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

A Dispersion-Constrained Martingale Transport Approach

Julien Guyon

Bloomberg L.P.
Quantitative Research

Advances in Financial Mathematics 2020
Paris, January 16, 2020

jguyon2@bloomberg.net

Inversion of Convex Ordering in the VIX Market (JG, 2019)

INVERSION OF CONVEX ORDERING IN THE VIX MARKET

JULIEN GUYON
QUANTITATIVE RESEARCH, BLOOMBERG L.P.

Abstract. 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 $\mathbb{E}[\sigma_t^2|S_t] = \sigma_{\mathrm{lv}}^2(t, S_t)$. 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 $\sigma_{\mathrm{lv}}^2(t, S_t)$ 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 squared is larger in convex order in the associated local volatility model than in the original model for short maturities. We argue and numerically demonstrate that, when the (typically negative) spot-vol correlation is large enough in absolute value, (a) traditional stochastic volatility 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.

Keywords. VIX, convex order, inversion of convex ordering, martingale transport, local volatility, stochastic volatility, mean reversion, rough volatility, smile calibration.

Inversion of Convex ordering: Local Volatility Does Not Maximize the Price of VIX Futures (with B. Acciaio, to appear in SIFIN)

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 this 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 SPX 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.

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 SPX index of the form

dS_t

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.
- 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 S&P 500 (SPX) options, VIX futures and VIX options.
- **Very challenging problem, especially for short maturities**.

Motivation

- 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**.
- For example the double mean-reverting model of Gatheral (2008), though it is very flexible, cannot perfectly fit both the negative at-the-money SPX skew (not large enough in absolute value) and the at-the-money VIX implied volatility (too large) for short maturities up to five months.
- One should decrease the volatility of volatility to decrease the latter, but this would also decrease the former, which is already too small.
- See JG (2017, 2018).

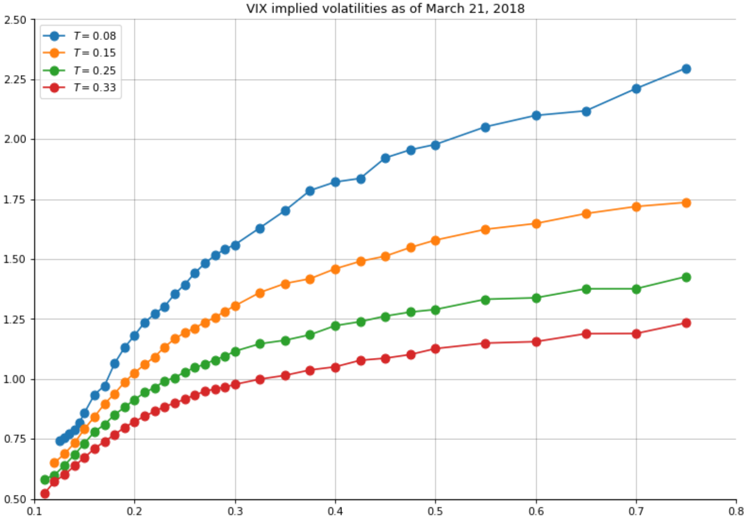
Skewed rough Bergomi model

- 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 (JG 2018):

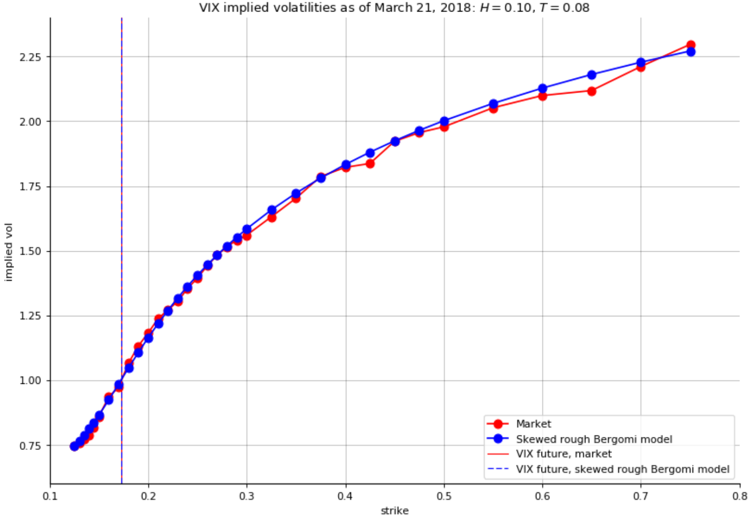
$$\sigma_t^2 = \xi_0^t \left((1-\lambda)\mathcal{E}\left(\nu_0 \int_0^t (t-s)^{H-\frac{1}{2}} dZ_s\right) + \lambda \mathcal{E}\left(\nu_1 \int_0^t (t-s)^{H-1/2} dZ_s\right)\right)$$

  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.

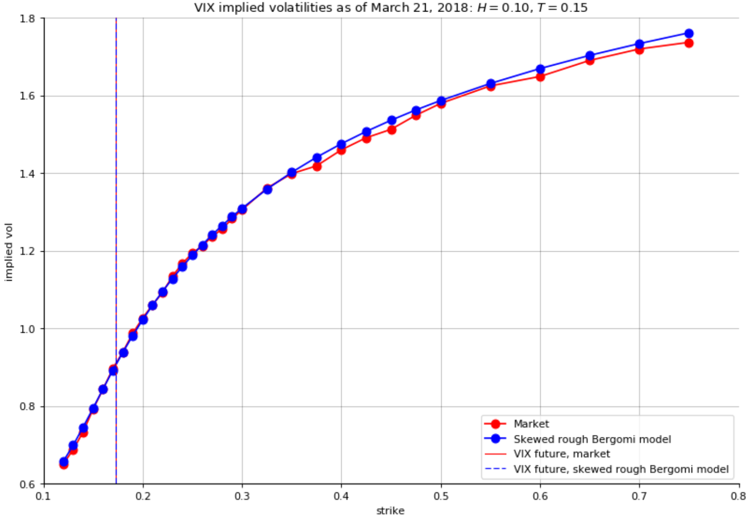
Skewed rough Bergomi: Calibration to VIX future and VIX options (March 21, 2018)



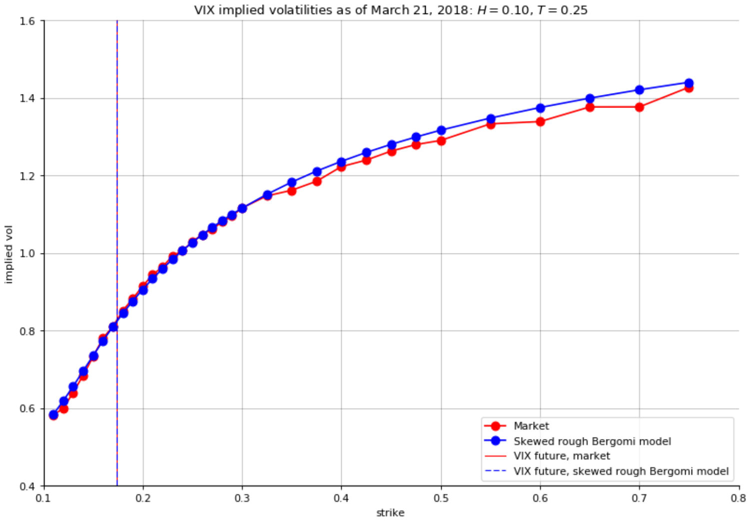
Skewed rough Bergomi: Calibration to VIX future and VIX options (March 21, 2018)



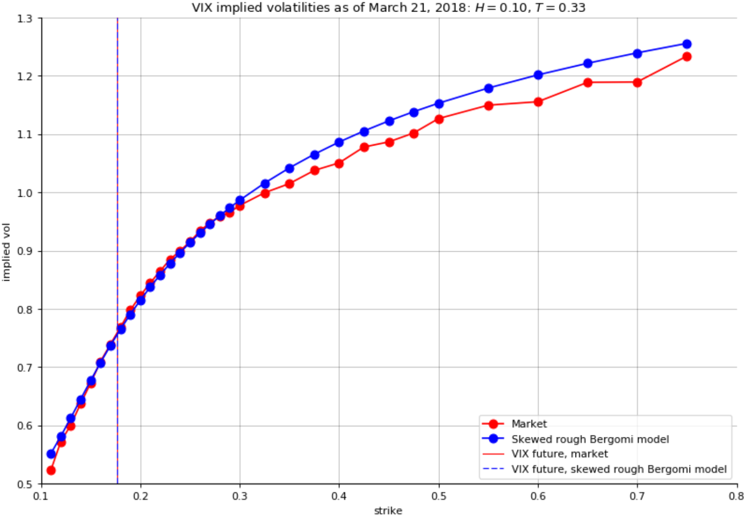
Skewed rough Bergomi: Calibration to VIX future and VIX options (March 21, 2018)



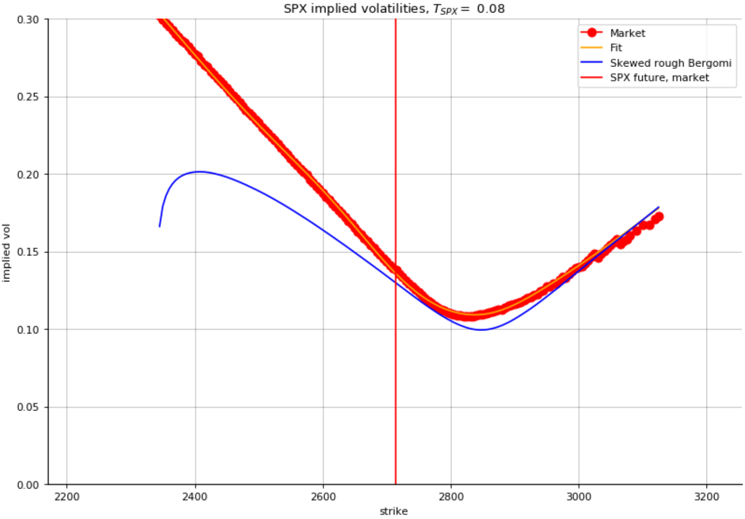
Skewed rough Bergomi: Calibration to VIX future and VIX options (March 21, 2018)



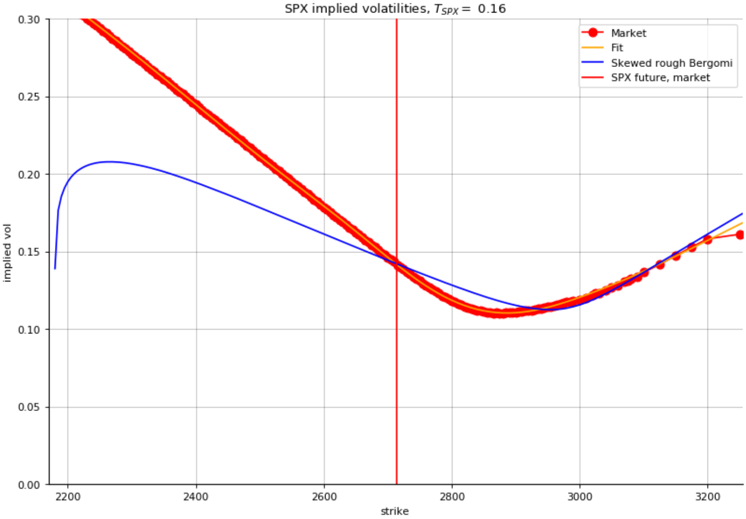
Skewed rough Bergomi: Calibration to VIX future and VIX options (March 21, 2018)



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



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



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.

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

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

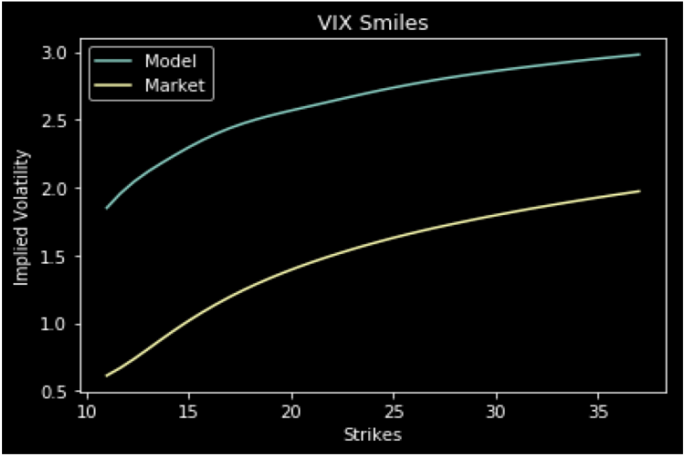
$$\frac{dS_t}{S_t} = \frac{a_t}{\sqrt{\mathbb{E}[a_t^2|S_t]}}\sigma_{\mathrm{loc}}(t, S_t)\, dW_t.$$

- Define

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

- Optimize stoch vol parameters to fit VIX options.

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



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

Related works with continuous models on the SPX

- Fouque-Saporito (2017), 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.
- 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?)."

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**.

Motivation

- 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 the **martingality constraint** on the SPX as well as the requirement that the VIX at T_1 is the implied volatility of the 30-day log-contract on the SPX (**consistency condition**).
- The discrete-time 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).

Setting and notation

- For simplicity: zero interest rates, repos, and dividends.
- $\mu_1 =$ risk-neutral distribution of $S_1 \longleftrightarrow$ market smile of S&P at T_1.
- $\mu_V =$ risk-neutral distribution of $V \longleftrightarrow$ market smile of VIX at T_1.
- $\mu_2 =$ risk-neutral distribution of $S_2 \longleftrightarrow$ market smile of S&P 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

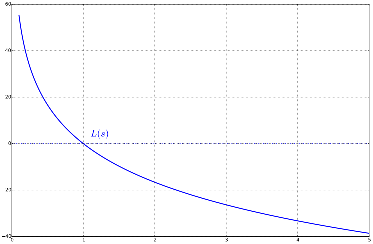
$$\mathbb{E}^i[S_i] = S_0, \quad \mathbb{E}^i[|\ln S_i|] < \infty, \quad i \in \{1,2\}; \qquad \mathbb{E}^V[V] = F_V, \quad \mathbb{E}^V[V^2] < \infty.$$

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

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.



Superreplication, duality

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 the upper and lower bounds of these prices is precisely the subject of **classical optimal transport**.
- 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**, see 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**.

Superreplication: primal problem

Following De Marco-Henry-Labordère (2015), G.-Menegaux-Nutz (2017):

Available instruments:

- 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: MktPrice$[u_1(S_1)]$ + MktPrice$[u_2(S_2)]$ + MktPrice$[u_V(V)]$
- 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:

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

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 integrable superreplicating portfolios, i.e., the set of all measurable functions $(u_1, u_V, u_2, \Delta_S, \Delta_L)$ with $u_1 \in L^1(\mu_1)$, $u_V \in L^1(\mu_V)$, $u_2 \in L^1(\mu_2)$, $\Delta_S, \Delta_L : \mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}$, that satisfy the superreplication constraint: $\forall (s_1, s_2, v) \in \mathbb{R}^2_{>0} \times \mathbb{R}_{\geq 0}$,

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

- Linear program.

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

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

- 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.

Superreplication: absence of a duality gap

Theorem

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*

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

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.*

Proof: straightforward adaptation of the proof of Theorem 1 in Beiglbock et al (martingale optimal transport, 2013).

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 the upper and lower bounds of these prices is precisely the subject of **classical optimal transport**.
- 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**, see 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.

Joint SPX/VIX arbitrage

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

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

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 two quantities in (3.1) do not purely depend on market data. The l.h.s. depends on an (arbitrage-free) extrapolation of the smile of V beyond the last quoted strikes, while the r.h.s. depends on (arbitrage-free) extrapolations of the SPX smile at maturities T_1 and T_2.
- The reported violations of (3.1) actually rely on some arbitrary smile extrapolations.
- JG (2018) explains how to build **consistent extrapolations of the VIX and SPX smiles** so that (3.1) holds.

Joint SPX/VIX arbitrage

Theorem (G., 2018)

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., $\mathbb{E}^\nu[f(S_1, L(S_1) + V^2)] \leq \mathbb{E}^2[f(S_2, L(S_2))]$ for any convex function $f : \mathbb{R}_{>0} \times \mathbb{R} \to \mathbb{R}$.*

(i) $\iff$ (ii): By duality (Theorem 1), we have $P_0 = D_0$. Now, by definition, the market is free of (S_1, S_2, V)-arbitrage if and only if $P_0 = 0$, and from Theorem 1, $\mathcal{P}(\mu_1, \mu_V, \mu_2) \neq \emptyset$ if and only if $D_0 \neq -\infty$, in which case $D_0 = 0$.

Joint SPX/VIX arbitrage

(ii) $\Longleftrightarrow$ (iii): Define $M_1 = (S_1, L(S_1) + V^2)$, $M_2 = (S_2, L(S_2))$, and

$$\mu_{M_2}(dx, dy) = \mu_2(dx)\delta_{L(x)}(dy).$$

Let $\Pi(\mu_1, \mu_V)$ denote the set of transport plans from μ_1 to μ_V, i.e., the set of all couplings of μ_1 and μ_V.
For $\nu \in \Pi(\mu_1, \mu_V)$, denote by $\mu^{\nu}_{M_1}$ the distribution of M_1 under ν and by $\mathcal{M}(\nu, \mu_2)$ the set of all probability measures μ on $\mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0}$ s.t.

$$M_1 \sim \mu^{\nu}_{M_1}, \qquad M_2 \sim \mu_{M_2}, \qquad \mathbb{E}^{\mu}\left[M_2 | M_1\right] = M_1.$$

Then

$$\mathcal{P}(\mu_1, \mu_V, \mu_2) = \bigcup_{\nu \in \Pi(\mu_1, \mu_V)} \mathcal{M}(\nu, \mu_2).$$

By Strassen's theorem, each $\mathcal{M}(\nu, \mu_2)$ is nonempty if and only if $\mu^{\nu}_{M_1}$ and μ_{M_2} are in convex order.

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 will 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).

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

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

- We explain how to numerically 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).

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

- $\mathcal{M}_1$: set of probability measures on $\mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0}$.
- $\mathcal{U}$: set of all integrable portfolios $u = (u_1, u_V, u_2, \Delta_S, \Delta_L)$.
- Introduce the Lagrange multipliers $u = (u_1, u_V, u_2, \Delta_S, \Delta_L)$ associated to the five constraints of $\mathcal{P}(\mu_1, \mu_V, \mu_2)$ and assume that the inf and sup operators can be swapped (absence of a duality gap):

$$
\begin{aligned}
D_{\bar{\mu}} &:= \inf_{\mu \in \mathcal{P}(\mu_1, \mu_V, \mu_2)} H(\mu, \bar{\mu}) \\
&= \inf_{\mu \in \mathcal{M}_1} \sup_{u \in \mathcal{U}} \Big\{ H(\mu, \bar{\mu}) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \\
&\quad - \mathbb{E}^\mu \Big[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(s_1, v, s_2) + \Delta_L^{(L)}(s_1, v, s_2) \Big] \Big\} \\
&= \sup_{u \in \mathcal{U}} \inf_{\mu \in \mathcal{M}_1} \Big\{ H(\mu, \bar{\mu}) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \\
&\quad - \mathbb{E}^\mu \Big[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(s_1, v, s_2) + \Delta_L^{(L)}(s_1, v, s_2) \Big] \Big\}
\end{aligned}
$$

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

$$\begin{aligned} D_{\bar\mu} &= \sup_{u\in\mathcal{U}} \inf_{\mu\in\mathcal{M}_1} \Big\{ H(\mu,\bar\mu) + \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_2(S_2)] \\ &\quad -\mathbb{E}^\mu\left[u_1(S_1) + u_V(V) + u_2(S_2) + \Delta_S^{(S)}(s_1,v,s_2) + \Delta_L^{(L)}(s_1,v,s_2)\right]\Big\} \end{aligned}$$

For any random variable X, denote by $\bar\mu_X$ the probability distribution defined by $\frac{d\bar\mu_X}{d\bar\mu} = \frac{e^X}{\mathbb{E}^{\bar\mu}[e^X]}$:

$$\begin{aligned} \inf_{\mu\in\mathcal{M}_1} \{H(\mu,\bar\mu) - \mathbb{E}^\mu[X]\} &= \inf_{\mu\in\mathcal{M}_1} \mathbb{E}^\mu\left[\ln\frac{d\mu}{d\bar\mu} - X\right] = \inf_{\mu\in\mathcal{M}_1} \mathbb{E}^\mu\left[\ln\frac{d\mu}{d\bar\mu_X} + \ln\frac{d\bar\mu_X}{d\bar\mu} - X\right] \\ &= \inf_{\mu\in\mathcal{M}_1} \mathbb{E}^\mu\left[\ln\frac{d\mu}{d\bar\mu_X} - \ln\mathbb{E}^{\bar\mu}[e^X]\right] = \inf_{\mu\in\mathcal{M}_1} H(\mu,\bar\mu_X) - \ln\mathbb{E}^{\bar\mu}[e^X] = -\ln\mathbb{E}^{\bar\mu}[e^X] \end{aligned}$$

and the infimum is attained at $\mu = \bar\mu_X$ since for all $\mu\in\mathcal{M}_1$, $H(\mu,\bar\mu_X)\geq 0$ and $H(\mu,\bar\mu_X) = 0$ if and only if $\mu = \bar\mu_X$.

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

$$D_{\bar{\mu}} = \sup_{u \in \mathcal{U}} \Psi_{\bar{\mu}}(u) =: P_{\bar{\mu}}$$

where for $u = (u_1, u_V, u_2, \Delta_S, \Delta_L) \in \mathcal{U}$, we have defined

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

- $D_{\bar{\mu}} \neq +\infty$ if and only if $\mathcal{P}(\mu_1, \mu_V, \mu_2) \neq \emptyset$, and in that case the infimum defining $D_{\bar{\mu}}$ is attained. Indeed, $\mu \mapsto H(\mu, \bar{\mu})$ is lower semicontinuous in the weak topology (Dembo-Zeitouni). Since $\mathcal{P}(\mu_1, \mu_V, \mu_2)$ is compact in this topology, the infimum is attained.
- If the supremum defining $P_{\bar{\mu}}$ is attained at $u^* = (u_1^*, u_V^*, u_2^*, \Delta_S^*, \Delta_L^*)$, the infimum defining $D_{\bar{\mu}}$ is reached at

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

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

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

- $\Psi_{\bar{\mu}}$ is invariant by translation of u_1, u_V, and u_2: for any constant $c \in \mathbb{R}$, $\Psi_{\bar{\mu}}(u_1 + c, u_V, u_2, \Delta_S, \Delta_L) = \Psi_{\bar{\mu}}(u_1, u_V, u_2, \Delta_S, \Delta_L)$ (and similarly with u_V and u_2); $c =$ cash position $\Longrightarrow$ We will always work with a normalized version of $u^* \in \mathcal{U}$ s.t.

$$\mathbb{E}^{\bar{\mu}}\left[e^{u_1^*(S_1)+u_V^*(V)+u_2^*(S_2)+\Delta_S^{*(S)}(S_1,V,S_2)+\Delta_L^{*(L)}(S_1,V,S_2)}\right] = 1. \qquad (4.1)$$

- **By duality, the initial, difficult problem of minimizing over $\mu \in \mathcal{P}(\mu_1, \mu_V, \mu_2)$ (constrained) has been reduced to the simpler problem of maximizing the strictly concave function $\Psi_{\bar{\mu}}$ over $u \in \mathcal{U}$ (unconstrained)**. If it exists, the optimum u^* cancels the gradient of $\Psi_{\bar{\mu}}$: $\frac{\partial \Psi_{\bar{\mu}}}{\partial u_1(s_1)} = \frac{\partial \Psi_{\bar{\mu}}}{\partial u_V(v)} = \frac{\partial \Psi_{\bar{\mu}}}{\partial u_2(s_2)} = \frac{\partial \Psi_{\bar{\mu}}}{\partial \Delta_S(s_1,v)} = \frac{\partial \Psi_{\bar{\mu}}}{\partial \Delta_L(s_1,v)} = 0.$

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

$$
\begin{array}{rrcl}
\forall s_1 > 0, & u_1(s_1) & = & \Phi_1(s_1; u_V, u_2, \Delta_S, \Delta_L) \\
\forall v \geq 0, & u_V(v) & = & \Phi_V(v; u_1, u_2, \Delta_S, \Delta_L) \\
\forall s_2 > 0, & u_2(s_2) & = & \Phi_2(s_2; u_1, u_V, \Delta_S, \Delta_L) \\
\forall s_1 > 0,\ \forall v \geq 0, & 0 & = & \Phi_{\Delta_S}(s_1, v; \Delta_S(s_1, v), \Delta_L(s_1, v)) \\
\forall s_1 > 0,\ \forall v \geq 0, & 0 & = & \Phi_{\Delta_L}(s_1, v; \Delta_S(s_1, v), \Delta_L(s_1, v))
\end{array}
\tag{4.2}
$$

where, imposing the normalization (4.1),

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

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

- Note that these are also the equations satisfied by the maximum of (no log)

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

- One could directly get that $D_{\bar{\mu}} = \sup_{u\in\mathcal{U}} \bar{\Psi}_{\bar{\mu}}(u)$ by using the set $\mathcal{M}_+$ of nonnegative measures instead of $\mathcal{M}_1$ in (4.1), and by computing the inner $\inf_{\mu\in\mathcal{M}_+}$ in (4.1) by differentiating w.r.t. $\frac{d\mu}{d\bar{\mu}}$.
- In any case, **the jointly calibrating model reads**

$$\mu^*(ds_1, dv, ds_2) = \bar{\mu}(ds_1, dv, ds_2)e^{u_1^*(s_1)+u_V^*(v)+u_2^*(s_2)+\Delta_S^{*(S)}(s_1,v,s_2)+\Delta_L^{*(L)}(s_1,v,s_2)}. \tag{4.3}$$

where $u^* = (u_1^*, u_V^*, u_2^*, \Delta_S^*, \Delta_L^*)$ is solution of (4.2).

- We could have simply postulated a model of the form (4.3)! Then the five conditions defining $\mathcal{P}(\mu_1, \mu_V, \mu_2)$ translate into the five equations (4.2).

Sinkhorn's algorithm

- Sinkhorn's algorithm (1967) was first used in the context of optimal transport by Cuturi (2013).
- In our context, Sinkhorn's algorithm is an exponentially fast **fixed point method that iterates computations of one-dimensional gradients to approximate the optimizer** u^*.
- Starting from an initial $u^{(0)} = (u_1^{(0)}, u_V^{(0)}, u_2^{(0)}, \Delta_S^{(0)}, \Delta_L^{(0)})$, we 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 five lines corresponds to a Bregman projection in the space of measures.

Sinkhorn's algorithm

- Sinkhorn's algorithm (1967) was first used in the context of optimal transport by Cuturi (2013).
- In our context, Sinkhorn's algorithm is an exponentially fast **fixed point method that iterates computions of one-dimensional gradients to approximate the optimizer u^***.
- Starting from an initial $u^{(0)} = (u_1^{(0)}, u_V^{(0)}, u_2^{(0)}, \Delta_S^{(0)}, \Delta_L^{(0)})$, we 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 five lines corresponds to a Bregman projection in the space of measures.

Sinkhorn's algorithm

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

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

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

Sinkhorn's algorithm

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

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

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

Sinkhorn's algorithm

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

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

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

Implementation details

Implementation details

- Natural choice: pick a reference measure $\bar{\mu}$ that satisfies all the constraints of $\mathcal{P}(\mu_1, \mu_V, \mu_2)$ except $S_2 \sim \mu_2$, i.e., pick $\bar{\mu}$ in the set $\mathcal{P}(\mu_1, \mu_V)$ of all the probability distributions

$$\mu(ds_1, dv, ds_2) = \nu(ds_1, dv)\, T(s_1, v, ds_2)$$

where ν is a coupling of μ_1 and μ_V and the transition kernel $T(s_1, v, ds_2)$ satisfies

$$\int s_2\, T(s_1, v, ds_2) = s_1, \qquad \int L(s_2)\, T(s_1, v, ds_2) = L(s_1) + v^2$$

for μ_1-a.e. $s_1 > 0$ and μ_V-a.e. $v \geq 0$.

- For instance, we may choose

$$\nu = \mu_1 \otimes \mu_V, \qquad T(s_1, v, ds_2) \text{ is the distribution of } s_1 \exp\left(v\sqrt{\tau}G - \frac{1}{2}v^2\tau\right),$$

where G denotes a standard Gaussian random variable.

Implementation details

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

$$\mu^*_{\mathcal{K}}(ds_1, dv, ds_2) = \bar{\mu}(ds_1, dv, ds_2) e^{c^* + \Delta^{0*}_S s_1 + \Delta^{0*}_V v + \sum_{K \in \mathcal{K}_1} a^{1*}_K (s_1 - K)_+} \\ e^{\sum_{K \in \mathcal{K}_V} a^{V*}_K (v-K)_+ + \sum_{K \in \mathcal{K}_2} a^{2*}_K (s_2 - K)_+ + \Delta^{*(S)}_S (s_1, v, s_2) + \Delta^{*(L)}_L (s_1, v, s_2)}$$

where $\theta^* := (c^*, \Delta^{0*}_S, \Delta^{0*}_V, a^{1*}, a^{V*}, a^{2*}, \Delta^*_S, \Delta^*_L)$ maximizes

$$\bar{\Psi}_{\bar{\mu}, \mathcal{K}}(\theta) := c + \Delta^0_S S_0 + \Delta^0_V F_V + \sum_{K \in \mathcal{K}_1} a^1_K C^1_K + \sum_{K \in \mathcal{K}_V} a^V_K C^V_K + \sum_{K \in \mathcal{K}_2} a^2_K C^2_K$$

$$- \mathbb{E}^{\bar{\mu}} \left[e^{c + \Delta^0_S S_1 + \Delta^0_V V + \sum_{\mathcal{K}_1} a^1_K (S_1 - K)_+ + \sum_{\mathcal{K}_V} a^V_K (V-K)_+ + \sum_{\mathcal{K}_2} a^2_K (S_2 - K)_+ + \Delta^{(S)}_S (\ldots) + \Delta^{(L)}_L (\ldots)} \right]$$

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

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

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

- 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 $\frac{\partial \bar{\Psi}_{\bar{\mu},\mathcal{K}}}{\partial \theta_i}(\theta) = 0$:

Implementation details

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

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

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

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

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

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

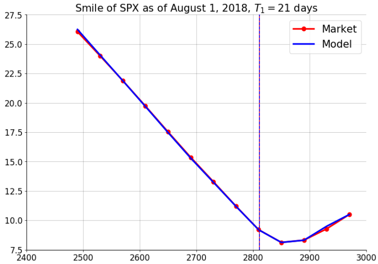
$$\frac{\partial\bar{\Psi}_{\bar{\mu},\mathcal{K}}}{\partial\Delta_L(s_1,v)} = 0 : \mathbb{E}^{\bar{\mu}}\left[\left(L\left(\frac{S_2}{S_1}\right)-V^2\right)\frac{d\mu^*_{\mathcal{K}}}{d\bar{\mu}}\middle|S_1=s_1,V=v\right] = 0, \quad \forall s_1\geq 0, v>0$$

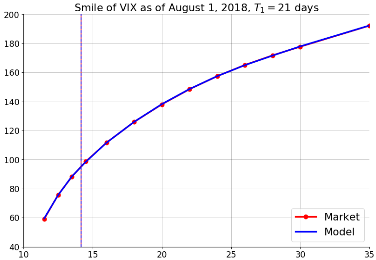
Implementation details

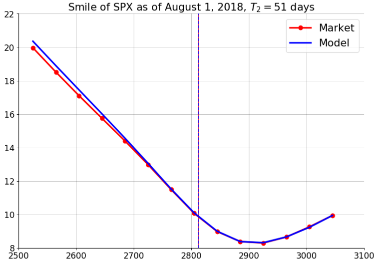
- We use $\theta^{(0)} = 0$ as the starting point of the Sinkhorn algorithm.
- Integrals estimated using Gaussian quadrature; Gauss-Legendre when we integrate over s_1 and v, and Gauss-Hermite when we integrate over s_2.
- While the expression for $c^{(n+1)}$ is explicit, computing the other parameters requires using a one-dimensional root solver; we use Newton's algorithm.
- As an exception, for each point s_1 and v in the quadrature, $(\Delta_S^{(n+1)}(s_1, v), \Delta_L^{(n+1)}(s_1, v))$ are jointly computed using the Levenberg-Marquardt algorithm.
- Enough accuracy is typically reached after about a hundred iterations and gives us θ^*, hence $\mu_{\mathcal{K}}^*$.
- **If the Sinkhorn algorithm diverges**, then $D_{\bar{\mu},\mathcal{K}} = +\infty$, so $P_{\bar{\mu},\mathcal{K}} = +\infty$, which means that $\mathcal{P}(\mathcal{K}) = \emptyset$, i.e., **there exists a joint SPX/VIX arbitrage** (based only on $\mathcal{K}$).

Numerical experiments

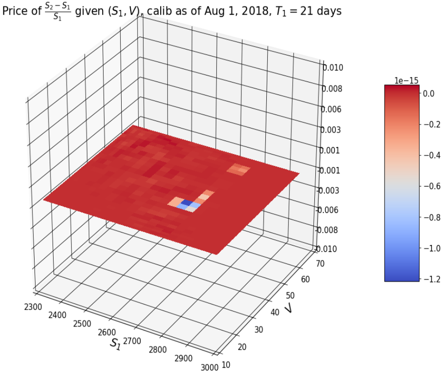
August 1, 2018, $T_1 = 21$ days

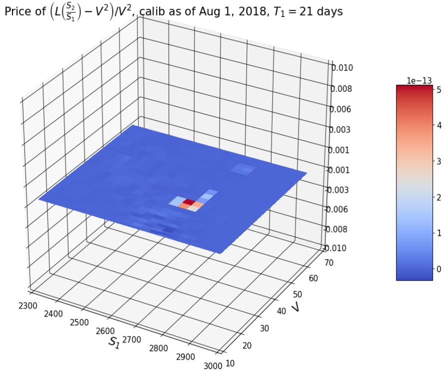






August 1, 2018, $T_1 = 21$ days





August 1, 2018, $T_1 = 21$ days

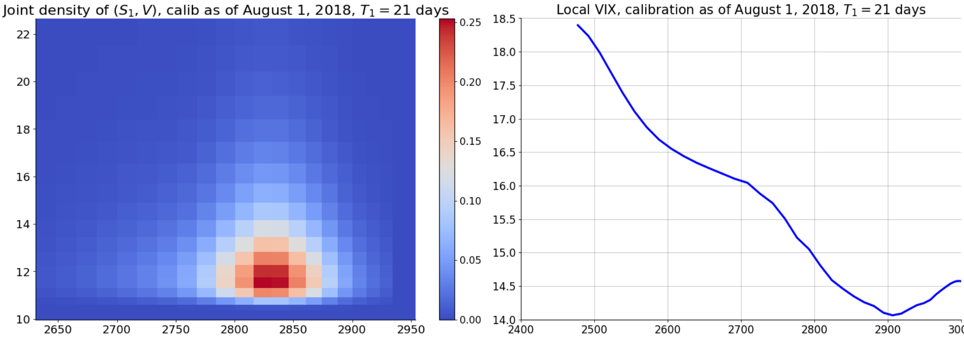


Figure: Joint distribution of (S_1, V) and local VIX function $\mathrm{VIX}_{\mathrm{loc}}(s_1)$

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

August 1, 2018, $T_1 = 21$ days

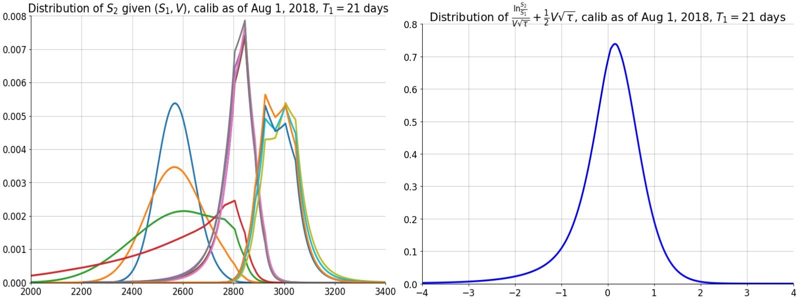


Figure: Conditional distribution of S_2 given (s_1, v) under $\mu^*_{\mathcal{K}}$ for different vales of (s_1, v): $s_1 \in \{2571, 2808, 3000\}$, $v \in \{10.10, 15.30, 23.20, 35.72\}\%$, and distribution of the normalized return $R := \frac{\ln(S_2/S_1)}{V\sqrt{\tau}} + \frac{1}{2}V\sqrt{\tau}$

August 1, 2018, $T_1 = 21$ days

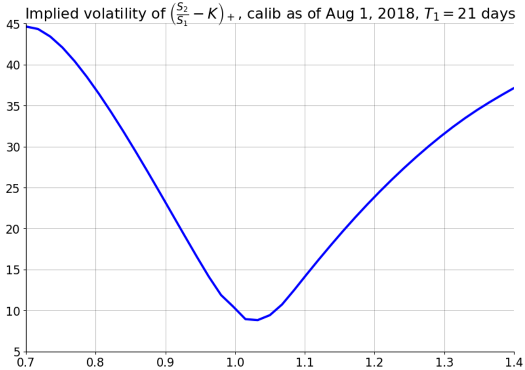
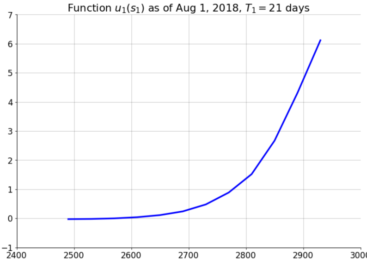
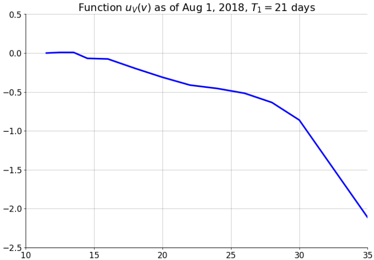
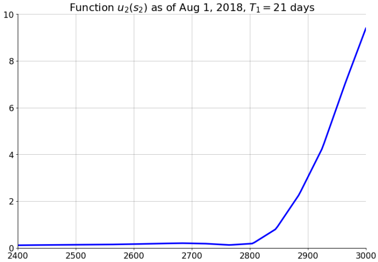


Figure: Smile of forward starting call options $(S_2/S_1 - K)_+$

August 1, 2018, $T_1 = 21$ days







August 1, 2018, $T_1 = 21$ days

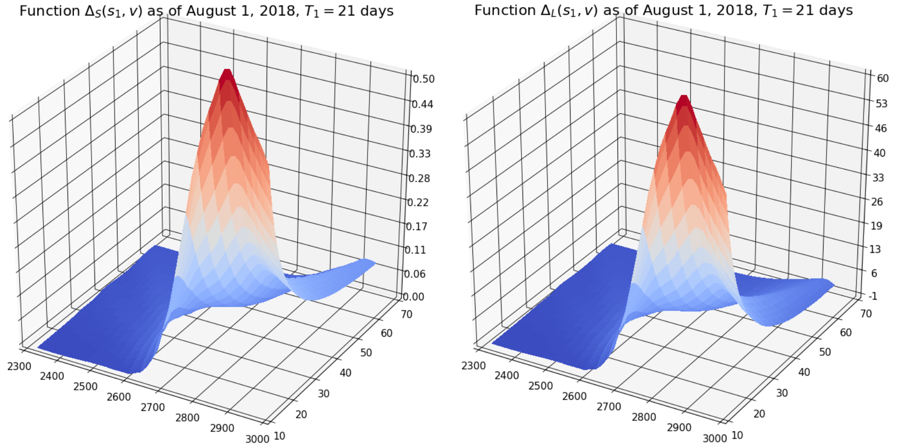
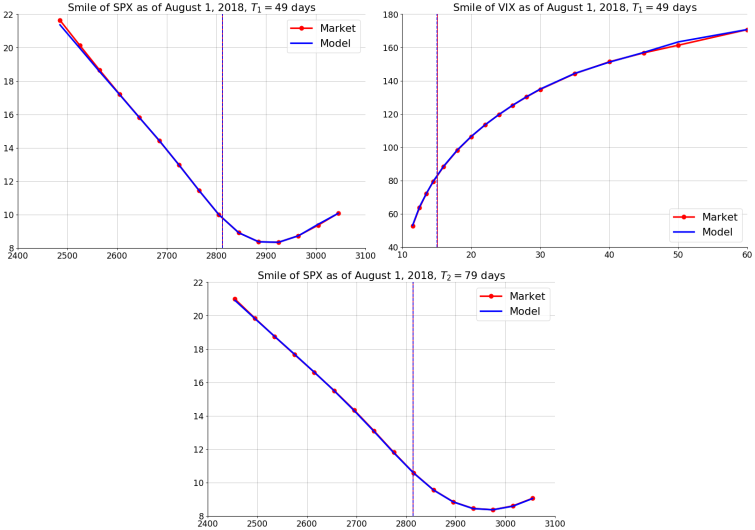
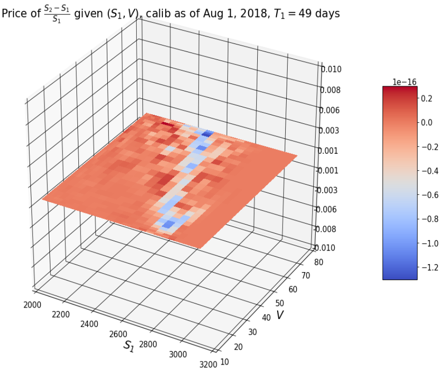


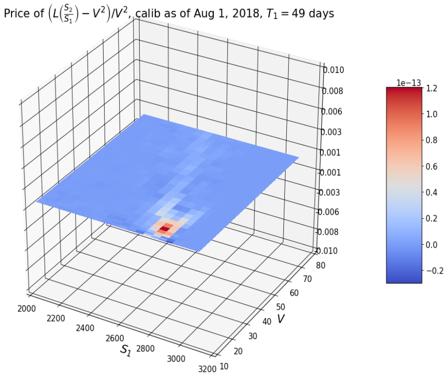
Figure: Optimal functions $\Delta_S^*(s_1, v)$ and $\Delta_L^*(s_1, v)$ for (s_1, v) in the quadrature grid

August 1, 2018, $T_1 = 49$ days



August 1, 2018, $T_1 = 49$ days





August 1, 2018, $T_1 = 49$ days

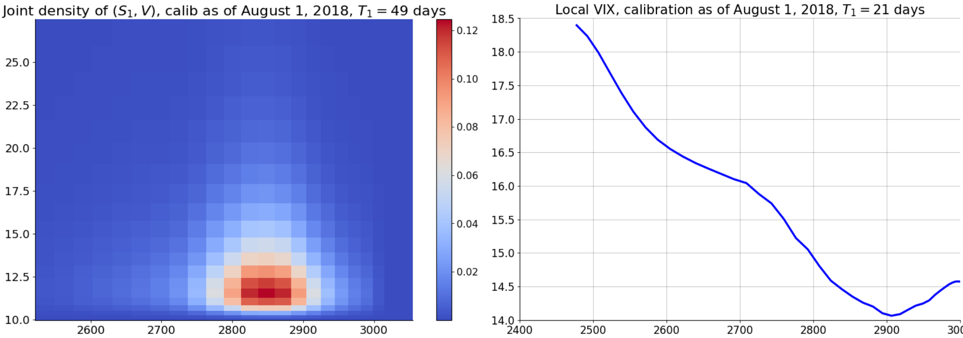


Figure: Joint distribution of (S_1, V) and local VIX function $\text{VIX}_{\text{loc}}(s_1)$

August 1, 2018, $T_1 = 49$ days

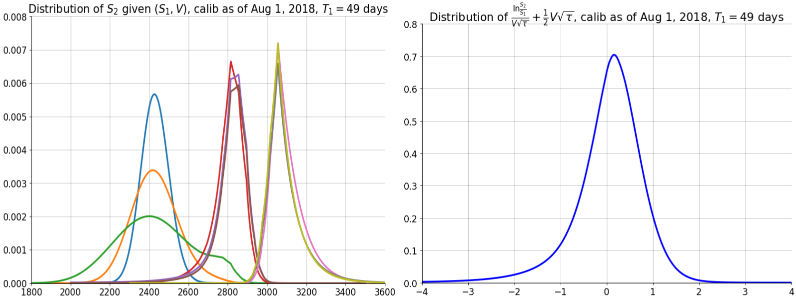


Figure: Conditional distribution of S_2 given (s_1, v) under $\mu^*_{\mathcal{K}}$ for different vales of (s_1, v): $s_1 \in \{2571, 2808, 3000\}$, $v \in \{10.10, 15.30, 23.20, 35.72\}\%$, and distribution of the normalized return $R := \frac{\ln(S_2/S_1)}{V\sqrt{\tau}} + \frac{1}{2}V\sqrt{\tau}$

August 1, 2018, $T_1 = 49$ days

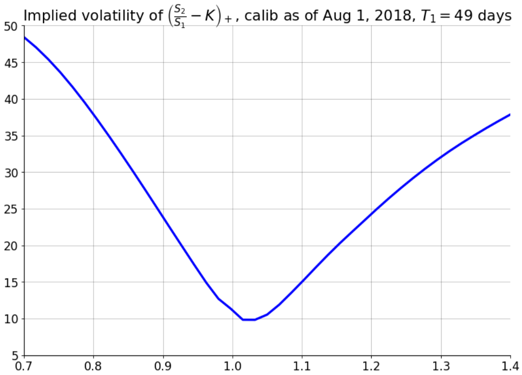
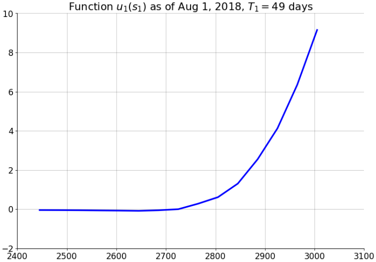
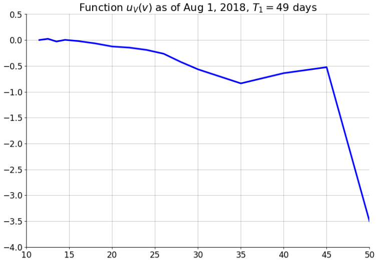
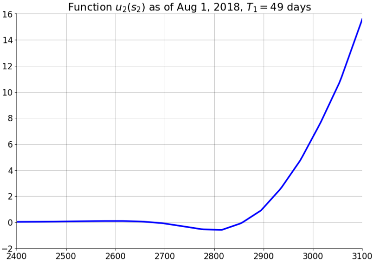


Figure: Smile of forward starting call options $(S_2/S_1 - K)_+$

August 1, 2018, $T_1 = 49$ days







August 1, 2018, $T_1 = 49$ days

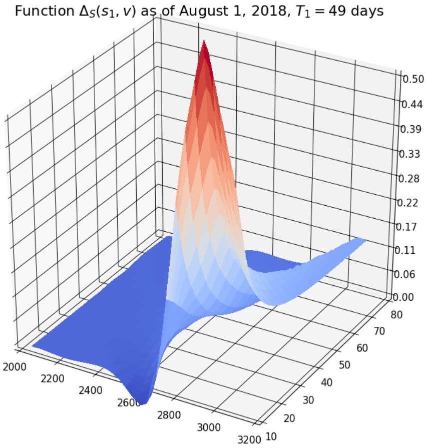
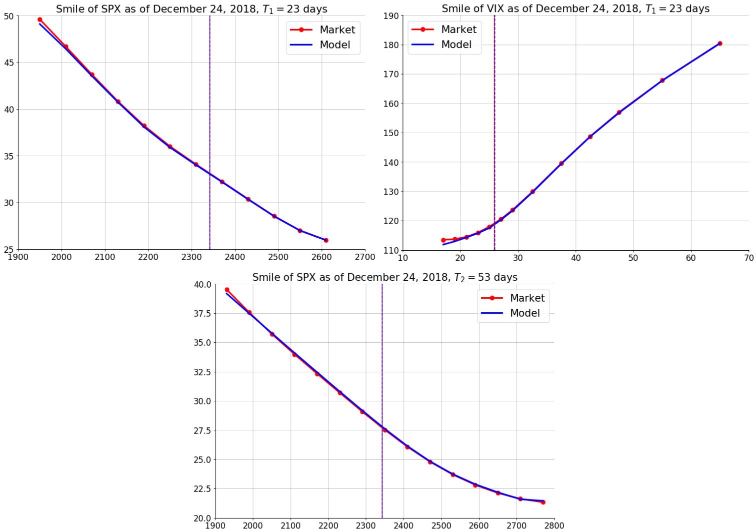
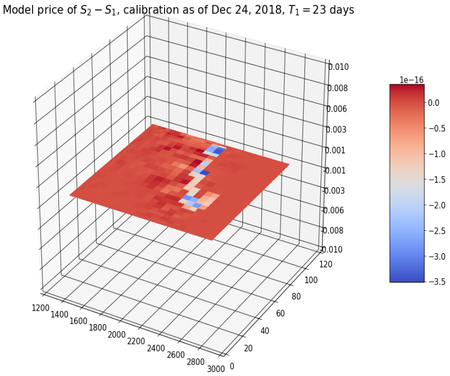


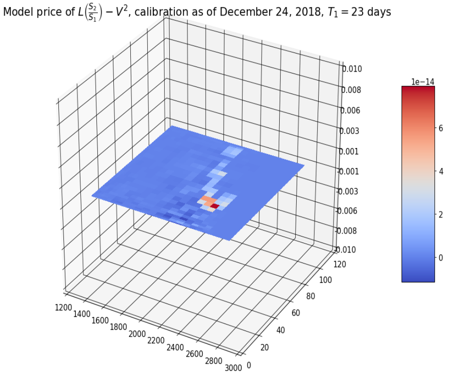
Figure: Optimal functions $\Delta_S^*(s_1, v)$ and $\Delta_L^*(s_1, v)$ for (s_1, v) in the quadrature grid

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



December 24, 2018, $T_1 = 23$ days





December 24, 2018, $T_1 = 23$ days

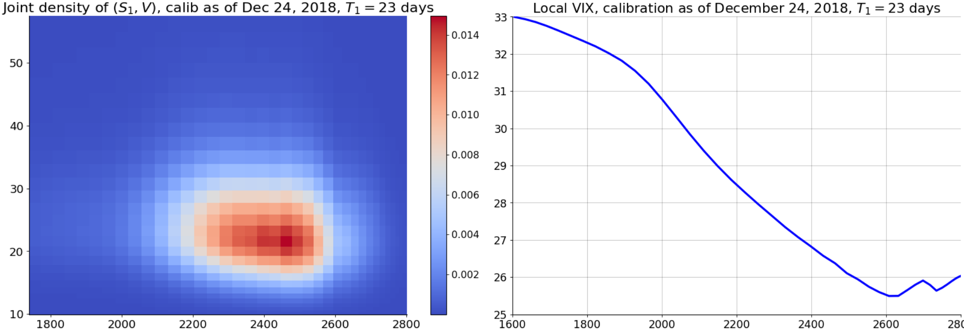


Figure: Joint distribution of (S_1, V) and local VIX function $\text{VIX}_{\text{loc}}(s_1)$

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

December 24, 2018, $T_1 = 23$ days

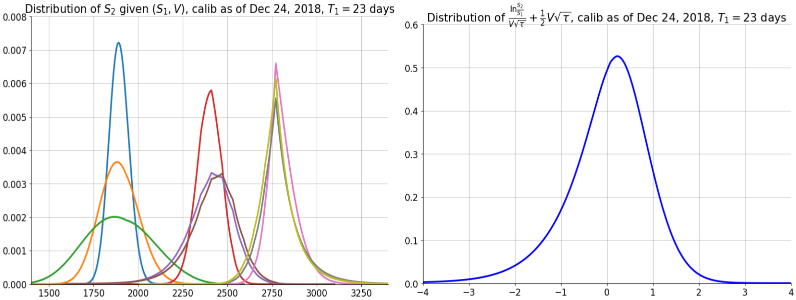


Figure: Conditional distribution of S_2 given (s_1, v) under $\mu^*_{\mathcal{K}}$ for different vales of (s_1, v): $s_1 \in \{2571, 2808, 3000\}$, $v \in \{10.10, 15.30, 23.20, 35.72\}\%$, and distribution of the normalized return $R := \frac{\ln(S_2/S_1)}{V\sqrt{\tau}} + \frac{1}{2}V\sqrt{\tau}$

December 24, 2018, $T_1 = 23$ days

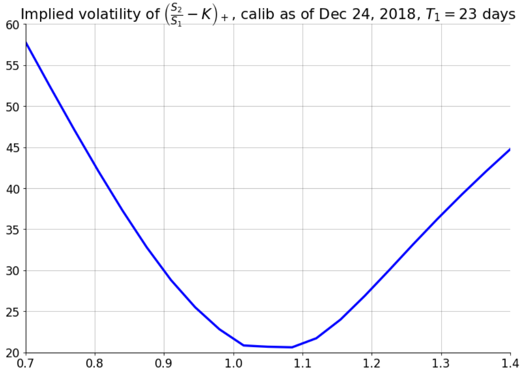
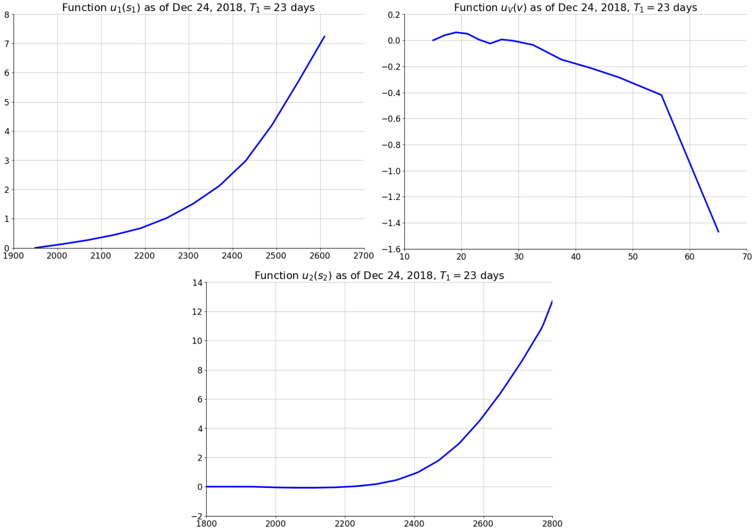


Figure: Smile of forward starting call options $(S_2/S_1 - K)_+$

December 24, 2018, $T_1 = 23$ days



December 24, 2018, $T_1 = 23$ days

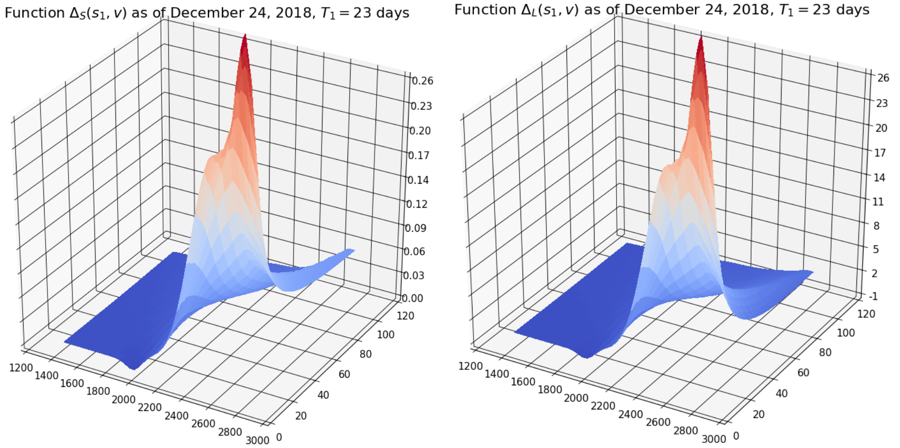


Figure: Optimal functions $\Delta_S^*(s_1, v)$ and $\Delta_L^*(s_1, v)$ for (s_1, v) in the quadrature grid

Maturity issue

- SPX options maturity $T_1' = T_1 + 2$ days or $T_1 - 5$ days.
- Rigorous treatment: introduce S_1' representing the value of the SPX index at time T_1'. If T_1' is two days after T_1, we consider the primal portfolios

$$u_1(s_1')+u_V(v)+u_2(s_2)+\Delta_S(s_1,v)(s_1'-s_1)+\Delta_S'(s_1,v,s_1')(s_2-s_1')+\Delta_L(s_1,v)\left(L\left(\frac{s_2}{s_1}\right)-$$

and the dual risk-neutral probability measures $V \sim \mu_V, S_1' \sim \mu_1, S_2 \sim \mu_2$,

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

- If T_1' is five days before T_1, the primal portfolios are

$$u_1(s_1')+u_V(v)+u_2(s_2)+\Delta_S'(s_1')(s_1-s_1')+\Delta_S(s_1',s_1,v)(s_2-s_1)+\Delta_L(s_1',s_1,v)\left(L\left(\frac{s_2}{s_1}\right)-$$

and the dual risk-neutral probability measures $V \sim \mu_V, S_1' \sim \mu_1, S_2 \sim \mu_2$,

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

- Approx: assume SPX options mature exactly at T_1; maturity interpolation of SPX data.

Extension to the multi-maturity case

Extension to the multi-maturity case

- Assume that monthly SPX options and VIX futures maturities T_i perfectly coincide and, for two consecutive months, are separated by exactly 30 days, $T_{i+1} - T_i = \tau$ for all $i \geq 1$.
- Assume that for each i we are able to build a jointly calibrating model ν_i using the Sinkhorn algorithm.
- Here ν_i denotes the joint distribution of (S_i, V_i, S_{i+1}) where S_i and V_i denote the SPX and VIX values at T_i.
- Then we can build a calibrated model on $(S_i, V_i)_{i \geq 1}$ as follows: $(S_1, V_1, S_2) \sim \nu_1$; recursively we 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}.
- It is easy to check that 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 $i \geq 1$,

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

Continuous time

(joint work with Pierre Henry-Labordère)

Continuous time

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

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

- 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. $\mathbb{P}_0$:

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

Continuous time

$$
\begin{aligned}
D &:= \inf_{\mathbb{P}\in\mathcal{P}} H(\mathbb{P},\mathbb{P}_0) \\
&= \inf_{\mathbb{P}\in\mathcal{M}_1} \sup_{u_1\in L^1(\mu_1),u_2\in L^1(\mu_2),u_V\in L^1(\mu_V),(\Delta_t)\,\mathcal{F}\text{-adapted}} \Big\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i, u_i) \\
&\qquad - \mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V\left(\sqrt{\mathbb{E}^{\mathbb{P}}\left[L\left(\frac{S_2}{S_1}\right)\middle|\mathcal{F}_1\right]}\right) + \int_0^{T_2} \Delta_t \, dS_t \right] \Big\} \\
&= \inf_{\mathbb{P}\in\mathcal{M}_1} \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Big\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i, u_i) \\
&\qquad - \mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V(V) + \int_0^{T_2} \Delta_t \, dS_t + \Delta^L\left(L\left(\frac{S_2}{S_1}\right) - V^2 \right) \right] \Big\} \\
&\overset{\text{(dual)}}{=} \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \inf_{\mathbb{P}\in\mathcal{M}_1} \Big\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i, u_i) \\
&\qquad - \mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V(V) + \int_0^{T_2} \Delta_t \, dS_t + \Delta^L\left(L\left(\frac{S_2}{S_1}\right) - V^2 \right) \right] \Big\} \\
&= \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Big\{ \sum_{i\in\{1,2,V\}} (\mu_i, u_i) \\
&\qquad - \ln \mathbb{E}^0\left[e^{u_1(S_1)+u_2(S_2)+u_V(V)+\int_0^{T_2} \Delta_t \, dS_t + \Delta^L\left(L\left(\frac{S_2}{S_1}\right) - V^2\right)} \right] \Big\} \\
&= \sup_{u_1,u_2,u_V} \sup_{(\Delta_t)_{t\in[0,T_1]}} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \sup_{(\Delta_t)_{t\in[T_1,T_2]}} \{\cdots\}
\end{aligned}
$$

Continuous time

$$
\begin{aligned}
D \quad &:= \quad \inf_{\mathbb{P}\in\mathcal{P}} H(\mathbb{P},\mathbb{P}_0) \\
&= \quad \inf_{\mathbb{P}\in\mathcal{M}_1} \sup_{u_1\in L^1(\mu_1),u_2\in L^1(\mu_2),u_V\in L^1(\mu_V),(\Delta_t)\,\mathcal{F}\text{-adapted}} \Bigg\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i,u_i) \\
&\qquad -\mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V\left(\sqrt{\mathbb{E}^{\mathbb{P}}\left[L\left(\frac{S_2}{S_1}\right)\middle|\mathcal{F}_1\right]}\right) + \int_0^{T_2} \Delta_t\,dS_t \right] \Bigg\} \\
&= \quad \inf_{\mathbb{P}\in\mathcal{M}_1} \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Bigg\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i,u_i) \\
&\qquad -\mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V(V) + \int_0^{T_2} \Delta_t\,dS_t + \Delta^L\left(L\left(\frac{S_2}{S_1}\right) - V^2\right)\right] \Bigg\} \\
&\overset{\text{(dual)}}{=} \quad \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \inf_{\mathbb{P}\in\mathcal{M}_1} \Bigg\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i,u_i) \\
&\qquad -\mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V(V) + \int_0^{T_2} \Delta_t\,dS_t + \Delta^L\left(L\left(\frac{S_2}{S_1}\right) - V^2\right)\right] \Bigg\} \\
&= \quad \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Bigg\{ \sum_{i\in\{1,2,V\}} (\mu_i,u_i) \\
&\qquad - \ln \mathbb{E}^0\left[e^{u_1(S_1)+u_2(S_2)+u_V(V)+\int_0^{T_2} \Delta_t\,dS_t + \Delta^L\left(L\left(\frac{S_2}{S_1}\right)-V^2\right)} \right] \Bigg\} \\
&= \quad \sup_{u_1,u_2,u_V} \sup_{(\Delta_t)_{t\in[0,T_1]}} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \sup_{(\Delta_t)_{t\in[T_1,T_2]}} \{\cdots\}
\end{aligned}
$$

Continuous time

$$
\begin{aligned}
D &:= \inf_{\mathbb{P}\in\mathcal{P}} H(\mathbb{P},\mathbb{P}_0) \\
&= \inf_{\mathbb{P}\in\mathcal{M}_1} \sup_{u_1\in L^1(\mu_1),u_2\in L^1(\mu_2),u_V\in L^1(\mu_V),(\Delta_t)\,\mathcal{F}\text{-adapted}} \Big\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i,u_i) \\
&\quad -\mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V\left(\sqrt{\mathbb{E}^{\mathbb{P}}\left[L\left(\frac{S_2}{S_1}\right) \middle| \mathcal{F}_1 \right]} \right) + \int_0^{T_2} \Delta_t\, dS_t \right] \Big\} \\
&= \inf_{\mathbb{P}\in\mathcal{M}_1} \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Big\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i,u_i) \\
&\quad -\mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V(V) + \int_0^{T_2} \Delta_t\, dS_t + \Delta^L\left(L\left(\frac{S_2}{S_1}\right) - V^2 \right) \right] \Big\} \\
&\overset{\text{(dual)}}{=} \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \inf_{\mathbb{P}\in\mathcal{M}_1} \Big\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i,u_i) \\
&\quad -\mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V(V) + \int_0^{T_2} \Delta_t\, dS_t + \Delta^L\left(L\left(\frac{S_2}{S_1}\right) - V^2 \right) \right] \Big\} \\
&= \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Big\{ \sum_{i\in\{1,2,V\}} (\mu_i,u_i) \\
&\quad -\ln \mathbb{E}^0\left[e^{u_1(S_1)+u_2(S_2)+u_V(V)+\int_0^{T_2}\Delta_t\, dS_t+\Delta^L\left(L\left(\frac{S_2}{S_1}\right)-V^2\right)} \right] \Big\} \\
&= \sup_{u_1,u_2,u_V} \sup_{(\Delta_t)_{t\in[0,T_1]}} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \sup_{(\Delta_t)_{t\in[T_1,T_2]}} \{\cdots\}
\end{aligned}
$$

Continuous time

$$
\begin{aligned}
D &:= \inf_{\mathbb{P}\in\mathcal{P}} H(\mathbb{P},\mathbb{P}_0)\\
&= \inf_{\mathbb{P}\in\mathcal{M}_1} \sup_{u_1\in L^1(\mu_1),u_2\in L^1(\mu_2),u_V\in L^1(\mu_V),(\Delta_t)\,\mathcal{F}\text{-adapted}} \Big\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i,u_i)\\
&\quad -\mathbb{E}^{\mathbb{P}}\left[u_1(S_1)+u_2(S_2)+u_V\left(\sqrt{\mathbb{E}^{\mathbb{P}}\left[L\left(\frac{S_2}{S_1}\right)\Big|\mathcal{F}_1\right]}\right)+\int_0^{T_2}\Delta_t\,dS_t\right]\Big\}\\
&= \inf_{\mathbb{P}\in\mathcal{M}_1} \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Big\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i,u_i)\\
&\quad -\mathbb{E}^{\mathbb{P}}\left[u_1(S_1)+u_2(S_2)+u_V(V)+\int_0^{T_2}\Delta_t\,dS_t+\Delta^L\left(L\left(\frac{S_2}{S_1}\right)-V^2\right)\right]\Big\}\\
&\overset{\text{(dual)}}{=} \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \inf_{\mathbb{P}\in\mathcal{M}_1} \Big\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i,u_i)\\
&\quad -\mathbb{E}^{\mathbb{P}}\left[u_1(S_1)+u_2(S_2)+u_V(V)+\int_0^{T_2}\Delta_t\,dS_t+\Delta^L\left(L\left(\frac{S_2}{S_1}\right)-V^2\right)\right]\Big\}\\
&= \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Big\{ \sum_{i\in\{1,2,V\}} (\mu_i,u_i)\\
&\quad -\ln\mathbb{E}^0\left[e^{u_1(S_1)+u_2(S_2)+u_V(V)+\int_0^{T_2}\Delta_t\,dS_t+\Delta^L\left(L\left(\frac{S_2}{S_1}\right)-V^2\right)}\right]\Big\}\\
&= \sup_{u_1,u_2,u_V} \sup_{(\Delta_t)_{t\in[0,T_1]}} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \sup_{(\Delta_t)_{t\in[T_1,T_2]}} \Big\{\cdots\Big\}
\end{aligned}
$$

Continuous time

$$
\begin{aligned}
D &:= \inf_{\mathbb{P}\in\mathcal{P}} H(\mathbb{P},\mathbb{P}_0) \\
&= \inf_{\mathbb{P}\in\mathcal{M}_1} \sup_{u_1\in L^1(\mu_1),u_2\in L^1(\mu_2),u_V\in L^1(\mu_V),(\Delta_t)\mathcal{F}\text{-adapted}} \Big\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i, u_i) \\
&\quad - \mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V\left(\sqrt{\mathbb{E}^{\mathbb{P}}\left[L\left(\frac{S_2}{S_1}\right)\middle|\mathcal{F}_1\right]}\right) + \int_0^{T_2} \Delta_t \, dS_t \right] \Big\} \\
&= \inf_{\mathbb{P}\in\mathcal{M}_1} \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Big\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i, u_i) \\
&\quad - \mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V(V) + \int_0^{T_2} \Delta_t \, dS_t + \Delta^L\left(L\left(\frac{S_2}{S_1}\right) - V^2\right)\right] \Big\} \\
&\stackrel{\text{(dual)}}{=} \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \inf_{\mathbb{P}\in\mathcal{M}_1} \Big\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i, u_i) \\
&\quad - \mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V(V) + \int_0^{T_2} \Delta_t \, dS_t + \Delta^L\left(L\left(\frac{S_2}{S_1}\right) - V^2\right)\right] \Big\} \\
&= \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Big\{ \sum_{i\in\{1,2,V\}} (\mu_i, u_i) \\
&\quad - \ln \mathbb{E}^0\left[e^{u_1(S_1)+u_2(S_2)+u_V(V)+\int_0^{T_2} \Delta_t \, dS_t + \Delta^L\left(L\left(\frac{S_2}{S_1}\right)-V^2\right)} \right] \Big\} \\
&= \sup_{u_1,u_2,u_V} \sup_{(\Delta_t)_{t\in[0,T_1]}} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \sup_{(\Delta_t)_{t\in[T_1,T_2]}} \{\cdots\}
\end{aligned}
$$

Continuous time

$$
\begin{aligned}
D &:= \inf_{\mathbb{P}\in\mathcal{P}} H(\mathbb{P},\mathbb{P}_0)\\
&= \inf_{\mathbb{P}\in\mathcal{M}_1} \sup_{u_1\in L^1(\mu_1),u_2\in L^1(\mu_2),u_V\in L^1(\mu_V),(\Delta_t)\,\mathcal{F}\text{-adapted}} \Bigg\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i, u_i)\\
&\qquad -\mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V\left(\sqrt{\mathbb{E}^{\mathbb{P}}\left[L\left(\frac{S_2}{S_1}\right)\middle|\mathcal{F}_1\right]}\right) + \int_0^{T_2} \Delta_t\, dS_t \right] \Bigg\}\\
&= \inf_{\mathbb{P}\in\mathcal{M}_1} \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Bigg\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i, u_i)\\
&\qquad -\mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V(V) + \int_0^{T_2} \Delta_t\, dS_t + \Delta^L\left(L\left(\frac{S_2}{S_1}\right) - V^2\right)\right] \Bigg\}\\
&\overset{\text{(dual)}}{=} \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \inf_{\mathbb{P}\in\mathcal{M}_1} \Bigg\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i, u_i)\\
&\qquad -\mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V(V) + \int_0^{T_2} \Delta_t\, dS_t + \Delta^L\left(L\left(\frac{S_2}{S_1}\right) - V^2\right)\right] \Bigg\}\\
&= \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Bigg\{ \sum_{i\in\{1,2,V\}} (\mu_i, u_i)\\
&\qquad - \ln \mathbb{E}^0\left[e^{u_1(S_1)+u_2(S_2)+u_V(V)+\int_0^{T_2} \Delta_t\, dS_t + \Delta^L\left(L\left(\frac{S_2}{S_1}\right)-V^2\right)} \right] \Bigg\}\\
&= \sup_{u_1,u_2,u_V} \sup_{(\Delta_t)_{t\in[0,T_1]}} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \sup_{(\Delta_t)_{t\in[T_1,T_2]}} \left\{ \cdots \right\}
\end{aligned}
$$

Continuous time

$$
\begin{aligned}
D &:= \inf_{\mathbb{P}\in\mathcal{P}} H(\mathbb{P},\mathbb{P}_0) \\
&= \inf_{\mathbb{P}\in\mathcal{M}_1} \sup_{u_1\in L^1(\mu_1),u_2\in L^1(\mu_2),u_V\in L^1(\mu_V),(\Delta_t)\,\mathcal{F}\text{-adapted}} \Big\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i,u_i) \\
&\quad -\mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V\left(\sqrt{\mathbb{E}^{\mathbb{P}}\left[L\left(\frac{S_2}{S_1}\right)\middle|\mathcal{F}_1\right]}\right) + \int_0^{T_2}\Delta_t\, dS_t \right] \Big\} \\
&= \inf_{\mathbb{P}\in\mathcal{M}_1} \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Big\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i,u_i) \\
&\quad -\mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V(V) + \int_0^{T_2}\Delta_t\, dS_t + \Delta^L\left(L\left(\frac{S_2}{S_1}\right) - V^2\right)\right] \Big\} \\
&\overset{\text{(dual)}}{=} \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \inf_{\mathbb{P}\in\mathcal{M}_1} \Big\{ H(\mathbb{P},\mathbb{P}_0) + \sum_{i\in\{1,2,V\}} (\mu_i,u_i) \\
&\quad -\mathbb{E}^{\mathbb{P}}\left[u_1(S_1) + u_2(S_2) + u_V(V) + \int_0^{T_2}\Delta_t\, dS_t + \Delta^L\left(L\left(\frac{S_2}{S_1}\right) - V^2\right)\right] \Big\} \\
&= \sup_{u_1,u_2,u_V,(\Delta_t)} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Big\{ \sum_{i\in\{1,2,V\}} (\mu_i,u_i) \\
&\quad - \ln \mathbb{E}^0\left[e^{u_1(S_1)+u_2(S_2)+u_V(V)+\int_0^{T_2}\Delta_t\, dS_t+\Delta^L\left(L\left(\frac{S_2}{S_1}\right)-V^2\right)} \right] \Big\} \\
&= \sup_{u_1,u_2,u_V} \sup_{(\Delta_t)_{t\in[0,T_1]}} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \sup_{(\Delta_t)_{t\in[T_1,T_2]}} \{\cdots\}
\end{aligned}
$$

Continuous time

- The inner $\inf_{\mathbb{P}\in\mathcal{M}_1}$ is reached at $\mathbb{P}^*$ defined by (renorm. $Z=1$ by cash adjustment of vanilla payoffs)

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

$$\begin{aligned}
D &= \sup_{u_1,u_2,u_V}\ \sup_{(\Delta_t)_{t\in[0,T_1]}}\ \inf_{V\in\mathcal{F}_1}\ \sup_{\Delta^L\in\mathcal{F}_1}\ \sup_{(\Delta_t)_{t\in[T_1,T_2]}}\Bigg\{\sum_{i\in\{1,2,V\}}(\mu_i,u_i) \\
&\qquad - \ln \mathbb{E}^0\left[e^{u_1(S_1)+u_2(S_2)+u_V(V)+\int_0^{T_2}\Delta_t dS_t+\Delta^L\left(L\left(\frac{S_2}{S_1}\right)-V^2\right)}\right]\Bigg\} \\
&= \sup_{u_1,u_2,u_V}\ \sup_{(\Delta_t)_{t\in[0,T_1]}}\ \inf_{V\in\mathcal{F}_1}\ \sup_{\Delta^L\in\mathcal{F}_1}\Bigg\{\sum_{i\in\{1,2,V\}}(\mu_i,u_i) \\
&\qquad - \ln \inf_{(\Delta_t)_{t\in[T_1,T_2]}} \mathbb{E}^0\left[e^{u_1(S_1)+u_2(S_2)+u_V(V)+\int_0^{T_2}\Delta_t dS_t+\Delta^L\left(L\left(\frac{S_2}{S_1}\right)-V^2\right)}\right]\Bigg\}
\end{aligned}$$

Continuous time

$$
\begin{aligned}
D \quad &= \quad \sup_{u_1,u_2,u_V} \sup_{(\Delta_t)_{t\in[0,T_1]}} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Big\{ \sum_{i\in\{1,2,V\}} (\mu_i, u_i) \\
&\quad - \ln \inf_{(\Delta_t)_{t\in[T_1,T_2]}} \mathbb{E}^0 \left[e^{u_1(S_1)+u_2(S_2)+u_V(V)+\int_0^{T_2}\Delta_t dS_t + \Delta^L\left(L\left(\frac{S_2}{S_1}\right)-V^2\right)} \right] \Big\} \\
&\overset{\text{(DPP)}}{=} \quad \sup_{u_1,u_2,u_V} \sup_{(\Delta_t)_{t\in[0,T_1]}} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Big\{ \sum_{i\in\{1,2,V\}} (\mu_i, u_i) \\
&\quad - \ln \mathbb{E}^0 \left[e^{u_1(S_1)+u_V(V)+\int_0^{T_1}\Delta_t dS_t - \Delta^L(L(S_1)+V^2)} U(T_1, S_1, a_1; \Delta^L) \right] \Big\}
\end{aligned}
$$

- Stochastic control:

$$
U(t, S_t, a_t; \Delta^L) := \inf_{(\Delta_r)_{r\in[t,T_2]}} \mathbb{E}^0 \left[e^{u_2(S_2)+\int_t^{T_2}\Delta_r dS_r + \Delta^L L(S_2)} \Big| S_t, a_t, \Delta^L \right], \quad t \in [T_1, T_2].
$$

Continuous time

- U is solution to the HJB PDE

$$\begin{aligned} \partial_t U + \mathcal{L}^0 U + \inf_{\Delta} \left\{ \frac{1}{2} \Delta^2 a^2 s^2 U + \Delta as \left(as\partial_s U + \rho\sigma(a)\partial_a U \right) \right\} &= 0, \\ U(T_2, s, a; \delta^L) &= e^{u_2(s)+\delta^L L(s)}. \end{aligned}$$

- Optimal delta:

$$\Delta_t^* = -\frac{\partial_s U(t, S_t, a_t) + \rho \frac{\sigma(a_t)}{a_t S_t} \partial_a U(t, S_t, a_t)}{U(t, S_t, a_t)},$$

- U satisfies

$$\partial_t U + \mathcal{L}^0 U - \frac{(as\partial_s U + \rho\sigma(a)\partial_a U)^2}{2U} = 0, \qquad U(T_2, s, a; \delta^L) = e^{u_2(s)+\delta^L L(s)}.$$

- $u := \ln U$ satisfies

$$\partial_t u + \mathcal{L}^0 u + \frac{1}{2}(1-\rho^2)\sigma(a)^2(\partial_a u)^2 = 0, \qquad u(T_2, s, a; \delta^L) = u_2(s) + \delta^L L(s).$$

Continuous time

$$
\begin{aligned}
D \quad = \quad & \sup_{u_1,u_2,u_V} \sup_{(\Delta_t)_{t\in[0,T_1]}} \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Bigg\{ \sum_{i\in\{1,2,V\}} (\mu_i, u_i) \\
& - \ln \mathbb{E}^0 \left[e^{u_1(S_1)+u_V(V)+\int_0^{T_1} \Delta_t dS_t - \Delta^L(L(S_1)+V^2)+u(T_1,S_1,a_1;\Delta^L)} \right] \Bigg\}
\end{aligned}
$$

Since S_1, a_1, and $\int_0^{T_1} \Delta_t dS_t$ are $\mathcal{F}_1$-measurable,

$$
\begin{aligned}
& \inf_{V\in\mathcal{F}_1} \sup_{\Delta^L\in\mathcal{F}_1} \Big\{ -\ln \mathbb{E}^0 \left[e^{u_1(S_1)+u_V(V)+\int_0^{T_1} \Delta_t dS_t - \Delta^L(L(S_1)+V^2)+u(T_1,S_1,a_1;\Delta^L)} \right] \\
= \quad & -\ln \sup_{V\in\mathcal{F}_1} \inf_{\Delta^L\in\mathcal{F}_1} \mathbb{E}^0 \left[e^{u_1(S_1)+u_V(V)+\int_0^{T_1} \Delta_t dS_t - \Delta^L(L(S_1)+V^2)+u(T_1,S_1,a_1;\Delta^L)} \right] \\
= \quad & -\ln \mathbb{E}^0 \left[e^{u_1(S_1)+\int_0^{T_1} \Delta_t dS_t + \Phi(S_1,a_1)} \right]
\end{aligned}
$$

$$
\Phi(s,a) := \sup_{v\geq 0} \inf_{\delta^L\in\mathbb{R}} \left\{ u_V(v) - \delta^L(L(s)+v^2) + u(T_1,s,a;\delta^L) \right\}.
$$

The optimal V and Δ^L are functions of (S_1,a_1): $v^*(S_1,a_1)$, $\delta_L^*(S_1,a_1)$.

Continuous time

$$\begin{aligned}
D &= \sup_{u_1,u_2,u_V} \sup_{(\Delta_t)_{t\in[0,T_1]}} \left\{ \sum_{i\in\{1,2,V\}} (\mu_i,u_i) - \ln \mathbb{E}^0\left[e^{u_1(S_1)+\int_0^{T_1}\Delta_t dS_t+\Phi(S_1,a_1)}\right]\right\} \\
&= \sup_{u_1,u_2,u_V} \left\{ \sum_{i\in\{1,2,V\}} (\mu_i,u_i) - \ln \inf_{(\Delta_t)_{t\in[0,T_1]}} \mathbb{E}^0\left[e^{u_1(S_1)+\int_0^{T_1}\Delta_t dS_t+\Phi(S_1,a_1)}\right]\right\} \\
&= \sup_{u_1,u_2,u_V} \left\{ \sum_{i\in\{1,2,V\}} (\mu_i,u_i) - \ln U(0,S_0,a_0)\right\} \\
&= \sup_{u_1,u_2,u_V} \left\{ \sum_{i\in\{1,2,V\}} (\mu_i,u_i) - u(0,S_0,a_0)\right\} =: P
\end{aligned}$$

where $U(t,S_t,a_t) := \inf_{(\Delta_r)_{r\in[t,T_1]}} \mathbb{E}^0\left[e^{u_1(S_1)+\int_t^{T_1}\Delta_r dS_r+\Phi(S_1,a_1)}\middle|S_t,a_t\right]$ satisfies

$$\partial_t U + \mathcal{L}^0 U - \frac{(as\partial_s U + \rho\sigma(a)\partial_a U)^2}{2U} = 0,\ t\in[0,T_1), \quad U(T_1,s,a) = e^{u_1(s)+\Phi(s,a)}$$

and $u := \ln U$ satisfies

$$\partial_t u + \mathcal{L}^0 u + \frac{1}{2}(1-\rho^2)\sigma(a)^2(\partial_a u)^2 = 0,\ t\in[0,T_1), \quad u(T_1,s,a) = u_1(s)+\Phi(s,a).$$

Continuous time

$$\begin{aligned}
D &= \sup_{u_1,u_2,u_V} \sup_{(\Delta_t)_{t\in[0,T_1]}} \left\{ \sum_{i\in\{1,2,V\}} (\mu_i,u_i) - \ln \mathbb{E}^0 \left[e^{u_1(S_1)+\int_0^{T_1} \Delta_t dS_t + \Phi(S_1,a_1)} \right] \right\} \\
&= \sup_{u_1,u_2,u_V} \left\{ \sum_{i\in\{1,2,V\}} (\mu_i,u_i) - \ln \inf_{(\Delta_t)_{t\in[0,T_1]}} \mathbb{E}^0 \left[e^{u_1(S_1)+\int_0^{T_1} \Delta_t dS_t + \Phi(S_1,a_1)} \right] \right\} \\
&= \sup_{u_1,u_2,u_V} \left\{ \sum_{i\in\{1,2,V\}} (\mu_i,u_i) - \ln U(0,S_0,a_0) \right\} \\
&= \sup_{u_1,u_2,u_V} \left\{ \sum_{i\in\{1,2,V\}} (\mu_i,u_i) - u(0,S_0,a_0) \right\} =: P
\end{aligned}$$

where $U(t,S_t,a_t) := \inf_{(\Delta_r)_{r\in[t,T_1]}} \mathbb{E}^0 \left[e^{u_1(S_1)+\int_t^{T_1} \Delta_r dS_r + \Phi(S_1,a_1)} \middle| S_t, a_t \right]$ satisfies

$$\partial_t U + \mathcal{L}^0 U - \frac{(as\partial_s U + \rho\sigma(a)\partial_a U)^2}{2U} = 0,\ t\in[0,T_1), \quad U(T_1,s,a) = e^{u_1(s)+\Phi(s,a)}$$

and $u := \ln U$ satisfies

$$\partial_t u + \mathcal{L}^0 u + \frac{1}{2}(1-\rho^2)\sigma(a)^2(\partial_a u)^2 = 0,\ t\in[0,T_1), \quad u(T_1,s,a) = u_1(s) + \Phi(s,a).$$

Continuous time

$$\begin{aligned}
D &= \sup_{u_1,u_2,u_V} \sup_{(\Delta_t)_{t\in[0,T_1]}} \left\{ \sum_{i\in\{1,2,V\}} (\mu_i,u_i) - \ln \mathbb{E}^0\left[e^{u_1(S_1)+\int_0^{T_1}\Delta_t dS_t+\Phi(S_1,a_1)}\right]\right\} \\
&= \sup_{u_1,u_2,u_V} \left\{ \sum_{i\in\{1,2,V\}} (\mu_i,u_i) - \ln \inf_{(\Delta_t)_{t\in[0,T_1]}} \mathbb{E}^0\left[e^{u_1(S_1)+\int_0^{T_1}\Delta_t dS_t+\Phi(S_1,a_1)}\right]\right\} \\
&= \sup_{u_1,u_2,u_V} \left\{ \sum_{i\in\{1,2,V\}} (\mu_i,u_i) - \ln U(0,S_0,a_0)\right\} \\
&= \sup_{u_1,u_2,u_V} \left\{ \sum_{i\in\{1,2,V\}} (\mu_i,u_i) - u(0,S_0,a_0)\right\} =: P
\end{aligned}$$

where $U(t,S_t,a_t) := \inf_{(\Delta_r)_{r\in[t,T_1]}} \mathbb{E}^0\left[e^{u_1(S_1)+\int_t^{T_1}\Delta_r dS_r+\Phi(S_1,a_1)}\middle| S_t,a_t\right]$ satisfies

$$\partial_t U + \mathcal{L}^0 U - \frac{(as\partial_s U + \rho\sigma(a)\partial_a U)^2}{2U} = 0,\ t\in[0,T_1), \quad U(T_1,s,a) = e^{u_1(s)+\Phi(s,a)}$$

and $u := \ln U$ satisfies

$$\partial_t u + \mathcal{L}^0 u + \frac{1}{2}(1-\rho^2)\sigma(a)^2(\partial_a u)^2 = 0,\ t\in[0,T_1), \quad u(T_1,s,a) = u_1(s)+\Phi(s,a).$$

Continuous time

- Assume $P < +\infty$ and (u_1^*, u_V^*, u_2^*) maximizes P. The probability $\mathbb{P}^*$ that minimizes $H(\mathbb{P}, \mathbb{P}_0)$ satisfies $(Z = 1)$

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

- Let $M_t := \mathbb{E}^0[M_{T_2}|\mathcal{F}_t]$. It is easy to check that $M_t = \mathcal{E}(L)_t$ with

$$dL_t = \sqrt{1-\rho^2}\sigma(a_t)\partial_a u^*(t, S_t, a_t)\, dW_t^{0,\perp}$$

- Girsanov $\Longrightarrow$ $(W^*, W^{*,\perp})$ is a standard $\mathbb{P}^*$-Brownian motion, where

$$W_t^* = W_t^0, \qquad W_t^{*,\perp} = W_t^{0,\perp} - \sqrt{1-\rho^2}\int_0^t \sigma(a_r)\partial_a u^*(r, S_r, a_r)\, dr.$$

- The model dynamics reads

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

Continuous time

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

- In particular, the drift of (a_t) under $\mathbb{P}^*$ is

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

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

- $\mathbb{P}^*\sim\mathbb{P}_0$ and under $\mathbb{P}^*$, $S_1\sim\mu_1$, $S_2\sim\mu_2$, $\sqrt{\mathbb{E}^{\mathbb{P}^*}[L(S_2/S_1)|\mathcal{F}_1]}\sim\mu_V$, and S is an $((\mathcal{F}_t),\mathbb{P})$-martingale.
- If $P=+\infty$, then $\mathcal{P}=\emptyset$.

Acknowledgements

I would like to thank Pierre Henry-Labordère for interesting discussions.

A few selected references

- Acciaio, B., Guyon, J.: *Inversion of Convex Ordering: Local Volatility Does Not Maximize the Price of VIX Futures*, to appear in SIFIN, 2019.
- 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.
- 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.
- 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. Finan. Math.

A few selected references

- Dembo, A., Zeitouni, O.: *Large Deviations Techniques and Applications*, 2nd ed., Springer, New York, 1998.
- Dupire, B.: *Pricing with a smile*, Risk, January, 1994.
- Gatheral, J.: *Consistent modeling of SPX and VIX options*, presentation at Bachelier Congress, 2008.
- 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*, SSRN preprint available at ssrn.com/abstract=3397382, 2019.
- Guyon, J.: *Inversion of Convex Ordering in the VIX Market*, SSRN preprint available at ssrn.com/abstract=3504022, 2019.

A few selected references

- 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.
- Jacquier, A., Martini, C., Muguruza, A.: *On the VIX futures in the rough Bergomi model*, preprint, 2017.
- Johansen, S.: The extremal convex functions. Math. Scand., 34, 61–68 (1974)
- 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.
- Villani, C.: *Topics in optimal transportation*, Graduate studies in Mathematics AMS, Vol 58.

