Rough Heston models: Pricing and hedging

Mathieu Rosenbaum

École Polytechnique

Cours Bachelier 2019

Table of contents

Table of contents

A well-know stochastic volatility model

The Heston model

A very popular stochastic volatility model for a stock price is the Heston model :

$$dS_t = S_t\sqrt{V_t}dW_t$$

$$dV_t = \lambda(\theta - V_t)dt + \lambda\nu\sqrt{V_t}dB_t, \quad \langle dW_t, dB_t\rangle = \rho dt.$$

Popularity of the Heston model

- Reproduces several important features of low frequency price data : leverage effect, time-varying volatility, fat tails,. . .
- Provides quite reasonable dynamics for the volatility surface.
- Explicit formula for the characteristic function of the asset log-price$\rightarrow$ very efficient model calibration procedures.

But...

Volatility is rough !

- In Heston model, volatility follows a Brownian diffusion.
- It is shown in Gatheral *et al.* that log-volatility time series behave in fact like a fractional Brownian motion, with Hurst parameter of order 0.1.
- More precisely, basically all the statistical stylized facts of volatility are retrieved when modeling it by a rough fractional Brownian motion.
- From Alos, Fukasawa and Bayer *et al.*, we know that such model also enables us to reproduce very well the behavior of the implied volatility surface, in particular the at-the-money skew (without jumps).

Modifying Heston model

Rough Heston model

It is natural to modify Heston model and consider its rough version :

$$dS_t = S_t \sqrt{V_t} dW_t$$

$$V_t = V_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda(\theta - V_s) ds + \frac{\lambda\nu}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \sqrt{V_s} dB_s,$$

with $\langle dW_t, dB_t \rangle = \rho dt$ and $\alpha \in (1/2, 1)$.

Pricing in Heston models

Classical Heston model

From simple arguments based on the Markovian structure of the model and Ito's formula, we get that in the classical Heston model, the characteristic function of the log-price $X_t = \log(S_t/S_0)$ satisfies

$$\mathbb{E}[e^{iaX_t}] = \exp\big(g(a,t) + V_0 h(a,t)\big),$$

where h is solution of the following Riccati equation :

$$\partial_t h = \frac{1}{2}(-a^2 - ia) + \lambda(ia\rho\nu - 1)h(a,s) + \frac{(\lambda\nu)^2}{2}h^2(a,s), \quad h(a,0) = 0,$$

and

$$g(a,t) = \theta\lambda \int_0^t h(a,s)ds.$$

Table of contents

Building the model

Necessary conditions for a good microscopic price model

We want :

- A tick-by-tick model.
- A model reproducing the stylized facts of modern electronic markets in the context of high frequency trading.
- A model helping us to understand the rough dynamics of the volatility from the high frequency behavior of market participants.
- A model helping us to understand leverage effect.
- A model helping us to derive a Heston like formula and hedging strategies.

Digression : How is leverage effect generated ?

Traditional macroscopic explanations for leverage effect

- Asset price declines$\rightarrow$ company becomes automatically more leveraged since the ratio of its debt with respect to the equity value becomes larger$\rightarrow$ risk of the asset (the volatility) should become more important.
- Forecast of an increase of the volatility should be compensated by a higher rate of return, which can only be obtained through a decrease in the asset value.

Microstructural component for leverage effect ?

- We want to address the following question : Can leverage effect be partly generated from high frequency features of the asset ?

Building the model

Stylized facts 1-2

- Markets are highly endogenous, meaning that most of the orders have no real economic motivations but are rather sent by algorithms in reaction to other orders, see Bouchaud *et al.*, Filimonov and Sornette.
- Mechanisms preventing statistical arbitrages take place on high frequency markets, meaning that at the high frequency scale, building strategies that are on average profitable is hardly possible.

Building the model

Stylized facts 3-4

- There is some asymmetry in the liquidity on the bid and ask sides of the order book. In particular, a market maker is likely to raise the price by less following a buy order than to lower the price following the same size sell order, see Brennan *et al.*, Brunnermeier and Pedersen, Hendershott and Seasholes.
- A large proportion of transactions is due to large orders, called metaorders, which are not executed at once but split in time.

Building the model

Hawkes processes

- Our tick-by-tick price model is based on Hawkes processes in dimension two, very much inspired by the approaches in Bacry *et al.* and Jaisson and R.
- A two-dimensional Hawkes process is a bivariate point process $(N_t^+, N_t^-)_{t\geq 0}$ taking values in $(\mathbb{R}^+)^2$ and with intensity $(\lambda_t^+, \lambda_t^-)$ of the form :

$$\begin{pmatrix}\lambda_t^+\\ \lambda_t^-\end{pmatrix} = \begin{pmatrix}\mu^+\\ \mu^-\end{pmatrix} + \int_0^t \begin{pmatrix}\varphi_1(t-s) & \varphi_3(t-s)\\ \varphi_2(t-s) & \varphi_4(t-s)\end{pmatrix} \cdot \begin{pmatrix}dN_s^+\\ dN_s^-\end{pmatrix}.$$

Building the model

The microscopic price model

- Our model is simply given by

$$P_t = N_t^+ - N_t^- .$$

- N_t^+ corresponds to the number of upward jumps of the asset in the time interval $[0, t]$ and N_t^- to the number of downward jumps. Hence, the instantaneous probability to get an upward (downward) jump depends on the location in time of the past upward and downward jumps.
- By construction, the price process lives on a discrete grid.
- Statistical properties of this model have been studied in details.

Encoding the stylized facts

The right parametrization of the model

- Recall that

$$\begin{pmatrix} \lambda_t^+ \\ \lambda_t^- \end{pmatrix} = \begin{pmatrix} \mu^+ \\ \mu^- \end{pmatrix} + \int_0^t \begin{pmatrix} \varphi_1(t-s) & \varphi_3(t-s) \\ \varphi_2(t-s) & \varphi_4(t-s) \end{pmatrix} \cdot \begin{pmatrix} dN_s^+ \\ dN_s^- \end{pmatrix}.$$

- High degree of endogeneity of the market$\rightarrow$ L^1 norm of the largest eigenvalue of the kernel matrix close to one.
- No arbitrage$\rightarrow$ $\varphi_1 + \varphi_3 = \varphi_2 + \varphi_4$.
- Liquidity asymmetry$\rightarrow$ $\varphi_3 = \beta\varphi_2$, with $\beta > 1$.
- Metaorders splitting$\rightarrow$ $\varphi_1(x),\ \varphi_2(x) \underset{x\rightarrow\infty}{\sim} K/x^{1+\alpha}$, $\alpha \approx 0.6$.

The scaling limit of the price model

Limit theorem

After suitable scaling in time and space, the long term limit of our price model satisfies the following rough Heston dynamics :

$$P_t = \int_0^t \sqrt{V_s} dW_s - \frac{1}{2} \int_0^t V_s ds,$$

$$V_t = V_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda(\theta - V_s) ds + \frac{\lambda\nu}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \sqrt{V_s} dB_s,$$

with

$$d\langle W, B\rangle_t = \frac{1-\beta}{\sqrt{2(1+\beta^2)}} dt.$$

The scaling limit of the price model

Comments on the theorem

- The Hurst parameter $H = \alpha - 1/2$.
- Hence stylized facts of modern market microstructure naturally give rise to fractional dynamics and leverage effect.
- One of the only cases of scaling limit of a non ad hoc "micro model" where leverage effect appears in the limit. Compare with Nelson's limit of GARCH models for example.
- Uniqueness of the limiting solution is a difficult result. The proof requires the use of recent results in SPDEs theory by Mytnik and Salisbury.
- Obtaining a non-zero starting value for the volatility is a tricky point. To do so, we in fact consider a time-dependent μ.

Table of contents

A general case

Multidimensional Hawkes process

- To obtain the characteristic function of our microscopic price process, we derive the characteristic function of multidimensional Hawkes processes.
- Let us consider a d-dimensional Hawkes process $N = (N^1, ..., N^d)$ with intensity

$$\lambda_t = \begin{pmatrix} \lambda_t^1 \\ \vdots \\ \lambda_t^d \end{pmatrix} = \mu(t) + \int_0^t \phi(t-s).dN_s.$$

Multidimensional Hawkes process

Population interpretation

- Migrants of type $k \in \{1, .., d\}$ arrive as a non-homogenous Poisson process with rate $\mu_k(t)$.
- Each migrant of type $k \in \{1, .., d\}$ gives birth to children of type $j \in \{1, .., d\}$ following a non-homogenous Poisson process with rate $\phi_{j,k}(t)$.
- Each child of type $k \in \{1, .., d\}$ also gives birth to other children of type $j \in \{1, .., d\}$ following a non-homogenous Poisson process with rate $\phi_{j,k}(t)$.

Multidimensional Hawkes process

Towards the characteristic function

- Let $(\tilde{N}^{k,j})_{1\leq j\leq d}$ be d multivariate Hawkes processes with migrant rate $(\phi_{j,k})_{1\leq j\leq d}$ (for given k) and kernel matrix ϕ.
- Let $N_t^{0,k}$ be the number of migrants of type k arrived up to time t of the initial Hawkes process.
- Let $T_1^k < ... < T^k_{N_t^{0,k}} \in [0,t]$ the arrival times of migrants of type k.
- We have

$$N_t^k \underset{law}{=} N_t^{0,k} + \sum_{1\leq j\leq d} \sum_{1\leq l\leq N_t^{0,j}} \tilde{N}^{j,k,(l)}_{t-T_l^j},$$

where the $(\tilde{N}^{j,k,(l)})$ are independent copies of $(\tilde{N}^{j,k})$.

Characteristic function of multidimensional Hawkes processes

Theorem

We have

$$\mathbb{E}[\exp(ia.N_t)] = \exp\Big(\int_0^t \big(C(a, t-s) - \mathbf{1}\big).\mu(s)ds\Big),$$

where $C : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{C}^d$ is solution of the following integral equation :

$$C(a, t) = \exp\Big(ia + \int_0^t \phi^*(s).(C(a, t-s) - \mathbf{1})ds\Big).$$

Deriving the characteristic function of the rough Heston model

Strategy

- From our last theorem, we are able to derive the characteristic function of our high frequency price model.
- We then pass to the limit.

Characteristic function of rough Heston models

We write :

$$I^{1-\alpha}f(x) = \frac{1}{\Gamma(1-\alpha)}\int_0^x \frac{f(t)}{(x-t)^\alpha}dt, \;\; D^\alpha f(x) = \frac{d}{dx}I^{1-\alpha}f(x).$$

Theorem

The characteristic function at time t for the rough Heston model is given by

$$\exp\Big(\int_0^t g(a,s)ds + \frac{V_0}{\theta\lambda}I^{1-\alpha}g(a,t)\Big),$$

with $g(a,)$ the unique solution of the fractional Riccati equation :

$$D^\alpha g(a,s) = \frac{\lambda\theta}{2}(-a^2 - ia) + \lambda(ia\rho\nu - 1)g(a,s) + \frac{\lambda\nu^2}{2\theta}g^2(a,s).$$

Table of contents

Conditional law of the rough Heston model

Theorem

The law of the process $(S_t^{t_0}, V_t^{t_0})_{t\geq 0} = (S_{t+t_0}, V_{t+t_0})_{t\geq 0}$ is that of a rough Heston model with the following dynamics :

$$dS_t^{t_0} = S_t^{t_0}\sqrt{V_t^{t_0}}dW_t^{t_0}; \quad S_0^{t_0} = S_{t_0},$$

$$V_t^{t_0}=V_{t_0}+\frac{1}{\Gamma(\alpha)}\int_0^t(t-s)^{\alpha-1}\lambda\big(\theta^{t_0}(s)-V_s^{t_0}\big)ds+\frac{\lambda\nu}{\Gamma(\alpha)}\int_0^t(t-s)^{\alpha-1}\sqrt{V_s^{t_0}}dB_s^{t_0},$$

where $(W_t^{t_0}, B_t^{t_0}) = (W_{t_0+t} - W_{t_0}, B_{t_0+t} - B_{t_0})$ and θ^{t_0} is an explicit $\mathcal{F}_{t_0}$-measurable process, depending on $(V_u)_{0\leq u\leq t_0}$.

Generalized rough Heston model

Generalized rough Heston

So we naturally generalize the definition of the rough Heston model as follows :

$$dS_t = S_t \sqrt{V_t} dW_t$$

$$V_t = V_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda(\theta^0(s) - V_s) ds + \frac{\lambda \nu}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \sqrt{V_s} dB_s,$$

with $\langle dW_t, dB_t \rangle = \rho dt$, $\alpha \in (1/2, 1)$.

Characteristic function of the generalized rough-Heston model

Using the Hawkes framework, we get the following result :

Theorem

The characteristic function of $\log(S_t/S_0)$ in the generalized rough Heston model is given by

$$\exp\Big(\int_0^t h(a,t-s)(\lambda\theta^0(s)+V_0\frac{s^{-\alpha}}{\Gamma(1-\alpha)}ds)\Big),$$

where h is the unique solution of the fractional Riccati equation

$$D^\alpha h(a,t)=\frac{1}{2}(-a^2-ia)+\lambda(ia\rho\nu-1)h(a,s)+\frac{(\lambda\nu)^2}{2}h^2(a,s).$$

Link between the characteristic function and the forward variance curve

Link between θ^0 and the forward variance curve

$$\theta^0_. = D^\alpha(\mathbb{E}[V_.] - V_0) + \mathbb{E}[V_.].$$

Suitable expression for the characteristic function

The characteristic function can be written as follows :

$$\exp\Big(\int_0^t g(a, t-s)\mathbb{E}[V_s]ds\Big),$$

with

$$g(a,t) = \frac{1}{2}(-a^2 - ia) + \lambda ia\rho\nu h(a,s) + \frac{(\lambda\nu)^2}{2}h^2(a,s).$$

The suitable state variables

Recall that

$$P_t^T(a) = \mathbb{E}[\exp(ia\log(S_T))|\mathcal{F}_t].$$

The conditional law of the rough Heston model being a generalized rough Heston, we deduce the following theorem :

Theorem

$$P_t^T(a) = \exp\left(ia\log(S_t) + \int_0^{T-t} g(a,s)\mathbb{E}[V_{T-s}|\mathcal{F}_t]ds\right)$$

and

$$dP_t^T(a) = iaP_t^T(a)\frac{dS_t}{S_t} + P_t^T(a)\int_0^{T-t} g(a,s)d\mathbb{E}[V_{T-s}|\mathcal{F}_t]ds.$$

We can perfectly hedge the option with the underlying stock and the forward variance curve ! (at least theoretically)

Calibration

We collect S&P implied volatility surface, from Bloomberg, for different maturities

$$T_j = \quad 0.25, \quad 0.5, \quad 1, \quad 1.5, \quad 2 \quad \text{years},$$

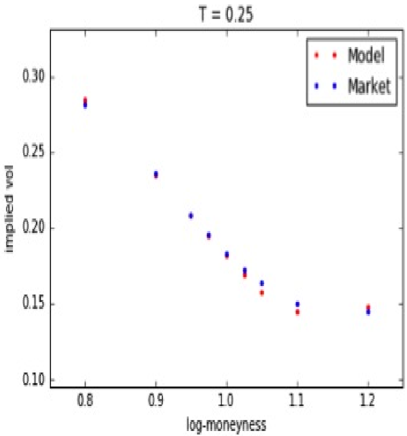
and different moneyness

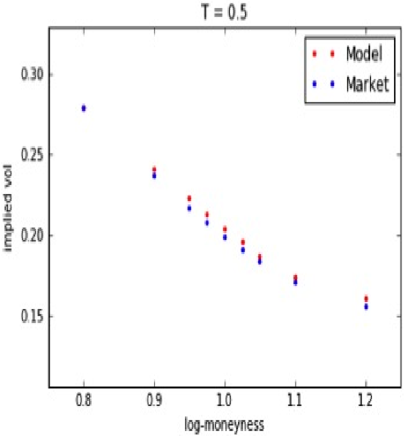
$$K/S_0 = 0.80,\ 0.90,\ 0.95,\ 0.975,\ 1.00,\ 1.025,\ 1.05,\ 1.10,\ 1.20.$$

Calibration results on data of 7 January 2010 :

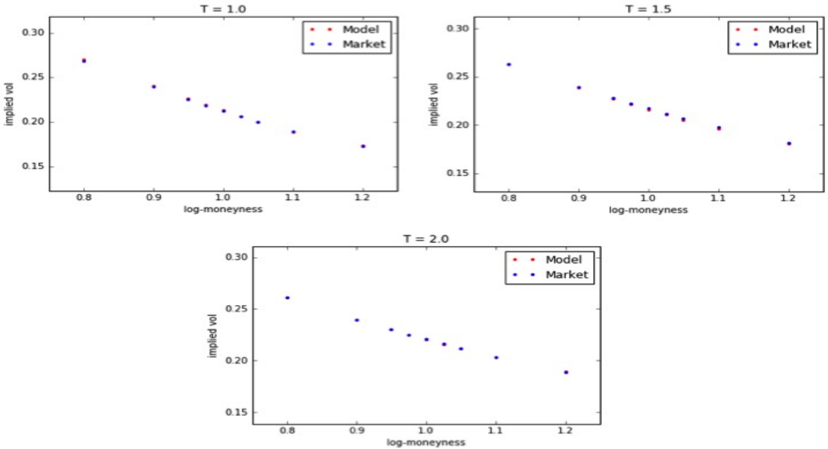
$$\rho = -0.68; \quad \nu = 0.305; \quad H = 0.09.$$

Calibration results : Market vs model implied volatilities, 7 January 2010

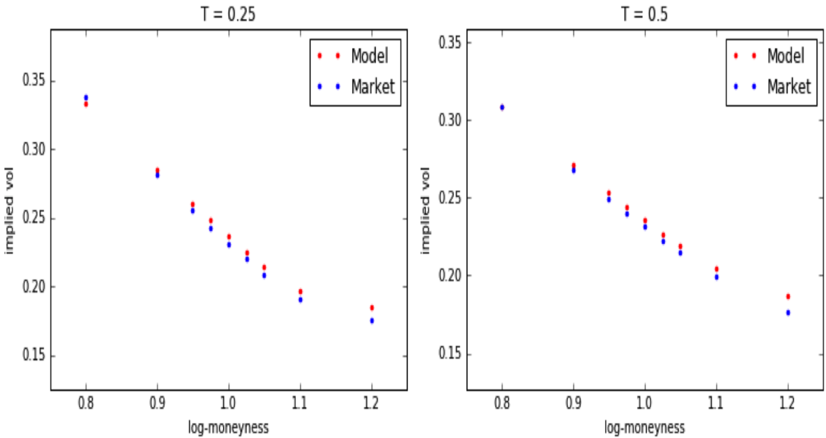




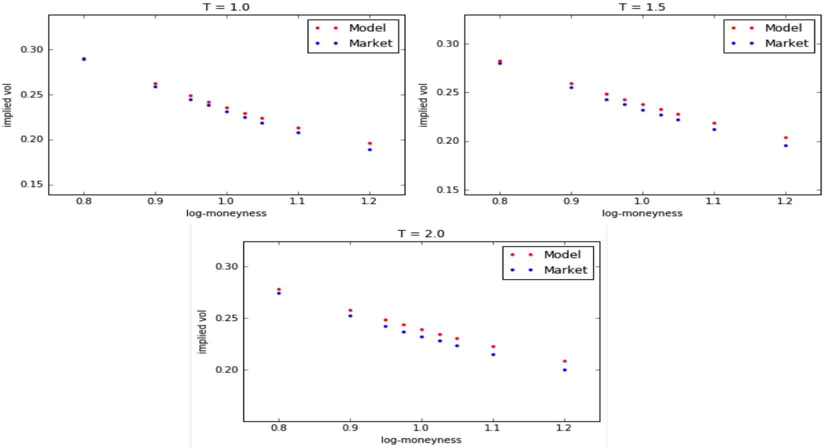
Calibration results : Market vs model implied volatilities, 7 January 2010



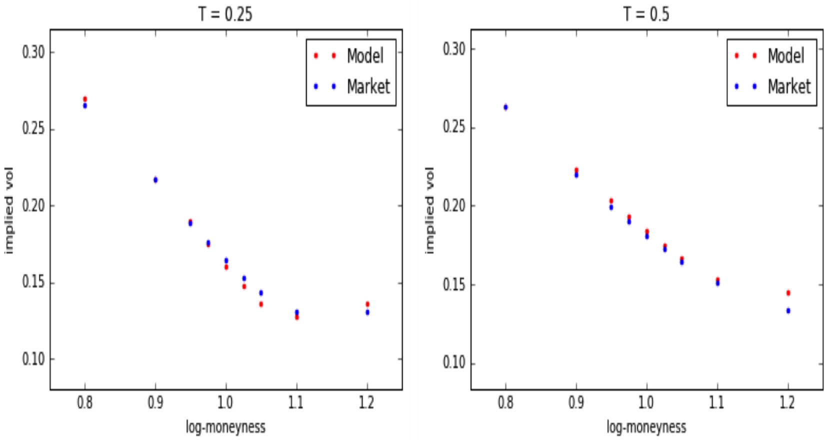
Stability : Results on 8 February 2010 (one month after calibration)



Stability : Results on 8 February 2010 (one month after calibration)



Stability : Results on 7 April 2010 (three months after calibration)



Stability : Results on 7 April 2010 (three months after calibration)