The Joint S&P 500/VIX Smile Calibration Puzzle Solved

Julien Guyon

Bloomberg, Quantitative Research
Columbia University, Department of Mathematics
NYU, Courant Institute of Mathematical Sciences

Bachelier Finance Society
One World Seminar
July 16, 2020

jguyon2@bloomberg.net, jg3601@columbia.edu, julien.guyon@nyu.edu

© 2020 Bloomberg Finance L.P. All rights reserved.

Risk, April 2020

Cutting edge: Volatility

The joint S&P 500/Vix smile calibration puzzle solved

Since Vix options started trading in 2006, many researchers have tried to build a model that jointly and exactly calibrates to the prices of Standard & Poor's 500 options, Vix futures and Vix options. In this article, Julien Guyon solves this long-standing puzzle by casting it as a discrete-time dispersion-constrained martingale transport problem, which he solves in a non-parametric way using Sinkhorn's algorithm

Volatility indexes, such as the Vix index, do not just serve as market-implied indicators of volatility. Futures and options on these indexes are also widely used as risk management tools to hedge the volatility exposure of options portfolios. The existence of a liquid market for these futures and options has led to the need for models that jointly calibrate to the prices of options on the underlying asset and the prices of volatility derivatives. Without such models, financial institutions could possibly arbitrage each other: even market-making desks within the same institution could do so, eg, the Vix desk could arbitrage the S&P 500 (SPX) desk. By using models that fail to correctly incorporate the prices of the hedging instruments, such as SPX options, Vix futures and Vix options, exotic desks may misprice options, especially (but not only) those with payoffs that involve both the underlying and its volatility index.

For this reason, since Vix options began trading in 2006, many researchers and practitioners have tried to build a model that jointly and exactly calibrates to the prices of SPX futures, SPX options, Vix futures and Vix options. This is known to be a very challenging problem, especially for short maturities. In particular, the very large negative skew of short-term SPX options, which in continuous models implies a very large volatility of volatility (vol-of-vol), seems inconsistent with the comparatively low levels of Vix implied and Vix smiles: that the distribution of the Dupire market local variance be smaller than the distribution of the (instantaneous) Vix squared in the convex order, at all times. He also reported that for short maturities the distribution of the true Vix squared in the market local volatility model is actually larger than the market-implied distribution of the true Vix squared in the convex order. Guyon showed numerically that when the (typically negative) spot-vol correlation is large enough in absolute value, both (a) traditional stochastic volatility models with large mean reversion and (b) rough volatility models with a small Hurst exponent can reproduce this inversion of convex ordering. Acciaio & Guyon (2020) provide a mathematical proof that the inversion of convex ordering can be produced by continuous models. However, the inversion of convex ordering is only a necessary condition for the joint SPX/Vix calibration of continuous models; it is not sufficient.

Since it looks to be very difficult to jointly calibrate the SPX and Vix smiles with continuous models, many authors have incorporated jumps in the dynamics of the SPX: see references in Guyon (2019a). Jumps offer extra degrees of freedom to partly decouple the ATM SPX skew and the ATM Vix implied volatility. However, the attempts so far have only produced imperfect, approximate fits.

Risk Digital Media Services

© 2020 Bloomberg Finance L.P. All rights reserved.

Motivation

- Volatility indices, such as the VIX index, are not only used as market-implied indicators of volatility.
- Futures and options on these indices are also widely used as risk-management tools to hedge the volatility exposure of options portfolios.
- Existence of a liquid market for these futures and options $\Longrightarrow$ need for models that jointly calibrate to the prices of options the underlying asset and prices of volatility derivatives.
- Calibration of stochastic volatility models to liquid hedging instruments: e.g., S&P 500 (SPX) options + VIX futures and options.
- Since VIX options started trading in 2006, many researchers and practitioners have tried to build a model that jointly and exactly calibrates to the prices of SPX options, VIX futures and VIX options.
- **Very challenging problem, especially for short maturities**.

© 2020 Bloomberg Finance L.P. All rights reserved.

Motivation

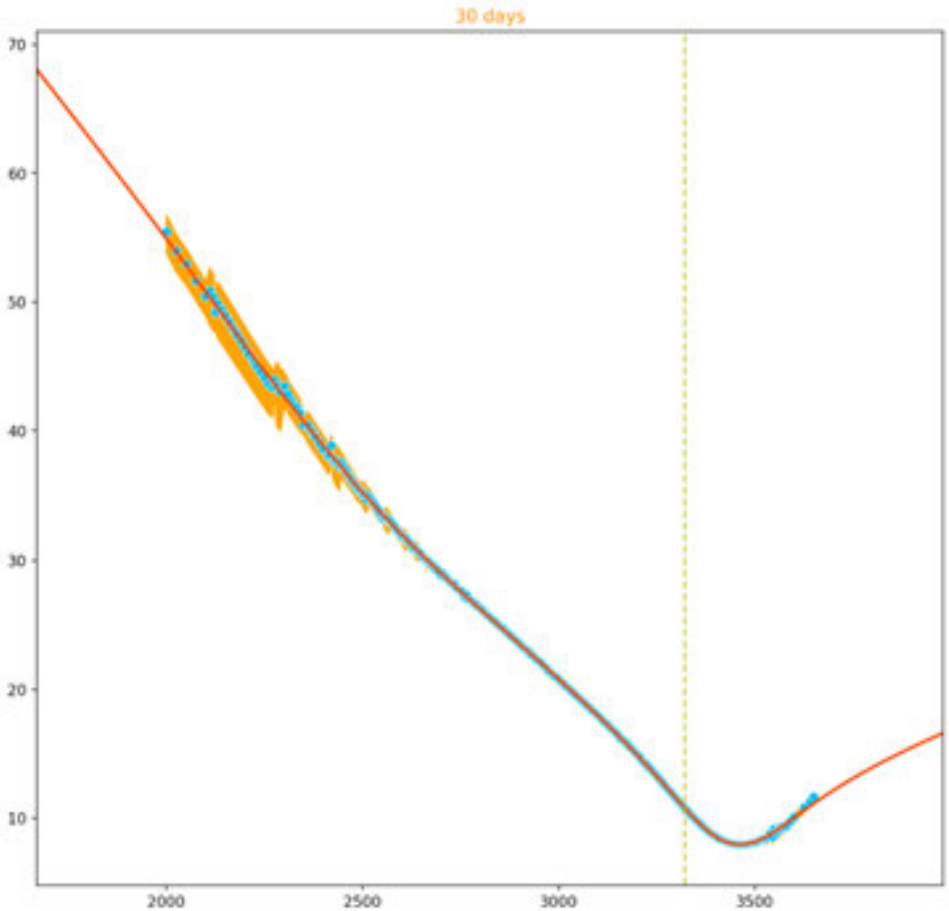


Figure: SPX smile as of January 22, 2020, $T = 30$ days

© 2020 Bloomberg Finance L.P. All rights reserved.

Motivation

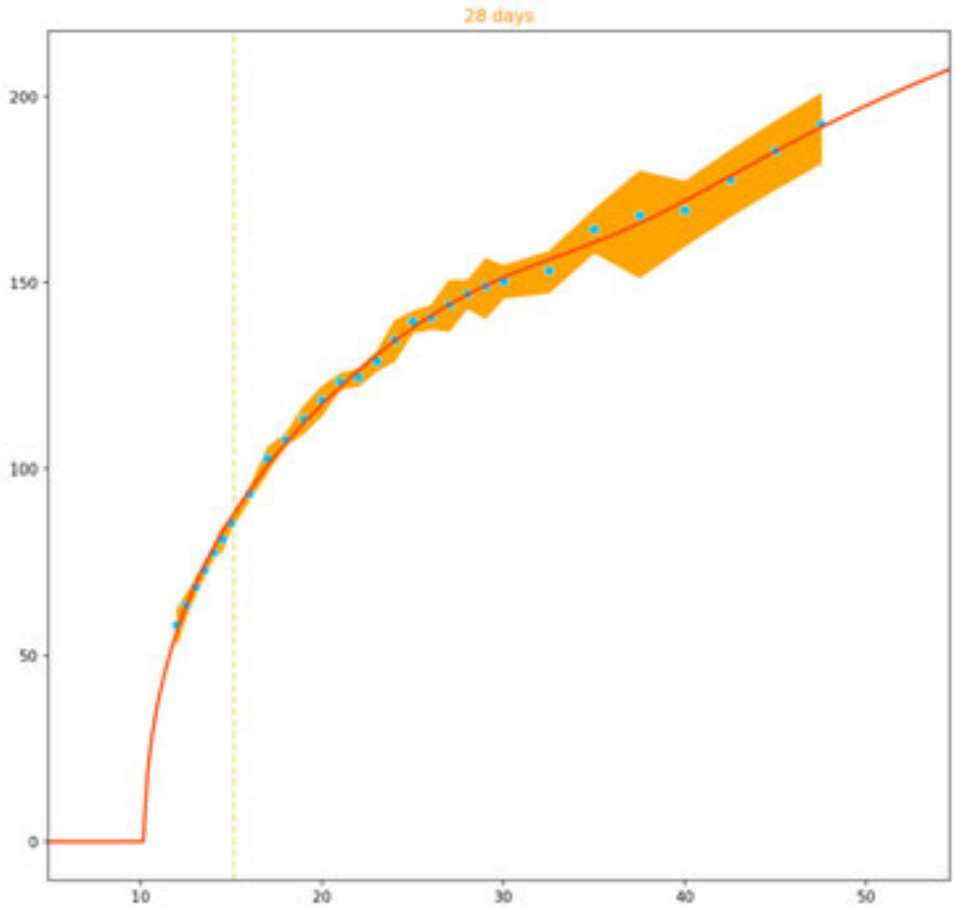


Figure: VIX smile as of January 22, 2020, $T = 28$ days

© 2020 Bloomberg Finance L.P. All rights reserved.

Motivation

- ATM skew:

$$\text{Definition:} \quad \mathcal{S}_T = \frac{d\sigma_{\mathrm{BS}}(K,T)}{\frac{dK}{K}}\Big|_{K=F_T}$$

$$\text{SPX, small } T\text{:} \quad \mathcal{S}_T \approx -1.5$$

$$\text{Classical one-factor SV model:} \quad \mathcal{S}_T \underset{T\to 0}{\longrightarrow} \frac{1}{2} \times \text{spot-vol correl} \times \text{vol-of-vol}$$

- Calibration to short-term ATM SPX skew $\Longrightarrow$

$$\text{vol-of-vol} \geq 3 = 300\% \gg \text{short-term ATM VIX implied vol}$$

The **very large negative skew of short-term SPX options**, which in continuous models implies a very large volatility of volatility, **seems inconsistent with the comparatively low levels of VIX implied volatilities**.

© 2020 Bloomberg Finance L.P. All rights reserved.

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

© 2020 Bloomberg Finance L.P. All rights reserved.

Double CEV dynamics

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

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

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*.

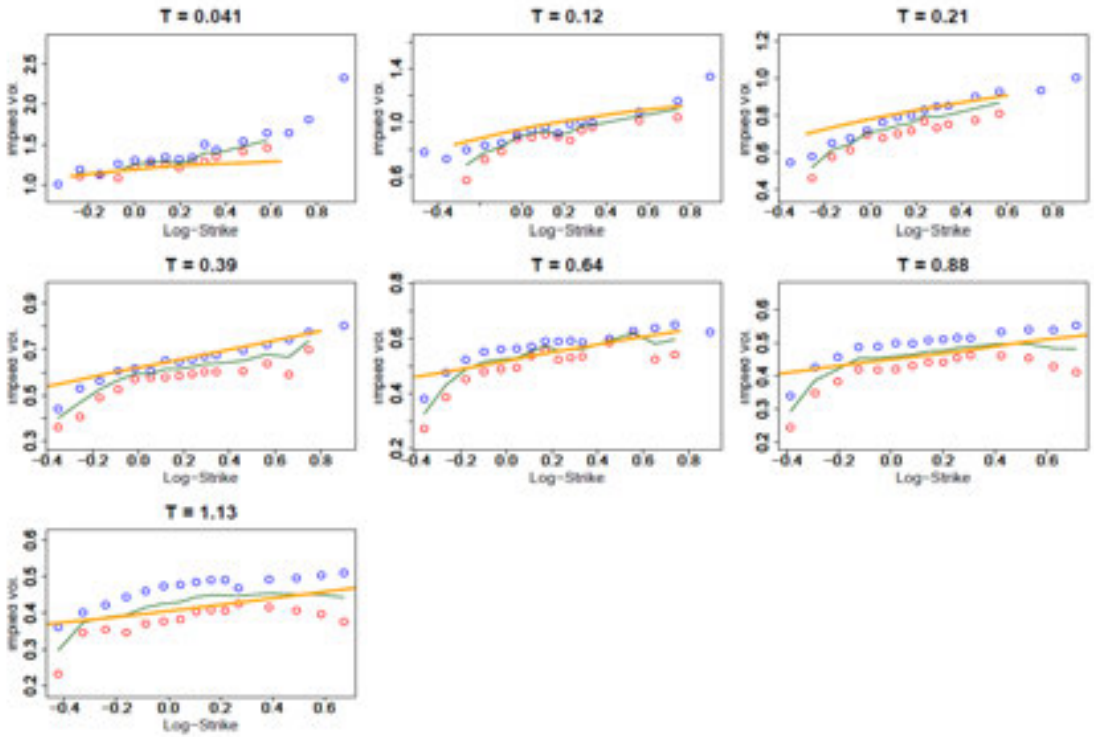
- All such models involve a short term variance level v that reverts to a moving level v' at rate κ. v' reverts to the long-term level z_3 at the slower rate $c < \kappa$.

© 2020 Bloomberg Finance L.P. All rights reserved.

Consistent Modeling of SPX and VIX options
The Double CEV model
Calibration of ξ_1, ξ_2 to VIX option prices

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):

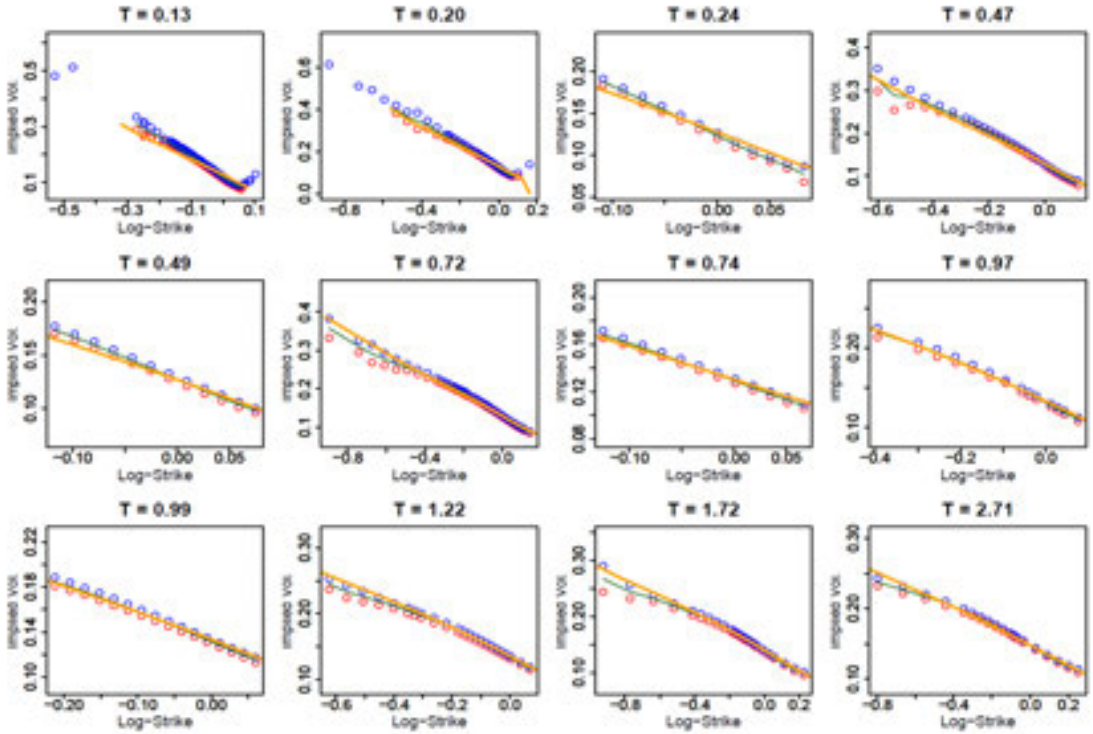


© 2020 Bloomberg Finance L.P. All rights reserved.

Consistent Modeling of SPX and VIX options
The Double CEV model
Calibration of ρ_1 and ρ_2 to SPX option prices

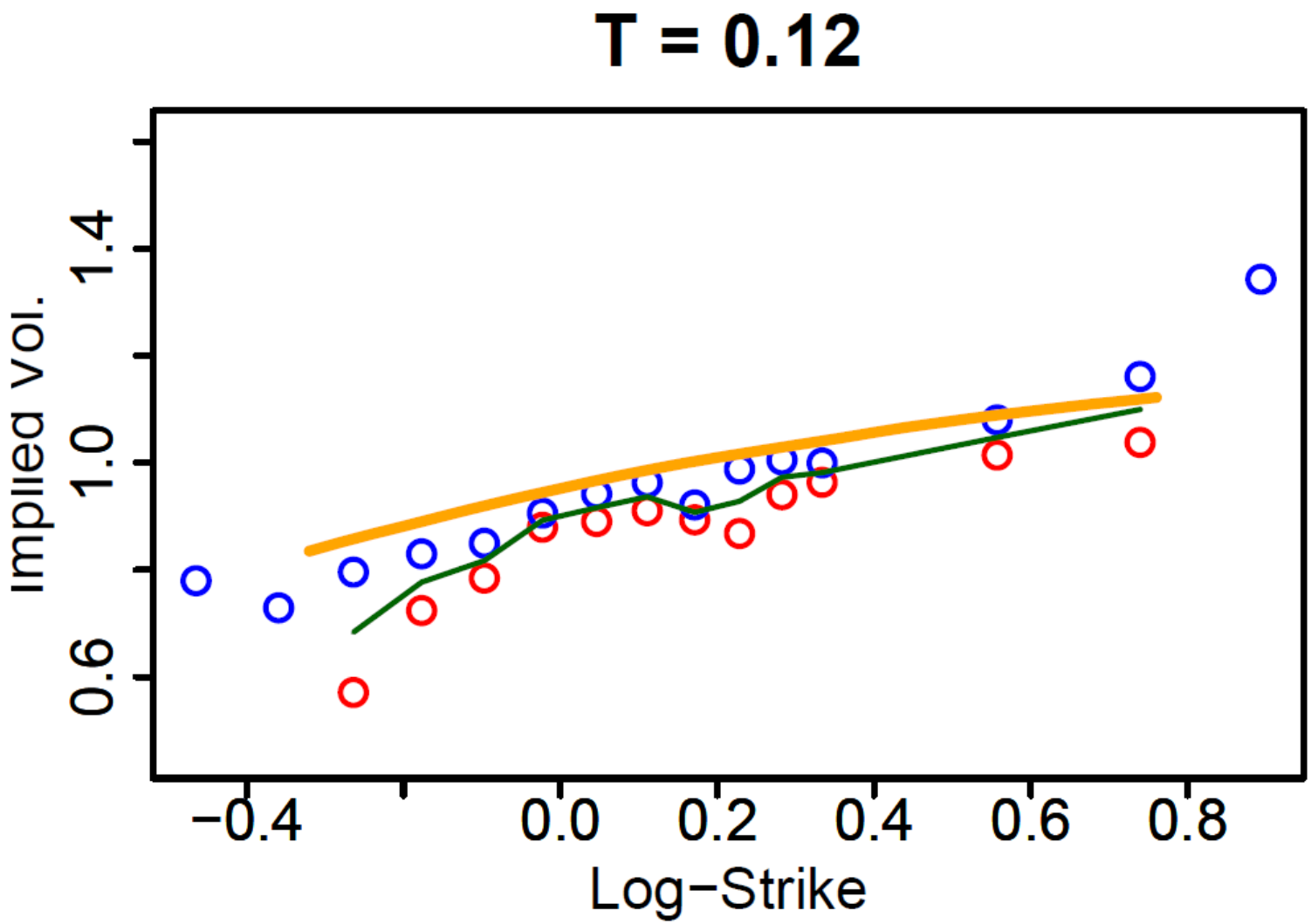
Double CEV fit to SPX options as of 03-Apr-2007

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):



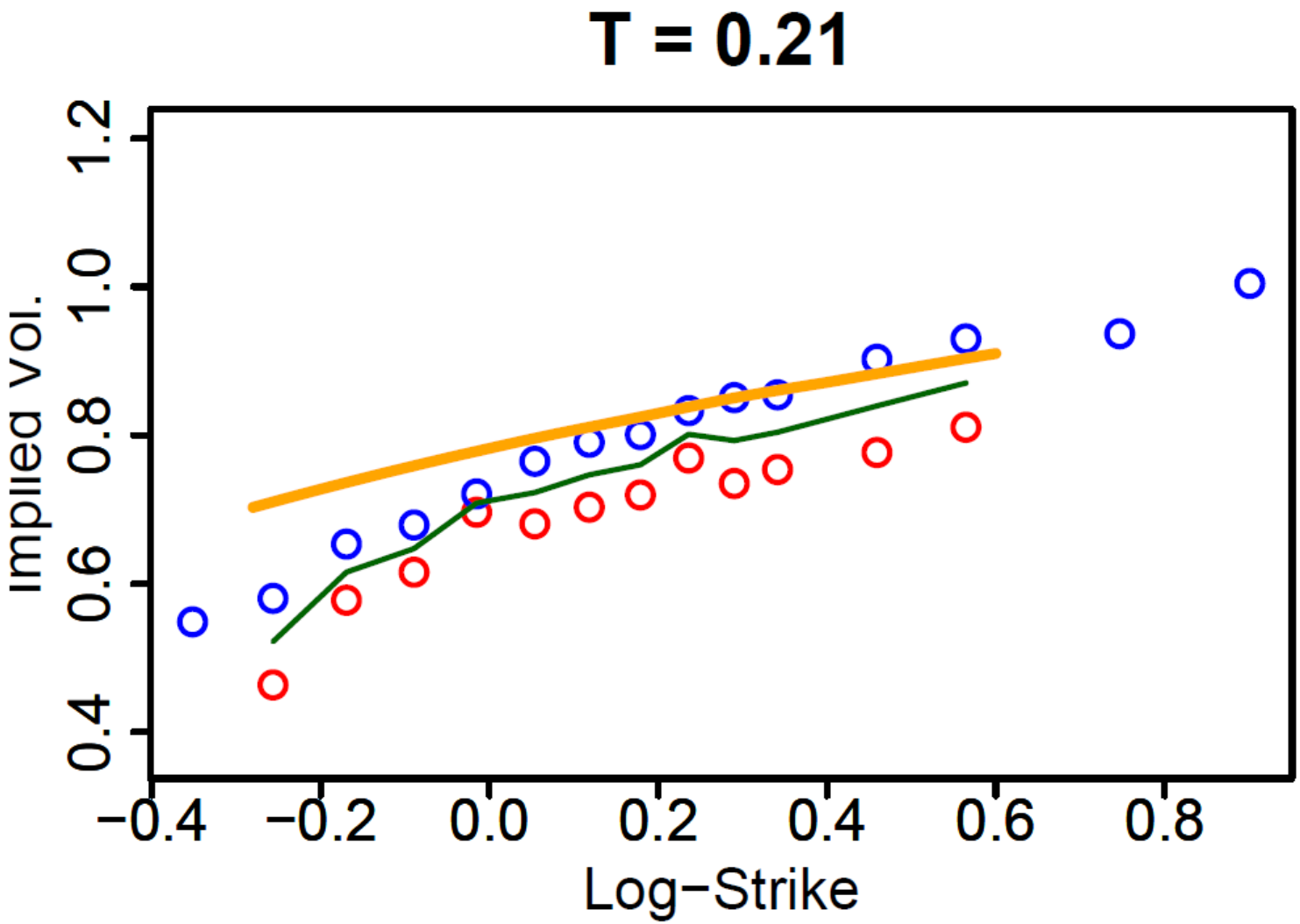
© 2020 Bloomberg Finance L.P. All rights reserved.

Fit to VIX options



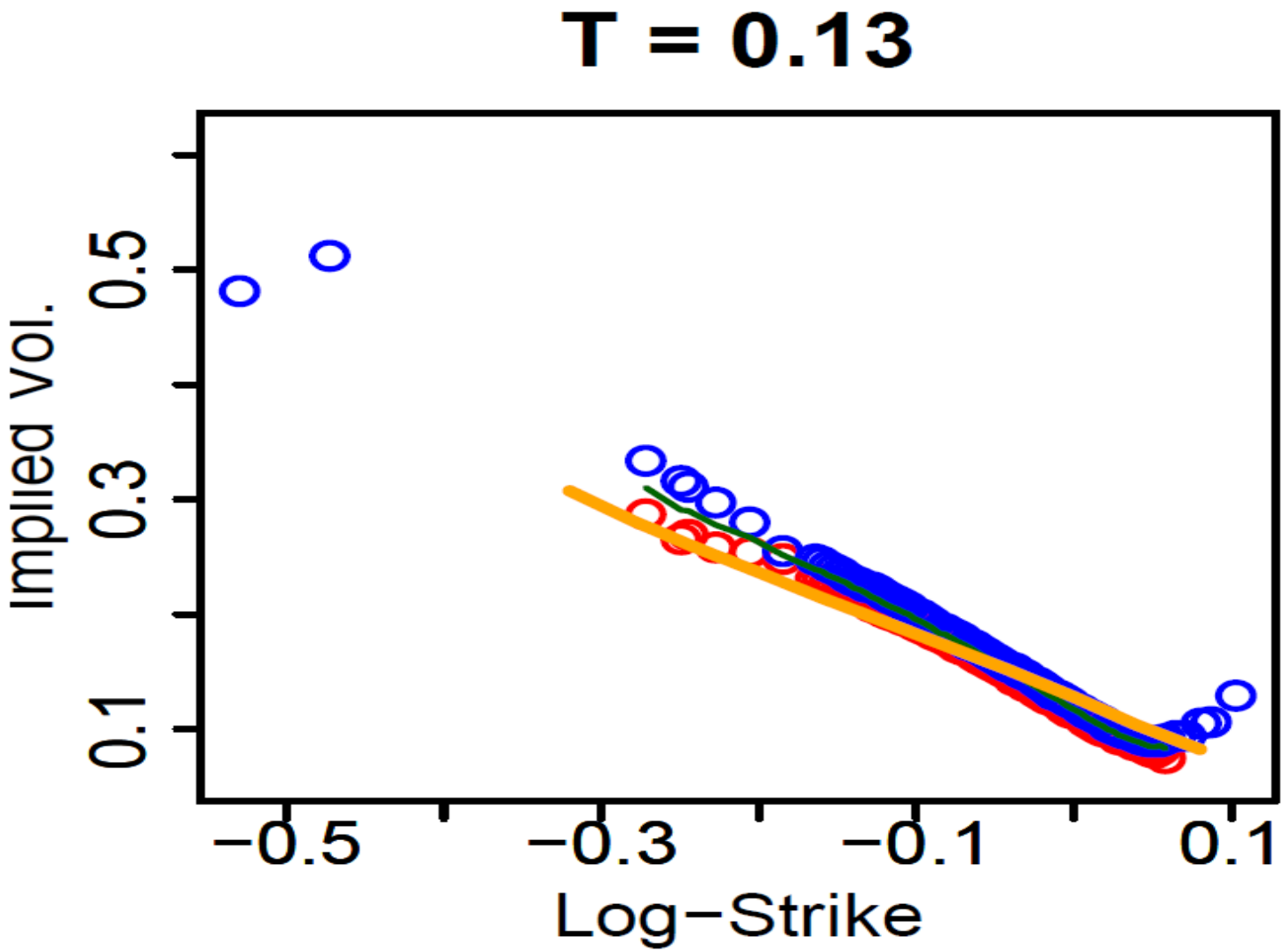
© 2020 Bloomberg Finance L.P. All rights reserved.

Fit to VIX options



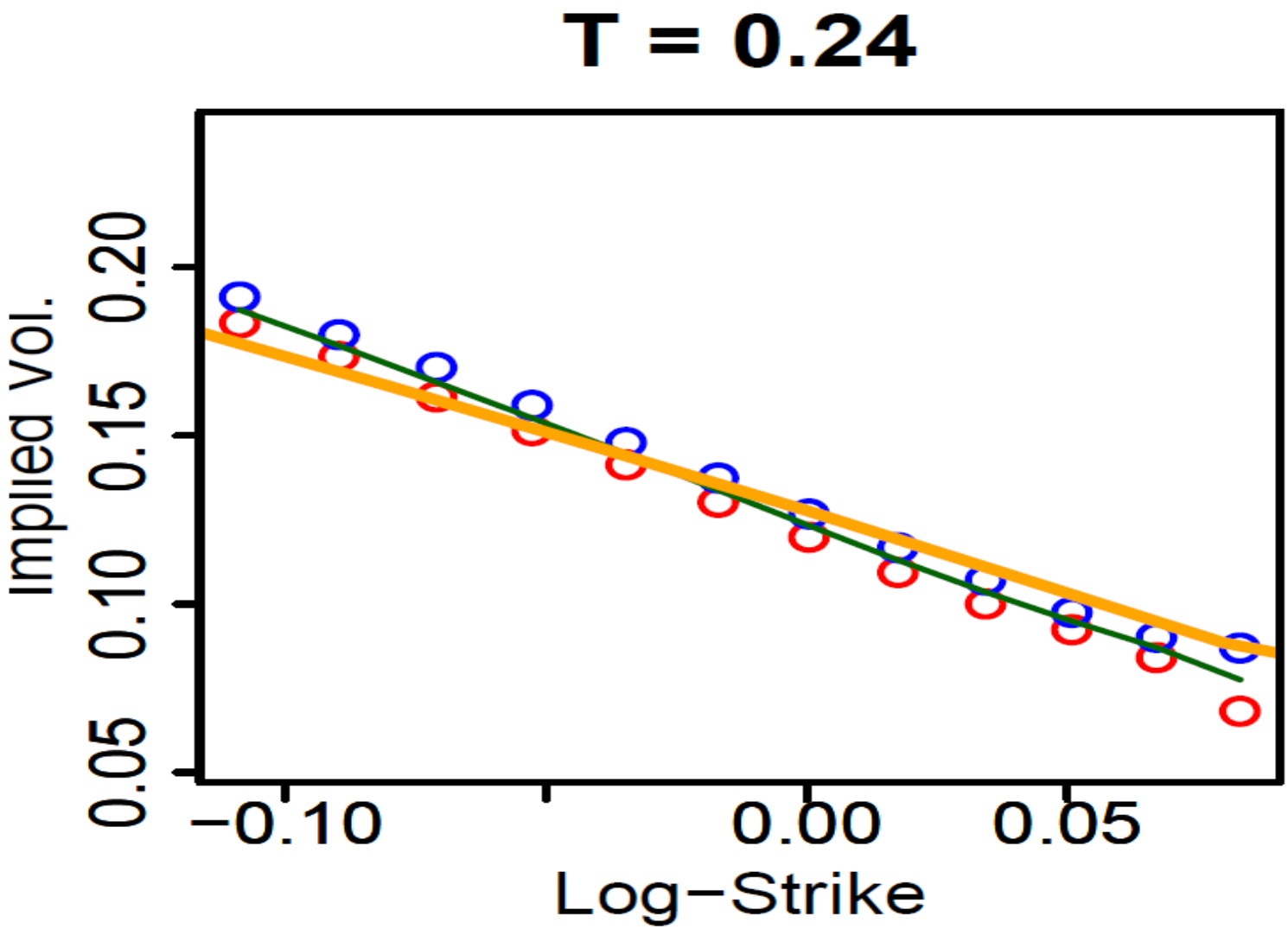
© 2020 Bloomberg Finance L.P. All rights reserved.

Fit to SPX options



© 2020 Bloomberg Finance L.P. All rights reserved.

Fit to SPX options



© 2020 Bloomberg Finance L.P. All rights reserved.

Skewed rough Bergomi model (G., 2018)

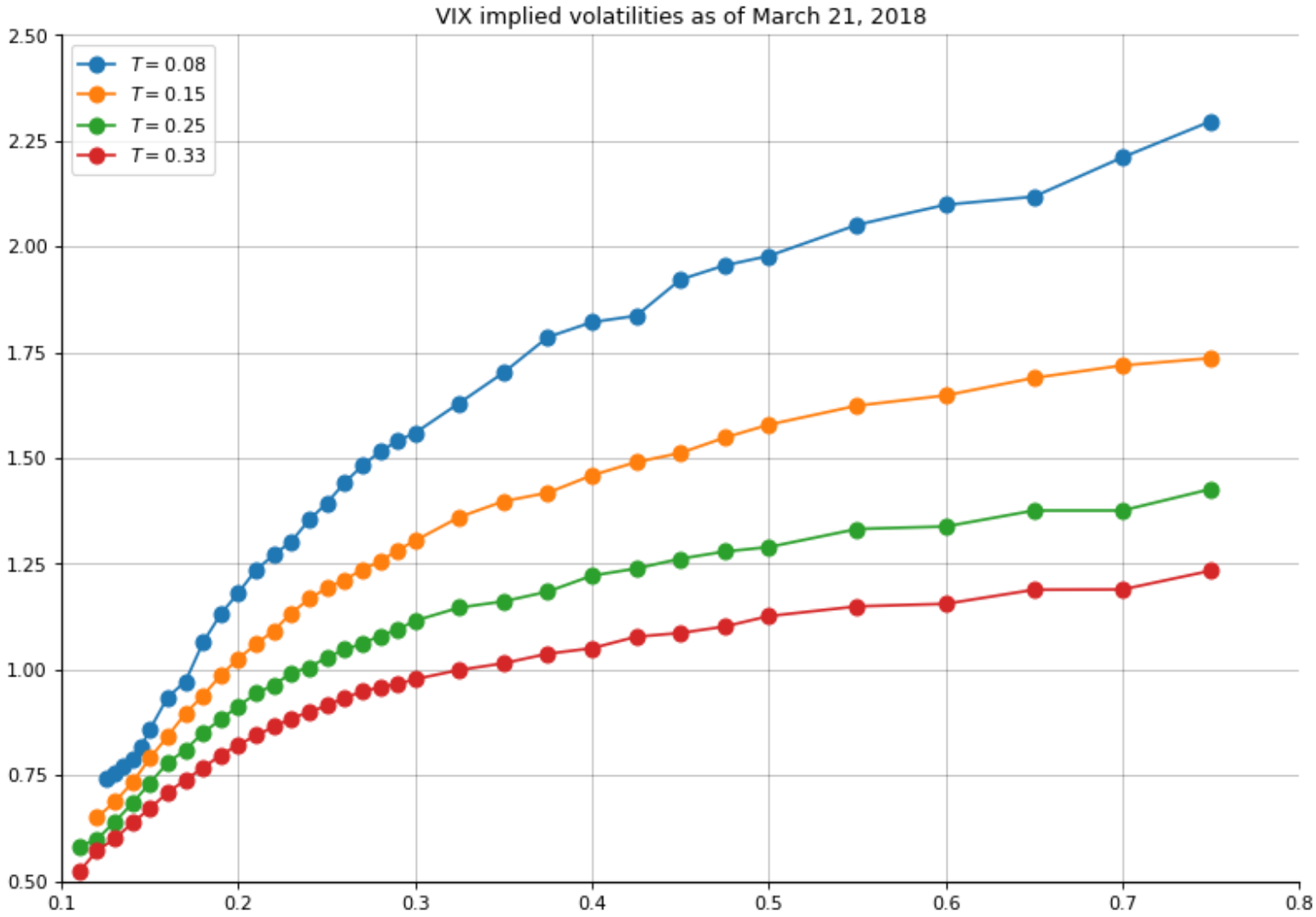
- Following Bergomi (2008), we suggested using a linear combination of two lognormal random variables to model the instantaneous variance σ_t^2 so as to generate positive VIX skew (G., 2018):

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

  with $\lambda \in [0,1]$.
- $\mathcal{E}(X)$ is simply a shorthand notation for $\exp\left(X - \frac{1}{2}\mathrm{Var}(X)\right)$.
- Also independently proposed by De Marco.

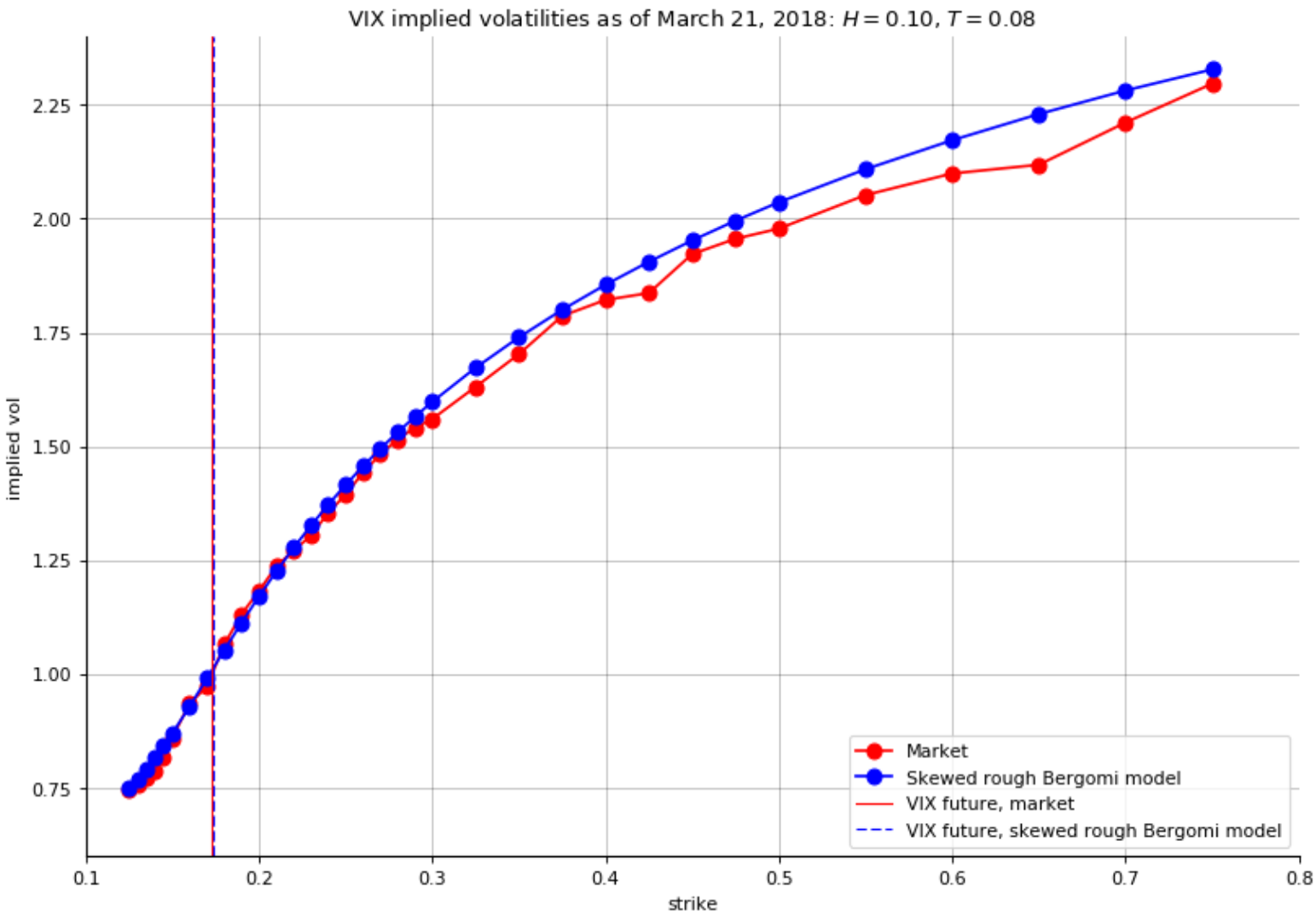
© 2020 Bloomberg Finance L.P. All rights reserved.

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



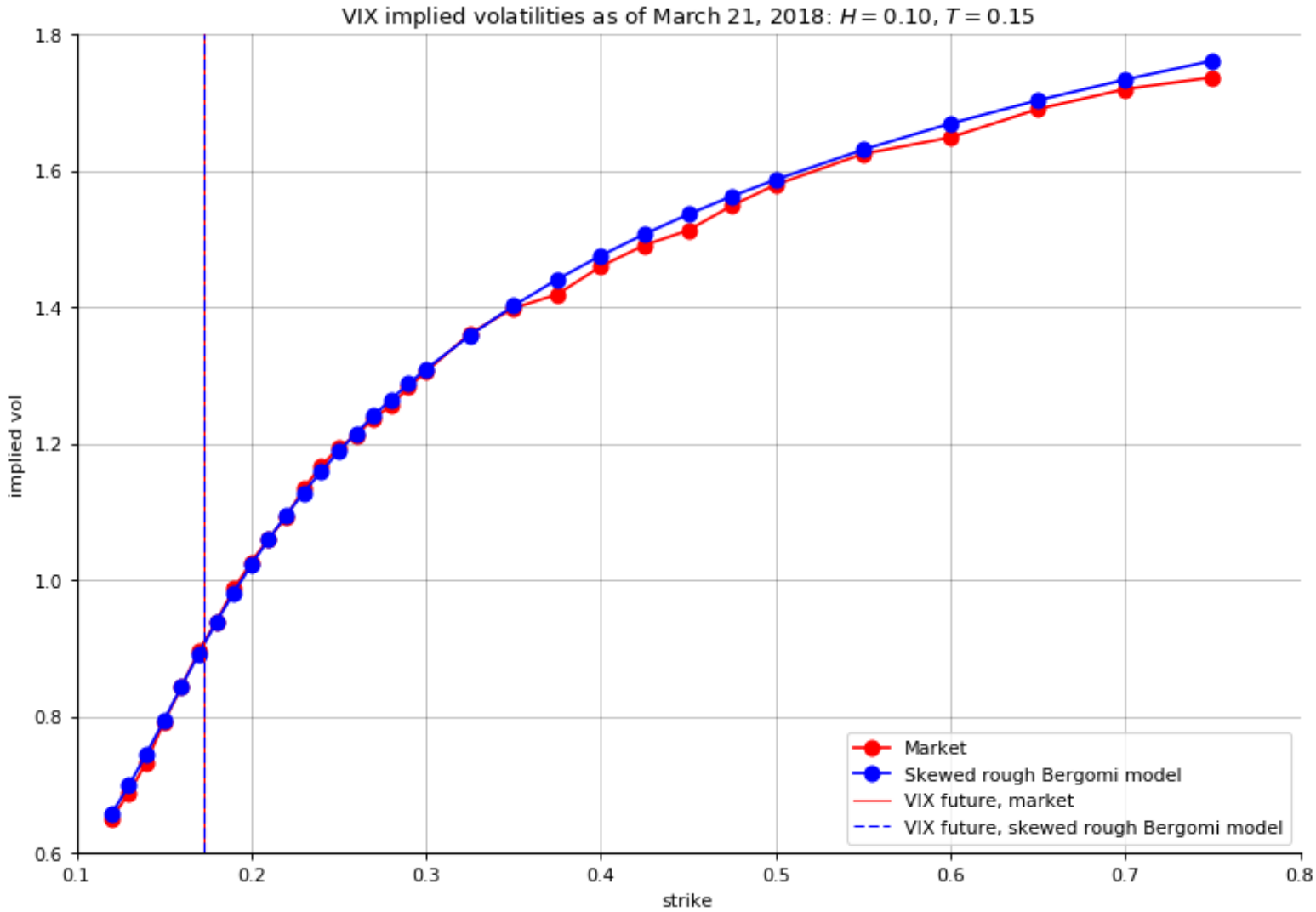
© 2020 Bloomberg Finance L.P. All rights reserved.

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



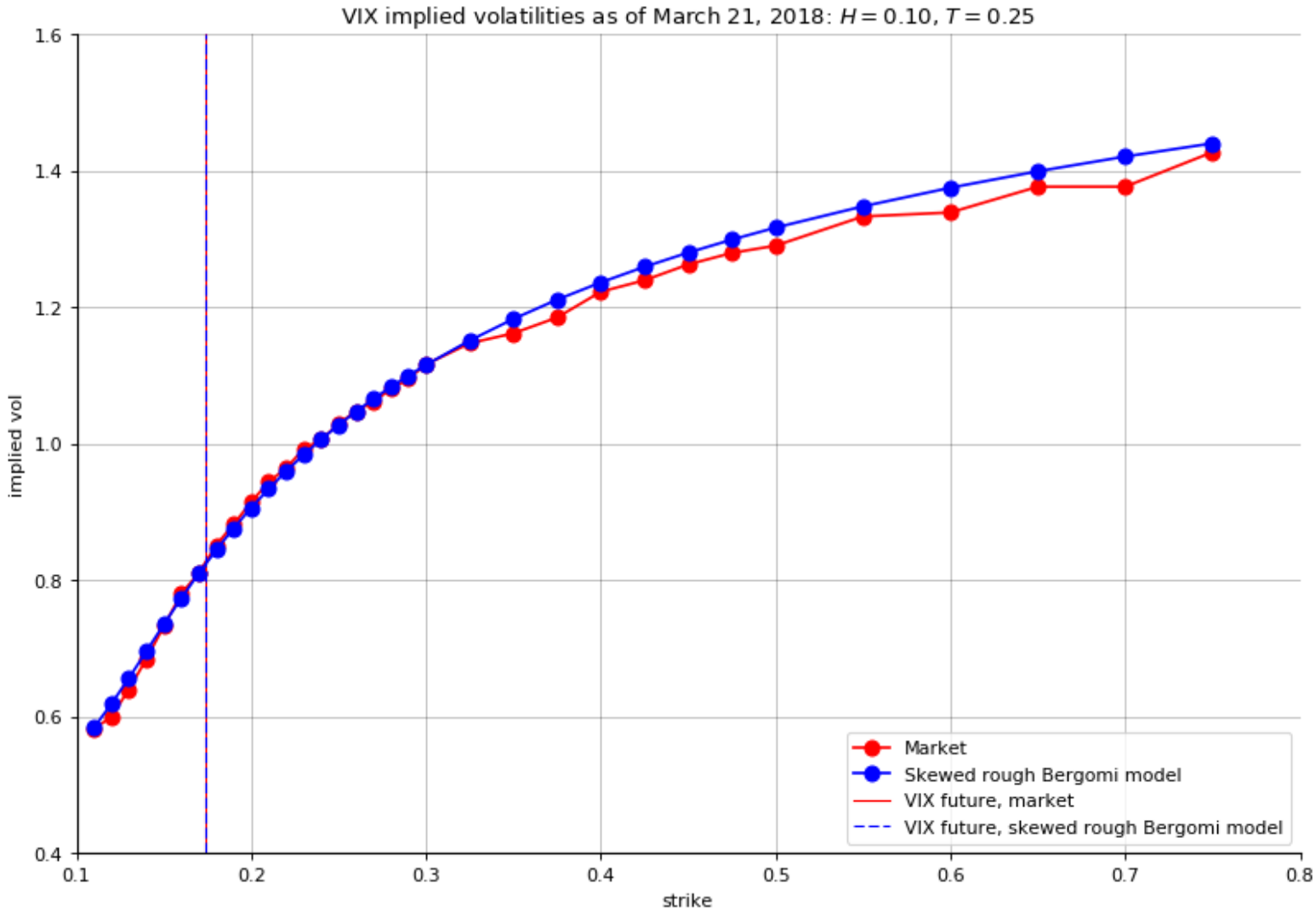
© 2020 Bloomberg Finance L.P. All rights reserved.

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



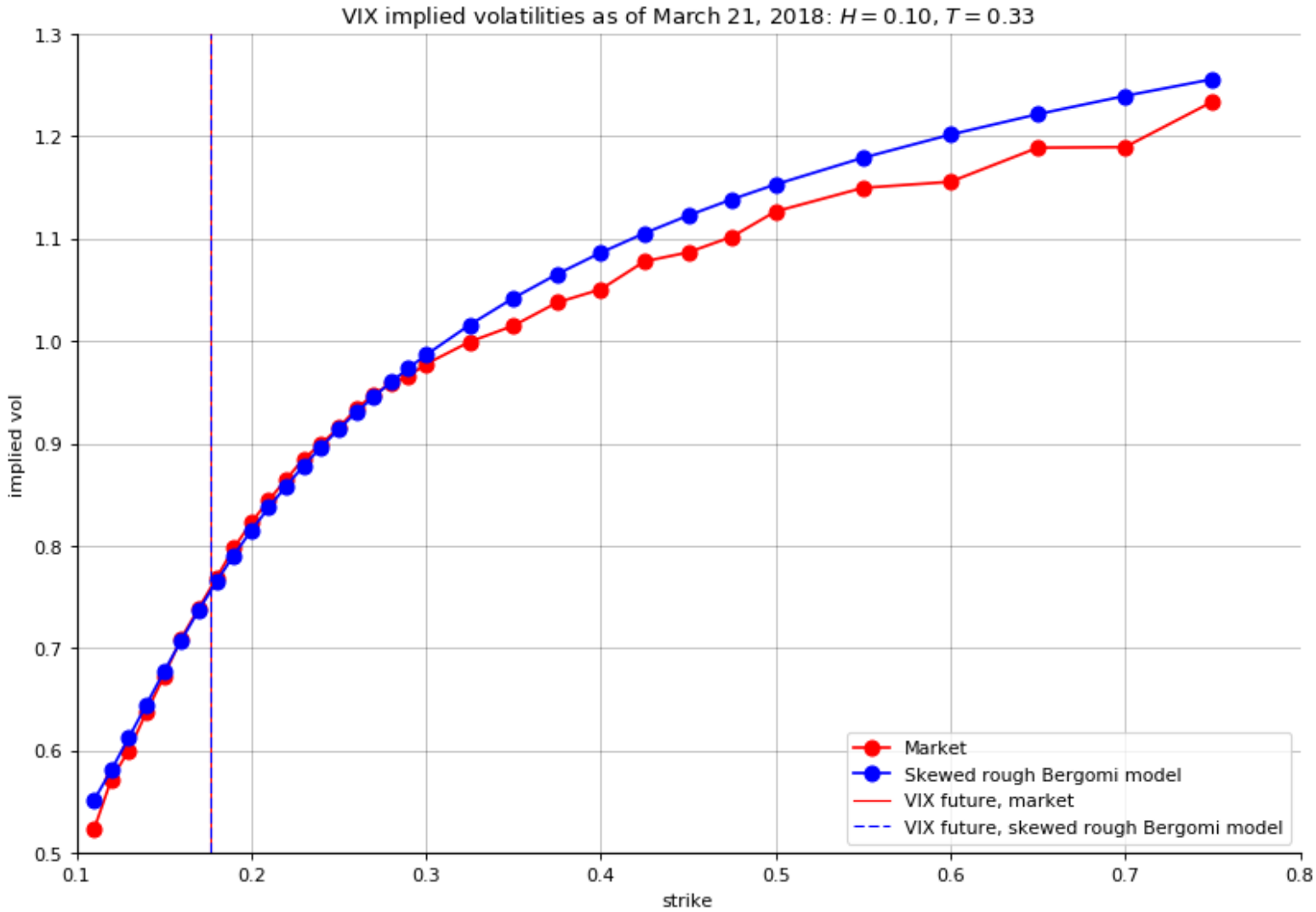
© 2020 Bloomberg Finance L.P. All rights reserved.

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



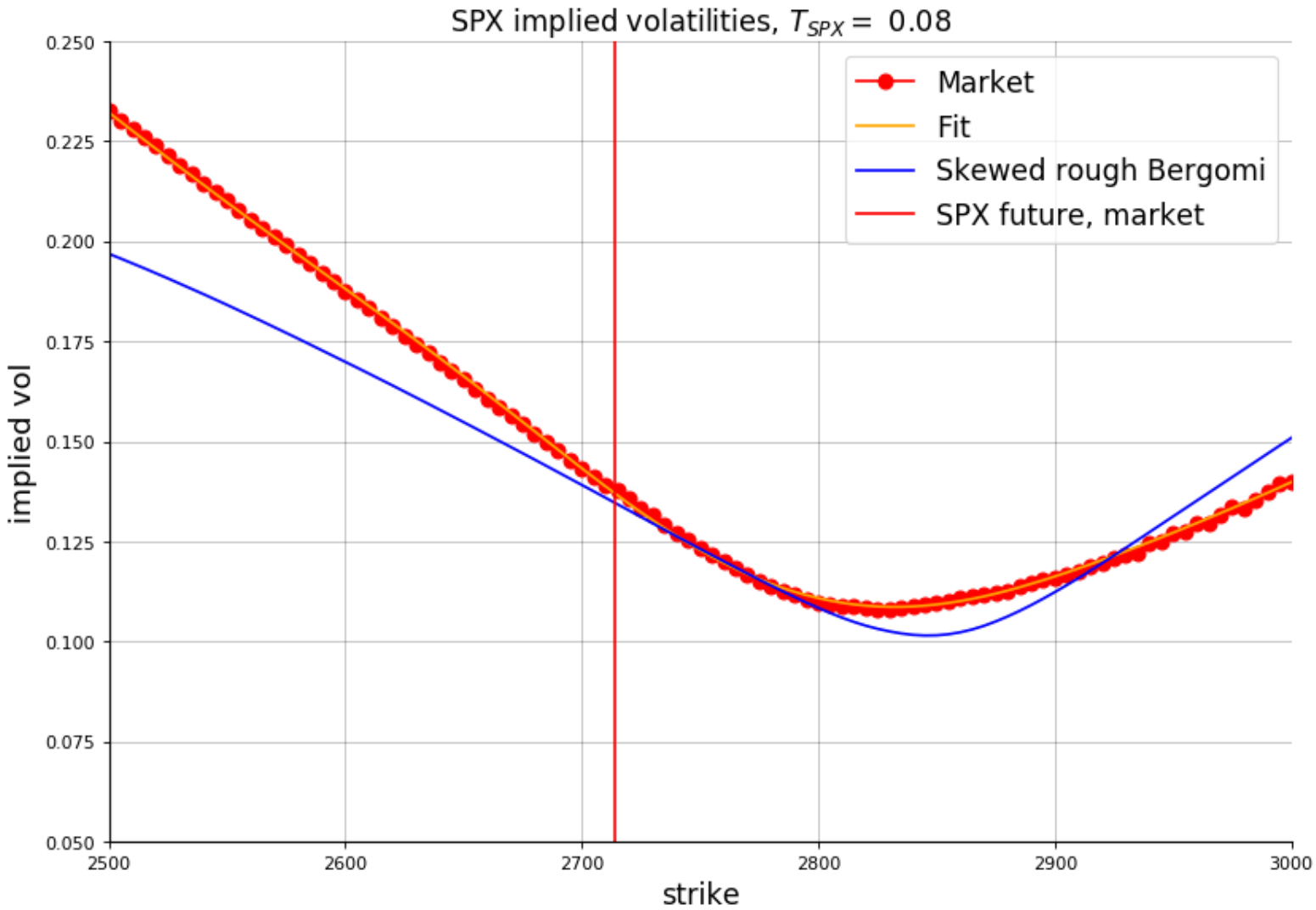
© 2020 Bloomberg Finance L.P. All rights reserved.

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



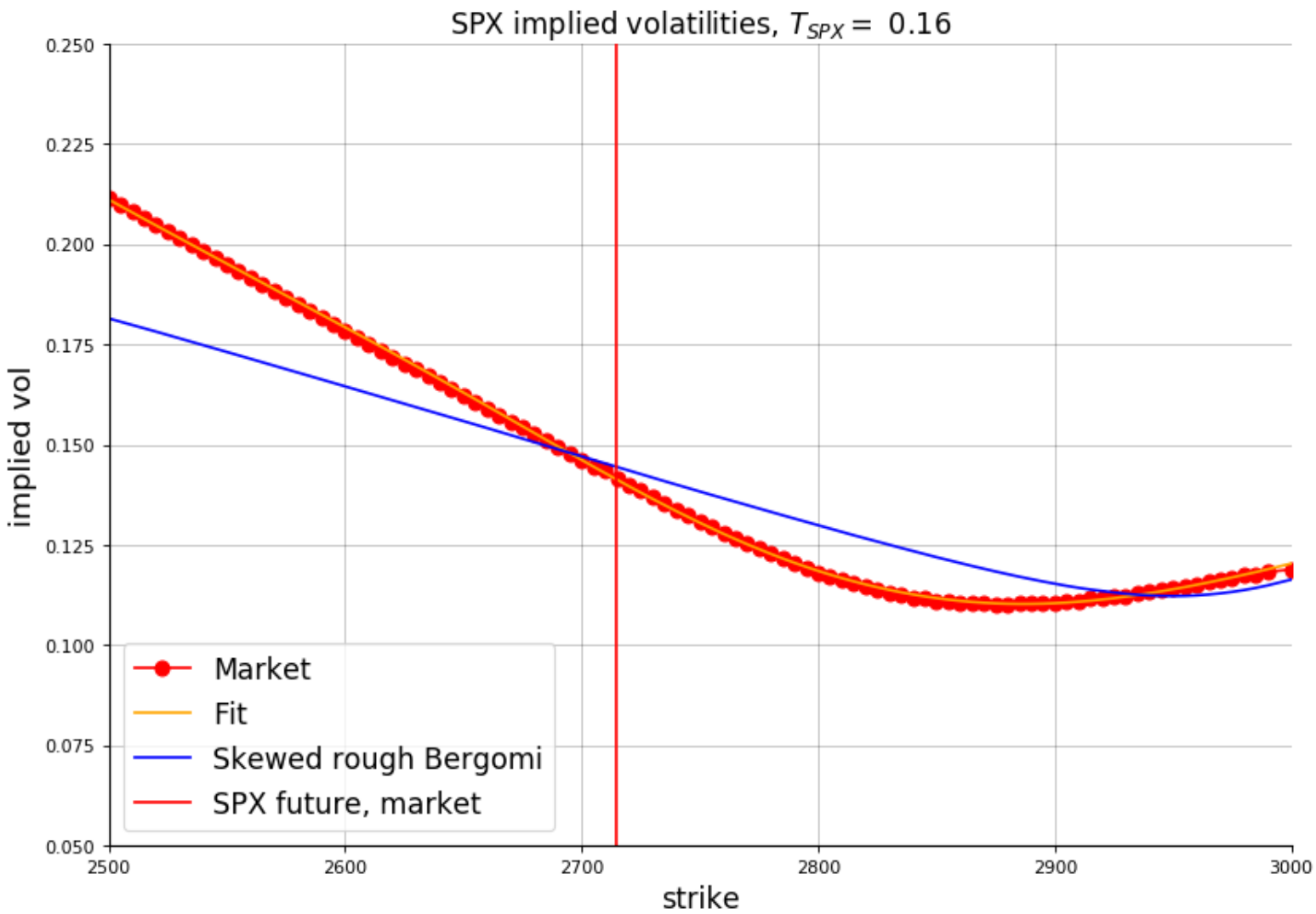
© 2020 Bloomberg Finance L.P. All rights reserved.

Skewed rough Bergomi calibrated to VIX: SPX smile (March 21, 2018)



© 2020 Bloomberg Finance L.P. All rights reserved.

Skewed rough Bergomi calibrated to VIX: SPX smile (March 21, 2018)



© 2020 Bloomberg Finance L.P. All rights reserved.

Skewed rough Bergomi calibrated to VIX: SPX smile

- **Not enough ATM skew for SPX**, despite pushing negative spot-vol correlation as much as possible.
- I get **similar results** when I use the **skewed 2-factor Bergomi model** instead of the skewed rough Bergomi model.

© 2020 Bloomberg Finance L.P. All rights reserved.

SLV calibrated to SPX: VIX smile (Aug 1, 2018)

- All continuous models on SPX that are calibrated to full SPX smile are of the form:

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

- They are stochastic local volatility (SLV) models

$$\frac{dS_t}{S_t} = a_t \ell(t,S_t)\,dW_t$$

with stochastic volatility (SV) (a_t) and leverage function

$$\ell(t,S_t) = \frac{\sigma_{\text{loc}}(t,S_t)}{\sqrt{\mathbb{E}[a_t^2|S_t]}}.$$

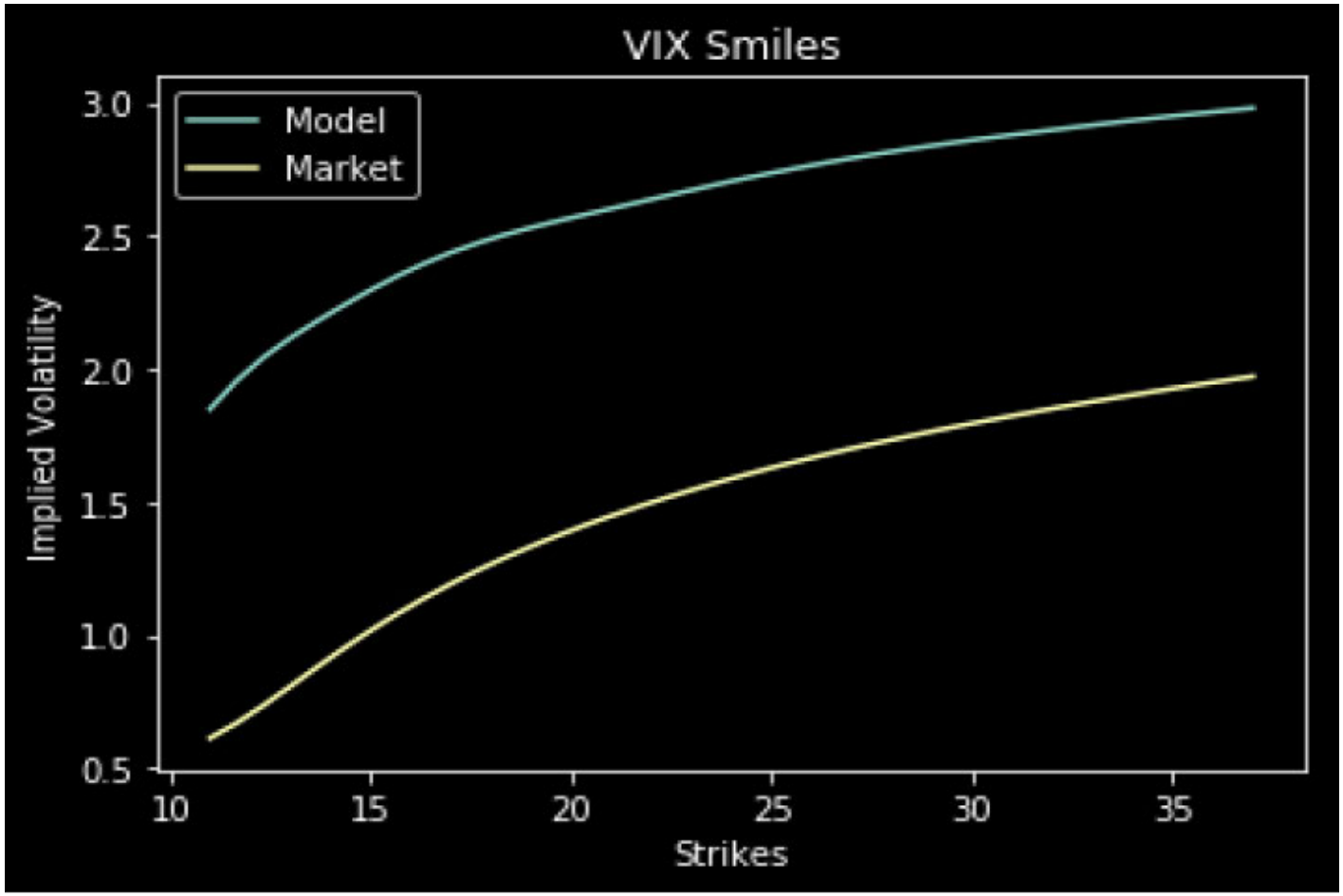
- In those models ($\tau := 30$ days)

$$\text{VIX}_T^2 \quad = \quad \frac{1}{\tau}\int_T^{T+\tau} \mathbb{E}\left[\frac{a_t^2}{\mathbb{E}[a_t^2|S_t]}\sigma_{\text{loc}}^2(t,S_t)\middle|\mathcal{F}_T\right]\,dt.$$

- Optimize SV parameters to fit VIX options.

© 2020 Bloomberg Finance L.P. All rights reserved.

SLV calibrated to SPX: VIX smile, $T = 21$ days (Aug 1, 2018)



SLV model, SV = skewed 2-factor Bergomi model
SV params optimized to fit VIX smile

© 2020 Bloomberg Finance L.P. All rights reserved.

Related works with continuous models on the SPX

- Fouque-Saporito (2018), Heston with stochastic vol-of-vol. Problem: their approach does not apply to short maturities (below 4 months), for which VIX derivatives are most liquid and the joint calibration is most difficult.
- Goutte-Ismail-Pham (2017), Heston with parameters driven by a Hidden Markov jump process.
- Gatheral-Jusselin-Rosenbaum (2020): quadratic rough Heston volatility model.
- Guo-Loeper-Obłłoj-Wang (2020): joint calibration via optimal transport—see later.
- Jacquier-Martini-Muguruza, *On the VIX futures in the rough Bergomi model* (2017):

  "Interestingly, we observe a 20% difference between the [vol-of-vol] parameter obtained through VIX calibration and the one obtained through SPX. This suggests that the volatility of volatility in the SPX market is 20% higher when compared to VIX, revealing potential data inconsistencies (arbitrage?)."

© 2020 Bloomberg Finance L.P. All rights reserved.

Motivation

- To try to jointly fit the SPX and VIX smiles, many authors have incorporated **jumps** in the dynamics of the SPX: Sepp, Cont-Kokholm, Papanicolaou-Sircar, Baldeaux-Badran, Pacati et al, Kokholm-Stisen, Bardgett et al...
- Jumps offer extra degrees of freedom to partly decouple the ATM SPX skew and the ATM VIX implied volatility.
- So far all the attempts at solving the joint SPX/VIX smile calibration problem only produced an **approximate fit**.

© 2020 Bloomberg Finance L.P. All rights reserved.

Our approach

- We solve this puzzle using a **completely different approach**: instead of postulating a parametric continuous-time (jump-)diffusion model on the SPX, we build a **nonparametric discrete-time model**:
  - Decouples SPX skew and VIX implied vol.
  - Perfectly fits the smiles.
- Given a VIX future maturity T_1, we build a **joint probability measure on** (S_1, V, S_2) which is **perfectly calibrated** to the SPX smiles at T_1 and $T_2 = T_1 + 30$ days, and the VIX future and VIX smile at T_1.
- S_1: SPX at T_1, V: VIX at T_1, S_2: SPX at T_2.
- Our model satisfies:
  - **Martingality constraint** on the SPX;
  - **Consistency condition**: the VIX at T_1 is the implied volatility of the 30-day log-contract on the SPX.
- Our model is cast as the solution of a **dispersion-constrained martingale transport problem** which is solved using the **Sinkhorn algorithm**, in the spirit of De March and Henry-Labordère (2019).

© 2020 Bloomberg Finance L.P. All rights reserved.

Setting and notation

- For simplicity: zero interest rates, repos, and dividends.
- μ_1 = risk-neutral distribution of $S_1 \longleftrightarrow$ market smile of SPX at T_1.
- μ_V = risk-neutral distribution of $V \longleftrightarrow$ market smile of VIX at T_1.
- μ_2 = risk-neutral distribution of $S_2 \longleftrightarrow$ market smile of SPX at T_2.
- F_V: value at time 0 of VIX future maturing at T_1.
- We denote $\mathbb{E}^i := \mathbb{E}^{\mu_i}$, $\mathbb{E}^V := \mathbb{E}^{\mu_V}$ and assume

$$\mathbb{E}^i[S_i] = S_0, \quad \mathbb{E}^i[|\ln S_i|] < \infty, \quad i \in \{1,2\}; \qquad \mathbb{E}^V[V] = F_V, \quad \mathbb{E}^V[V^2] < \infty.$$

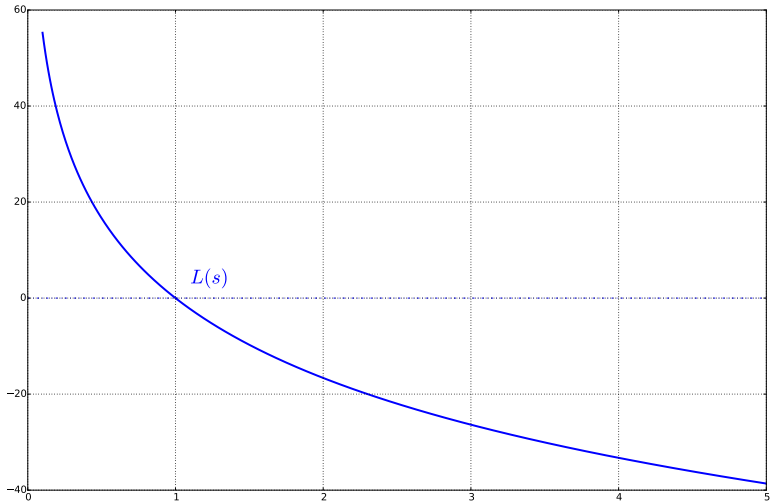
- No calendar arbitrage $\Longleftrightarrow \mu_1 \leq_c \mu_2$ (convex order)

© 2020 Bloomberg Finance L.P. All rights reserved.

Setting and notation

$$V^2 := (\mathrm{VIX}_{T_1})^2 := -\frac{2}{\tau}\mathrm{Price}_{T_1}\left[\ln\left(\frac{S_2}{S_1}\right)\right] = \mathrm{Price}_{T_1}\left[L\left(\frac{S_2}{S_1}\right)\right]$$

- $\tau := 30$ days.
- $L(x) := -\frac{2}{\tau}\ln x$: convex, decreasing.

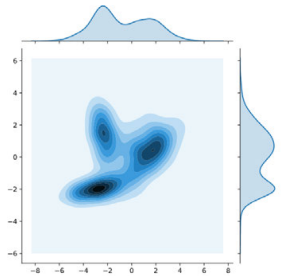


© 2020 Bloomberg Finance L.P. All rights reserved.

Superreplication, duality

© 2020 Bloomberg Finance L.P. All rights reserved.

Superreplication of forward-starting options



- The knowledge of μ_1 and μ_2 gives little information on the prices $\mathbb{E}^{\mu}[g(S_1, S_2)]$, e.g., prices of forward-starting options $\mathbb{E}^{\mu}[f(S_2/S_1)]$.
- Computing upper and lower bounds of these prices:
  Optimal transport (Monge, 1781; Kantorovich)
- Adding the no-arbitrage constraint that (S_1, S_2) is a martingale leads to more precise bounds, as this provides information on the conditional average of S_2/S_1 given S_1:
  Martingale optimal transport (Henry-Labordère, 2017)
- When $S =$ SPX: Adding VIX market data information produces even more precise bounds, as it gives information on the conditional dispersion of S_2/S_1, which is controlled by the VIX V:
  Dispersion-constrained martingale optimal transport

© 2020 Bloomberg Finance L.P. All rights reserved.

Classical optimal transport

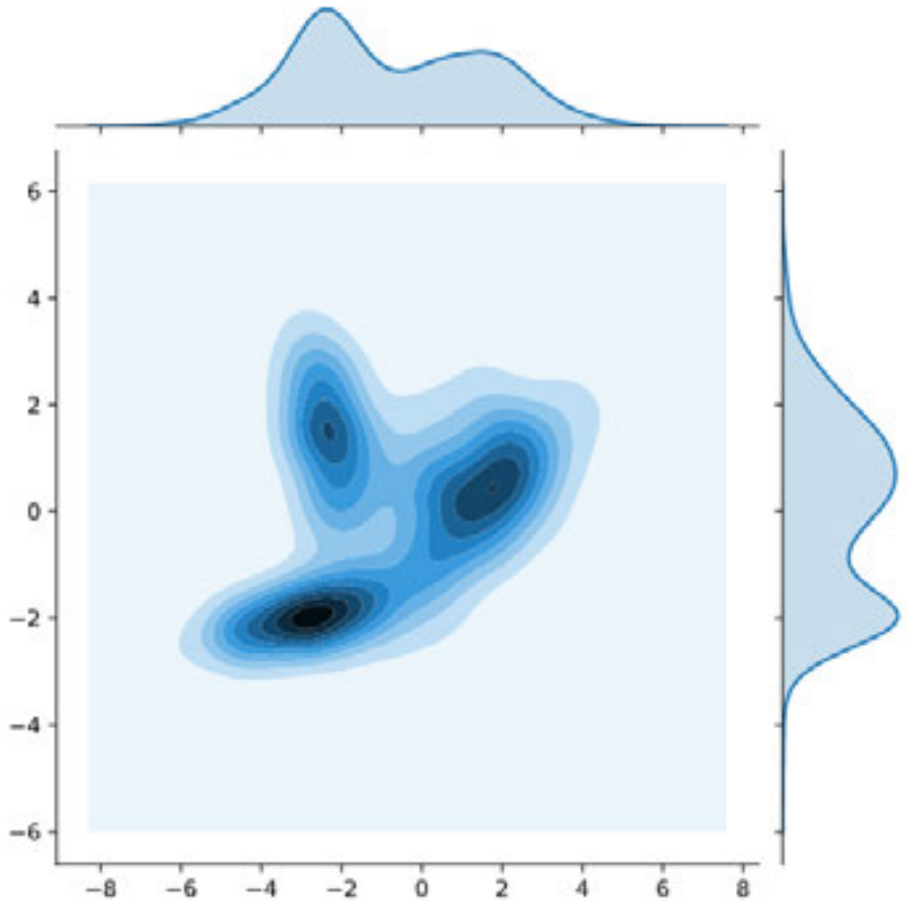


Figure: Example of a transport plan. Source: Wikipedia

© 2020 Bloomberg Finance L.P. All rights reserved.

Superreplication: primal problem

Fundamental principle: Upper bound for the price of payoff $f(S_1, V, S_2)$ = smallest price at time 0 of a superreplicating portfolio.

Following De Marco-Henry-Labordère (2015), G.-Menegaux-Nutz (2017), the available instruments for superreplication are:

- At time 0:
  - $u_1(S_1)$: SPX vanilla payoff maturity T_1 (including cash)
  - $u_2(S_2)$: SPX vanilla payoff maturity T_2
  - $u_V(V)$: **VIX vanilla payoff maturity T_1**

  Cost: $\text{MktPrice}[u_1(S_1)] + \text{MktPrice}[u_2(S_2)] + \text{MktPrice}[u_V(V)]$
  $= \quad \mathbb{E}^1[u_1(S_1)] \quad + \quad \mathbb{E}^2[u_2(S_2)] \quad + \quad \mathbb{E}^V[u_V(V)]$
- At time T_1:
  - $\Delta_S(S_1, V)(S_2 - S_1)$: delta hedge
  - $\Delta_L(S_1, V)(L(S_2/S_1) - V^2)$: buy $\Delta_L(S_1, V)$ log-contracts

  Cost: 0

Shorthand notation:

$$\Delta^{(S)}(s_1, v, s_2) := \Delta(s_1, v)(s_2 - s_1), \quad \Delta^{(L)}(s_1, v, s_2) := \Delta(s_1, v)\left(L\left(\frac{s_2}{s_1}\right) - v^2\right)$$

© 2020 Bloomberg Finance L.P. All rights reserved.

Superreplication: primal problem

- The model-independent no-arbitrage upper bound for the derivative with payoff $f(S_1, V, S_2)$ is the smallest price at time 0 of a superreplicating portfolio:

$$P_f := \inf_{\mathcal{U}_f} \left\{ \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \right\}.$$

- $\mathcal{U}_f$: set of superreplicating portfolios, i.e., the set of all functions $(u_1, u_V, u_2, \Delta_S, \Delta_L)$ that satisfy the superreplication constraint:

$$u_1(s_1) + u_V(v) + u_2(s_2) + \Delta_S^{(S)}(s_1, v, s_2) + \Delta_L^{(L)}(s_1, v, s_2) \geq f(s_1, v, s_2).$$

- Linear program.

© 2020 Bloomberg Finance L.P. All rights reserved.

Superreplication: dual problem

- $\mathcal{P}(\mu_1, \mu_V, \mu_2)$: set of all the probability measures μ on $\mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0}$ such that

$$S_1 \sim \mu_1, \quad V \sim \mu_V, \quad S_2 \sim \mu_2, \quad \mathbb{E}^\mu [S_2 | S_1, V] = S_1, \quad \mathbb{E}^\mu \left[L\left(\frac{S_2}{S_1}\right) \middle| S_1, V \right] = V^2.$$

- Dual problem:

$$D_f := \sup_{\mu \in \mathcal{P}(\mu_1, \mu_V, \mu_2)} \mathbb{E}^\mu [f(S_1, V, S_2)].$$

- **Dispersion-constrained martingale optimal transport problem**.
- $\mathbb{E}^\mu[S_2|S_1, V] = S_1$: martingality condition of the SPX index, condition on the average of the distribution of S_2 given S_1 and V.
- $\mathbb{E}^\mu[L(S_2/S_1)|S_1, V] = V^2$: consistency condition, condition on dispersion around the average.

© 2020 Bloomberg Finance L.P. All rights reserved.

Superreplication: absence of a duality gap

Theorem (G., 2020)

Let $f : \mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0} \to \mathbb{R}$ *be upper semicontinuous and satisfy*

$$|f(s_1, v, s_2)| \leq C\big(1 + s_1 + s_2 + |L(s_1)| + |L(s_2)| + v^2\big)$$

for some constant $C > 0$. *Then*

$$P_f := \inf_{\mathcal{U}_f} \Big\{ \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \Big\}$$
$$= \sup_{\mu \in \mathcal{P}(\mu_1, \mu_V, \mu_2)} \mathbb{E}^\mu[f(S_1, V, S_2)] =: D_f.$$

Moreover, $D_f \neq -\infty$ *if and only if* $\mathcal{P}(\mu_1, \mu_V, \mu_2) \neq \emptyset$, *and in that case the supremum is attained.*

© 2020 Bloomberg Finance L.P. All rights reserved.

Superreplication of forward-starting options

- The knowledge of μ_1 and μ_2 gives little information on the prices $\mathbb{E}^{\mu}[g(S_1, S_2)]$, e.g., prices of forward-starting options $\mathbb{E}^{\mu}[f(S_2/S_1)]$.
- Computing upper and lower bounds of these prices:
  Optimal transport (Monge, 1781; Kantorovich)
- Adding the no-arbitrage constraint that (S_1, S_2) is a martingale leads to more precise bounds, as this provides information on the conditional average of S_2/S_1 given S_1:
  Martingale optimal transport (Henry-Labordère, 2017)
- When $S =$ SPX: Adding VIX market data information produces even more precise bounds, as it information on the conditional dispersion of S_2/S_1, which is controlled by the VIX V:
  Dispersion-constrained martingale optimal transport
- **Adding VIX market data may possibly reveal a joint SPX/VIX arbitrage. Corresponds to $\mathcal{P}(\mu_1, \mu_V, \mu_2) = \emptyset$** (see next slides).
- In the limiting case where $\mathcal{P}(\mu_1, \mu_V, \mu_2) = \{\mu_0\}$ is a singleton, the joint SPX/VIX market data information completely specifies the joint distribution of (S_1, S_2), hence the price of forward starting options.

© 2020 Bloomberg Finance L.P. All rights reserved.

Joint SPX/VIX arbitrage

© 2020 Bloomberg Finance L.P. All rights reserved.

Joint SPX/VIX arbitrage

- $\mathcal{U}_0$ = the portfolios $(u_1, u_2, u_V, \Delta^S, \Delta^L)$ superreplicating 0:

$$u_1(s_1)+u_2(s_2)+u_V(v)+\Delta^S(s_1,v)(s_2-s_1)+\Delta^L(s_1,v)\left(L\left(\frac{s_2}{s_1}\right)-v^2\right) \geq 0$$

- An (S_1, S_2, V)-arbitrage is an element of $\mathcal{U}_0$ with negative price:

$$\text{MktPrice}[u_1(S_1)] + \text{MktPrice}[u_2(S_2)] + \text{MktPrice}[u_V(V)] < 0$$

- Equivalently, there is an (S_1, S_2, V)-arbitrage if and only if

$$\inf_{\mathcal{U}_0} \{\text{MktPrice}[u_1(S_1)] + \text{MktPrice}[u_2(S_2)] + \text{MktPrice}[u_V(V)]\} = -\infty$$

© 2020 Bloomberg Finance L.P. All rights reserved.

Consistent extrapolation of SPX and VIX smiles

- If $\mathbb{E}^V[V^2] \neq \mathbb{E}^2[L(S_2)] - \mathbb{E}^1[L(S_1)]$, there is a trivial (S_1, S_2, V)-arbitrage. For instance, if $\mathbb{E}^V[V^2] < \mathbb{E}^2[L(S_2)] - \mathbb{E}^1[L(S_1)]$, pick

$$u_1(s_1) = L(s_1), \quad u_2(s_2) = -L(s_2), \quad u_V(v) = v^2, \quad \Delta_S(s_1, v) = 0, \quad \Delta_L(s_1, v) = 1.$$

- $\Longrightarrow$ We assume that

$$\mathbb{E}^V[V^2] = \mathbb{E}^2[L(S_2)] - \mathbb{E}^1[L(S_1)]. \tag{3.1}$$

- Violations of (3.1) in the market have been reported, suggesting arbitrage opportunities, see, e.g., Section 7.7.4 in Bergomi (2016).
- However, the quantities in (3.1) do not purely depend on market data. They depend on smile extrapolations.
- The reported violations of (3.1) actually rely on some arbitrary smile extrapolations.
- G. (2018) explains how to build **consistent extrapolations of the VIX and SPX smiles** so that (3.1) holds.

© 2020 Bloomberg Finance L.P. All rights reserved.

Joint SPX/VIX arbitrage

Theorem (G., 2020)

The following assertions are equivalent:

(i) *The market is free of* (S_1, S_2, V)*-arbitrage,*

(ii) $\mathcal{P}(\mu_1, \mu_V, \mu_2) \neq \emptyset$,

(iii) *There exists a coupling* ν *of* μ_1 *and* μ_V *such that* $\mathrm{Law}_\nu(S_1, L(S_1) + V^2)$ *and* $\mathrm{Law}_{\mu_2}(S_2, L(S_2))$ *are in convex order, i.e., for any convex function* $f : \mathbb{R}_{>0} \times \mathbb{R} \to \mathbb{R}$,

$$\mathbb{E}^\nu[f(S_1, L(S_1) + V^2)] \leq \mathbb{E}^2[f(S_2, L(S_2))].$$

© 2020 Bloomberg Finance L.P. All rights reserved.

Joint SPX/VIX arbitrage

(i) The market is free of (S_1, S_2, V)-arbitrage,

(ii) $\mathcal{P}(\mu_1, \mu_V, \mu_2) \neq \emptyset$,

(iii) There exists a coupling ν of μ_1 and μ_V such that $\mathrm{Law}_\nu(S_1, L(S_1) + V^2)$ and $\mathrm{Law}_{\mu_2}(S_2, L(S_2))$ are in convex order.

- Directly solving the linear problem associated to (i) is not easy as one needs to try all possible $(u_1, u_V, u_2, \Delta_S, \Delta_V)$ and check the superreplication constraints for all $s_1, s_2 > 0$ and $v \geq 0$.
- Checking (iii) numerically is difficult as, in dimension two, the extreme rays of the convex cone of convex functions are dense in the cone (Johansen 1974), contrary to the case of dimension one where the extreme rays are the call and put payoffs (Blaschke-Pick 1916).
- Instead, we verify absence of (S_1, S_2, V)-arbitrage by building – numerically, but with high accuracy – an element of $\mathcal{P}(\mu_1, \mu_V, \mu_2)$, thus checking (ii).

© 2020 Bloomberg Finance L.P. All rights reserved.

Build a model in $\mathcal{P}(\mu_1, \mu_V, \mu_2)$

© 2020 Bloomberg Finance L.P. All rights reserved.

Build a model in $\mathcal{P}(\mu_1, \mu_V, \mu_2)$

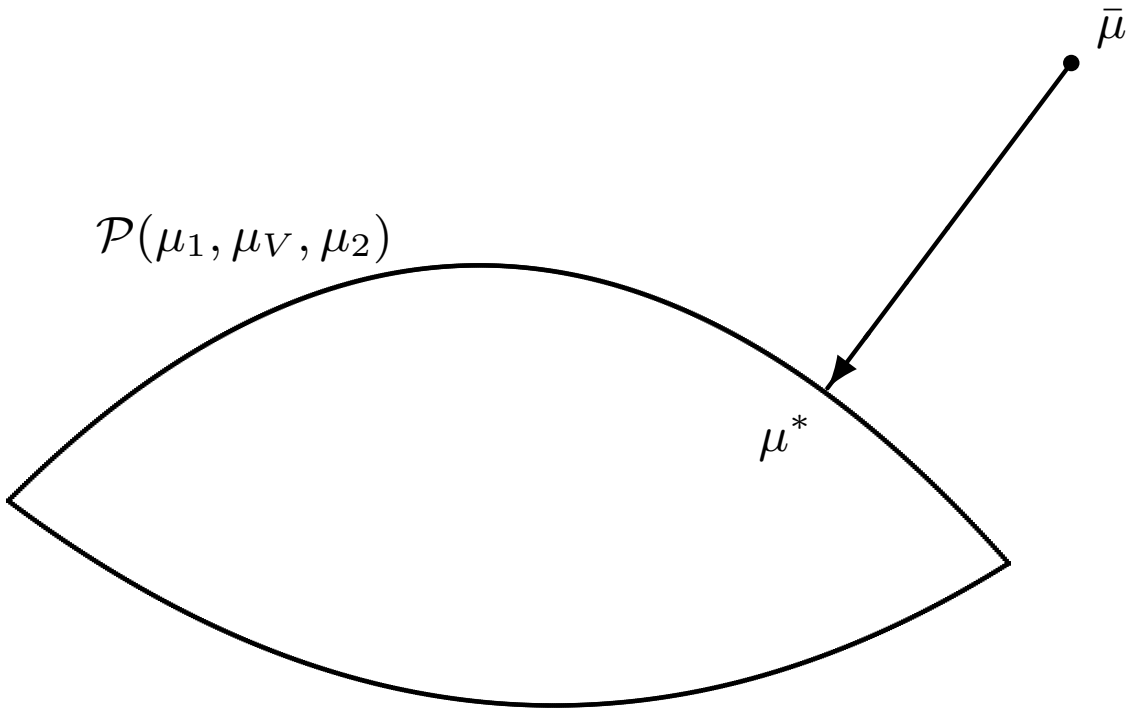
- We build a model $\mu \in \mathcal{P}(\mu_1, \mu_V, \mu_2)$.
- **We thus solve a longstanding puzzle in derivatives modeling: build an arbitrage-free model that jointly calibrates to the prices of SPX options, VIX futures and VIX options.**
- Our strategy is inspired by the recent work of De March and Henry-Labordère (2019).
- We assume that $\mathcal{P}(\mu_1, \mu_V, \mu_2) \neq \emptyset$ and try to build an element μ in this set. To this end, we fix a **reference probability measure** $\bar{\mu}$ on $\mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0}$ and look for the measure $\mu \in \mathcal{P}(\mu_1, \mu_V, \mu_2)$ that **minimizes the relative entropy** $H(\mu, \bar{\mu})$ of μ w.r.t. $\bar{\mu}$, also known as the Kullback-Leibler divergence:

$$D_{\bar{\mu}} := \inf_{\mu \in \mathcal{P}(\mu_1, \mu_V, \mu_2)} H(\mu, \bar{\mu}), \quad H(\mu, \bar{\mu}) := \begin{cases} \mathbb{E}^{\mu}\left[\ln \frac{d\mu}{d\bar{\mu}}\right] = \mathbb{E}^{\bar{\mu}}\left[\frac{d\mu}{d\bar{\mu}} \ln \frac{d\mu}{d\bar{\mu}}\right] & \text{if } \mu \ll \bar{\mu}, \\ +\infty & \text{otherwise.} \end{cases}$$

- This is a **strictly convex problem that can be solved after dualization using Sinkhorn's fixed point iteration** (Sinkhorn, 1967).

© 2020 Bloomberg Finance L.P. All rights reserved.

Build a model in $\mathcal{P}(\mu_1, \mu_V, \mu_2)$



© 2020 Bloomberg Finance L.P. All rights reserved.

Reminder on Lagrange multipliers

$$\begin{aligned}\inf_{g(x,y)=c} f(x,y) &= \inf_{x,y} \sup_{\lambda\in\mathbb{R}} \{f(x,y)-\lambda(g(x,y)-c)\} \\ &= \sup_{\lambda\in\mathbb{R}} \inf_{x,y} \{f(x,y)-\lambda(g(x,y)-c)\}\end{aligned}$$

- To compute the inner inf over x, y unconstrained, simply solve $\nabla f(x,y) = \lambda \nabla g(x,y)$: easy!
- Then maximize the result over λ unconstrained: easy!
- Constraint $g(x,y) = c \iff \frac{\partial}{\partial\lambda}\{f(x,y)-\lambda(g(x,y)-c)\} = 0$.

$$\begin{aligned}\inf_{\mu \text{ s.t. } S_1\sim\mu_1} H(\mu,\bar{\mu}) &= \inf_{\mu} \sup_{u_1(\cdot)} \left\{H(\mu,\bar{\mu}) + \mathbb{E}^1[u_1(S_1)] - \mathbb{E}^{\mu}[u_1(S_1)]\right\} \\ \inf_{\mu \text{ s.t. } \mathbb{E}^{\mu}[S_2|S_1,V]=S_1} H(\mu,\bar{\mu}) &= \inf_{\mu} \sup_{\Delta_S(\cdot,\cdot)} \left\{H(\mu,\bar{\mu}) - \mathbb{E}^{\mu}[\Delta_S(S_1,V)(S_2-S_1)]\right\}\end{aligned}$$

© 2020 Bloomberg Finance L.P. All rights reserved.

Reminder on Lagrange multipliers

$$\begin{aligned}\inf_{g(x,y)=c} f(x,y) &= \inf_{x,y} \sup_{\lambda\in\mathbb{R}} \{f(x,y)-\lambda(g(x,y)-c)\} \\ &= \sup_{\lambda\in\mathbb{R}} \inf_{x,y} \{f(x,y)-\lambda(g(x,y)-c)\}\end{aligned}$$

- To compute the inner inf over x, y unconstrained, simply solve $\nabla f(x,y) = \lambda \nabla g(x,y)$: easy!
- Then maximize the result over λ unconstrained: easy!
- Constraint $g(x,y) = c \iff \frac{\partial}{\partial\lambda}\{f(x,y)-\lambda(g(x,y)-c)\} = 0.$

$$\inf_{\mu \text{ s.t. } S_1\sim\mu_1} H(\mu,\bar{\mu}) = \inf_{\mu} \sup_{u_1(\cdot)} \left\{ H(\mu,\bar{\mu}) + \mathbb{E}^1[u_1(S_1)] - \mathbb{E}^{\mu}[u_1(S_1)] \right\}$$

$$\inf_{\mu \text{ s.t. } \mathbb{E}^{\mu}[S_2|S_1,V]=S_1} H(\mu,\bar{\mu}) = \inf_{\mu} \sup_{\Delta_S(\cdot,\cdot)} \left\{ H(\mu,\bar{\mu}) - \mathbb{E}^{\mu}[\Delta_S(S_1,V)(S_2-S_1)] \right\}$$

© 2020 Bloomberg Finance L.P. All rights reserved.

Build a model in $\mathcal{P}(\mu_1, \mu_V, \mu_2)$

- $\mathcal{M}_1$: set of probability measures on $\mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0}$: **unconstrained**
- $\mathcal{U}$: set of portfolios $u = (u_1, u_V, u_2, \Delta_S, \Delta_L)$: **Lagrange multipliers**

$$
\begin{aligned}
D_{\bar{\mu}} \quad &:= \quad \inf_{\mu \in \mathcal{P}(\mu_1, \mu_V, \mu_2)} H(\mu, \bar{\mu}) \\
&= \quad \inf_{\mu \in \mathcal{M}_1} \sup_{u \in \mathcal{U}} \Big\{ H(\mu, \bar{\mu}) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \\
&\qquad -\mathbb{E}^\mu \Big[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(S_1, V, S_2) + \Delta_L^{(L)}(S_1, V, S_2) \Big] \Big\} \\
&= \quad \sup_{u \in \mathcal{U}} \inf_{\mu \in \mathcal{M}_1} \Big\{ H(\mu, \bar{\mu}) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \\
&\qquad -\mathbb{E}^\mu \Big[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(S_1, V, S_2) + \Delta_L^{(L)}(S_1, V, S_2) \Big] \Big\}
\end{aligned}
$$

© 2020 Bloomberg Finance L.P. All rights reserved.

Build a model in $\mathcal{P}(\mu_1, \mu_V, \mu_2)$

- $\mathcal{M}_1$: set of probability measures on $\mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0}$: **unconstrained**
- $\mathcal{U}$: set of portfolios $u = (u_1, u_V, u_2, \Delta_S, \Delta_L)$: **Lagrange multipliers**

$$
\begin{aligned}
D_{\bar{\mu}} \quad &:= \quad \inf_{\mu \in \mathcal{P}(\mu_1, \mu_V, \mu_2)} H(\mu, \bar{\mu}) \\
&= \quad \inf_{\mu \in \mathcal{M}_1} \sup_{u \in \mathcal{U}} \Big\{ H(\mu, \bar{\mu}) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \\
&\qquad -\mathbb{E}^\mu \Big[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(S_1, V, S_2) + \Delta_L^{(L)}(S_1, V, S_2) \Big] \Big\} \\
&= \quad \sup_{u \in \mathcal{U}} \inf_{\mu \in \mathcal{M}_1} \Big\{ H(\mu, \bar{\mu}) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \\
&\qquad -\mathbb{E}^\mu \Big[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(S_1, V, S_2) + \Delta_L^{(L)}(S_1, V, S_2) \Big] \Big\}
\end{aligned}
$$

© 2020 Bloomberg Finance L.P. All rights reserved.

Build a model in $\mathcal{P}(\mu_1, \mu_V, \mu_2)$

- $\mathcal{M}_1$: set of probability measures on $\mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0}$: **unconstrained**
- $\mathcal{U}$: set of portfolios $u = (u_1, u_V, u_2, \Delta_S, \Delta_L)$: **Lagrange multipliers**

$$
\begin{aligned}
D_{\bar{\mu}} \quad &:= \quad \inf_{\mu \in \mathcal{P}(\mu_1, \mu_V, \mu_2)} H(\mu, \bar{\mu}) \\
&= \quad \inf_{\mu \in \mathcal{M}_1} \sup_{u \in \mathcal{U}} \Big\{ H(\mu, \bar{\mu}) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \\
&\qquad -\mathbb{E}^\mu \Big[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(S_1, V, S_2) + \Delta_L^{(L)}(S_1, V, S_2) \Big] \Big\} \\
&= \quad \sup_{u \in \mathcal{U}} \inf_{\mu \in \mathcal{M}_1} \Big\{ H(\mu, \bar{\mu}) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \\
&\qquad -\mathbb{E}^\mu \Big[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(S_1, V, S_2) + \Delta_L^{(L)}(S_1, V, S_2) \Big] \Big\}
\end{aligned}
$$

© 2020 Bloomberg Finance L.P. All rights reserved.

Build a model in $\mathcal{P}(\mu_1, \mu_V, \mu_2)$

- $\mathcal{M}_1$: set of probability measures on $\mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0}$: **unconstrained**
- $\mathcal{U}$: set of portfolios $u = (u_1, u_V, u_2, \Delta_S, \Delta_L)$: **Lagrange multipliers**

$$
\begin{aligned}
D_{\bar{\mu}} \quad &:= \quad \inf_{\mu \in \mathcal{P}(\mu_1, \mu_V, \mu_2)} H(\mu, \bar{\mu}) \\
&= \quad \inf_{\mu \in \mathcal{M}_1} \sup_{u \in \mathcal{U}} \Big\{ H(\mu, \bar{\mu}) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \\
&\qquad -\mathbb{E}^\mu \Big[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(S_1, V, S_2) + \Delta_L^{(L)}(S_1, V, S_2) \Big] \Big\} \\
&= \quad \sup_{u \in \mathcal{U}} \inf_{\mu \in \mathcal{M}_1} \Big\{ H(\mu, \bar{\mu}) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \\
&\qquad -\mathbb{E}^\mu \Big[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(S_1, V, S_2) + \Delta_L^{(L)}(S_1, V, S_2) \Big] \Big\}
\end{aligned}
$$

© 2020 Bloomberg Finance L.P. All rights reserved.

Build a model in $\mathcal{P}(\mu_1, \mu_V, \mu_2)$

- $\mathcal{M}_1$: set of probability measures on $\mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0}$: **unconstrained**
- $\mathcal{U}$: set of portfolios $u = (u_1, u_V, u_2, \Delta_S, \Delta_L)$: **Lagrange multipliers**

$$
\begin{aligned}
D_{\bar{\mu}} \quad &:= \quad \inf_{\mu \in \mathcal{P}(\mu_1, \mu_V, \mu_2)} H(\mu, \bar{\mu}) \\
&= \quad \inf_{\mu \in \mathcal{M}_1} \sup_{u \in \mathcal{U}} \Big\{ H(\mu, \bar{\mu}) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \\
&\qquad -\mathbb{E}^\mu \Big[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(S_1, V, S_2) + \Delta_L^{(L)}(S_1, V, S_2) \Big] \Big\} \\
&= \quad \sup_{u \in \mathcal{U}} \inf_{\mu \in \mathcal{M}_1} \Big\{ H(\mu, \bar{\mu}) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \\
&\qquad -\mathbb{E}^\mu \Big[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(S_1, V, S_2) + \Delta_L^{(L)}(S_1, V, S_2) \Big] \Big\}
\end{aligned}
$$

© 2020 Bloomberg Finance L.P. All rights reserved.

Build a model in $\mathcal{P}(\mu_1, \mu_V, \mu_2)$

- $\mathcal{M}_1$: set of probability measures on $\mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0}$: **unconstrained**
- $\mathcal{U}$: set of portfolios $u = (u_1, u_V, u_2, \Delta_S, \Delta_L)$: **Lagrange multipliers**

$$
\begin{aligned}
D_{\bar{\mu}} \quad &:= \quad \inf_{\mu \in \mathcal{P}(\mu_1, \mu_V, \mu_2)} H(\mu, \bar{\mu}) \\
&= \quad \inf_{\mu \in \mathcal{M}_1} \sup_{u \in \mathcal{U}} \Big\{ H(\mu, \bar{\mu}) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \\
&\qquad -\mathbb{E}^\mu \Big[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(S_1, V, S_2) + \Delta_L^{(L)}(S_1, V, S_2) \Big] \Big\} \\
&= \quad \sup_{u \in \mathcal{U}} \inf_{\mu \in \mathcal{M}_1} \Big\{ H(\mu, \bar{\mu}) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \\
&\qquad -\mathbb{E}^\mu \Big[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(S_1, V, S_2) + \Delta_L^{(L)}(S_1, V, S_2) \Big] \Big\}
\end{aligned}
$$

© 2020 Bloomberg Finance L.P. All rights reserved.

Build a model in $\mathcal{P}(\mu_1, \mu_V, \mu_2)$

- $\mathcal{M}_1$: set of probability measures on $\mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0}$: **unconstrained**
- $\mathcal{U}$: set of portfolios $u = (u_1, u_V, u_2, \Delta_S, \Delta_L)$: **Lagrange multipliers**

$$
\begin{aligned}
D_{\bar{\mu}} \quad &:= \quad \inf_{\mu \in \mathcal{P}(\mu_1, \mu_V, \mu_2)} H(\mu, \bar{\mu}) \\
&= \quad \inf_{\mu \in \mathcal{M}_1} \sup_{u \in \mathcal{U}} \Big\{ H(\mu, \bar{\mu}) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \\
&\qquad -\mathbb{E}^\mu \Big[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(S_1, V, S_2) + \Delta_L^{(L)}(S_1, V, S_2) \Big] \Big\} \\
&= \quad \sup_{u \in \mathcal{U}} \inf_{\mu \in \mathcal{M}_1} \Big\{ H(\mu, \bar{\mu}) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \\
&\qquad -\mathbb{E}^\mu \Big[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(S_1, V, S_2) + \Delta_L^{(L)}(S_1, V, S_2) \Big] \Big\}
\end{aligned}
$$

© 2020 Bloomberg Finance L.P. All rights reserved.

Build a model in $\mathcal{P}(\mu_1, \mu_V, \mu_2)$

$$D_{\bar{\mu}} = \sup_{u \in \mathcal{U}} \inf_{\mu \in \mathcal{M}_1} \Big\{ H(\mu, \bar{\mu}) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] - \mathbb{E}^\mu \Big[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(S_1, V, S_2) + \Delta_L^{(L)}(S_1, V, S_2) \Big] \Big\}$$

- **Remarkable fact: The inner infimum can be exactly computed:**

$$\inf_{\mu \in \mathcal{M}_1} \{ H(\mu, \bar{\mu}) - \mathbb{E}^\mu[X] \} = -\ln \mathbb{E}^{\bar{\mu}} \Big[e^X \Big]$$

and **the infimum is attained at $\mu = \bar{\mu}_X$ defined by** (Gibbs type)

$$\frac{d\bar{\mu}_X}{d\bar{\mu}} = \frac{e^X}{\mathbb{E}^{\bar{\mu}}[e^X]}.$$

- That is why we like (and chose) the "distance" $H(\mu, \bar{\mu})$!

© 2020 Bloomberg Finance L.P. All rights reserved.

Build a model in $\mathcal{P}(\mu_1, \mu_V, \mu_2)$

$$D_{\bar{\mu}} := \inf_{\mu \in \mathcal{P}(\mu_1, \mu_V, \mu_2)} H(\mu, \bar{\mu}) \;=\; \sup_{u \in \mathcal{U}} \Psi_{\bar{\mu}}(u) =: P_{\bar{\mu}}$$

$$\Psi_{\bar{\mu}}(u) := \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] - \ln \mathbb{E}^{\bar{\mu}} \left[e^{u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(S_1, V, S_2) + \Delta_L^{(L)}(S_1, V, S_2)} \right].$$

- $\inf_{\mu \in \mathcal{P}(\mu_1, \mu_V, \mu_2)}$: **constrained** optimization, **difficult**.
- $\sup_{u \in \mathcal{U}}$: **unconstrained** optimization, **easy!** To find the optimum $u^* = (u_1^*, u_V^*, u_2^*, \Delta_S^*, \Delta_L^*)$, simply cancel the gradient of $\Psi_{\bar{\mu}}$.
- Most important, $\inf_{\mu \in \mathcal{P}(\mu_1, \mu_V, \mu_2)} H(\mu, \bar{\mu})$ is reached at

$$\mu^*(ds_1, dv, ds_2) = \bar{\mu}(ds_1, dv, ds_2) \frac{e^{u_1^*(s_1) + u_V^*(v) + u_2^*(s_2) + \Delta_S^{*(S)}(s_1, v, s_2) + \Delta_L^{*(L)}(s_1, v, s_2)}}{\mathbb{E}^{\bar{\mu}} \left[e^{u_1^*(S_1) + u_V^*(V) + u_2^*(S_2) + \Delta_S^{*(S)}(S_1, V, S_2) + \Delta_L^{*(L)}(S_1, V, S_2)} \right]}.$$

- **Problem solved:** $\mu^* \in \mathcal{P}(\mu_1, \mu_V, \mu_2)$**!**

© 2020 Bloomberg Finance L.P. All rights reserved.

Equations for $u^* = (u_1^*, u_V^*, u_2^*, \Delta_S^*, \Delta_L^*)$

$$
\begin{aligned}
\tfrac{\partial \Psi_{\bar\mu}}{\partial u_1(s_1)} = 0: \quad & \forall s_1 > 0, & u_1(s_1) &= \Phi_1(s_1; u_V, u_2, \Delta_S, \Delta_L) \\
\tfrac{\partial \Psi_{\bar\mu}}{\partial u_V(v)} = 0: \quad & \forall v \geq 0, & u_V(v) &= \Phi_V(v; u_1, u_2, \Delta_S, \Delta_L) \\
\tfrac{\partial \Psi_{\bar\mu}}{\partial u_2(s_2)} = 0: \quad & \forall s_2 > 0, & u_2(s_2) &= \Phi_2(s_2; u_1, u_V, \Delta_S, \Delta_L) \\
\tfrac{\partial \Psi_{\bar\mu}}{\partial \Delta_S(s_1,v)} = 0: \quad & \forall s_1 > 0,\ \forall v \geq 0, & 0 &= \Phi_{\Delta_S}(s_1, v; \Delta_S(s_1,v), \Delta_L(s_1,v)) \\
\tfrac{\partial \Psi_{\bar\mu}}{\partial \Delta_L(s_1,v)} = 0: \quad & \forall s_1 > 0,\ \forall v \geq 0, & 0 &= \Phi_{\Delta_L}(s_1, v; \Delta_S(s_1,v), \Delta_L(s_1,v))
\end{aligned}
$$

- We could have simply postulated a model of the form

$$
\mu(ds_1, dv, ds_2) = \bar\mu(ds_1, dv, ds_2) \frac{e^{u_1(s_1)+u_V(v)+u_2(s_2)+\Delta_S^{(S)}(s_1,v,s_2)+\Delta_L^{(L)}(s_1,v,s_2)}}{\mathbb{E}^{\bar\mu}\left[e^{u_1(S_1)+u_V(V)+u_2(S_2)+\Delta_S^{(S)}(S_1,V,S_2)+\Delta_L^{(L)}(S_1,V,S_2)}\right]}.
$$

- Then the 5 conditions defining $\mathcal{P}(\mu_1, \mu_V, \mu_2)$ translate into the 5 above equations.

© 2020 Bloomberg Finance L.P. All rights reserved.

Sinkhorn's algorithm

- Sinkhorn's algorithm (1967) was first used in the context of optimal transport by Cuturi (2013).
- In our context: **fixed point method that iterates computions of one-dimensional gradients to approximate the optimizer u^*.**
- Start from initial guess $u^{(0)} = (u_1^{(0)}, u_V^{(0)}, u_2^{(0)}, \Delta_S^{(0)}, \Delta_L^{(0)})$, recursively define $u^{(n+1)}$ knowing $u^{(n)}$ by

$$
\begin{aligned}
\forall s_1 > 0, \quad & u_1^{(n+1)}(s_1) &&= \Phi_1(s_1; u_V^{(n)}, u_2^{(n)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall v \geq 0, \quad & u_V^{(n+1)}(v) &&= \Phi_V(v; u_1^{(n+1)}, u_2^{(n)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall s_2 > 0, \quad & u_2^{(n+1)}(s_2) &&= \Phi_2(s_2; u_1^{(n+1)}, u_V^{(n+1)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall s_1 > 0, \ \forall v \geq 0, \quad & 0 &&= \Phi_{\Delta_S}(s_1, v; u_2^{(n+1)}, \Delta_S^{(n+1)}(s_1, v), \Delta_L^{(n)}(s_1, v)) \\
\forall s_1 > 0, \ \forall v \geq 0, \quad & 0 &&= \Phi_{\Delta_L}(s_1, v; u_2^{(n+1)}, \Delta_S^{(n+1)}(s_1, v), \Delta_L^{(n+1)}(s_1, v))
\end{aligned}
$$

  until convergence.
- Each of the above 5 lines corresponds to a Bregman projection in the space of measures.

© 2020 Bloomberg Finance L.P. All rights reserved.

Sinkhorn's algorithm

- Sinkhorn's algorithm (1967) was first used in the context of optimal transport by Cuturi (2013).
- In our context: **fixed point method that iterates computions of one-dimensional gradients to approximate the optimizer** u^*.
- Start from initial guess $u^{(0)} = (u_1^{(0)}, u_V^{(0)}, u_2^{(0)}, \Delta_S^{(0)}, \Delta_L^{(0)})$, recursively define $u^{(n+1)}$ knowing $u^{(n)}$ by

$$
\begin{aligned}
\forall s_1 > 0, \quad & u_1^{(n+1)}(s_1) &&= \Phi_1(s_1; u_V^{(n)}, u_2^{(n)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall v \geq 0, \quad & u_V^{(n+1)}(v) &&= \Phi_V(v; u_1^{(n+1)}, u_2^{(n)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall s_2 > 0, \quad & u_2^{(n+1)}(s_2) &&= \Phi_2(s_2; u_1^{(n+1)}, u_V^{(n+1)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall s_1 > 0, \ \forall v \geq 0, \quad & 0 &&= \Phi_{\Delta_S}(s_1, v; u_2^{(n+1)}, \Delta_S^{(n+1)}(s_1, v), \Delta_L^{(n)}(s_1, v)) \\
\forall s_1 > 0, \ \forall v \geq 0, \quad & 0 &&= \Phi_{\Delta_L}(s_1, v; u_2^{(n+1)}, \Delta_S^{(n+1)}(s_1, v), \Delta_L^{(n+1)}(s_1, v))
\end{aligned}
$$

  until convergence.
- Each of the above 5 lines corresponds to a Bregman projection in the space of measures.

© 2020 Bloomberg Finance L.P. All rights reserved.

Sinkhorn's algorithm

- Sinkhorn's algorithm (1967) was first used in the context of optimal transport by Cuturi (2013).
- In our context: **fixed point method that iterates computions of one-dimensional gradients to approximate the optimizer** u^*.
- Start from initial guess $u^{(0)} = (u_1^{(0)}, u_V^{(0)}, u_2^{(0)}, \Delta_S^{(0)}, \Delta_L^{(0)})$, recursively define $u^{(n+1)}$ knowing $u^{(n)}$ by

$$
\begin{aligned}
\forall s_1 > 0, \quad & u_1^{(n+1)}(s_1) &&= \Phi_1(s_1; u_V^{(n)}, u_2^{(n)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall v \geq 0, \quad & u_V^{(n+1)}(v) &&= \Phi_V(v; u_1^{(n+1)}, u_2^{(n)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall s_2 > 0, \quad & u_2^{(n+1)}(s_2) &&= \Phi_2(s_2; u_1^{(n+1)}, u_V^{(n+1)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall s_1 > 0, \ \forall v \geq 0, \quad & 0 &&= \Phi_{\Delta_S}(s_1, v; u_2^{(n+1)}, \Delta_S^{(n+1)}(s_1, v), \Delta_L^{(n)}(s_1, v)) \\
\forall s_1 > 0, \ \forall v \geq 0, \quad & 0 &&= \Phi_{\Delta_L}(s_1, v; u_2^{(n+1)}, \Delta_S^{(n+1)}(s_1, v), \Delta_L^{(n+1)}(s_1, v))
\end{aligned}
$$

  until convergence.
- Each of the above 5 lines corresponds to a Bregman projection in the space of measures.

© 2020 Bloomberg Finance L.P. All rights reserved.

Sinkhorn's algorithm

- Sinkhorn's algorithm (1967) was first used in the context of optimal transport by Cuturi (2013).
- In our context: **fixed point method that iterates computions of one-dimensional gradients to approximate the optimizer** u^*.
- Start from initial guess $u^{(0)} = (u_1^{(0)}, u_V^{(0)}, u_2^{(0)}, \Delta_S^{(0)}, \Delta_L^{(0)})$, recursively define $u^{(n+1)}$ knowing $u^{(n)}$ by

$$
\begin{aligned}
\forall s_1 > 0, \quad u_1^{(n+1)}(s_1) &= \Phi_1(s_1; u_V^{(n)}, u_2^{(n)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall v \geq 0, \quad u_V^{(n+1)}(v) &= \Phi_V(v; u_1^{(n+1)}, u_2^{(n)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall s_2 > 0, \quad u_2^{(n+1)}(s_2) &= \Phi_2(s_2; u_1^{(n+1)}, u_V^{(n+1)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall s_1 > 0, \ \forall v \geq 0, \quad 0 &= \Phi_{\Delta_S}(s_1, v; u_2^{(n+1)}, \Delta_S^{(n+1)}(s_1, v), \Delta_L^{(n)}(s_1, v)) \\
\forall s_1 > 0, \ \forall v \geq 0, \quad 0 &= \Phi_{\Delta_L}(s_1, v; u_2^{(n+1)}, \Delta_S^{(n+1)}(s_1, v), \Delta_L^{(n+1)}(s_1, v))
\end{aligned}
$$

  until convergence.
- Each of the above 5 lines corresponds to a Bregman projection in the space of measures.

© 2020 Bloomberg Finance L.P. All rights reserved.

Sinkhorn's algorithm

- Sinkhorn's algorithm (1967) was first used in the context of optimal transport by Cuturi (2013).
- In our context: **fixed point method that iterates computions of one-dimensional gradients to approximate the optimizer** u^*.
- Start from initial guess $u^{(0)} = (u_1^{(0)}, u_V^{(0)}, u_2^{(0)}, \Delta_S^{(0)}, \Delta_L^{(0)})$, recursively define $u^{(n+1)}$ knowing $u^{(n)}$ by

$$
\begin{aligned}
\forall s_1 > 0, \quad u_1^{(n+1)}(s_1) &= \Phi_1(s_1; u_V^{(n)}, u_2^{(n)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall v \geq 0, \quad u_V^{(n+1)}(v) &= \Phi_V(v; u_1^{(n+1)}, u_2^{(n)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall s_2 > 0, \quad u_2^{(n+1)}(s_2) &= \Phi_2(s_2; u_1^{(n+1)}, u_V^{(n+1)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall s_1 > 0, \ \forall v \geq 0, \quad 0 &= \Phi_{\Delta_S}(s_1, v; u_2^{(n+1)}, \Delta_S^{(n+1)}(s_1, v), \Delta_L^{(n)}(s_1, v)) \\
\forall s_1 > 0, \ \forall v \geq 0, \quad 0 &= \Phi_{\Delta_L}(s_1, v; u_2^{(n+1)}, \Delta_S^{(n+1)}(s_1, v), \Delta_L^{(n+1)}(s_1, v))
\end{aligned}
$$

  until convergence.
- Each of the above 5 lines corresponds to a Bregman projection in the space of measures.

© 2020 Bloomberg Finance L.P. All rights reserved.

Sinkhorn's algorithm

- Sinkhorn's algorithm (1967) was first used in the context of optimal transport by Cuturi (2013).
- In our context: **fixed point method that iterates computions of one-dimensional gradients to approximate the optimizer** u^*.
- Start from initial guess $u^{(0)} = (u_1^{(0)}, u_V^{(0)}, u_2^{(0)}, \Delta_S^{(0)}, \Delta_L^{(0)})$, recursively define $u^{(n+1)}$ knowing $u^{(n)}$ by

$$
\begin{aligned}
\forall s_1 > 0, \quad & u_1^{(n+1)}(s_1) &&= \Phi_1(s_1; u_V^{(n)}, u_2^{(n)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall v \geq 0, \quad & u_V^{(n+1)}(v) &&= \Phi_V(v; u_1^{(n+1)}, u_2^{(n)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall s_2 > 0, \quad & u_2^{(n+1)}(s_2) &&= \Phi_2(s_2; u_1^{(n+1)}, u_V^{(n+1)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall s_1 > 0, \ \forall v \geq 0, \quad & 0 &&= \Phi_{\Delta_S}(s_1, v; u_2^{(n+1)}, \Delta_S^{(n+1)}(s_1, v), \Delta_L^{(n)}(s_1, v)) \\
\forall s_1 > 0, \ \forall v \geq 0, \quad & 0 &&= \Phi_{\Delta_L}(s_1, v; u_2^{(n+1)}, \Delta_S^{(n+1)}(s_1, v), \Delta_L^{(n+1)}(s_1, v))
\end{aligned}
$$

  until convergence.
- Each of the above 5 lines corresponds to a **Bregman projection** in the space of measures.

© 2020 Bloomberg Finance L.P. All rights reserved.

Sinkhorn's algorithm

- Sinkhorn's algorithm (1967) was first used in the context of optimal transport by Cuturi (2013).
- In our context: **fixed point method that iterates computions of one-dimensional gradients to approximate the optimizer** u^*.
- Start from initial guess $u^{(0)} = (u_1^{(0)}, u_V^{(0)}, u_2^{(0)}, \Delta_S^{(0)}, \Delta_L^{(0)})$, recursively define $u^{(n+1)}$ knowing $u^{(n)}$ by

$$
\begin{aligned}
\forall s_1 > 0, \quad & u_1^{(n+1)}(s_1) &&= \Phi_1(s_1; u_V^{(n)}, u_2^{(n)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall v \geq 0, \quad & u_V^{(n+1)}(v) &&= \Phi_V(v; u_1^{(n+1)}, u_2^{(n)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall s_2 > 0, \quad & u_2^{(n+1)}(s_2) &&= \Phi_2(s_2; u_1^{(n+1)}, u_V^{(n+1)}, \Delta_S^{(n)}, \Delta_L^{(n)}) \\
\forall s_1 > 0, \ \forall v \geq 0, \quad & 0 &&= \Phi_{\Delta_S}(s_1, v; u_2^{(n+1)}, \Delta_S^{(n+1)}(s_1, v), \Delta_L^{(n)}(s_1, v)) \\
\forall s_1 > 0, \ \forall v \geq 0, \quad & 0 &&= \Phi_{\Delta_L}(s_1, v; u_2^{(n+1)}, \Delta_S^{(n+1)}(s_1, v), \Delta_L^{(n+1)}(s_1, v))
\end{aligned}
$$

- Each of the above five lines corresponds to a **Bregman projection** in the space of measures.
- **If the algorithm diverges**, then $P_{\bar{\mu}} = +\infty$, so $D_{\bar{\mu}} = +\infty$, i.e., $\mathcal{P}(\mu_1, \mu_V, \mu_2) \cap \{\mu \in \mathcal{M}_1 | \mu \ll \bar{\mu}\} = \emptyset$. In practice, when $\bar{\mu}$ has full support, this is a sign that **there likely exists a joint SPX/VIX arbitrage**.

© 2020 Bloomberg Finance L.P. All rights reserved.

Numerical experiments

© 2020 Bloomberg Finance L.P. All rights reserved.

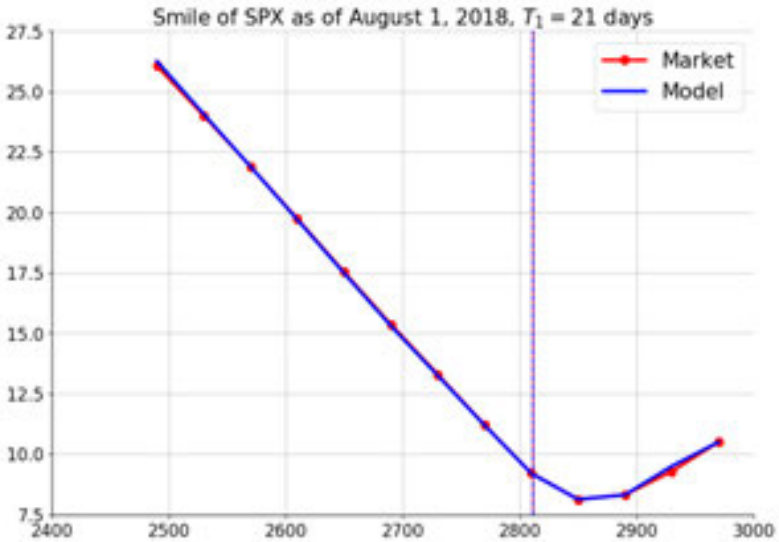
Implementation details

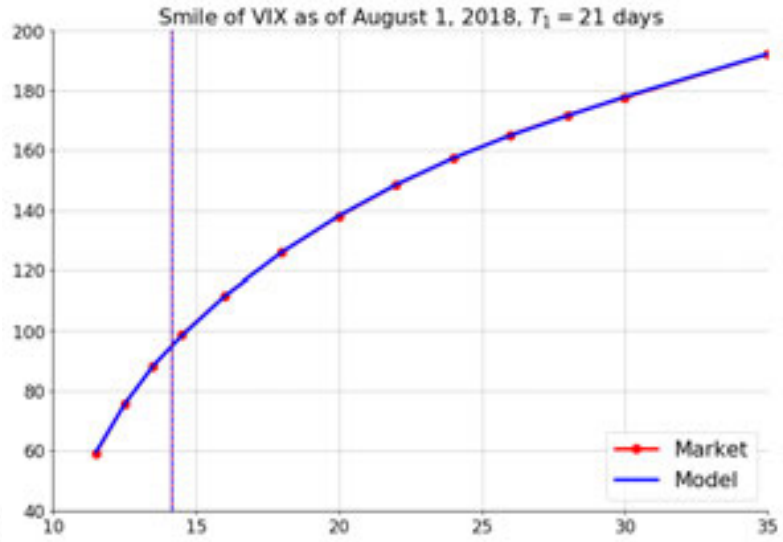
- Choice of $\bar{\mu}$:
  - $S_1 \sim \mu_1$ and $V \sim \mu_V$ independent;
  - Conditional on (S_1, V), S_2 lognormal with mean S_1 and variance V.

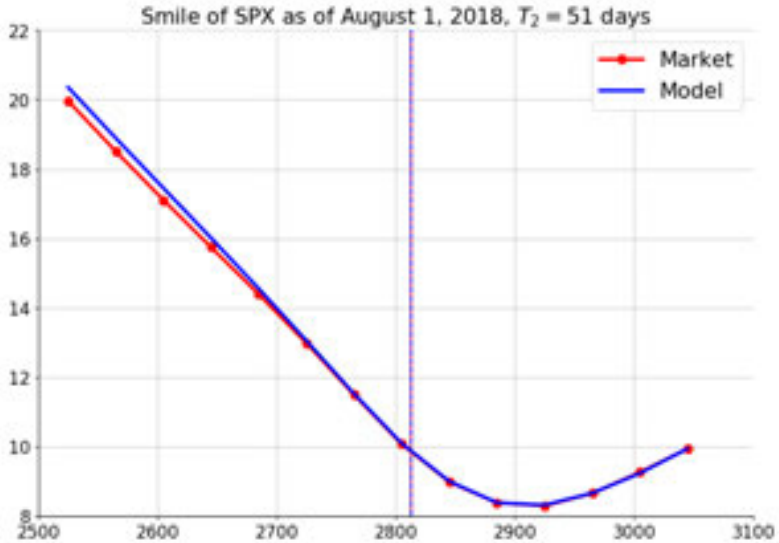
  Under $\bar{\mu}$, $S_2 \not\sim \mu_2$.
- Instead of abstract payoffs u_1, u_V, u_2, we work with market strikes and market prices of vanilla options on S_1, V, and S_2.
- Initial guess of the Sinkhorn algorithm: zero.
- Integrals in the expressions of $\Phi_1, \Phi_V, \Phi_2, \Phi_{\Delta_S}, \Phi_{\Delta_L}$ estimated using Gaussian quadrature.
- Enough accuracy is typically reached after ≈ 100 iterations.

© 2020 Bloomberg Finance L.P. All rights reserved.

August 1, 2018, $T_1 = 21$ days

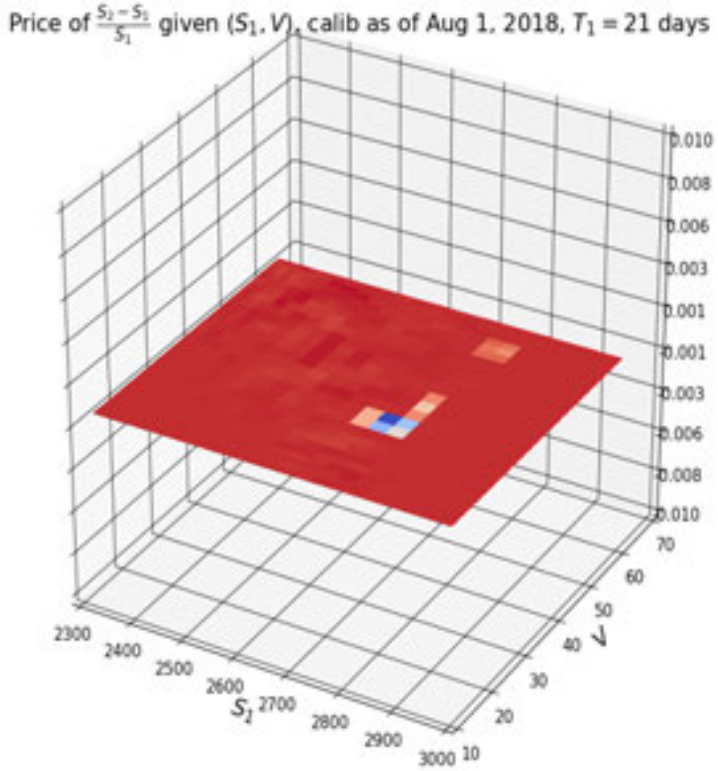


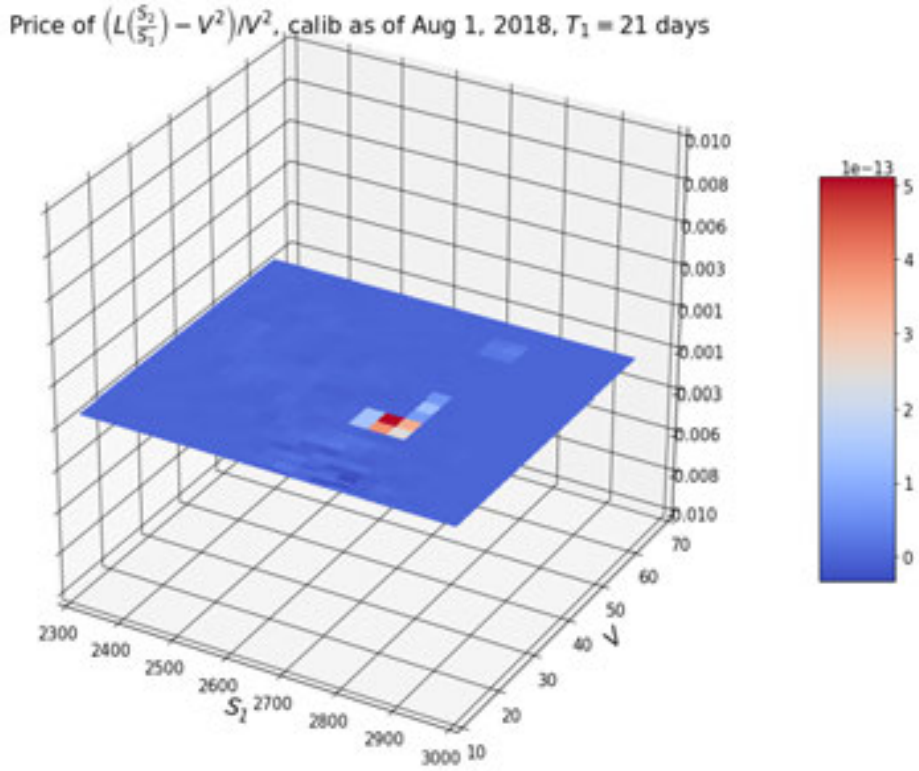




© 2020 Bloomberg Finance L.P. All rights reserved.

August 1, 2018, $T_1 = 21$ days





© 2020 Bloomberg Finance L.P. All rights reserved.

August 1, 2018, $T_1 = 21$ days

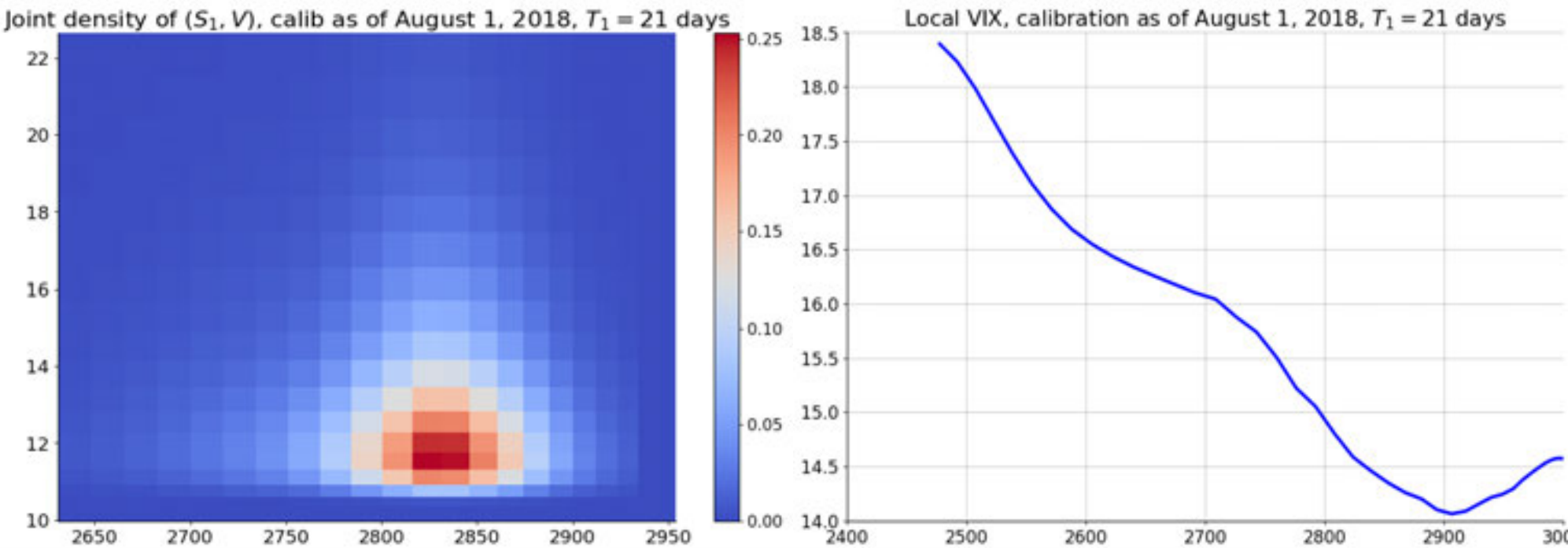
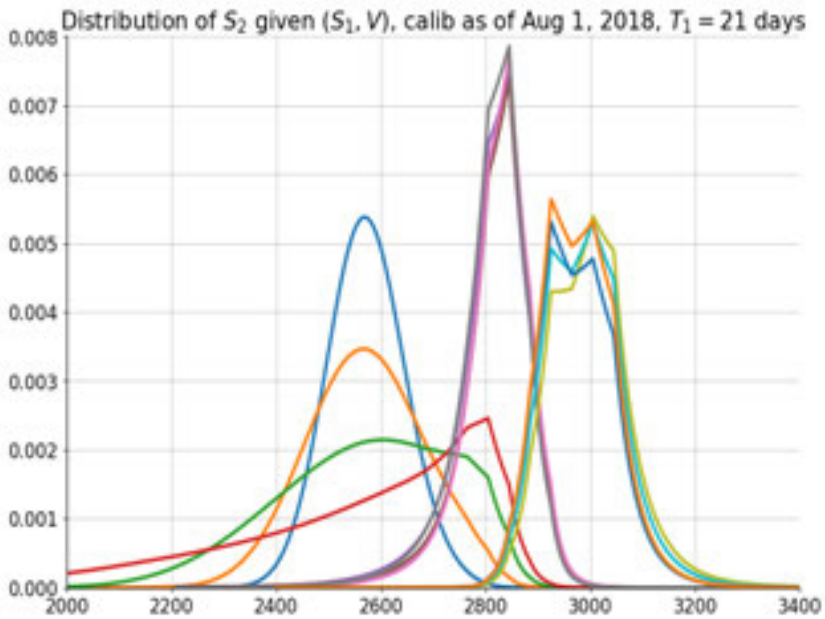


Figure: Joint distribution of (S_1, V) and local VIX function $\mathsf{VIX}_{\mathsf{loc}}(s_1)$

$$\mathsf{VIX}^2_{\mathsf{loc}}(S_1) := \mathbb{E}^{\mu^*}\left[V^2 \middle| S_1\right]$$

© 2020 Bloomberg Finance L.P. All rights reserved.

August 1, 2018, $T_1 = 21$ days



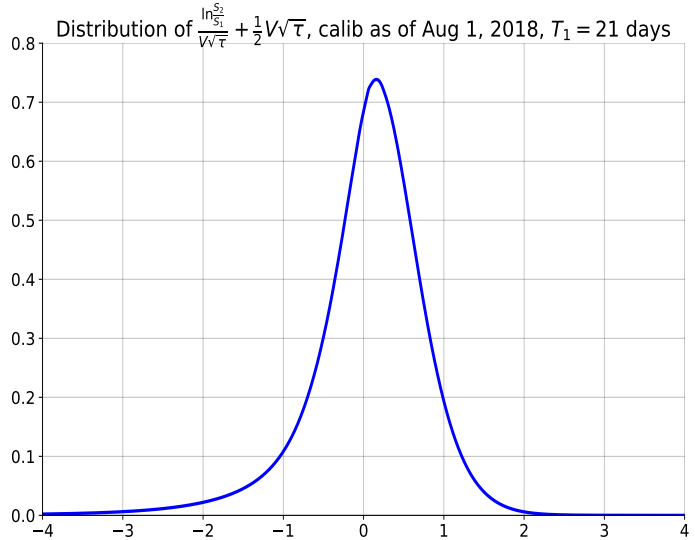


Figure: Conditional distribution of S_2 given (s_1, v) under μ^* for different vales of (s_1, v): $s_1 \in \{2571, 2808, 3000\}$, $v \in \{10.10, 15.30, 23.20, 35.72\}\%$, and distribution of the normalized return $R := \frac{\ln(S_2/S_1)}{V\sqrt{\tau}} + \frac{1}{2}V\sqrt{\tau}$

© 2020 Bloomberg Finance L.P. All rights reserved.

August 1, 2018, $T_1 = 21$ days

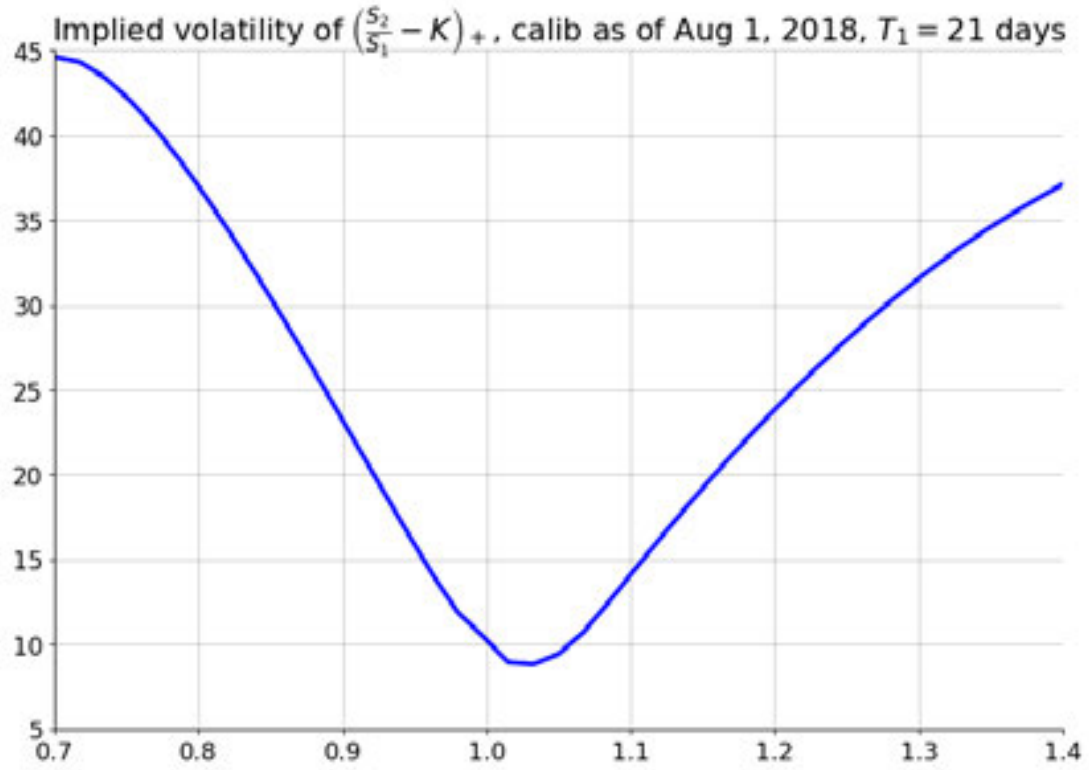
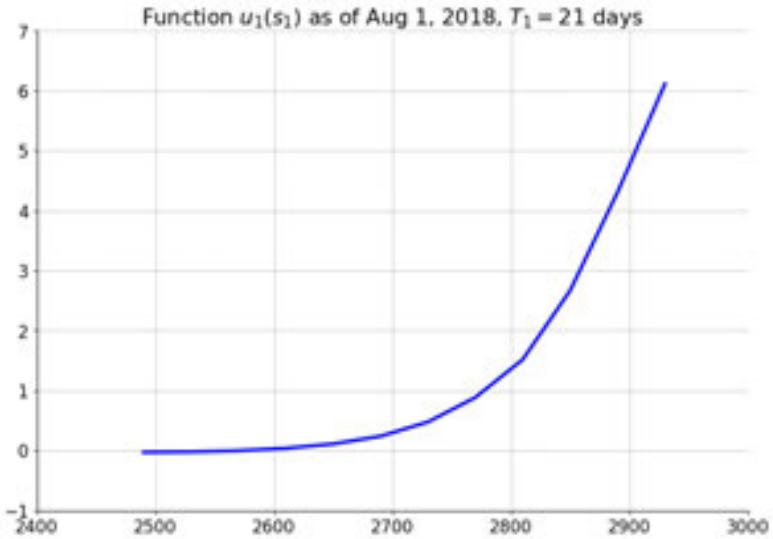
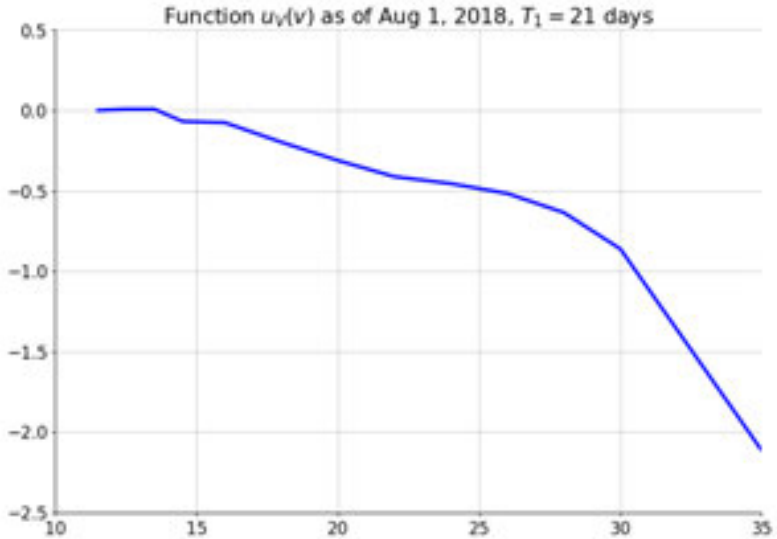


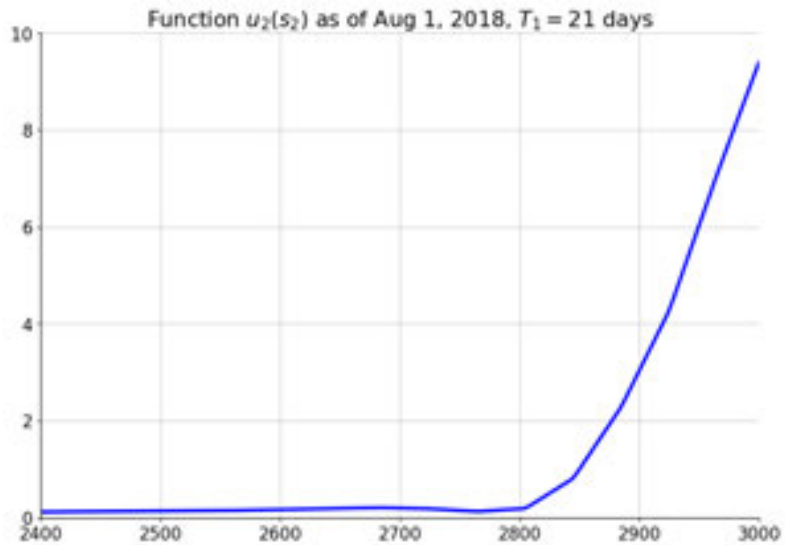
Figure: Smile of forward starting call options $(S_2/S_1 - K)_+$

© 2020 Bloomberg Finance L.P. All rights reserved.

August 1, 2018, $T_1 = 21$ days

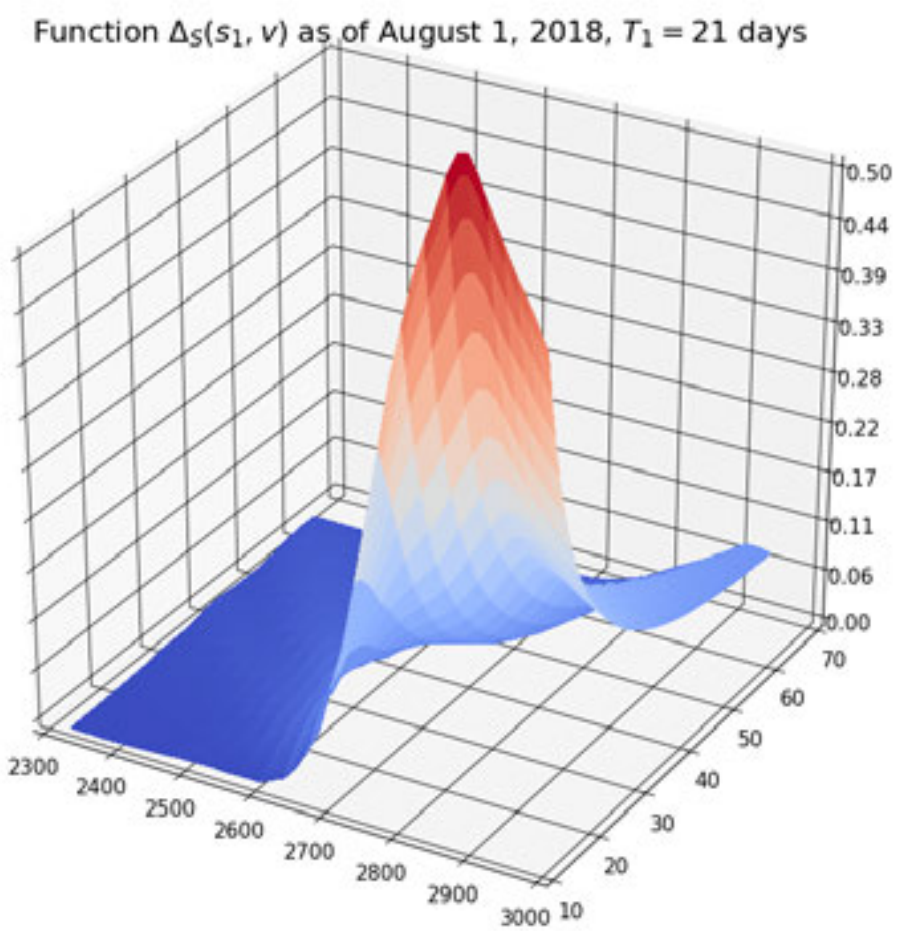






© 2020 Bloomberg Finance L.P. All rights reserved.

August 1, 2018, $T_1 = 21$ days



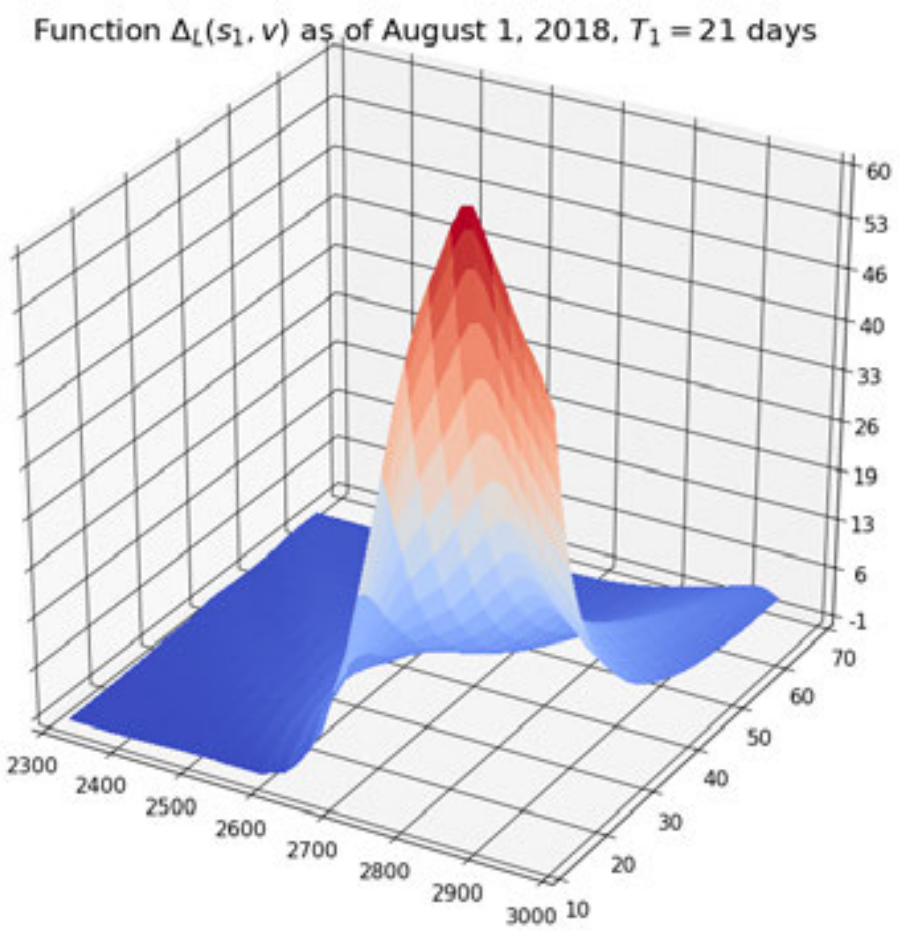
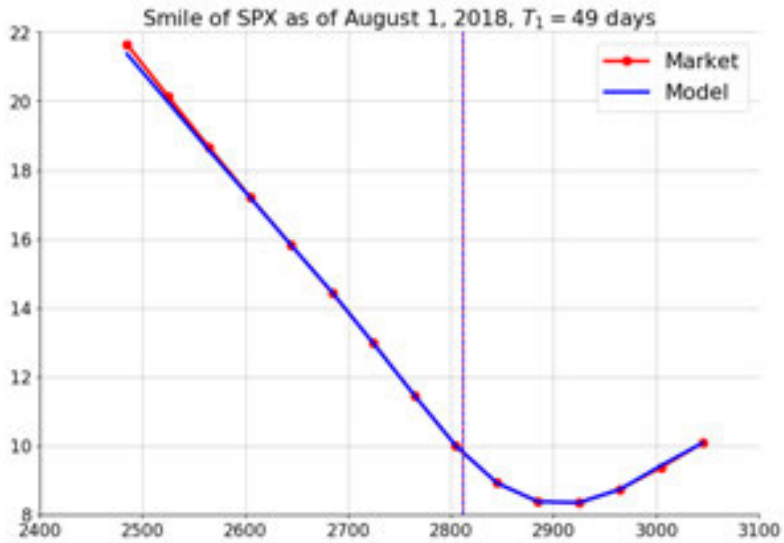
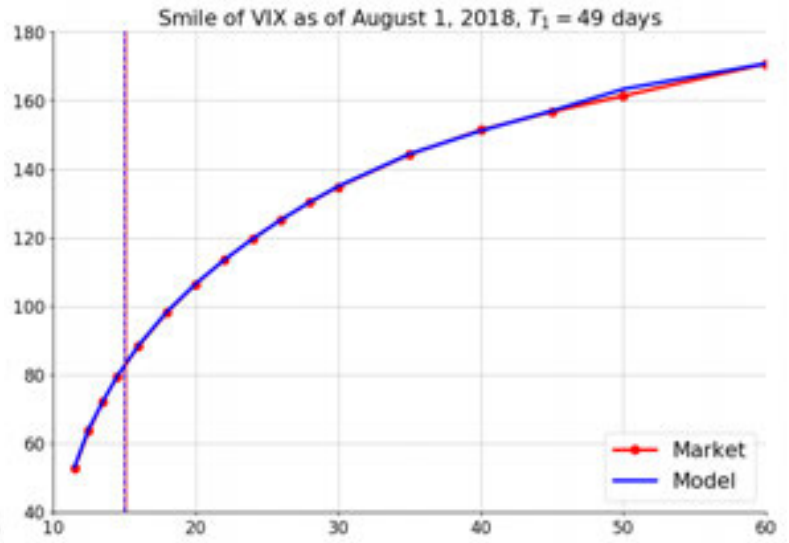


Figure: Optimal functions $\Delta_S^*(s_1, v)$ and $\Delta_L^*(s_1, v)$ for (s_1, v) in the quadrature grid

© 2020 Bloomberg Finance L.P. All rights reserved.

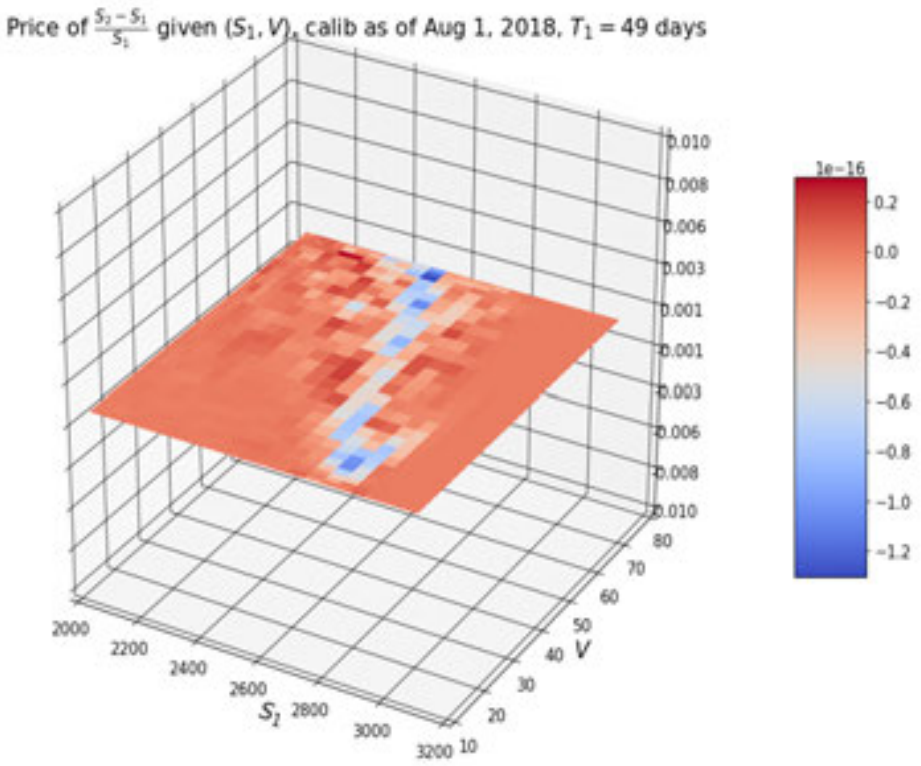
August 1, 2018, $T_1 = 49$ days

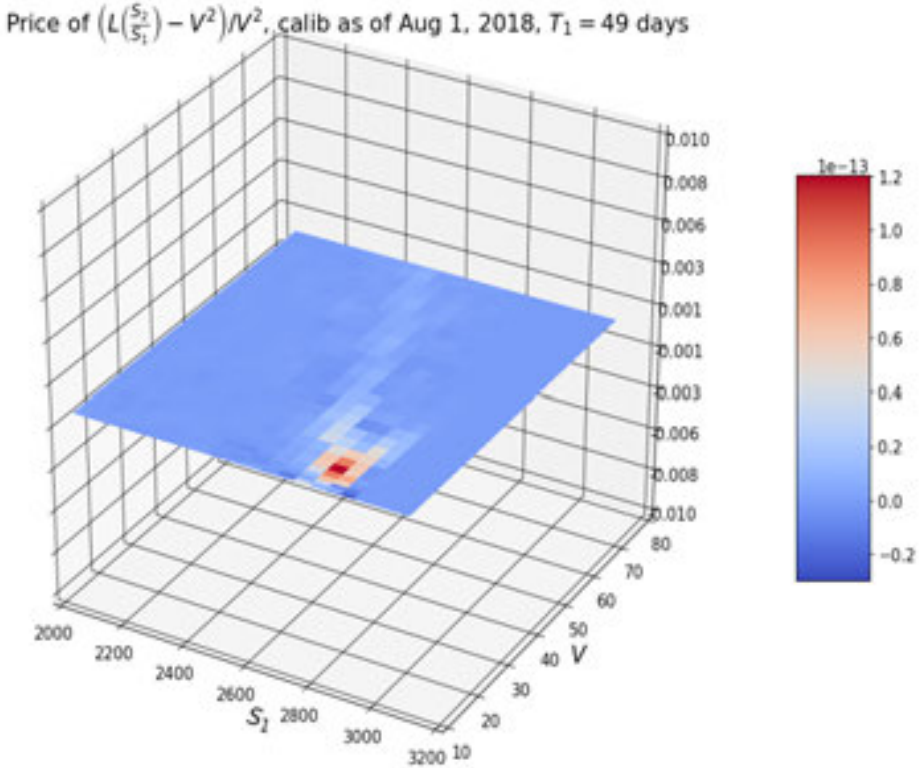




© 2020 Bloomberg Finance L.P. All rights reserved.

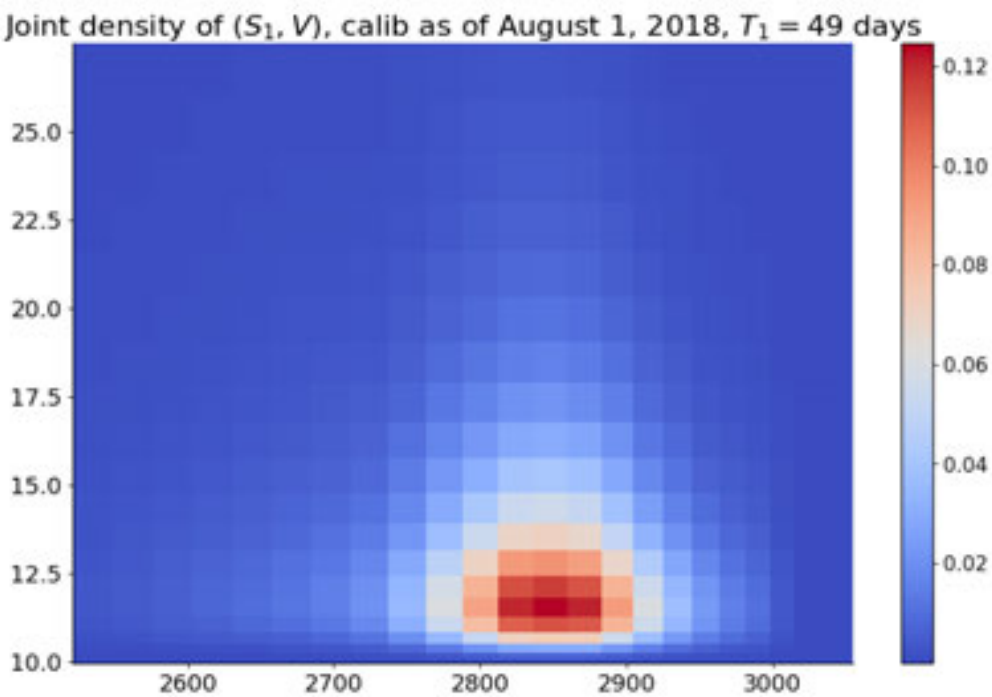
August 1, 2018, $T_1 = 49$ days





© 2020 Bloomberg Finance L.P. All rights reserved.

August 1, 2018, $T_1 = 49$ days



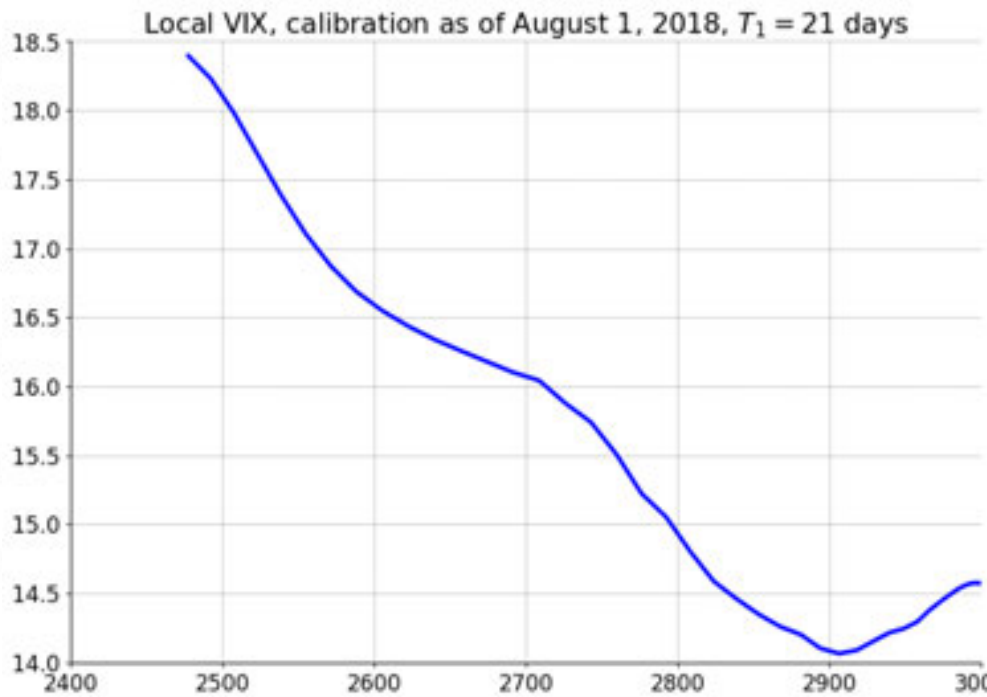


Figure: Joint distribution of (S_1, V) and local VIX function $\mathrm{VIX}_{\mathrm{loc}}(s_1)$

© 2020 Bloomberg Finance L.P. All rights reserved.

August 1, 2018, $T_1 = 49$ days

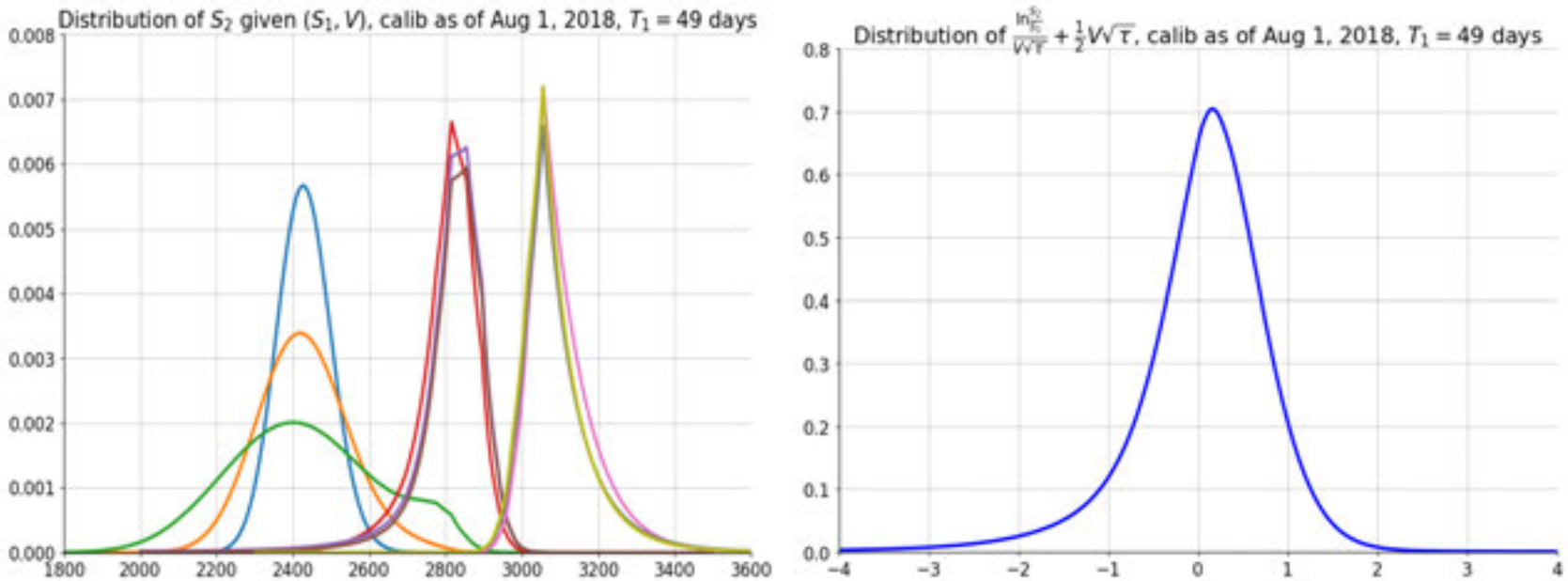


Figure: Conditional distribution of S_2 given (s_1, v) under $\mu^*_{\mathcal{K}}$ for different vales of (s_1, v): $s_1 \in \{2571, 2808, 3000\}$, $v \in \{10.10, 15.30, 23.20, 35.72\}\%$, and distribution of the normalized return $R := \frac{\ln(S_2/S_1)}{V\sqrt{\tau}} + \frac{1}{2}V\sqrt{\tau}$

© 2020 Bloomberg Finance L.P. All rights reserved.

August 1, 2018, $T_1 = 49$ days

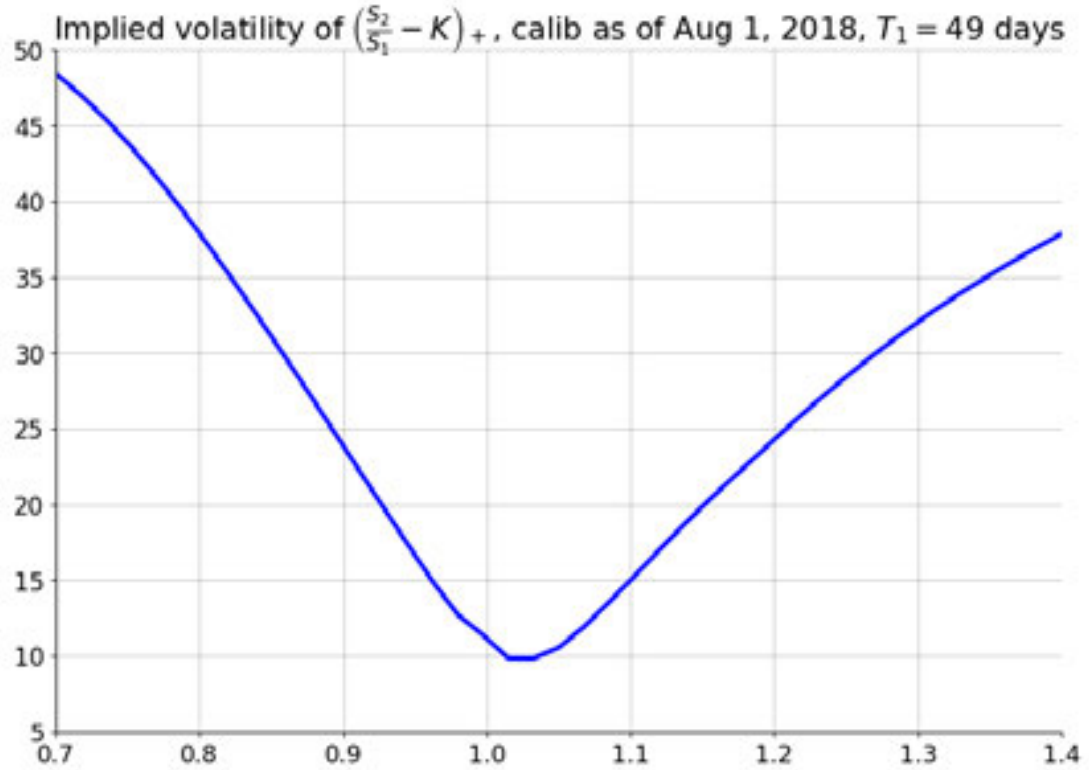
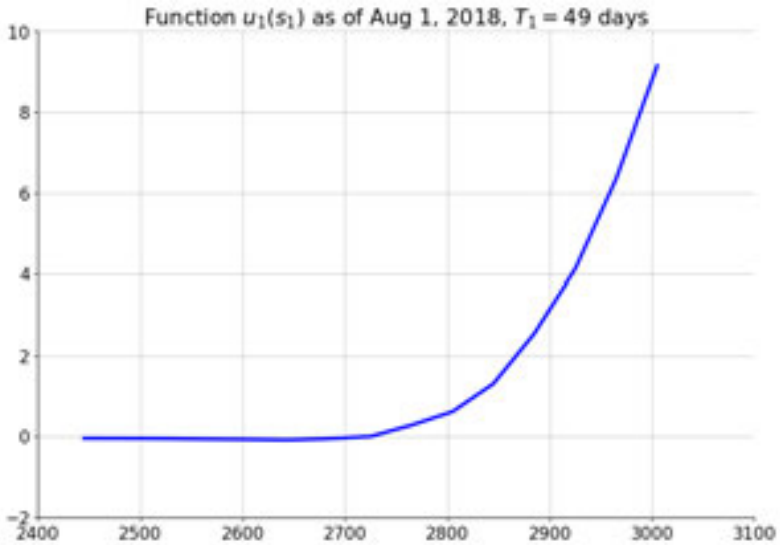
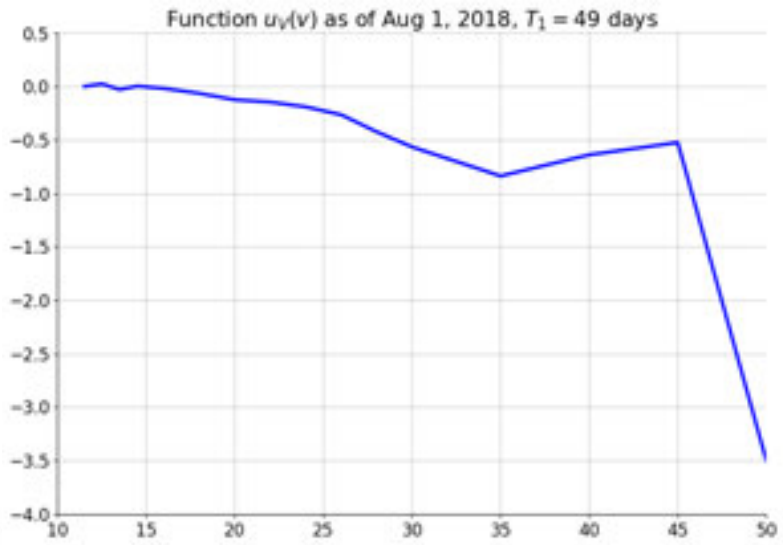


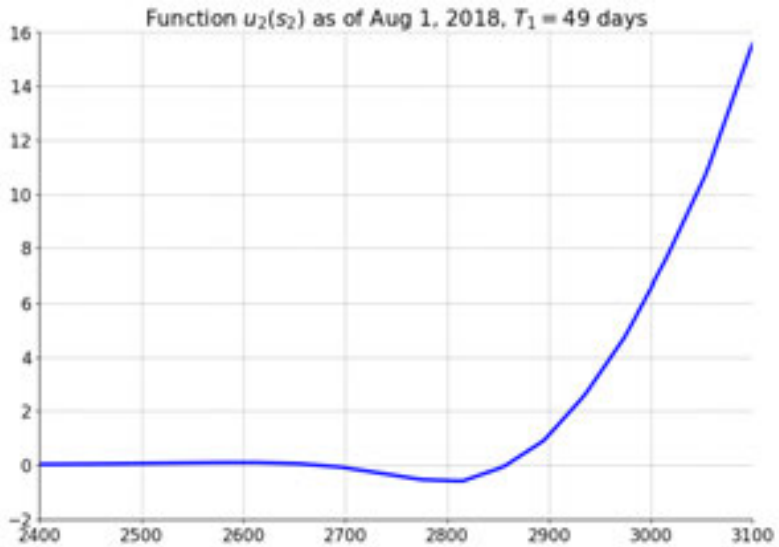
Figure: Smile of forward starting call options $(S_2/S_1 - K)_+$

© 2020 Bloomberg Finance L.P. All rights reserved.

August 1, 2018, $T_1 = 49$ days







© 2020 Bloomberg Finance L.P. All rights reserved.

August 1, 2018, $T_1 = 49$ days

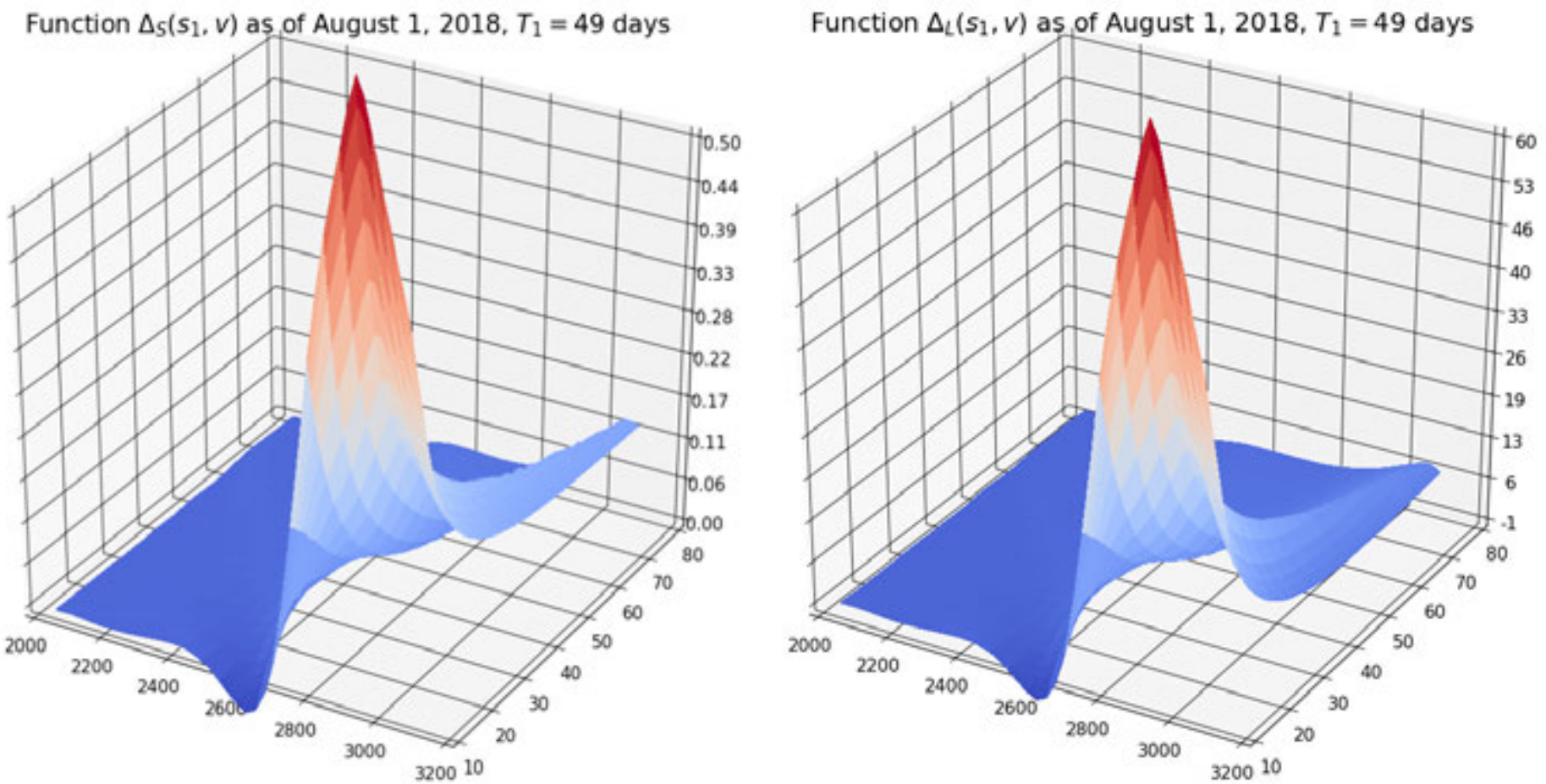
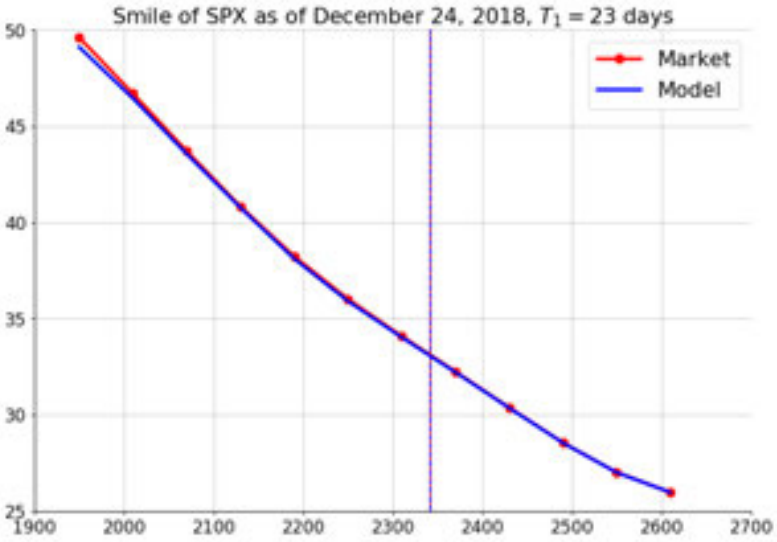
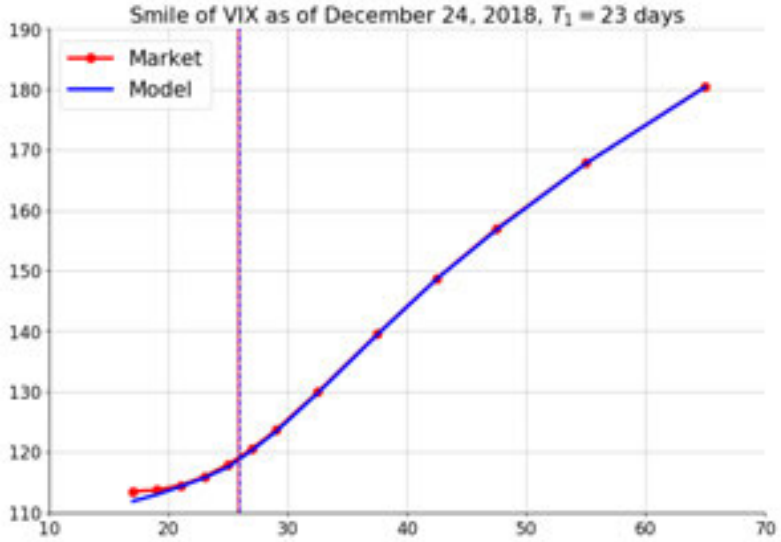


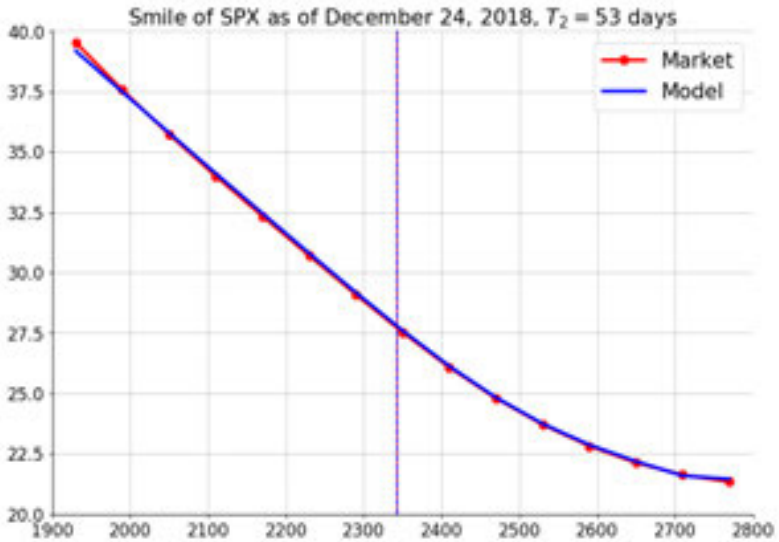
Figure: Optimal functions $\Delta_S^*(s_1, v)$ and $\Delta_L^*(s_1, v)$ for (s_1, v) in the quadrature grid

© 2020 Bloomberg Finance L.P. All rights reserved.

December 24, 2018, $T_1 = 23$ days: large VIX, $F_V \approx 26\%$

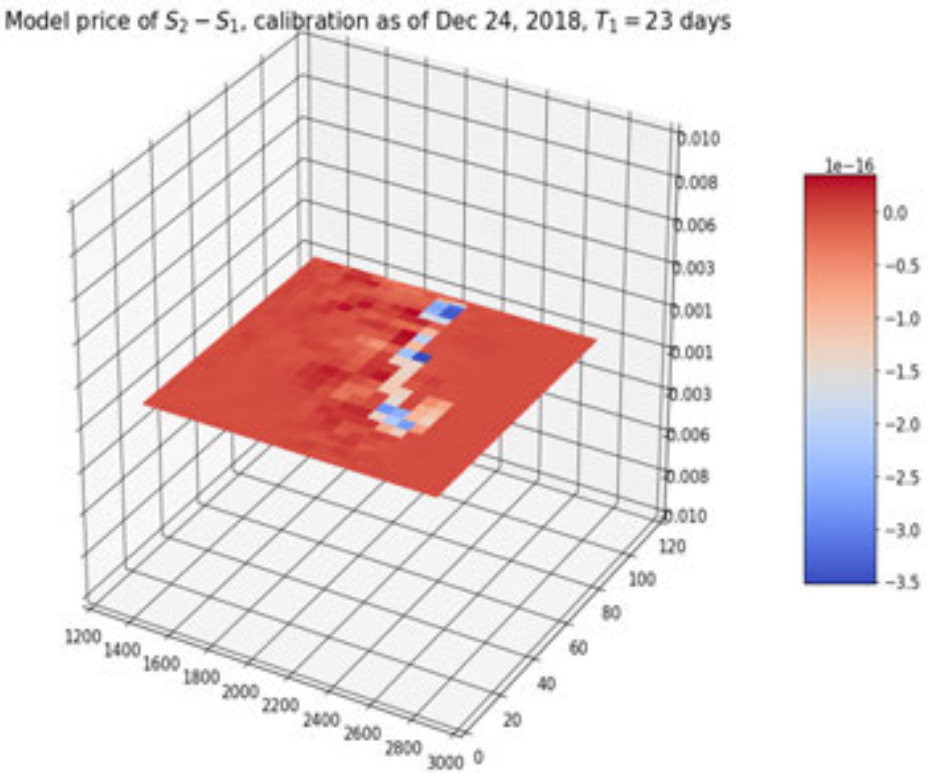


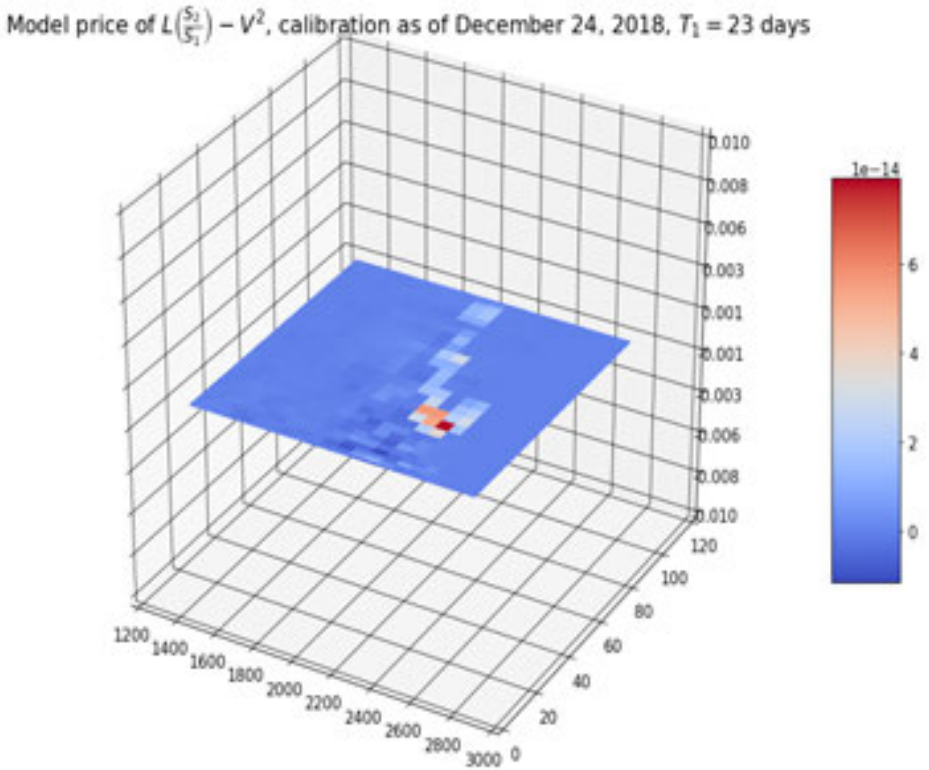




© 2020 Bloomberg Finance L.P. All rights reserved.

December 24, 2018, $T_1 = 23$ days





© 2020 Bloomberg Finance L.P. All rights reserved.

December 24, 2018, $T_1 = 23$ days

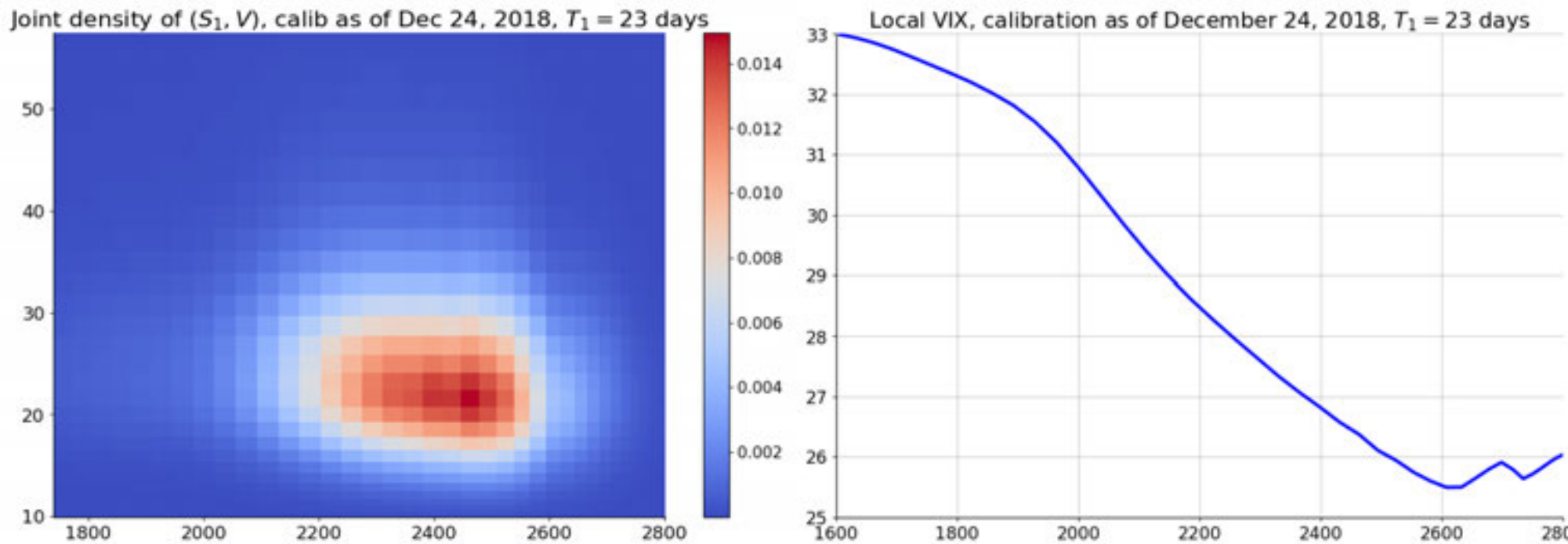


Figure: Joint distribution of (S_1, V) and local VIX function $\mathsf{VIX}_{\mathsf{loc}}(s_1)$

© 2020 Bloomberg Finance L.P. All rights reserved.

December 24, 2018, $T_1 = 23$ days

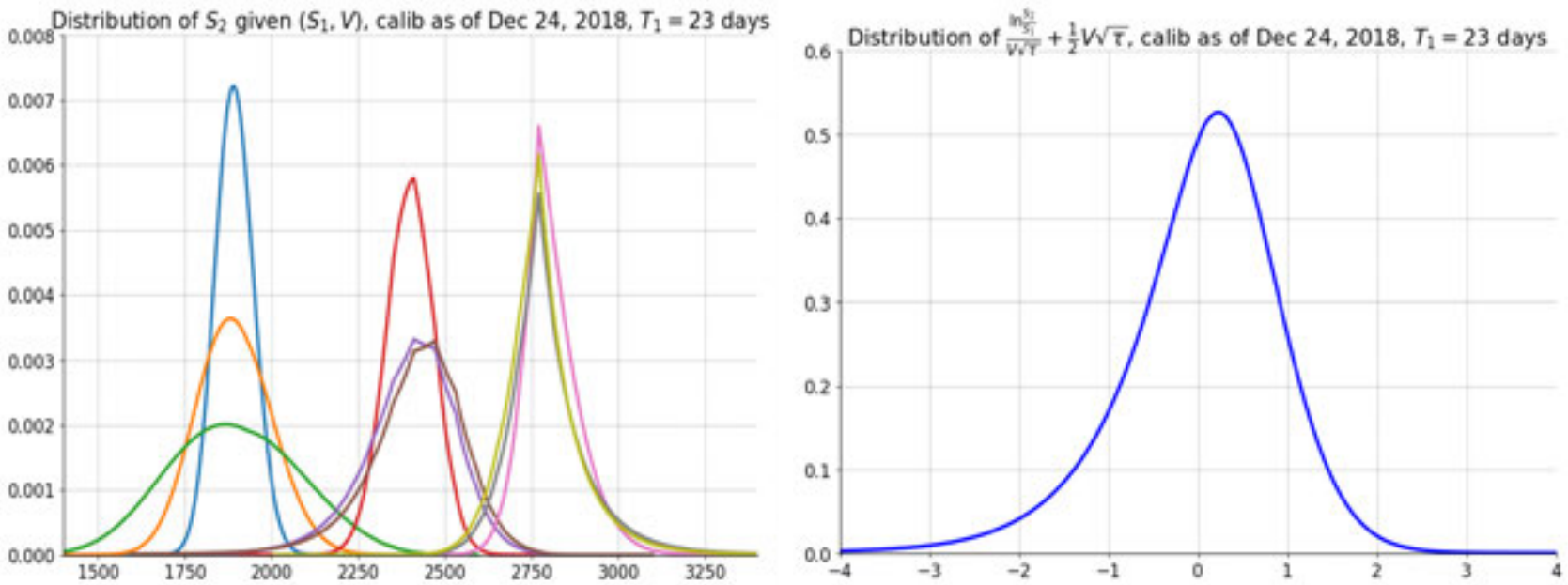


Figure: Conditional distribution of S_2 given (s_1, v) under μ^* for different vales of (s_1, v): $s_1 \in \{2571, 2808, 3000\}$, $v \in \{10.10, 15.30, 23.20, 35.72\}\%$, and distribution of the normalized return $R := \frac{\ln(S_2/S_1)}{V\sqrt{\tau}} + \frac{1}{2}V\sqrt{\tau}$

© 2020 Bloomberg Finance L.P. All rights reserved.

December 24, 2018, $T_1 = 23$ days

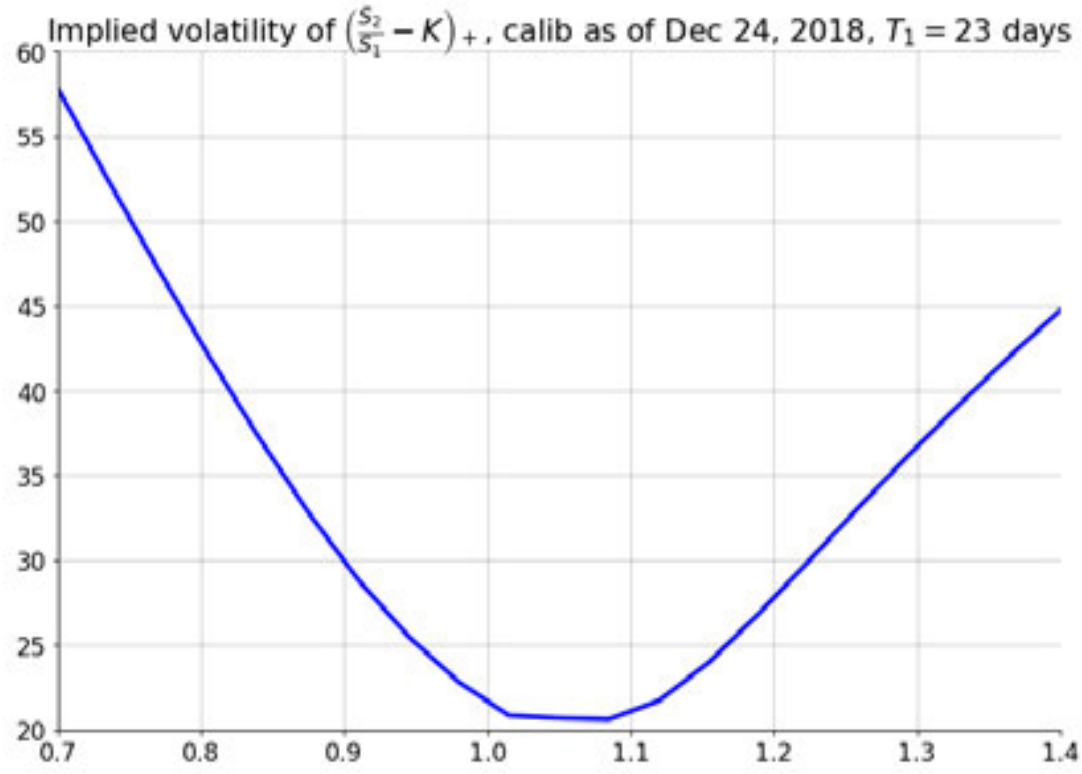
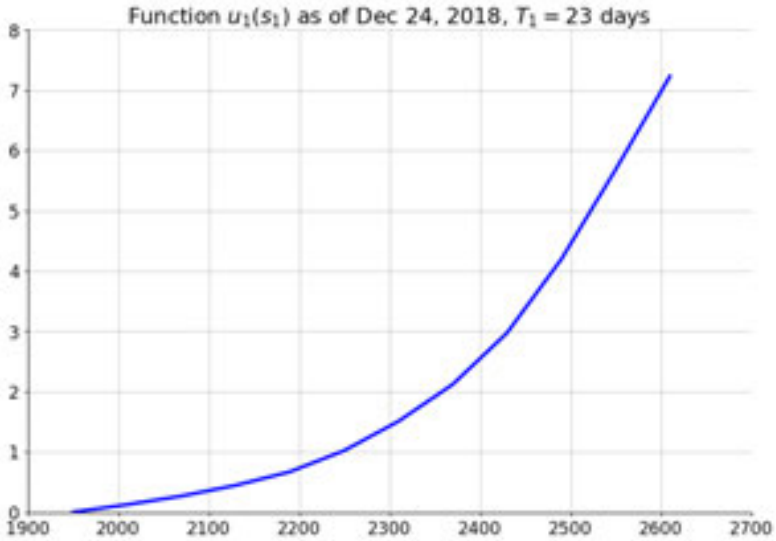
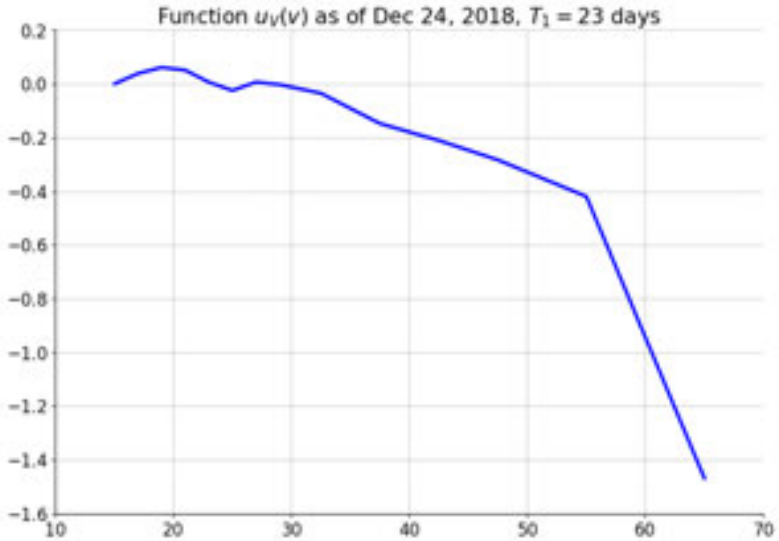


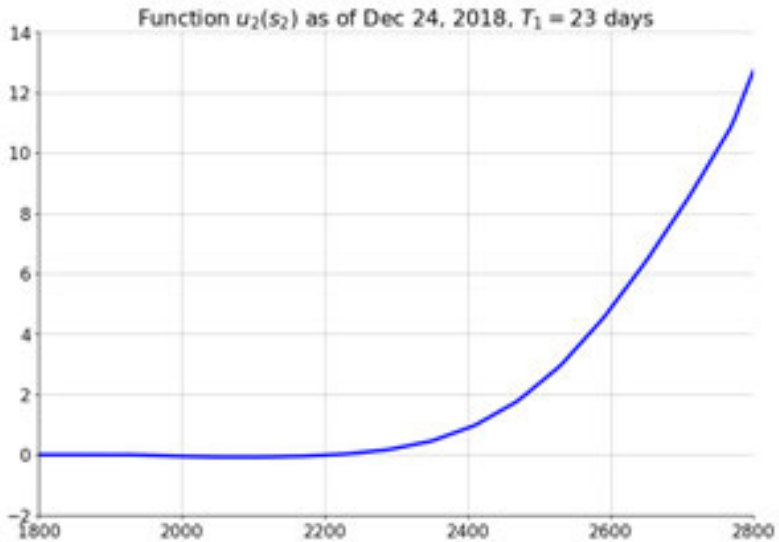
Figure: Smile of forward starting call options $(S_2/S_1 - K)_+$

© 2020 Bloomberg Finance L.P. All rights reserved.

December 24, 2018, $T_1 = 23$ days







© 2020 Bloomberg Finance L.P. All rights reserved.

December 24, 2018, $T_1 = 23$ days

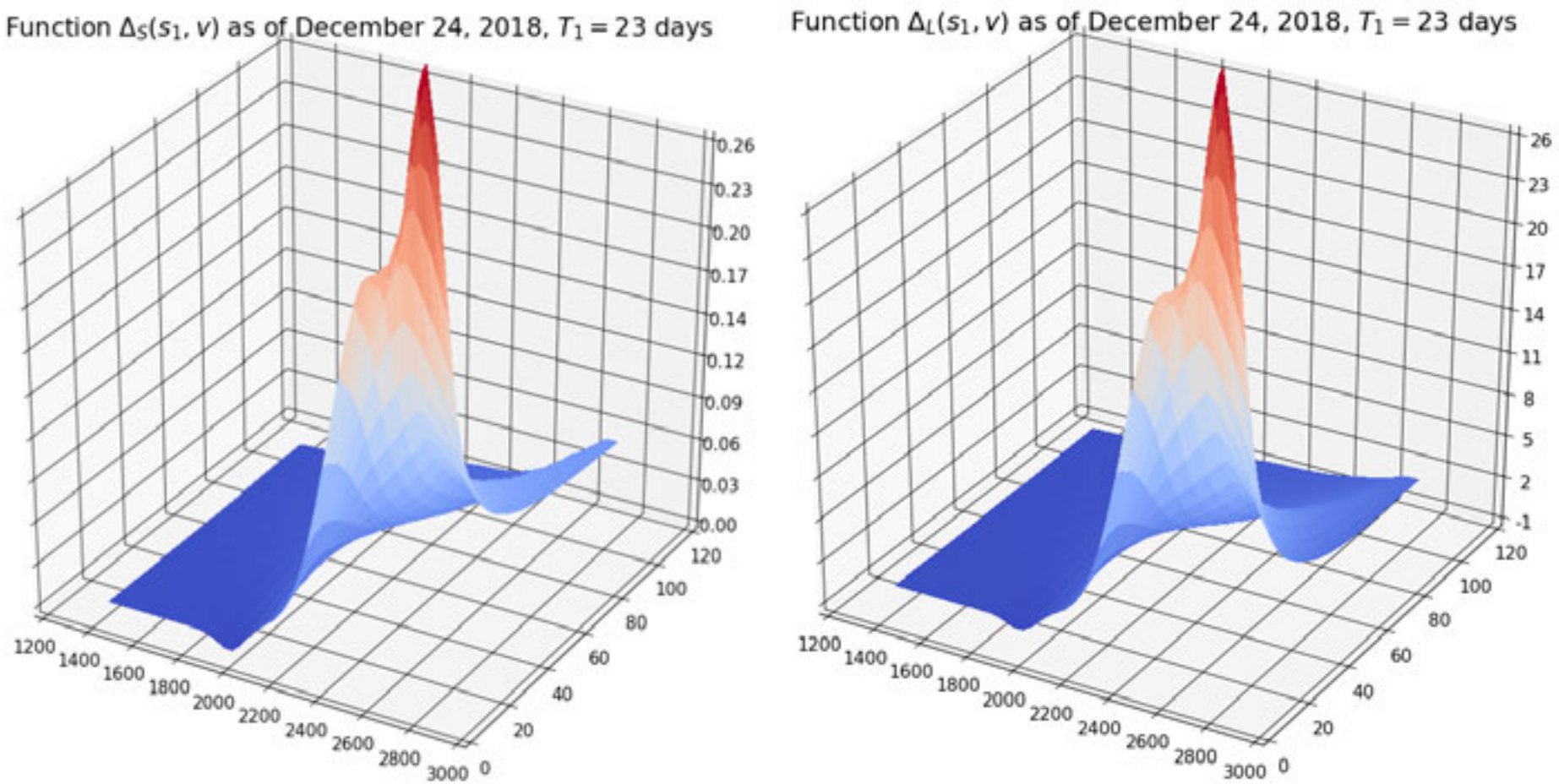


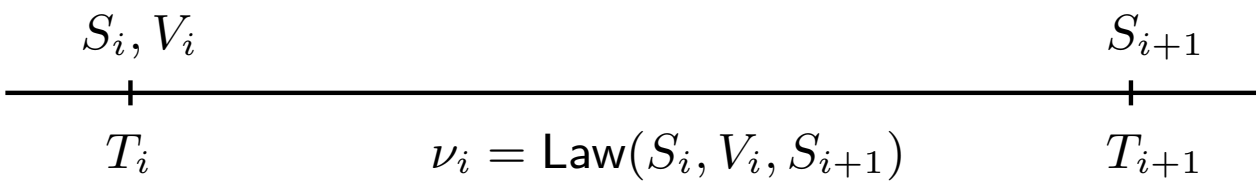
Figure: Optimal functions $\Delta_S^*(s_1, v)$ and $\Delta_L^*(s_1, v)$ for (s_1, v) in the quadrature grid

© 2020 Bloomberg Finance L.P. All rights reserved.

Extension to the multi-maturity case

© 2020 Bloomberg Finance L.P. All rights reserved.

Extension to the multi-maturity case



- Assume for simplicity that monthly SPX options and VIX futures maturities T_i perfectly coincide: $T_{i+1} - T_i = \tau$ for all $i \geq 1$.
- For each i we build a jointly calibrating model ν_i using the Sinkhorn algorithm.
- **Glue the models together**: Build a calibrated model on $(S_i, V_i)_{i \geq 1}$ as follows: $(S_1, V_1, S_2) \sim \nu_1$; recursively define the distribution of (V_{i+1}, S_{i+2}) given $(S_1, V_1, S_2, V_2, \ldots, S_i, V_i, S_{i+1})$ as the conditional distribution of (V_{i+1}, S_{i+2}) given S_{i+1} under ν_{i+1}.
- The resulting model ν is arbitrage-free, consistent, and calibrated to all the SPX and VIX monthly market smiles μ_{S_i} and μ_{V_i}: for all $i \geq 1$,

$$S_i \sim \mu_{S_i}, \quad V_i \sim \mu_{V_i}, \quad \mathbb{E}^\nu \left[S_{i+1} | (S_j, V_j)_{1 \leq j \leq i} \right] = S_i, \quad \mathbb{E}^\nu \left[L \left(\frac{S_{i+1}}{S_i} \right) \middle| (S_j, V_j)_{1 \leq j \leq i} \right] = V_i^2.$$

© 2020 Bloomberg Finance L.P. All rights reserved.

Continuous time

© 2020 Bloomberg Finance L.P. All rights reserved.

Continuous time

- Same technique: Pick a reference measure $\mathbb{P}_0 \longleftrightarrow$ a particular SV model:

$$\begin{aligned} \frac{dS_t}{S_t} &= a_t \, dW_t^0 \\ da_t &= b(a_t)\, dt + \sigma(a_t) \left(\rho \, dW_t^0 + \sqrt{1-\rho^2} dW_t^{0,\perp} \right) \end{aligned}$$

- We want to prove that $\mathcal{P} \neq \emptyset$ and build $\mathbb{P} \in \mathcal{P}$, where

$$\mathcal{P} := \{ \mathbb{P} \ll \mathbb{P}_0 | S_1 \sim \mu_1, S_2 \sim \mu_2, \sqrt{\mathbb{E}^{\mathbb{P}}[L(S_2/S_1)|\mathcal{F}_1]} \sim \mu_V, S \text{ is a } \mathbb{P}\text{-martingale} \}.$$

- We look for $\mathbb{P} \in \mathcal{P}$ that minimizes the relative entropy w.r.t. $\mathbb{P}_0$:

$$D := \inf_{\mathbb{P} \in \mathcal{P}} H(\mathbb{P}, \mathbb{P}_0)$$

- Inspired by Henry-Labordère (2019): *From (Martingale) Schrödinger Bridges to a New Class of Stochastic Volatility Models*.

© 2020 Bloomberg Finance L.P. All rights reserved.

Continuous time

$$\begin{aligned}\frac{dS_t}{S_t} &= a_t\, dW_t^* \\ da_t &= \left(b(a_t)+(1-\rho^2)\sigma(a_t)^2\partial_a u^*(t,S_t,a_t)\right)dt + \sigma(a_t)\left(\rho\, dW_t^* + \sqrt{1-\rho^2}dW_t^{*,\perp}\right)\end{aligned}$$

- Let $P := \sup_{u_1,u_V,u_2}\left\{\sum_{i\in\{1,2,V\}}(\mu_i,u_i) - u(0,S_0,a_0)\right\}$ where u is solution to a nonlinear Hamilton-Jacobi-Bellman PDE:

$$\begin{aligned} u(T_2,s,a;\delta^L) &= u_2(s)+\delta^L L(s), \\ \partial_t u + \mathcal{L}^0 u + \frac{1}{2}(1-\rho^2)\sigma(a)^2(\partial_a u)^2 &= 0, \quad t\in(T_1,T_2), \\ \Phi(s,a) := \sup_{v\geq 0}\inf_{\delta^L\in\mathbb{R}}\Big\{u_V(v) &- \delta^L(L(s)+v^2) + u(T_1,s,a;\delta^L)\Big\}, \\ u(T_1,s,a) &= u_1(s)+\Phi(s,a), \\ \partial_t u + \mathcal{L}^0 u + \frac{1}{2}(1-\rho^2)\sigma(a)^2(\partial_a u)^2 &= 0, \quad t\in[0,T_1). \end{aligned}$$

- Assume $P<+\infty$ and (u_1^*,u_V^*,u_2^*) maximizes $P \longrightarrow u^*$

© 2020 Bloomberg Finance L.P. All rights reserved.

Continuous time

$$\begin{aligned}\frac{dS_t}{S_t} &= a_t\, dW_t^* \\ da_t &= \left(b(a_t)+(1-\rho^2)\sigma(a_t)^2\partial_a u^*(t,S_t,a_t)\right)dt + \sigma(a_t)\left(\rho\, dW_t^* + \sqrt{1-\rho^2}dW_t^{*,\perp}\right)\end{aligned}$$

- Optimal deltas:

$$\Delta_t^* = -\partial_s u^*(t,S_t,a_t) - \rho\frac{\sigma(a_t)}{a_t S_t}\partial_a u^*(t,S_t,a_t); \quad \Delta^{*,L} = \delta^{*,L}(S_1,a_1)$$

- The probability $\mathbb{P}^* \in \mathcal{P}$ that minimizes $H(\mathbb{P},\mathbb{P}_0)$ satisfies

$$\frac{d\mathbb{P}^*}{d\mathbb{P}_0} = Z e^{u_1^*(S_1)+u_2^*(S_2)+u_V^*(V^*)+\int_0^{T_2}\Delta_t^* dS_t + \Delta^{*,L}\left(L\left(\frac{S_2}{S_1}\right)-(V^*)^2\right)}.$$

- The drift of (a_t) under $\mathbb{P}^*$ also reads as

$$b(a_t) + (1-\rho^2)\sigma(a_t)^2\partial_a \ln \mathbb{E}^0[e^{u_1^*(S_1)+\int_t^{T_1}\Delta^*(r,S_r,a_r)dS_r+\Phi^*(S_1,a_1)}|S_t,a_t], \quad t\in[0,T_1],$$

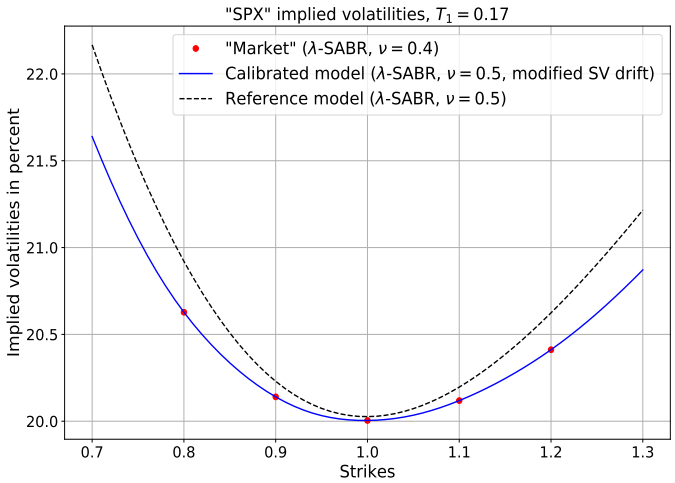
$$b(a_t) + (1-\rho^2)\sigma(a_t)^2\partial_a \ln \mathbb{E}^0[e^{u_2^*(S_2)+\int_t^{T_2}\Delta^*(r,S_r,a_r)dS_r+\delta^{*,L}(S_1,a_1)L(S_2)}|S_t,a_t], \quad t\in[T_1,T_2].$$

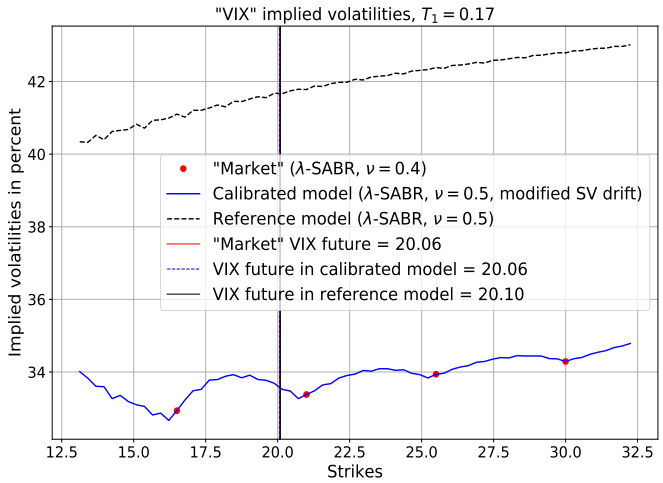
- It is path-dependent on $[T_1,T_2]$ so as to match the market VIX smile.
- If $P=+\infty$, then $\mathcal{P}=\emptyset$.

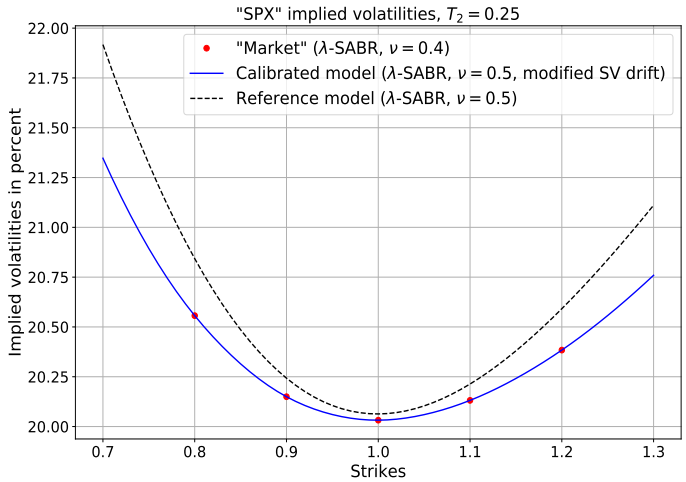
© 2020 Bloomberg Finance L.P. All rights reserved.

$da_t = -k(a_t - m)\,dt + \nu a_t\,dZ_t$. 'Market': $\nu = 0.4$, $\mathbb{P}_0 : \nu = 0.5$

$$k = 1.5, \quad a_0 = m = 0.2, \quad \rho = 0$$

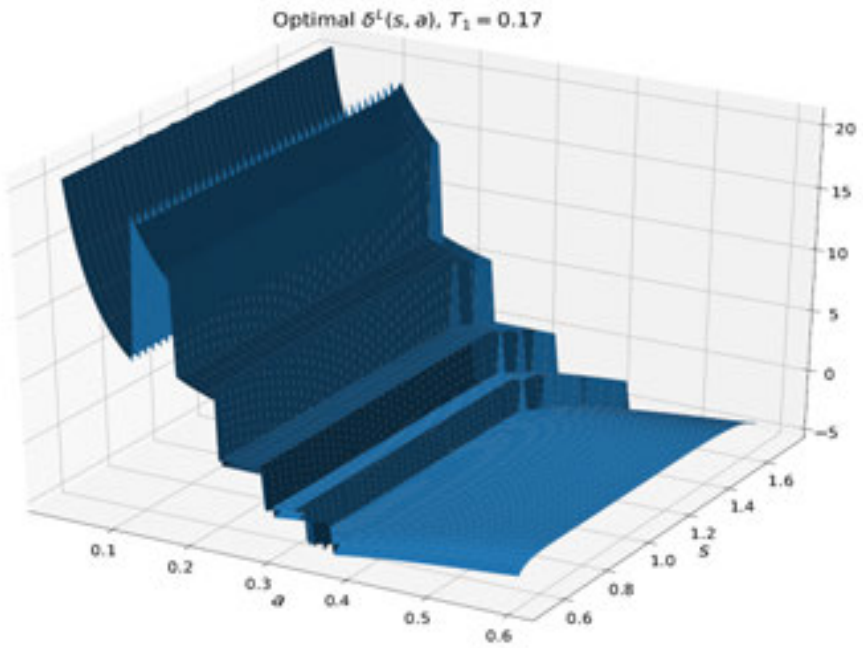


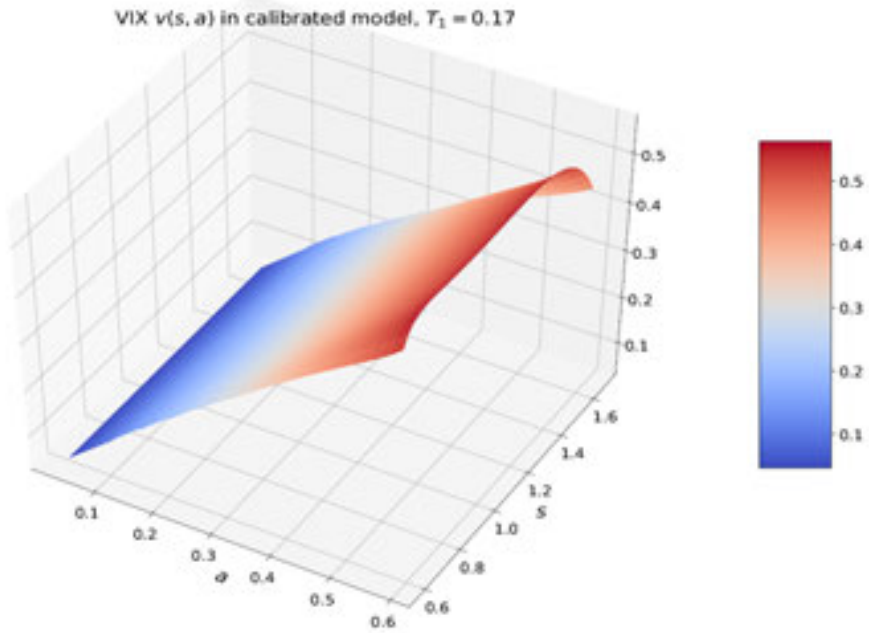




© 2020 Bloomberg Finance L.P. All rights reserved.

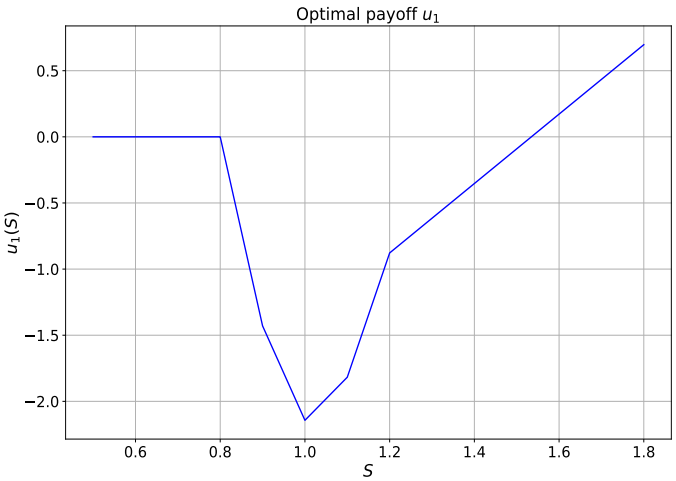
$da_t = -k(a_t - m)\,dt + \nu a_t\,dZ_t$. 'Market': $\nu = 0.4$, $\mathbb{P}_0 : \nu = 0.5$

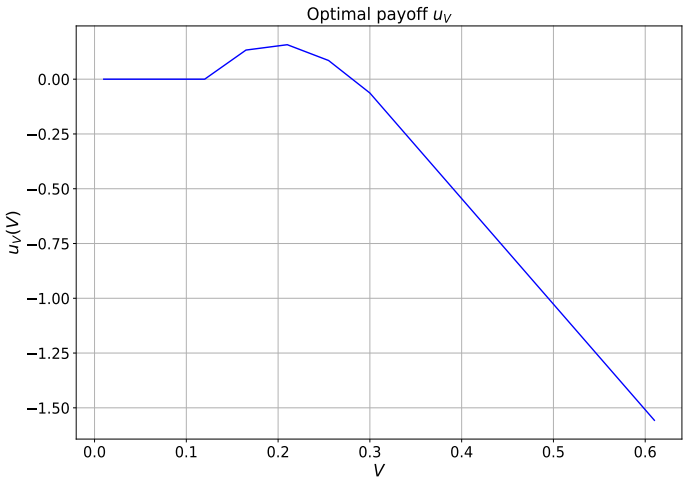


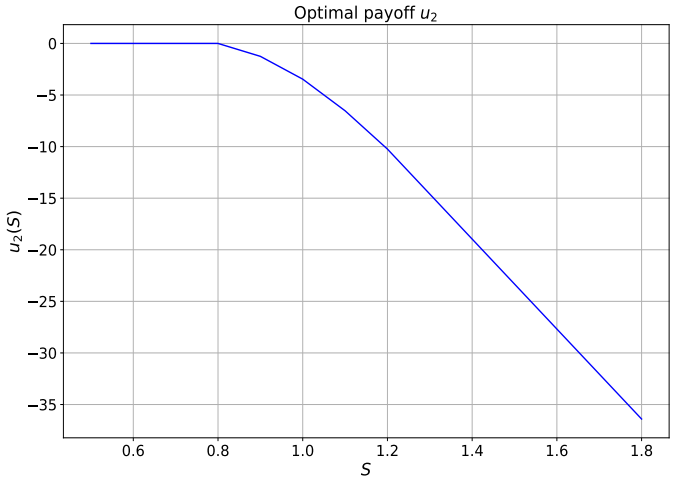


© 2020 Bloomberg Finance L.P. All rights reserved.

$da_t = -k(a_t - m)\,dt + \nu a_t\,dZ_t$. 'Market': $\nu = 0.4$, $\mathbb{P}_0 : \nu = 0.5$

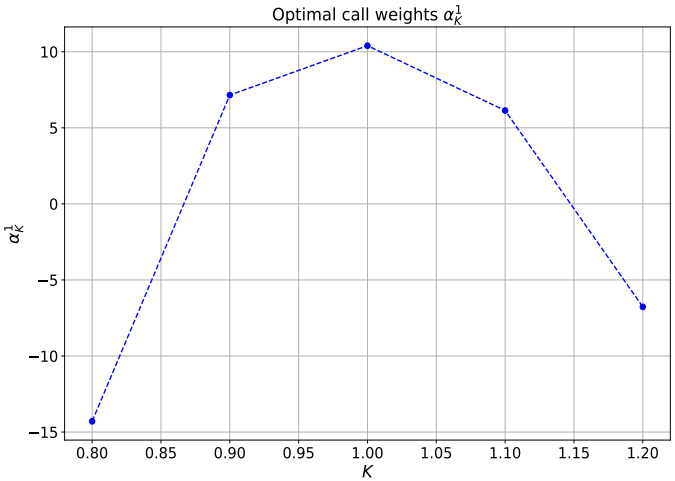


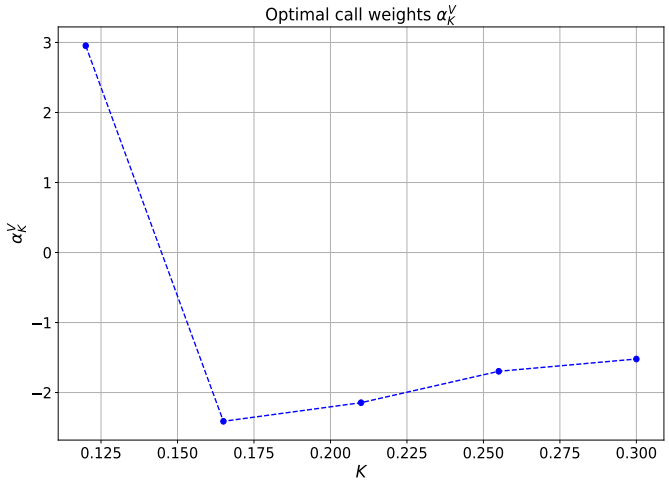


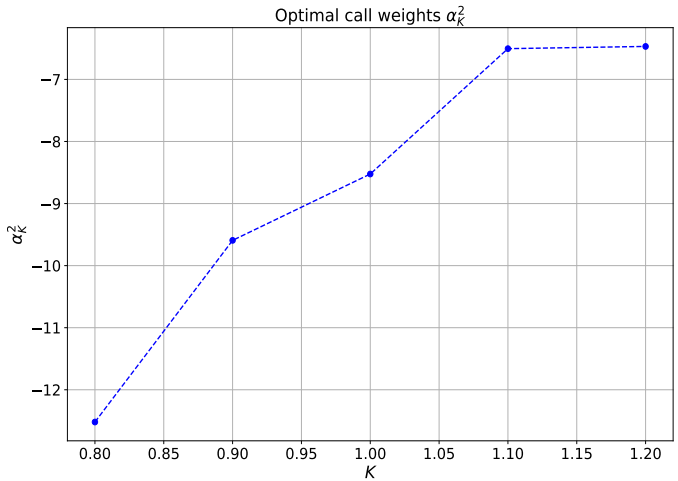


© 2020 Bloomberg Finance L.P. All rights reserved.

$da_t = -k(a_t - m)\,dt + \nu a_t\,dZ_t$. 'Market': $\nu = 0.4$, $\mathbb{P}_0 : \nu = 0.5$







© 2020 Bloomberg Finance L.P. All rights reserved.

Other approaches

- Cont-Kokholm (2013): Bergomi-like model with simultaneous jumps on SPX and VIX.
  - Best fit
  - An approximation of the VIX in the model is used
- Gatheral-Jusselin-Rosenbaum (2020): quadratic rough Heston volatility model.
  - Best fit
  - VIX smile well calibrated, not enough ATM SPX skew
- Guo-Loeper-Obłoj-Wang (2020): joint calibration via optimal transport
  - More general cost function: volatilities are allowed to be modified from reference model
  - Model (S_t, Y_t) instead of (S_t, a_t) where Y is the price of the log payoff: (S_{T_2}, Y_{T_2}) live on a graph
  - Numerical tests: calibration to SPX smile and VIX future
- Fouque-Saporito (2018): Heston Stochastic Vol-of-Vol Model
- Pacati-Pompa-Renò (2018): displacement of multi-factor affine models with jumps (Heston++)
- Papanicolaou-Sircar (2014), Goutte-Amine-Pham (2017): regime-switching Heston model (with or without jumps)

© 2020 Bloomberg Finance L.P. All rights reserved.

Inversion of convex ordering in the VIX market

© 2020 Bloomberg Finance L.P. All rights reserved.

Quantitative Finance (accepted April 3, 2020)

Quantitative Finance, 2020
https://doi.org/10.1080/14697688.2020.1753885

Routledge
Taylor & Francis Group
Check for updates

Inversion of convex ordering in the VIX market

JULIEN GUYON*

Quantitative Research, Bloomberg L.P., 731 Lexington Ave, New York, NY 10022, USA

(*Received 14 December 2019; accepted 3 April 2020*)

We investigate conditions for the existence of a *continuous* model on the S&P 500 index (SPX) that jointly calibrates to a *full surface* of SPX implied volatilities and to the VIX smiles. We present a novel approach based on the SPX smile calibration condition $\mathbb{E}[\sigma_t^2|S_t] = \sigma_{\text{lv}}^2(t, S_t)$. In the limiting case of instantaneous VIX, a novel application of martingale transport to finance shows that such model exists if and only if, for each time t, the local variance $\sigma_{\text{lv}}^2(t, S_t)$ is smaller than the instantaneous variance σ_t^2 in convex order. The real case of a 30-day VIX is more involved, as averaging over 30 days and projecting onto a filtration can undo convex ordering.

We show that in usual market conditions, and for reasonable smile extrapolations, the distribution of VIX_T^2 in the market local volatility model is *larger* than the market-implied distribution of VIX_T^2 in convex order for short maturities T, and that the two distributions are not rankable in convex order for intermediate maturities. In particular, a *necessary* condition for continuous models to jointly calibrate to the SPX and VIX markets is the *inversion of convex ordering* property: the fact that, even though associated local variances are smaller than instantaneous variances in convex order, the VIX

© 2020 Bloomberg Finance L.P. All rights reserved.

Continuous model on SPX calibrated to SPX options

$$\frac{dS_t}{S_t} = \sigma_t \, dW_t, \qquad S_0 = x. \tag{8.1}$$

- Corresponding local volatility function σ_{loc}: $\sigma^2_{\mathrm{loc}}(t, S_t) := \mathbb{E}[\sigma_t^2 | S_t]$.
- Corresponding local volatility model:

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

- From Gyöngy (1986): $\forall t \geq 0, \quad S_t^{\mathrm{loc}} \overset{(d)}{=} S_t$.
- Using Dupire (1994), we conclude that Model (8.1) is calibrated to the full SPX smile if and only if $\sigma_{\mathrm{loc}} = \sigma_{\mathrm{lv}}$ (market local volatility computed using Dupire's formula).
- Market local volatility model:

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

© 2020 Bloomberg Finance L.P. All rights reserved.

VIX

- By definition, the (idealized) VIX at time $T \geq 0$ is the implied volatility of a 30 day log-contract on the SPX index starting at T. For continuous models (8.1), this translates into

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

- 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,

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

© 2020 Bloomberg Finance L.P. All rights reserved.

Reminder on convex order

- (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** (dimension 1).

© 2020 Bloomberg Finance L.P. All rights reserved.

The case of instantaneous VIX: $\tau \to 0$

- Assume SV model is calibrated to the SPX smile: $\mathbb{E}[\sigma_t^2 | S_t] = \sigma_{\text{lv}}^2(t, S_t)$.
- As observed by Dupire (2005), by conditional Jensen, $\sigma_{\text{lv}}^2(t, S_t) \leq_c \sigma_t^2$, i.e.,

$$\text{mkt local var}_t \leq_c \text{instVIX}_t^2.$$

- Conversely, if mkt local $\text{var}_t \leq_c \text{instVIX}_t^2$, there exists a jointly calibrating SPX/instVIX model (G., 2017).
- $\Longrightarrow$ **Convex order condition is necessary and sufficient for instVIX**.
- Proof uses a **new type of application of martingale transport to finance**: martingality constraint applies to (mkt local var_t, instVIX_t^2) at a single date, instead of (S_1, S_2).

© 2020 Bloomberg Finance L.P. All rights reserved.

The real VIX: $\tau = 30$ days

- In reality, squared VIX are not instantaneous variances but the **fair strikes** of **30-day** realized variances.
- Let us look at market data (August 1, 2018). We compare the market distributions of

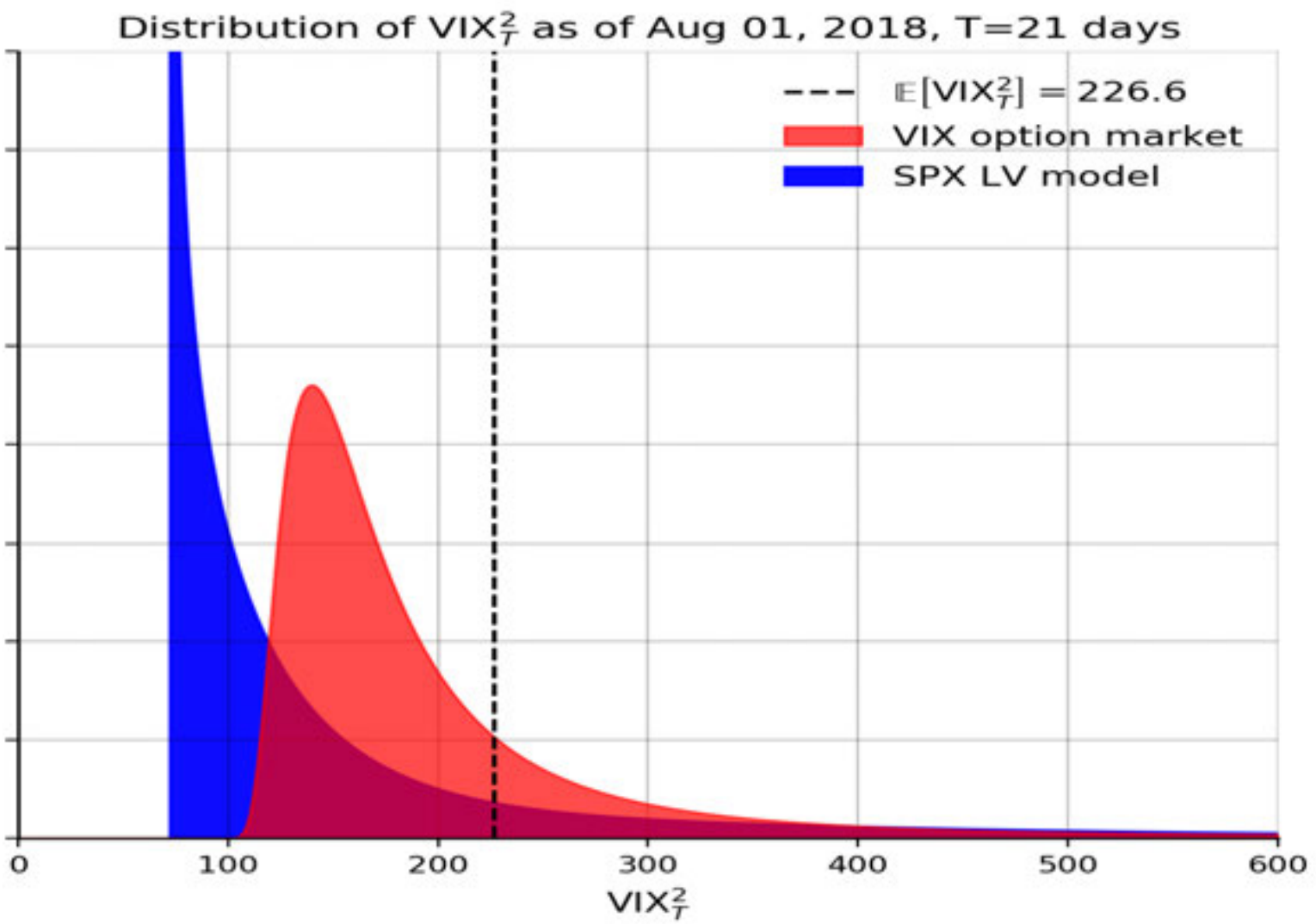
$$\mathrm{VIX}^2_{\mathrm{lv},T} := \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma^2_{\mathrm{lv}}(t, S^{\mathrm{lv}}_t)\, dt \middle| S^{\mathrm{lv}}_T\right]$$

and

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

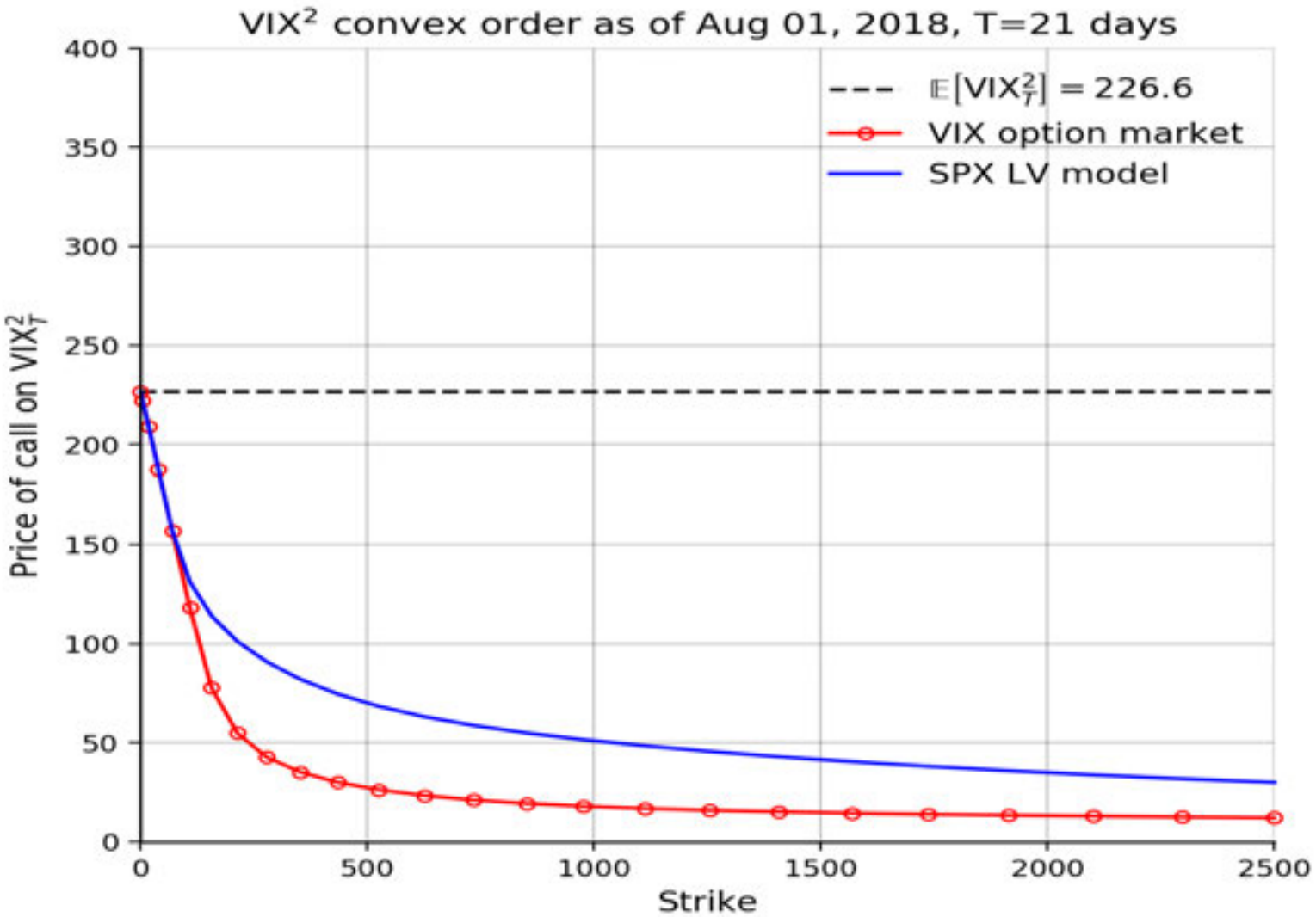
© 2020 Bloomberg Finance L.P. All rights reserved.

$T = 21$ days



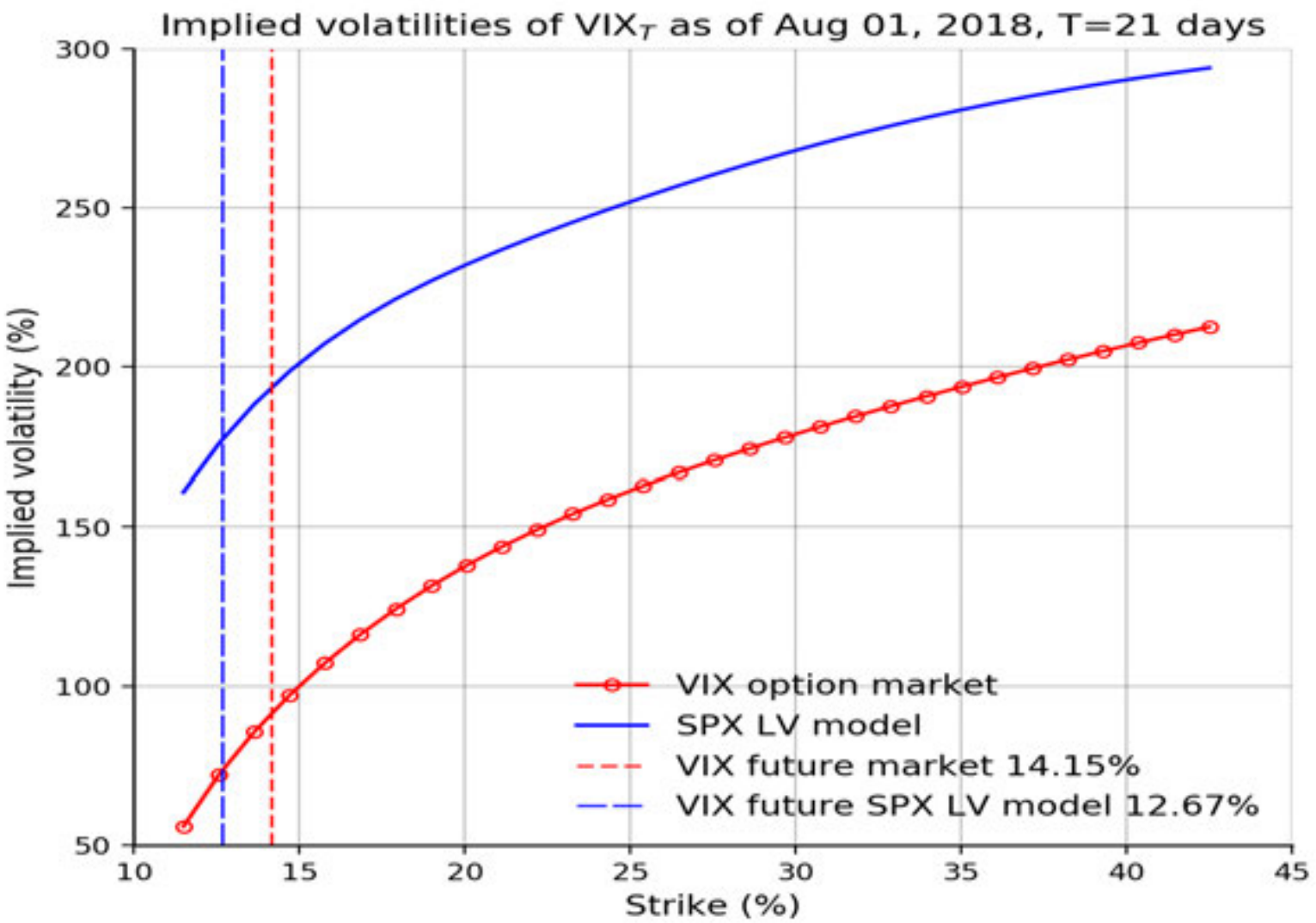
© 2020 Bloomberg Finance L.P. All rights reserved.

$T = 21$ days



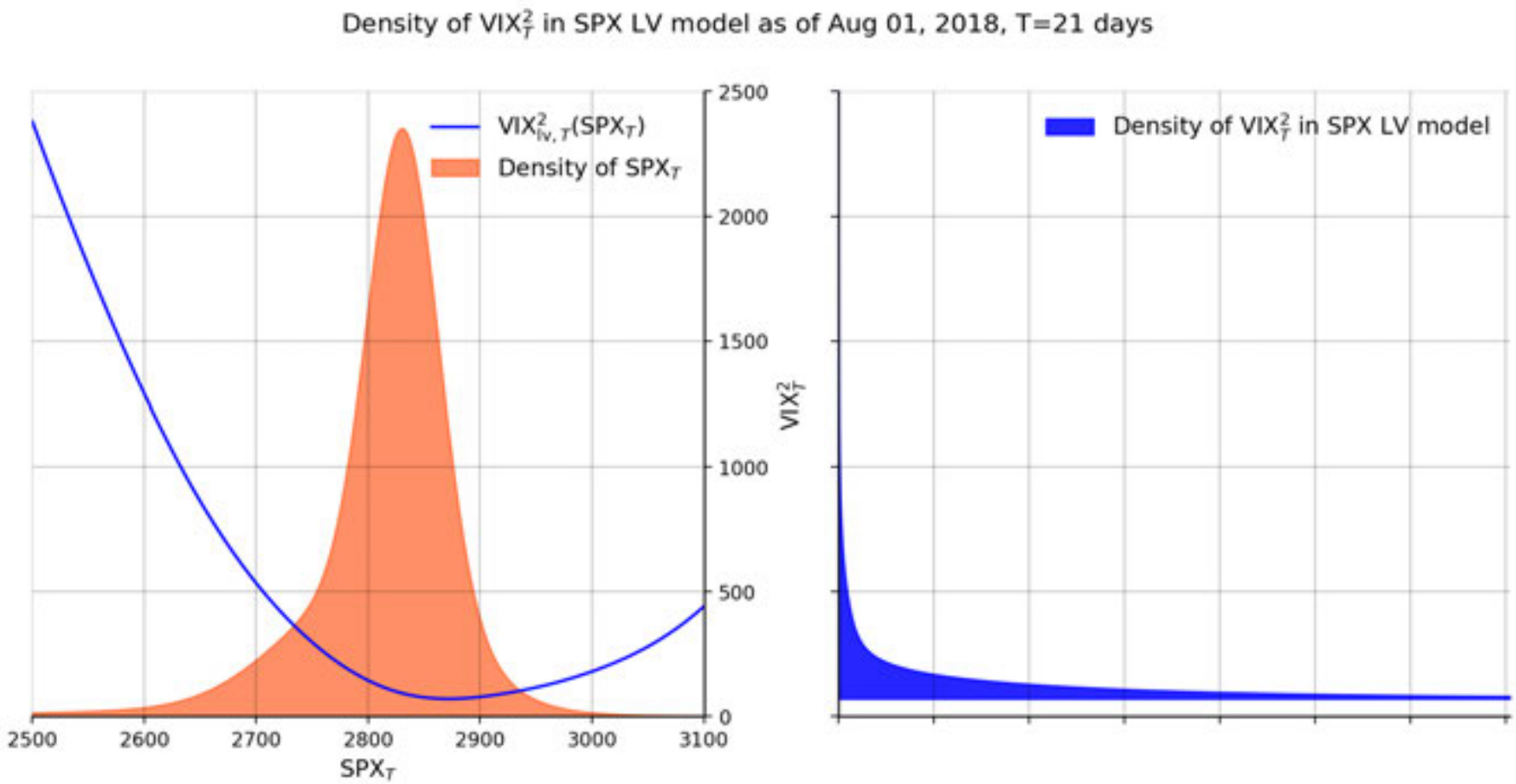
© 2020 Bloomberg Finance L.P. All rights reserved.

$T = 21$ days



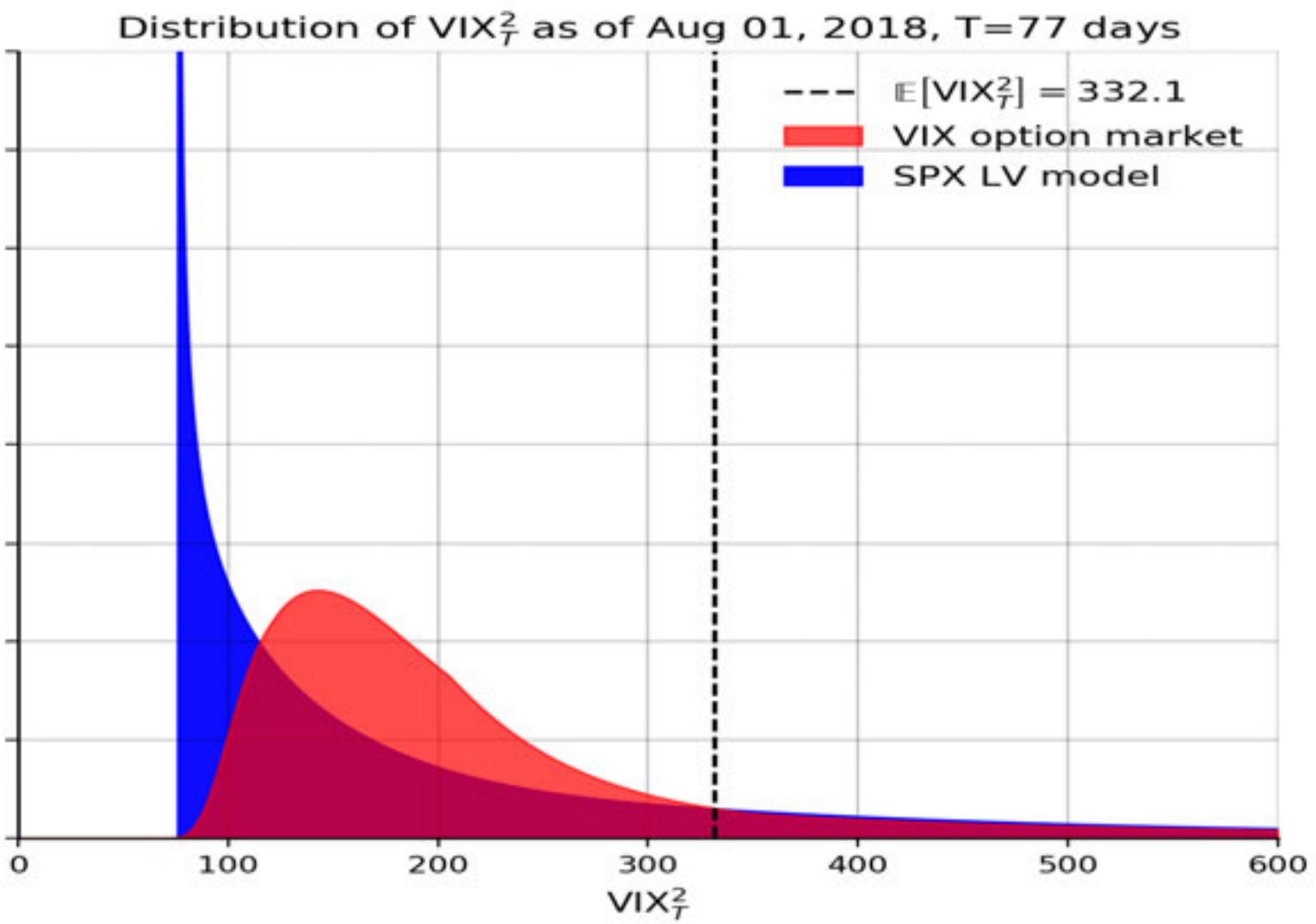
© 2020 Bloomberg Finance L.P. All rights reserved.

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



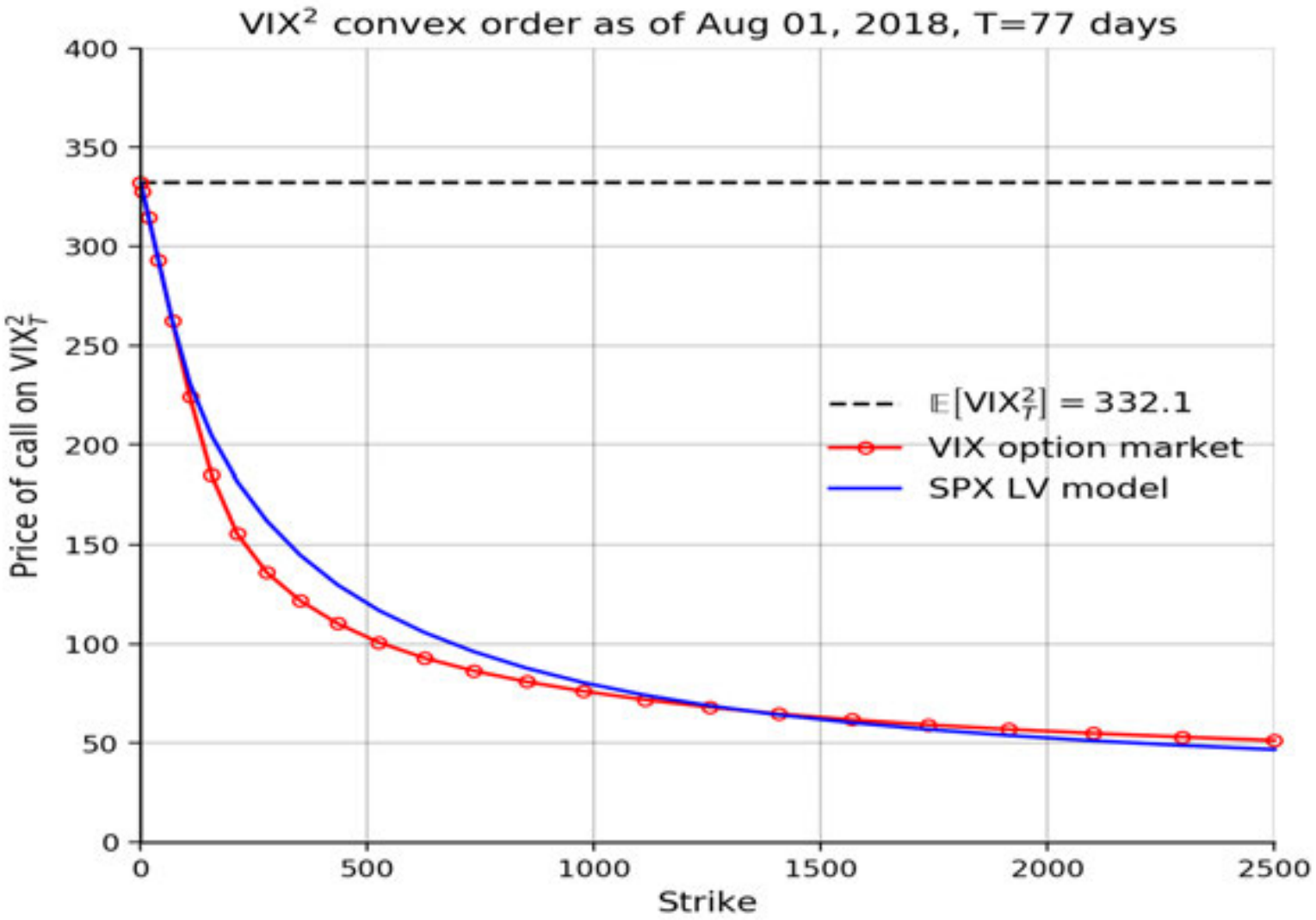
© 2020 Bloomberg Finance L.P. All rights reserved.

$T = 77$ days



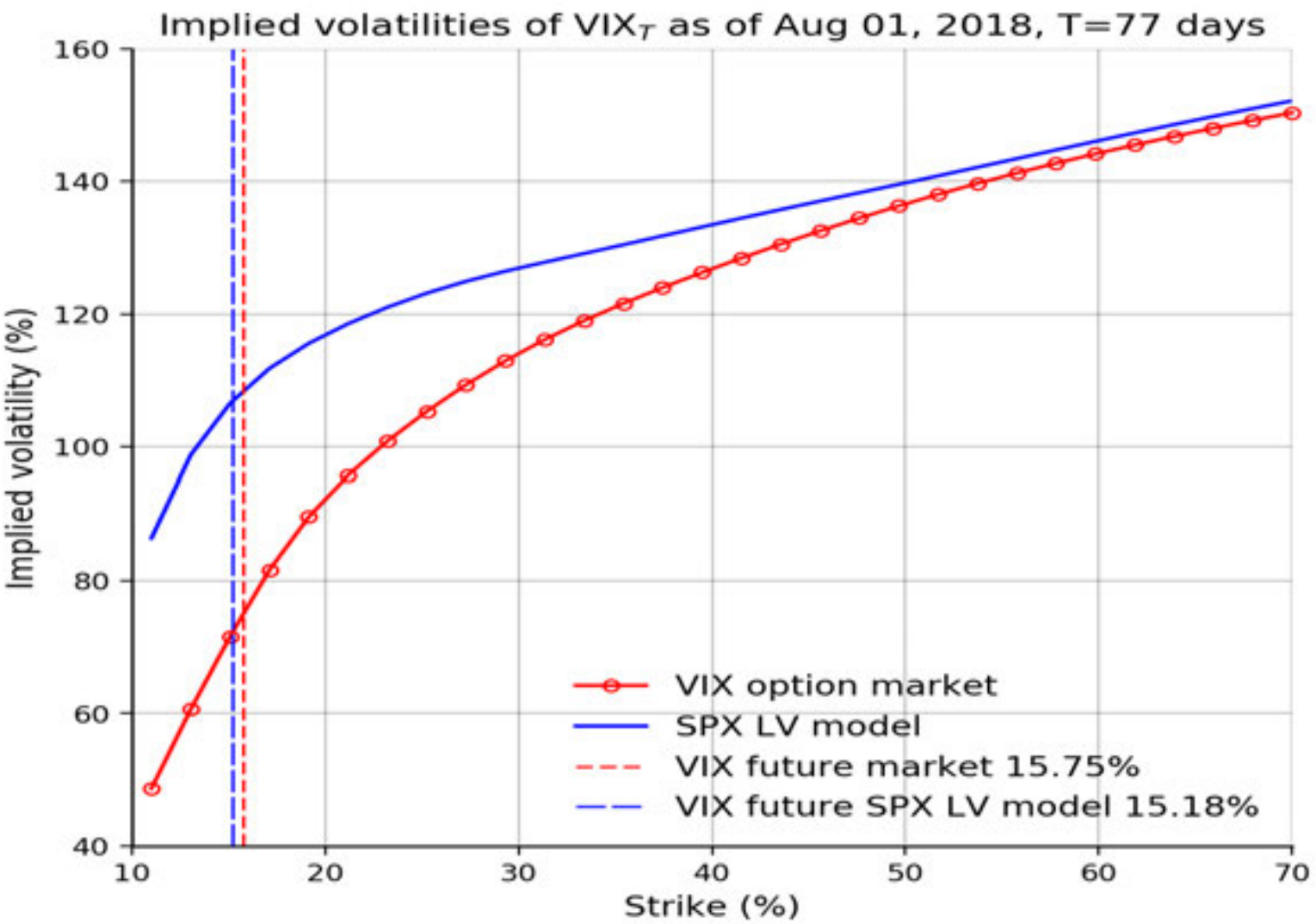
© 2020 Bloomberg Finance L.P. All rights reserved.

$T = 77$ days



© 2020 Bloomberg Finance L.P. All rights reserved.

$T = 77$ days



© 2020 Bloomberg Finance L.P. All rights reserved.

Inversion of convex ordering

- **Inversion of convex ordering**: the fact that, for small T, $\mathrm{VIX}^2_{\mathrm{loc},T} \geq_c \mathrm{VIX}^2_T$ **despite the fact that for all** t, $\sigma^2_{\mathrm{loc}}(t, S_t) \leq_c \sigma^2_t$.
- A **necessary** condition for continuous models to jointly calibrate to the SPX and VIX markets.
- In the paper, we numerically show that when the spot-vol correlation is large enough in absolute value,
  (a) traditional SV models with **large mean reversion**, and
  (b) rough volatility models with **small Hurst exponent**

  satisfy the inversion of convex ordering property, and more generally can reproduce the market term-structure of convex ordering of the local and stochastic squared VIX.
- Not a sufficient condition though.
- Actually we have proved that **inversion of convex ordering can be produced by a continuous SV model**.
- In such models, for small T, $\mathrm{VIX}^2_{\mathrm{loc},T} >_c \mathrm{VIX}^2_T$ so ($x \mapsto \sqrt{x}$ concave)

$$\mathbb{E}[\mathrm{VIX}_T] > \mathbb{E}\left[\mathrm{VIX}_{\mathrm{loc},T}\right] :$$

Local volatility does NOT maximize the price of VIX futures.

© 2020 Bloomberg Finance L.P. All rights reserved.

SIAM J. Financial Math. 11(1):SC1–SC13, 2020 (with B. Acciaio)

SIAM J. FINANCIAL MATH.
Vol. 11, No. 1, pp. SC1–SC13

© 2020 Society for Industrial and Applied Mathematics

Short Communication: Inversion of Convex Ordering: Local Volatility Does Not Maximize the Price of VIX Futures*

Beatrice Acciaio† and Julien Guyon‡

Abstract. It has often been stated that, within the class of continuous stochastic volatility models calibrated to vanillas, the price of a VIX future is maximized by the Dupire local volatility model. In this article we prove that this statement is incorrect: we build a continuous stochastic volatility model in which a VIX future is *strictly more expensive* than in its associated local volatility model. More generally, in our model, strictly convex payoffs on a squared VIX are strictly cheaper than in the associated local volatility model. This corresponds to an *inversion of convex ordering* between local and stochastic variances, when moving from instantaneous variances to squared VIX, as convex payoffs on instantaneous variances are always cheaper in the local volatility model. We thus prove that this inversion of convex ordering, which is observed in the S&P 500 market for short VIX maturities, can be produced by a continuous stochastic volatility model. We also prove that the model can be extended so that, as suggested by market data, the convex ordering is preserved for long maturities.

Key words. VIX, VIX futures, stochastic volatility, local volatility, convex order, inversion of convex ordering

AMS subject classifications. 91G20, 91G80, 60H30

DOI. 10.1137/19M129303X

1. Introduction. 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 stochastic volatility models on the S&P 500 index (SPX) of the form

© 2020 Bloomberg Finance L.P. All rights reserved.

Acknowledgements

I would like to thank Bruno Dupire, Pierre Henry-Labordère, Stefano de Marco, and Bryan Liang for interesting discussions, and Bryan Liang for providing some graphs.

© 2020 Bloomberg Finance L.P. All rights reserved.

A few selected references

Acciaio, B., Guyon, J.: *Inversion of Convex Ordering: Local Volatility Does Not Maximize the Price of VIX Futures*, SIAM J. Financial Math. 11(1):SC1–SC13, 2020.

Beiglböck, M., Henry-Labordère, P., Penkner, F.: *Model-independent bounds for option prices: A mass-transport approach*, Finance Stoch., 17(3):477–501, 2013.

Blaschke, W, Pick, G.: *Distanzschätzungen im Funktionenraum II*, Math. Ann. 77:277–302, 1916.

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.

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

Cuturi, M.: *Sinkhorn distances: Lightspeed computation of optimal transport*, Advances in neural information processing systems, 2292–2300, 2013.

De March, A.: *Entropic approximation for multi-dimensional martingale optimal transport*, preprint, arXiv, 2018.

© 2020 Bloomberg Finance L.P. All rights reserved.

A few selected references

De March, A., Henry-Labordère, P.: *Building arbitrage-free implied volatility: Sinkhorn's algorithm and variants* , preprint, SSRN, 2019.

De Marco, S., Henry-Labordere, P.: *Linking vanillas and VIX options: A constrained martingale optimal transport problem*, SIAM J. Financial Math. 6:1171–1194, 2015.

Dupire, B.: *Pricing with a smile*, Risk, January, 1994.

Fouque, J.-P., Saporito, Y.: *Heston Stochastic Vol-of-Vol Model for Joint Calibration of VIX and S&P 500 Options*, Quantitative Finance 18(6):1003–1016, 2018.

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

Gatheral, J., Jusselin, P., Rosenbaum, M.: *The quadratic rough Heston model and the joint calibration problem*, preprint, 2020.

Goutte, S, Amine, I., and Pham, H.: *Regime-switching stochastic volatility model: estimation and calibration to VIX options*, Applied Math. Finance 24(1):38–75, 2017.

© 2020 Bloomberg Finance L.P. All rights reserved.

A few selected references

Guo, I., Loeper, G., Obłłoj, J., Wang, S.: *Joint Modelling and Calibration of SPX and VIX by Optimal Transport*, preprint, 2020.

Guyon, J., Menegaux, R., Nutz, M.: *Bounds for VIX futures given S&P 500 smiles*, Finan. & Stoch. 21(3):593–630, 2017.

Guyon, J.: *On the joint calibration of SPX and VIX options*, Conference in honor of Jim Gatheral's 60th birthday, NYU Courant, 2017. And Finance and Stochastics seminar, Imperial College London, 2018.

Guyon, J.: *The Joint S&P 500/VIX Smile Calibration Puzzle Solved*, Risk, April 2020.

Guyon, J.: *Inversion of convex ordering in the VIX market*, accepted in Quantitative Finance, 2019.

Henry-Labordère, P.: *Automated Option Pricing: Numerical Methods*, Intern. Journ. of Theor. and Appl. Finance 16(8), 2013.

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

© 2020 Bloomberg Finance L.P. All rights reserved.

A few selected references

Jacquier, A., Martini, C., Muguruza, A.: *On the VIX futures in the rough Bergomi model*, Quantitative Finance 18(1):45–61, 2018.

Johansen, S.: The extremal convex functions. Math. Scand., 34, 61–68 (1974)

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.

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.

Pacati, C., Pompa, P., Renò, R.: *Smiling Twice: The Heston++ Model*, Journal of Banking & Finance, Elsevier, 96(C):185–206, 2018.

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

Sinkhorn, R.: *Diagonal equivalence to matrices with prescribed row and column sums*, The American Mathematical Monthly, 74(4):402–405, 1967.

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

© 2020 Bloomberg Finance L.P. All rights reserved.

Equations for $u^* = (u_1^*, u_V^*, u_2^*, \Delta_S^*, \Delta_L^*)$

$$
\begin{aligned}
\tfrac{\partial \Psi_{\bar\mu}}{\partial u_1(s_1)} = 0: \quad & \forall s_1 > 0, & u_1(s_1) &= \Phi_1(s_1; u_V, u_2, \Delta_S, \Delta_L) \\
\tfrac{\partial \Psi_{\bar\mu}}{\partial u_V(v)} = 0: \quad & \forall v \geq 0, & u_V(v) &= \Phi_V(v; u_1, u_2, \Delta_S, \Delta_L) \\
\tfrac{\partial \Psi_{\bar\mu}}{\partial u_2(s_2)} = 0: \quad & \forall s_2 > 0, & u_2(s_2) &= \Phi_2(s_2; u_1, u_V, \Delta_S, \Delta_L) \\
\tfrac{\partial \Psi_{\bar\mu}}{\partial \Delta_S(s_1,v)} = 0: \quad & \forall s_1 > 0,\ \forall v \geq 0, & 0 &= \Phi_{\Delta_S}(s_1, v; \Delta_S(s_1,v), \Delta_L(s_1,v)) \\
\tfrac{\partial \Psi_{\bar\mu}}{\partial \Delta_L(s_1,v)} = 0: \quad & \forall s_1 > 0,\ \forall v \geq 0, & 0 &= \Phi_{\Delta_L}(s_1, v; \Delta_S(s_1,v), \Delta_L(s_1,v))
\end{aligned}
$$

$$
\begin{aligned}
\Phi_1(s_1; u_V, \Delta_S, \Delta_L) &:= \ln \mu_1(s_1) - \ln \left(\int \bar\mu(s_1, dv, ds_2) e^{u_V(v) + u_2(s_2) + \Delta_S^{(S)}(s_1,v,s_2) + \Delta_L^{(L)}(s_1,v,s_2)} \right) \\
\Phi_V(v; u_1, \Delta_S, \Delta_L) &:= \ln \mu_V(v) - \ln \left(\int \bar\mu(ds_1, v, ds_2) e^{u_1(s_1) + u_2(s_2) + \Delta_S^{(S)}(s_1,v,s_2) + \Delta_L^{(L)}(s_1,v,s_2)} \right) \\
\Phi_2(s_2; u_1, u_V, \Delta_S, \Delta_L) &:= \ln \mu_2(s_2) - \ln \left(\int \bar\mu(ds_1, dv, s_2) e^{u_1(s_1) + u_V(v) + \Delta_S^{(S)}(s_1,v,s_2) + \Delta_L^{(L)}(s_1,v,s_2)} \right) \\
\Phi_{\Delta_S}(s_1, v; u_2, \delta_S, \delta_L) &:= \int \bar\mu(s_1, v, ds_2)(s_2 - s_1) e^{u_2(s_2) + \delta_S(s_2 - s_1) + \delta_L \left(L\left(\frac{s_2}{s_1}\right) - v^2 \right)} \\
\Phi_{\Delta_L}(s_1, v; u_2, \delta_S, \delta_L) &:= \int \bar\mu(s_1, v, ds_2) \left(L\left(\frac{s_2}{s_1}\right) - v^2 \right) e^{u_2(s_2) + \delta_S(s_2 - s_1) + \delta_L \left(L\left(\frac{s_2}{s_1}\right) - v^2 \right)}.
\end{aligned}
$$

© 2020 Bloomberg Finance L.P. All rights reserved.

Implementation details

Practically, we consider market strikes $\mathcal{K} := (\mathcal{K}_1, \mathcal{K}_V, \mathcal{K}_2)$ and market prices (C_K^1, C_K^V, C_K^2) of vanilla options on S_1, V, and S_2, and we build the model

$$\mu_{\mathcal{K}}^*(ds_1, dv, ds_2) = \bar{\mu}(ds_1, dv, ds_2) e^{c^* + \Delta_S^{0*} s_1 + \Delta_V^{0*} v + \sum_{K \in \mathcal{K}_1} a_K^{1*}(s_1 - K)_+}$$
$$e^{\sum_{K \in \mathcal{K}_V} a_K^{V*}(v-K)_+ + \sum_{K \in \mathcal{K}_2} a_K^{2*}(s_2 - K)_+ + \Delta_S^{*(S)}(s_1, v, s_2) + \Delta_L^{*(L)}(s_1, v, s_2)}$$

where $\theta^* := (c^*, \Delta_S^{0*}, \Delta_V^{0*}, a^{1*}, a^{V*}, a^{2*}, \Delta_S^*, \Delta_L^*)$ maximizes

$$\bar{\Psi}_{\bar{\mu}, \mathcal{K}}(\theta) := c + \Delta_S^0 S_0 + \Delta_V^0 F_V + \sum_{K \in \mathcal{K}_1} a_K^1 C_K^1 + \sum_{K \in \mathcal{K}_V} a_K^V C_K^V + \sum_{K \in \mathcal{K}_2} a_K^2 C_K^2$$
$$- \mathbb{E}^{\bar{\mu}} \left[e^{c + \Delta_S^0 S_1 + \Delta_V^0 V + \sum_{\mathcal{K}_1} a_K^1 (S_1 - K)_+ + \sum_{\mathcal{K}_V} a_K^V (V-K)_+ + \sum_{\mathcal{K}_2} a_K^2 (S_2 - K)_+ + \Delta_S^{(S)}(\ldots) + \Delta_L^{(L)}(\ldots)} \right]$$

over the set Θ of portfolios $\theta := (c, \Delta_S^0, \Delta_V^0, a^1, a^V, a^2, \Delta_S, \Delta_L)$ such that $c, \Delta_S^0, \Delta_V^0 \in \mathbb{R}$, $a^1 \in \mathbb{R}^{\mathcal{K}_1}$, $a^V \in \mathbb{R}^{\mathcal{K}_V}$, $a^2 \in \mathbb{R}^{\mathcal{K}_2}$, and $\Delta_S, \Delta_L : \mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ are measurable functions of (s_1, v).

© 2020 Bloomberg Finance L.P. All rights reserved.

Implementation details

- This corresponds to solving the entropy minimization problem

$$P_{\bar{\mu},\mathcal{K}} := \inf_{\mu \in \mathcal{P}(\mathcal{K})} H(\mu, \bar{\mu}) = \sup_{\theta \in \Theta} \bar{\Psi}_{\bar{\mu},\mathcal{K}}(\theta) =: D_{\bar{\mu},\mathcal{K}}$$

where $\mathcal{P}(\mathcal{K})$ denotes the set of probability measures μ on $\mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0}$ such that

$$\mathbb{E}^\mu[S_1] = S_0, \quad \mathbb{E}^\mu[V] = F_V, \quad \forall K \in \mathcal{K}_1,\ \mathbb{E}^\mu\left[(S_1 - K)_+\right] = C_K^1,$$
$$\forall K \in \mathcal{K}_V,\ \mathbb{E}^\mu\left[(V - K)_+\right] = C_K^V, \quad \forall K \in \mathcal{K}_2,\ \mathbb{E}^\mu\left[(S_2 - K)_+\right] = C_K^2,$$
$$\mathbb{E}^\mu\left[S_2 | S_1, V\right] = S_1, \quad \mathbb{E}^\mu\left[L\left(\frac{S_2}{S_1}\right)\middle| S_1, V\right] = V^2.$$

- One can directly check that model $\mu^*_{\mathcal{K}}$ is an arbitrage-free model that jointly calibrates the prices of SPX futures, options, VIX future, and VIX options. Indeed, if $\bar{\Psi}_{\bar{\mu},\mathcal{K}}$ reaches its maximum at θ^*, then θ^* is solution to $\frac{\partial \bar{\Psi}_{\bar{\mu},\mathcal{K}}}{\partial \theta_i}(\theta) = 0$:

© 2020 Bloomberg Finance L.P. All rights reserved.

Implementation details

$$\bar{\Psi}_{\bar{\mu},\mathcal{K}}(\theta) := c + \Delta_S^0 S_0 + \Delta_V^0 F_V + \sum_{K\in\mathcal{K}_1} a_K^1 C_K^1 + \sum_{K\in\mathcal{K}_V} a_K^V C_K^V + \sum_{K\in\mathcal{K}_2} a_K^2 C_K^2$$

$$-\mathbb{E}^{\bar{\mu}}\left[e^{c+\Delta_S^0 S_1+\Delta_V^0 V+\sum_{\mathcal{K}_1} a_K^1 (S_1-K)_+ +\sum_{\mathcal{K}_V} a_K^V (V-K)_+ +\sum_{\mathcal{K}_2} a_K^2 (S_2-K)_+ +\Delta_S^{(S)}(\ldots)+\Delta_L^{(L)}(\ldots)}\right]$$

$$\frac{\partial\bar{\Psi}_{\bar{\mu},\mathcal{K}}}{\partial c} = 0 : \mathbb{E}^{\bar{\mu}}\left[\frac{d\mu_{\mathcal{K}}^*}{d\bar{\mu}}\right] = 1 \qquad \frac{\partial\bar{\Psi}_{\bar{\mu},\mathcal{K}}}{\partial \Delta_S^0} = 0 : \mathbb{E}^{\bar{\mu}}\left[S_1\frac{d\mu_{\mathcal{K}}^*}{d\bar{\mu}}\right] = S_0$$

$$\frac{\partial\bar{\Psi}_{\bar{\mu},\mathcal{K}}}{\partial \Delta_V^0} = 0 : \mathbb{E}^{\bar{\mu}}\left[V\frac{d\mu_{\mathcal{K}}^*}{d\bar{\mu}}\right] = F_V \qquad \frac{\partial\bar{\Psi}_{\bar{\mu},\mathcal{K}}}{\partial a_K^1} = 0 : \mathbb{E}^{\bar{\mu}}\left[(S_1-K)_+\frac{d\mu_{\mathcal{K}}^*}{d\bar{\mu}}\right] = C_K^1$$

$$\frac{\partial\bar{\Psi}_{\bar{\mu},\mathcal{K}}}{\partial a_K^V} = 0 : \mathbb{E}^{\bar{\mu}}\left[(V-K)_+\frac{d\mu_{\mathcal{K}}^*}{d\bar{\mu}}\right] = C_K^V \qquad \frac{\partial\bar{\Psi}_{\bar{\mu},\mathcal{K}}}{\partial a_K^2} = 0 : \mathbb{E}^{\bar{\mu}}\left[(S_2-K)_+\frac{d\mu_{\mathcal{K}}^*}{d\bar{\mu}}\right] = C_K^2$$

$$\frac{\partial\bar{\Psi}_{\bar{\mu},\mathcal{K}}}{\partial \Delta_S(s_1,v)} = 0 : \mathbb{E}^{\bar{\mu}}\left[(S_2-S_1)\frac{d\mu_{\mathcal{K}}^*}{d\bar{\mu}}\middle| S_1=s_1, V=v\right] = 0, \quad \forall s_1\geq 0, v>0$$

$$\frac{\partial\bar{\Psi}_{\bar{\mu},\mathcal{K}}}{\partial \Delta_L(s_1,v)} = 0 : \mathbb{E}^{\bar{\mu}}\left[\left(L\left(\frac{S_2}{S_1}\right)-V^2\right)\frac{d\mu_{\mathcal{K}}^*}{d\bar{\mu}}\middle| S_1=s_1, V=v\right] = 0, \quad \forall s_1\geq 0, v>0$$

© 2020 Bloomberg Finance L.P. All rights reserved.