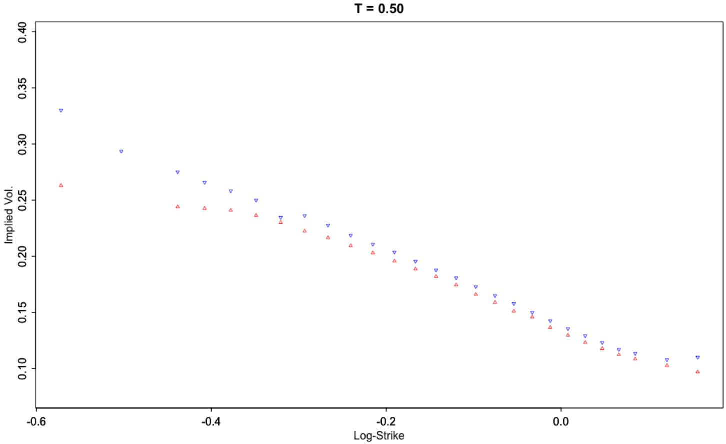

Figure 1: The March expiry smile as of 15-Sep-2005.

SPX volatility smiles as of 15-Sep-2005

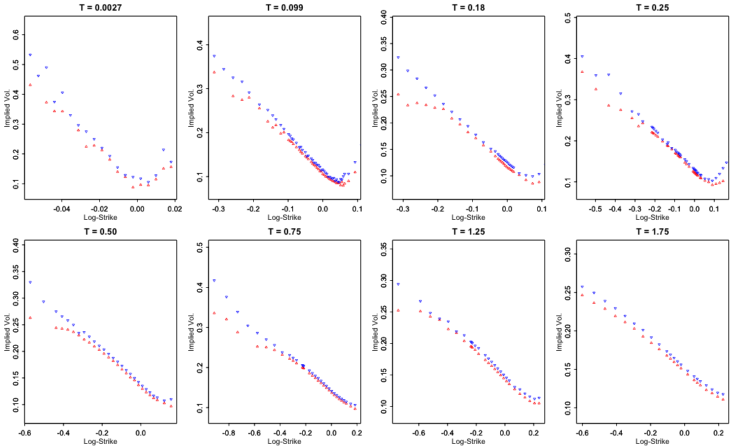

Figure 2: SPX volatility smiles as of 15-Sep-2005.

SPX volatility smiles as of 15-Sep-2005

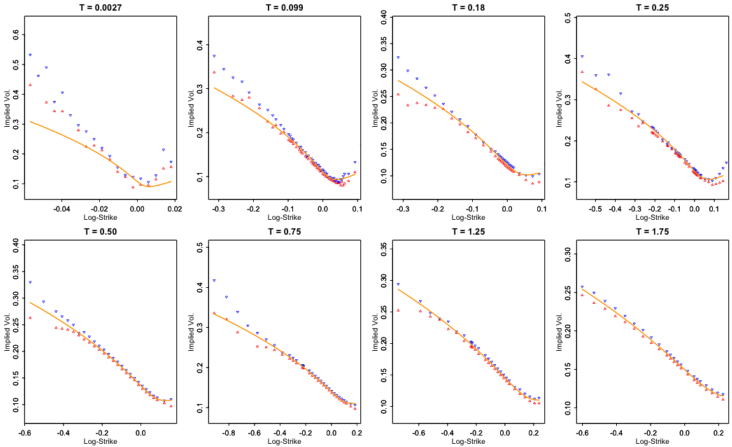

Figure 3: SVI fit superimposed on smiles.

The SPX volatility surface as of 15-Sep-2005

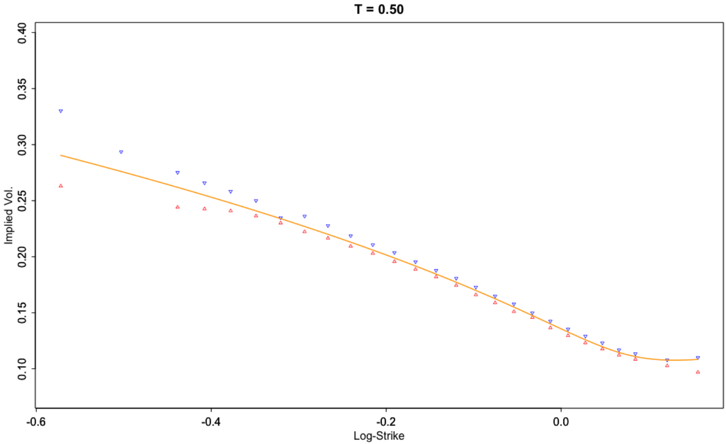

Figure 4: The March expiry smile as of 15-Sep-2005 – the SVI fit looks OK!

The SPX volatility surface as of 15-Sep-2005

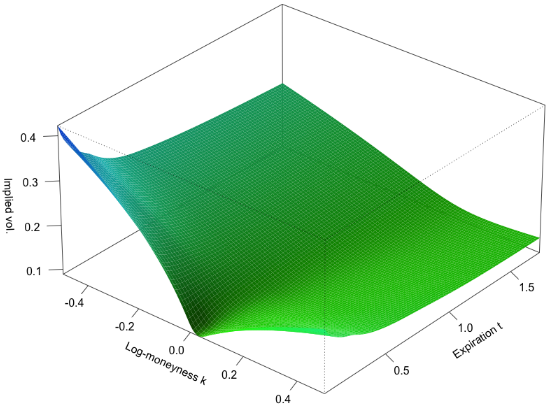

Figure 5: The SPX volatility surface as of 15-Sep-2005 (Figure 3.2 of The Volatility Surface).

Interpreting the 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.
    - Heston for example fits a maximum of two smiles simultaneously.
    - SABR can only fit one smile at a time.

Term structure of at-the-money skew

- What really distinguishes between models is how the generated smile depends on time to expiration.
  - In particular, their predictions for the term structure of ATM volatility skew defined as

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

Term structure of SPX ATM skew as of 15-Sep-2005

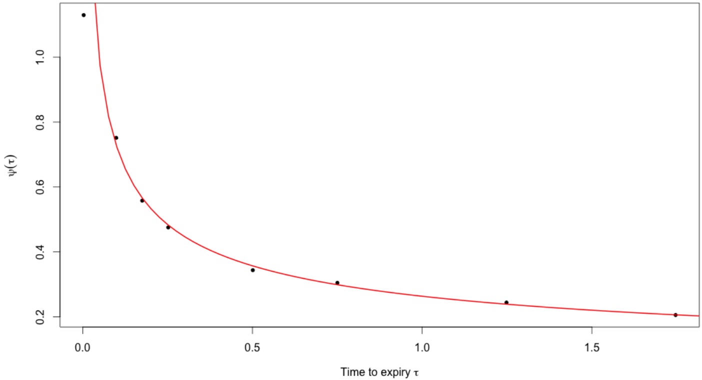

Figure 6: Term structure of ATM skew as of 15-Sep-2005, with power law fit $\tau^{-0.44}$ superimposed in red.

SPX volatility surfaces from 2005 to 2011

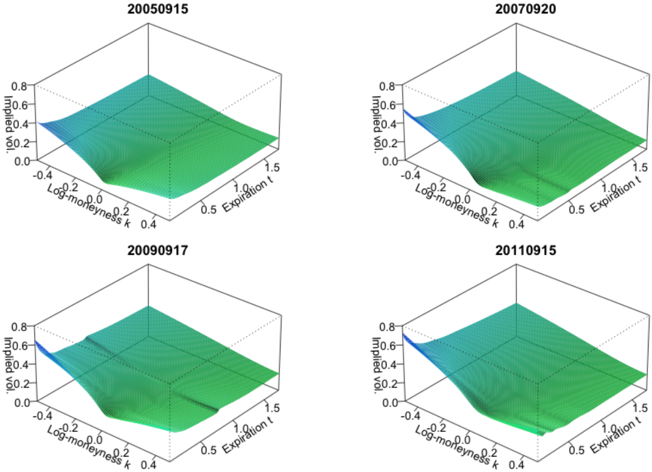

Figure 7: SPX volatility surfaces over the years as of the close before September SQ.

Observations

- We note that although the levels and orientations of the volatility surfaces change over time, their rough shape stays very much the same.

- Let's now look at the term structure of ATM skew on these dates.

Term structure of SPX ATM skew as over the years

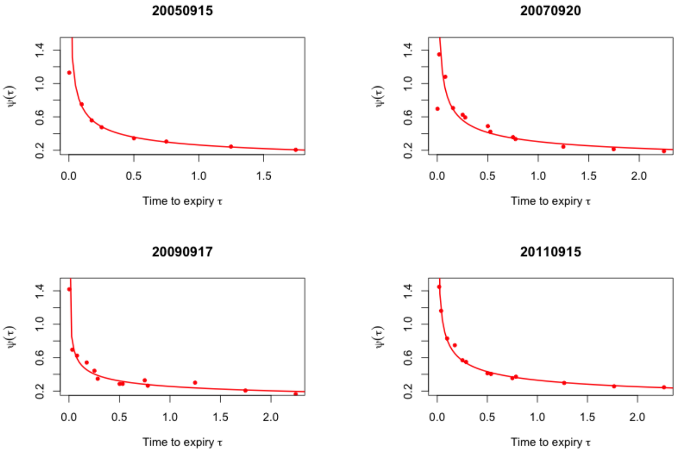

Figure 8: SPX ATM skew over the years as of the close before September SQ. Power-laws fits are superimposed.

Conclusion

- The shape of the volatility surface seems to be more-or-less stable.
  - It's then natural to look for a time-homogeneous model.
- The term structure of ATM volatility skew

$$\psi(\tau) \sim \frac{1}{\tau^{\alpha}}$$

with $\alpha \in (0.3, 0.5)$.

Motivation for Rough Volatility I: Better fitting stochastic volatility models

- Conventional stochastic volatility models generate volatility surfaces that are inconsistent with the observed volatility surface.
  - In stochastic volatility models, the ATM volatility skew is constant for short dates and inversely proportional to T for long dates.
  - Empirically, we find that the term structure of ATM skew is proportional to $1/T^{\alpha}$ for some $0 < \alpha < 1/2$ over a very wide range of expirations.
- The conventional solution is to introduce more volatility factors, as for example in the DMR and Bergomi models.
- One could imagine the power-law decay of ATM skew to be the result of adding (or averaging) many sub-processes, each of which is characteristic of a trading style with a particular time horizon.

Fitting the term structure of ATM skew

- According to (3.21) of [The Volatility Surface], the term structure of ATM skew in a conventional one-factor stochastic volatility model is roughly proportional to

$$\psi(\kappa, \tau) := \frac{1}{\kappa\,\tau} \left\{ 1 - \frac{1 - e^{-\kappa\,\tau}}{\kappa\,\tau} \right\}.$$

- In Figure 9, we show that this function cannot fit the empirically observed term structure of ATM skew but that adding another such term (as a proxy for adding another factor) generates an excellent fit.

Empirical SPX ATM skew term structure with fits

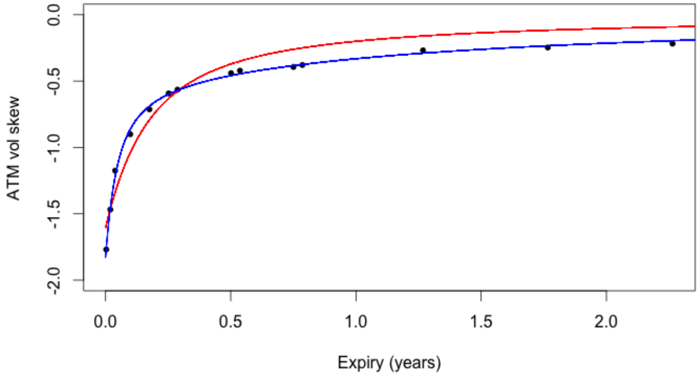

Figure 9: The points in black are SPX ATM skews as of Sep 15, 2011. The red line is the best fit of $A\,\psi(\kappa,\tau)$. The blue line is the best fit of $A_1\,\psi(\kappa_1,\tau)+A_2\,\psi(\kappa_2,\tau)$.

Bergomi Guyon

- Define the forward variance curve $\xi_t(u) = \mathbb{E}\left[v_u \right| \mathcal{F}_t]$.
- According to [Bergomi and Guyon], in the context of a variance curve model, implied volatility may be expanded as

$$\sigma_{\rm BS}(k,T) = \sigma_0(T) + \sqrt{\frac{w}{T}}\,\frac{1}{2\,w^2}\,C^{x\,\xi}\,k + O(\eta^2) \qquad (1)$$

where η is volatility of volatility, $w = \int_0^T \xi_0(s)\,ds$ is total variance to expiration T, and

$$C^{x\,\xi} = \int_0^T dt \int_t^T du\,\frac{\mathbb{E}\left[dx_t\,d\xi_t(u)\right]}{dt}. \qquad (2)$$

  - Thus, given a stochastic model, defined in terms of an SDE, we can easily (at least in principle) compute this smile approximation.

Connecting the time series with options prices

- Suppose for a moment that the pricing measure $\mathbb{Q}$ is the same as the historical (or physical) measure $\mathbb{P}$.
- Then equation (2) also connects the prices of options with statistics of the historical time series of volatility.

ATM volatility and the autocorrelation of volatility

- We may write $\xi_t(u) \approx \beta\, \xi_t(t) + \epsilon$ where $\epsilon \perp \xi_t(t)$ and

$$\beta = \frac{\text{cov}(\xi_t(u)\, \xi_t(t))}{\text{var}(\xi_t(t))} = \frac{\text{cov}(v_u, v_t)}{\text{var}(v_t)}$$

which is just the variance autocorrelation $\rho_v(u-t)$.

- Then

$$C^{x\,\xi} \approx \mathbb{E}\left[\frac{\mathbb{E}\left[dx_t\, d\xi_t(t)\right]}{dt}\right] \int_0^T dt \int_t^T du\, \rho_v(u-t).$$

- Thus, the ATM volatility skew

$$\psi(T) := \partial_k \sigma_{\rm BS}(k,T)|_{k=0} \sim \frac{1}{T^2} \int_0^T dt \int_t^T du\, \rho_v(u-t)$$

  - The term structure of ATM skew thus reflects the (historical) variance autocorrelation function, which may be estimated from the volatility time series.

The Bergomi model

- The n-factor Bergomi variance curve model reads:

$$\xi_t(u) = \xi_0(u) \exp\left\{ \sum_{i=1}^{n} \eta_i \int_0^t e^{-\kappa_i (t-s)} \, dW_s^{(i)} + \text{ drift } \right\}. \tag{3}$$

- To achieve a decent fit to the observed volatility surface, and to control the forward smile, we need at least two factors.
  - In the two-factor case, there are 8 parameters.
- When calibrating, we find that the two-factor Bergomi model is already over-parameterized. Any combination of parameters that gives a roughly $1/\sqrt{T}$ ATM skew fits well enough.
  - Moreover, the calibrated correlations between the Brownian increments $dW_s^{(i)}$ tend to be high.

ATM skew in the Bergomi model

- The Bergomi model generates a term structure of volatility skew $\psi(\tau)$ that is something like

$$\psi(\tau) = \sum_i \frac{1}{\kappa_i\,\tau}\left\{1 - \frac{1 - e^{-\kappa_i\,\tau}}{\kappa_i\,\tau}\right\}.$$

  - This functional form is related to the term structure of the autocorrelation function.
  - Which is in turn driven by the exponential kernel in the exponent in (3).
- The observed $\psi(\tau) \sim \tau^{-\alpha}$ for some α.
- It's tempting to replace the exponential kernels in (3) with a power-law kernel.

Tinkering with the Bergomi model

- This would give a model of the form

$$\xi_t(u) = \xi_0(u)\, \exp\left\{\eta \int_0^t \frac{dW_s}{(t-s)^\gamma} + \text{ drift } \right\}$$

which looks similar to

$$\xi_t(u) = \xi_0(u)\, \exp\left\{\eta\, W_t^H + \text{ drift } \right\}$$

where W_t^H is fractional Brownian motion.

Conversely

- Suppose the true model were something like

$$\xi_t(u) = \xi_0(u)\,\exp\left\{\eta \int_0^t \frac{dW_s}{(t-s)^\gamma} + \text{ drift }\right\}$$

- Then, using a discrete Laplace transform, we could approximate the kernel as

$$(t-s)^{-\gamma} \approx \sum_{i=1}^{n} \alpha_i\, e^{-\kappa_i\,(t-s)}$$

  for some coefficients α_i.
  - Then we would have the Bergomi model back (but with all Brownians perfectly correlated).

Power-laws from averaging: A toy example

The following example, adapted from [Comte and Renault], illustrates how power-law behavior can emerge from the averaging of short memory processes.

- Consider the following OU process ($X_t = \log \sigma_t$ say) indexed by κ:

$$X_t(\kappa) = \int_0^t e^{-\kappa\,(t-s)}\, dW_s.$$

  Then $X_t \sim N(0, \Sigma(\kappa)^2)$ with $\Sigma(\kappa)^2 = \int_0^t e^{-2\,\kappa\,(t-s)}\, ds$.

- Consider a multiplicity of such processes with gamma-distributed κ. Explicitly,

$$p_\Gamma(\kappa) = \frac{\kappa^{\alpha-1}\, e^{-\kappa/\theta}}{\theta^\alpha\, \Gamma(\alpha)}$$

  for some $\alpha > 0$ and $\theta > 0$.

- Then, the average $\bar{X} \sim N(0, \bar{\Sigma}^2)$ with

$$\bar{\Sigma}^2 = \int_0^\infty p_\Gamma(\kappa) \int_0^t e^{-2\,\kappa\,(t-s)}\,d\kappa\,ds = \int_0^t \frac{1}{[1 + 2\,\theta\,(t-s)]^\alpha}\,ds$$

and

$$\bar{X}_t = \int_0^t \frac{dW_s}{[1 + \theta\,(t-s)]^{\alpha/2}}.$$

- Thus, averaging short memory volatility processes (with exponential kernels) over different timescales can generate a volatility process with a power-law kernel

Motivation for Rough Volatility II: Power-law scaling of the volatility process

- The Oxford-Man Institute of Quantitative Finance makes historical realized variance (RV) estimates freely available at http://realized.oxford-man.ox.ac.uk. These estimates are updated daily.
- Using daily RV estimates as proxies for instantaneous variance, we may investigate the time series properties of v_t empirically.

SPX realized variance from 2000 to 2014

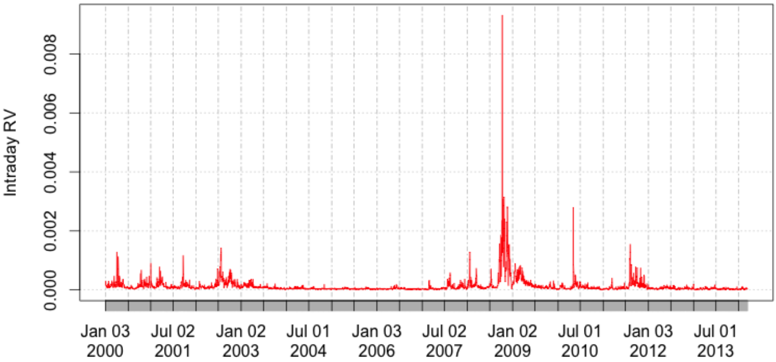

Figure 10: KRV estimates of SPX realized variance from 2000 to 2014.

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[1]:

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

- For example

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

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

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

Scaling of $m(q, \Delta)$ with lag Δ

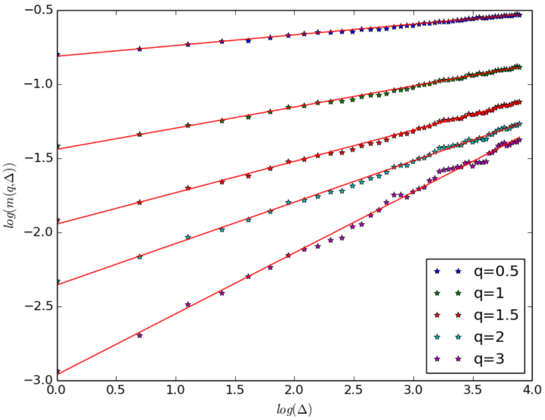

Figure 11: $\log m(q, \Delta)$ as a function of $\log \Delta$, SPX.

Monofractal scaling result

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

$$\zeta_q = q\,H$$

with $H \approx 0.14$.

  - Note however that H does vary over time, in a narrow range.
  - 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.

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

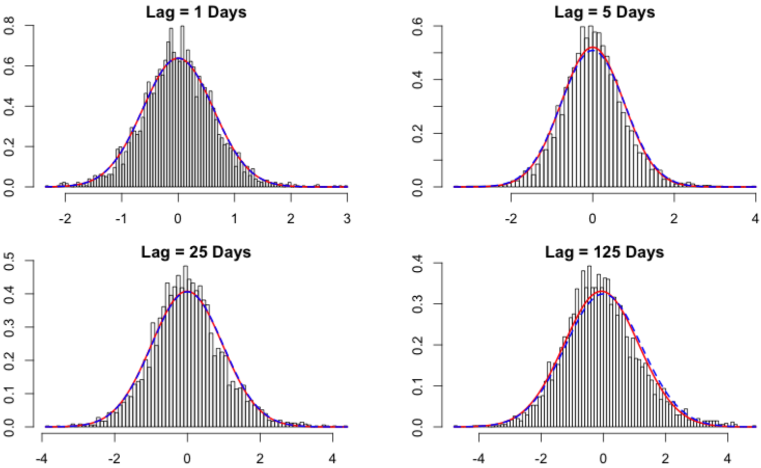

Figure 12: Histograms of $(\log \sigma_{t+\Delta} - \log \sigma_t)$ for various lags Δ; normal fit in red; $\Delta = 1$ normal fit scaled by $\Delta^{0.14}$ in blue.

Estimated H for all indices

Repeating this analysis for all 21 indices in the Oxford-Man dataset yields:

Index	$\zeta_{0.5}/0.5$	ζ_1	$\zeta_{1.5}/1.5$	$\zeta_2/2$	$\zeta_3/3$
SPX2.rv	0.128	0.126	0.125	0.124	0.124
FTSE2.rv	0.132	0.132	0.132	0.131	0.127
N2252.rv	0.131	0.131	0.132	0.132	0.133
GDAXI2.rv	0.141	0.139	0.138	0.136	0.132
RUT2.rv	0.117	0.115	0.113	0.111	0.108
AORD2.rv	0.072	0.073	0.074	0.075	0.077
DJI2.rv	0.117	0.116	0.115	0.114	0.113
IXIC2.rv	0.131	0.133	0.134	0.135	0.137
FCHI2.rv	0.143	0.143	0.142	0.141	0.138
HSI2.rv	0.079	0.079	0.079	0.080	0.082
KS11.rv	0.133	0.133	0.134	0.134	0.132
AEX.rv	0.145	0.147	0.149	0.149	0.149
SSMI.rv	0.149	0.153	0.156	0.158	0.158
IBEX2.rv	0.138	0.138	0.137	0.136	0.133
NSEI.rv	0.119	0.117	0.114	0.111	0.102
MXX.rv	0.077	0.077	0.076	0.075	0.071
BVSP.rv	0.118	0.118	0.119	0.120	0.120
GSPTSE.rv	0.106	0.104	0.103	0.102	0.101
STOXX50E.rv	0.139	0.135	0.130	0.123	0.101
FTSTI.rv	0.111	0.112	0.113	0.113	0.112
FTSEMIB.rv	0.130	0.132	0.133	0.134	0.134

Table 1: Estimates of ζ_q for all indices in the Oxford-Man dataset.

Universality?

- [Gatheral, Jaisson and Rosenbaum] compute daily realized variance estimates over one hour windows for SPX and NASDAQ, finding similar scaling relationships.
- We have also checked that Gold and Crude Oil futures give similar results.

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:

$$\log \sigma_{t+\Delta} - \log \sigma_t = \nu \left(W^H_{t+\Delta} - W^H_t \right) \tag{4}$$

  where W^H is fractional Brownian motion.
- In our paper [Gatheral, Jaisson and Rosenbaum], we refer to a stationary version of (4) 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

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

  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.

Representations of fBm

There are infinitely many possible representations of fBm in terms of Brownian motion. For example, with $\gamma = \frac{1}{2} - H$,

Mandelbrot-Van Ness

$$W_t^H = C_H \left\{ \int_{-\infty}^{t} \frac{dW_s}{(t-s)^\gamma} - \int_{-\infty}^{0} \frac{dW_s}{(-s)^\gamma} \right\}.$$

where the choice

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

ensures that

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

Another representation of fBm

Define

$$K_H(t,s) = C_H \, F\left(\gamma, -\gamma, 1-\gamma, 1-\frac{t}{s}\right) \frac{1}{(t-s)^{\gamma}}.$$

where $F(\cdot)$ is Gauss's hypergeometric function. Then, fBm can also be represented as:

Molchan-Golosov

$$W_t^H = \int_0^t K_H(t,s)\, dW_s.$$

- The Mandelbrot-Van Ness representation uses the entire history of the Brownian motion $\{W_s; s \leq t\}$.
- The Molchan-Golosov representation uses only the history of the Brownian motion from time 0.

Why "fractional"?

Denote the differentiation operator $\frac{d}{dt}$ by D. Then

$$D^{-1}f(t) = \int_0^t f(s)\, ds.$$

The Cauchy formula for repeated integration gives for any integer $n > 0$,

$$D^{-n}f(t) = \int_0^t \frac{1}{n!}\,(t-s)^{n-1}\, f(s)\, ds.$$

The generalization of this formula to real ν gives the definition of the fractional integral:

$$D^{-\nu}f(t) = \frac{1}{\Gamma(\nu)}\int_0^t (t-s)^{\nu-1}\, f(s)\, ds.$$

Note in particular that $D^0 f(t) = f(t)$.

Comte and Renault: FSV model

- [Comte and Renault] were perhaps the first to model volatility using fractional Brownian motion.
- In their fractional stochastic volatility (FSV) model,

$$\begin{aligned} \frac{dS_t}{S_t} &= \sigma_t \, dZ_t \\ d \log \sigma_t &= -\alpha \, (\log \sigma_t - \theta) \, dt + \gamma \, d\hat{W}_t^H \end{aligned} \tag{5}$$

with

$$\hat{W}_t^H = \int_0^t \frac{(t-s)^{H-1/2}}{\Gamma(H+1/2)} \, dW_s, \quad 1/2 \leq H < 1$$

and $\mathbb{E}\left[dW_t \, dZ_t\right] = \rho \, dt$.

- The FSV model is a generalization of the Hull-White stochastic volatility model.

Integral formulation

Solving (5) formally gives

$$\sigma_t = \exp\left\{\theta + e^{-\alpha\, t}\,(\log\sigma_0 - \theta) + \gamma\int_0^t e^{-\alpha\,(t-s)}\, d\hat{W}_s^H\right\}. \quad (6)$$

- $H > 1/2$ to ensure long-memory.
- Stationarity is achieved with the exponential kernel $e^{-\alpha\,(t-s)}$ at the cost of introducing an explicit timescale α^{-1}.

RFSV and FSV

- The model (4):

$$\log \sigma_{t+\Delta} - \log \sigma_t = \nu \left(W^H_{t+\Delta} - W^H_t \right) \tag{7}$$

is not stationary.

  - Stationarity is desirable both for mathematical tractability and also to ensure reasonableness of the model at very large times.

- The RFSV model (the stationary version of (4)) is formally identical to the FSV model. Except that
  - $H < 1/2$ in RFSV vs $H > 1/2$ in FSV.
  - $\alpha\, T \gg 1$ in RFSV vs $\alpha\, T \sim 1$ in FSV

where T is a typical timescale of interest.

FSV and long memory

- Why did [Comte and Renault] choose $H > 1/2$?
  - Because it has been a widely-accepted stylized fact that the volatility time series exhibits long memory.
- In this technical sense, *long memory* means that the autocorrelation function of volatility decays as a power-law.
- One of the influential papers that established this was [Andersen et al.] which estimated the degree d of fractional integration from daily realized variance data for the 30 DJIA stocks.
  - Using the GPH estimator, they found d around 0.35 which implies that the ACF $\rho(\tau) \sim \tau^{2\,d-1} = \tau^{-0.3}$ as $\tau \to \infty$.
- But every statistical estimator assumes the validity of some underlying model!

Correlogram and test of scaling

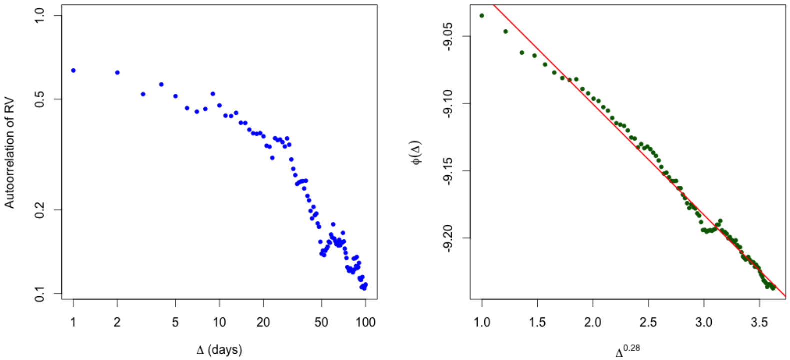

Figure 13: The LH plot is a conventional correlogram of RV; the RH plot is of $\phi(\Delta) := \langle \log\left(\text{cov}(\sigma_{t+\Delta}, \sigma_t) + \langle \sigma_t \rangle^2\right)\rangle$ vs Δ^{2H} with $H = 0.14$. The RH plot again supports the scaling relationship $m(2, \Delta) \propto \Delta^{2H}$.

Heuristic derivation of autocorrelation function

We assume that $\sigma_t = \bar\sigma_t \, e^{\eta \, W^H_t}$. Then

$$\begin{aligned} & \operatorname{cov}\left[\sigma_t, \sigma_{t+\Delta}\right] \\ = \quad & \bar\sigma_t \, \bar\sigma_{t+\Delta} \, \left[\exp\left\{\frac{\eta^2}{2}\,\left(t^{2\,H} + (t+\Delta)^{2\,H} - \Delta^{2\,H}\right)\right\} - 1\right] \\ \sim \quad & \bar\sigma_t \, \bar\sigma_{t+\Delta} \, \exp\left\{\frac{\eta^2}{2}\,\left(t^{2\,H} + (t+\Delta)^{2\,H} - \Delta^{2\,H}\right)\right\} \text{ as } t \to \infty. \end{aligned}$$

Similarly,

$$\operatorname{var}\left[\sigma_t\right] \sim \bar\sigma_t^2 \, \exp\left\{\eta^2 \, t^{2\,H}\right\}.$$

Thus

$$\rho(\Delta) = \frac{\operatorname{cov}\left[\sigma_t, \sigma_{t+\Delta}\right]}{\sqrt{\operatorname{var}\left[\sigma_t\right]\,\operatorname{var}\left[\sigma_{t+\Delta}\right]}} \sim \exp\left\{-\frac{\eta^2}{2}\,\Delta^{2\,H}\right\}.$$

Model vs empirical autocorrelation functions

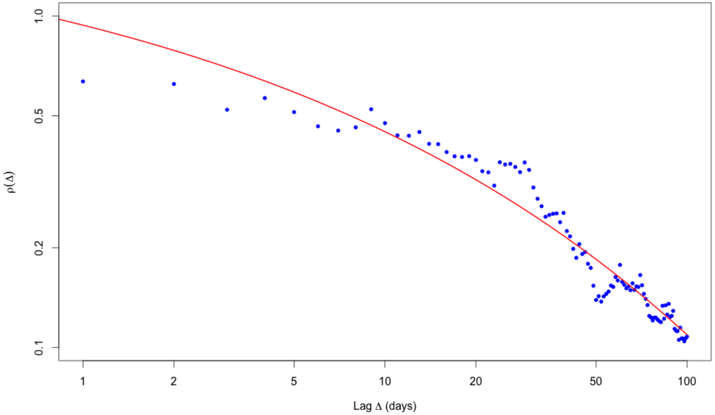

Figure 14: Here we superimpose the predicted functional form of $\rho(\Delta)$ (in red) on the empirical curve (in blue).

Volatility is not long memory

- It's clear from Figures 13 and 14 that volatility is not long memory.
- Moreover, the RFSV model reproduces the observed autocorrelation function very closely.
- [Gatheral, Jaisson and Rosenbaum] further simulate volatility in the RFSV model and apply estimators such as GPH to the simulated data.
- Real data and simulated data generate very similar plots and similar estimates of the long memory parameter to those found in the prior literature.
- The RSFV model does not have the long memory property.
    - Classical estimation procedures seem to identify spurious long memory of volatility.

FSV covariance computation

We can compute the FSV autocovariance function explicitly:

Define $y_t = \log \sigma_t$. We have

$$\mathrm{cov}(y_t, y_{t+\Delta}) \propto \int_{-\infty}^{0} e^{\alpha\, s}\, ds \int_{-\infty}^{\Delta} e^{\alpha\,(s'-\Delta)}\, ds'\, |s - s'|^{2\,H-2}.$$

Then $\mathbb{E}\left[(y_{t+\Delta} - y_t)^2\right] = 2\,\mathrm{var}[y_t] - 2\,\mathrm{cov}(y_t, y_{t+\Delta})$ where

$$\begin{aligned} &\mathrm{cov}(y_t, y_{t+\Delta}) \\ &\propto \frac{e^{-k\Delta}}{2k^{2H}} \int_0^{k\Delta} \frac{e^u\, du}{u^{2-2H}} + \frac{e^{-k\Delta}}{2k^{2H}} \Gamma(2H-1) + \frac{e^{k\Delta}}{2k^{2H}} \int_{k\Delta}^{+\infty} \frac{e^{-u}\, du}{u^{2-2H}}. \end{aligned}$$

Incompatibility of FSV with realized variance (RV) data

- In Figure 15, we demonstrate graphically that long memory volatility models such as FSV with $H > 1/2$ are not compatible with the RV data.
- In the FSV model, the autocorrelation function $\rho(\Delta) \propto \Delta^{2H-2}$. Then, for long memory, we must have $1/2 < H < 1$.
    - For $\Delta \gg 1/\alpha$, stationarity kicks in and $m(2,\Delta)$ tends to a constant as $\Delta \to \infty$.
    - For $\Delta \ll 1/\alpha$, the exponential decay in (6) is not significant and $m(2,\Delta) \propto \Delta^{2H}$.

Incompatibility of FSV with RV data

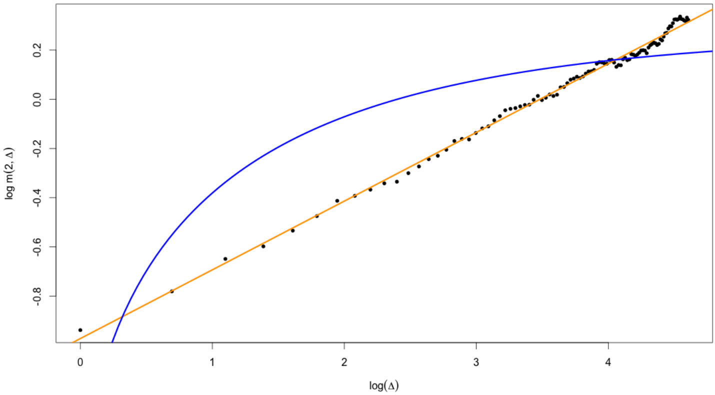

Figure 15: Black points are empirical estimates of $m(2, \Delta)$; the blue line is the FSV model with $\alpha = 0.5$ and $H = 0.53$; the orange line is the RFSV model with $\alpha = 0$ and $H = 0.14$.

Does simulated RSFV data look real?

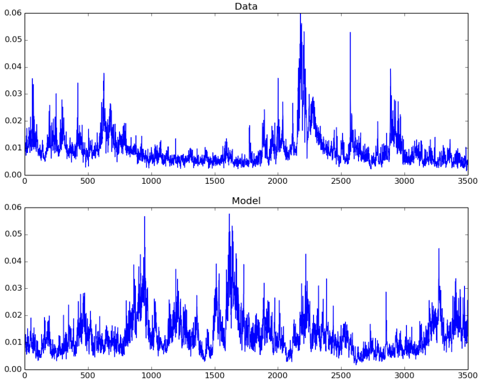

Figure 16: Volatility of SPX (above) and of the RFSV model (below).

Remarks on the comparison

- The simulated and actual graphs look very alike.
  - 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, [Bacry and Muzy].
  - In particular, their Multifractal Random Walk (MRW) can be understood as the limit of the RSFV model as $H \to 0$.

Pricing under rough volatility

The foregoing behavior suggest the following model for volatility under the real (or historical or physical) measure $\mathbb{P}$:

$$y_t := \log \sigma_t = \nu \, W_t^H.$$

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

$$W_t^H = C_H \left\{ \int_{-\infty}^{t} \frac{dW_s^{\mathbb{P}}}{(t-s)^\gamma} - \int_{-\infty}^{0} \frac{dW_s^{\mathbb{P}}}{(-s)^\gamma} \right\}$$

where the choice

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

ensures that

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

Pricing under rough volatility

Then

$$\begin{aligned} & y_u - y_t \\ = \quad & \nu\, C_H \left\{ \int_t^u \frac{1}{(u-s)^\gamma}\, dW_s^{\mathbb{P}} + \int_{-\infty}^t \left[\frac{1}{(u-s)^\gamma} - \frac{1}{(t-s)^\gamma} \right] dW_s^{\mathbb{P}} \right\} \\ =: \quad & \nu\, C_H \left[M_t(u) + Z_t(u) \right]. \end{aligned} \tag{8}$$

- 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.
- In terms of $v_t = \sigma_t^2$,

$$\log v_u - \log v_t = 2\,\nu\, C_H \left[M_t(u) + Z_t(u) \right]. \tag{9}$$

- 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 forward variance curve

Taking the exponential of (9) gives

$$v_u = v_t \exp\left\{2\,\nu\,C_H\,\left[M_t(u) + Z_t(u)\right]\right\}$$

Ignoring the difference between $\mathbb{P}$ and $\mathbb{Q}$, and computing the conditional expectation gives

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}\left[v_u \middle| \mathcal{F}_t\right] &= \xi_t(u)\\ &= v_t\,\exp\left\{2\,\nu\,C_H\,Z_t(u)\right\}\,\mathbb{E}\left[\exp\left\{2\,\nu\,C_H\,M_t(u)\right\}\middle|\mathcal{F}_t\right]\end{aligned}$$

where (by definition) $\xi_t(u)$ is the forward variance curve at time t.

- The $Z_t(u)$ are encoded in the forward variance curve $\xi_t(u)$!

The rough Bergomi model

Define $\eta = 2\,\nu\,C_H/\sqrt{2\,H}$. Rewriting gives

$$v_u = \xi_t(u)\,\mathcal{E}\left(\eta\,\sqrt{2\,H}\,\int_t^u \frac{dW_s}{(u-s)^\gamma}\right) \tag{10}$$

where $\mathcal{E}(\cdot)$ denotes the stochastic exponential.

- We could call this a *rough Bergomi* or *rBergomi* model.

Features of the rough Bergomi model

- The forward variance curve

$$\xi_u(t) = \mathbb{E}\left[v_u \middle| \mathcal{F}_t \right] = v_t \, \exp\left\{ \eta \sqrt{2H} \, Z_t(u) + \frac{1}{2} \eta^2 \, (u-t)^{2H} \right\}.$$

  depends on the historical path $\{W_s, s < t\}$ of the Brownian motion since inception ($s = -\infty$ say).
- The rough Bergomi model is non-Markovian:

$$\mathbb{E}\left[v_u \middle| \mathcal{F}_t \right] \neq \mathbb{E}[v_u | v_t].$$

- However, given the (infinite) state vector $\xi_t(u)$, which can in principle be computed from option prices, the dynamics of the model are well-determined.

Re-interpretation of the conventional Bergomi model

- A conventional n-factor Bergomi model is not self-consistent for an arbitrary choice of the initial forward variance curve $\xi_t(u)$.
  - $\xi_t(u) = \mathbb{E}\left[v_u \middle| \mathcal{F}_t \right]$ should be consistent with the assumed dynamics.
- Viewed from the perspective of the fractional Bergomi model however:
  - The initial curve $\xi_t(u)$ reflects the history $\{W_s; s < t\}$ of the driving Brownian motion up to time t.
  - The exponential kernels in the exponent of (3) approximate more realistic power-law kernels.
- The conventional two-factor Bergomi model is then justified in practice as a tractable Markovian engineering approximation to a more realistic fractional Bergomi model.

An fBm-like Volterra process

Define

$$\tilde{W}_t^H = \sqrt{2\,H} \int_0^t \frac{1}{(u-s)^\gamma}\, dW_s^{\mathbb{P}}.$$

- Note that $\mathrm{var}[\tilde{W}_t^H] = t^{2\,H}$ just like fBm.
- However, the covariance $\mathbb{E}\left[\tilde{W}_t^H\,\tilde{W}_s^H\right]$ is very different.

The rBergomi model (10) may be rewritten in terms of this Volterra process (as of time 0) as:

$$v_t = \xi_0(t)\,\mathcal{E}\left(\eta\,\tilde{W}_t^H\right).$$

The stock price process

- The observed anticorrelation between price moves and volatility moves may be modeled naturally by anticorrelating the Brownian motion W that drives the volatility process with the Brownian motion driving the price process.
- Then

$$\frac{dS_t}{S_t} = \sqrt{v_t}\, dZ_t$$

with

$$dZ_t = \rho\, dW_t + \sqrt{1-\rho^2}\, dW_t^{\perp}$$

where ρ is the correlation between volatility moves and price moves.

Simulation of the Bergomi and rSABR models

- First, for each Monte Carlo path, generate the correlated Brownian increments ΔW_t and ΔZ_t.
- The $\tilde{W}_t^H$ may then be constructed by appropriately discretizing the definition

$$\tilde{W}_t^H = \sqrt{2\,H} \int_0^t \frac{dW_s}{(t-s)^\gamma}.$$

Estimating H and η

- We could in principle estimate H either from the term structure of ATM SPX skew, or from the term structure of *ATM* VIX volatilities.
  - Implied volatility of VIX should be “volatility of SPX volatility”!
- Fast calibration of the Bergomi model is work in progress.

SPX smiles in the rBergomi model

- In Figure 17, we show how a rBergomi model simulation is consistent with the SPX option market as of 04-Feb-2010, a day when the ATM volatility term structure happened to be pretty flat.
- rSABR parameters were: $\bar{\sigma} = 0.235$, $\eta = 1.7$, $H = 0.1$, $\rho = -0.85$.
    - Note in particular that we have obtained a good fit to the whole volatility surface using a model with very few parameters!
- In Figure 18, we see that the empirical SPX skews are very consistent with the rBergomi model, in particular with $H = 0.1$.
    - Whatever else is allowed to change, H must be the same under $\mathbb{P}$ and $\mathbb{Q}$ if the model is well-specified.

rBergomi fits to SPX smiles as of 04-Feb-2010

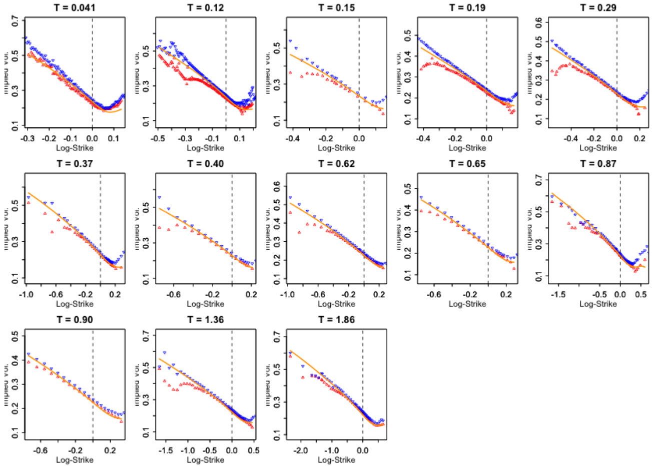

Figure 17: Red and blue points represent bid and offer SPX implied volatilities; orange smiles are from the rBergomi simulation.

Term structure of ATM skew in the rBergomi model

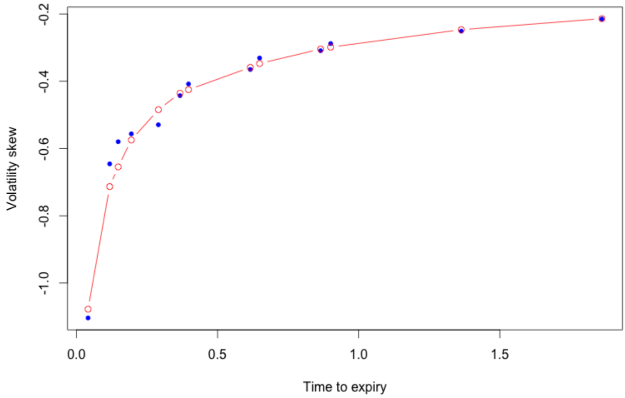

Figure 18: Blue points are empirical skews; the red points are from the rBergomi simulation.

Summary

- We uncovered a remarkable monofractal scaling relationship in historical volatility.
    - A corollary is that volatility is not a long memory process, as widely believed.
- This leads to a natural non-Markovian stochastic volatility model.
- This model fits the observed volatility surface surprisingly well with very few parameters.
- For perhaps the first time, we have a simple consistent model of historical and implied volatility.

References

- Torben G Andersen, Tim Bollerslev, Francis X Diebold, and Heiko Ebens, The distribution of realized stock return volatility, *Journal of Financial Economics* **61** (1) 43–76 (2001).
- Emmanuel Bacry and Jean-François Muzy, Log-infinitely divisible multifractal processes, *Communications in Mathematical Physics* **236**(3) 449–475 (2003).
- Lorenzo Bergomi, Smile dynamics II, *Risk Magazine* 67–73 (October 2005).
- Lorenzo Bergomi and Julien Guyon, The smile in stochastic volatility models, *SSRN* (2011).
- Fabienne Comte and Eric Renault, Long memory in continuous-time stochastic volatility models, *Mathematical Finance* **8** 29–323(1998).
- Jim Gatheral, *The Volatility Surface: A Practitioner's Guide*, Wiley Finance (2006).
- Jim Gatheral and Antoine Jacquier, Arbitrage-free SVI volatility surfaces, *Quantitative Finance*, **14**(1) 59–71 (2014).
- Jim Gatheral, Thibault Jaisson and Mathieu Rosenbaum, Volatility is rough, *Available at SSRN 2509457*, (2014).