Recent Advances in Fractional Stochastic Volatility Models

Alexandra Chronopoulou

Industrial & Enterprise Systems Engineering
University of Illinois at Urbana-Champaign

IPAM National Meeting
of Women in Financial Mathematics

April 28, 2017

Stochastic Volatility

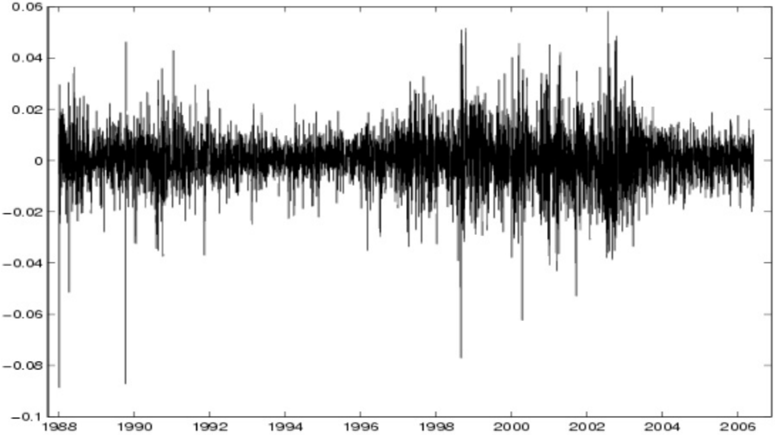
- $\{S_t; t \in [0,T]\}$: Stock Price Process

$$dS_t = \mu S_t dt + \sigma(X_t)\, S_t\, dW_t,$$

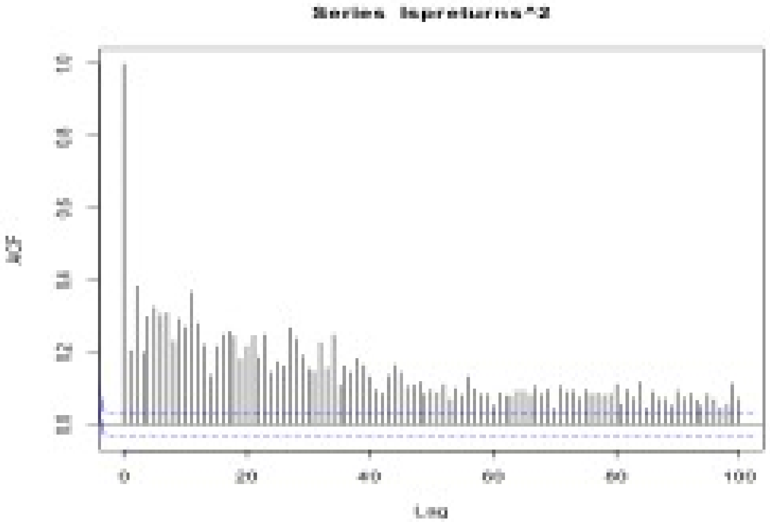
  $\sigma(\cdot)$ is a deterministic function and X_t is a stochastic process.

- Assumption: X_t is modeled by a diffusion driven by noise other than W.
  - *Log-Normal*: $dX_t = c_1\, X_t\, dt + c_2\, X_t\, dZ_t$,
  - *Mean Reverting OU*: $dX_t = \alpha\,(m - X_t)\, dt + \beta\, dZ_t$,
  - *Feller/Cox–Ingersoll–Ross (CIR)*: $dX_t = k\,(\nu - X_t)\, dt + v\sqrt{X_t}\, dZ_t$.

Stylized Fact: Volatility Persistence



ACF of S&P 500 Returns^2



Stochastic Process with Long Memory

Long-memory

If $\{X_t\}_{t\in\mathbb{N}}$ is a stationary process and there exists $H \in (\frac{1}{2}, 1)$ such that

$$\lim_{t\to\infty} \frac{Corr(X_t, X_1)}{c\, t^{2-2H}} = 1$$

then $\{X_t\}$ has long memory (long-range dependence).

Equivalently,

$$\begin{cases} \text{Long Memory:} \quad \sum_{t=1}^{\infty} Corr(X_t, X_1) = \infty, \ \text{when } H > 1/2 \\ \text{Antipersistence:} \quad \sum_{t=1}^{\infty} Corr(X_t, X_1) < \infty, \ \text{when } H < 1/2 \end{cases}$$

Stochastic Process with Long Memory

Long-memory

If $\{X_t\}_{t\in\mathbb{N}}$ is a stationary process and there exists $H \in (\frac{1}{2}, 1)$ such that

$$\lim_{t\to\infty} \frac{Corr(X_t, X_1)}{c\, t^{2-2H}} = 1$$

then $\{X_t\}$ has long memory (long-range dependence).

Equivalently,

$$\begin{cases} \text{Long Memory: } \sum_{t=1}^{\infty} Corr(X_t, X_1) = \infty, \text{ when } H > 1/2 \\ \text{Antipersistence: } \sum_{t=1}^{\infty} Corr(X_t, X_1) < \infty, \text{ when } H < 1/2 \end{cases}$$

Fractional Stochastic Volatility Model

Log-Returns: $\{Y_t, t \in [0,T]\}$

$$dY_t = \left(r - \frac{{\sigma_t}^2}{2}\right)\, dt + \sigma_t\, dW_t.$$

where $\sigma_t = \sigma(X_t)$ and X_t is described by:

$$dX_t = \alpha\,(m - X_t)\, dt + \beta\, dB_t^H$$

– B_t^H is a fractional Brownian motion with Hurst parameter $H \in (0,1)$.
– X_t is a fractional Ornstein-Uhlenbeck process (fOU).

Fractional Stochastic Volatility Model

Log-Returns: $\{Y_t, t \in [0,T]\}$

$$dY_t = \left(r - \frac{{\sigma_t}^2}{2} \right) \, dt + \sigma_t \, dW_t.$$

where $\sigma_t = \sigma(X_t)$ and X_t is described by:

$$dX_t = \alpha \, (m - X_t) \, dt + \beta \, dB_t^H$$

- B_t^H is a fractional Brownian motion with Hurst parameter $H \in (0,1)$.
- X_t is a fractional Ornstein-Uhlenbeck process (fOU).

Fractional Brownian Motion

Definition

A centered Gaussian process $B^H = \{B_t^H, t \geq 0\}$ is called fractional Brownian motion (fBm) with selfsimilarity parameter $H \in (0,1)$, if it has the following covariance function

$$\mathbf{E}\left(B_t^H B_s^H\right) = \frac{1}{2}\left\{t^{2H} + s^{2H} - |t-s|^{2H}\right\}.$$

and a.s. continuous paths.

– When $H = \frac{1}{2}$, B^H is a **standard** Brownian motion.

Fractional Brownian Motion

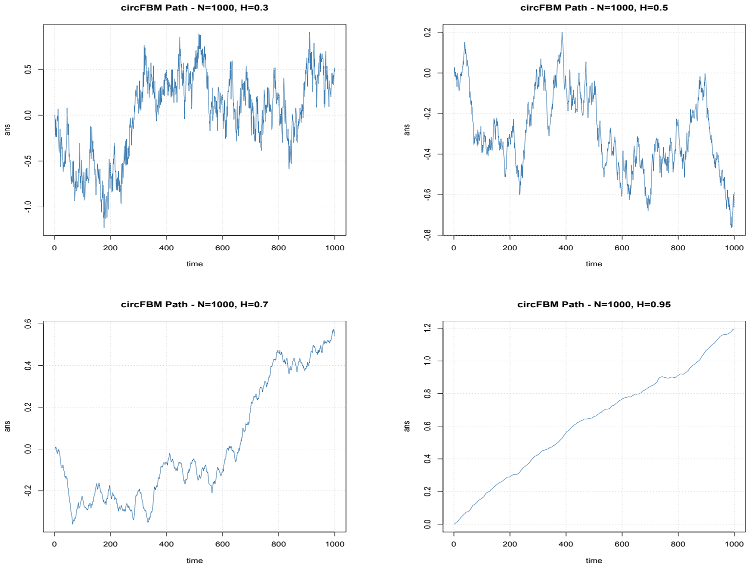
Definition

A centered Gaussian process $B^H = \{B_t^H, t \geq 0\}$ is called fractional Brownian motion (fBm) with selfsimilarity parameter $H \in (0,1)$, if it has the following covariance function

$$\mathbf{E}\left(B_t^H B_s^H\right) = \frac{1}{2}\left\{t^{2H} + s^{2H} - |t-s|^{2H}\right\}.$$

and a.s. continuous paths.

- When $H = \frac{1}{2}$, B^H is a **standard** Brownian motion.


circFBM Path - N=1000, H=0.3
circFBM Path - N=1000, H=0.5
circFBM Path - N=1000, H=0.7
circFBM Path - N=1000, H=0.95
ans
time

Properties of FBM

The increments of fBm, $\{B_n - B_{n-1}\}_{n\in\mathbb{N}}$, are

- *stationary*, i.e. $\mathbb{E}[(B_n - B_{n-1})\,(B_{n+h} - B_{n+h-1})] = \gamma(h)$.
- *H–selfsimilar*, i.e. $c^{-H}\,(B_n - B_{n-1}) \sim^{\mathcal{D}} (B_{c\,n} - B_{c(n-1)})$
- *dependent*

Properties of FBM

The increments of fBm, $\{B_n - B_{n-1}\}_{n\in\mathbb{N}}$, are

- *stationary*, i.e. $\mathbb{E}[(B_n - B_{n-1})\,(B_{n+h} - B_{n+h-1})] = \gamma(h)$.
- *H–selfsimilar*, i.e. $c^{-H}\,(B_n - B_{n-1}) \sim^{\mathcal{D}} (B_{c\,n} - B_{c(n-1)})$
- *dependent*
  - When $H > \frac{1}{2}$, the increments exhibit *long-memory*, i.e.

$$\sum \mathbb{E}\left[B_1\,(B_n - B_{n-1})\right] = +\infty.$$

  - When $H < \frac{1}{2}$, the increments exhibit *antipersistence*, i.e.

$$\sum \mathbb{E}[B_1\,(B_n - B_{n-1})] < +\infty.$$

Properties of FBM

The increments of fBm, $\{B_n - B_{n-1}\}_{n\in\mathbb{N}}$, are

- *stationary*, i.e. $\mathbb{E}[(B_n - B_{n-1})\,(B_{n+h} - B_{n+h-1})] = \gamma(h)$.
- *H–selfsimilar*, i.e. $c^{-H}\,(B_n - B_{n-1}) \sim^{\mathcal{D}} (B_{c\,n} - B_{c(n-1)})$
- *dependent*
  - When $H > \frac{1}{2}$, the increments exhibit *long-memory*, i.e.

$$\sum \mathbb{E}\left[B_1\left(B_n - B_{n-1}\right)\right] = +\infty.$$

  - When $H < \frac{1}{2}$, the increments exhibit *antipersistence*, i.e.

$$\sum \mathbb{E}\left[B_1\left(B_n - B_{n-1}\right)\right] < +\infty.$$

Properties of FBM

The increments of fBm, $\{B_n - B_{n-1}\}_{n\in\mathbb{N}}$, are

- *stationary*, i.e. $\mathbb{E}[(B_n - B_{n-1})\,(B_{n+h} - B_{n+h-1})] = \gamma(h)$.
- *H–selfsimilar*, i.e. $c^{-H}\,(B_n - B_{n-1}) \sim^{\mathcal{D}} (B_{c\,n} - B_{c(n-1)})$
- *dependent*
  - When $H > \frac{1}{2}$, the increments exhibit *long-memory*, i.e.

$$\sum \mathbb{E}\left[B_1\left(B_n - B_{n-1}\right)\right] = +\infty.$$

  - When $H < \frac{1}{2}$, the increments exhibit *antipersistence*, i.e.

$$\sum \mathbb{E}\left[B_1\left(B_n - B_{n-1}\right)\right] < +\infty.$$

Properties of FBM

The increments of fBm, $\{B_n - B_{n-1}\}_{n\in\mathbb{N}}$, are

- *stationary*, i.e. $\mathbb{E}[(B_n - B_{n-1})\,(B_{n+h} - B_{n+h-1})] = \gamma(h)$.
- *H–selfsimilar*, i.e. $c^{-H}\,(B_n - B_{n-1}) \sim^{\mathcal{D}} (B_{c\,n} - B_{c(n-1)})$
- *dependent*
  - When $H > \frac{1}{2}$, the increments exhibit *long-memory*, i.e.

$$\sum \mathbb{E}\left[B_1\left(B_n - B_{n-1}\right)\right] = +\infty.$$

  - When $H < \frac{1}{2}$, the increments exhibit *antipersistence*, i.e.

$$\sum \mathbb{E}\left[B_1\left(B_n - B_{n-1}\right)\right] < +\infty.$$

Properties of FBM

- It has a.s. *Hölder-continuous* sample paths of any order $\gamma < H$.

- Its $\frac{1}{H}$-variation on $[0, t]$ is finite. In particular, fBm has an infinite quadratic variation for $H < 1/2$.

- When $H \neq \frac{1}{2}$, B_t^H is not a semimartingale.

Properties of FBM

- It has a.s. *Hölder-continuous* sample paths of any order $\gamma < H$.
- Its $\frac{1}{H}$-variation on $[0, t]$ is finite. In particular, fBm has an infinite quadratic variation for $H < 1/2$.
- When $H \neq \frac{1}{2}$, B_t^H is not a semimartingale.

Properties of FBM

- It has a.s. *Hölder-continuous* sample paths of any order $\gamma < H$.
- Its $\frac{1}{H}$-variation on $[0, t]$ is finite. In particular, fBm has an infinite quadratic variation for $H < 1/2$.
- When $H \neq \frac{1}{2}$, B_t^H is not a semimartingale.

Stochastic Integral wrt fBm

- *Pathwise Riemann-Stieltjes integral*, when $H > 1/2$. (Lin; Dai and Heyde).
- *Stochastic calculus of variations* with respect to a Gaussian process. (Decreusefond and Üstünel; Carmona and Coutin; Alòs, Mazet and Nualart; Duncan, Hu and Pasik-Duncan and Hu and Oksendal).
- *Pathwise* stochastic integral interpreted in the Young sense, when $H > 1/2$ (Young; Gubinelli).
- *Rough path*-theoretic approach by T. Lyons.

Fractional Stochastic Volatility Model

$$\begin{cases} dY_t &= \left(r - \frac{\sigma^2(X_t)}{2}\right)\, dt + \sigma(X_t)\, dW_t, \\ dX_t &= \alpha\,(m - X_t)\, dt + \beta\, dB_t^H \end{cases}$$

Some Properties

- *Hölder continuity*: Hölder continuous of order γ, for all $\gamma < H$.
- *Self-similarity*: Self-similar in the sense that

$$\{B_{ct}^H; t \in \mathbb{R}\} \sim^{\mathcal{D}} \{c^H B_t^H; t \in \mathbb{R}\}, \quad \forall c \in \mathbb{R}$$

This property is approximately inherited by the fOU, for scales smaller than $1/\alpha$.

Fractional Stochastic Volatility Model

Hurst Index	Model
$H > 1/2$	Long-Memory SV: *persistence* ACF Decay $\sim dt^{2H-2}$
$H < 1/2$	Rough Volatility: *anti-persistence* ACF Decay $\sim dt^{H}$
$H = 1/2$	Classical SV

Long Memory Stochastic Volatility Models

- Comte and Renault (1998)
- Comte, Coutin and Renault (2010)
- C. and Viens (2010, 2012)
- Gulisashvili, Viens and Zhang (2015)
- Guennoun, Jacquier, and Roome (2015)
- Garnier & Solna (2015, 2016)
- Bezborodov, Di Persio & Mishura (20176)
- Fouque & Hu (2017)

Rough Stochastic Volatility Models

- Gatheral, Jaisson, and Rosenbaum (2014)
- Bayer, Friz, and Gatheral (2015)
- Forde, Zhang (2015)
- El Euch, Rosenbaum (2016, 2017)

Related Literature

- Willinger, Taqqu, Teverovsky (1999): LRD in the stock market
- Bayraktar, Poor, Sircar (2003): Estimation of fractal dimension of S&P 500.
- Björk, Hult (2005): Fractional Black-Scholes market.
- Cheriditio (2003): Arbitrage in fractional BS market.
- Guasoni (2006): No arbitral under transaction cost with fBm.

Research Questions

1. Option Pricing
2. Statistical Inference
3. Hedging

Option Pricing

Option Pricing

- In practice, we have access to discrete-time observations of historical stock prices, while the volatility is *unobserved*.

Two-steps

1. Estimate the *empirical stochastic volatility distribution.*

Option Pricing

– In practice, we have access to discrete-time observations of historical stock prices, while the volatility is *unobserved*.

Two-steps

1. Estimate the *empirical stochastic volatility distribution*.

2. Construct a *multinomial recombining tree* to compute the option prices.

Option Pricing

- In practice, we have access to discrete-time observations of historical stock prices, while the volatility is *unobserved*.

Two-steps

1. Estimate the *empirical stochastic volatility distribution*.
   - Through adjusted particle filtering algorithms.
2. Construct a *multinomial recombining tree* to compute the option prices.

Option Pricing

- In practice, we have access to discrete-time observations of historical stock prices, while the volatility is *unobserved*.

Two-steps

1. Estimate the *empirical stochastic volatility distribution*.
   - Through adjusted particle filtering algorithms.
2. Construct a *multinomial recombining tree* to compute the option prices.

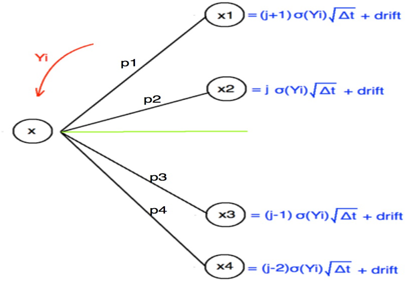
Option Pricing

- In practice, we have access to discrete-time observations of historical stock prices, while the volatility is *unobserved*.

Two-steps

1. Estimate the *empirical stochastic volatility distribution*.
   - Through adjusted particle filtering algorithms.
2. Construct a *multinomial recombining tree* to compute the option prices.

Multinomial Recombining Tree

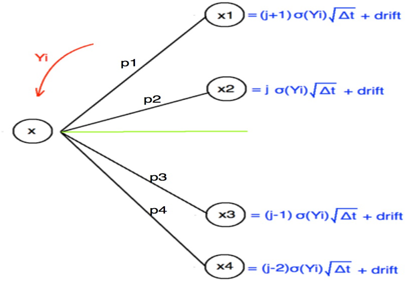


(i) At each step, *a value for the volatility is drawn from the volatility filter*.

(ii) A standard pricing technique using *backward induction* can be used to compute the option price.

(iii) We iterate this procedure by using N repeated volatility samples, constructing a different tree with each sample and averaging over all prices.

Multinomial Recombining Tree

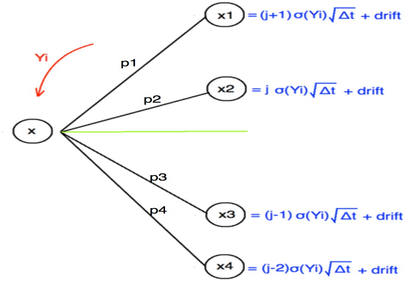


(i) At each step, *a value for the volatility is drawn from the volatility filter*.

(ii) A standard pricing technique using *backward induction* can be used to compute the option price.

(iii) We iterate this procedure by using N repeated volatility samples, constructing a different tree with each sample and averaging over all prices.

Multinomial Recombining Tree

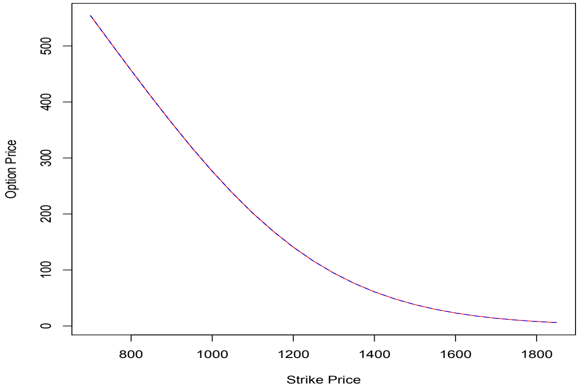


(i) At each step, *a value for the volatility is drawn from the volatility filter*.

(ii) A standard pricing technique using *backward induction* can be used to compute the option price.

(iii) We iterate this procedure by using N repeated volatility samples, constructing a different tree with each sample and averaging over all prices.

S&P 500 Data


Option Price
0
100
200
300
400
500
800
1000
1200
1400
1600
1800
Strike Price

Statistical Inference

Inference for FSV Models

$$\begin{cases} dY_t &= \left(\mu - \frac{\sigma^2(X_t)}{2}\right) dt + \sigma(X_t)\, dW_t, \\ dX_t &= \alpha\,(m - X_t)\, dt + \beta\, dB_t^H \end{cases}$$

- We can also assume that $Corr(B_t^H, W_t) = \rho$ (leverage effects).
- Parameters to estimate: $\theta = (\alpha, m, \beta, \mu, \rho)$ and H.

Inference for FSV Models

Remark

- The estimation of H is decoupled from the estimation of the drift components, but not from the estimation of the "diffusion" terms.

Framework

- Observations:
  Historical stock prices $\rightarrow$ Discrete, even when in high-frequency.
- Unobserved State:
  Stochastic Volatility with non-Markovian structure is hidden.

Inference for FSV Models

Extension of classical statistical methods

- For H known:

  μ, m, α and β can be estimated with standard techniques (Fouque, Papanicolaou and Sircar, 2000) using *high-frequency data*.

- For H unknown:

  - Traditional volatility proxies: Squared returns, logarithm of squared returns.
  - Use a non-parametric method to estimate H through the squared returns and then go back to classical techniques and estimate the remaining parameters.

Inference for FSV Models

Extension of classical statistical methods

– For H known:

μ, m, α and β can be estimated with standard techniques (Fouque, Papanicolaou and Sircar, 2000) using *high-frequency data*.

– For H unknown:

- Traditional volatility proxies: Squared returns, logarithm of squared returns.
- Use a non-parametric method to estimate H through the squared returns and then go back to classical techniques and estimate the remaining parameters.

Inference for FSV Models

Extension of classical statistical methods

– For H known:

μ, m, α and β can be estimated with standard techniques (Fouque, Papanicolaou and Sircar, 2000) using *high-frequency data*.

– For H unknown:

- Traditional volatility proxies: Squared returns, logarithm of squared returns.
- Use a non-parametric method to estimate H through the squared returns and then go back to classical techniques and estimate the remaining parameters.

Inference for FSV Models

Extension of classical statistical methods

– For H known:

μ, m, α and β can be estimated with standard techniques (Fouque, Papanicolaou and Sircar, 2000) using *high-frequency data*.

– For H unknown:

- Traditional volatility proxies: Squared returns, logarithm of squared returns.
- Use a non-parametric method to estimate H through the squared returns and then go back to classical techniques and estimate the remaining parameters.

Inference for FSV Models

Simulation Based Methods

- Employ a Sequential Monte Carlo (SMC) method to estimate the unobserved state along with the unknown parameters.
- Denote θ the vector of all parameters, except for H.
  - Observation equation: $f(Y_t|X_t;\theta)$
  - State equation: $f(X_t|X_{t-1},\ldots,X_1;\theta)$
- *Key Idea*: Sequentially compute

$$f(X_{1:t};\theta|Y_{1:t}) \propto \quad f(X_1)\cdot f(X_2|X_1;\theta)\cdot\ldots\cdot f(X_n|X_{n-1},\ldots,X_1;\theta)$$
$$\cdot\prod_{i=1}^{t} f(Y_i|X_i;\theta)\cdot f(\theta|Y_t),$$

Inference for FSV Models

Simulation Based Methods

- Employ a Sequential Monte Carlo (SMC) method to estimate the unobserved state along with the unknown parameters.
- Denote θ the vector of all parameters, except for H.
  - Observation equation: $f(Y_t|X_t;\theta)$
  - State equation: $f(X_t|X_{t-1},\ldots,X_1;\theta)$
- *Key Idea*: Sequentially compute

$$f(X_{1:t};\theta|Y_{1:t}) \propto \quad f(X_1)\cdot f(X_2|X_1;\theta)\cdot\ldots\cdot f(X_n|X_{n-1},\ldots,X_1;\theta)$$
$$\cdot\prod_{i=1}^{t} f(Y_i|X_i;\theta)\cdot f(\theta|Y_t),$$

Inference for FSV Models

Simulation Based Methods

- Employ a Sequential Monte Carlo (SMC) method to estimate the unobserved state along with the unknown parameters.
- Denote θ the vector of all parameters, except for H.
  - Observation equation: $f(Y_t|X_t;\theta)$
  - State equation: $f(X_t|X_{t-1},\dots,X_1;\theta)$
- *Key Idea*: Sequentially compute

$$f(X_{1:t};\theta|Y_{1:t}) \propto f(X_1)\cdot f(X_2|X_1;\theta)\cdot\ldots\cdot f(X_n|X_{n-1},\dots,X_1;\theta) \cdot \prod_{i=1}^{t} f(Y_i|X_i;\theta)\cdot f(\theta|Y_t),$$

Inference for FSV Models

Simulation Based Methods

- Employ a Sequential Monte Carlo (SMC) method to estimate the unobserved state along with the unknown parameters.
- Denote θ the vector of all parameters, except for H.
  - Observation equation: $f(Y_t|X_t;\theta)$
  - State equation: $f(X_t|X_{t-1},\ldots,X_1;\theta)$
- *Key Idea*: Sequentially compute

$$f(X_{1:t};\theta|Y_{1:t}) \propto \quad f(X_1)\cdot f(X_2|X_1;\theta)\cdot\ldots\cdot f(X_n|X_{n-1},\ldots,X_1;\theta) \cdot \prod_{i=1}^{t} f(Y_i|X_i;\theta)\cdot f(\theta|Y_t),$$

Inference for FSV Models

Simulation Based Methods

- Employ a Sequential Monte Carlo (SMC) method to estimate the unobserved state along with the unknown parameters.
- Denote θ the vector of all parameters, except for H.
  - Observation equation: $f(Y_t|X_t;\theta)$
  - State equation: $f(X_t|X_{t-1},\ldots,X_1;\theta)$
- *Key Idea*: Sequentially compute

$$f(X_{1:t};\theta|Y_{1:t}) \propto \quad f(X_1)\cdot f(X_2|X_1;\theta)\cdot\ldots\cdot f(X_n|X_{n-1},\ldots,X_1;\theta) \cdot \prod_{i=1}^{t} f(Y_i|X_i;\theta)\cdot f(\theta|Y_t),$$

Learning θ sequentially

Filtering for states and parameter(s): *Learning X_t and θ sequentially*.

$$\text{Posterior at } t: f(X_t|\theta, Y_t)\ f(\theta|Y_t)$$
$$\Downarrow$$
$$\text{Prior at } t+1: f(X_{t+1}|\theta, Y_t)\ f(\theta|Y_t)$$
$$\Downarrow$$
$$\text{Posterior at } t+1: f(X_{t+1}|\theta, Y_{t+1})\ f(\theta|Y_{t+1})$$

Advantages

1. Sequential updates of $f(\theta|Y_t)$, $f(X_t|Y_t)$ and $f(\theta, X_t|Y_t)$.
2. Sequential h-step ahead forecasts $f(Y_{t+h}|Y_t)$
3. Sequential approximations for $f(Y_t|Y_{t-1})$.

Learning θ sequentially

Filtering for states and parameter(s): *Learning X_t and θ sequentially*.

$$
\begin{array}{c}
\text{Posterior at } t : f(X_t|\theta, Y_t)\ f(\theta|Y_t) \\
\Downarrow \\
\text{Prior at } t+1 : f(X_{t+1}|\theta, Y_t)\ f(\theta|Y_t) \\
\Downarrow \\
\text{Posterior at } t+1 : f(X_{t+1}|\theta, Y_{t+1})\ f(\theta|Y_{t+1})
\end{array}
$$

Advantages

1. Sequential updates of $f(\theta|Y_t)$, $f(X_t|Y_t)$ and $f(\theta, X_t|Y_t)$.
2. Sequential h-step ahead forecasts $f(Y_{t+h}|Y_t)$
3. Sequential approximations for $f(Y_t|Y_{t-1})$.

Learning θ sequentially

Filtering for states and parameter(s): *Learning X_t and θ sequentially*.

$$\text{Posterior at } t: f(X_t|\theta, Y_t)\ f(\theta|Y_t)$$
$$\Downarrow$$
$$\text{Prior at } t+1: f(X_{t+1}|\theta, Y_t)\ f(\theta|Y_t)$$
$$\Downarrow$$
$$\text{Posterior at } t+1: f(X_{t+1}|\theta, Y_{t+1})\ f(\theta|Y_{t+1})$$

Advantages

1. Sequential updates of $f(\theta|Y_t)$, $f(X_t|Y_t)$ and $f(\theta, X_t|Y_t)$.
2. Sequential h-step ahead forecasts $f(Y_{t+h}|Y_t)$
3. Sequential approximations for $f(Y_t|Y_{t-1})$.

Artificial Evolution of θ

- Draw θ from a mixture of Normals:
  - $\forall t$ update $f(\theta|Y_t)$
  - Compute a Monte Carlo approximation of $f(\theta|Y_t)$, by using samples $\theta_t^{(j)}$ and weights $w_t^{(j)}$.
  - Smooth kernel density approximation

$$f(\theta|Y_t) \approx \sum_{j=1}^{N} \omega_t^{(j)} \mathcal{N}(\theta|m_t^{(j)}, h^2 V_t)$$

Artificial Evolution of θ

- Draw θ from a mixture of Normals:
  - $\forall t$ update $f(\theta|Y_t)$
  - Compute a Monte Carlo approximation of $f(\theta|Y_t)$, by using samples $\theta_t^{(j)}$ and weights $w_t^{(j)}$.
  - Smooth kernel density approximation

$$f(\theta|Y_t) \approx \sum_{j=1}^{N} \omega_t^{(j)} \mathcal{N}(\theta|m_t^{(j)}, h^2 V_t)$$

Artificial Evolution of θ

- Draw θ from a mixture of Normals:
  - $\forall t$ update $f(\theta|Y_t)$
  - Compute a Monte Carlo approximation of $f(\theta|Y_t)$, by using samples $\theta_t^{(j)}$ and weights $w_t^{(j)}$.
  - Smooth kernel density approximation

$$f(\theta|Y_t) \approx \sum_{j=1}^{N} \omega_t^{(j)} \mathcal{N}(\theta|m_t^{(j)}, h^2 V_t)$$

Artificial Evolution of θ

- Draw θ from a mixture of Normals:
  - $\forall t$ update $f(\theta|Y_t)$
  - Compute a Monte Carlo approximation of $f(\theta|Y_t)$, by using samples $\theta_t^{(j)}$ and weights $w_t^{(j)}$.
  - Smooth kernel density approximation

$$f(\theta|Y_t) \approx \sum_{j=1}^{N} \omega_t^{(j)} \mathcal{N}(\theta|m_t^{(j)}, h^2 V_t)$$

Filter convergence

Let $\phi : \mathcal{X} \mapsto \mathbb{R}$ be an appropriate test function and assume that we want estimate

$$\bar{\phi}_t = \int \phi_t(x_{1:t}) p(x_{1:t}, \theta_{(t)} | Y_{1:t}) dx_{1:t} d\theta_{(t)}.$$

The SISR algorithm provides us with the estimator

$$\hat{\phi}_t^N = \int \phi_t(x_{1:t}) \pi^N(dx_{1:t}) = \sum_{i=1}^{N} W_t^i \phi_t \left(X_{1:t-1,t-1}^{(i)}, \tilde{X}_{t,t}^{(i)} \right)$$

CLT for the filter (C. and Spiliopoulos)

$$\sqrt{N} \left(\hat{\phi}_t^N - \bar{\phi}_t \right) \Rightarrow \mathcal{N}\left(0, \sigma^2(\phi_t)\right)$$

as $N \to \infty$.

Convergence of the Parameter

$$\bar{\theta}^N_{(t)} = \sum_{i=1}^{N} W_t^{(i)} \theta_{(t)}^{(N,i)}, \quad \text{where} \quad \theta_{(t)}^{(N,i)} \sim \mathcal{N}(m_{t-1}^{(N,i)} | h^2 V_{t-1}^N)$$

with $m_{t-1}^N = \alpha \theta_{(t-1)}^{(N,i)} + (1-\alpha)\bar{\theta}^N_{(t-1)}$, $V_{t-1}^N = \frac{1}{N-1}\sum_{i=1}^N \left(W_{t-1}^{(i)} \theta_{(t-1)}^{(N,i)} - \bar{\theta}^N_{(t-1)} \right)^2$.

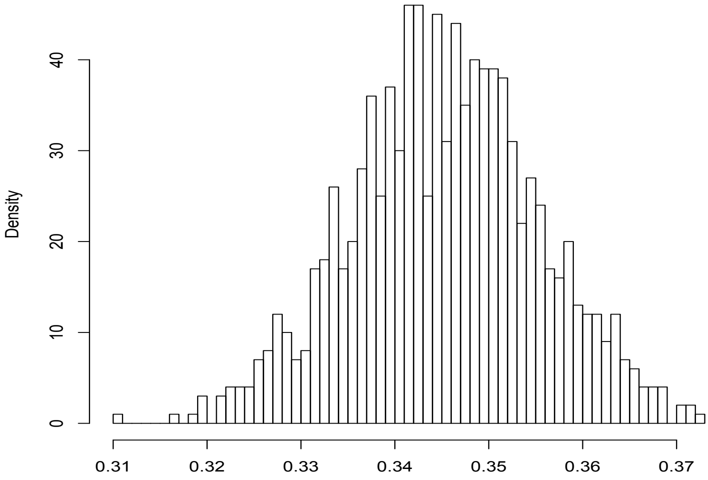
CLT for the parameter

Assuming that $\mathbb{E}_{\pi_t^\theta} \|W_t\|^{2+\delta} < \infty$:

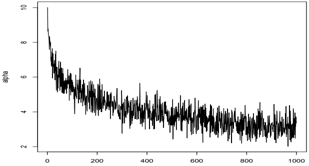
$$\sqrt{N}\left(\bar{\theta}_{(t)}^{(N)} - \bar{\theta}_{(t)} \right) \Rightarrow \mathcal{N}\left(0, \sigma^2(\theta_{(t)})\right), \quad \text{as} \quad N \to \infty \qquad (1)$$

Moreover, if the model $\mathbb{P}_\theta$ is identifiable, then the posterior mean $\bar{\theta}_{(t)}$ consistently estimates the true parameter value θ, as $t \to \infty$.

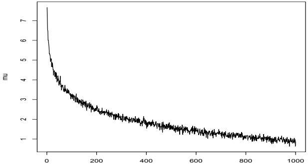
S& P 500: *Volatility Particle Filter*



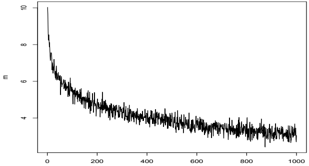
S& P 500: *Parameter Estimators*



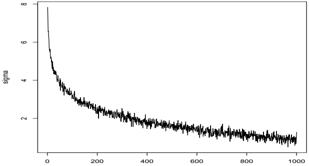
(a) Estimator of μ



(b) Estimator of α

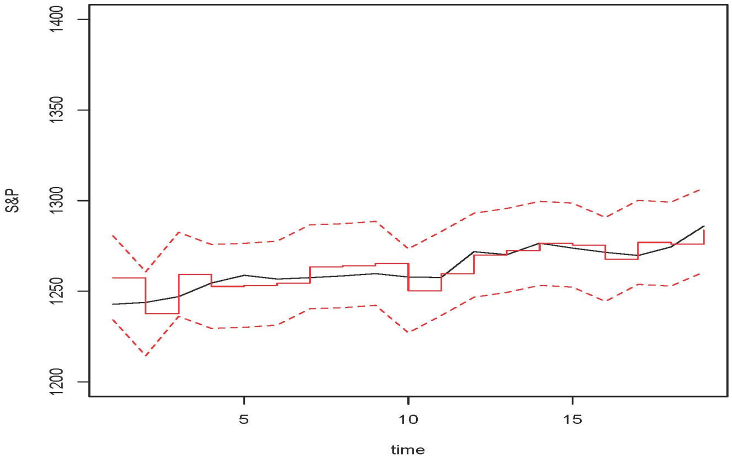


(c) Estimator of m

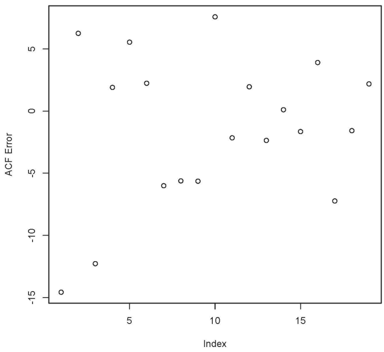


(d) Estimator of β

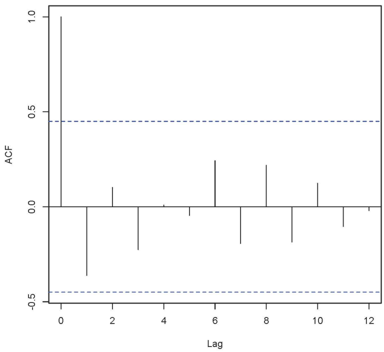
Model Validation: *1-Step Ahead Prediction*



Model Validation: *Residuals*



(a) Residuals



(b) ACF of Residuals

Hedging

Hedging

- Conditionally on the past and the *entire volatility path*

$$\ln S_T/S_t \sim \mathcal{N}\left(r - (1/2)\int_t^T \sigma_s^2 ds,\ \int_t^T \sigma_s^2 ds\right).$$

So,

$$C_t = S_t\left\{\mathbb{E}^Q\left[\Phi\left(\frac{x_t}{V_{t,T}} + \frac{V_{t,T}}{2}\right)\right] - e^{-x_t}\mathbb{E}^Q\left[\Phi\left(\frac{x_t}{V_{t,T}} - \frac{V_{t,T}}{2}\right)\right]\right\},$$

where $x_t = \ln\left(e^{-rT}S_t/K\right)$ and $V_{t,T} = \left(\int_t^T \sigma_s^2 ds\right)^{1/2}$.

- Imperfect *Delta-Sigma* hedging strategy

$$\Delta_t(x_t, \sigma_t) = \mathbb{E}^Q\left[\Phi\left(\frac{x_t}{V_{t,T}} + \frac{V_{t,T}}{2}\right)\right]$$

Hedging

- Conditionally on the past and the *entire volatility path*

$$\ln S_T/S_t \sim \mathcal{N}\left(r-(1/2)\int_t^T \sigma_s^2 ds,\ \int_t^T \sigma_s^2 ds\right).$$

  So,

$$C_t = S_t\left\{\mathbb{E}^Q\left[\Phi\left(\frac{x_t}{V_{t,T}}+\frac{V_{t,T}}{2}\right)\right]-e^{-x_t}\mathbb{E}^Q\left[\Phi\left(\frac{x_t}{V_{t,T}}-\frac{V_{t,T}}{2}\right)\right]\right\},$$

  where $x_t = \ln\left(e^{-rT}S_t/K\right)$ and $V_{t,T}=\left(\int_t^T \sigma_s^2 ds\right)^{1/2}$.

- Imperfect *Delta-Sigma* hedging strategy

$$\Delta_t(x_t,\sigma_t)=\mathbb{E}^Q\left[\Phi\left(\frac{x_t}{V_{t,T}}+\frac{V_{t,T}}{2}\right)\right]$$

Hedging

- Conditionally on the past and the *entire volatility path*

$$\ln S_T/S_t \sim \mathcal{N}\left(r-(1/2)\int_t^T \sigma_s^2 ds,\ \int_t^T \sigma_s^2 ds\right).$$

  So,

$$C_t = S_t\left\{\mathbb{E}^Q\left[\Phi\left(\frac{x_t}{V_{t,T}}+\frac{V_{t,T}}{2}\right)\right]-e^{-x_t}\mathbb{E}^Q\left[\Phi\left(\frac{x_t}{V_{t,T}}-\frac{V_{t,T}}{2}\right)\right]\right\},$$

  where $x_t = \ln\left(e^{-rT}S_t/K\right)$ and $V_{t,T} = \left(\int_t^T \sigma_s^2 ds\right)^{1/2}$.

- Imperfect *Delta-Sigma* hedging strategy

$$\Delta_t(x_t,\sigma_t) = \mathbb{E}^Q\left[\Phi\left(\frac{x_t}{V_{t,T}}+\frac{V_{t,T}}{2}\right)\right]$$

Hedging

Two notions of Implied Volatility

1. Black-Scholes Implied Volatility:
   The unique solution to

$$C_t(x, \sigma) = C_t^{BS}(x, \sigma^i(x, \sigma)),$$

   i.e. the volatility parameter that equates the BS price to the HW.

2. Hedging Volatility:
   The unique solution to

$$\Delta_t(x, \sigma) = \Delta_t^{BS}(x, \sigma^h(x, \sigma))$$

   i.e. the volatility parameter that equates the BS hedge ratio against the underlying asset variations to the HW one.

Hedging

Two notions of Implied Volatility

1. Black-Scholes Implied Volatility:
   The unique solution to

$$C_t(x,\sigma) = C_t^{BS}(x, \sigma^i(x,\sigma)),$$

   i.e. the volatility parameter that equates the BS price to the HW.

2. Hedging Volatility:
   The unique solution to

$$\Delta_t(x,\sigma) = \Delta_t^{BS}(x, \sigma^h(x,\sigma))$$

   i.e. the volatility parameter that equates the BS hedge ratio against the underlying asset variations to the HW one.

Hedging

Hedging Bias *(Definition)*

The difference between the BS implied volatility-based hedging ratio and the HW one:

$$\text{Bias } = \Delta_t^{BS}(x, \sigma^i(x,\sigma)) - \Delta_t(x,\sigma)$$

Hedging

Theorem

(a) Sign of Hedging Bias:

$$\sigma^h(-x,\sigma) \leq \sigma^i(-x,\sigma) = \sigma^i(x,\sigma) \leq \sigma^h(x,\sigma)$$

(b) Accuracy of approximation of partial hedging ratio by the BS implicit volatility-based hedging ratio:

$$\Delta^{BS}(x,\sigma) \leq \Delta(x,\sigma)$$
$$\Delta^{BS}(-x,\sigma) \geq \Delta(-x,\sigma)$$
$$\Delta^{BS}(0,\sigma) = \Delta(0,\sigma)$$

Hedging

Theorem

(a) Sign of Hedging Bias:

$$\sigma^h(-x,\sigma) \le \sigma^i(-x,\sigma) = \sigma^i(x,\sigma) \le \sigma^h(x,\sigma)$$

(b) Accuracy of approximation of partial hedging ratio by the BS implicit volatility-based hedging ratio:

$$\begin{aligned} \Delta^{BS}(x,\sigma) &\le \Delta(x,\sigma) \\ \Delta^{BS}(-x,\sigma) &\ge \Delta(-x,\sigma) \\ \Delta^{BS}(0,\sigma) &= \Delta(0,\sigma) \end{aligned}$$

Thank you!