QuantMinds International

Intercontinential O2 London, November 17, 2025

Rough volatility workshop

Lecture 1: Econometrics

Jim Gatheral
Department of Mathematics

Baruch COLLEGE
The City University of New York

Outline of Lecture 1

- Shape of the volatility surface
- Scaling of implied volatility smiles
- Monofractal scaling of realized variance
- Estimation of H
- Realized variance forecasting

What is R? (http://cran.r-project.org)

From Wikipedia:

- In computing, R is a programming language and software environment for statistical computing and graphics. It is an implementation of the S programming language with lexical scoping semantics inspired by Scheme.
- R was created by Ross Ihaka and Robert Gentleman at the University of Auckland, New Zealand, and is now developed by the R Development Core Team. It is named partly after the first names of the first two R authors (Robert Gentleman and Ross Ihaka), and partly as a play on the name of S. The R language has become a de facto standard among statisticians for the development of statistical software.
- R is widely used for statistical software development and data analysis. R is part of the GNU project, and its source code is freely available under the GNU General Public License, and pre-compiled binary versions are provided for various operating systems. R uses a command line interface, though several graphical user interfaces are available.

The IPython Notebook (http://ipython.org/notebook.html)

From ipython.org:

The IPython Notebook is a web-based interactive computational environment where you can combine code execution, text, mathematics, plots and rich media into a single document:

The IPython notebook with embedded text, code, math and figures. These notebooks are normal files that can be shared with colleagues, converted to other formats such as HTML or PDF, etc. You can share any publicly available notebook by using the IPython Notebook Viewer service which will render it as a static web page. This makes it easy to give your colleagues a document they can read immediately without having to install anything.

http://nbviewer.ipython.org/github/dboyliao/cookbook-code/blob/master/notebooks/chapter07_stats/08_r.ipynb has instructions on using R with iPython notebook.

The SPX volatility surface as of 15-Feb-2023

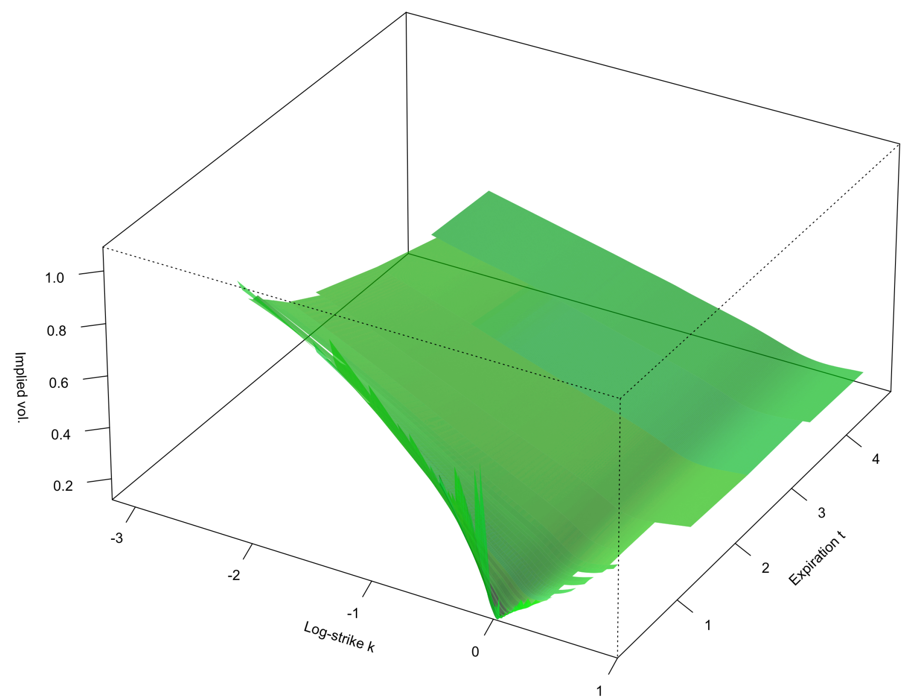


Figure 1. The SPX volatility surface as of 15-Feb-2023 (data from OptionMetrics via WRDS).

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!
- 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

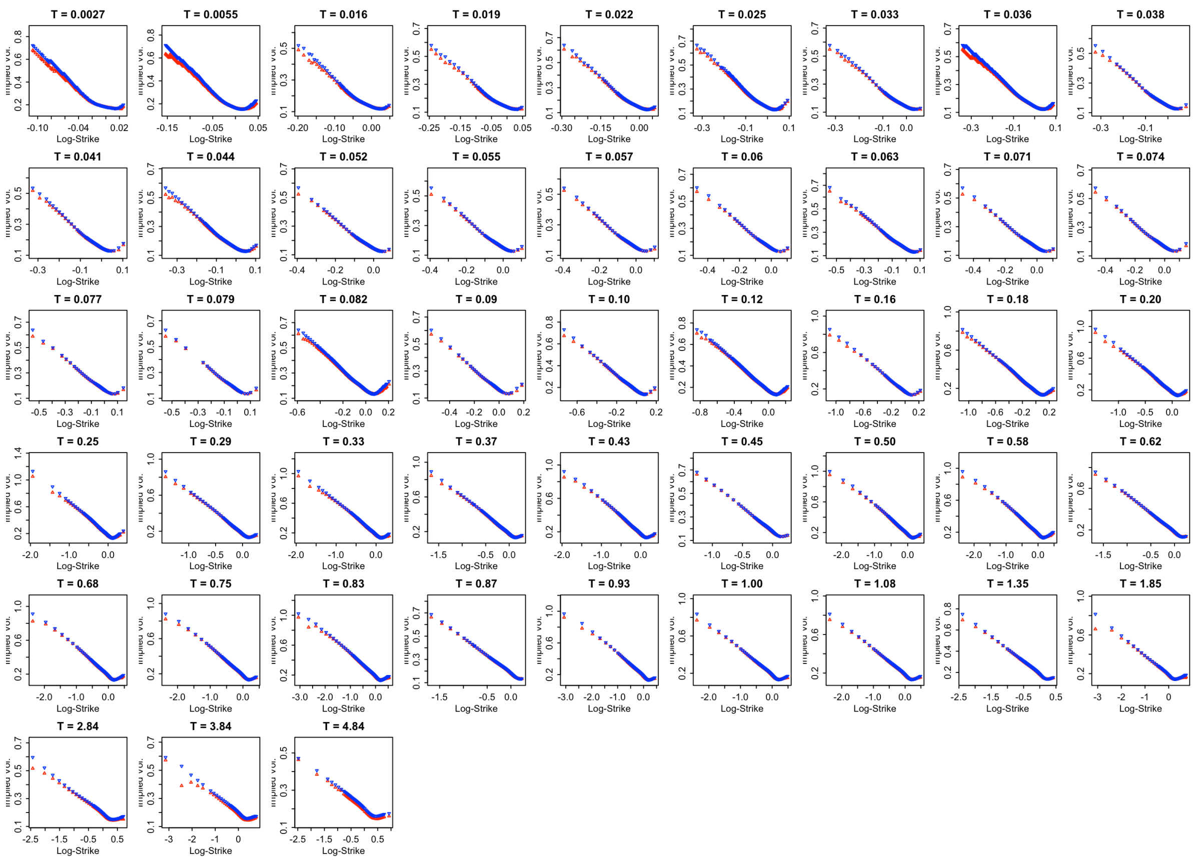


Figure 2. SPX volatility smiles as of 15-Feb-2023.

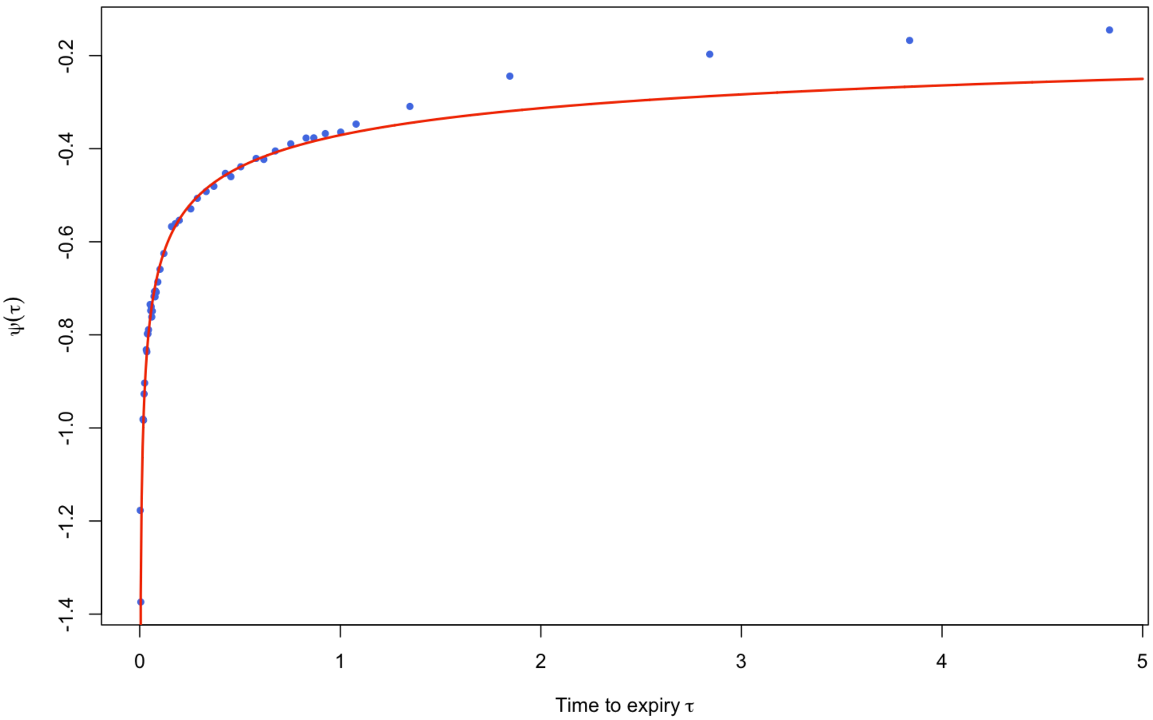
Term structure of at-the-money skew

- 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

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

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



Fugre 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.

Alòs and Fukasawa

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.][1] and subsequently [Fukasawa][8] to generate a short-dated ATM skew of the form

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

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

Rough volatility

- 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

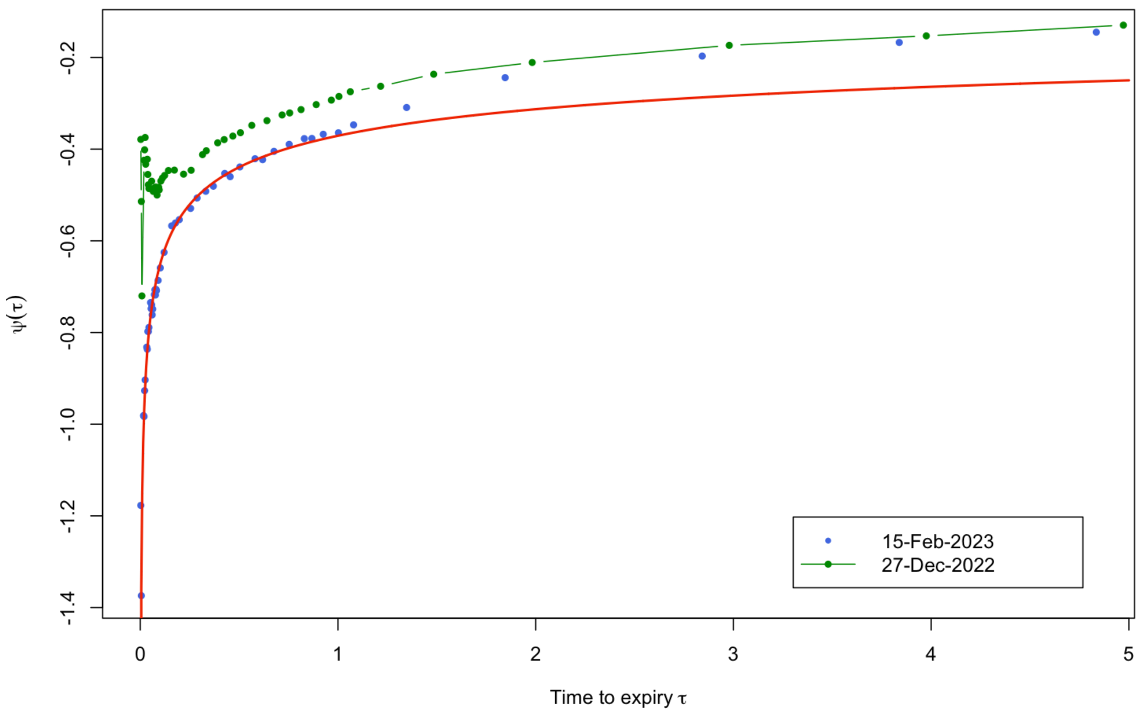


Figure 4. ATM skew term structure on two different dates. On 27-Dec-2022, the skew term structure is not even monotonic!

Total variance plot

- Define the implied total variance $w(k,\tau) := \sigma_{\text{BS}}(k,\tau)^2\,\tau$.
- 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.

Total variance plot as of 27-Dec-2022

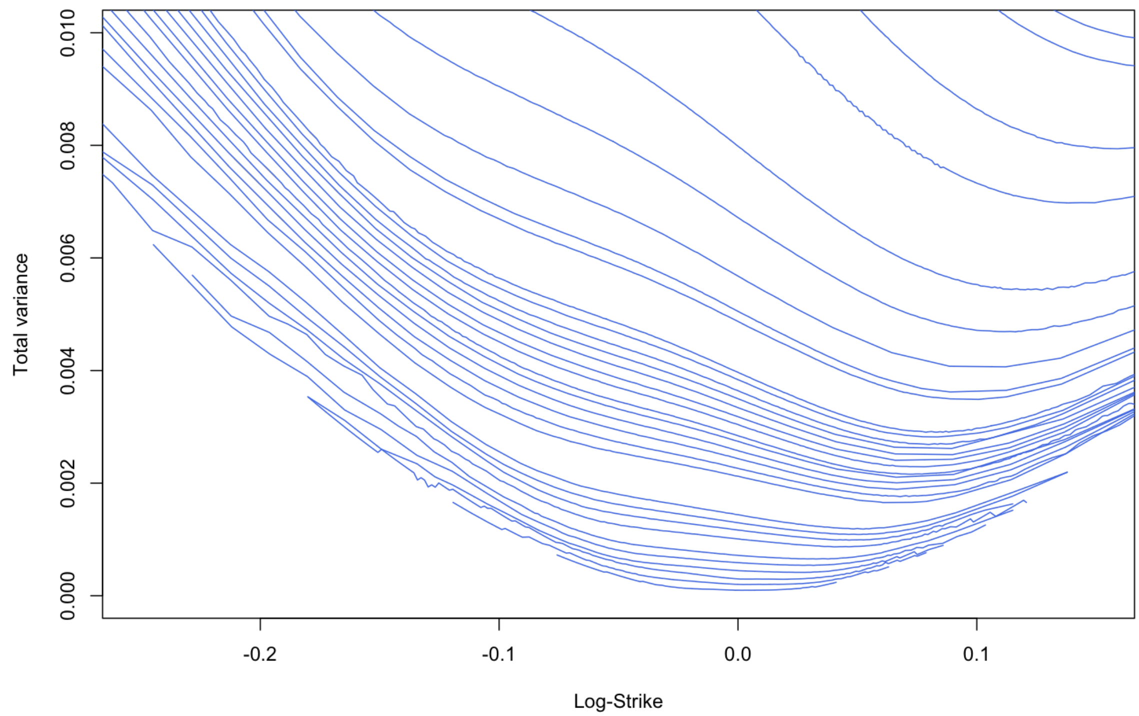


Figure 5. On 27-Dec-2022, no calendar spread arbitrage. However, some individual total variance curves are W-shaped.

Scaling of total variance

- The rough SABR formula of \cite{fukasawa2022rough} 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 6. does suggest close-to-power-law scaling, even in the 27-Dec-2022 case.

Scaling of total variance

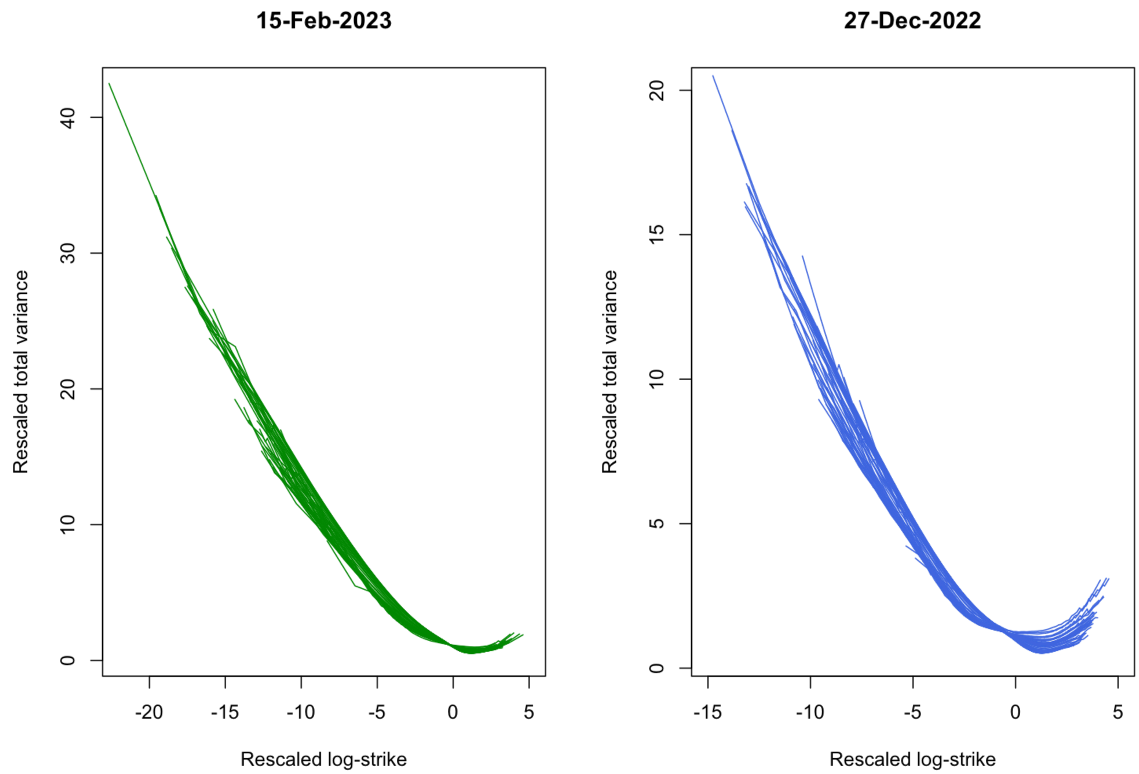


Figure 6. ATM skew term structure on two different dates. On 27-Dec-2022, the skew term structure is not even monotonic!

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.

Set up the environment

In [1]:
```
# Download .zip from GitHub
release_url <- "https://github.com/jgatheral/RoughVolatilityWorkshop2025/raw/main/QM_2025.zi
local_zip <- "QM_2025.zip"

download.file(release_url, local_zip, mode = "wb")
unzip(local_zip)
```

In [2]:
```
# The required packages

library(quantmod)
library(repr)
library(colorspace)
options(repr.plot.width=10,repr.plot.height=7,repr.plot.res=150)
```

```
Loading required package: xts

Loading required package: zoo


Attaching package: ‘zoo’


The following objects are masked from ‘package:base’:

    as.Date, as.Date.numeric


Loading required package: TTR

Registered S3 method overwritten by 'quantmod':
  method            from
  as.zoo.data.frame zoo
```

Set up nice colors

In [3]:
```
my.col <- sequential_hcl(5, palette="Batlow")
bl <- "royalblue"
rd <- "red2"
pk <- "hotpink1"
gr <- "green4"
br <-"brown"
pu <- "purple"
or <- "orange"
```

The time series of realized variance

- We would like to study the time series of instantaneous variance V_t but of course cannot because V_t is latent.
- On the other hand, integrated variance $\frac{1}{\delta}\int_t^{t+\delta} V_s\,ds$ may (in principle) be estimated arbitrarily accurately given enough price data.
  - In practice, market microstructure noise makes estimation harder at very high frequency.
  - Sophisticated estimators of integrated variance have been developed to adjust for market microstructure noise. See Gatheral and Oomen [11] (for example) for details of these.

The Oxford-Man dataset

- 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.
  - Each day, for 31 different indices, all trades and quotes were used to estimate realized (or integrated) variance over the trading day from open to close.
- Using daily RV estimates as proxies for instantaneous variance, we may investigate the time series properties of integrated variance empirically.

In [4]:
```
load("OxfordRV.rData")
names(rv.list)
```

'.AEX' · '.AORD' · '.BFX' · '.BSESN' · '.BVLG' · '.BVSP' · '.DJI' · '.FCHI' · '.FTMIB' · '.FTSE' · '.GDAXI' · '.GSPTSE' · '.HSI' · '.IBEX' · '.IXIC' · '.KS11' · '.KSE' · '.MXX' · '.N225' · '.NSEI' · '.OMXC20' · '.OMXHPI' · '.OMXSPI' · '.OSEAX' · '.RUT' · '.SMSI' · '.SPX' · '.SSEC' · '.SSMI' · '.STI' · '.STOXX50E'

Let's plot SPX realized variance.

In [5]:
```
spx.rk <- rv.list[[".SPX"]]
stoxx.rk <- rv.list[[".STOXX50E"]]
ftse.rk <- rv.list[[".FTSE"]]
```

In [6]:
```
plot(log(spx.rk), main="Log of SPX realized variance",col=rd)
```

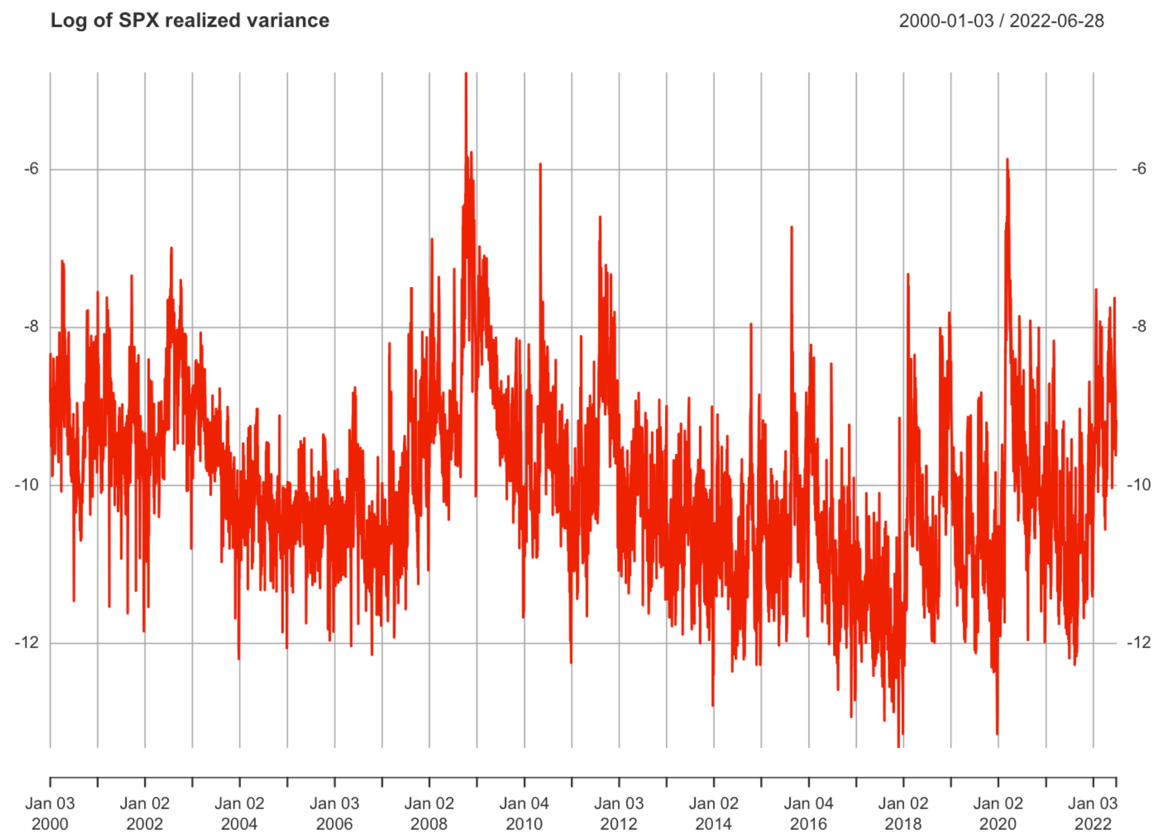


Figure 7: Oxford-Man Log KRV estimates of SPX realized variance from January 2000 to June 2022.

In [7]:
```
print(head(spx.rk))
print(tail(spx.rk))
```

```
                    [,1]
2000-01-03 1.301572e-04
2000-01-04 1.622259e-04
2000-01-05 2.398365e-04
2000-01-06 1.322324e-04
2000-01-07 9.486773e-05
2000-01-10 1.121113e-04
                    [,1]
2022-06-17 2.234366e-04
2022-06-22 1.367900e-04
2022-06-23 1.499486e-04
2022-06-24 6.649679e-05
2022-06-27 9.746335e-05
2022-06-28 1.039309e-04
```

Scaling of the volatility process

For $q \geq 0$, we define the qth sample moment of differences of log-volatility at a given lag Δ.($\langle\cdot\rangle$ denotes the sample average):

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

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.

Scaling of $m(q,\Delta)$ with lag Δ

In [8]:
```
sig <- sqrt(as.numeric(spx.rk))

mq.del.Raw <- function(q,lag){mean(abs(diff(log(sig),lag=lag))^q)}
mq.del <- function(x,q){sapply(x,function(x){mq.del.Raw(q,x)})}

# Plot mq.del(1:100,q) for various q
x <- 1:100
ylab <- expression(paste(log," ",m(q,Delta)))
xlab <- expression(paste(log, " ", Delta))

qVec <- c(.5,1,1.5,2,3)
zeta.q <- numeric(5)
q <- qVec[1]

options(repr.plot.height=7, repr.plot.width=10)
```

In [9]:
```
plot(log(x),log(mq.del(x,q)),pch=20,cex=.5,
         ylab=ylab, xlab=xlab,ylim=c(-3,-.5))
fit.lm <- lm(log(mq.del(x,q)) ~ log(x))
abline(fit.lm, col=my.col[1],lwd=2)
zeta.q[1] <- coef(fit.lm)[2]

for (i in 2:5){
    q <- qVec[i]
    points(log(x),log(mq.del(x,q)),pch=20,cex=.5)
    fit.lm <- lm(log(mq.del(x,q)) ~ log(x))
    abline(fit.lm, col=my.col[i],lwd=2)
    zeta.q[i] <- coef(fit.lm)[2]
    }
 legend("bottomright", c("q = 0.5","q = 1.0","q = 1.5","q = 2.0","q = 3.0"),
        inset=0.05, lty=1, col = my.col)

print(zeta.q)
```

```
[1] 0.08377466 0.16687154 0.24956083 0.33209300 0.49728148
```

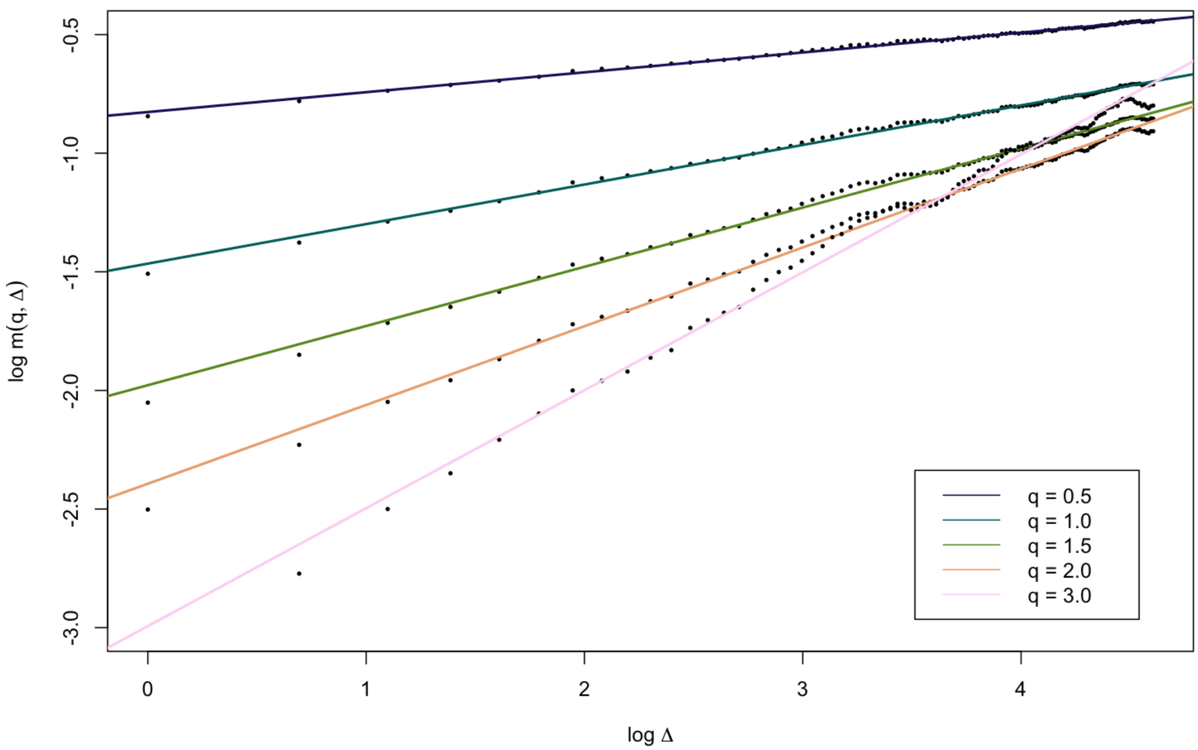


Figure 8: $\log m(q, \Delta)$ as a function of $\log \Delta$, SPX.

Monofractal scaling result

- From the above log-log plot, we see that for each q, $m(q, \Delta) \propto \Delta^{\zeta_q}$.
- How does ζ_q scale with q?

Scaling of ζ_q with q

In [10…

```
plot(qVec,zeta.q,xlab="q",ylab=expression(zeta[q]),pch=20,col=bl,cex=2)
fit.lm <- lm(zeta.q[1:4] ~ qVec[1:4]+0)
abline(fit.lm, col=rd,lwd=2)
(h.est <- coef(fit.lm)[1])
```

qVec[1:4]: 0.16630481565354

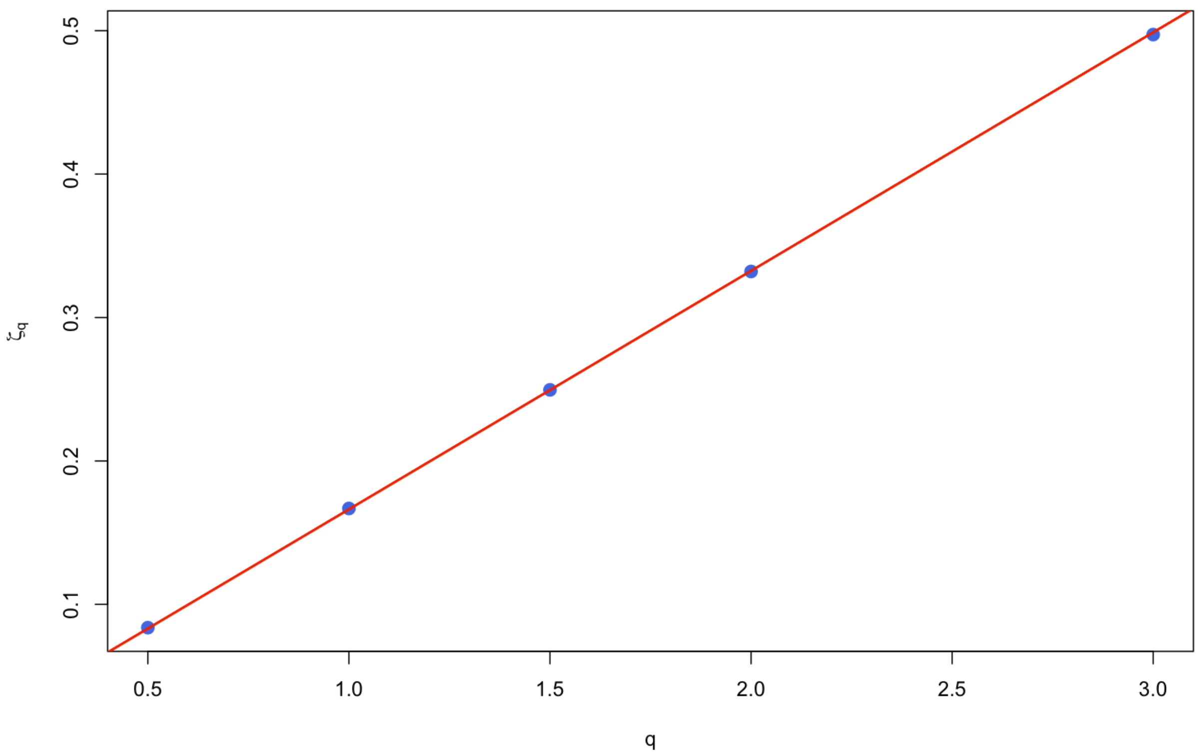


Figure 9: Scaling of ζ_q with q.

We find the monofractal scaling relationship

$$\zeta_q = q\,H$$

with $H \approx 0.166$.

- Note however that H does vary over time, in a narrow range, as we will see later.
- 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 trading day.
  - On the other hand, the time series of realized variance is noisy and this causes our estimate of H to be biased low.

- This scaling property as $\Delta \to 0$ is equivalent to H-Hölder continuity of paths of the volatility.
  - Since $H \ll 1/2$, *volatility is rough*!

Estimated H for all indices

We now repeat this analysis for all 31 indices in the Oxford-Man dataset.

In [11…

```
n <- length(rv.list)
h <- numeric(n) # H is estimated as half of the slope
nu <- numeric(n)

for (i in 1:n){ # Run all the regressions
  v <- rv.list[[i]]
  sig1 <- sqrt(abs(as.numeric(v)))
```

```
  x <- 1:100
  dlsig2 <- function(lag){mean((diff(log(sig1),lag=lag))^2)}
  dlsig2Vec <- function(x){sapply(x,dlsig2)}

  fit.lm <- lm(log(dlsig2Vec(x)) ~ log(x))

  nu[i] <- sqrt(exp(coef(fit.lm)[1]))
  h[i] <- coef(fit.lm)[2]/2

}
```

In [12…

```
(OxfordH <- data.frame(names(rv.list),h.est=h,nu.est=nu))
```

A data.frame: 31 × 3

names.rv.list.	h.est	nu.est
<chr>	<dbl>	<dbl>
.AEX	0.15116126	0.2794888
.AORD	0.11051645	0.3037204
.BFX	0.14397740	0.2590815
.BSESN	0.13365442	0.2906722
.BVLG	0.14173291	0.2370132
.BVSP	0.13744512	0.2932294
.DJI	0.16790232	0.2838851
.FCHI	0.14176726	0.2870727
.FTMIB	0.14089962	0.2739568
.FTSE	0.13898212	0.2823674
.GDAXI	0.15540606	0.2670089
.GSPTSE	0.14469026	0.3000855
.HSI	0.12339313	0.2305324
.IBEX	0.13171773	0.2693158
.IXIC	0.15594794	0.2897655
.KS11	0.12479648	0.2773506
.KSE	0.11244817	0.3891245
.MXX	0.09265060	0.2851999
.N225	0.13216697	0.2993542
.NSEI	0.12837502	0.3185341
.OMXC20	0.11454790	0.2984794
.OMXHPI	0.12444437	0.3141314
.OMXSPI	0.13371935	0.3058484
.OSEAX	0.13926292	0.2614015
.RUT	0.12839201	0.3496840
.SMSI	0.11575621	0.3198488
.SPX	0.16604650	0.3022089
.SSEC	0.13577949	0.3194991
.SSMI	0.18768451	0.1886540
.STI	0.06749902	0.2394070
.STOXX50E	0.11488373	0.3543745

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

Having established these beautiful scaling results for the moments, how do the histograms look?

In [13…

```
plotScaling <- function(j,scaleFactor){
  v <- as.numeric(rv.list[[j]])
```

```
  x <- 1:100

  xDel <- function(x,lag){diff(x,lag=lag)}
  sd1 <- sd(xDel(log(v),1))
  sdl <- function(lag){sd(xDel(log(v),lag))}

  h <- OxfordH$h.est[j]

  plotLag <- function(lag){
    y <- xDel(log(v),lag)
    hist(y,breaks=100,freq=F,main=paste("Lag =",lag,"Days"),xlab=NA)# Very long tailed!
    curve(dnorm(x,mean=mean(y),sd=sd(y)),add=T,col=rd,lwd=2)
    curve(dnorm(x,mean=0,sd=sd1*lag^h),add=T,lty=2,lwd=2,col=bl)
  }

  (lags <- scaleFactor^(0:3))
  print(names(rv.list)[j])
  par(mfrow=c(2,2))
  par(mar=c(3,2,1,3))
  for (i in 1:4){plotLag(lags[i])}
  par(mfrow=c(1,1))
}
```

In [14…

```
options(repr.plot.height=7, repr.plot.width=10)
```

In [15…

```
plotScaling(27,5)
```

```
[1] ".SPX"
```

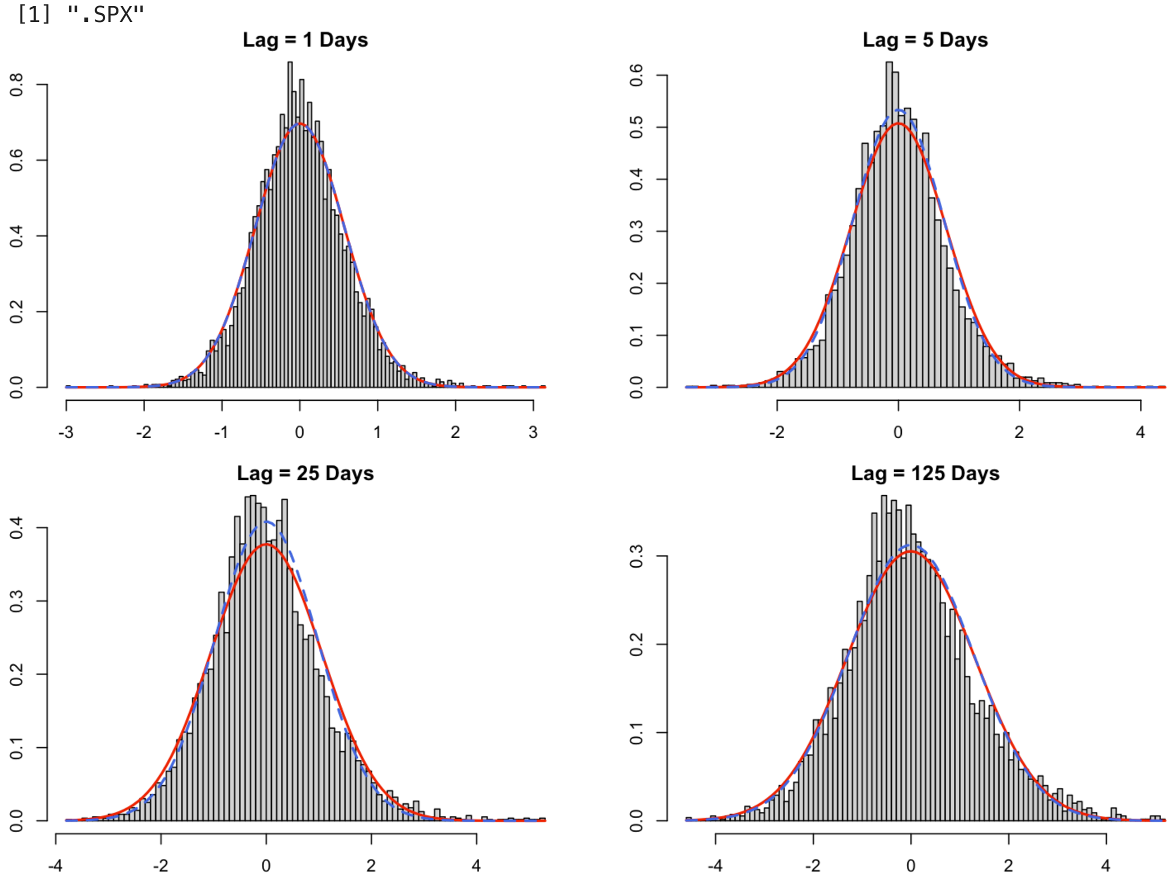


Figure 10: Histograms of $(\log \sigma_{t+\Delta} - \log \sigma_t)$ for various lags Δ; normal fit in red; $\Delta = 1$ normal fit scaled by Δ^H in blue.

Universality?

- [Gatheral, Jaisson and Rosenbaum][10] 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.
- [Bennedsen et al.][2], estimate volatility time series for more than five thousand individual US equities, finding rough volatility in every case.

A microstructural explanation: A Hawkes model of price formation

- Why might rough volatility be universal?
- [Jaisson and Rosenbaum][12] show that rough volatility can be obtained as a scaling limit of a simple model of price dynamics in terms of Hawkes processes.
- Remarkably, [El Euch and Rosenbaum][7] were able to compute the characteristic function of the resulting *rough Heston* model.

A natural model of realized volatility

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

  (1)

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

  where W^H is fractional Brownian motion.

- Indeed, if H is constant, (1) is the *unique* model consistent with Gaussianity of log differences, the observed scaling, and continuity of the volatility process.

Fractional Brownian motion (fBm)

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

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

  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 ("trending").
  - If $H < 1/2$, increments are negatively correlated ("reverting").

More sophisticated estimators of H

- Numerous authors have pointed out that the estimates of H by linear regression in [Gatheral, Jaisson and Rosenbaum][9] make sense only if estimation error is not too 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!
    - Easily rejected by examining the magnitude of ν.
- More sophisticated estimators of H include
    - The ACF estimator of [Bennedsen et al.][4]
    - The Whittle estimator of [Fukasawa and Takabatake][9]
    - The GMM estimator of [Bolko et al.][5]
    - The TDML estimator of [Wang et al.][14]
- All of these papers conclude that volatility of SPX is indeed rough.

Heuristic derivation of the ACF estimator

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 $\mathrm{var}[y_t] = t^{2H}$. and

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

Dividing one by the other 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 $H \ll \frac{1}{2}$.
- Also, when $H = \frac{1}{2}$, we have $\rho(\Delta) = 1$ as we would expect for Brownian motion.

The ACF estimator

Taking logs of each side, we obtain

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

- Thus H can be estimated efficiently by regression.

In [16…

```
h.acf <- function(path){
    y.acf <- acf(path,plot=F)
    log.del <- log(y.acf$lag[-1])
    log.lhs <- log(1-y.acf$acf[-1])
    fit.lm <- lm(log.lhs ~ log.del)
```

```
    return(fit.lm$coef[2]/2)
        }
```

Estimates of H for two different periods of history

First 2017:

In [17…
```
yPath <- spx.rk["2017-01-01::2017-12-31"]
plot(log(yPath),col=bl)
```

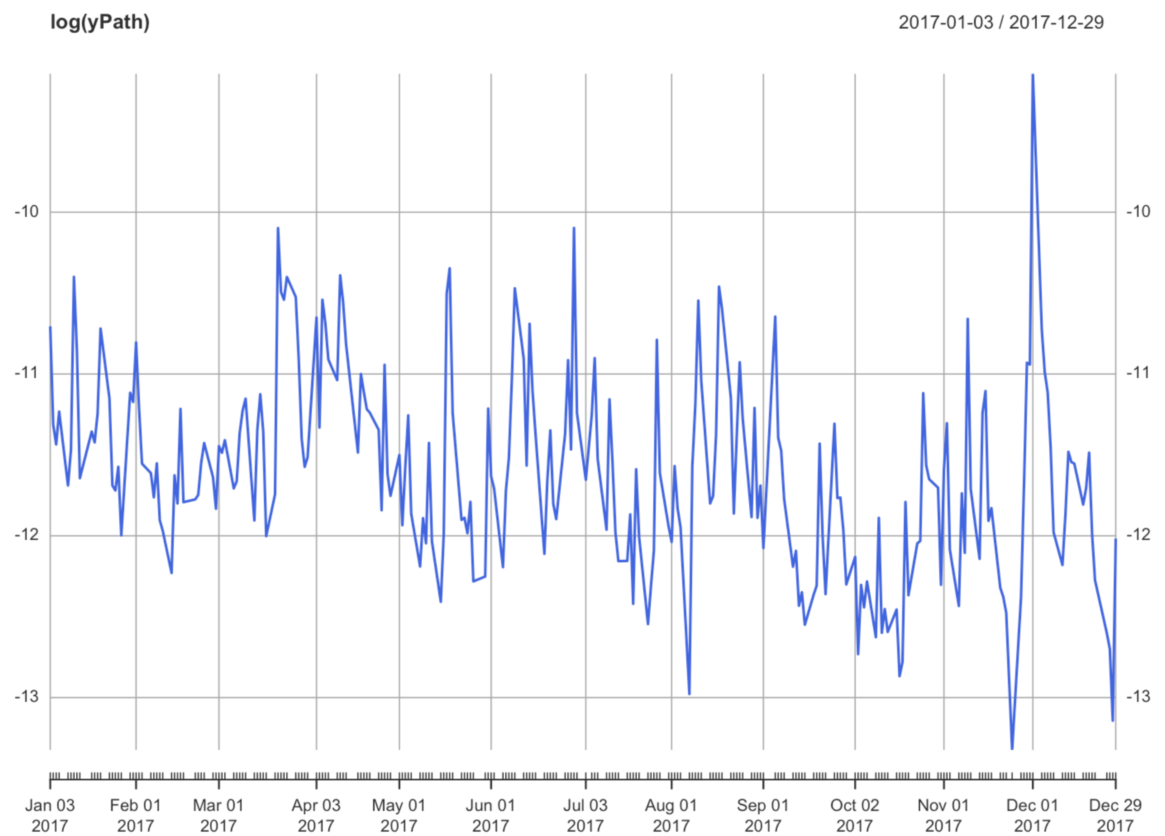


Figure 11: Log of SPX realized kernel estimates of integrated variance for 2017.

In [18…
```
h.acf(as.numeric(yPath))
```

log.del: 0.0592206840376

Then 2020:

In [19…
```
yPath <- spx.rk["2020-01-01::2020-12-31"]
plot(log(yPath),col=bl)
```

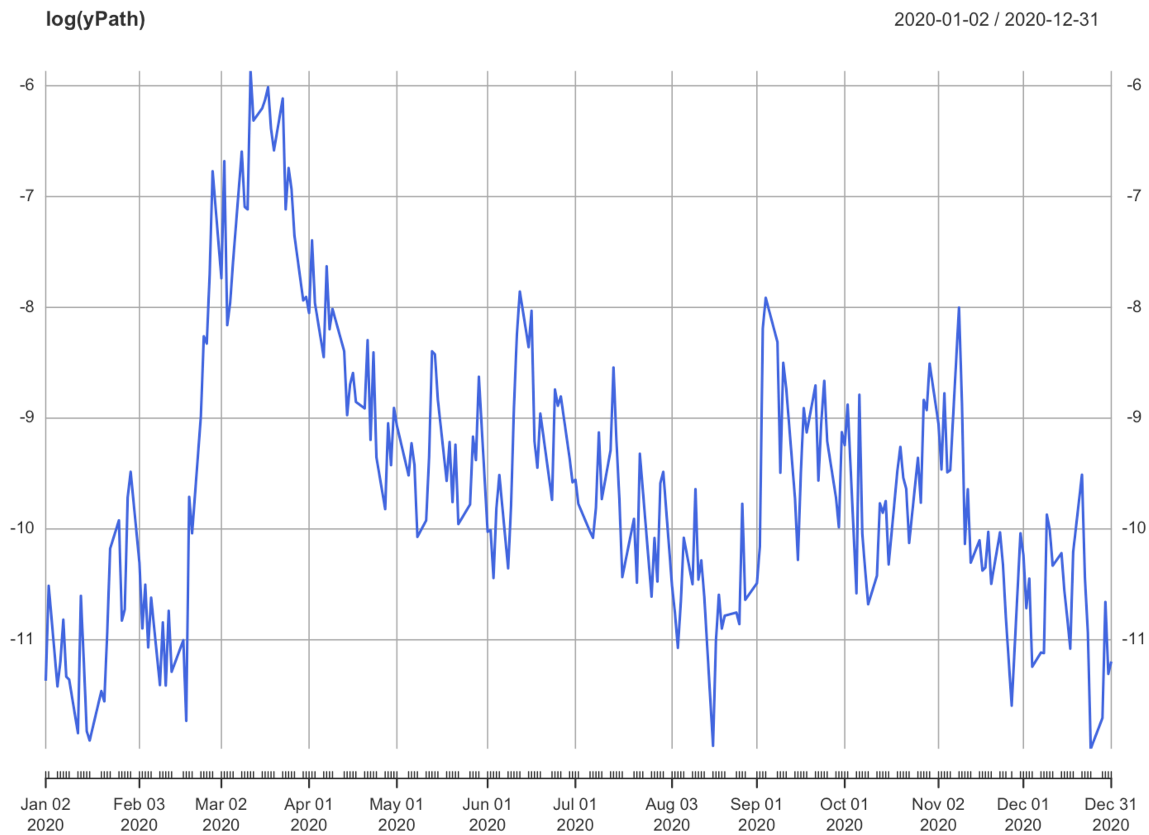


Figure 12: Log of SPX realized kernel estimates of integrated variance for 2020.

In [20…
```
h.acf(as.numeric(yPath))
```

log.del: 0.34899344707183

- Two very different estimates of H for different periods.

Time series of H using ACF

- We now give code to compute the time series of H using the ACF estimator.

In [21…
```
h.acf.i <- function(series)function(del)function(i){
    rk.path <- as.numeric(series[(i-del):i])
    h.acf(rk.path)
    }

h.acf.series <- function(series)function(del){
    require(xts)
    n <- length(series)
    res <- sapply((1+del):n,h.acf.i(series)(del))
    return(xts(res,order.by=index(series[(1+del):length(series)]),tzone = Sys.getenv("TZ")))
}
```

Compare the two estimates of H over the whole dataset

In [22…
```
rownum <- which(OxfordH[,1]==".SPX")
n.spx <- length(spx.rk)
h.spx.acf <- as.numeric(h.acf.series(spx.rk)(n.spx-1))
h.spx.regression <- OxfordH$h.est[rownum]
```

```
nu.spx.regression <- OxfordH$nu.est[rownum]
data.frame(h.spx.acf,h.spx.regression)
```

A data.frame: 1 × 2

h.spx.acf	h.spx.regression
<dbl>	<dbl>
0.1337375	0.1660465

- Looking again at the log-log plots of $m_q(\Delta)$ against Δ, we note that the points don't quite lie on a straight line.
- A more careful analysis that takes account of the bias due to averaging and the noisiness of the time series of realized variance gives us an estimate of H more consistent with the ACF estimate.

Time series of H for SPX

- Here $\alpha = H - \frac{1}{2}$. Estimates use 15-minute data.

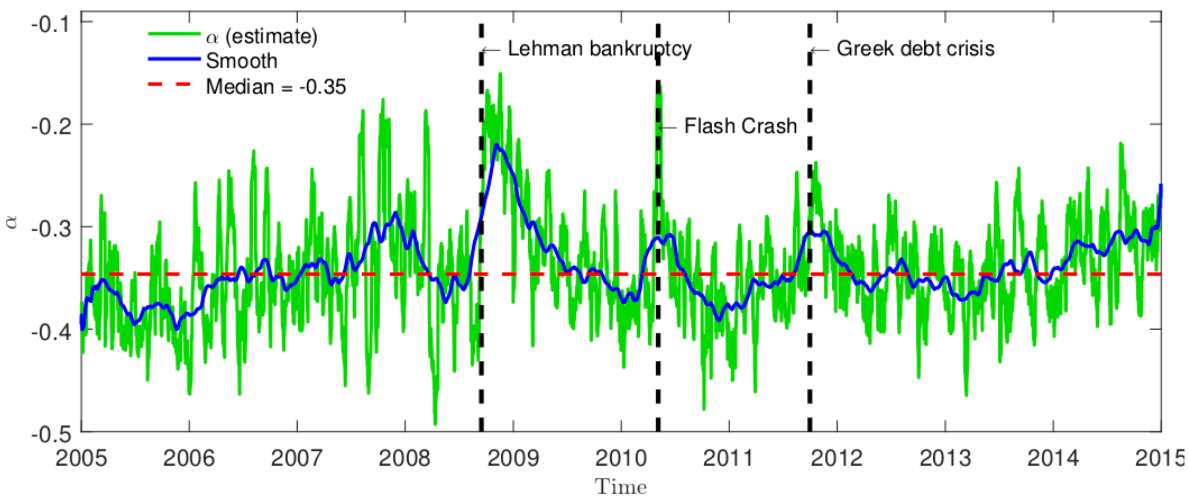


Figure 13: Time series of H from [Bennedsen et al.][4].

Observations

- H tends to spike when the market is under stress.
  - And seems close to zero when the market is calm.
  - Could H be related to underlying market liquidity?
- Note the following peaks
  - The Greek debt crisis in late 2011.
  - The Brexit vote in 2015. In this case H rises with uncertainty then collapses.
- When the market crashes, H rises. But often H rises without the market crashing.
- In particular, H of the volatility time series seems to be a meaningful and relevant statistic.

Time series of ACF-estimated H for SPX

In [23…

```
h.spx.61 <-  h.acf.series(spx.rk["20170101::"])(61)
```

In [24…
```
options(repr.plot.width=14,repr.plot.height=7)
plot(h.spx.61,main="SPX",ylab="H",col="red")
```

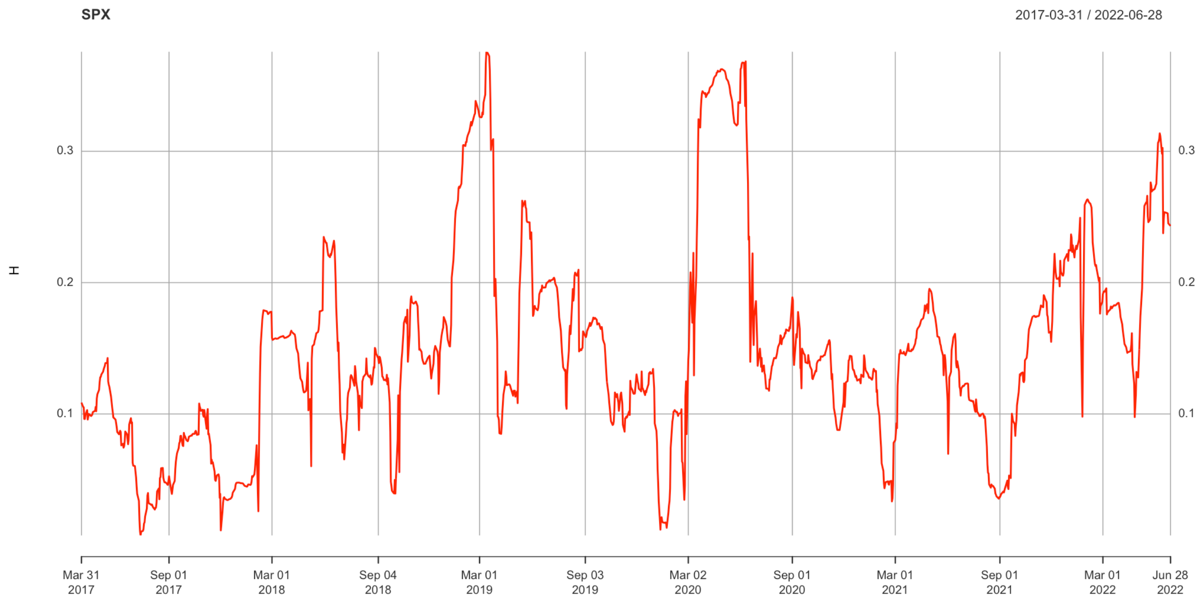


Figure 14: Time series of H using data realized kernel estimates.

Time series of H for STOXX50

In [25…
```
h.stoxx.61 <- h.acf.series(stoxx.rk["20170101::"])(61)
```

In [26…
```
plot(h.stoxx.61,main="STOXX50",ylab="H",col=bl)
```

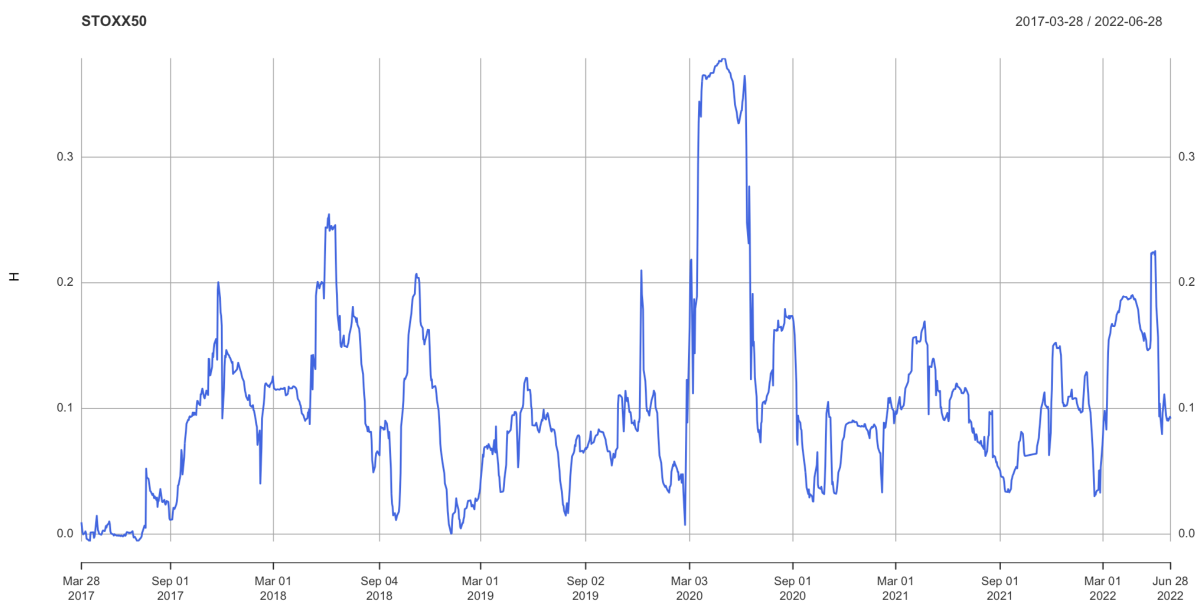


Figure 15: Time series of H for STOXX50 using data realized kernel estimates.

Plot both together

In [27…
```
plot(cbind(h.spx.61,h.stoxx.61),main="SPX plus STOXX50",col=c(rd,bl),lwd=2,major.ticks= "yea
          minor.ticks = FALSE)
legend((x="topleft"), legend = c("SPY", "STOXX50"),lty = 1,lwd=2,col = c(rd,bl))
```

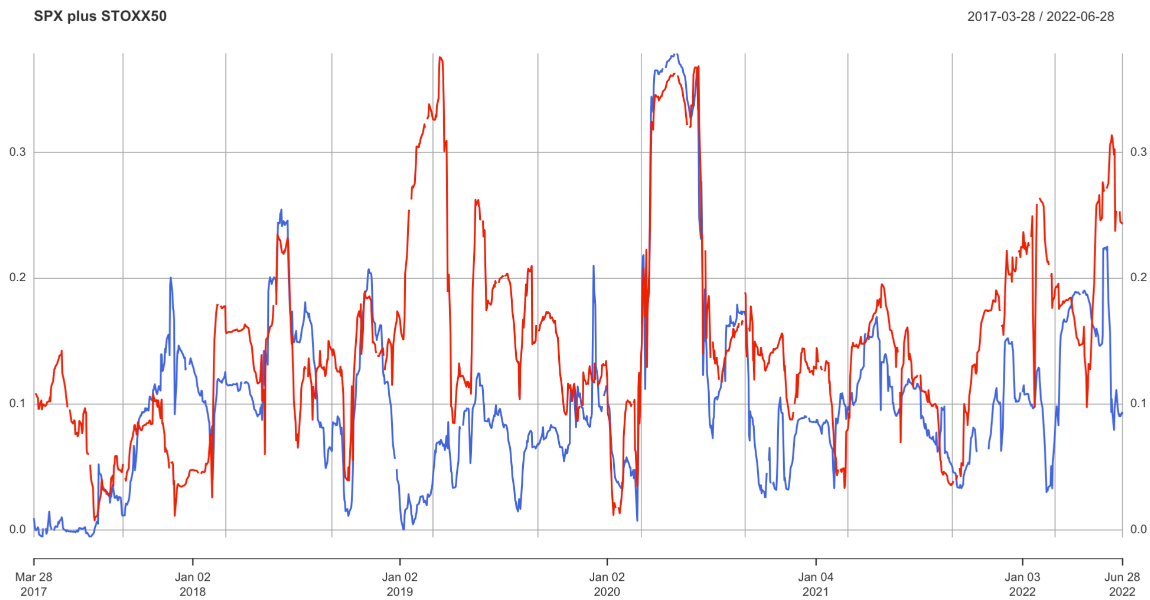


Figure 16: Sometimes the peaks line up, and sometimes not.

Line up time series of H with VIX

- First we use `quantmod` to download VIX data.

In [28…

```
options("getSymbols.warning4.0"=FALSE,"getSymbols.yahoo.warning"=FALSE)
getSymbols('^VIX',from="2017-01-01",to="2022-06-28")
```

```
Warning message:
“^VIX contains missing values. Some functions will not work if objects contain missing values
in the middle of the series. Consider using na.omit(), na.approx(), na.fill(), etc to remove
or replace them.”
```

'VIX'

Superimpose VIX and H time series

In [29…

```
plot(as.zoo(log(Cl(VIX))),col=bl,yaxt="n",ylab="",xlab="Date")
par(new=TRUE)
plot(as.zoo(h.spx.61),col=rd, xaxt="n", yaxt="n",  xlab="", ylab="",lwd=2)
```

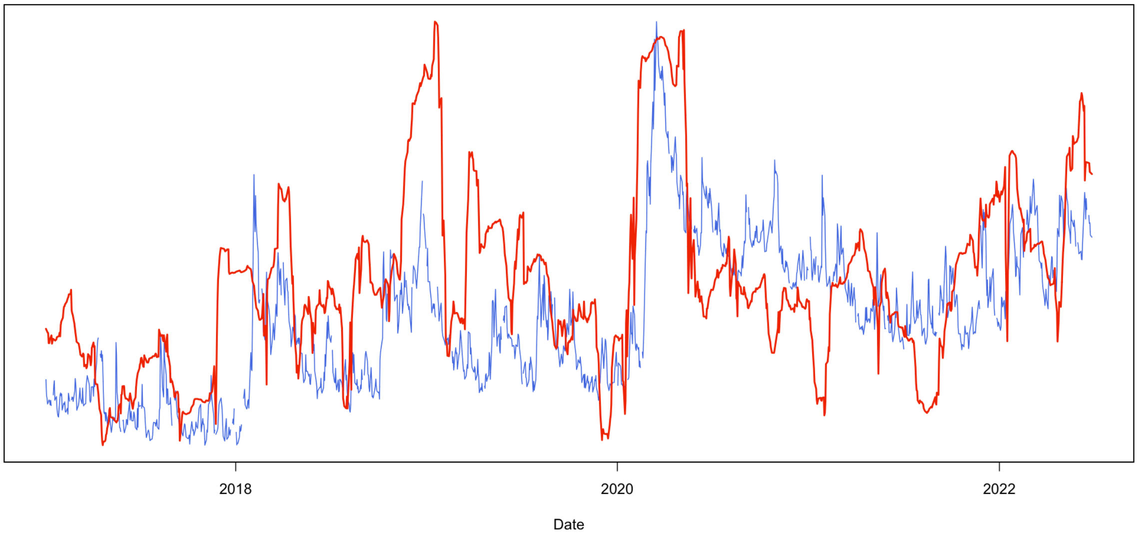


Figure 17: log(VIX) in blue; H in red. Sometimes H increases with VIX and sometimes not.

Comte and Renault: FSV model

[Comte and Renault][6] were perhaps the first to model volatility using fractional Brownian motion.

In their fractional stochastic volatility (FSV) model,

$$\begin{aligned}\frac{dS_t}{S_t} &= \sigma_t\, dZ_t \\ d\log\sigma_t &= -\alpha\,(\log\sigma_t - \theta)\, dt + \gamma\, d\hat{W}_t^H\end{aligned} \tag{1}$$

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

RFSV and FSV

- The model (1):

  $$\log\sigma_{t+\Delta} - \log\sigma_t = \nu\left(W_{t+\Delta}^H - W_t^H\right)$$

  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 (1)) 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][6] 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, Bollerslev, Diebold and Ebens][2] which estimated the degree d of fractional integration from daily realized variance data for the 30 DJIA stocks. They effectively tried to fit something like FIGARCH.
  - Using the GPH (Geweke-Porter-Hudak) estimator, they found d around 0.35 which implies that the ACF $\rho(\tau) \sim \tau^{2\,d-1} = \tau^{-0.3}$ as $\tau \to \infty$.

Log-log plot of empirical autocorrelation of volatility (correlogram)

In [30…

```
v <- as.numeric(rv.list[[".SPX"]]  )
ac.sig <- acf(sqrt(v),lag=100,plot=F)
plot(log(ac.sig$lag[-1]),log(ac.sig$acf[-1]),pch=20,
     ylab=expression(rho[sigma](Delta)),xlab=expression(paste(Delta," (days)")),col=bl)
```

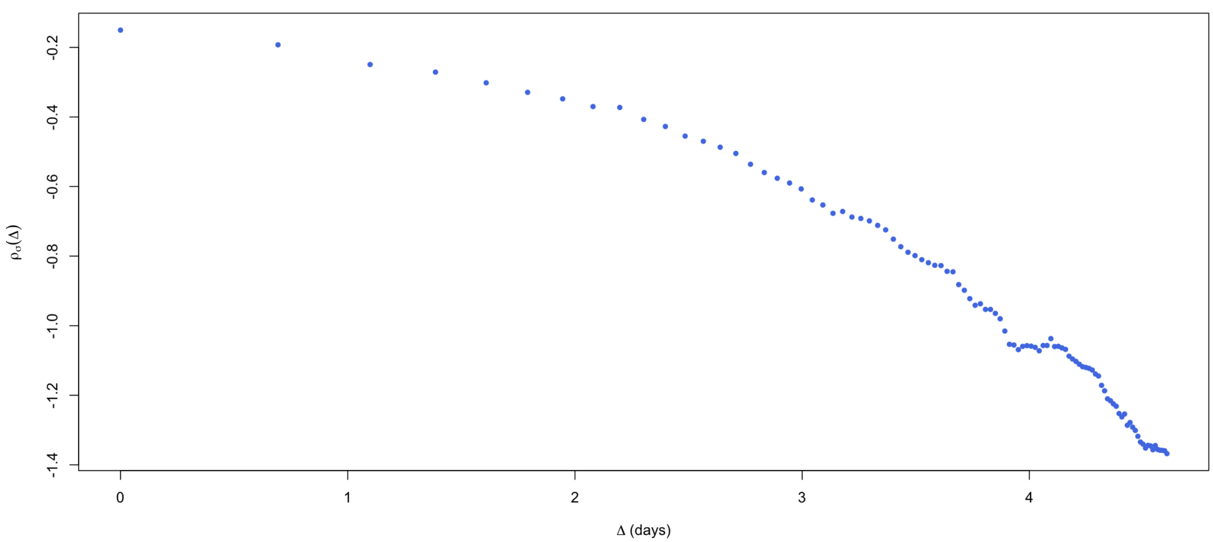


Figure 18: A correlogram of $\sigma_t = \sqrt{RV_t}$; it doesn't look linear!

Power-law fit

- We exclude the first 20 points so as to fit the tail.

In [31…

```
(fit.lm <- lm(log(ac.sig$acf[-1][-(1:20)]) ~ log(ac.sig$lag[-1][-(1:20)])))
```

```
Call:
lm(formula = log(ac.sig$acf[-1][-(1:20)]) ~ log(ac.sig$lag[-1][-(1:20)]))

Coefficients:
                 (Intercept)  log(ac.sig$lag[-1][-(1:20)])
                      0.9087                       -0.4889
```

In [32…

```
plot(log(ac.sig$lag[-1]),log(ac.sig$acf[-1]),pch=20,
     ylab=expression(rho(Delta)),xlab=expression(paste(Delta," (days)")),col=bl)
```

```
abline(fit.lm,col=rd,lwd=2)
```

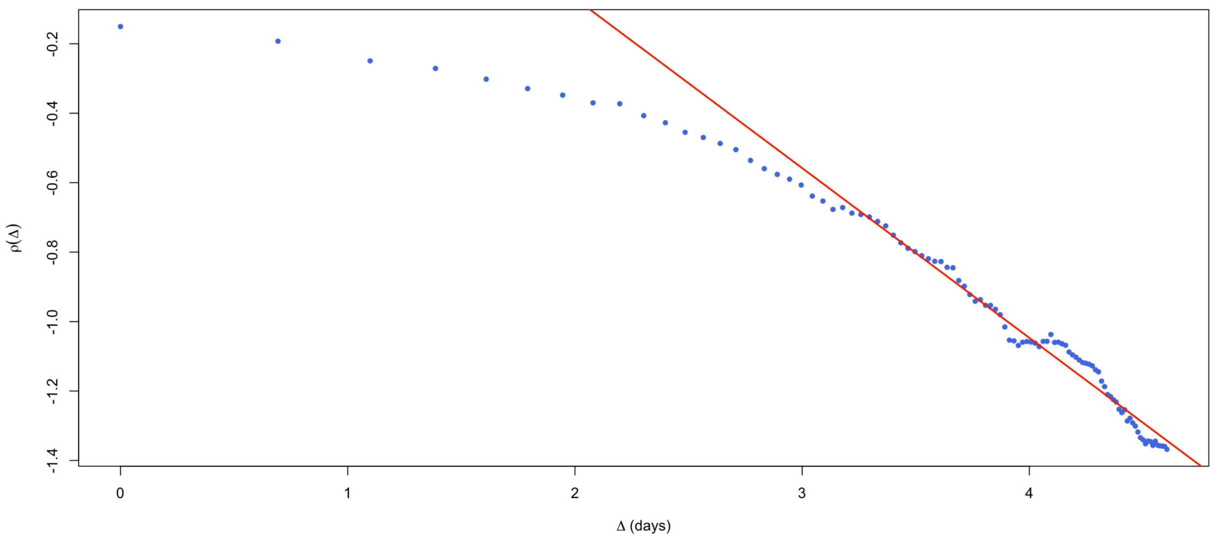


Figure 19: Correlogram of $\sigma_t = \sqrt{RV_t}$ with power-law fit.

- In other words, just fitting a straight line to the log-log plot of the autocorrelation $\rho_\sigma(\Delta)$ of the volatility we get

$$\rho_\sigma(\Delta) \sim \Delta^{-0.49}$$

as $\Delta \to \infty$.

- This corresponds to $d = 0.25$, consistent with the $d = 0.35$ found by [Andersen, Bollerslev, Diebold and Ebens][1].
- Note however that the correlogram does not look like a straight line on the log-log plot!

Plot vs Δ^{2H}

- Again, we have $\log \sigma_t = \nu \, W_t^H + \text{const.}$ so

$$\text{cov}\left[\log \sigma_t, \log \sigma_{t+\Delta}\right] = \text{var}\left[\log \sigma_t\right] - \nu^2 \, t^{2\,H} \, \Delta^{2H}.$$

- Thus $\text{cov}\left[\log \sigma_t, \log \sigma_{t+\Delta}\right]$ should be a linear function of Δ^{2H}.

In [33…

```
sig.cov <- acf(sig,lag.max=100,type="covariance",plot=F)$acf[-1]
x <- (1:100)^(2*h.spx.regression)
plot(x,sig.cov,pch=20,col="dark green",ylab= expression(paste("Covariance of log ",sigma)),
     xlab=expression(Delta^0.33 ))
abline(lm(sig.cov~x),col="red",lwd=2)
```

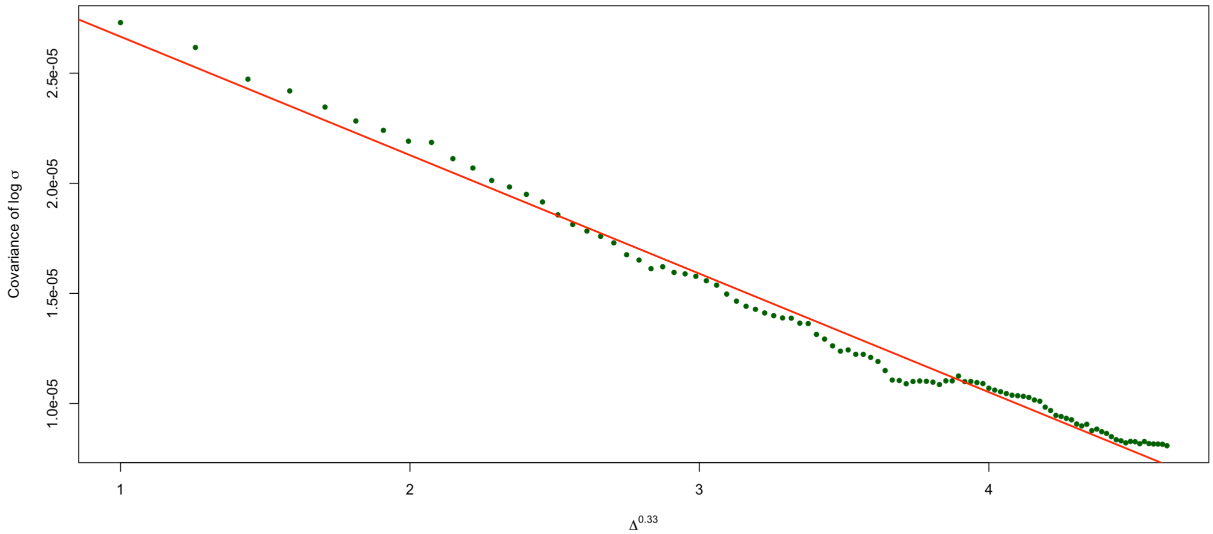


Figure 20: The data is very consistent with the RFSV model.

Long memory of volatility may be spurious

- Figures 5, 6, and 7 all demonstrate consistency of the realized kernel data with RFSV and are inconsistent with power-law decay of the autocorrelation function.
  - RFSV does not have this long memory property.
- Moreover, [Gatheral, Jaisson and Rosenbaum][6] simulate volatility in the RFSV model and apply standard estimators 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.

- Classical estimation procedures seem to identify spurious long memory of volatility.

- Here is a quote from [Bennedsen, Lunde and Pakkanen][4]:

> Having examined intraday volatility measurements on the E-mini S&P 500 futures contract, we can conclude that volatility is rough, highly persistent, and non-Gaussian. However, we were unable to distinguish between genuine long memory and persistence, yet technically short memory in the data.

- The potential mis-indentification of long memory is further explored in a nice paper by [Li et al.][13]

Incompatibility of FSV with realized variance (RV) data

- In Figure 9, 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 $\rho(\Delta) \propto \Delta^{2\,H-2}$. 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 $\Delta \ll 1/\alpha$, mean reversion is not significant and $m(2,\Delta) \propto \Delta^{2\,H}$.

RFSV vs FSV

- We can compute $m(2, \Delta)$ explicitly in both the FSV and RFSV models.
- The smallest possible value of H in FSV is $H = 1/2$. One empirical estimate in the literature says that $H \approx 0.53$ some time in 2008.

- Let's see how the theoretical estimates of $m(2, \Delta)$ compare with data.

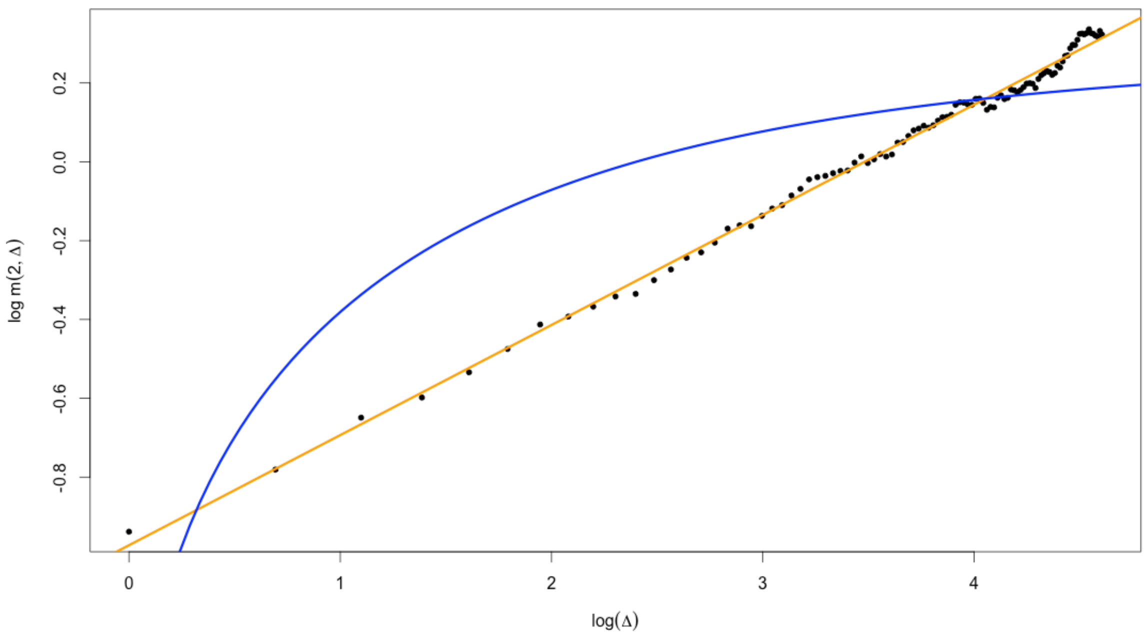
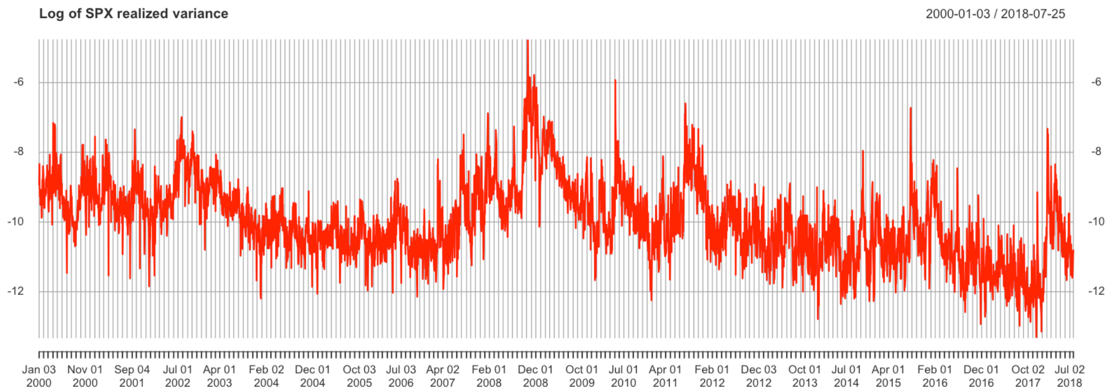


Figure 21: Black points are empirical estimates of $m(2, \Delta)$; the blue line is the FSV model with $\alpha = 0.5$ and $H = 0.53$; the orange line is the RFSV model with $\alpha = 0$ and $H = 0.14$.

Does simulated RSFV data look real?



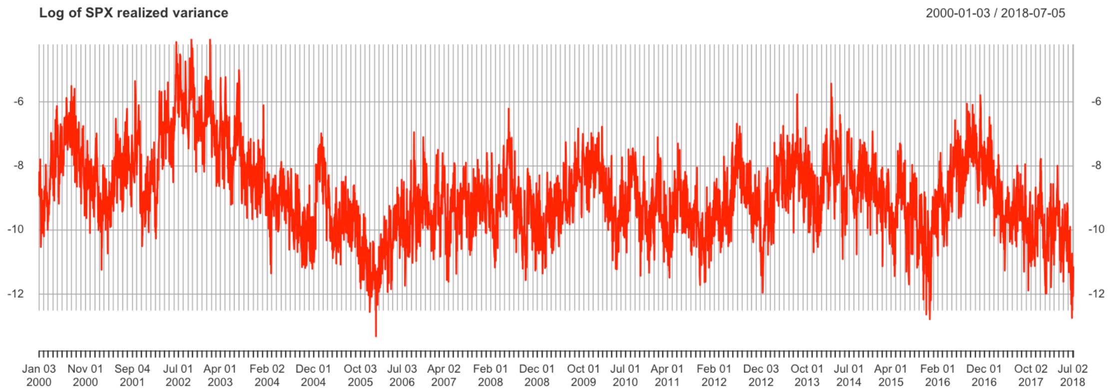


Figure 22: Volatility of SPX (above) and of the RFSV model (below).

Remarks on the comparison

- In respect of roughness, 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][3] .
  - In particular, their Multifractal Random Walk (MRW) is related to a limiting case of the RSFV model as $H \to 0$.

Applications

- What is this rough volatility model good for?
- If we could change measure from $\mathbb{P}$ to $\mathbb{Q}$, we would be able to price options.
- Another obvious application is to volatility forecasting.

Forecasting fBm

- In the RFSV model (1), $\log \sigma_t \approx \nu\, W^H_t + C$ for some constant C.

- [Nuzman and Poor][14] 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

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

  where

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

A heuristic explanation of the formula

- The forecast formula comes from regressing $W^H_{t+\Delta}$ against the W^H_s with $s<t$.

- Let

$$
\beta(u,\Delta) = \frac{\cos(H\pi)}{\pi} \Delta^{H+1/2} \frac{1}{(u+\Delta)\, u^{H+1/2}}.
$$

  Then, for $t, \Delta > 0$ and $0

$$
\int_0^\infty \beta(u,\Delta)\, |t-u|^{2H}\, du = (t+\Delta)^{2H}.
$$

  - In particular,

$$
\int_0^\infty \beta(u,\Delta)\, du = 1.
$$

- With $\beta(u,\Delta)$ thus defined and for $s<t$,

$$
\mathbb{E}\left[W^H_s \left(W^H_{t+\Delta} - \int_{-\infty}^t \beta(t-u,\Delta)\, W^H_u\, du \right) \right] = 0.
$$

  - In other words, the $\beta(t-u,\Delta)$ are the normal regression coefficients.

The forecast formula

Using that W^H is a Gaussian random variable, we get that

Variance forecast formula

(3)

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

where

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

Discretization of the forecast formula

In [Gatheral, Jaisson, Rosenbaum][4], we discretize the integral by taking mid-points as in

$$\mathbb{E}^{\mathbb{P}}\left[\log V_{t+\Delta}\right|\mathcal{F}_t] \approx \frac{1}{A}\sum_{j=0}^{L}\frac{\log V_{t-j}}{\left(j+\frac{1}{2}+\Delta\right)\,(j+\frac{1}{2})^{H+1/2}}.$$

where L is the maximum number of lags and the normalizing constant A is given by

$$A=\sum_{j=0}^{L}\frac{1}{\left(j+\frac{1}{2}+\Delta\right)\,(j+\frac{1}{2})^{H+1/2}}.$$

Inspired by [Bennedsen, Lunde and Pakkanen][4], we approximate the first term in the sum more accurately as follows.

$$\mathbb{E}^{\mathbb{P}}\left[\log V_{t+\Delta}\right|\mathcal{F}_t] \approx \frac{1}{A}\left\{\frac{\log V_t}{(s^\star+\Delta)\,(s^\star)^{H+1/2}}+\sum_{j=1}^{L}\frac{\log V_{t-j}}{(j+\frac{1}{2}+\Delta)\,(j+\frac{1}{2})^{H+1/2}}\right\}$$

where $s^\star$ is chosen such that

$$\frac{1}{\gamma}=\int_0^1\frac{ds}{s^{H+\frac{1}{2}}}=\frac{1}{{s^\star}^{H+\frac{1}{2}}}=\frac{1}{{s^\star}^{1-\gamma}}$$

where $\gamma=\frac{1}{2}-H$. Thus

$$s^\star=\gamma^{\frac{1}{1-\gamma}}.$$

Implement variance forecast in R

In [34…
```
# Find all of the dates
dateIndex <- substr(as.character(index(spx.rk)),1,10) # Create index of dates

cTilde <- function(h){gamma(3/2-h)/(gamma(h+1/2)*gamma(2-2*h))} # Factor because we are comp

# XTS compatible version of forecast
rv.forecast.XTS <- function(rvdata,h,date,nLags,delta,nu){
    gam <- 1/2-h
    j <- (1:nLags)-1
    cf <- 1/((j+1/2)^(h+1/2)*(j+1/2+delta)) # Lowest number should apply to latest date
    s.star <- gam^(1/(1-gam))
    cf[1] <- 1/(s.star^(h+1/2)*(s.star+delta))
    datepos <- which(dateIndex==date)
    ldata <- log(as.numeric(rvdata[datepos-j])) # Note that this object is ordered from earl
    pick <- which(!is.na(ldata))
    norm <- sum(cf[pick])
    fcst <- cf[pick]%*%ldata[rev(pick)]/norm # Most recent dates get the highest weight
```

```
    return(exp(fcst+2*nu^2*cTilde(h)*delta^(2*h)))

    }
```

SPX actual vs forecast variance

- In order to forecast using (3), we need estimates of H and ν.
    - We use our estimates of H and ν from the regressions rather than from the ACF estimator.
    - The choice does not seem to make much difference.

In [35…

```
var.forecast.spx <- function(h,nu)function(del){
    n <- length(spx.rk)
    nLags <- 200

    range <- nLags:(n-del)
    rv.predict <- sapply(dateIndex[range],function(d){rv.forecast.XTS(rvdata=spx.rk,h,d,nLag
    rv.actual <- spx.rk[range+del]
    return(list(rv.predict=rv.predict,rv.actual=rv.actual))
    }
```

- From experiment, we found that around 200 lags works best.

Scatter plot of delta days ahead predictions

In [36…

```
del <- 1
vf <- var.forecast.spx(h=h.spx.regression,nu=nu.spx.regression)(del)
rv.predict <- vf$rv.predict
rv.actual <- vf$rv.actual
vol.predict <- sqrt(as.numeric(rv.predict))
vol.actual <- sqrt(as.numeric(rv.actual))
vol.actual <-  sqrt(as.numeric(rv.actual))
```

In [37…

```
c(mean(vol.actual-vol.predict),sd(vol.actual-vol.predict))
```

-0.000353060171100924 · 0.0027320209272416

In [38…

```
plot(vol.predict,vol.actual,col=bl,pch=20, ylab="Actual vol.", xlab="Predicted vol.")
abline(coef=c(0,1),col=rd)
```

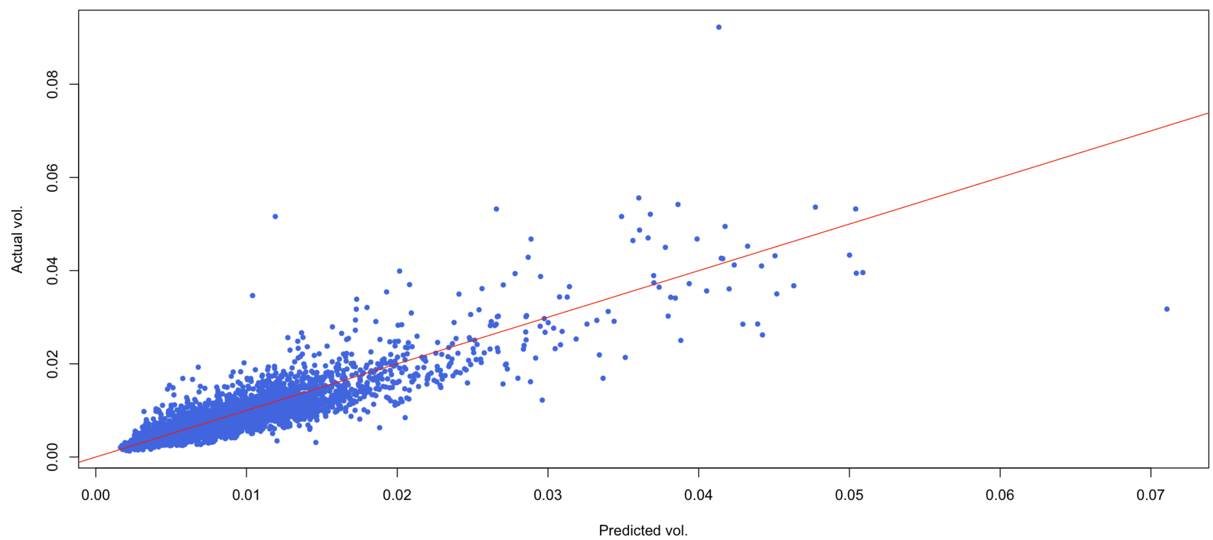


Figure 23: Actual vols vs predicted vols.

Which point is the outlier?

In [39…

```
rv.actual[which(as.numeric(vol.actual)>.09)]
```

```
                   [,1]
2008-10-10 0.008509655
```

In [40…

```
rv.predict[which(as.numeric(vol.predict)>.06)]
```

2008-10-10: 0.00505010537195742

Superimpose actual and predicted vols

In [41…

```
plot(vol.actual, col=bl,type="l")
lines(vol.predict, col=rd,type="l")
```

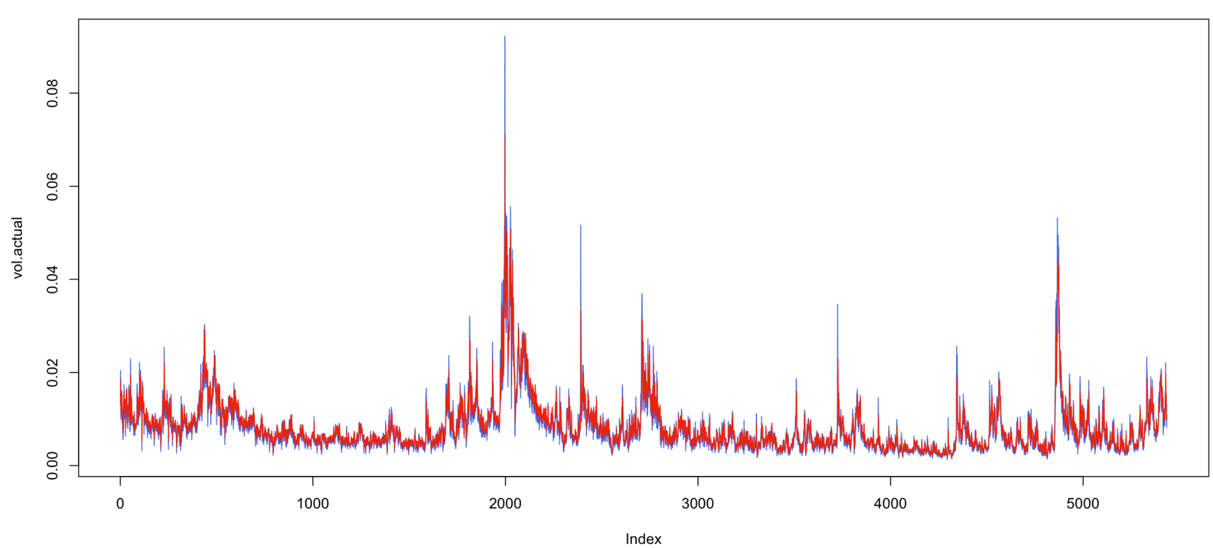
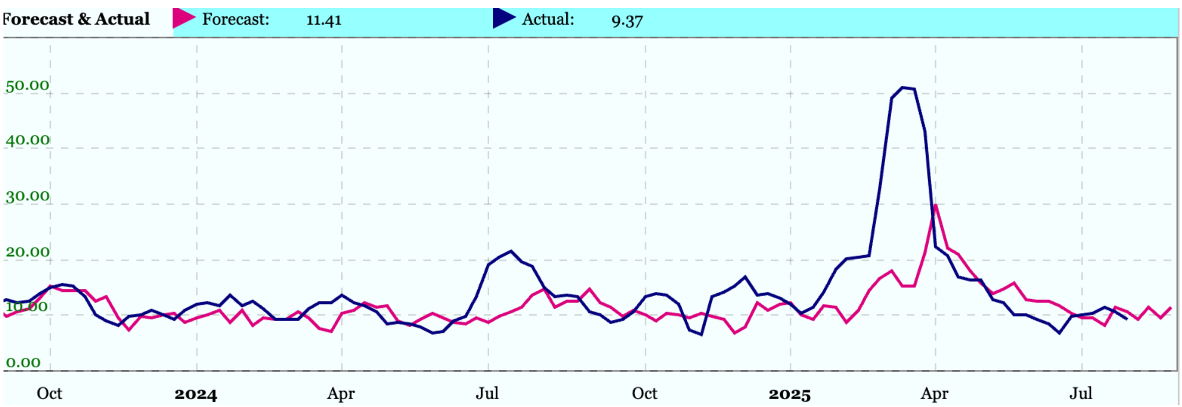


Figure 24: Actual volatilities in blue; predicted vols in red. Note that volatilities are in daily terms.

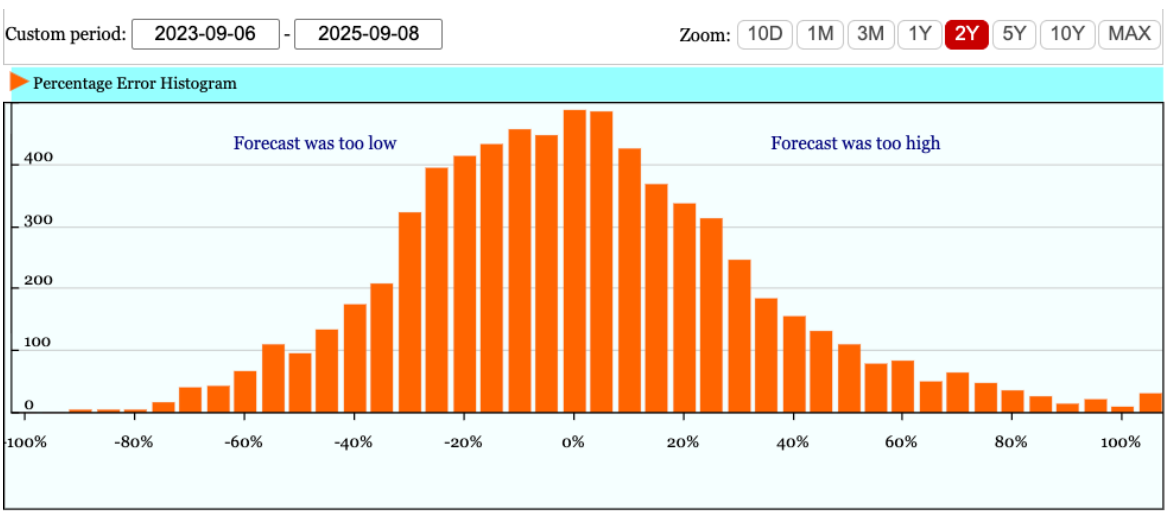
VolX

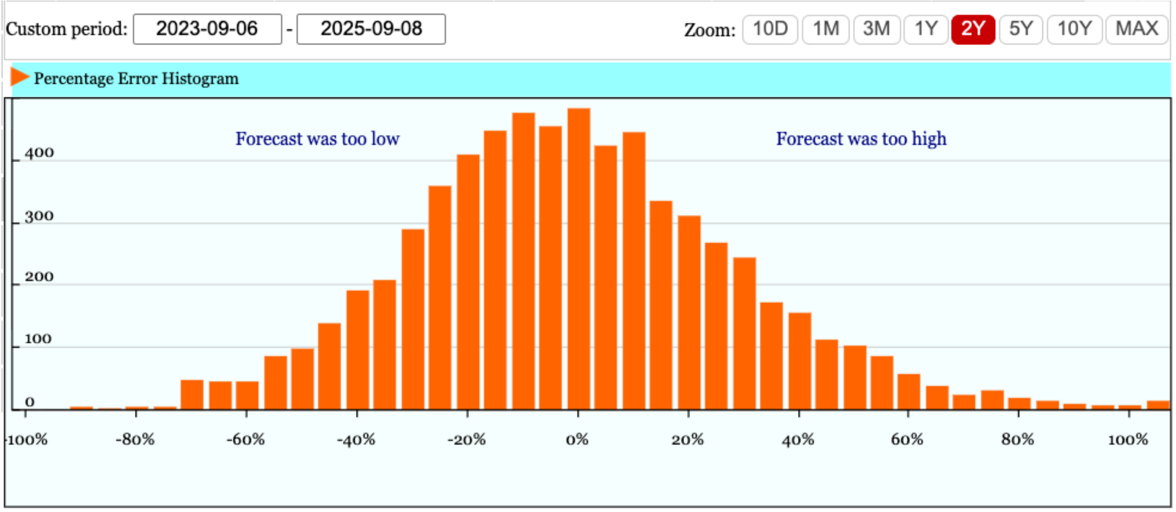
- The commercial company VolX (http://volx.us) has developed a number of RealVol Instruments and RealVol Indices based on realized volatility as defined by the RealVol Formulas.
  - Their business model is to license these indices to exchanges and information providers.
- They publish daily forecasts of RV using HARK (which is HAR-RV with Kalman filtering, and RVOL, an implementation of the Rough Volatility forecast.
- You can compare forecast versus actual volatility for the two estimators here: http://www.volx.us/volatilitycharts.shtml?2&SPY&PRED.

VolX screenshots



Rough Volatility and HAR error histograms





Conditional and unconditional variances

- The HAR and rough volatility forecasts are both impressive.
    - Much superior to alternatives such as GARCH.
- However, HAR is a regression and rough volatility is a proper model.
- One practical consequence is that we can put error bars on our volatility forecasts.

So how good is the forecast?

Specifically, by how much is the variance of the future variance reduced by taking into account the whole history of the fBm?

- In practice of course, we only consider some finite history, 200 points say.
- We know this again from [Nuzman and Poor][10] who showed that the ratio of the conditional to the unconditional variance of the $\log V_t$ is

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

- We can compute this ratio empirically and compare with the model prediction.

Unconditional and conditional variance vs lag Δ

First we compute the time series of prediction errors.

In [42…

```
log.vol.err <- function(del){
    vf <- var.forecast.spx(h=h.spx.regression,nu=nu.spx.regression)(del)
    rv.predict <- vf$rv.predict
    rv.actual <- vf$rv.actual
    vol.predict <- sqrt(as.numeric(rv.predict))
    vol.actual <-  sqrt(as.numeric(rv.actual))
    err <- log(vol.actual)-log(vol.predict)
    return(err)
    }

var.log.err <- function(del){
    var(log.vol.err(del))
}
```

```
var.log.err(10)
```

0.151181721494837

The following code takes too long to run. You can run it by uncommenting the code.

In [43…

```
del <- 1:100
# system.time(var.log.err.del <- sapply(del,var.log.err))
```

In [44…

```
# save(var.log.err.del ,file="varerr202206.rData")
load(file="varerr202206.rData")
```

Plot of conditional and unconditional variance

- The unconditional variance of differences in log-vol is given by

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

- The conditional variance is given by `var.log.err` (Δ).

In [45…

```
plot(del,mq.del(del,2),pch=20,cex=1,ylab=expression(Variance),
     xlab=expression(Delta),col=bl,ylim=c(0,.45),
     main= "Unconditional and conditional variance")
curve(nu.spx.regression^2*x^(2*h.spx.regression),from=0,to=100,add=T,col="red",lwd=2,n=1000)
points(del,var.log.err.del,col=gr,pch=20)
curve(cTilde(h.spx.regression)* nu.spx.regression^2*x^(2*h.spx.regression),from=0,to=100,
      add=T,col=or,lwd=2,n=1000)
```

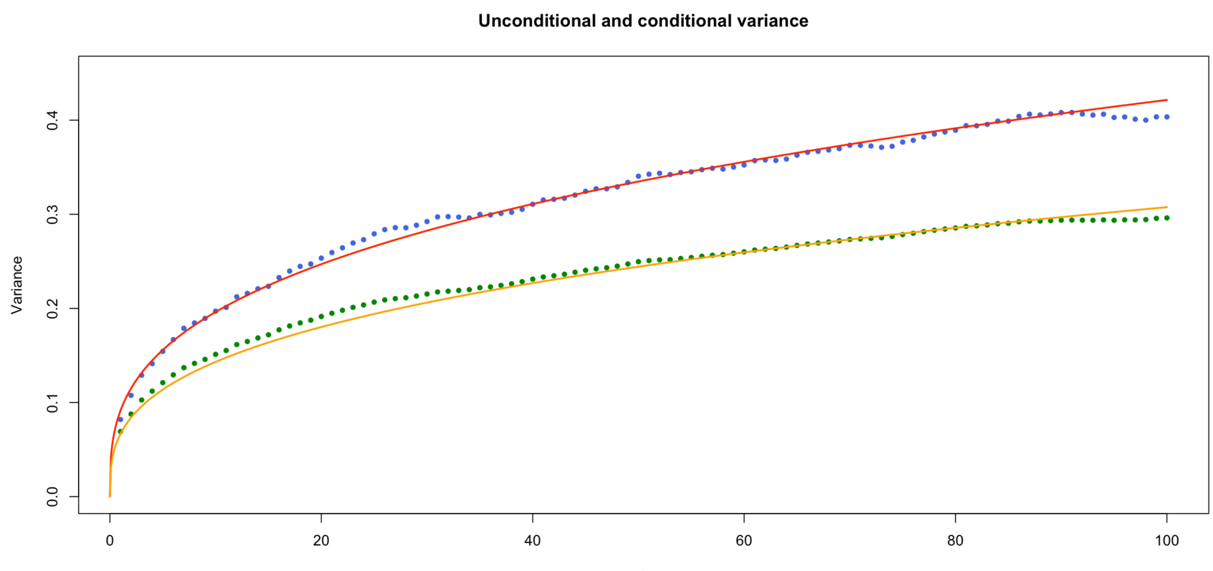


Figure 25: Actual unconditional variance in blue, rough volatility prediction in red; Actual conditional variance in green, rough volatility prediction in orange.

Amazing agreement between data and model

- We observe that the ratio of conditional to unconditional variance is more or less *exactly* as predicted by the model!

Pricing under rough volatility

Following [Bayer, Friz and Gatheral][4], the foregoing behavior suggest the following model for volatility under the real (or historical or physical) measure $\mathbb{P}$:

$$\log \sigma_u - \log \sigma_t = \nu \, \left(W^H_u - W^H_t \right), \quad u > t.$$

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

$$\mathbb{E}\left[W_t^H\,W_s^H \right] = \frac12 \, \left\{t^{2H} + s^{2H} - |t-s|^{2H} \right\}.$$

Then

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

- Note that $\mathbb{E}^{\mathbb{P}}\left[\left. M_t(u) \right| \mathcal{F}_t \right] = 0$ and $Z_t(u)$ is $\mathcal{F}_t$-measurable.
  - To specify the process, 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 model under $\mathbb{P}$

With $\tilde \eta := 2\,\nu\,C_H$, and denoting the stochastic exponential by $\mathcal{E}(\cdot)$, we may write

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

- The conditional distribution of V_u depends on $\mathcal{F}_t$ only through the variance forecasts $\mathbb{E}^{\mathbb{P}}\left[\left. V_u \right| \mathcal{F}_t \right]$,
- To specify the volatility process, one does not need to know $\mathcal{F}_t$, the entire history of the Brownian motion $W_s^\mathbb{P}$ for $s<t$.

Summary of Lecture 1

- We uncovered a remarkable monofractal scaling relationship in historical volatility.

- Conventional long memory models are inconsistent with this scaling relationship.
- Prior work indicating long memory in volatility time series is not supported.

- The Hurst exponent H varies over time.
  - Peaks typically correspond to periods of market stress.
- The resulting RFSV model can be used to forecast realized variance.

References

1. ^ Alòs, Elisa, 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).
2. ^ 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).
3. ^ Emmanuel Bacry and Jean-François Muzy, Log-infinitely divisible multifractal processes, *Communications in Mathematical Physics* **236**(3) 449-475 (2003).
4. ^ 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, (2021).
5. ^ 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).
6. ^ Fabienne Comte and Eric Renault, Long memory in continuous-time stochastic volatility models, *Mathematical Finance* **8** 29-323(1998).
7. ^ Omar El Euch and Mathieu Rosenbaum, The characteristic function of rough Heston models, *Mathematical Finance* **29**(1) 3-38 (2019).
8. ^ Masaaki Fukasawa, Asymptotic analysis for stochastic volatility: Martingale expansion, *Finance and Stochastics* **15** 635-654 (2011).
9. ^ Masaaki Fukasawa and Tetsuya Takabatake, Asymptotically efficient estimators for self-similar stationary Gaussian noises under high frequency observations, *Bernoulli* **25**(3) 1970-1900 (2019).
10. ^ Jim Gatheral, Thibault Jaisson and Mathieu Rosenbaum, Volatility is rough, *Quantitative Finance*, **18**(6) 933-949 (2018).
11. ^ Jim Gatheral and Roel Oomen, Zero-intelligence realized variance estimation, *Finance and Stochastics* **14**(2) 249-283 (2010).
12. ^ Thibault Jaisson and Mathieu Rosenbaum, Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes, *The Annals of Applied Probability* **26**(5) 2860-2882 (2016).
13. ^ Jia Ji, Peter C. B. Phillips, Shuping Shi and Jun Yu, Weak Identification of Long Memory with Implications for Inference, *SSRN* 4140818 (2022).
14. ^ Xiaohu Wang, Weilin Xiao and Jun Yu, Modeling and forecasting realized volatility with the fractional Ornstein–Uhlenbeck process, *Journal of Econometrics* **232**(2) 389-415 (2024).