18th Winter school on Mathematical Finance

Lunteren, January 21-23, 2019

Rough volatility

Lecture 2: Pricing

Jim Gatheral
Department of Mathematics

Baruch COLLEGE
The City University of New York

Outline of Lecture 2

- The volatility surface: Stylized facts
- The RFSV model under $\mathbb{Q}$
- The rough Bergomi model
- VIX futures under rough volatility
- Relating historical and implied model parameters

The SPX volatility surface as of 15-Sep-2005

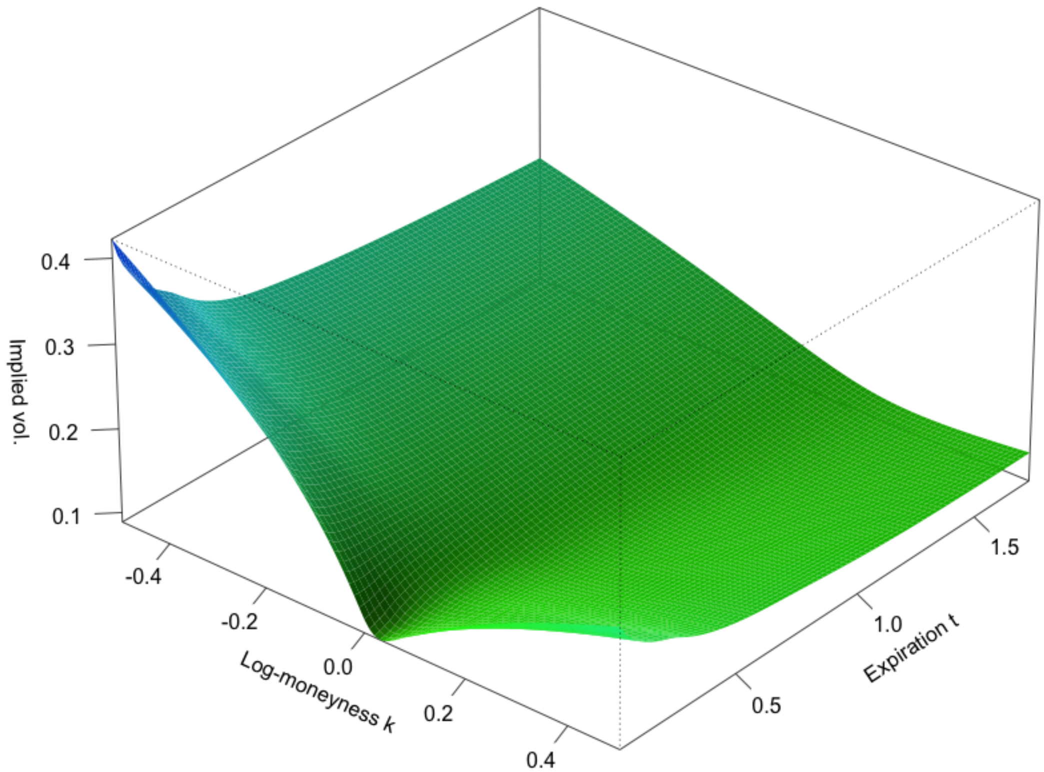

Figure 1: The SPX volatility surface as of 15-Sep-2005 (Figure 3.2 of The Volatility Surface).

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.

The 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

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

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

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

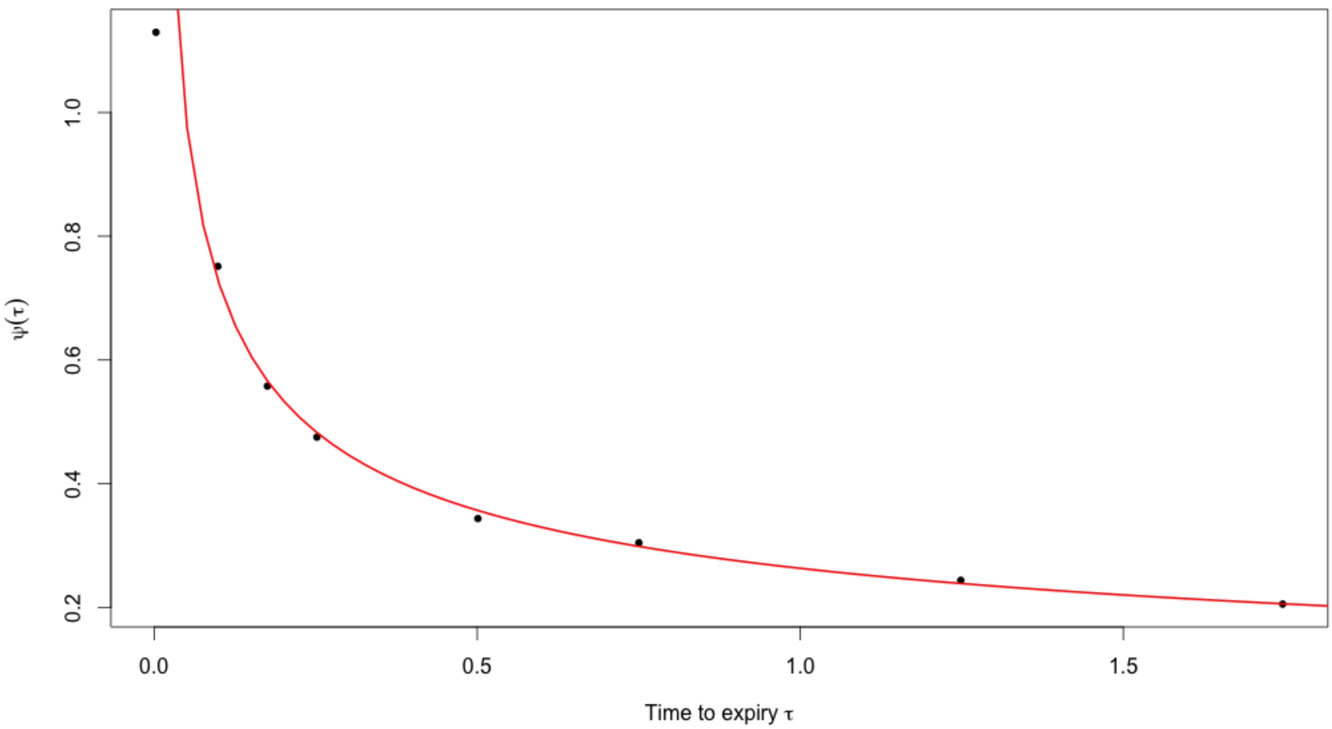

Figure 2: Term structure of ATM skew as of 15-Sep-2005, with power law fit $\tau^{-0.44}$ superimposed in red.

Stylized facts

- Although the levels and orientations of the volatility surfaces change over time, their rough shape stays very much the same.
  - 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 $\alpha \in (0.3, 0.5)$.

Conventional 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.

Forward variance curve models

Inspired by the HJM approach to interest rate modeling, [Bergomi and Guyon][6]</sup> originally suggested that it is natural to express stochastic volatility models in forward variance form. Specifically let

$$\begin{aligned}\frac{dS_t}{S_t} &= \sqrt{v_t}\, dZ_t \\ d\xi_t(u) &= \lambda(t, u, \xi_t)\, dW_t.\end{aligned}$$

where v_t denotes instantaneous variance and the $\xi_t(u) = \mathbb{E}\left[v_u \,|\, \mathcal{F}_t\right], u \in (t, T]$ are forward variances.

Forward variance curve models and perfect hedging

- As noted by [El Euch and Rosenbaum][7]</sup> , models written in forward variance form are explicitly Markovian in the asset price S_t and the (infinite-dimensional) forward variance curve ξ_t.
  - European payoffs V may be perfectly hedged.
  - The delta-hedging strategy involves holding $\partial_S V$ in the asset and $\partial_\xi V$ in forward variance contracts where ∂_ξ denotes the Fréchet derivative of V with respect to the forward variance curve.

Bergomi Guyon

- According to [Bergomi and Guyon][5]</sup> , in the context of a variance curve model, implied volatility may be expanded as

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

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

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

- Thus, given a stochastic volatility model in forward variance form, we can easily (at least in principle) compute this smile approximation.

The Bergomi model

- The n-factor Bergomi variance curve model reads:

(1)

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

- 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 (1).

- 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 (1).

Tinkering with the Bergomi model

- Empirically, $\psi(\tau) \sim \tau^{-\alpha}$ for some α.

- It's tempting to replace the exponential kernels in (1) with a power-law kernel.

- This would give a model of the form

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

  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.

History of fractional stochastic volatility models

More formally, the model

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

belongs to a larger class of fractional stochastic volatility models that was originally shown by [Alòs et al.][1]
</sup> and then by [Fukasawa][8]</sup> to generate a short-dated ATM skew of the form

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

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

Further motivation from the time series of realized volatility

- In Lecture 1, we saw that distributions of differences in log realized variance are close to Gaussian.
  - This motivates us to model v_t (and so σ_t) as a lognormal random variable.
- Moreover, the scaling property of variance of RV differences suggests the (RFSV) model:

(2)

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

where W^H is fractional Brownian motion.

Fractional Brownian motion (fBm) again

- *Fractional Brownian motion* (fBm) $\{W^H_t; t \in \mathbb{R}\}$ is the unique Gaussian process with mean zero and autocovariance function
$$\mathbb{E}\left[W^H_t\, W^H_s\right] = \frac{1}{2}\left\{|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right\}$$
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.

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

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

The RFSV model again

Then from the definition (2) of the model, with the Mandelbrot-Van Ness representation of fBm,

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

- Note that $\mathbb{E}^{\mathbb{P}}\,[M_t(u)|\,\mathcal{F}_t] = 0$ 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

Pricing under $\mathbb{P}$

Let

$$\tilde{W}_t^{\mathbb{P}}(u) := \sqrt{2H} \int_t^u \frac{dW_s^{\mathbb{P}}}{(u-s)^\gamma}$$

With $\eta := 2\,\nu\,C_H/\sqrt{2\,H}$ we have $2\,\nu\,C_H\,M_t(u) = \eta\,\tilde{W}_t^{\mathbb{P}}(u)$ so denoting the stochastic exponential by $\mathcal{E}(\cdot)$, we may write

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

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

Pricing under $\mathbb{Q}$

Our model under $\mathbb{P}$ reads:

(3)

$$v_u = \mathbb{E}^{\mathbb{P}}\left[v_u \,|\, \mathcal{F}_t\right] \mathcal{E}\left(\eta\,\tilde{W}_t^{\mathbb{P}}(u)\right).$$

Consider some general change of measure

$$dW_s^{\mathbb{P}} = dW_s^{\mathbb{Q}} + \lambda_s\,ds,$$

where $\{\lambda_s : s > t\}$ has a natural interpretation as the price of volatility risk. We may then rewrite (2) as

$$v_u = \mathbb{E}^{\mathbb{P}}\left[v_u \,|\, \mathcal{F}_t\right] \mathcal{E}\left(\eta\,\tilde{W}_t^{\mathbb{Q}}(u)\right) \exp\left\{\eta\,\sqrt{2H} \int_t^u \frac{\lambda_s}{(u-s)^\gamma}\,ds\right\}.$$

- Although the conditional distribution of v_u under $\mathbb{P}$ is lognormal, it will not be lognormal in general under $\mathbb{Q}$.
    - The upward sloping smile in VIX options means λ_s cannot be deterministic in this picture.

The rough Bergomi (rBergomi) model

Let's nevertheless consider the simplest change of measure

$$dW_s^{\mathbb{P}} = dW_s^{\mathbb{Q}} + \lambda(s)\,ds,$$

where $\lambda(s)$ is a deterministic function of s. Then from (2), we would have

$$\begin{aligned} v_u &= \mathbb{E}^{\mathbb{P}}\left[v_u \,|\, \mathcal{F}_t\right] \mathcal{E}\left(\eta\, \tilde{W}_t^{\mathbb{Q}}(u)\right) \exp\left\{\eta\,\sqrt{2H}\int_t^u \frac{1}{(u-s)^{\gamma}}\,\lambda(s)\,ds\right\} \\ &= \xi_t(u)\,\mathcal{E}\left(\eta\,\tilde{W}_t^{\mathbb{Q}}(u)\right) \end{aligned}$$

where the forward variances $\xi_t(u) = \mathbb{E}^{\mathbb{Q}}\left[v_u \,|\, \mathcal{F}_t\right]$ are (at least in principle) tradable and observed in the market.

- $\xi_t(u)$ is the product of two terms:
- $\mathbb{E}^{\mathbb{P}}\left[v_u \,|\, \mathcal{F}_t\right]$ which depends on the historical path $\{W_s, s
- a term which depends on the price of risk $\lambda(s)$.

Features of the rough Bergomi model

- The rBergomi model is a non-Markovian generalization of the Bergomi model:

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

  - The rBergomi model is Markovian in the (infinite-dimensional) state vector $\mathbb{E}^{\mathbb{Q}}\left[v_u \,|\, \mathcal{F}_t\right] = \xi_t(u)$.

- We have achieved our earlier aim of replacing the exponential kernels in the Bergomi model with a power-law kernel.

- We may therefore expect that the rBergomi model will generate a realistic term structure of ATM volatility skew.

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 $\xi_t(u)$.
  - $\xi_t(u) = \mathbb{E}\left[v_u \,|\, \mathcal{F}_t\right]$ 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
  - The exponential kernels in the exponent of the conventional Bergomi model 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.

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.

- Thus
$$\frac{dS_t}{S_t} = \sqrt{v_t}\, dZ_t$$
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 rBergomi model

- In [Bayer, Friz and Gatheral][2]</sup>, we performed an exact simulation of the Volterra process $\tilde{W}$.
- This simulation was very slow!

Hybrid simulation of BSS processes

- The Rough Bergomi variance process is a special case of a Brownian Semistationary (BSS) process.

- [Bennedsen, Lunde and Pakkanen][4]</sup> show how to simulate such processes more efficiently.

- More recently, [McCrickerd and Pakkanen][10] show how to massively increasing the efficiency of the hybrid scheme.
  - Moreover, they provide a sample Jupyter notebook!

- Their idea is roughly as follows:

$$\begin{aligned}\int_t^u \frac{dW_s}{(u-s)^\gamma} &= \sum_{k=1}^{n} \int_{t_{k+1}}^{t_k} \frac{dW_s}{(u-s)^\gamma} \\ &\approx \sum_{k=1}^{\kappa} \int_{t_{k+1}}^{t_k} \frac{dW_s}{(u-s)^\gamma} + \sum_{k=\kappa+1}^{n} \frac{1}{(u-s_k)^\gamma} \int_{t_{k+1}}^{t_k} dW_s \\ &= \sum_{k=1}^{\kappa} \int_{t_{k+1}}^{t_k} \frac{dW_s}{(u-s)^\gamma} + \sum_{k=\kappa+1}^{n} \frac{1}{(u-s_k)^\gamma} Z_k \sqrt{\frac{u-t}{n}}\end{aligned}$$

where $t_k = u - \frac{k}{n}(u-t)$, the Z_k are iid $N(0,1)$ random variables and the s_k are such that

$$\int_{t_{k+1}}^{t_k} \frac{ds}{(u-s)^\gamma} = \frac{1}{(u-s_k)^\gamma}.$$

- The choice $\kappa = 1$ works well in practice.
- The choice $\kappa = 0$ corresponds to the Euler scheme which as expected performs poorly.

Some R-code

In [1]:
```
setwd("./LRV")
```

In [2]:
```
source("BlackScholes.R")
source("hybridSimulation.R")
source("plotIvols.R")
```

R-implementation of the hybrid scheme

In [3]:
```
hybridScheme
```

```
function (xi, params)
function(N, steps, expiries) {
    eta <- params$eta
    H <- params$H
    rho <- params$rho
    W <- matrix(rnorm(N * steps), nrow = steps, ncol = N)
    Wperp <- matrix(rnorm(N * steps), nrow = steps, ncol = N)
    Z <- rho * W + sqrt(1 - rho * rho) * Wperp
    Wtilde <- Wtilde.sim(W, Wperp, H)
    S <- function(expiry) {
        dt <- expiry/steps
        ti <- (1:steps) * dt
        Wtilde.H <- expiry^H * Wtilde
        xi.t <- xi(ti)
        v1 <- xi.t * exp(eta * Wtilde.H - 1/2 * eta^2 * ti^(2 *
            H))
        v0 <- rep(xi(0), N)
        v <- rbind(v0, v1[-steps, ])
        logs <- apply(sqrt(v * dt) * Z - v/2 * dt, 2, sum)
        s <- exp(logs)
        return(s)
    }
    st <- t(sapply(expiries, S))
    return(st)
}
```

Run the hybrid BSS scheme

We will use R parallel processing functionality.

In [4]:
```
library(foreach)
library(doParallel)
```

```
Loading required package: iterators
Loading required package: parallel
```

In [5]:
```
paths <- 1e5
steps <- 200
```

In [6]:
```
params.rBergomi <- list(H=0.05, eta=1.9, rho=-0.9)
xiCurve <- function(t){0.16^2+0*t}
```

In [7]:
```
expiries <- c(.25,1)
```

In [8]:
```
t0<-proc.time()

#number of iterations
iters<- max(1,floor(paths/1000))

#setup parallel backend
cl.num <- detectCores() # This number is 8 on my MacBook Pro
cl<-makeCluster(cl.num)
registerDoParallel(cl)

#loop
ls <- foreach(icount(iters)) %dopar% {
        hybridScheme(xiCurve,params.rBergomi)(N=1000, steps=steps, expir
ies=expiries)
        }

stopCluster(cl)
mcMatrix1 <- do.call(cbind, ls) #Bind all of the submatrices into one bi
g matrix

print(proc.time()- t0)
```

```
   user  system elapsed
  0.115   0.028   4.800
```

Plot the 3-month smile

In [9]:
```
plotSmile <- function(mcMatrix,expiries,slice)function(kmin,kmax,yrange)
{

            t <- expiries[slice]
            spots <- mcMatrix[slice, ]
            s0 <- mean(spots)
            curve(bsOut(spots, t, s0 * exp(x))$BSV,from=kmin,to=kmax,
                 xlab="Log Strike",ylab="Implied Vol.",col="red",lwd=2,y
lim=yrange)

        }
```

In [10]:
```
options(repr.plot.width=11,repr.plot.height=6)
```

```
In [11]: plotSmile(mcMatrix1,expiries,1)(-.25,.25, yrange=c(.1,.26))
```

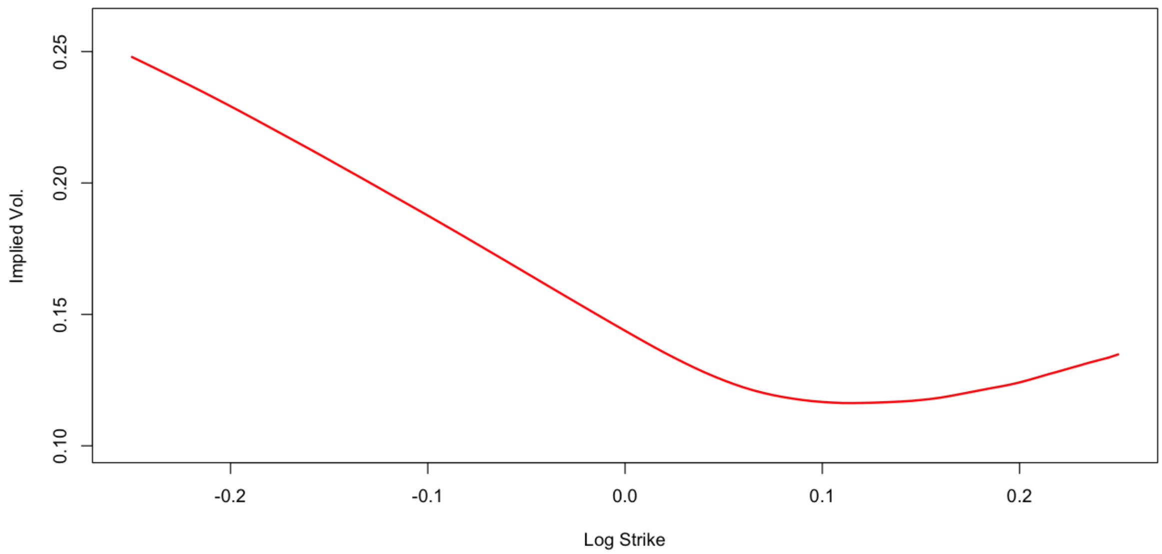

Figure 3: 3-month rough Bergomi smile with parameters `params.rBergomi`.

Guessing rBergomi model parameters

- The rBergomi model has only three parameters: H, η and ρ.
- If simulation were fast enough, we could just iterate on these parameters to find the best fit to observed option prices.
  - The BSS scheme is not yet fast enough, at least in my R implementation.

- However, the model parameters H, η and ρ have very direct interpretations:
  - H controls the decay of ATM skew $\psi(\tau)$ for very short expirations.
  - The product $\rho\,\eta$ sets the level of the ATM skew for longer expirations.
    - Keeping $\rho\,\eta$ constant but decreasing ρ (so as to make it more negative) pushes the minimum of each smile towards higher strikes.

- So we can guess parameters in practice.
  - A couple of examples of the results of guessing are given in [Bayer, Friz and Gatheral][2] </sup>.

Calibration using machine learning

- Recently, [Bayer and Stemper][3]</sup> showed how to calibrate the rough Bergomi model to the volatility surface using machine learning.
  - A neural network is trained to approximate the implied volatility map.

H from VIX futures

- Rather than brute-force fitting a rough volatility model to the volatility surface, following [Jacquier, Martini and Muguruza], one can try to fix H from the term structure of the convexity adjustment between variance swaps and VIX futures.
- Once the Volterra process $\tilde{W}$ has been simulated for this H, iterating on the parameters η and ρ to fit the observed volatility surface is relatively fast.
- The main practical trick is to fit normalized smiles
$$\hat{\sigma}(k,T) = \frac{\sigma(k,T)}{\sigma(0,T)}.$$
  ATM volatilities can then be fitted by iterating on the forward variance curve as explained above.

Rough Bergomi parameters under $\mathbb{P}$ and under $\mathbb{Q}$

- We might wonder whether implied model parameters are consistent with historical parameters.
- It is shown in [Bayer, Friz and Gatheral][2]</sup> that the volatility of volatility parameter η in the rough Bergomi model and the volatility of volatility ν in the historical time series should be related as follows.
$$\tilde{\eta} := \eta\,\sqrt{2\,H} = 2\,\nu\,C_H$$
  with
$$C_H = \sqrt{\frac{2\,H\,\Gamma(3/2-H)}{\Gamma(H+1/2)\,\Gamma(2-2\,H)}}.$$

Parameter estimates under $\mathbb{Q}$

In Section 5.2 of [Bayer, Friz and Gatheral][2]</sup>, parameter guesses for the SPX implied volatility surface on two particular dates in history are given as follows:

Date	H	η	$\tilde{\eta}$
February 4, 2010	0.07	1.9	0.7109
August 14, 2013	0.05	2.3	0.7273

- Estimates of $\tilde{\eta}$ seem more stable than estimates of η and H separately.

- We observe the same phenomenon when estimating ν and H from historical RV data.
  - Estimates of the product $\nu\,\sqrt{H}$ are more stable than estimates of the two parameters separately.

Parameter estimates under $\mathbb{P}$

- From our analysis of the SPX realized variance time series in Lecture 1, we estimated

$$H \approx 0.15, \quad \nu \approx 0.30.$$

- Plugging these estimates into the formula (from above)

$$\tilde{\eta}_1 = 2\,\nu\,\sqrt{\frac{2\,H\,\Gamma(3/2-H)}{\Gamma(H+1/2)\,\Gamma(2-2\,H)}} \approx 0.25.$$

In [12]:
```
h.est <- 0.15
nu.est <- 0.3
(nu.tilde <- 2*nu.est*sqrt(2*h.est*gamma(3/2-h.est)/gamma(h.est+1/2)*gam
ma(2-2*h.est)))
```

0.251298848335933

- Seemingly inconsistent with the implied estimate of around 0.7.

- However, the historical estimate is in daily terms and the implied estimate in annualized terms.

- To convert, we need to multiply the historical estimate by the annualization factor $(252)^H$, to get

$$\tilde{\eta} \approx \tilde{\eta}_1 \times (252)^H = 0.58.$$

  - At least by physicists' standards, the historical and implied estimates are consistent.

- It is not unexpected for implied volatility of volatility to be higher than historical to reflect the volatility of the volatility risk premium.

More rough volatility models

This form suggests many other rough volatility models of the form

$$\frac{dS_t}{S_t} = \sqrt{\xi_t(t)}\, dZ_t$$
$$d\xi_t(u) = \lambda(\xi)\, \kappa(u-t)\, dW_t$$

where both the function λ and the kernel κ depend on the model.

- As long as $\kappa(\tau) \sim \tau^{-\gamma}$ as $\tau \to 0$, the model will be rough in the sense that sample paths of instantaneous variance will be Hölder continuous with exponent $H = \frac{1}{2} - \gamma$.

Rough volatility and long memory

- In [Bennedsen, Lunde and Pakkanen][5]</sup>, the authors show how we can both have our cake and eat it by choosing different kernels.
- In particular, with appropriate choices of γ and β the kernel

$$\kappa(\tau) = \frac{1}{\tau^\gamma\, (1+\tau)^\beta}$$

  generates a model that exhibits both rough volatility and power-law decay of the autocorrelation function.
  - That is rough volatility plus long memory.
- Models with with more parameters may of course also fit the volatility surface better.

Dynamics of the volatility surface: Model dependence

- All rough stochastic volatility models have essentially the same implications for the shape of the volatility surface.
- At first it might therefore seem that it would be hard to differentiate between models.
  - That would certainly be the case if we were to confine our attention to the shape of the volatility surface today.

- If instead we were to study the dynamics of the volatility skew – in particular, how the observed volatility skew depends on the overall level of volatility, we would be able to differentiate between models.
- As explained in [The Volatility Surface][9] , we expect the ATM volatility skew to be roughly independent of the ATM volatility in a lognormal model such as rough Bergomi.
- In Figure 4, we see how the ATM skew varies with ATM volatility under rough Bergomi, with the above parameters and compare with empirical estimates.

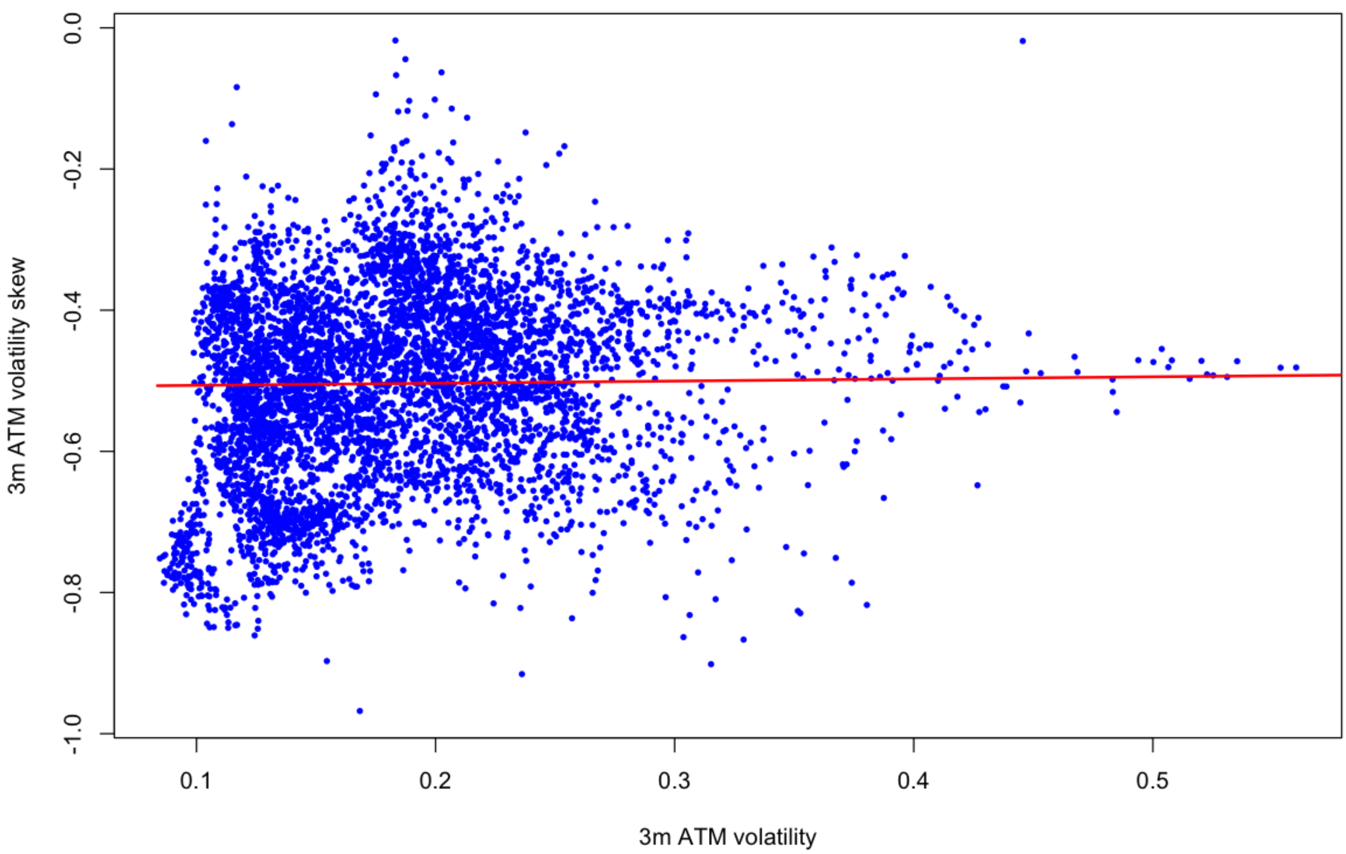

Figure 4: Blue points are empirical 3-month ATM volatilities and skews (from Jan-1996 to today); the red line is the rough Bergomi computation with the above parameters.

Summary

- In Lecture 1, scaling properties of the time series of historical volatility suggested a natural non-Markovian stochastic volatility model under $\mathbb{P}$.
- The simplest specification of $\frac{d\mathbb{Q}}{d\mathbb{P}}$ gives the rough Bergomi model, a non-Markovian generalization of the Bergomi model.
  - The history of the Brownian motion $\lbrace W_s, s
  - Efficient computations are possible using the hybrid BSS scheme.

- Rough Bergomi is a simple tractable stochastic volatility model consistent with both the historical time series of volatility and the implied volatility surface.
  - Moreover, rough Bergomi dynamics seem to be reasonable.

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. ^ Christian Bayer, Peter Friz and Jim Gatheral, Pricing under rough volatility, *Quantitative Finance* **16**(6) 887-904 (2016).
3. ^ Christian Bayer, and Benjamin Stemper, Deep calibration of rough stochastic volatility models, available at https://arxiv.org/abs/1810.03399, (2018).
4. ^ Mikkel Bennedsen, Asger Lunde, and Mikko S. Pakkanen, Hybrid scheme for Brownian semistationary processes, *Finance and Stochastics* **21**(4) 931-965 (2017).
5. ^ Mikkel Bennedsen, Asger Lunde, and Mikko S. Pakkanen, Decoupling the short-and long-term behavior of stochastic volatility, available at https://arxiv.org/abs/1610.00332, (2016).
6. ^ Lorenzo Bergomi and Julien Guyon, Stochastic Volatility's Orderly Smiles, *Risk Magazine* 60-66 (May 2012).
7. ^ Omar El Euch and Mathieu Rosenbaum, Perfect hedging in rough Heston models, *The Annals of Applied Probability* **28**(6) 3813-3856 (2018).
8. ^ Masaaki Fukasawa, Asymptotic analysis for stochastic volatility: Martingale expansion, *Finance and Stochastics* **15** 635-654 (2011).
9. ^Jim Gatheral, *The Volatility Surface: A Practitioner's Guide*, John Wiley and Sons, Hoboken, NJ (2006).
10. ^ Ryan McCrickerd and Mikko S Pakkanen, Turbocharging Monte Carlo Pricing for the Rough Bergomi Model, *Quantitative Finance* **18**(11) 1877-1886 (2018).


```
In [ ]:
```