On the Joint Calibration of SPX and VIX Options

Julien Guyon

Bloomberg L.P.
Quantitative Research

QuantMinds
Lisbon, May 16, 2018

jguyon2@bloomberg.net

Motivation

- VIX options started trading in 2006.
- How to build a model for the SPX that jointly calibrates to SPX options, VIX futures, and VIX options?
- In 2008, Gatheral was one of the first to investigate this question, and showed that a diffusive model (the double mean-reverting model) could approximately match both markets.
- Later, others have argued that jumps in SPX are needed to (approximately) fit both markets.
- In this talk, I revisit this problem, trying to answer the following questions:

Does there exist a continuous model on the SPX that jointly calibrates to SPX options, VIX futures, and VIX options?

If so, how to build one such model? If not, why?

A necessary condition

Necessary condition

Price of VIX_T^2 from VIX options = Price of VIX_T^2 from SPX options

From VIX options:

- Vanilla payoff VIX_T^2 replicated using VIX future, OTM VIX calls, and OTM VIX puts (weight 2)

From SPX options:

- Price of VIX_T^2 = Price of a calendar spread of SPX log-contracts T, $T + 30$ days
- Each SPX log-contract replicated using SPX future, OTM SPX calls and OTM SPX puts (weight $1/K^2$)

Consistent extrapolations of SPX smile and VIX smile

Not a problem! Build consistent extrapolations of SPX smile and VIX smile

- Price of SPX log-contract can be made arbitrarily large: make SPX put prices linear in K when $K \to 0$:

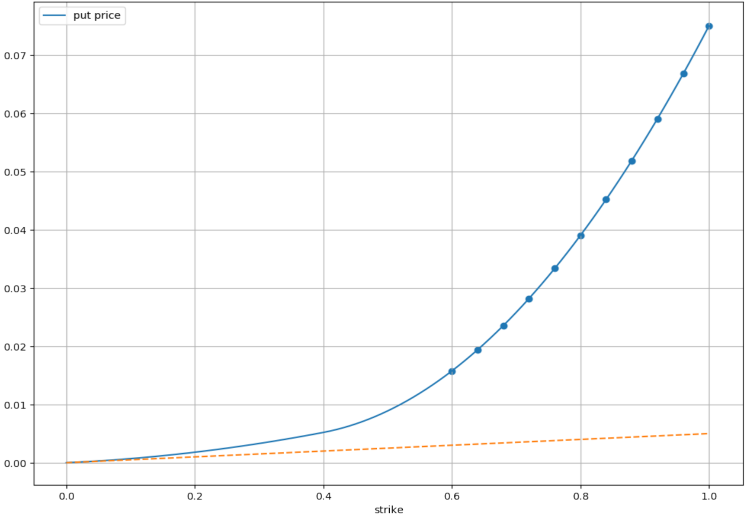
$$\int_0 P_{\mathsf{SPX}}(K)\frac{dK}{K^2} = \infty$$

- Price of VIX_T^2 can be made arbitrarily large: make $K \mapsto C_{\mathsf{VIX}}(K)$ non-integrable when $K \to \infty$:

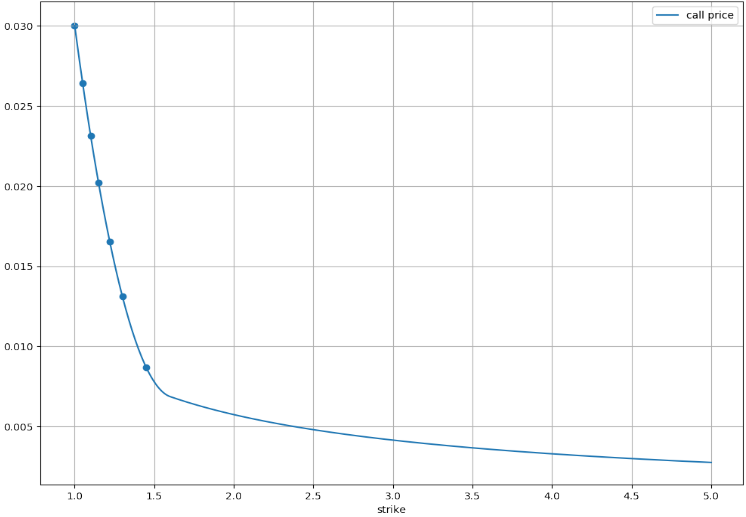
$$\int^{\infty} C_{\mathsf{VIX}}(K)\, dK = \infty$$

- Lower bounds for price of SPX log-contract and price of VIX_T^2 can also be computed:
  - given a finite number of implied volatilities
  - given an interpolation of implied volatilities between K_{min} and K_{max} (easier)

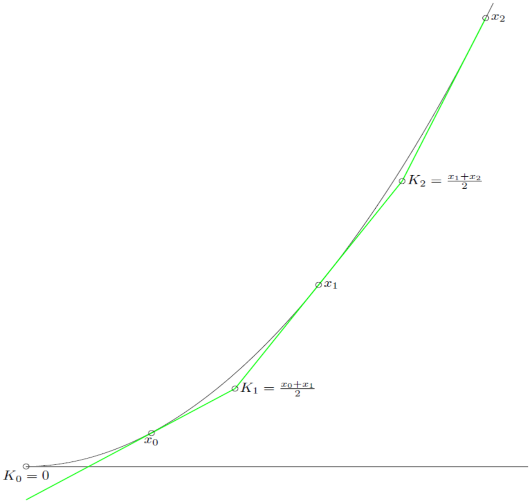
Extrapolation of OTM SPX put prices



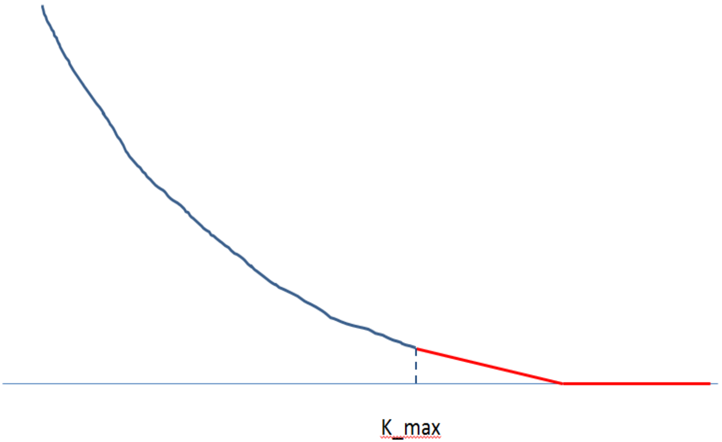
Extrapolation of OTM VIX call prices



Lower bound on VIX^2 given VIX implied vols: subreplication of parabola



Lower bound on VIX^2 given an interpolation of VIX implied vols



Necessary condition

Be careful of those traps:

- VIX smile extrapolations can be modified independently for different maturities, since options with different maturities have different underlyings (the VIX futures with those expiries)
- SPX smile extrapolations cannot be modified independently for different maturities, since options with different maturities have the same underlying (the SPX): $T \mapsto \sigma^2_{\mathsf{BS}}(T, K/F_T)$ must be nondecreasing for all K
- When T is a VIX option/future maturity:
  - $T + 30$ days is an SPX option maturity (3rd Friday of the month)
  - T is never an SPX option maturity: either 2 days (2 business days) before or 5 days (3 business days) after
  - When T is large, $T + 30$ days may not yet be a listed SPX option maturity

Past attempts

Gatheral (2008)

Consistent Modeling of SPX and VIX options

Consistent Modeling of SPX and VIX options

Jim Gatheral

Merrill Lynch

The Fifth World Congress of the Bachelier Finance Society
London, July 18, 2008

Double CEV dynamics

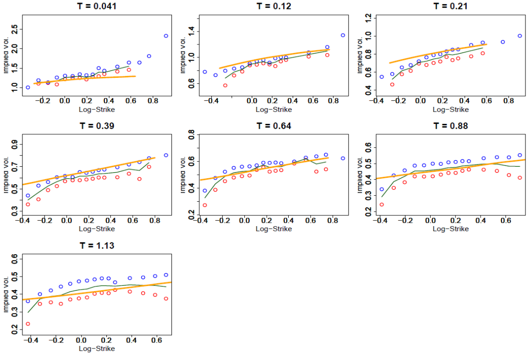
- Buehler's affine variance curve functional is consistent with double mean reverting dynamics of the form:

$$\begin{aligned} \frac{dS}{S} &= \sqrt{v}\, dW \\ dv &= -\kappa\,(v - v')\, dt + \eta_1\, v^{\alpha}\, dZ_1 \\ dv' &= -c\,(v' - z_3)\, dt + \eta_2\, v'^{\beta}\, dZ_2 \end{aligned} \qquad (2)$$

  for any choice of $\alpha,\ \beta \in [1/2, 1]$.
  - We will call the case $\alpha = \beta = 1/2$ *Double Heston*,
  - the case $\alpha = \beta = 1$ *Double Lognormal*,
  - and the general case *Double CEV*.
- All such models involve a short term variance level v that reverts to a moving level v' at rate κ. v' reverts to the long-term level z_3 at the slower rate $c < \kappa$.

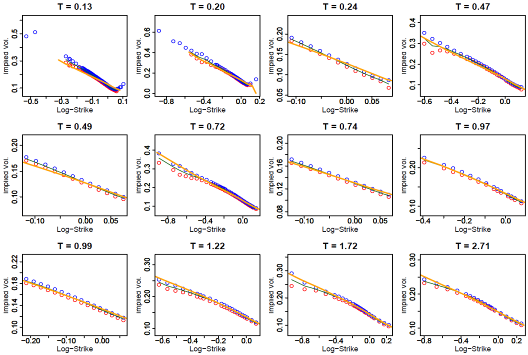
Double CEV fit to VIX options as of 03-Apr-2007

Setting the correlation ρ between volatility factors z_1 and z_2 to its historical average (see later) and iterating on the volatility of volatility parameters ξ_1 and ξ_2 to minimize the differences between model and market VIX option prices, we obtain the following fits (orange lines):

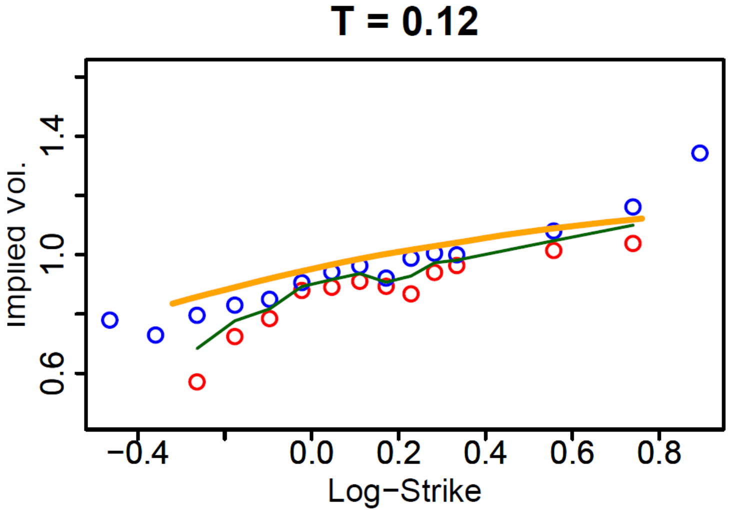


Double CEV fit to SPX options as of 03-Apr-2007

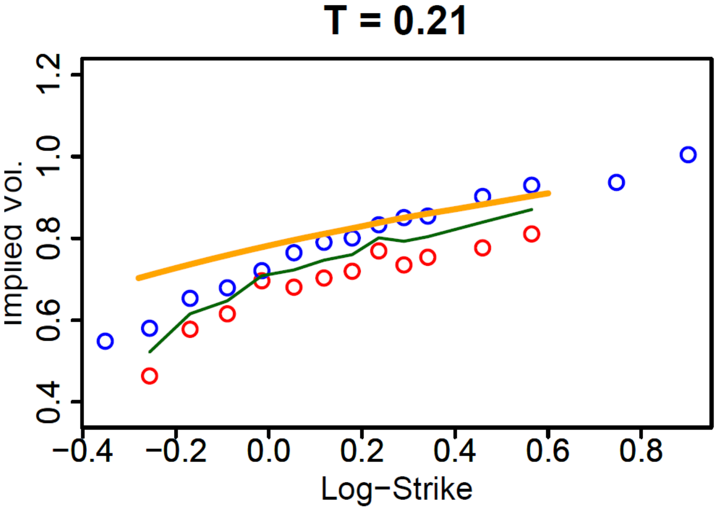
Minimizing the differences between model and market SPX option prices, we find $\rho_1 = -0.9$, $\rho_2 = -0.7$ and obtain the following fits to SPX option prices (orange lines):



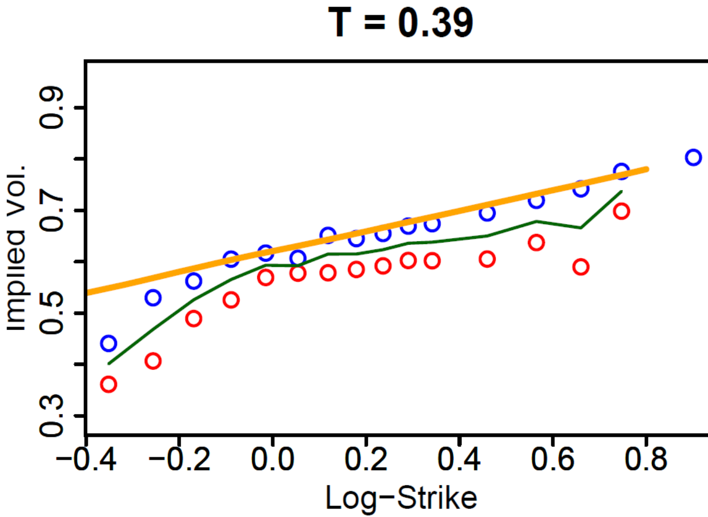
Fit to VIX options



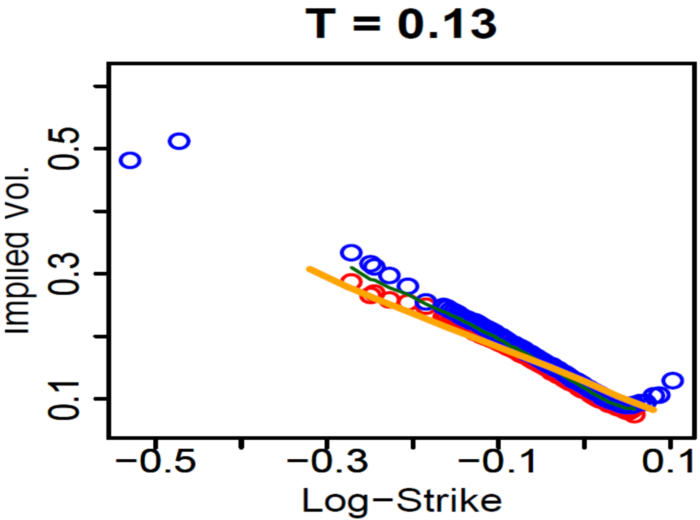
Fit to VIX options



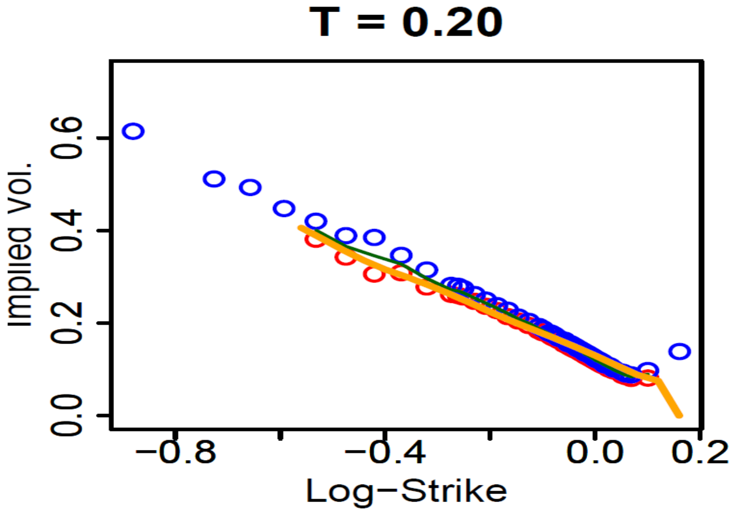
Fit to VIX options



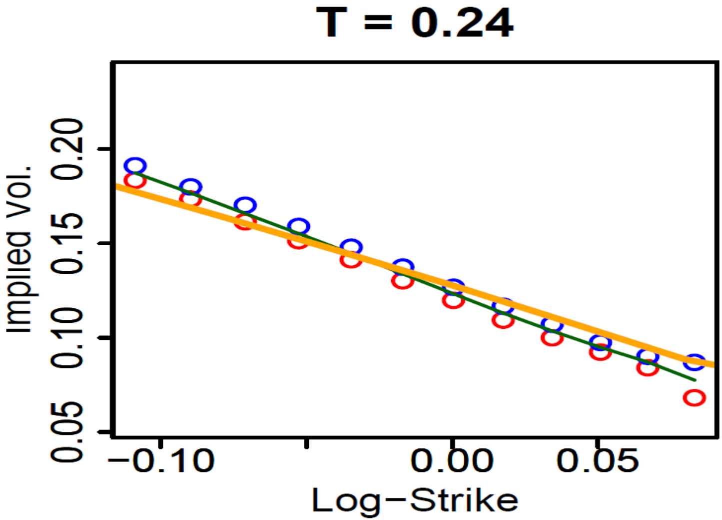
Fit to SPX options



Fit to SPX options



Fit to SPX options



- Joint calibration not so good for short maturities (up to 6 months).
- Unfortunate as these are the most liquid maturities for VIX futures and options.
- Vol-of-vol is either too large for VIX market, or too small for SPX market (or both).

Trying with jumps in SPX

Sepp (2012)

Part I. Joint calibration of SPX and VIX skews using jumps

I consider several volatility models to reproduce the volatility skew observed in equity options on the S&P500 index:
Local volatility model (LV)
Jump-diffusion model (JD)
Stochastic volatility model (SV)
Local stochastic volatility model (LSV) with jumps

For each model, I analyze its implied skew for options on the VIX

I show that LV, JD and SV without jumps are not consistent with the implied volatility skew observed in option on the VIX

I show that:
Only the SV model with appropriately chosen jumps can fit the implied VIX skew

Importantly, that only the LSV model with jumps can fit both Equity and VIX option skews

7

Baldeaux-Badran (2014)

Consistent Modelling of VIX and Equity Derivatives Using a $3/2$ plus Jumps Model

Jan Baldeaux and Alexander Badran

Abstract

The paper demonstrates that a pure-diffusion $3/2$ model is able to capture the observed upward-sloping implied volatility skew in VIX options. This observation contradicts a common perception in the literature that jumps are required for the consistent modelling of equity and VIX derivatives. The pure-diffusion model, however, struggles to reproduce the smile in the implied volatilities of short-term index options. One remedy to this problem is to augment the model by introducing jumps in the index. The resulting $3/2$ plus jumps model turns out to be as tractable as its pure-diffusion counterpart when it comes to pricing equity, realized variance and VIX derivatives, but accurately captures the smile in implied volatilities of short-term index options.

Keywords: Stochastic volatility plus jumps model, $3/2$ model, VIX derivatives

Baldeaux-Badran (2014)

$$dS_t = S_{t-}\left((r-\lambda\bar{\mu})dt + \rho\sqrt{V_t}dW_t^1 + \sqrt{1-\rho^2}\sqrt{V_t}dW_t^2 + (e^{\xi}-1)dN_t\right), \quad (3)$$

$$dV_t = \kappa V_t(\theta - V_t)dt + \epsilon(V_t^{3/2})dW_t^1, \quad (4)$$

where we denote by N a Poisson process at constant rate λ, by e^{ξ} the relative jump size of the stock and N is adapted to a filtration $(\mathcal{F}_t)_{t\in[0,T]}$. The distribution of ξ is assumed to be normal with mean μ and variance σ^2. The parameters μ, $\bar{\mu}$, and σ satisfy the following relationship

$$\mu = \log(1+\bar{\mu}) - \frac{1}{2}\sigma^2.$$

Kokholm-Stisen (2015)

$$\frac{dS_t}{S_t} = (r - q - \overline{\mu}\lambda)dt + \sqrt{V_t}dW_t + (e^{J^S} - 1)dN_t \tag{1}$$

$$dV_t = \kappa(\theta - V_t)dt + \eta\sqrt{V_t}dZ_t + J^V dN_t \tag{2}$$

where, W_t and Z_t are Wiener processes correlated with coefficient $\rho \in [-1, 1]$ and $\theta, \kappa, \eta \geq 0$. The price and volatility processes have simultaneous jumps with constant arrival intensity $\lambda \geq 0$. The jumps in volatility are independent and identically exponentially distributed with mean $\mu_v \geq 0$. Conditionally, on the jump in volatility, the jump in the price process is normally distributed with:

$$J^V \sim \exp(\mu_v), \quad J^S \,|\, J^V = y \sim N(\mu_s + \rho_J y, \sigma^2) \tag{3}$$

where $\sigma \geq 0$, $\rho_J \in [-1, 1]$, $\mu_s \in \mathbb{R}$. The martingale condition on the discounted price process imposes that:

$$\overline{\mu} = \frac{e^{\mu_s + \frac{1}{2}\sigma^2}}{1 - \rho_J \mu_v} - 1 \tag{4}$$

Kokholm-Stisen (2015)

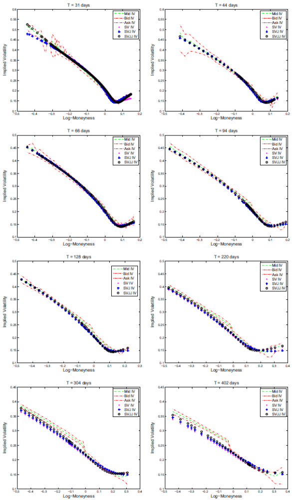


Figure 6.
Fit to the SPX option smiles on May 16, 2012, of the SV, SVJ and SVJJ models calibrated to SPX options and VIX derivatives without the Feller condition imposed

Kokholm-Stisen (2015)

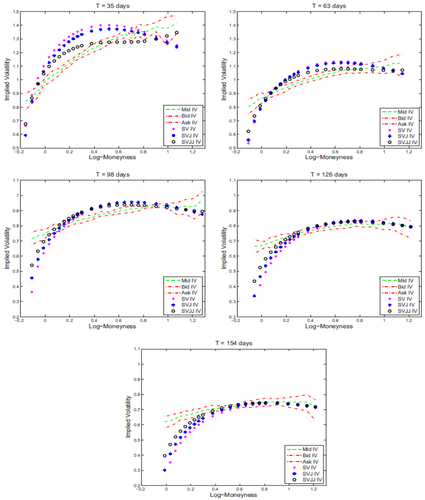


Figure 7.
Fit to the VIX option smiles on May 16, 2012, of the SV, SVJ and SVJJ models calibrated to SPX options and VIX derivatives without the Feller condition imposed

Kokholm-Stisen (2015)

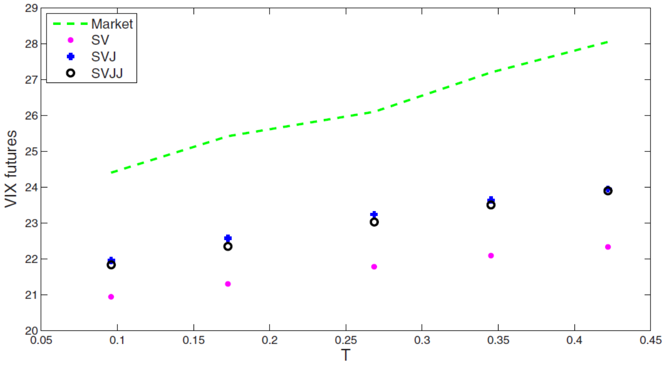


Figure 8. Fit to the VIX futures on May 16, 2012, of the SV, SVJ and SVJJ models calibrated to SPX options and VIX derivatives without the Feller condition imposed

Bardgett-Gourier-Leippold (2015)

$$dY_t = [-\lambda^{Y v}(v_{t-}, m_{t-})(\theta_Z(1,0,0) - 1) - \frac{1}{2}v_{t-}]dt + \sqrt{v_{t-}}dW_t^Y + dJ_t^Y,$$

$$dv_t = \kappa_v(m_{t-} - v_{t-})dt + \sigma_v\sqrt{v_{t-}}dW_t^v + dJ_t^v,$$

$$dm_t = \kappa_m(\theta_m - m_{t-})dt + \sigma_m\sqrt{m_{t-}}dW_t^m + dJ_t^m,$$

Bardgett-Gourier-Leippold (2015)

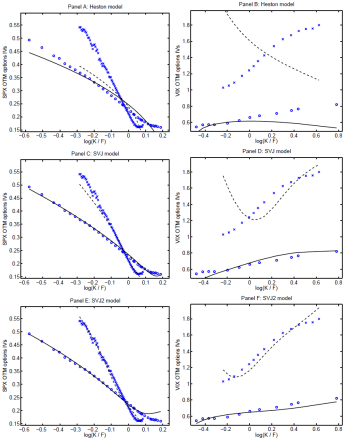


Figure 2: Market and model IVs on May 5, 2010, obtained by a joint calibration on the S&P 500 and VIX option market. Circles represent the market IV for $T = 0.05$ (S&P 500) and $T = 0.04$ (VIX). Crosses represent the market IV for $T = 0.3$ (S&P 500) and $T = 0.36$ (VIX). The dashed line corresponds to the model fit for $T = 0.05$ (S&P 500) and $T = 0.04$ (VIX) while the solid line corresponds to the model fit for $T = 0.3$ (S&P 500) and $T = 0.36$ (VIX). Panels A (S&P 500) and B (VIX) plot the model IVs based on the Heston model. Panels C and D display the corresponding results for the SVJ model, while Panels E and F do so for the SVJ2 model.

Papanicolaou-Sircar (2014)

- Use a regime-switching stochastic volatility model
- Hidden regime θ: continuous time Markov chain

$$
\begin{aligned}
dX_t &= \left(r - \frac{1}{2} f^2(\theta_t) Y_t - \delta\nu(\theta_{t\text{-}})\right) dt + f(\theta_t)\sqrt{Y_t}\, dW_t - \lambda(\theta_t) J_t\, dN_t \ , \\
dY_t &= \kappa(\bar{Y} - Y_t)dt + \gamma\sqrt{Y_t}\, dB_t \ , \\
dN_t &= \mathbb{1}_{[\theta_t \neq \theta_{t-}]},
\end{aligned}
$$

Papanicolaou-Sircar (2014)

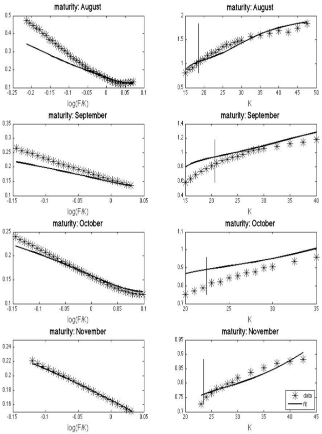


Figure 9: *The implied volatilities of July 27th SPX options (left column) and VIX options (right column), plotted alongside those of a fitted Heston with jumps. The fitted parameter values are given in Table 4. The vertical lines in the plots on the right mark the VIX futures price on the date of maturity.*

Cont-Kokholm (2013)

- Framework *à la* Bergomi:
  1. Model dynamics of forward variances $V_t^{[T_i,T_{i+1}]}$
  2. Given $V_{T_i}^{[T_i,T_{i+1}]}$, model dynamics of SPX
- Simultaneous (Lévy) jumps on forward variances and SPX
- First time a model seems to be able to jointly fit SPX skew and VIX level even for short maturities

Cont-Kokholm (2013)

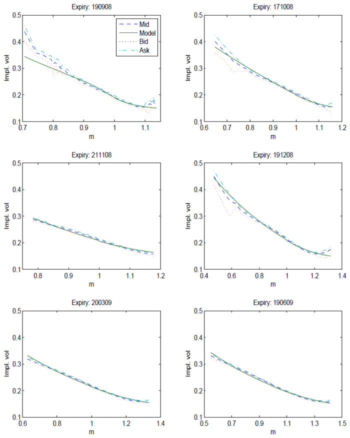


Figure 6: S&P 500 implied volatility smiles on August 20th 2008 for the model with normally distributed jump sizes plotted against moneyness $m = K/S_t$ on the horizontal axis.

Cont-Kokholm (2013)

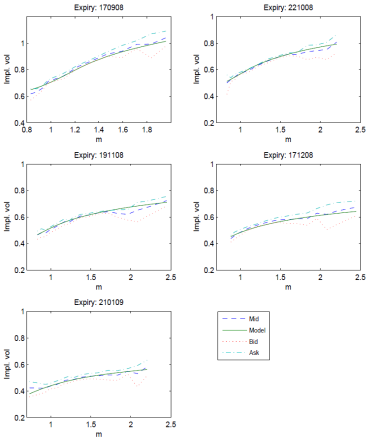


Figure 4: VIX implied volatility smiles on August 20th 2008 for the model with normally distributed jump sizes plotted against moneyness $m = K/VIX_t$ on the horizontal axis.

Pacati-Pompa-Renò (2015)

$$\begin{cases} dx_t = \left[r - q - \lambda\bar{\mu} - \frac{1}{2}\left(\sigma_{1,t}^2 + \phi_t + \sigma_{2,t}^2\right)\right]dt + \sqrt{\sigma_{1,t}^2 + \phi_t}\, dW_{1,t}^S + \sigma_{2,t} dW_{2,t}^S + c_x dN_t \\ d\sigma_{1,t}^2 = \alpha_1(\beta_1 - \sigma_{1,t}^2)dt + \Lambda_1 \sigma_{1,t} dW_{1,t}^\sigma + c_\sigma dN_t + c'_\sigma dN'_t \\ d\sigma_{2,t}^2 = \alpha_2(\beta_2 - \sigma_{2,t}^2)dt + \Lambda_2 \sigma_{2,t} dW_{2,t}^\sigma \end{cases}$$

Pacati-Pompa-Renò (2015)

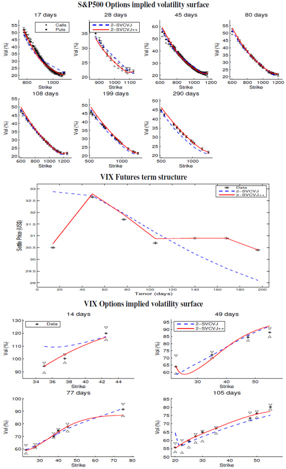


Figure 1: This figure reports market and model implied volatilities for S&P500 (plot at the top) and VIX (plot at the bottom) options, together with the term structure of VIX futures (plot in the middle) on September 2, 2009 obtained calibrating jointly on the three markets the 2-SVCVJ (blue dashed line) and 2-SVCVJ++ (red line). Maturities and tenors are expressed in days and volatilities are in % points and VIX futures settle prices are in US$.

Pacati-Pompa-Renò (2015)

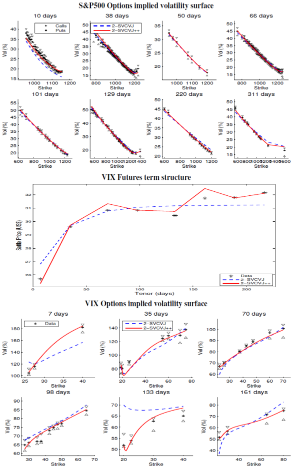


Figure 2: This figure reports market and model implied volatilities for S&P500 (plot at the top) and VIX (plot at the bottom) options, together with the term structure of VIX futures (plot in the middle) on August 11, 2010 obtained calibrating jointly on the three markets the 2-SVCVJ (blue dashed line) and 2-SVCVJ++ (red line). Maturities and tenors are expressed in days and volatilities are in % points and VIX futures settle prices are in US$.

Trying again with no jumps in SPX

Goutte-Ismail-Pham (2017)

- Also use a regime-switching stochastic volatility model
- Hidden regime Z: continuous time Markov chain

$$\begin{cases} dS_t = S_t(r dt + \sqrt{V_t} dW_t^1), \quad S_0 = s \\ dV_t = \kappa(Z_t)(\theta(Z_t) - V_t)dt + \xi(Z_t)\sqrt{V_t} dW_t^2, \quad V_0 = v_0. \end{cases}$$

Goutte-Ismail-Pham (2017)

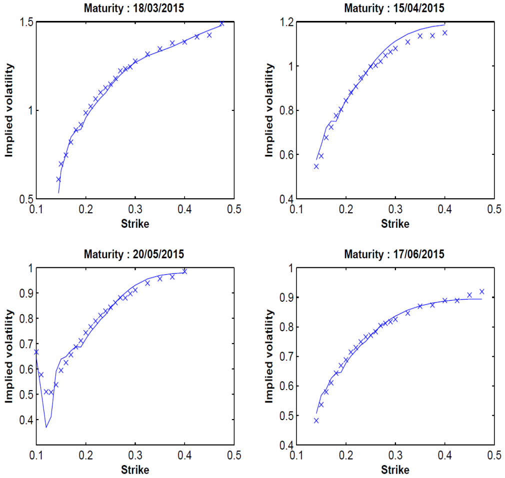


Figure 8: Implied volatilities of February 13, 2015, for VIX call options and the calibrated smile.

Goutte-Ismail-Pham (2017)

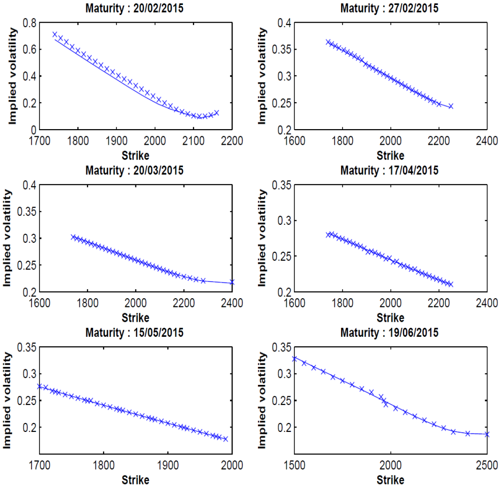


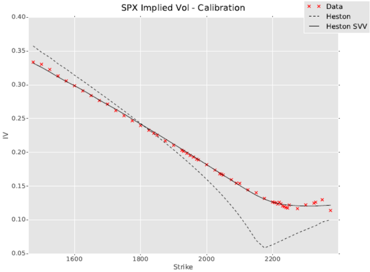
Figure 9: Implied volatilities of February 13, 2015, for S&P 500 call options and the calibrated smile.

Goutte-Ismail-Pham (2017)

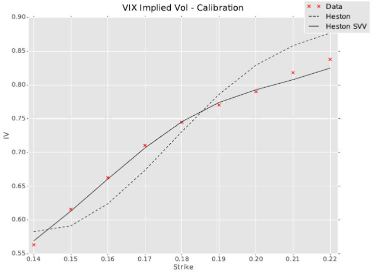
...but problem with SPX market data

Fouque-Saporito (2017)

- Based on Heston model with stochastic vol of vol
- No jumps
- Good fit to both SPX and VIX options... but only for maturities ≥ 4 months



(a) S&P 500



(b) VIX

So does there exist a continuous model on the SPX that jointly calibrates to SPX options, VIX futures, and VIX options?

No answer yet...

Continuous model on SPX

Continuous model on SPX calibrated to SPX options

- For simplicity, let us assume zero interest rates, repos, and dividends.
- Let $\mathcal{F}_t$ denote the market information available up to time t.
- We consider continuous models on the SPX index:

$$\frac{dS_t}{S_t} = \sigma_t \, dW_t, \qquad S_0 = x. \tag{3.1}$$

- W denotes a standard one-dimensional $(\mathcal{F}_t)$-Brownian motion, (σ_t) is an $(\mathcal{F}_t)$-adapted process such that for all $t \geq 0$, $\int_0^t \sigma_s^2 \, ds < \infty$ a.s.
- The local volatility function corresponding to Model (3.1) is the function σ_{loc} defined by

$$\sigma_{\text{loc}}^2(t, S_t) := \mathbb{E}[\sigma_t^2 | S_t]. \tag{3.2}$$

- The corresponding local volatility model is defined by:

$$\frac{dS_t^{\text{loc}}}{S_t^{\text{loc}}} = \sigma_{\text{loc}}(t, S_t^{\text{loc}}) \, dW_t, \qquad S_0^{\text{loc}} = x.$$

Continuous model on SPX calibrated to SPX options

- From Gyöngy (1986), the marginal distributions of the processes $(S_t, t \geq 0)$ and $(S_t^{\mathrm{loc}}, t \geq 0)$ agree:

$$\forall t \geq 0, \quad S_t^{\mathrm{loc}} \overset{(d)}{=} S_t. \tag{3.3}$$

- Using Dupire (1994), we conclude that Model (3.1) is calibrated to the full SPX smile if and only if

$$\sigma_{\mathrm{loc}} = \sigma_{\mathrm{lv}} \tag{3.4}$$

where σ_{lv} is the local volatility function derived from market prices of vanilla options on the SPX using Dupire's formula.

- We denote by S^{lv} the market local volatility model is defined by:

$$\frac{dS_t^{\mathrm{lv}}}{S_t^{\mathrm{lv}}} = \sigma_{\mathrm{lv}}(t, S_t^{\mathrm{lv}})\, dW_t, \qquad S_0^{\mathrm{lv}} = x.$$

VIX

- Let $T \geq 0$. By definition, the (idealized) VIX at time T is the implied volatility of a 30 day log-contract on the SPX index starting at T. For continuous models (3.1), this translates into

$$\mathrm{VIX}_T^2 = \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma_t^2\, dt \middle| \mathcal{F}_T\right] = \frac{1}{\tau}\int_T^{T+\tau} \mathbb{E}\left[\sigma_t^2 \middle| \mathcal{F}_T\right] dt \qquad (3.5)$$

  where $\tau = \frac{30}{365}$ (30 days).

- Since $\mathbb{E}[\sigma_{\mathrm{loc}}^2(t, S_t^{\mathrm{loc}})|\mathcal{F}_T] = \mathbb{E}[\sigma_{\mathrm{loc}}^2(t, S_t^{\mathrm{loc}})|S_T^{\mathrm{loc}}]$, $\mathrm{VIX}_{\mathrm{loc},T}$ satisfies

$$\mathrm{VIX}_{\mathrm{loc},T}^2 = \frac{1}{\tau}\int_T^{T+\tau} \mathbb{E}[\sigma_{\mathrm{loc}}^2(t, S_t^{\mathrm{loc}})|S_T^{\mathrm{loc}}]\, dt = \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma_{\mathrm{loc}}^2(t, S_t^{\mathrm{loc}})\, dt \middle| S_T^{\mathrm{loc}}\right].$$

- Similarly,

$$\mathrm{VIX}_{\mathrm{lv},T}^2 = \frac{1}{\tau}\int_T^{T+\tau} \mathbb{E}[\sigma_{\mathrm{lv}}^2(t, S_t^{\mathrm{lv}})|S_T^{\mathrm{lv}}]\, dt = \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma_{\mathrm{lv}}^2(t, S_t^{\mathrm{lv}})\, dt \middle| S_T^{\mathrm{lv}}\right].$$

VIX

- The prices at time 0 of the VIX future and the VIX call options with common maturity T in Model (3.1) are respectively given by

$$\mathrm{VIX}_0^{\mathrm{model}}(T) = \mathbb{E}\left[\sqrt{\mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau}\sigma_t^2\,dt\middle|\mathcal{F}_T\right]}\right], \tag{3.6}$$

$$C_{\mathrm{VIX}}^{\mathrm{model}}(T,K) = \mathbb{E}\left[\left(\sqrt{\mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau}\sigma_t^2\,dt\middle|\mathcal{F}_T\right]}-K\right)_+\right]. \tag{3.7}$$

- We observe market prices for those instruments, for a list of liquid monthly VIX future maturities T_i, denoted by $\mathrm{VIX}_0^{\mathrm{mkt}}(T_i)$ and $C_{\mathrm{VIX}}^{\mathrm{mkt}}(T_i,K)$, with the most liquid maturities lying below 6 months.
- **Can we find a model satisfying (3.1)-(3.4) and such that for all T_i and K, $\mathrm{VIX}_0^{\mathrm{model}}(T_i) = \mathrm{VIX}_0^{\mathrm{mkt}}(T_i)$ and $C_{\mathrm{VIX}}^{\mathrm{model}}(T_i,K) = C_{\mathrm{VIX}}^{\mathrm{mkt}}(T_i,K)$?**

The case of instantaneous VIX

The case of instantaneous VIX

$\tau \to 0$: The realized variance over 30 days is replaced by the instantaneous variance, and (3.6)-(3.7) boil down to

$$\mathrm{instVIX}_0^{\mathrm{model}}(T) = \mathbb{E}\left[\sigma_T\right], \quad (4.1)$$

$$C_{\mathrm{instVIX}}^{\mathrm{model}}(T,K) = \mathbb{E}\left[(\sigma_T - K)_+\right]. \quad (4.2)$$

- Reminder: (The distributions of) two random variables X and Y are said to be in convex order if and only if, for any convex function f, $\mathbb{E}[f(X)] \leq \mathbb{E}[f(Y)]$.
- Denoted by $X \leq_c Y$.
- Both distributions have same mean, but distribution of Y is more "spread" than that of X.
- In financial terms: X and Y have the same forward value, but calls (puts) on Y are more expensive than calls (puts) on X.

The case of instantaneous VIX

- Assume $\sigma_{\rm loc} = \sigma_{\rm lv}$: $\mathbb{E}[\sigma_t^2|S_t] = \sigma_{\rm lv}^2(t, S_t)$. By conditional Jensen,

$$\forall t \geq 0, \qquad \sigma_{\rm lv}^2(t, S_t) \leq_c \sigma_t^2.$$

- Conversely, if $\sigma_{\rm lv}^2(t, S_t) \leq_c \sigma_t^2$, then there exists a joint distribution π_t of (S_t, σ_t) such that $\mathbb{E}[\sigma_t^2|S_t] = \sigma_{\rm lv}^2(t, S_t)$ for all t.
- Indeed, let us denote

$$X_t := \sigma_{\rm lv}^2(t, S_t) \qquad \text{and} \qquad Y_t := \sigma_t^2$$

From Strassen's theorem (1965), there exists a joint distribution π'_t of (X_t, Y_t) such that $\mathbb{E}[Y_t|X_t] = X_t$ (**martingale transport**). Define π_t as follows:

  - $S_t \sim$ risk-neutral distribution of the SPX for maturity t
  - Given S_t, $X_t = \sigma_{\rm lv}^2(t, S_t)$ is known and $\sigma_t^2 \sim \mathcal{L}_{\pi'}(Y_t|X_t)$.

By construction, $\mathbb{E}[\sigma_t^2|S_t] = \mathbb{E}[Y_t|X_t] = X_t = \sigma_{\rm lv}^2(t, S_t)$.

The case of instantaneous VIX

Necessary and sufficient condition for a jointly calibrating continuous model on the SPX to exist (G., 2017):

$$\forall t \geq 0, \qquad \sigma^2_{\mathrm{lv}}(t, S_t) \leq_c \sigma^2_t$$

- If $\mathrm{instVIX}_0^{\mathrm{mkt}}(t)$ and $C^{\mathrm{mkt}}_{\mathrm{instVIX}}(t,K)$ were accessible, we could imply from the market the distribution of σ_t^2, and compare it to the risk-neutral distribution of $\sigma^2_{\mathrm{lv}}(t, S_t)$.
- This general construction does not address the issue of the dynamics of (σ_t): σ_t and $\sigma_{t'}$ could be very loosely related for arbitrarily close t and t'.
- In this limiting case, **it might be impossible to build a continuous model on the SPX that jointly calibrates to SPX and VIX options**. This happens if and only if for some t the market-implied distribution of $\sigma^2_{\mathrm{lv}}(t, S_t)$ is "more spread" than that of the instantaneous VIX squared.

The case of instantaneous VIX

- In practice, to build a calibrating process, discretize time and recursively solve **martingale transport problems**:

$$\mathcal{L}\left(\sigma_{\mathrm{lv}}^2(t_k, S_{t_k})\right) \text{ and } \mathcal{L}\left(\sigma_{t_k}^2\right) \text{given,} \quad \mathbb{E}[\sigma_{t_k}^2 | \sigma_{\mathrm{lv}}^2(t_k, S_{t_k})] = \sigma_{\mathrm{lv}}^2(t_k, S_{t_k}).$$

- Solutions π'_{t_k} to those martingale transport problems include left- and right-curtains (Beiglböck-Juillet, Henry-Labordère), forward-starting solutions to the Skorokhod embedding problems (Dupire), and the local variance gamma model of Carr and Nadtochiy.
- This is a **new type of application of martingale transport to finance**:
  - Usually, the martingality constraint applies to the underlying at two different dates (Henry-Labordère, Beiglböck, Penkner, Nutz, Touzi, Martini, De Marco, Dolinsky, Soner, Obłój, Stebegg, JG...)
  - Here it applies to **two types of instantaneous variances at a single date**, ensuring that the SPX smile is matched.

Real (30-day) VIX

The real case

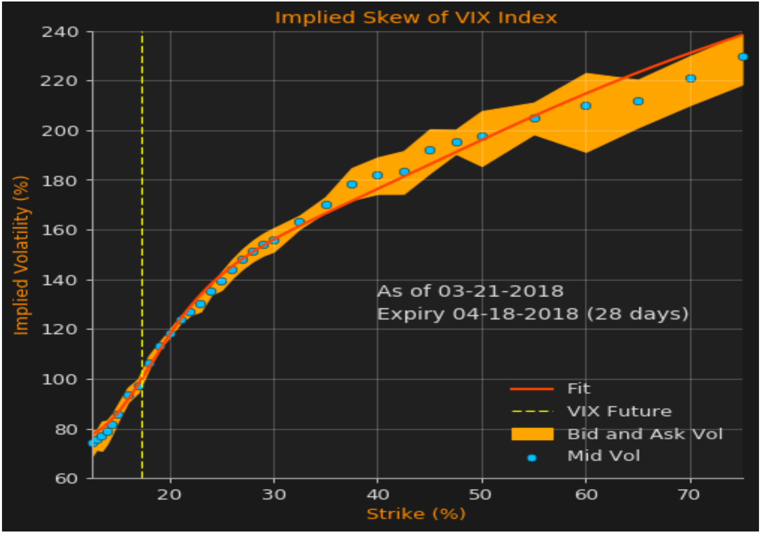
- In reality, squared VIX are not instantaneous variances but the **fair strikes** of **30-day** realized variances.
- Let us look at market data (March 21, 2018). We compare the market distributions of

$$\mathrm{VIX}^2_{\mathrm{lv},T} := \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma^2_{\mathrm{lv}}(t, S^{\mathrm{lv}}_t)\,dt \middle| S^{\mathrm{lv}}_T\right]$$

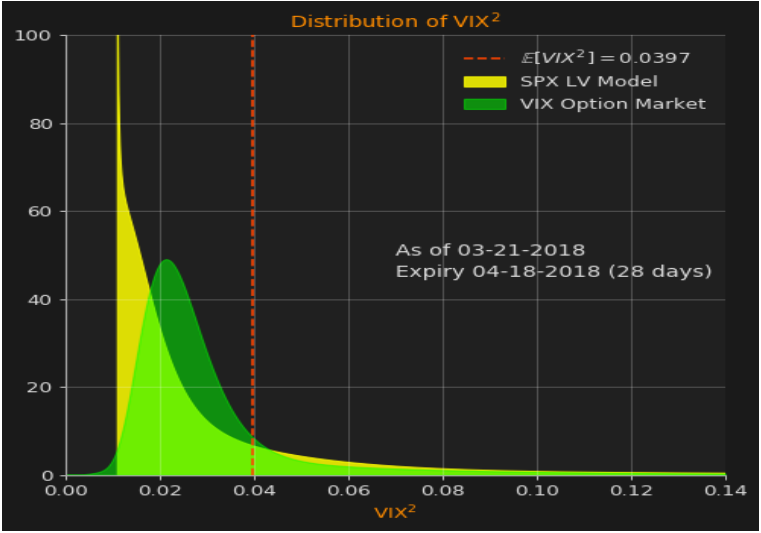
and

$$\mathrm{VIX}^2_{\mathrm{mkt},T} \quad \left(\longleftrightarrow \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma^2_t\,dt \middle| \mathcal{F}_T\right]\right)$$

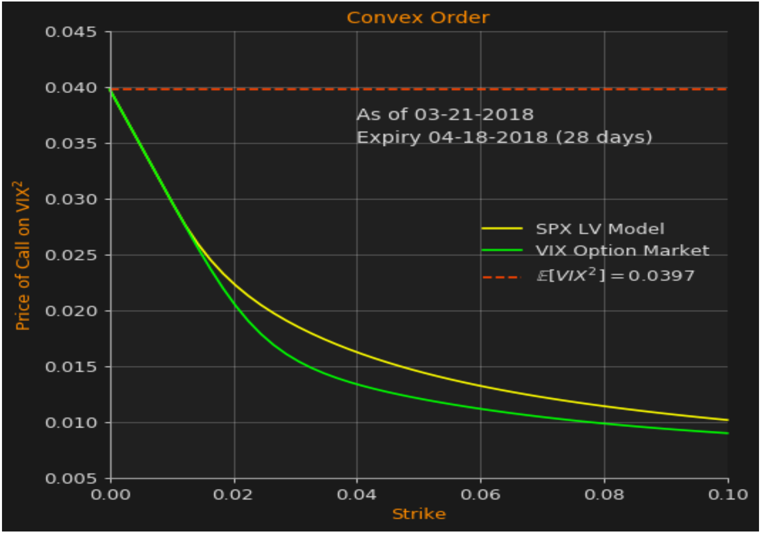
$T = 1$ month


Implied Skew of VIX Index
Implied Volatility (%)
240
220
200
180
160
140
120
100
80
60
20
30
40
50
60
70
Strike (%)
As of 03-21-2018
Expiry 04-18-2018 (28 days)
Fit
VIX Future
Bid and Ask Vol
Mid Vol

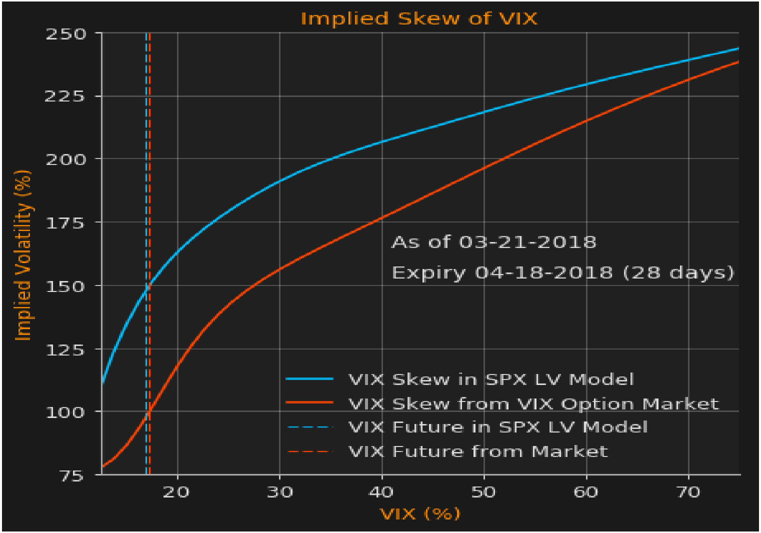
$T = 1$ month


Distribution of VIX²
$\mathbb{E}[VIX^2] = 0.0397$
SPX LV Model
VIX Option Market
As of 03-21-2018
Expiry 04-18-2018 (28 days)
100
80
60
40
20
0
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14
VIX²

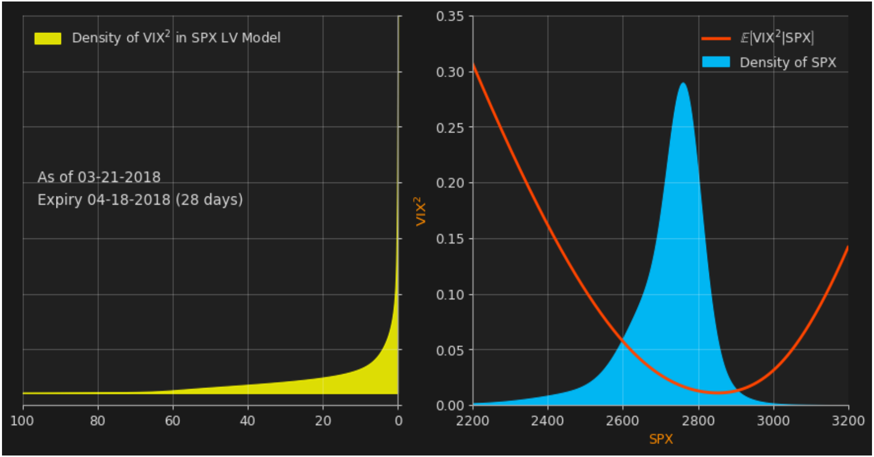
$T = 1$ month


Convex Order
As of 03-21-2018
Expiry 04-18-2018 (28 days)
SPX LV Model
VIX Option Market
$\mathbb{E}[VIX^2] = 0.0397$
Price of Call on VIX^2
Strike
0.045
0.040
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.00
0.02
0.04
0.06
0.08
0.10

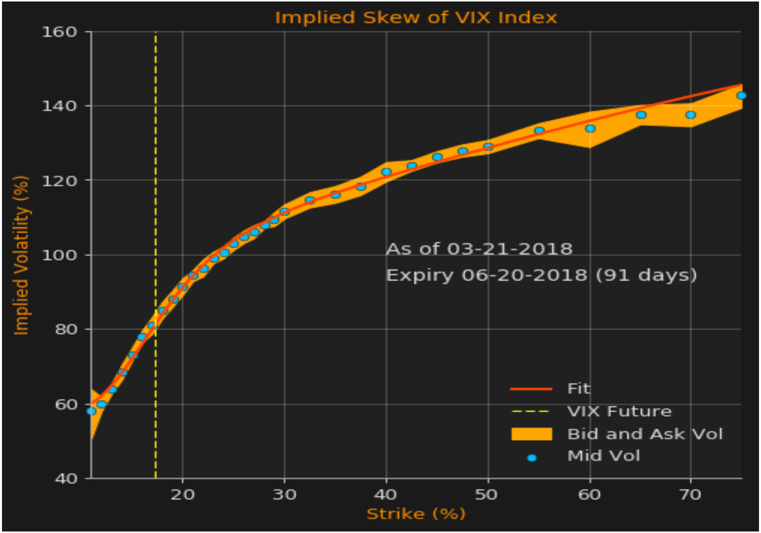
$T = 1$ month


Implied Skew of VIX
Implied Volatility (%)
VIX (%)
250
225
200
175
150
125
100
75
20
30
40
50
60
70
As of 03-21-2018
Expiry 04-18-2018 (28 days)
VIX Skew in SPX LV Model
VIX Skew from VIX Option Market
VIX Future in SPX LV Model
VIX Future from Market

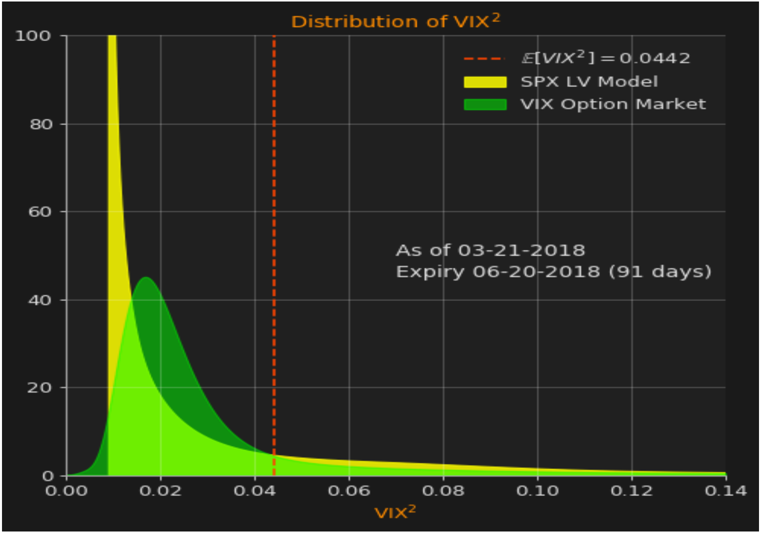
$\mathrm{VIX}^2_{\mathrm{lv},T}(S_T^{\mathrm{lv}})$



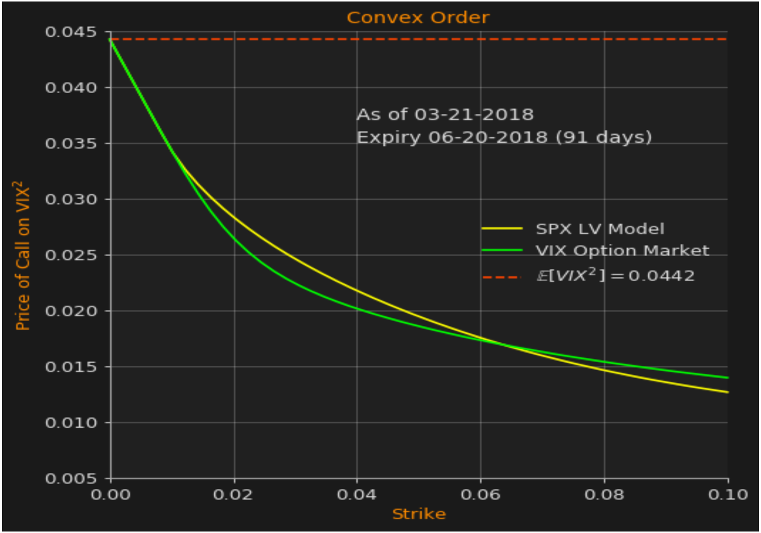
$T = 3$ months


Implied Skew of VIX Index
Implied Volatility (%)
Strike (%)
160
140
120
100
80
60
40
20
30
40
50
60
70
As of 03-21-2018
Expiry 06-20-2018 (91 days)
Fit
VIX Future
Bid and Ask Vol
Mid Vol

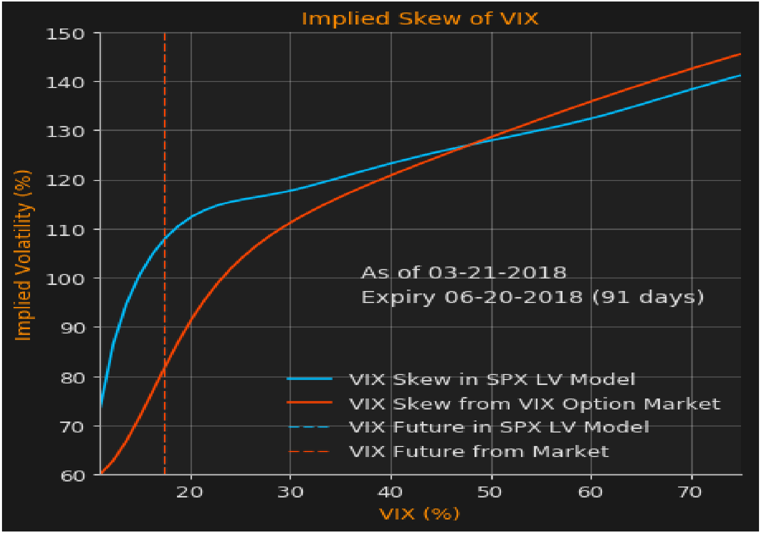
$T = 3$ months


Distribution of VIX²
100
80
60
40
20
0
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14
VIX²
E[VIX²] = 0.0442
SPX LV Model
VIX Option Market
As of 03-21-2018
Expiry 06-20-2018 (91 days)

$T = 3$ months


Convex Order
As of 03-21-2018
Expiry 06-20-2018 (91 days)
SPX LV Model
VIX Option Market
$\mathbb{E}[VIX^2] = 0.0442$
Price of Call on VIX²
Strike
0.045
0.040
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.00
0.02
0.04
0.06
0.08
0.10

$T = 3$ months


Implied Skew of VIX
Implied Volatility (%)
150
140
130
120
110
100
90
80
70
60
20
30
40
50
60
70
VIX (%)
As of 03-21-2018
Expiry 06-20-2018 (91 days)
VIX Skew in SPX LV Model
VIX Skew from VIX Option Market
VIX Future in SPX LV Model
VIX Future from Market

The real case

$$\begin{aligned}\mathrm{VIX}^2_{\mathrm{lv},T} &:= \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau}\sigma^2_{\mathrm{lv}}(t,S^{\mathrm{lv}}_t)\,dt\,\middle|\,S^{\mathrm{lv}}_T\right]\\ \mathrm{VIX}^2_T &= \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau}\sigma^2_t\,dt\,\middle|\,\mathcal{F}_T\right]\end{aligned}$$

- In typical market conditions:

$$\begin{aligned}1-2\ \text{months}: &\quad \mathrm{VIX}^2_{\mathrm{mkt},T} \le_c \mathrm{VIX}^2_{\mathrm{lv},T} \qquad (5.1)\\ 3-4\ \text{months}: &\quad \mathrm{VIX}^2_{\mathrm{lv},T}\ \text{ and }\ \mathrm{VIX}^2_{\mathrm{mkt},T}\ \text{ not rankable}\\ 5+\ \text{months}: &\quad \mathrm{VIX}^2_{\mathrm{lv},T} \le_c \mathrm{VIX}^2_{\mathrm{mkt},T}\end{aligned}$$

For short maturities, the local volatility model generates a distribution of VIX^2 that is "more spread" than the distribution of VIX^2 implied from VIX futures and options.

The real case

$$\begin{aligned} \mathrm{VIX}^2_{\mathrm{lv},T} &:= \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma^2_{\mathrm{lv}}(t, S^{\mathrm{lv}}_t)\, dt \,\middle|\, S^{\mathrm{lv}}_T\right] \\ \mathrm{VIX}^2_T &= \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma^2_t\, dt \,\middle|\, \mathcal{F}_T\right] \end{aligned}$$

- From the case of instantaneous VIX, one may be tempted to believe that there exists a model of the form $\frac{dS_t}{S_t} = \sigma_t\, dW_t$ calibrated to the SPX smile and to all VIX options if and only if for all T_i, $\mathrm{VIX}^2_{\mathrm{lv},T_i} \leq_c \mathrm{VIX}^2_{\mathrm{mkt},T_i}$ — and then conclude that such a model does not exist.
- However:

$$\sigma^2_{\mathrm{lv}}(t, S_t) \leq_c \sigma^2_t \;\;\not\Longrightarrow\;\; \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma^2_{\mathrm{lv}}(t, S_t)\, dt \,\middle|\, \mathcal{F}_T\right] \leq_c \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma^2_t\, dt \,\middle|\, \mathcal{F}_T\right]$$

- **Sum over 30 days and conditioning on $\mathcal{F}_T$ may undo convex ordering**.

Convex order is not preserved under sum

Example

- A trivial almost counterexample:

$$Y_0 = X_0 + Z, \qquad Y_1 = X_1 - Z$$

  with $\mathbb{E}[Z|X_0] = \mathbb{E}[Z|X_1] = 0$ (e.g., Z has zero mean and is independent from (X_0, X_1)).
- Y_0 can be much larger than X_0 in the convex order and Y_1 can be much larger than X_1 in the convex order, if Z has large variance.
- However, $Y_0 + Y_1 = X_0 + X_1$.

Example

- $X_0 = W_{t_1}$, $X_1 = -W_{t_2}$, $Y_0 = W_{t_3}$, and $Y_1 = -W_{t_3}$, with $0 < t_1 < t_2 < t_3$.
- $\mathbb{E}[Y_0|X_0] = X_0$, $\mathbb{E}[Y_1|X_1] = X_1$, hence $X_0 \leq_c Y_0$ and $X_1 \leq_c Y_1$, yet $0 = Y_0 + Y_1 <_c X_0 + X_1$.

Convex order is not preserved under sum

Example

- We generalize the previous example: $G = (X_0, Y_0, X_1, Y_1)$ Gaussian vector.
- We assume that $\mathbb{E}[Y_0|X_0] = X_0$ and $\mathbb{E}[Y_1|X_1] = X_1$, and look for necessary and sufficient conditions under which $X_0 + X_1 \leq_c Y_0 + Y_1$.[1]
- $m_X := \mathbb{E}[X]$, σ_X std dev of X, ρ_{XY} the correlation between X and Y.
- Since G is Gaussian, $\mathbb{E}[Y_i|X_i] = m_{Y_i} + \rho_{X_iY_i}\frac{\sigma_{Y_i}}{\sigma_{X_i}}(X_i - m_{X_i})$ so

$$m_{X_i} = m_{Y_i} \quad \text{and} \quad \sigma_{X_i} = \rho_{X_iY_i}\sigma_{Y_i}. \tag{5.2}$$

In particular, $\rho_{X_iY_i} > 0$. As a consequence, $m_{X_0+X_1} = m_{Y_0+Y_1}$, and since $X_0 + X_1$ and $Y_0 + Y_1$ are Gaussian,

$$X_0 + X_1 \leq_c Y_0 + Y_1 \iff \mathrm{Var}(X_0 + X_1) \leq \mathrm{Var}(Y_0 + Y_1).$$

[1] We ignore trivial cases by assuming that all components of G have positive variance.

Convex order is not preserved under sum

Example

- Now, using the second equation in (5.2), we have

$$\begin{aligned}\mathrm{Var}(X_0+X_1) &= \sigma_{X_0}^2+\sigma_{X_1}^2+2\rho_{X_0X_1}\sigma_{X_0}\sigma_{X_1}\\ &= \rho_{X_0Y_0}^2\sigma_{Y_0}^2+\rho_{X_1Y_1}^2\sigma_{Y_i}^2+2\rho_{X_0X_1}\rho_{X_0Y_0}\sigma_{Y_0}\rho_{X_1Y_1}\sigma_{Y_1}\end{aligned}$$

so $X_0+X_1\leq_c Y_0+Y_1$ if and only if

$$\sigma_{Y_0}^2(1-\rho_{X_0Y_0}^2)+\sigma_{Y_1}^2(1-\rho_{X_1Y_1}^2)+2\sigma_{Y_0}\sigma_{Y_1}(\rho_{Y_0Y_1}-\rho_{X_0X_1}\rho_{X_0Y_0}\rho_{X_1Y_1})\geq 0.$$

In particular, if $\sigma_{Y_0}=\sigma_{Y_1}$, $\rho_{X_iY_i}\neq 1$ for $i\in\{0,1\}$, and

$$\chi := \frac{\rho_{Y_0Y_1}-\rho_{X_0X_1}\rho_{X_0Y_0}\rho_{X_1Y_1}}{\sqrt{1-\rho_{X_0Y_0}^2}\sqrt{1-\rho_{X_1Y_1}^2}} < -1$$

then $X_0+X_1\not\leq_c Y_0+Y_1$.

Convex order is not preserved under conditioning

- Conditioning with respect to $\mathcal{F}_T$ may undo convex ordering too.
- Simple counterexample: if $X \leq_c Y$ with X $\mathcal{F}$-measurable and not constant, and Y independent of $\mathcal{F}$, then $\mathbb{E}[Y] = \mathbb{E}[Y|\mathcal{F}] <_c \mathbb{E}[X|\mathcal{F}] = X$.
- Intuition: **Fast mean reversion in (σ_t) may undo convex ordering**:
  - σ_t quickly forgets the information contained in $\mathcal{F}_T$.
  - $\sigma_{\text{loc}}(t, S_t^{\text{loc}})$ does not mean revert: S_t^{loc} is a martingale and does not forget the information contained in S_T^{loc}.

  However, the larger the mean reversion, the flatter $S \mapsto \sigma_{\text{loc}}(t, S)$.
- **Rough volatility** models with small Hurst exponent H **may also undo convex ordering**, since volatility increments are negatively correlated.

Inversion of convex ordering

- For Model (3.1) to fit both SPX and VIX option prices, it **must** satisfy

$$\mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau}\sigma_u^2\,du\middle|\mathcal{F}_T\right] \leq_c \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau}\sigma_{\text{loc}}^2(u,S_u^{\text{loc}})\,du\middle|S_T^{\text{loc}}\right] \quad (5.3)$$

for T up to a few months, despite the fact that for all $u \geq 0$, $\sigma_{\text{loc}}^2(u,S_u^{\text{loc}}) \leq_c \sigma_u^2$ (**inversion of convex ordering**).

- One natural way to achieve (5.3) is to require that

$$\mathbb{E}[\sigma_u^2|\mathcal{F}_T] \leq_c \mathbb{E}[\sigma_{\text{loc}}^2(u,S_u^{\text{loc}})|S_T^{\text{loc}}] \quad (5.4)$$

for many $u \in (T, T+\tau]$ and hope that this convex ordering of forward instantaneous variances will be preserved when we sum over u.

- When $u = T$, $\mathbb{E}[\sigma_u^2|\mathcal{F}_T] = \sigma_T^2 \geq_c \sigma_{\text{loc}}^2(T,S_T^{\text{loc}}) = \mathbb{E}[\sigma_{\text{loc}}^2(u,S_u^{\text{loc}})|S_T^{\text{loc}}] \Longrightarrow$ We will require that (5.4) holds for all $T \leq \bar{T}$ and $u \in [T+\varepsilon, T+\tau]$.
- When (5.4) holds, the **convex ordering** $\sigma_{\text{loc}}^2(u,S_u^{\text{loc}}) \leq_c \sigma_u^2$ **is actually reversed after conditioning on** $\mathcal{F}_T$:

$$\mathbb{E}[\sigma_u^2|\mathcal{F}_T] \leq_c \mathbb{E}[\sigma_{\text{loc}}^2(u,S_u^{\text{loc}})|\mathcal{F}_T] \leq_c \sigma_{\text{loc}}^2(u,S_u^{\text{loc}}) \leq_c \sigma_u^2. \quad (5.5)$$

Inversion of convex ordering in forward variance models

One-factor lognormal forward instantaneous variance models

- $\mathbb{E}[\sigma_u^2|\mathcal{F}_T]$ is a **forward instantaneous variance**. We denote it by
$$\xi_T^u := \mathbb{E}[\sigma_u^2|\mathcal{F}_T].$$
- It is well known (Dupire, Bergomi, Buehler) that forward instantaneous variances are driftless.
- Second generation stochastic volatility models directly model the dynamics of $(\xi_t^u, t \in [0,u])$ under a risk-neutral measure. The only requirement is that these processes, indexed by u, be **nonnegative** and **driftless** under risk-neutral measures.
- For simplicity, we assume that forward instantaneous variances are lognormal and all driven by a single standard one-dimensional $(\mathcal{F}_t)$-Brownian motion Z, correlated with W:
$$\frac{d\xi_t^u}{\xi_t^u} = K(t,u)\,dZ_t. \tag{6.1}$$
- The kernel K is deterministic.
- The SPX dynamics simply reads as (3.1) with $\sigma_t^2 := \xi_t^t$:
$$\frac{dS_t}{S_t} = \sigma_t\,dW_t, \qquad \sigma_t^2 := \xi_t^t, \qquad d\langle W, Z\rangle_t = \rho\,dt.$$

One-factor lognormal forward instantaneous variance models

- The solution to (6.1) is simply

$$\xi_t^u = \xi_0^u \exp\left(\int_0^t K(s,u)\,dZ_s - \frac{1}{2}\int_0^t K(s,u)^2\,ds\right) \tag{6.2}$$

which yields

$$\sigma_u^2 = \xi_0^u \exp\left(\int_0^u K(s,u)\,dZ_s - \frac{1}{2}\int_0^u K(s,u)^2\,ds\right). \tag{6.3}$$

- For simplicity, let us choose a time-homogeneous kernel

$$K(s,u) = K(u-s).$$

- Financially, we expect the kernel $K : \mathbb{R}_+ \to \mathbb{R}_+$ to be decreasing: The further away the instantaneous forward variance maturity u, the less volatile the instantaneous forward variance.

Ingredients needed for inversion of convex ordering

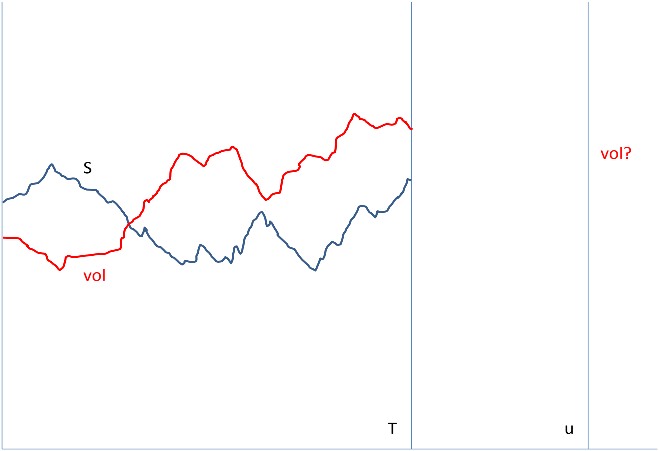
- **Can we choose a kernel K such that**

$$\mathbb{E}[\sigma_u^2|\mathcal{F}_T] \leq_c \mathbb{E}[\sigma_{\text{loc}}^2(u, S_u^{\text{loc}})|S_T^{\text{loc}}]$$

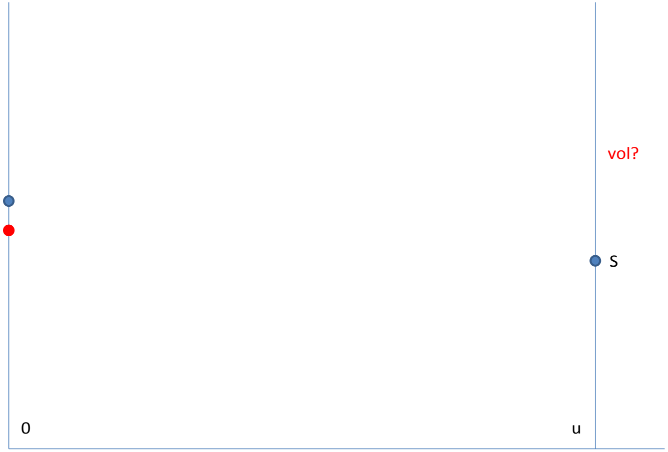
for all $T \leq \bar{T}$ and $u \in [T+\varepsilon, T+\tau]$?

- **Ingredient 1**: The knowledge of $\mathcal{F}_T := \sigma(W_s, Z_s, 0 \leq s \leq T)$ should give little information on σ_u^2, so the distribution of $\mathbb{E}[\sigma_u^2|\mathcal{F}_T]$ is narrow.
- **Ingredient 2**: The knowledge of S_u should give enough information on σ_u^2, so that $S \mapsto \sigma_{\text{loc}}^2(u, S) = \mathbb{E}[\sigma_u^2|S_u = S]$ varies enough with S and the distribution of $\mathbb{E}[\sigma_{\text{loc}}^2(u, S_u^{\text{loc}})|S_T^{\text{loc}}]$ is not as narrow as that of $\mathbb{E}[\sigma_u^2|\mathcal{F}_T]$.
- Loosely speaking, S_u **should give more information than $\mathcal{F}_T$ on σ_u^2**.

Ingredients needed for inversion of convex ordering



Ingredients needed for inversion of convex ordering



Ingredients 1 and 2 are antagonistic

$$\sigma_u^2 = \xi_0^u \exp\left(\int_0^u K(u-s)\, dZ_s - \frac{1}{2} \int_0^u K(u-s)^2\, ds \right).$$

- **Ingredient 1**: $K(u-s)$ **must be small for all** $s \in [0,T]$.
- **Ingredient 2**: $K(u-s)$ **must be large for at least some** $s \in [0,u]$. The knowledge of S_u gives partial information on $(W_s, 0 \le s \le u)$, which is passed to $(Z_s, 0 \le s \le u)$ through the correlation ρ; this gives enough information on σ_u^2 only if $K(u-s)$ is large for at least some $s \in [0,u]$.[2]
- **Ingredient 1** + **Ingredient 2**: $K(\theta)$ **must be large for** $\theta \in [0, u-T]$ **and small for** $\theta \in [u-T, u]$.

$\Longrightarrow$ **Fast decreasing kernel**

[2] The knowledge of S_u may also give partial direct information on $(\sigma_s, 0 \le s \le u)$. Indeed, if S_u is extremely large, then many σ_s, $0 \le s \le u$, must have been very large, and σ_u is likely to be large. This explains why the smile has a positive slope at large strikes in stochastic volatility models even if $\rho < 0$. However, for values of S_u close to S_0, the knowledge of S_u is transferred to σ_u^2 mostly through the paths of W and Z up to u.

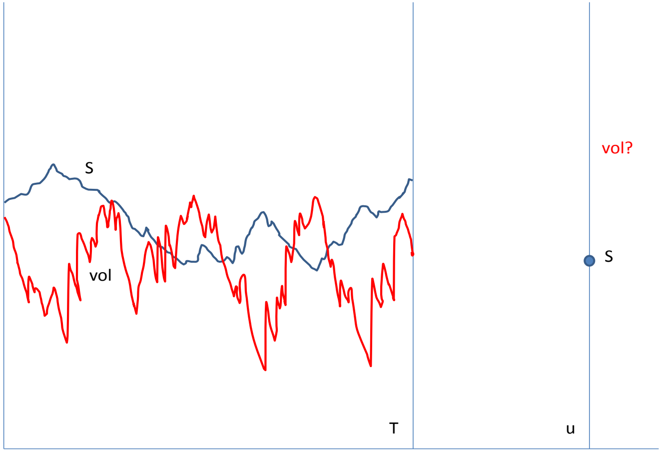
Ingredients needed for inversion of convex ordering

$$K(\theta) = \begin{cases} \text{large if } \theta \le \varepsilon, \\ \text{small if } \theta > \varepsilon \end{cases} \qquad \varepsilon \le \tau \tag{6.4}$$

- **Ingredient 1** holds for all $u \in [T+\varepsilon, T+\tau]$.
- **Ingredient 2** will hold only if $K(\theta)$ is large enough for $\theta \in [0,\varepsilon]$; this is needed for the limited information that S_u gives on $(dZ_s, u-\varepsilon \le s \le u)$ to be amplified enough by the kernel K so that it impacts $\mathbb{E}[\sigma_u^2|S_u]$.
- **Such a fast decreasing kernel K is reminiscent of fast mean-reversion**, with characteristic time $\varepsilon \le \tau = 30$ days.
- The fact that $K(\theta)$ must be very large for small θ means **very high volatility of instantaneous volatility**.
- **Rough volatility models** also have a fast decreasing kernel $(H < \frac{1}{2})$: $K(\theta) \sim \theta^{H-\frac{1}{2}}$, $K(0+) = +\infty$.

$\Longrightarrow$ **Fast mean-reversion and very large vol-of-vol** and **rough vol** are good candidates for inversion of convex ordering

Fast mean-reversion



Two remarks

- The smaller T, the more information S_u gives on $(Z_s, s \in [T,u])$ hence on σ_u^2 $(u \in [T, T+\tau) \Longrightarrow$ In those models, inversion of convex order is more likely to hold for small T. **This is precisely in line with market data**.
- Since $\sigma_{\text{loc}}(t, S_t^{\text{loc}})$ **does not mean revert**, contrary to σ_t, we expect $\mathbb{E}[\sigma_{\text{loc}}^2(u, S_u^{\text{loc}})|S_T^{\text{loc}}]$ to be only slightly smaller than $\sigma_{\text{loc}}^2(u, S_u^{\text{loc}})$ in convex order.

Exponential kernel

$$K(\theta) = \omega \exp(-k\theta), \qquad \omega \geq 0,\ k > 0$$

- One-factor Bergomi model (Dupire 1993, Bergomi 2005).
- ξ_t^u admits a one-dimensional Markov representation: $\xi_t^u = \xi_0^u f^u(t, X_t)$

$$f^u(t,x) = \exp\left(\omega e^{-k(u-t)} x - \frac{\omega^2}{2} e^{-2k(u-t)} \mathrm{Var}(X_t)\right), \quad \mathrm{Var}(X_t) = \frac{1 - e^{-2kt}}{2k}$$

Ornstein-Uhlenbeck process $X_t = \int_0^t e^{-k(t-s)} dZ_s$:

$$dX_t = -kX_t\, dt + dZ_t, \quad X_0 = 0. \tag{6.5}$$

- (S_t, X_t) Markov process.
- $k =$ mean-reversion.
- $\omega =$ vol of variance: it is the instantaneous (lognormal) volatility of the instantaneous variance $\sigma_t^2 := \xi_t^t$.
- **Inversion of convex ordering:** k **large,** $k \geq \frac{1}{\tau} \approx 12$**;** ω **large**.

Power-law kernel

$$K(\theta) = \nu\theta^{H-\frac{1}{2}}, \qquad \nu \geq 0, \qquad H \in \left(0, \frac{1}{2}\right) \tag{6.6}$$

- Rough Bergomi model (Bayer, Friz, Gatheral, 2014).
- $\lim_{\theta \to 0^+} K(\theta) = +\infty$.
- H = Hurst exponent.

$$\begin{aligned} \xi_t^u &= \xi_0^u \exp\left(\nu X_t^u - \frac{\nu^2}{2}\mathrm{Var}(X_t^u)\right) \\ X_t^u &= \int_0^t (u-s)^{H-\frac{1}{2}}\, dZ_s, \qquad \mathrm{Var}(X_t^u) = \frac{u^{2H} - (u-t)^{2H}}{2H}. \end{aligned}$$

- $H > 0$ ensures that $\mathrm{Var}(X_u^u)$ is finite.
- **Inversion of convex ordering: H small.**

Inversion of convex ordering: Numerical experiments

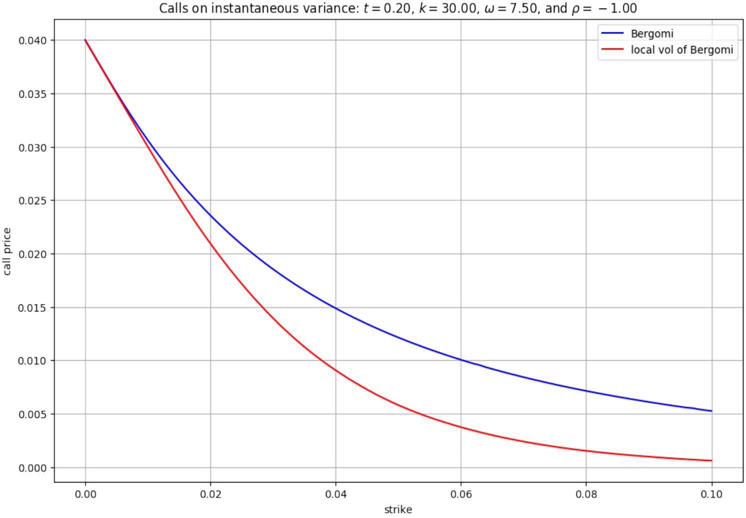
One-factor Bergomi model

$k = 30,\ \omega = 7.5,\ \rho = -1$

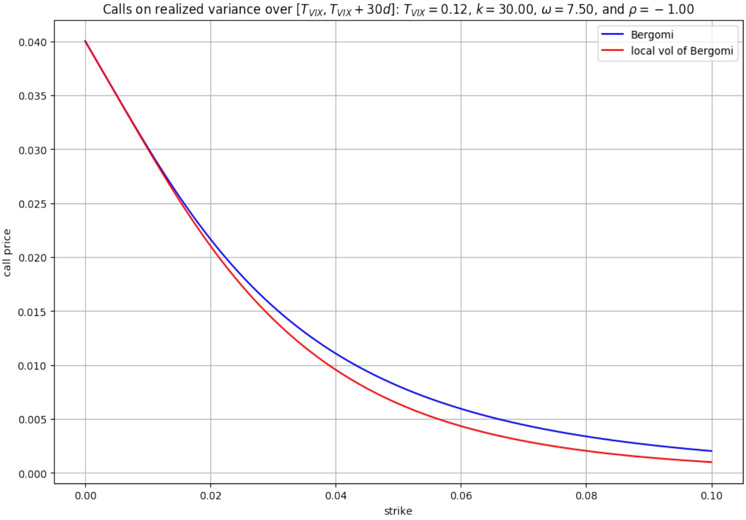
(Similar results in the rough Bergomi model with small H)

$$T = 0.12$$

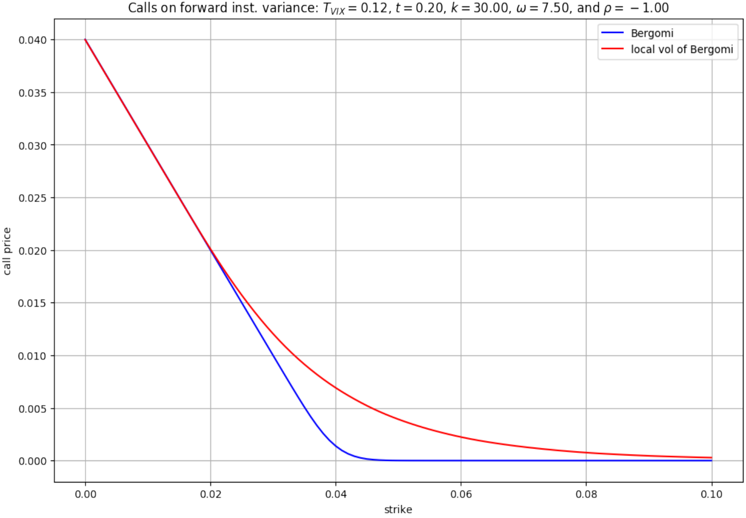
Call on instantaneous variance $\sigma^2_{T+\tau}$



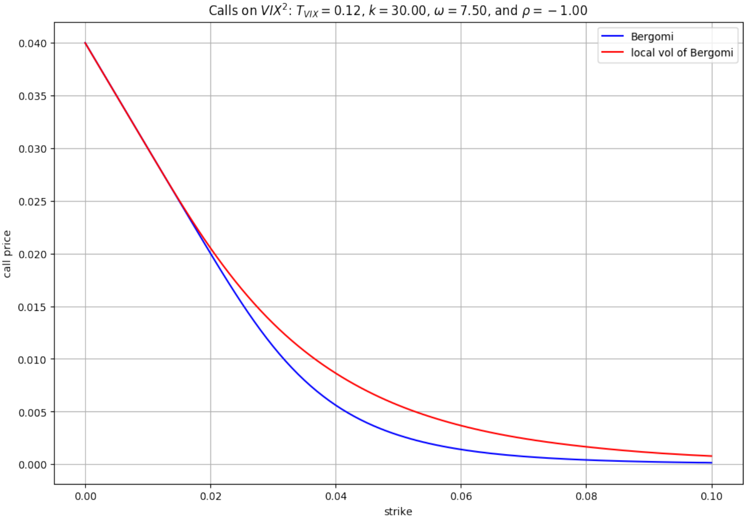
Call on realized variance $\frac{1}{\tau}\int_T^{T+\tau}\sigma_t^2\,dt$ (Sum)



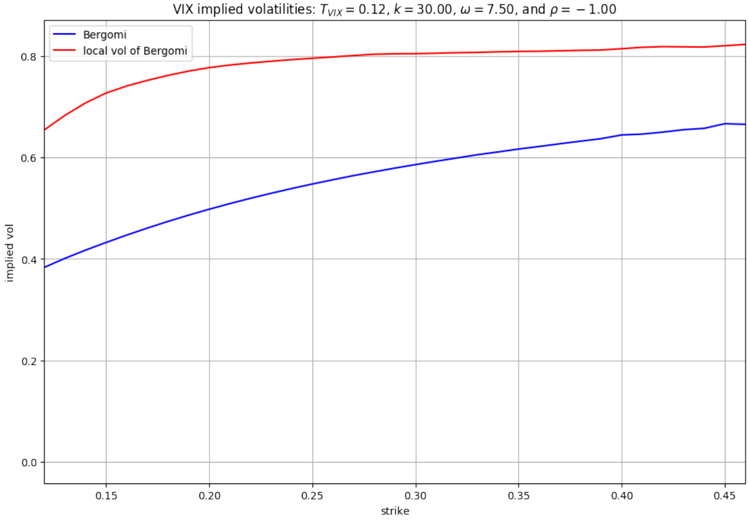
Call on forward instantaneous variance $\mathbb{E}[\sigma^2_{T+\tau}|\mathcal{F}_T]$ (Conditioning on $\mathcal{F}_T$)



Call on $\mathrm{VIX}_T^2 = \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau}\sigma_t^2\,dt \middle| \mathcal{F}_T\right]$ (Sum + Conditioning on $\mathcal{F}_T$)

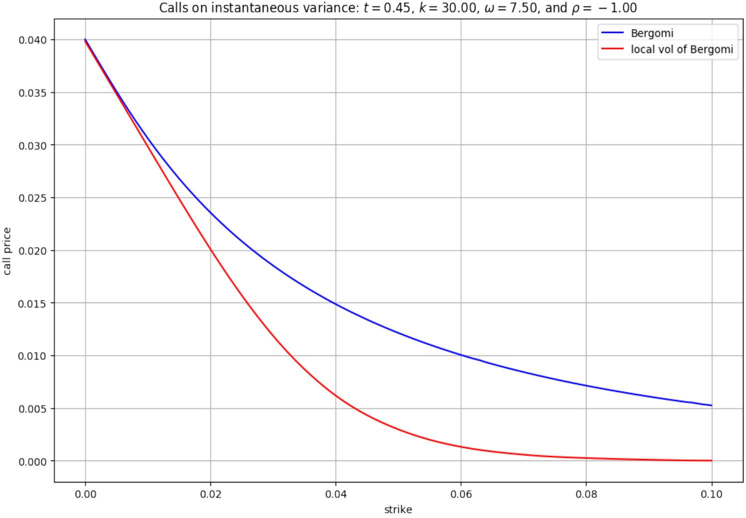


VIX implied volatility

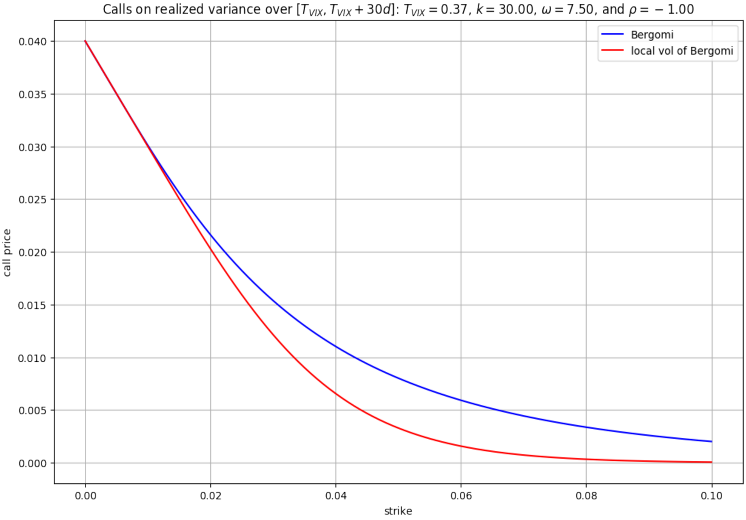


$$T = 0.37$$

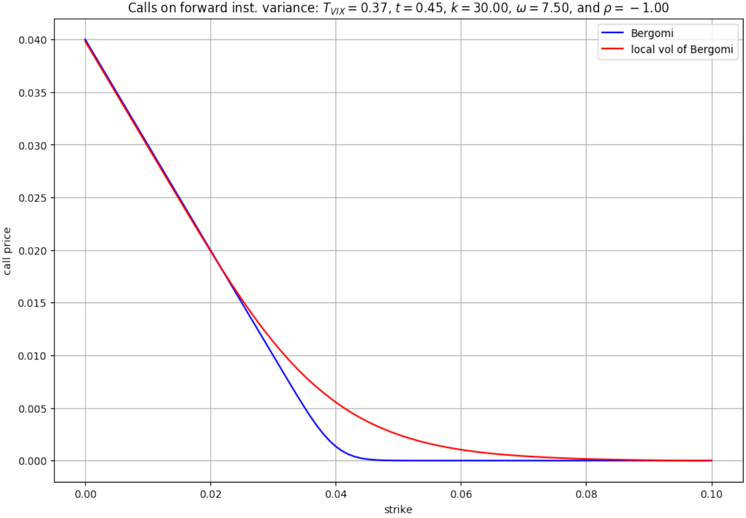
Call on instantaneous variance $\sigma^2_{T+\tau}$



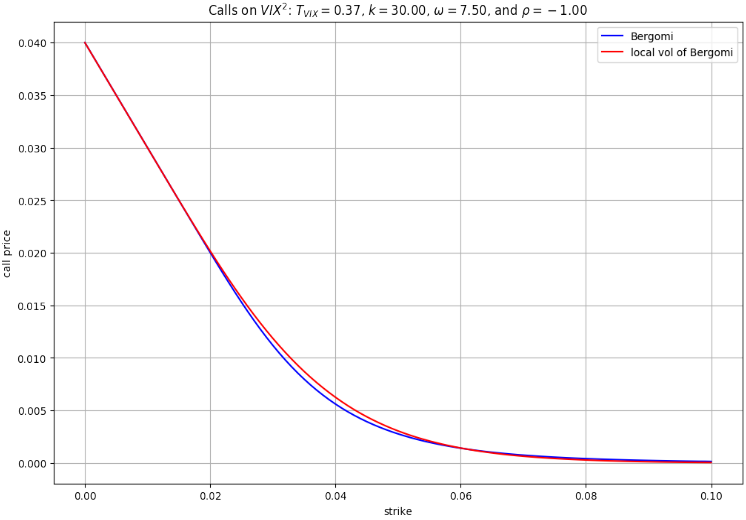
Call on realized variance $\frac{1}{\tau}\int_T^{T+\tau}\sigma_t^2\,dt$ (Sum)



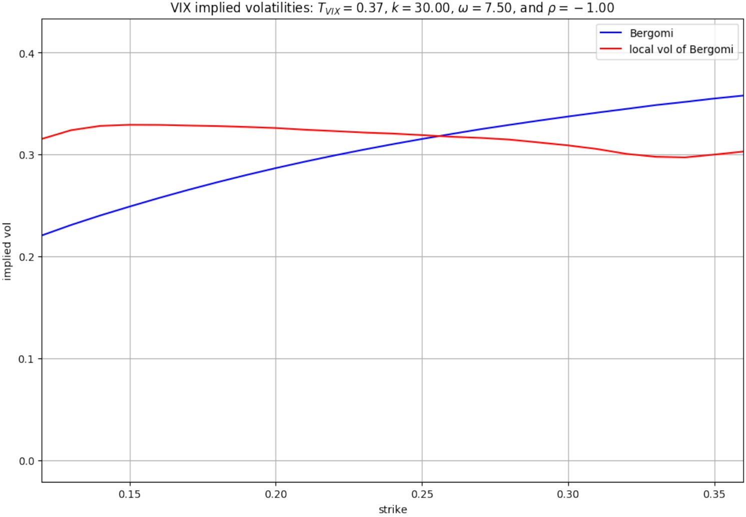
Call on forward instantaneous variance $\mathbb{E}[\sigma^2_{T+\tau}|\mathcal{F}_T]$ (Conditioning on $\mathcal{F}_T$)



Call on $\text{VIX}_T^2 = \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma_t^2\,dt \middle| \mathcal{F}_T\right]$ (Sum + Conditioning on $\mathcal{F}_T$)



VIX implied volatility



Small vol-of-vol asymptotics

Exponential kernel

$$K(\theta) = \omega \exp(-k\theta), \qquad \omega \geq 0,\ k > 0$$

Objectives:

- Use directly approximate formulas of SPX skew and VIX futures in the one-factor Bergomi model to prove that in the one-factor Bergomi model joint calibration requires large k and ω.
- Make this statement more precise: How big should be $\frac{\omega}{k}$? $\frac{\omega^2}{k}$?

Reminder on the ergodic regime:

- The limiting regime where k and ω tend to $+\infty$ while $\frac{\omega^2}{k}$ is kept **constant** corresponds to an **ergodic** limit where (ωX_t) quickly reaches its stationary distribution $\mathcal{N}(0, \frac{\omega^2}{2k})$. Cf Fouque, Papanicolaou and Sircar (2000).
- Only regime where k, ω are large and the variance of σ_t^2 has a finite limit, which is the natural regime in finance.

The SPX smile in the one-factor Bergomi model

Bergomi-G. expansion (2012) gives the smile of generic stochastic volatility models at order 2 in vol-of-vol:

$$\widehat{\sigma}(T,K) \;=\; \widehat{\sigma}_T^{\mathrm{ATM}} + \mathcal{S}_T \ln\left(\frac{K}{S_0}\right) + \mathcal{C}_T \ln^2\left(\frac{K}{S_0}\right) + O(\omega^3) \qquad (8.1)$$

In the case of the one-factor Bergomi model with a flat initial term structure of variance swaps ($\xi_0^t \equiv \xi$), coefficients are explicit functions of ρ, ω, k, ξ, T. In particular, the ATM skew

$$\mathcal{S}_T = \frac{\rho\omega}{2}\mathcal{J}(kT) + \frac{\rho^2\omega^2\sqrt{\xi}T}{8}\left(2\mathcal{H}(kT) + 4\frac{\mathcal{J}(kT)-\mathcal{J}(2kT)}{kT} - 3\mathcal{J}(kT)^2\right)$$

where

$$\mathcal{I}(\alpha) = \frac{1-e^{-\alpha}}{\alpha}, \qquad \mathcal{J}(\alpha) = \frac{\alpha-1+e^{-\alpha}}{\alpha^2}$$

$$\mathcal{K}(\alpha) = \frac{1-e^{-\alpha}-\alpha e^{-\alpha}}{\alpha^2}, \qquad \mathcal{H}(\alpha) = \frac{\mathcal{J}(\alpha)-\mathcal{K}(\alpha)}{\alpha}$$

The VIX future in the one-factor Bergomi model (G., 2018)

The price of VIX futures in the one-factor Bergomi model satisfies (G., 2018):

$$\mathbb{E}[\mathrm{VIX}_T] = \sqrt{\Xi_0}\left\{1 + \alpha_1\omega^2 V_T + \alpha_2(\omega^2 V_T)^2 + \alpha_3(\omega^2 V_T)^3\right\} + O(\omega^7)$$

where $V_T = \mathsf{Var}(X_T) = \frac{1-e^{-2kT}}{2k} = T\mathcal{I}(2kT)$,

$$\begin{aligned}
\alpha_1 &= -\frac{1}{8}I_1^2,\\
\alpha_2 &= -\frac{1}{16}I_2^2 + \frac{3}{16}I_1^2 I_2 - \frac{15}{128}I_1^4,\\
\alpha_3 &= -\frac{1}{48}I_3^2 + \frac{1}{16}I_2^3 + \frac{3}{16}I_1 I_2 I_3 - \frac{75}{128}I_1^2 I_2^2 - \frac{5}{32}I_1^3 I_3 + \frac{105}{128}I_1^4 I_2 - \frac{315}{1024}I_1^6
\end{aligned}$$

and

$$\Xi_n := \frac{1}{\tau}\int_T^{T+\tau} \xi_0^u e^{-nk(u-T)}du, \qquad I_n := \frac{\Xi_n}{\Xi_0}, \qquad n \in \mathbb{N}$$

When $\xi_0^u \equiv \xi$, $\Xi_n = \xi\mathcal{I}(nk\tau)$ and $I_n = \mathcal{I}(nk\tau)$ are known in closed form.

$$\mathcal{S}_T \approx \frac{\rho\omega}{2}\frac{kT-1+e^{-kT}}{(kT)^2}$$

$$\sqrt{\xi}-\mathbb{E}[\mathrm{VIX}_T] \approx \frac{1}{8}\sqrt{\xi}\left(\frac{1-e^{-k\tau}}{k\tau}\right)^2\omega^2\frac{1-e^{-2kT}}{2k}$$

Small mean-reversion: cannot jointly calibrate

- $\mathcal{S}_T \approx \frac{\rho\omega}{4}$. Calibration to short-term SPX smile: $\mathcal{S}_T \approx -1.5$ $\Longrightarrow \rho\omega \approx -6 \Longrightarrow \omega \geq 6$.
- $\sqrt{\xi}-\mathbb{E}[\mathrm{VIX}_T] \approx \frac{1}{8}\sqrt{\xi}\omega^2 T \Longrightarrow$ implied vol of $\mathrm{VIX}_T^2 \approx \omega \geq 6$: too large compared to market data!

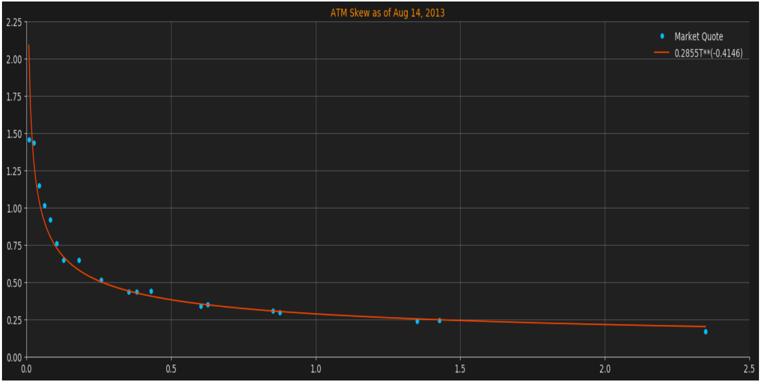
Large mean-reversion:

- $\mathcal{S}_T \approx \frac{\rho\omega}{2kT}$, $kT \gg 1$. Calibration to SPX smile, $T=\frac{1}{4}$: $\frac{\rho\omega}{2kT} \approx -0.6 \Longrightarrow 2\frac{\rho\omega}{k} \approx -0.6 \Longrightarrow \frac{\omega}{k} \geq 0.3$: ω and k are large. Numerical example: $k=30, \rho=-1 \Longrightarrow \omega=9$
- $\sqrt{\xi}-\mathbb{E}[\mathrm{VIX}_T] \approx \frac{1}{16}\sqrt{\xi}\frac{\omega^2}{k^3\tau^2}$ behaves like $\frac{\omega^2}{k^3}$ instead of $\omega^2 T$! Because of mean-reversion, implied vol of VIX_T^2 is much smaller. Numerical example: $\sqrt{\xi}-\mathbb{E}[\mathrm{VIX}_T] \approx \frac{1}{16}\sqrt{\xi}\frac{9^2}{30^3(\frac{1}{12})^2} = 0.027\sqrt{\xi}$.

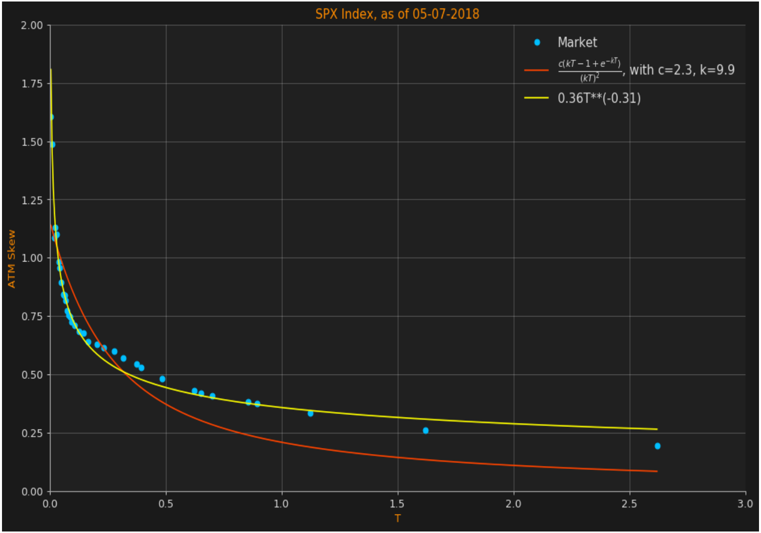
$\Longrightarrow$ Both ω and k must be large, with $\omega \approx k$ so $\frac{\omega^2}{k}$ large!
Very large stationary vol of instantaneous vol.

Term-structure of SPX ATM skew

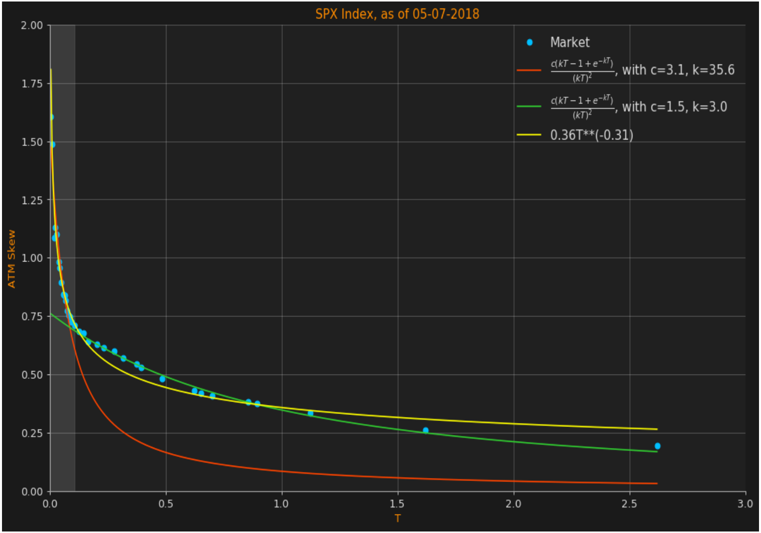
One-factor Bergomi model with large mean-reversion and vol-of-vol: $\mathcal{S}_T \sim \frac{1}{T}$. To mimic or produce a power-law decay $\mathcal{S}_T \sim \frac{1}{T^\alpha}$: 2-factor Bergomi model and rough volatility model.



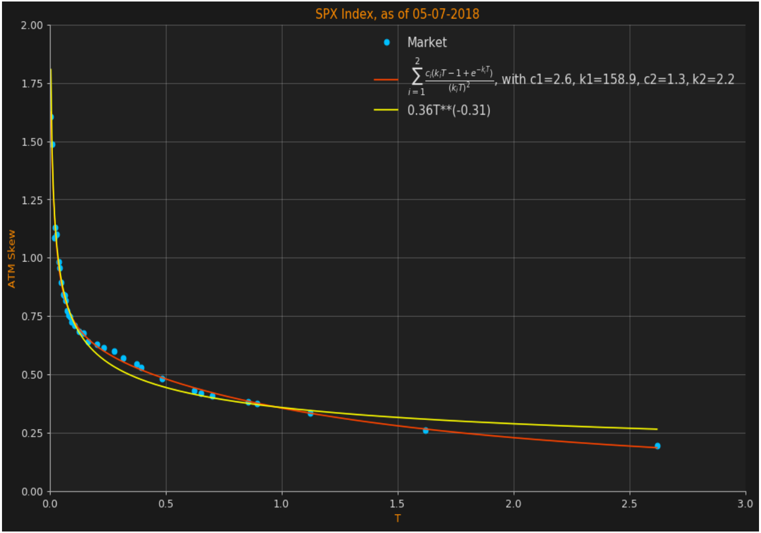
SPX ATM skew, May 7, 2018



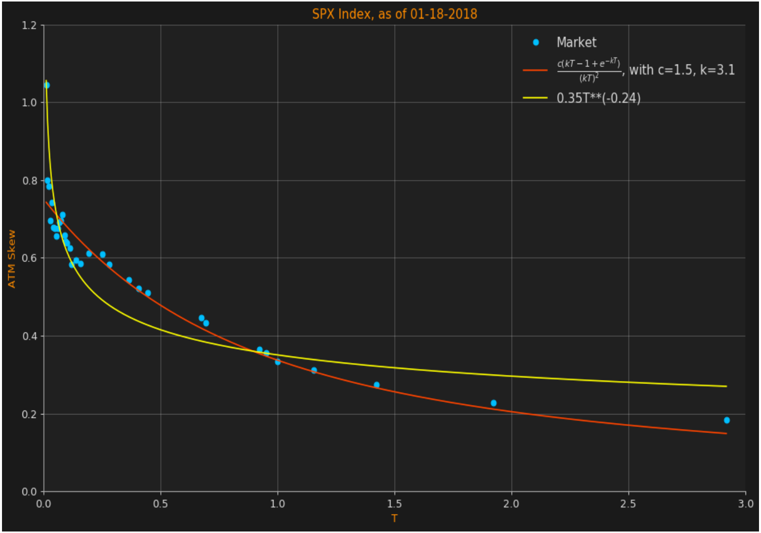
SPX ATM skew, May 7, 2018



SPX ATM skew, May 7, 2018



However... SPX ATM skew, Jan 18, 2018



Power-law kernel: $K(\theta) = \nu\theta^{H-\frac{1}{2}}$

- **No Markov representation** for ξ_t^u.
- **Instantaneous variance $\sigma_t^2 := \xi_t^t$ is not a semimartingale**. One cannot write Itô dynamics $d\xi_t^t = \cdots dt + \cdots dZ_t$ for the instantaneous variance. **No notion of a dynamic volatility of instantaneous spot variance**.
- However we can compare the values of $\mathrm{Var}\left(\ln \frac{\xi_t^u}{\xi_0^u}\right)$ in the power-law and exponential kernel models:

$$\nu^2 \frac{u^{2H} - (u-t)^{2H}}{2H} \quad \longleftrightarrow \quad \omega^2 e^{-2k(u-t)} \frac{1 - e^{-2kt}}{2k} \tag{8.2}$$

$$u = t \to 0 : \quad \nu^2 \frac{t^{2H}}{2H} \quad \longleftrightarrow \quad \omega^2 \frac{1 - e^{-2kt}}{2k} \approx \omega^2 t \tag{8.3}$$

$$\nu \frac{t^{H-\frac{1}{2}}}{\sqrt{2H}} \quad \longleftrightarrow \quad \omega \tag{8.4}$$

- $\nu \frac{t^{H-\frac{1}{2}}}{\sqrt{2H}}$ **can be interpreted as a short term volatility of instantaneous spot variance.**
- $[\nu] = \text{time}^{-H}$; $\left[\nu\theta^{H-\frac{1}{2}}\right] = \left[\nu \frac{t^{H-\frac{1}{2}}}{\sqrt{2H}}\right] =$ vol.

Power-law kernel: $K(\theta) = \nu\theta^{H-\frac{1}{2}}$

- Short-term ATM skew in SV models $\sim \rho\omega$. **Explains why the ATM skew in such rough volatility models behaves like $T^{H-\frac{1}{2}}$ for short maturities** T (Alós, Fukasawa...), which is one of the reasons why this model has been introduced (Gatheral, Jaisson, Rosenbaum, Friz, Bayer).
- In the limit $H \to 0$, for fixed ν, $\nu^2 \frac{t^{2H}}{2H} \to +\infty$ for any $t > 0$.
- In order for $\mathrm{Var}(\sigma_t^2)$ to tend to a finite limit, we must impose that $\frac{\nu^2}{2H}$ tend to a finite limit $\Longrightarrow$ A natural limiting regime, analogous to the **ergodic regime** described above for the exponential kernel, is **$H, \nu \to 0$, with $\frac{\nu^2}{2H}$ kept constant**.
- However in this ergodic limit the SPX skew is $\sim \sqrt{H}T^{H-\frac{1}{2}}$...

Calibrating first to VIX market

Skewing the models on ξ_t^u:

- Following Bergomi (2008), we use a linear combination of two lognormal random variables to model the instantaneous variance σ_t^2 so as to generate positive VIX skew:

$$\sigma_t^2 = \xi_0^t \left((1-\lambda)\mathcal{E}\left(\omega_0 \int_0^t e^{-k(t-s)} dZ_s\right) + \lambda\mathcal{E}\left(\omega_1 \int_0^t e^{-k(t-s)} dZ_s\right)\right)$$

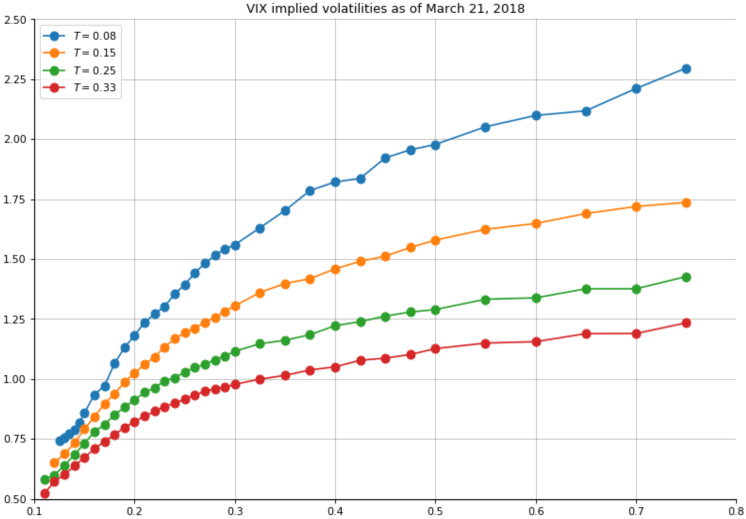
or

$$\sigma_t^2 = \xi_0^t \left((1-\lambda)\mathcal{E}\left(\nu_0 \int_0^t (t-s)^{H-\frac{1}{2}} dZ_s\right) + \lambda\mathcal{E}\left(\nu_1 \int_0^t (t-s)^{H-1/2} dZ_s\right)\right)$$

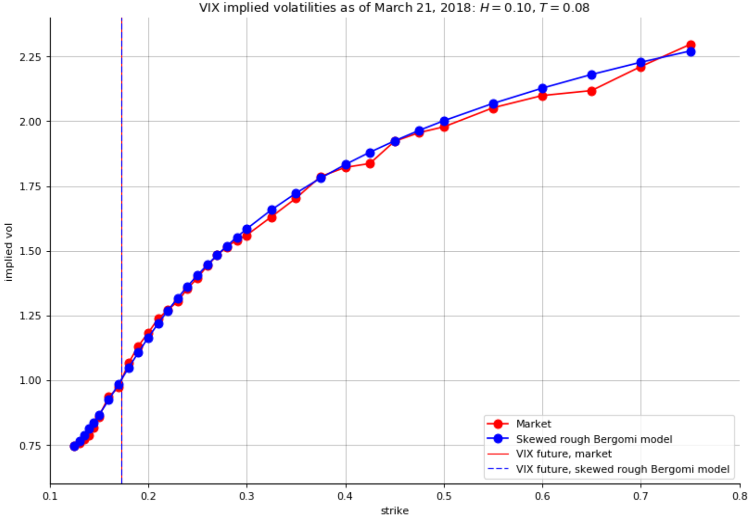
with $\lambda \in [0,1]$.

- $\mathcal{E}(X)$ is simply a shorthand notation for $\exp\left(X - \frac{1}{2}\mathrm{Var}(X)\right)$.
- Similar idea recently used and developed by De Marco.

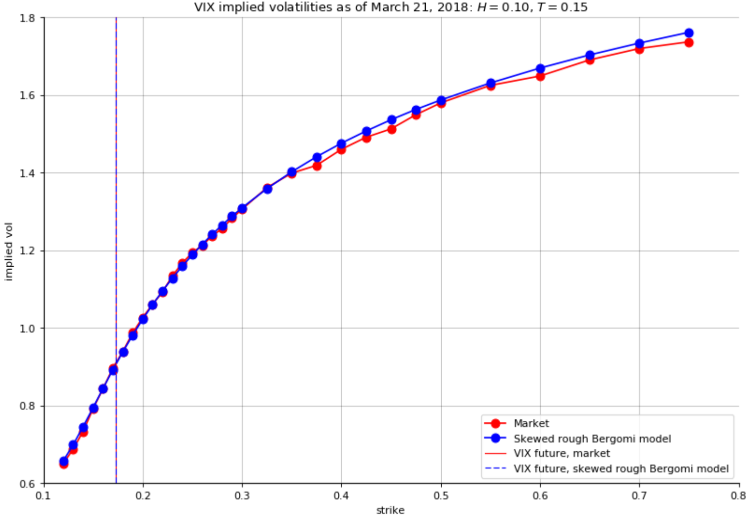
Skewed rough Bergomi: Calibration to VIX future and VIX options (March 21, 2018



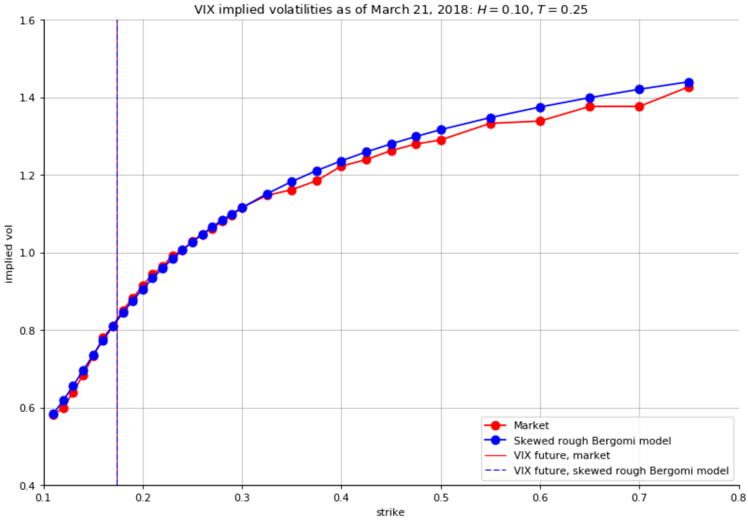
Skewed rough Bergomi: Calibration to VIX future and VIX options (March 21, 2018



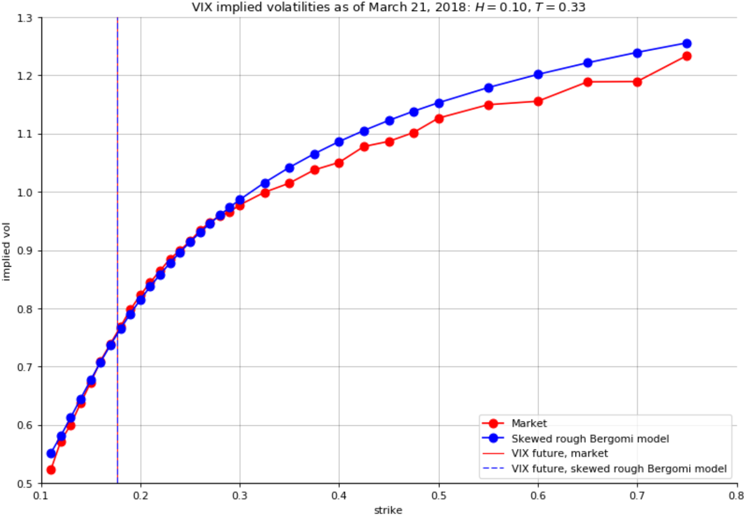
Skewed rough Bergomi: Calibration to VIX future and VIX options



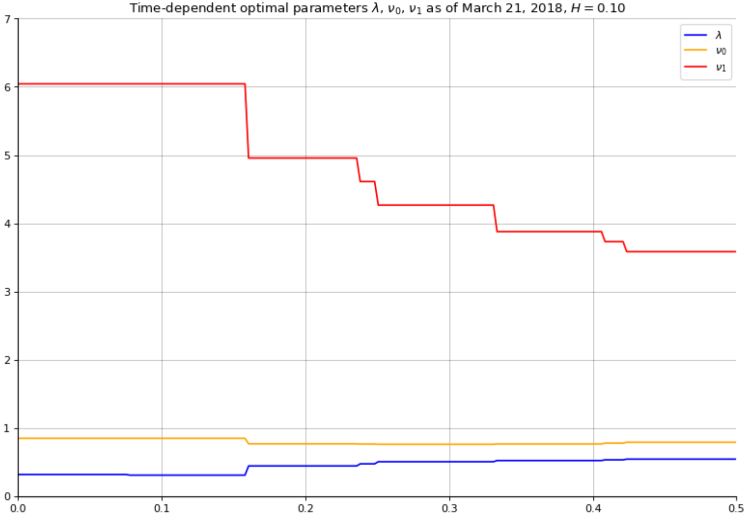
Skewed rough Bergomi: Calibration to VIX future and VIX options



Skewed rough Bergomi: Calibration to VIX future and VIX options

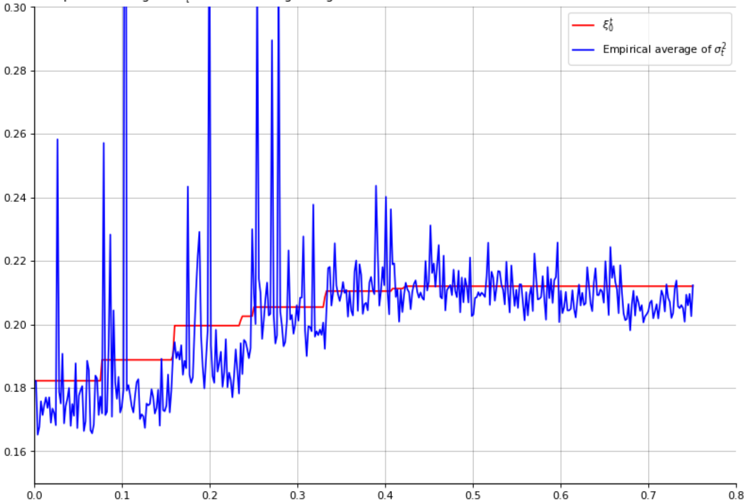


Skewed rough Bergomi: Calibration to VIX future and VIX options (March 21, 2018

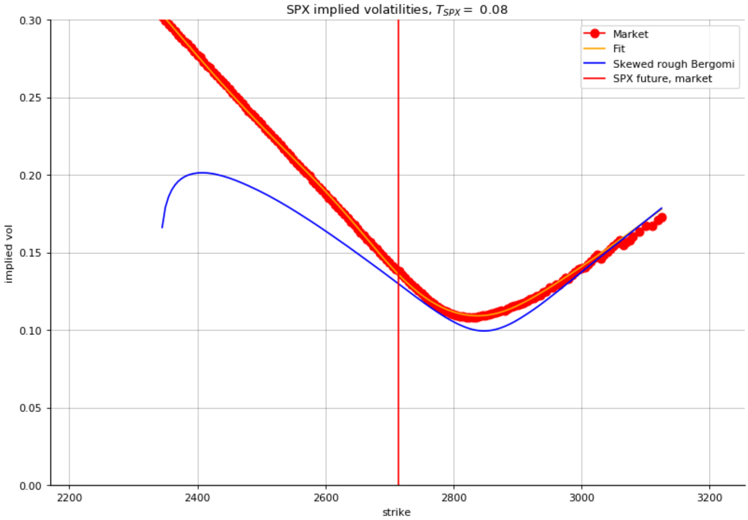


Skewed rough Bergomi: Calibration to VIX future and VIX options (March 21, 2018

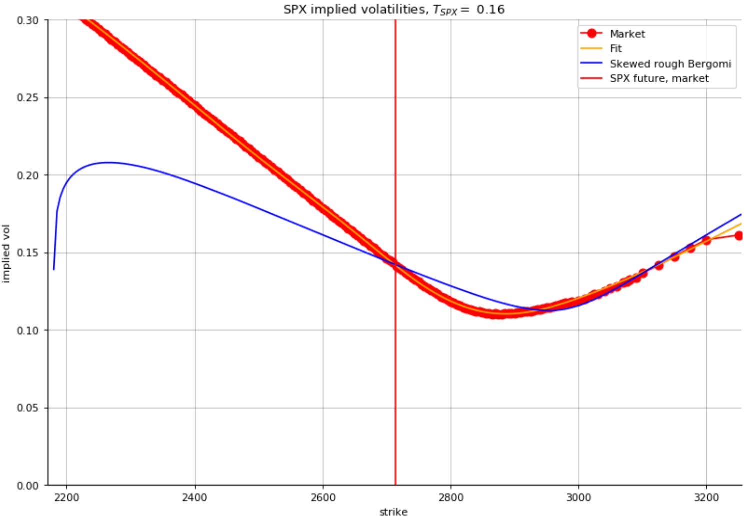
Empirical average of σ_t^2 in skewed rough Bergomi model calibrated on March 21, 2018, $H = 0.10$, N=400000



Skewed rough Bergomi calibrated to VIX: SPX smile



Skewed rough Bergomi calibrated to VIX: SPX smile



Skewed rough Bergomi calibrated to VIX: SPX smile

Skewed rough Bergomi calibrated to VIX: SPX smile

Calibrating first to SPX market (work in progress)

Consider only continuous models on SPX that are calibrated to SPX smile:

$$\frac{dS_t}{S_t} = \frac{a_t}{\sqrt{\mathbb{E}[a_t^2|S_t]}}\sigma_{\mathrm{lv}}(t, S_t)\, dW_t$$

and optimize on (a_t) so as to match VIX options — or compute the infimum of VIX implied vols within those models.

Example: $a_t = \sigma_i(X_t),\ t \in [T_i, T_i + \tau]$

- The **leverage function**

$$l(t, S_t) = \frac{\sigma_{\mathrm{lv}}(t, S_t)}{\sqrt{\mathbb{E}[a_t^2|S_t]}}$$

  does not mean revert; it **fights against inversion of convex ordering**.
- Numerically difficult to estimate

$$\mathrm{VIX}_{T_i}^2 = \frac{1}{\tau}\int_{T_i}^{T_{i+1}} \mathbb{E}\left[\frac{\sigma_i(X_t)^2}{\mathbb{E}[\sigma_i(X_t)^2|S_t]}\sigma_{\mathrm{loc}}(t, S_t)^2 \middle| \mathcal{F}_{T_i}\right] dt$$

  (use neural networks)

Why jumps can help

- For a continuous model to calibrate jointly to SPX and VIX options, the distribution of $\mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau}\sigma_t^2\,dt\middle|\mathcal{F}_T\right]$ should be as narrow as possible, but without killing the SPX skew. The problem of ergodic/stationary (σ_t) is that they produce flat SPX skew.
- Jump-Lévy processes are precisely examples of processes that can generate deterministic realized variance together with a smile on the underlying.
- This explains why jumps have proved useful in this problem.

Acknowledgements

I would like to thank Bryan Liang for his help, fruitful discussions, and producing some of the graphs.

A few selected references

Beiglböck, M., Juillet, N.: *On a problem of optimal transport under marginal martingale constraints*, Ann. Probab. 44(1):42–106, 2016.

Baldeaux, J, Badran, A.: *Consistent Modelling of VIX and Equity Derivatives Using a 3/2 plus Jumps Model*, Appl. Math. Finance 21(4):299–312, 2014.

Bayer, C., Friz, P., Gatheral, J.: *Pricing under rough volatility*, Quantitative Finance, 16(6):887–904, 2016.

Bennedsen, M., Lunde, A., Pakkanen, M.: *Hybrid scheme for Brownian semistationary processes*, Finance and Stochastics 21(4):931–965, 2017.

Bergomi, L.: *Smile dynamics II*, Risk, October, 2005.

Bergomi, L.: *Smile Dynamics III*, Risk, March 2008.

Bergomi, L., Guyon, J.: *Stochastic volatility's orderly smiles*, Risk, May, 2012.

Carr, P., Madan, D.: *Joint modeling of VIX and SPX options at a single and common maturity with risk management applications*, IIE Transactions 46(11):1125–1131, 2014.

A few selected references

Bergomi L., Guyon, J.: *Stochastic volatility's orderly smiles*, Risk Magazine, May 2012.

Carr, P., Nadtochiy, S.: *Local variance gamma and explicit calibration to option prices*, Math. Finance 27(1):151–193, 2017.

Cont, R., Kokholm, T.: *A consistent pricing model for index options and volatility derivatives*, Math. Finance 23(2):248–274, 2013.

De Marco, S., Henry-Labordere, P.: *Linking vanillas and VIX options: A constrained martingale optimal transport problem*, SSRN, 2015.

Fouque, J.-P., Saporito, Y.: *Heston Stochastic Vol-of-Vol Model for Joint Calibration of VIX and S&P 500 Options*, arXiv preprint, 2017. Available at arxiv.org/abs/1706.00873.

Gatheral, J.: *Consistent modeling of SPX and VIX options*, presentation at Bachelier Congress, 2008.

Gatheral, J., Jaisson, T., Rosenbaum, M.: *Volatility is rough*, to appear in Quantitative Finance, 2014.

Guyon, J.: *On the joint calibration of SPX and VIX options*, Conference in honor of Jim Gatheral's 60th birthday, NYU Courant, 2017.

Horvath, B., Jacquier, A., Tankov, P.: *Volatility options in rough volatility models*, preprint, 2017.

Jacquier, A., Martini, C., Muguruza, A.: *On the VIX futures in the rough Bergomi model*, preprint, 2017.

Kokholm, T., Stisen, M.: *Joint pricing of VIX and SPX options with stochastic volatility and jump models*, The Journal of Risk Finance 16(1):27–48, 2015.

A few selected references

Pacati, C., Pompa, P., Renò, R.: *Smiling Twice: The Heston++ Model*, SSRN preprint, 2015. Available at ssrn.com/abstract=2697179.

Papanicolaou, A., Sircar, R.: *A regime-switching Heston model for VIX and S&P 500 implied volatilities*, Quantitative Finance 14(10):1811–1827, 2014.

Henry-Labordère, P.: *Model-free Hedging: A Martingale Optimal Transport Viewpoint*, Chapman & Hall/CRC Financial Mathematics Series, 2017.

Sepp, A.: *Achieving consistent modeling of VIX and equity derivatives*, Imperial College Mathematical Finance Seminar, November 2, 2011.

Strassen, V.: *The existence of probability measures with given marginals*, Ann. Math. Statist., 36:423–439, 1965.