UNIVERSITY OF OXFORD

Mathematical Institute

OT Methodology for non-parametric calibration & Martinagle Benamou-Brenier problems

Jan Obłój
Mathematical Institute
University of Oxford

based on joint works with
Ivan Guo, Benjamin Joseph, Grégoire Loeper and Leo Wang

CREST Doctoral Course
ENSAE Paris, May 2025

Oxford Mathematics

St John's College Oxford

Calibration

... selection of a model which reproduces market prices.

Parametric: find the best model out of a given family $\mathbb{P}_\theta : \theta \in \Theta$,

$$\inf_{\theta \in \Theta} \left\| \text{MarketPrices} - \text{ModelPrices}(\theta) \right\|$$

Calibration

... selection of a model which reproduces market prices.

Parametric: find the best model out of a given family $\mathbb{P}_\theta : \theta \in \Theta$,

$$\inf_{\theta \in \Theta} \left\| \mathsf{MarketPrices} - \mathsf{ModelPrices}(\theta) \right\|$$

Non-parametric: find a fully calibrated model *close* to a reference one

Calibration

... selection of a model which reproduces market prices.

Parametric: find the best model out of a given family $\mathbb{P}_\theta : \theta \in \Theta$,

$$\inf_{\theta \in \Theta} \Big\| \text{MarketPrices} - \text{ModelPrices}(\theta) \Big\|$$

Non-parametric: find a fully calibrated model *close* to a reference one

- Fix a favourite reference model $\bar{\mathbb{P}}$
- Consider a cost given by

$$J(\mathbb{P}) = \begin{cases} \mathrm{dist}(\mathbb{P}, \bar{\mathbb{P}}) & \text{if } \mathbb{P} \text{ is market calibrated,} \\ +\infty & \text{otherwise.} \end{cases}$$

- Minimise $J(\mathbb{P})$ over all $\mathbb{P}$
  - pioneered by Avellaneda et al. '97 (via relative entropy)
  - here via Stochastic Optimal Transport

Transfer material from one site to another while minimising transportation costs.

- Monge (1781), Kantorovich (1948): Monge-Kantorovich problem
- Benamou & Brenier (2000): continuous-time formulation

Optimal transport, continuous-time formulation

Minimising the cost function F under given initial density ρ_0 and final density ρ_1

$$\inf_{\rho,v} \int_{\mathbb{R}^d} \int_0^1 \rho(t,x) F(v(t,x))\, dtdx,$$

subject to the continuity equation

$$\partial_t \rho(t,x) + \nabla \cdot (\rho(t,x) v(t,x)) = 0,$$

and the initial and final distributions

$$\rho(0,x) = \rho_0, \quad \rho(1,x) = \rho_1.$$

Tan & Touzi (2013) (also Mikami & Thieullen (2006), Huesmann & Trevisan (2017), Backhoff et al. (2017)): Consider probability measures $\mathbb{P}$ such that X is a semimartingale,

$$dX_t = \beta_t^{\mathbb{P}}\, dt + (\alpha_t^{\mathbb{P}})^{1/2}\, dW_t^{\mathbb{P}}.$$

Stochastic optimal transport problem

We want to minimise

$$V(\mu_0, \mu_1) = \inf_{\mathbb{P} \in \mathcal{P}(\mu_0, \mu_1)} \mathbb{E}^{\mathbb{P}} \int_0^1 F(\alpha^{\mathbb{P}}, \beta^{\mathbb{P}})\, dt,$$

where $\mathcal{P}(\mu_0, \mu_1)$ contains probability measures satisfying

$$\mathbb{P} \circ X_0^{-1} = \mu_0, \quad \mathbb{P} \circ X_1^{-1} = \mu_1.$$

Note that the cost function F is convex and may depend on (t, X) as well.

Tan & Touzi (2013) established the following duality result

Dual formulation

The primal problem is equivalent to

$$V(\mu_0,\mu_1) = \sup_{\phi_1} \int \phi_1\, d\mu_1 - \phi_0\, d\mu_0,$$

where

$$\phi_0(x) := \sup_{\mathbb{P}\in\mathcal{P}(\delta_x)} \mathbb{E}^{\mathbb{P}}\left(\phi_1(X_1) - \int_0^1 F(\alpha^{\mathbb{P}},\beta^{\mathbb{P}})\, dt\right).$$

and for $F_t = F(t, X_t, \alpha_t^{\mathbb{P}}, \beta_t^{\mathbb{P}})$ characterised ϕ_0 via PDEs.

Tan & Touzi (2013) established the following duality result

Dual formulation

The primal problem is equivalent to

$$V(\mu_0, \mu_1) = \sup_{\phi_1} \int \phi_1 \, d\mu_1 - \phi_0 \, d\mu_0,$$

where

$$\phi_0(x) := \sup_{\mathbb{P} \in \mathcal{P}(\delta_x)} \mathbb{E}^{\mathbb{P}} \left(\phi_1(X_1) - \int_0^1 F(\alpha^{\mathbb{P}}, \beta^{\mathbb{P}}) \, dt \right).$$

and for $F_t = F(t, X_t, \alpha_t^{\mathbb{P}}, \beta_t^{\mathbb{P}})$ characterised ϕ_0 via PDEs.

Guo and Loeper (2018) extended this to path dependent constraints and cost.
Path-dependent PDEs & functional Itô used to describe the dual.

Stochastic OT & Calibration

SOT induces a projection onto a subset of (semi)-martingales.

Use for calibration:

- Gather market data $\mathcal{G}$
- Fix a favourite reference model $\bar{\mathbb{P}}$
- Consider a cost F given by

$$F(\mathbb{P}) = \begin{cases} \mathrm{dist}(\mathbb{P}, \bar{\mathbb{P}}) & \text{if } \mathbb{P} \text{ is calibrated to } \mathcal{G}, \\ +\infty & \text{otherwise.} \end{cases}$$

- ensuring convexity to get duality
- Solve the dual via a non-linear (P)PDE
- $\mathbb{P}^*$ recovered via $\nabla F^*(\ldots)$.

SPX & VIX Calibration

w/ I. Guo, G. Loeper, and S. Wang

SPX and VIX

- S&P 500 Index (SPX): a stock market index that measures the stock performance of 500 large companies listed in the US stock market.
- CBOE Volatility Index (VIX): a volatility index that measures the market's expectation of the volatility of SPX over the following 30 days.

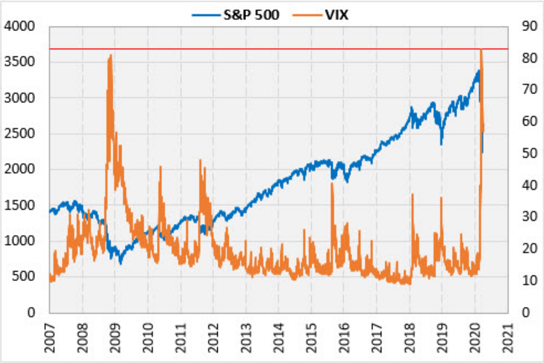


Figure: Historical SPX and VIX data. (Source: Schaeffer's Investment Research)

Why joint calibration?

- VIX futures and options are very popular hedging instruments.
  e.g., Szado (2009) shows that VIX call options are better than S&P 500 put options as a hedging instrument against the financial crisis in 2008.
- An arbitrage argument (Guyon 2020): existence of a liquid market
  $\Rightarrow$ need for models that jointly calibrate to the option prices of SPX and VIX
  $\Rightarrow$ avoid arbitrage between financial institutions (or even within the same institution)
- Joint calibration problem: build a (stochastic volatility) model that jointly calibrates to the prices of SPX options, VIX futures and VIX options.
- Very challenging problem, especially for short maturities.

Selected previous works

Previous works:

- Continuous-time diffusion models (without jump):
  - Gatheral (2008): double CEV model
  - Goutte–Ismail–Pham (2017): Regime-switching Heston model
  - Fouque–Saporito (2018): Heston stochastic vol-of-vol
- Continuous-time jump-diffusion models: many works including
  - Cont–Kokholm (2013), Lian–Zhu (2013), Baldeaux–Badran (2014), Kokholm–Stisen (2015), Pacati–Pompa–Reno (2018), ...

However, even with jumps, these models have yet to achieve an exact fit.

Recent works:

- Guyon (2020): nonparametric discrete-time model calibrated by martingale optimal transport
- Gatheral–Jusselin–Rosenbaum (2020): (parametric) quadratic rough Heston model (no efficient calibration method yet)
- $\Rightarrow$ This work: nonparametric continuous-time model calibrated by semimartingale optimal transport

Assumption: zero interest rates & dividends.

Let S_t be the SPX price:

$$S_t = S_0 + \int_0^t \sigma_s S_s \, dW_s.$$

Consider a time grid $0 < t_0 < t_1 < \cdots < t_n = T$ and an annualisation factor AF, e.g., if t_i corresponds to daily observations, then $AF = 100^2 \times 252/n$.

The *realised variance* of S_t during $[t_0, T]$:

$$AF \sum_{i=1}^{n} \left(\log \frac{S_{t_i}}{S_{t_i - 1}} \right)^2 \to \frac{100^2}{T - t_0} \int_{t_0}^{T} \sigma_t^2 \, dt, \quad a.s.$$

The VIX index at t_0:

$$VIX(t_0, T) = \sqrt{\mathbb{E}\left(\frac{100^2}{T - t_0} \int_{t_0}^{T} \sigma_t^2 \, dt \,\middle|\, \mathcal{F}_{t_0} \right)}$$

Modelling — Standard approach

Underlying assets:

$$S_t = S_0 + \int_0^t \sigma_s S_s \, dW_s$$

$$VIX(t_0, T) = \sqrt{\mathbb{E}\left(\frac{100^2}{T - t_0} \int_{t_0}^T \sigma_t^2 \, dt \,\middle|\, \mathcal{F}_{t_0} \right)}$$

Calibrating instruments:

$$\begin{aligned}
\text{SPX calls:} \quad & u^{SPX,c} = \mathbb{E}((S_T - K)^+) \\
\text{SPX puts:} \quad & u^{SPX,p} = \mathbb{E}((K - S_T)^+) \\
\text{VIX futures:} \quad & u^{VIX,f} = \mathbb{E}(VIX_{t_0}) \\
\text{VIX calls:} \quad & u^{VIX,c} = \mathbb{E}((VIX_{t_0} - K)^+) \\
\text{VIX puts:} \quad & u^{VIX,p} = \mathbb{E}((K - VIX_{t_0})^+)
\end{aligned}$$

Many previous works involve modelling (S_t, σ_t) or (S_t, σ_t^2)
$\Rightarrow$ the term VIX is a square root of conditional expectation
$\Rightarrow$ numerically difficult to compute the prices of VIX futures and VIX options.

Modelling — Our approach

Consider a two dimensional stochastic process $X = (X^1, X^2)$, let X^1 be the logarithm of S_t:

$$X_t^1 := \log S_t = X_0^1 - \frac{1}{2}\int_0^t \sigma_s^2\, ds + \int_0^t \sigma_s\, dW_s.$$

Let X^2 be a half of the expected forward quadratic variation of X^1 over $[t, T]$ observed at t:

$$X_t^2 = \mathbb{E}\left(\frac{1}{2}\int_t^T \sigma_s^2\, ds \,\middle|\, \mathcal{F}_t\right).$$

Calibrating instruments: for $\tau \leq T$,

$$\begin{aligned}
\text{SPX calls:}\quad & u^{SPX,c} = \mathbb{E}((\exp(X_\tau^1) - K)^+) =: \mathbb{E}(G^{SPX,c}(X_\tau)) \\
\text{SPX puts:}\quad & u^{SPX,p} = \mathbb{E}((K - \exp(X_\tau^1))^+) =: \mathbb{E}(G^{SPX,p}(X_\tau)) \\
\text{VIX futures:}\quad & u^{VIX,f} = \mathbb{E}(100\sqrt{2X_{t_0}^2/(T-t_0)}) =: \mathbb{E}(G^{VIX,f}(X_{t_0})) \\
\text{VIX calls:}\quad & u^{VIX,c} = \mathbb{E}((100\sqrt{2X_{t_0}^2/(T-t_0)} - K)^+) =: \mathbb{E}(G^{VIX,c}(X_{t_0})) \\
\text{VIX puts:}\quad & u^{VIX,p} = \mathbb{E}((K - 100\sqrt{2X_{t_0}^2/(T-t_0)})^+) =: \mathbb{E}(G^{VIX,p}(X_{t_0}))
\end{aligned}$$

All payoffs depend on only the marginal distributions of X at fixed times
$\Rightarrow$ suitable for the calibration framework via optimal transport.

The Heston model:

$$\begin{aligned} dS_t &= \sqrt{\nu_t} S_t \, dW_t^1, \\ d\nu_t &= -\kappa(\nu_t - \theta)\, dt + \omega \sqrt{\nu_t}\, dW_t^2, \\ \langle dW^1, dW^2 \rangle_t &= \eta \, dt. \end{aligned}$$

We can derive that

$$X_t^2 = \mathbb{E}\left(\frac{1}{2} \int_t^T \nu_s \, ds \,\middle|\, \mathcal{F}_t \right) = \frac{1 - e^{-\kappa(T-t)}}{2\kappa} (\nu_t - \theta) + \frac{1}{2}\theta(T - t).$$

Define $A(t, \kappa) := (1 - e^{-\kappa(T-t)})/\kappa$ and $\nu(t, X_t^2, \kappa, \theta) := A(t, \kappa)^{-1}(2X_t^2 - \theta(T - t)) + \theta$, then the Heston model in terms of (X^1, X^2) is

$$\begin{aligned} dX_t^1 &= -\frac{1}{2}\nu(t, X_t^2, \kappa, \theta)\, dt + \sqrt{\nu(t, X_t^2, \kappa, \theta)}\, dW_t^1, \\ dX_t^2 &= -\frac{1}{2}\nu(t, X_t^2, \kappa, \theta)\, dt + \frac{1}{2} A(t, \kappa)\omega \sqrt{\nu(t, X_t^2, \kappa, \theta)}\, dW_t^2, \\ \langle dW_t^1, dW_t^2 \rangle &= \eta \, dt. \end{aligned}$$

Framework — Semimartingale optimal transport

Consider probability measures $\mathbb{P}$ under which X is a semimartingale:

$$dX_t = \alpha_t^{\mathbb{P}}\, dt + (\beta_t^{\mathbb{P}})^{\frac{1}{2}}\, dW_t^{\mathbb{P}}.$$

Semimartingale optimal transport with discrete constraints

Minimise

$$\inf_{\mathbb{P}\in\mathcal{P}(X_0,\tau,G,c)} \mathbb{E}^{\mathbb{P}} \int_0^T F(\alpha_t^{\mathbb{P}}, \beta_t^{\mathbb{P}})\, dt,$$

where $\mathcal{P}(X_0,\tau,G,c)$ contains probability measures $\mathbb{P}$ satisfying

$$\mathbb{P}\circ X_0^{-1} = \delta_{X_0} \quad \text{and} \quad \mathbb{E}^{\mathbb{P}} G_i(X_{\tau_i}) = c_i, \quad i = 1,\dots,m.$$

Note that the cost function F is convex in $(\alpha^{\mathbb{P}}, \beta^{\mathbb{P}})$. It may depend on (t, X) as well.

The cost function plays a regularisation role to ensure that X has the correct dynamics.

We want X to have the following dynamics:

$$X_t^1 = X_0^1 - \frac{1}{2}\int_0^t \sigma_s^2\, ds + \int_0^t \sigma_s\, dW_s, \quad X_t^2 = \mathbb{E}\left(\frac{1}{2}\int_t^T \sigma_s^2\, ds \,\middle|\, \mathcal{F}_t\right).$$

The above dynamics can be captured by $\mathbb{P}$ such that

$$(\alpha_t^{\mathbb{P}}, \beta_t^{\mathbb{P}}) = \left(\begin{bmatrix} -\frac{1}{2}\sigma_t^2 \\ -\frac{1}{2}\sigma_t^2 \end{bmatrix}, \begin{bmatrix} \sigma_t^2 & (\beta_t)_{12} \\ (\beta_t)_{12} & (\beta_t)_{22} \end{bmatrix}\right), \quad 0 \le t \le T,$$

where $(\beta_t)_{12} = d\langle X^1, X^2\rangle_t \,/\, dt$ and $(\beta_t)_{22} = d\langle X^2\rangle_t \,/\, dt$ and with the additional property that $X_T^2 = 0$ $\mathbb{P}$-a.s.

Given $\bar{\beta}$, a reference for β, define the cost function:

$$F(\alpha,\beta) = \begin{cases} \sum_{i,j=1}^{2} (\beta_{ij} - \bar{\beta}_{ij})^2 & \text{if } \alpha_1 = \alpha_2 = -\frac{1}{2}\beta_{11}, \\ +\infty & \text{otherwise.} \end{cases}$$

Framework — Discrete constraints

The additional property $X_T^2 = 0$, $\mathbb{P}$-a.s. and the prices of calibrating instruments are imposed on X as discrete constraints $\Rightarrow$ exact calibration

We want to calibrate X to:

- m number of SPX options with payoffs $G = (G_1, \ldots, G_m)$, maturities $\tau \in (0, T]^m$ and prices $u^{SPX} \in \mathbb{R}_+^m$, e.g.,

$$\mathbb{E}^{\mathbb{P}} G_i(X_{\tau_i}) = u_i^{SPX}, \quad i = 1, \ldots, m,$$

- a VIX futures with payoff $J(x) = 100\sqrt{2x_2/(T - t_0)}$, maturity t_0 and price $u^{VIX,f} \in \mathbb{R}$, e.g.,

$$\mathbb{E}^{\mathbb{P}} J(X_{t_0}) = u^{VIX,f},$$

- n number of VIX options with payoffs $H = (H_1, \ldots, H_n)$, maturity t_0 and prices $u^{VIX} \in \mathbb{R}_+^m$, e.g.,

$$\mathbb{E}^{\mathbb{P}} (H_i \circ J)(X_{t_0}) = u_i^{VIX}, \quad i = 1, \ldots, n,$$

- a contract with payoff $\xi(x) = 1 - \exp(-(x_2)^2)$, maturity T and zero price, e.g.,

$$\mathbb{E}^{\mathbb{P}} \xi(X_T) = 0.$$

The last calibrating instrument ensures that $X_T^2 = 0$, $\mathbb{P}$-a.s. Since its price is always zero, we call it a *singular contract*.

Framework — Reformulation of the joint calibration problem

For simplicity, we represent all the discrete constraints by

$$\mathbb{E}^{\mathbb{P}}\mathcal{G}_i(X_{\mathcal{T}_i}) = c_i, \quad i = 1, \ldots, m+n+2,$$

where

$$\mathcal{G} = (\underbrace{G_1, \ldots, G_m}_{m \text{ SPX options}}, \underbrace{H_1 \circ J, \ldots, H_n \circ J}_{n \text{ VIX options}}, \underbrace{J}_{\text{VIX futures}}, \underbrace{\xi}_{\text{singular contract}}),$$

and $\mathcal{T}$ and c are defined in a similar manner.

Define a set of the probability measures $\mathcal{P}_{joint}$ such that

$$\mathcal{P}_{joint} := \{\mathbb{P} : \ \mathbb{P} \circ X_0^{-1} = \delta_{X_0} \text{ and } \mathbb{E}^{\mathbb{P}}\mathcal{G}_i(X_{\mathcal{T}_i}) = c_i, \ i = 1, \ldots, m+n+2\}$$

The joint calibration problem

Minimise $$V := \inf_{\mathbb{P} \in \mathcal{P}_{joint}} \mathbb{E}^{\mathbb{P}} \int_0^T F(\alpha_t^{\mathbb{P}}, \beta_t^{\mathbb{P}})\, dt.$$

If we find an optimal solution $\tilde{\mathbb{P}}$ and $V < +\infty$, then we have a well-calibrated model

$$X_t = X_0 + \int_0^t \alpha_s^{\tilde{\mathbb{P}}}\, ds + \int_0^t (\beta_s^{\tilde{\mathbb{P}}})^{\frac{1}{2}}\, dW_s^{\mathbb{P}}.$$

Framework — Markovian projection

Markovian projection: use a (Markovian) diffusion process mimic an Itô process by matching its marginals at fixed times. (Gyöngy (1986) and Brunick–Shreve (2013))

Lemma (Figalli (2008) and Trevisan (2016))

Let $\rho_t^{\mathbb{P}} = \mathbb{P} \circ X_t^{-1}$ be the marginal distribution of X_t under $\mathbb{P}$, $t \leq T$, then $\rho^{\mathbb{P}}$ is a weak solution to the Fokker–Planck equation:

$$\begin{cases} \partial_t \rho_t^{\mathbb{P}} + \nabla_x \cdot (\rho_t^{\mathbb{P}} \mathbb{E}_{t,x}^{\mathbb{P}} \alpha_t^{\mathbb{P}}) - \frac{1}{2} \sum_{i,j} \partial_{ij} (\rho_t^{\mathbb{P}} (\mathbb{E}_{t,x}^{\mathbb{P}} \beta_t^{\mathbb{P}})_{ij}) = 0 & \text{in } [0,T] \times \mathbb{R}^2, \\ \rho_0^{\mathbb{P}} = \delta_{X_0} & \text{in } \mathbb{R}^2. \end{cases}$$

Moreover, there exists another probability measure $\mathbb{P}'$ under which X has the same marginals, $\rho^{\mathbb{P}'} = \rho^{\mathbb{P}}$, and is a Markov process solving

$$dX_t = \alpha^{\mathbb{P}'}(t, X_t)dt + (\beta^{\mathbb{P}'}(t, X_t))^{\frac{1}{2}} \, dW_t^{\mathbb{P}'}, \quad 0 \leq t \leq T,$$

where $W^{\mathbb{P}'}$ is a $\mathbb{P}'$-Brownian motion, $\alpha^{\mathbb{P}'}(t,x) = \mathbb{E}_{t,x}^{\mathbb{P}} \alpha_t^{\mathbb{P}}$ and $\beta^{\mathbb{P}'}(t,x) = \mathbb{E}_{t,x}^{\mathbb{P}} \beta_t^{\mathbb{P}}$.

Notation: $\mathbb{E}_{t,x}^{\mathbb{P}} = \mathbb{E}^{\mathbb{P}}(\cdot \mid X_t = x)$.

Let $\mathcal{P}^{loc}_{joint}$ be a subset of $\mathcal{P}_{joint}$ such that, under any $\mathbb{P} \in \mathcal{P}^{loc}_{joint}$, X is a Markov process that solves

$$dX_t = \alpha^{\mathbb{P}}(t, X_t)dt + (\beta^{\mathbb{P}}(t, X_t))^{\frac{1}{2}}\, dW^{\mathbb{P}}_t, \quad 0 \le t \le T,$$

and X is fully calibrated to the calibrating instruments.

Proposition

$$V = \inf_{\mathbb{P}\in\mathcal{P}_{joint}} \mathbb{E}^{\mathbb{P}} \int_0^T F(\alpha^{\mathbb{P}}_t, \beta^{\mathbb{P}}_t)\, dt = \inf_{\mathbb{P}\in\mathcal{P}^{loc}_{joint}} \mathbb{E}^{\mathbb{P}} \int_0^T F(\alpha^{\mathbb{P}}_t(t, X_t), \beta^{\mathbb{P}}_t(t, X_t))\, dt$$

Proof: "$\geq$" follows by convexity of F via Jensen's inequality:

$$\begin{aligned}\mathbb{E}^{\mathbb{P}} \int_0^T F(\alpha^{\mathbb{P}}_t, \beta^{\mathbb{P}}_t)\, dt &= \mathbb{E}^{\mathbb{P}} \int_0^T \left(\mathbb{E}^{\mathbb{P}}_{t,x} F(\alpha^{\mathbb{P}}_t, \beta^{\mathbb{P}}_t)\right)\, dt \\ &\geq \mathbb{E}^{\mathbb{P}} \int_0^T F(\mathbb{E}^{\mathbb{P}}_{t,x}\alpha^{\mathbb{P}}_t, \mathbb{E}^{\mathbb{P}}_{t,x}\beta^{\mathbb{P}}_t)\, dt.\end{aligned}$$

"$\leq$" is clear since $\mathcal{P}^{loc}_{joint} \subset \mathcal{P}_{joint}$.

Framework — PDE formulation

The problem can be made convex by introducing $A = \rho\alpha$ and $B = \rho\beta$, since

$$\rho F(\alpha, \beta) = \rho F\left(\frac{A}{\rho}, \frac{B}{\rho}\right) = \sup_{r + F^*(a,b) \leq 0} \{\rho r + A \cdot a + B : b\},$$

is convex in (ρ, A, B), where $F^*(a, b) = \sup_{\alpha,\beta}\{a \cdot \alpha + b : \beta - F(\alpha, \beta)\}$ is the convex conjugate of F, and $B : b = \mathrm{Tr}(Bb)$.

PDE formulation

Minimise

$$V = \inf_{\rho, A, B} \int_0^T \int_{\mathbb{R}^2} \rho F(A/\rho, B/\rho)\, dxdt,$$

subject to constraints

$$\partial_t \rho + \nabla_x \cdot A - \frac{1}{2} \sum_{i,j} \partial_{ij} B_{ij} = 0,$$

$$\int_{\mathbb{R}^2} \mathcal{G}_i \rho(t, \cdot)\, dx = c_i, \quad i = 1, \ldots, m + n + 2$$

$$\rho(0, \cdot) = \delta_{X_0}.$$

Framework — Duality

Introducing Lagrange multipliers $\phi \in C_c^\infty([0,T] \times \mathbb{R}^2)$ and $\lambda \in \mathbb{R}^{m+n+2}$, the problem can be formulated as:

$$V = \inf_{\rho,A,B} \sup_{\phi,\lambda} \Big\{ \int_0^T \int_{\mathbb{R}^2} \left(\rho F\left(\frac{A}{\rho}, \frac{B}{\rho}\right) - (\partial_t \phi \rho + \nabla_x \phi \cdot A + \frac{1}{2} \nabla_x^2 \phi : B) - \sum_{i=1}^{m+n+2} \lambda_i \mathcal{G}_i \delta(t - \mathcal{T}_i)\rho \right) dx dt + \lambda \cdot c - \phi(0, X_0) \Big\}$$

$$= \sup_{\phi,\lambda} \inf_{\rho,A,B} \Big\{ \underbrace{\int_0^T \int_{\mathbb{R}^2} \left(\rho F\left(\frac{A}{\rho}, \frac{B}{\rho}\right) \right.}_{\text{objective of the primal}} - (\partial_t \phi \rho + \nabla_x \phi \cdot A + \frac{1}{2} \nabla_x^2 \phi : B) - \left. \sum_{i=1}^{m+n+2} \lambda_i \mathcal{G}_i \delta(t - \mathcal{T}_i)\rho \right) dx dt + \underbrace{\lambda \cdot c - \phi(0, X_0)}_{\text{objective of the dual}} \Big\}$$

The interchange of $\inf$ and $\sup$ can be formally established by the Fenchel–Rockafellar duality theorem.

By applying the Fenchel–Rockafellar duality theorem and a smoothing technique:

Dual formulation

Maximise

$$V = \sup_{\lambda \in \mathbb{R}^{m+n+2}} \lambda \cdot c - \phi(0, X_0),$$

where ϕ is the viscosity solution to the HJB equation:

$$\partial_t \phi + F^*(\nabla_x \phi, \frac{1}{2}\nabla_x^2 \phi) = - \sum_{i=1}^{m+n+2} \lambda_i \mathcal{G}_i \delta(t - \mathcal{T}_i),$$

with the terminal condition $\phi(T, \cdot) = 0$. If the supremum is attained and the associated solution to the HJB equation is $\tilde{\phi} \in BV([0,T], C_b^2(\mathbb{R}^2))$, then an optimal (α, β) of the PDE formulation can be found by

$$(\alpha, \beta) = \nabla F^*(\nabla_x \tilde{\phi}, \frac{1}{2}\nabla_x^2 \tilde{\phi}).$$

Note: $F^*(a, b) = \sup_{\alpha,\beta}\{a \cdot \alpha + b : \beta - F(\alpha, \beta)\}$ is the convex conjugate of F.

Framework — Gradient

Given $\lambda \in \mathbb{R}^{m+n+2}$ with the associated solution ϕ^λ, let $\mathbb{P}(\lambda)$ be the probability measure under which X has $(\alpha, \beta) = (\alpha^\lambda, \beta^\lambda) := \nabla F^*(\nabla_x \phi^\lambda, \frac{1}{2}\nabla_x^2 \phi^\lambda)$.

Define

$$L(\lambda) := \lambda \cdot c - \phi^\lambda(0, X_0).$$

The gradients of the objective can be formulated as the difference between the market prices and the model prices:

$$\partial_{\lambda_i} L(\lambda) = \underbrace{c_i}_{\text{market price}} - \underbrace{\mathbb{E}^{\mathbb{P}(\lambda)} \mathcal{G}_i(X_{\mathcal{T}_i})}_{\text{model price}}, \quad i = 1, \dots, m.$$

The model price $\mathbb{E}^{\mathbb{P}(\lambda)} \mathcal{G}_i(X_{\mathcal{T}_i}) = \phi'(0, X_0)$ where ϕ' satisfies

$$\begin{cases} \partial_t \phi' + \alpha^\lambda \cdot \nabla_x \phi' + \dfrac{1}{2} \beta^\lambda : \nabla_x^2 \phi' = 0, & \text{in } [0, \mathcal{T}_i) \times \mathbb{R}^2, \\ \phi'(\mathcal{T}_i, \cdot) = \mathcal{G}_i. \end{cases}$$

Note: For the calculation of different gradients, the PDEs are the same but with different terminal conditions. The inversion of the linear operator is only required once for all gradients.

Dual formulation:

$$\text{maximise} \quad V = \sup_{\lambda \in \mathbb{R}^{m+n+2}} \lambda \cdot c - \phi^{\lambda}(0, X_0),$$

$$\text{subject to} \quad \partial_t \phi^{\lambda} + F^*(\nabla_x \phi^{\lambda}, \frac{1}{2}\nabla_x^2 \phi^{\lambda}) = -\sum_{i=1}^{m+n+2} \lambda_i \mathcal{G}_i \delta(t - \mathcal{T}_i), \quad \phi(T, \cdot) = 0.$$

Numerical solution:

1. Set an initial λ (e.g., $\lambda = \mathbf{0}$),
2. Solve the HJB equation backward to get $\phi^{\lambda}(0, X_0)$ (see next slide),
3. Solve the linear PDEs and calculate all gradients,
4. Update λ by gradient descent.

HJB: $$\partial_t\phi+\sup_{\alpha,\beta}\{\alpha\cdot\nabla_x\phi+\frac{1}{2}\beta:\nabla_x^2\phi-F(\alpha,\beta)\}=-\sum_{i=1}^{m+n+2}\lambda_i\mathcal{G}_i\delta(t-\mathcal{T}_i),\quad \phi(T,\cdot)=0$$

Algorithm 1: Solving the HJB equation

for $k=N-1,\dots,0$ **do**

```
/* Handling the source term */
```

$\phi_{t_{k+1}}\leftarrow\phi_{t_{k+1}}+\sum_{i=1}^{m+n+2}\lambda_i\mathcal{G}_i\mathbb{1}(t_{k+1}=\mathcal{T}_i)$

```
/* Policy iteration */
```

$\phi_{t_k}^{new}\leftarrow\phi_{t_{k+1}}$

do

$\phi_{t_k}^{old}\leftarrow\phi_{t_k}^{new}$

Approximate the optimal $(\alpha_{t_k},\beta_{t_k})$ by solving the supremum with $\phi_{t_k}^{old}$

Solve the linearised HJB equation with $(\alpha_{t_k},\beta_{t_k})$ by a fully implicit finite difference method, and set the solution to $\phi_{t_k}^{new}$

while $\|\phi_{t_k}^{new}-\phi_{t_k}^{old}\|_\infty>\epsilon$

$\phi_{t_k}\leftarrow\phi_{t_k}^{new}$

end

Scaling the discrete constraints with proper scales might improve the stability and convergence.

$$\mathbb{E}^{\mathbb{P}} \hat{\mathcal{G}}(X_T) := \mathbb{E}^{\mathbb{P}} \frac{1}{\Gamma} \mathcal{G}(X_T) = \frac{c}{\Gamma} =: \hat{c}$$

Recommended values of Γ:

- for SPX and VIX options, set Γ to their Black–Scholes Vega
  $\Rightarrow$ 1e-4 error of $\hat{c} \approx$ 1 bp error in implied vol,
- for VIX futures, set $\Gamma = 100$
  $\Rightarrow$ 1e-4 error of $\hat{c} \approx$ 1 cent error in price.

Numerical method — Smoothing

So far we have ignored the significance of the reference model $\bar{\beta}$.
When the gaps between strikes are too large or $\bar{\beta}$ is too far away from the β that describes the actual market dynamics, there might be spikes in the volatility surfaces, which might cause hump-shaped model volatility skews.

Smoothing technique:

1. Set an initial reference $\bar{\beta}$
2. Solve the dual formulation to get an optimal $\beta = \beta^*$
3. Smooth β^* by a smoothing method and set the result to $\bar{\beta}$
4. Repeat steps 2-4 with the new $\bar{\beta}$

In the numerical example, we smooth β^* by the simple moving average method over (X^1, X^2) with bandwidths of $(3, 3)$.

Simulated data example — Generating model

Simulated calibrating instruments:

- SPX call options maturing at 44 days and 79 days
- VIX futures maturing at 49 days
- VIX call options maturing at 49 days

Prices of the above instruments are *generated* using *Heston dynamics* and parameters $(\kappa, \theta, \omega, \eta) = (0.6, 0.09, 0.4, -0.5)$, i.e., X satisfies

$$X_t = X_0 + \int_0^t \alpha_s^{\mathbb{P}} \, ds + \int_0^t (\beta_s^{\mathbb{P}})^{\frac{1}{2}} \, dW_s^{\mathbb{P}},$$

and

$$(\alpha_t^{\mathbb{P}}, \beta_t^{\mathbb{P}}) = \left(\begin{bmatrix} -\frac{1}{2}\nu(t, X_t^2, \kappa, \theta) \\ -\frac{1}{2}\nu(t, X_t^2, \kappa, \theta) \end{bmatrix}, \begin{bmatrix} \nu(t, X_t^2, \kappa, \theta) & \frac{1}{2}\eta\omega A(t,\kappa)\nu(t, X_t^2, \kappa, \theta) \\ \frac{1}{2}\eta\omega A(t,\kappa)\nu(t, X_t^2, \kappa, \theta) & \frac{1}{4}\omega^2 A(t,\kappa)^2 \nu(t, X_t^2, \kappa, \theta) \end{bmatrix} \right),$$

where $A(t,\kappa) := (1 - e^{-\kappa(T-t)})/\kappa$ and $\nu(t, X_t^2, \kappa, \theta) := A(t,\kappa)^{-1}(2X_t^2 - \theta(T-t)) + \theta$.

$\Rightarrow$ Solution exists!

Simulated data example — OT-Model

Recall our joint calibration problem is

$$\inf_{\mathbb{P}\in\mathcal{P}^{loc}_{joint}} \mathbb{E}^{\mathbb{P}} \int_0^T F(\alpha^{\mathbb{P}}_t, \beta^{\mathbb{P}}_t)\, dt, \quad \text{where } F(\alpha,\beta) = \begin{cases} \sum_{i,j=1}^{2} (\beta_{ij} - \bar{\beta}_{ij})^2 & \text{if } \alpha_1 = \alpha_2 = -\frac{1}{2}\beta_{11}, \\ +\infty & \text{otherwise.} \end{cases}$$

We consider two references:

(a) a Heston reference with parameters $(\bar{\kappa}, \bar{\theta}, \bar{\omega}, \bar{\eta}) = (0.9, 0.04, 0.6, -0.3)$:

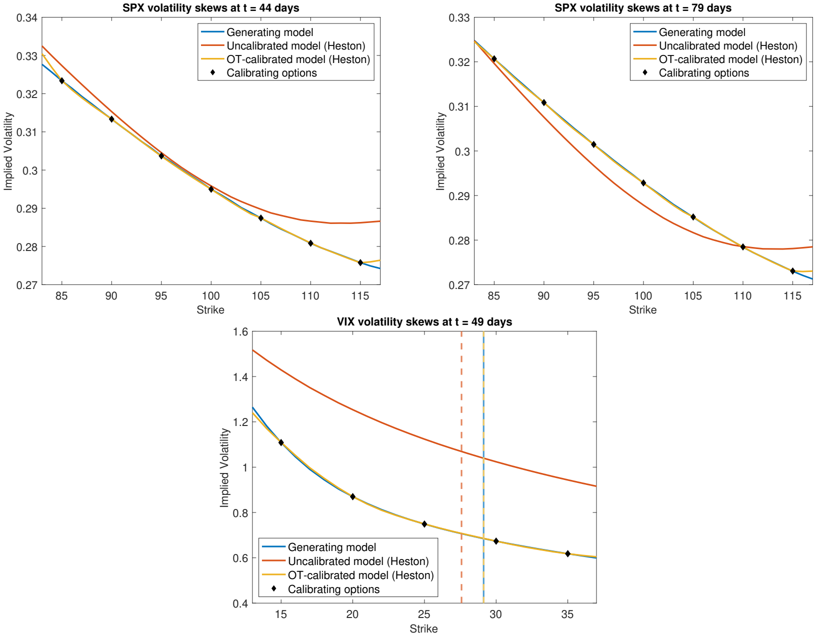
$$\bar{\beta}(t, X^1_t, X^2_t) = \begin{bmatrix} \nu(t, X^2_t, \bar{\kappa}, \bar{\theta}) & \frac{1}{2}\bar{\eta}\bar{\omega}A(t,\bar{\kappa})\nu(t, X^2_t, \bar{\kappa}, \bar{\theta}) \\ \frac{1}{2}\bar{\eta}\bar{\omega}A(t,\bar{\kappa})\nu(t, X^2_t, \bar{\kappa}, \bar{\theta}) & \frac{1}{4}\bar{\omega}^2 A(t,\bar{\kappa})^2\nu(t, X^2_t, \bar{\kappa}, \bar{\theta}) \end{bmatrix};$$

(b) a constant reference:

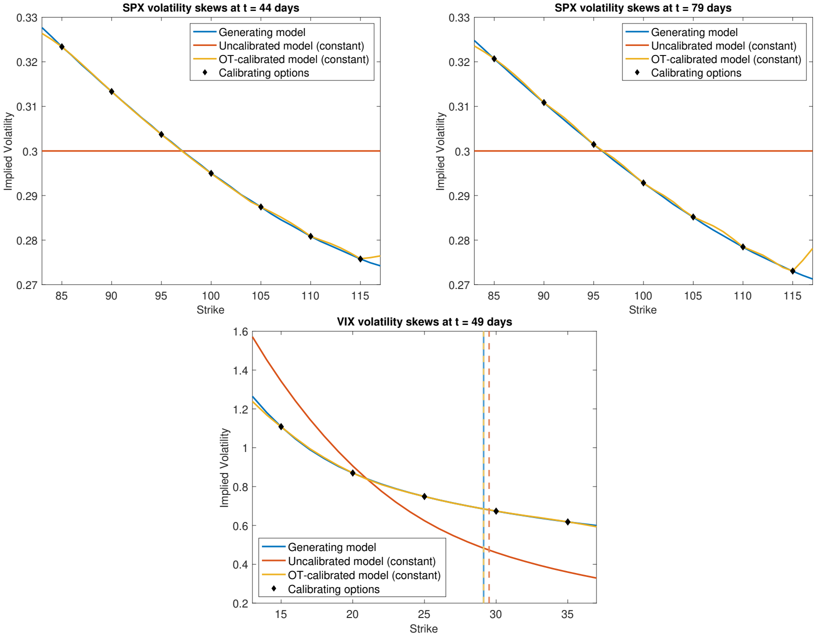
$$\bar{\beta}(t, X^1_t, X^2_t) = \begin{bmatrix} 0.09 & -0.01 \\ -0.01 & 0.04 \end{bmatrix}.$$

Rk: if in (a) we took the reference to be the generating model, $(\bar{\kappa}, \bar{\theta}, \bar{\omega}, \bar{\eta}) = (\kappa, \theta, \omega, \eta)$, then the algorithm quickly recovers OT-model = generating model by $\lambda = \mathbf{0}$, and $V = 0$.

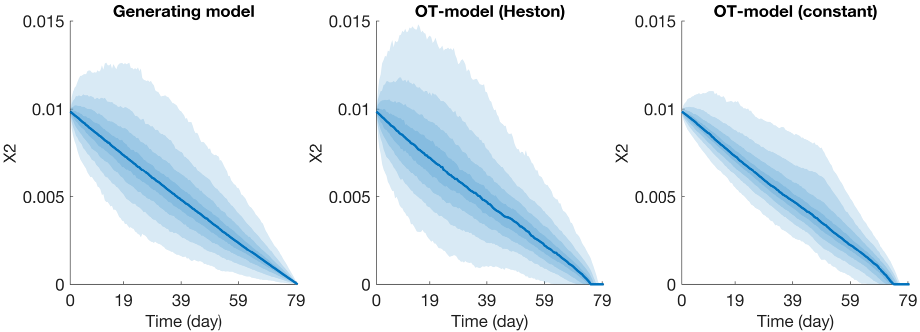
Simulated data example — Calibration results for Heston reference



Simulated data example — Calibrating results for constant reference



Simulated data example — Simulation of X^2



Market data example

Market data as of 1st September 2020:

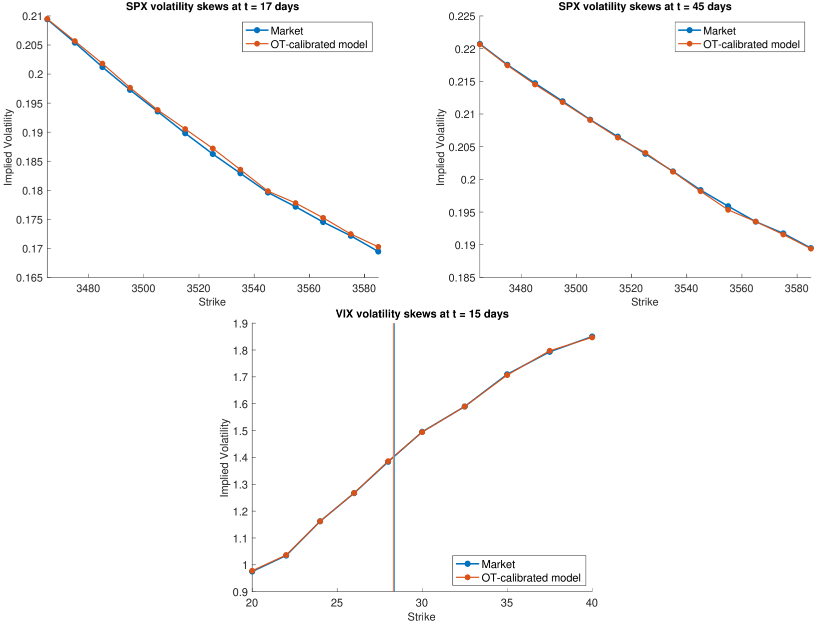
- SPX call options maturing at 17 days and 45 days
- VIX futures maturing at 15 days
- VIX call option maturing at 15 days

These are the shortest maturities, which is known as the most challenging case!

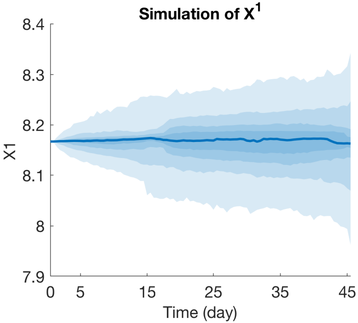
We calibrate the OT-model with a Heston reference $\bar{\beta}$. The parameters $(\bar{\kappa}, \bar{\theta}, \bar{\omega}, \bar{\eta}) = (4.99, 0.038, 0.52, -0.99)$ are obtained by (roughly) calibrating a standard Heston model to the SPX option prices.

Remark. Interest rates and dividends are NOT zero
$\Rightarrow$ model X^1 as the log of T-forward SPX price (instead of the spot price)
$\Rightarrow$ $\mathbb{P}$ are T-forward measures under which $\exp(X^1)$ is still a martingale.

Market data example — Calibration results



Market data example — Simulation of X^1 and X^2



SPX & Interest Rates Calibration

w/ B. Joseph and G. Loeper

Case I: A pre-calibrated short rate model fitting the term structure, zero dividends

Take a two dimensional stochastic process $X = (X^1, X^2)$, let X^1 log-stock price of some underlying asset and X^2 represent the short rate

$$X_t^1 = X_0^1 + X_t^2 - \frac{1}{2}\int_0^t \sigma_s^2 \,\mathrm{d}s + \int_0^t \sigma_s \,\mathrm{d}W_s^1,$$

we assume that X^2 is a Hull-White short rate process given by

$$X_t^2 = X_0^2 + \int_0^t (\theta(s) - a(s)X_s^2)\,\mathrm{d}s + \int_0^t \sigma_r(s)\,\mathrm{d}W_s^2.$$

We assume that W_t^1 and W_t^2 are correlated standard Brownian motions such that

$$\langle W_\cdot^1, W_\cdot^2 \rangle_t = \int_0^t \xi_s \,\mathrm{d}s.$$

Note that since r_t is assumed to be pre-calibrated, the parameters θ, a, and σ_r are all assumed to be known. We calibrate σ and ξ using Call options on the underlying at 60 and 120 days.

Given n Call options observed in the market with prices u_i, strikes K_i and maturities τ_i, our calibration constraints become

$$\mathbb{E}\left[e^{-\int_0^{\tau_i} X_s^2\,\mathrm{d}s}\left(e^{X_{\tau_i}^1} - K_i\right)\right] = u_i, \qquad i = 1, \ldots, n.$$

We therefore consider the set $\mathcal{P}(X_0, \tau, K, u)$ containing measures $\mathbb{P}$ such that X is a semimartingale and satisfies the calibration constraints.
Moreover, we may localise using Markovian projection and consider the subset $\mathcal{P}_{\text{loc}}(X_0, \tau, K, u) \subset \mathcal{P}(X_0, \tau, K, u)$ such that under the mimicking measure $\mathbb{P}' \in \mathcal{P}_{\text{loc}}(X_0, \tau, K, u)$, X is a Markov process satisfying

$$\mathrm{d}X_t = \alpha(t, X_t)\mathrm{d}t + (\beta(t, X_t))^{\frac{1}{2}}\mathrm{d}W_t,$$

where W is a $\mathbb{P}'$ Brownian motion.

The discount term $e^{-\int_0^{\tau_i} X_s^2 \, \mathrm{d}s}$ is path dependent and thus incompatible with our PDE formulation framework.
We could add an extra state variable, but that would increase the computational complexity when solving the HJB equation, so we provide a conditioning argument.

Discounted Density Transformation

Let $\bar{\rho}$ be the joint law of X_t^1, X_t^2 and $\int_0^t X_s^2 \, \mathrm{d}s$ and $\eta_{t,x}(y)$ the law of $\int_0^t X_s^2 \, \mathrm{d}s$ conditional on $X_t = [x^1, x^2]^\intercal$.
Define the 'discounted density' $\tilde{\rho}(t,x) = \left(\int_{\mathbb{R}} e^{-y} \eta_{t,x}(\mathrm{d}y)\right) \rho(t,x)$, $(t,x) \in [0,T] \times \mathbb{R}^2$.
Then $\tilde{\rho}$ satisfies for $(t,x) \in [0,T] \times \mathbb{R}^2$:

$$\partial_t \tilde{\rho}(t,x) + \nabla_x \cdot (\alpha(t,x)\tilde{\rho}(t,x)) - \frac{1}{2}\nabla_x^2 : (\beta(t,x)\tilde{\rho}(t,x)) + x_2 \tilde{\rho}(t,x) = 0.$$

A modification of the superposition principle allows us to reduce to Markov models in (t, X_t^1).

Framework — PDE Formulation

Primal Problem

Minimise

$$V = \inf_{\rho, A, B} \int_0^T \int_{\mathbb{R}^2} \rho F\left(\frac{A}{\rho}, \frac{B}{\rho}\right) \mathrm{d}x\mathrm{d}t,$$

subject to the constraints

$$\partial_t \rho + \nabla_x \cdot A - \frac{1}{2}\nabla^2 : B + x_2 \rho = 0$$

$$\int_{\mathbb{R}^2} \left(e^{x_1} - K_i\right)^+ \rho(\tau_i, \mathrm{d}x) = u_i, \qquad i = 1, \ldots, n$$

$$\rho(0, \cdot) = \delta_{X_0}$$

Framework — Dual Problem

Introduce the Lagrange multipliers $\phi \in C_c^\infty([0,T] \times \mathbb{R}^2)$ and $\lambda \in \mathbb{R}^n$, then

$$V = \inf_{\rho,A,B} \sup_{\phi,\lambda} \left\{ \int_0^T \int_{\mathbb{R}^2} \left(\rho F\left(\frac{A}{\rho}, \frac{B}{\rho}\right) - \left(\partial_t \phi \rho + \nabla_x \phi \cdot A + \frac{1}{2} \nabla_x^2 \phi : B - x_2 \phi \rho \right) \right. \right.$$
$$\left. \left. - \sum_{i=1}^n \lambda_i (e^{x_1} - K_i)^+ \delta_{\tau_i} \rho \right) \mathrm{d}x\mathrm{d}t + \lambda \cdot u - \phi(0, X_0) \right\}$$

Dual Problem

Maximise

$$V = \sup_{\lambda \in \mathbb{R}^n} \lambda \cdot u - \phi(0, X_0),$$

where ϕ is the viscosity solution to the HJB equation:

$$\partial_t \phi - x_2 \phi + F^*(\nabla_x \phi, \frac{1}{2} \nabla_x^2 \phi) + \sum_{i=1}^n \lambda_i (e_1^x - K_i)^+ \delta_{\tau_i} = 0$$

with the terminal condition $\phi(T, \cdot) = 0$. If the supremum is attained and the associated solution to the HJB equation is $\tilde{\phi} \in \mathrm{BV}([0,T], C_b^2(\mathbb{R}^2))$, then an optimal (α, β) of the PDE formulation can be found by

$$(\alpha, \beta) = \nabla F^*(\nabla_x \tilde{\phi}, \frac{1}{2} \nabla_x^2 \tilde{\phi}).$$

Choose a reference correlation $\bar{\xi}(t)$ and require $\xi(t, Z_t, r_t) = \frac{\sigma_r(t)}{\sigma(t, Z_t, r_t)}\bar{\xi}(t)$, for $t \in [0, T]$.
Define for $p > 1$

$$H(x, \bar{x}, s) = \begin{cases} (p-1)\left(\frac{x-s}{\bar{x}-s}\right)^{1+p} + (p+1)\left(\frac{x-s}{\bar{x}-s}\right)^{1-p} - 2p, & \text{if } x, \bar{x} > s, \\ +\infty, & \text{otherwise.} \end{cases}$$

Notice that the coefficients are chosen such that H is minimised over x at $x = \bar{x}$ with $\min H = 0$. Also define the convex set

$$\Gamma(t, X_t) = \Big\{(\alpha, \beta) \in \mathbb{R}^2 \times \mathbb{S}^2 \ : \ \alpha_1 = X_t^2 - \frac{1}{2}\beta_{11},\ \alpha_2 = (b(t) - aX_t^2),$$
$$\beta_{12} = \beta_{21} = \bar{\xi}\sigma_r(t),\ \beta_{22} = \sigma_r^2\Big\}$$

Define the cost function $F(\alpha, \beta) = \begin{cases} H(\beta_{11}, \bar{\sigma}^2, \bar{\xi}^2\sigma_r^2), & \text{if } (\alpha, \beta) \in \Gamma(t, X_t), \\ +\infty, & \text{otherwise.} \end{cases}$

$\bar{\sigma}^2 = \bar{\sigma}^2(t, X_t)$ is some reference value for the volatility

HJB Equation and Model Prices

HJB Equation

$$\sum_{i=1}^{n} \lambda_i (\exp(x_1) - K_i)^+ \delta_{\tau_i} + \partial_t \phi + \sup_{\beta_{11}} \Bigg(\left(x_2 - \frac{1}{2}\beta_{11} \right) \partial_{x_1} \phi$$
$$+ (b(t) - a x_2) \partial_{x_2} \phi + \frac{1}{2} \beta_{11} \partial^2_{x_1 x_1} \phi + \bar{\xi} \sigma_r \partial^2_{x_1 x_2} \phi + \frac{1}{2} \sigma_r^2 \partial^2_{x_2 x_2} \phi - x_2 \phi$$
$$- H(\beta_{11}, \bar{\sigma}^2, \bar{\xi}^2 \sigma_r^2) \Bigg) = 0, \quad (t, x) \in [0, T] \times \mathbb{R}^2.$$

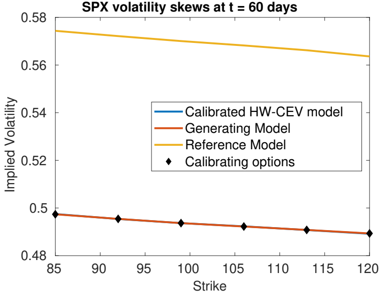
Given λ with associated solution $\mathbb{P}^\lambda$ of the dual problem, let $\mathbb{P}(\lambda)$ be the probability measure under which X has the characteristics $(\alpha^\lambda, \beta^\lambda) = \nabla F^* (\nabla_x \phi^\lambda, \frac{1}{2} \nabla^2_x \phi^\lambda)$. Then the model price of an instrument with payoff $\mathcal{G}$ and maturity $\mathcal{T}$ is given by $\mathbb{E}^{\mathbb{P}(\lambda)} \left[e^{-\int_0^{\mathcal{T}} X_s^2 \, \mathrm{d}s} \mathcal{G}(X_{\mathcal{T}}) \right] = \phi'(0, X_0)$, where ϕ' solves

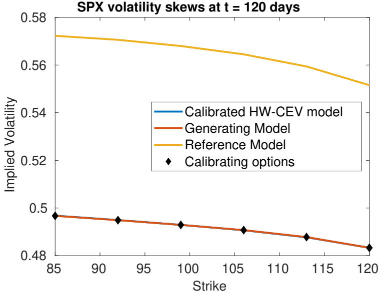
$$\begin{cases} \partial_t \phi + \alpha^\lambda \cdot \nabla_x \phi' + \frac{1}{2} \beta^\lambda : \nabla^2_x \phi' - x_2 \phi' = 0, & (t, x) \in [0, \mathcal{T}) \times \mathbb{R}^2 \\ \phi'(\mathcal{T}, \cdot) = \mathcal{G}(\cdot) \end{cases}$$

The numerical method is analogous in this case, and we may analytically compute the optimal β_{11} in the HJB equation with our chosen cost function.

Simulated Data Example

We used a CEV-Hull-White reference and generating model with the interest rate parameters the same in both. This gave us that $\bar{\sigma}(t,x) = \sigma \exp(x_1)^{\gamma-1}$. The generating model had parameters $(\sigma, \gamma, a, \sigma_r, \xi) = (0.78, 0.9, 0.4, 0.005, -0.6)$, and the "good" reference had $(\bar{\sigma}, \bar{\gamma}, \bar{\xi}) = (0.9, 0.9, -0.4)$, whereas the "bad" reference had $(\bar{\sigma}, \bar{\gamma}, \bar{\xi}) = (1.2, 0.78, 0.4)$





Simulated Data Example — Plots of Characteristics

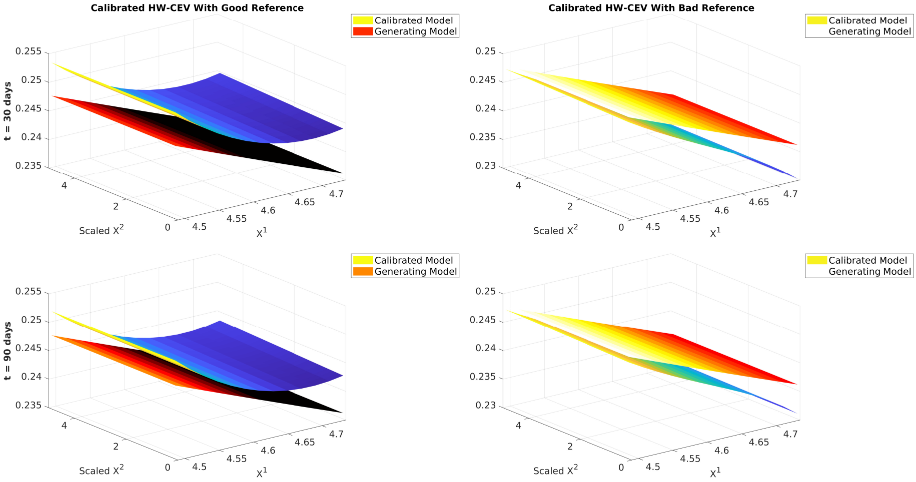


Figure: Comparison of β_{11} with the generating vol surface for a ‘good’ and a ‘bad’ reference model

Simulated Data Example — Plots of Characteristics

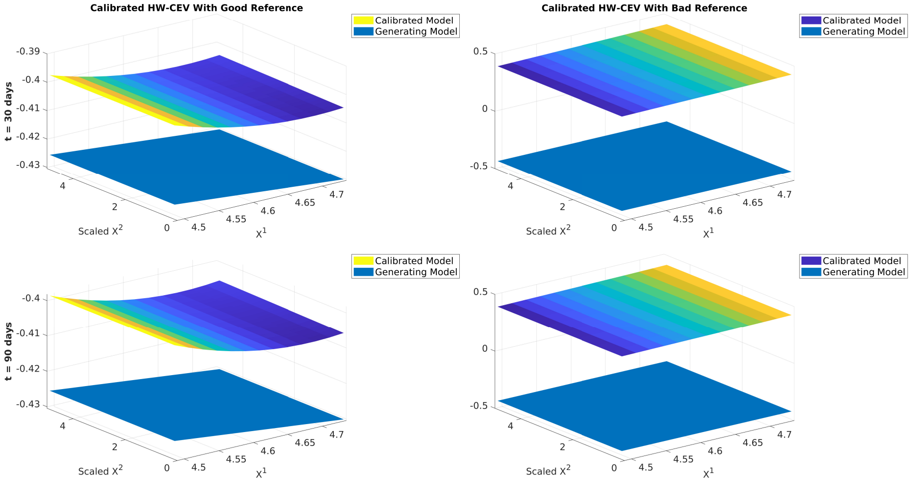


Figure: Comparison of ξ with the generating vol surface for a 'good' and a 'bad' reference model

Cost Function for Joint Calibration

Case II: Joint & simultaneous calibration exercise, zero dividends

Now assume we have no prior knowledge of the interest rate, our characteristics for the log-stock and short rate are therefore given by:

$$\alpha_t = \begin{bmatrix} X_t^2 - \frac{1}{2}(\beta_t)_{11} \\ (\alpha_t)_2 \end{bmatrix}, \qquad \beta_t = \begin{bmatrix} (\beta_t)_{11} & (\beta_t)_{12} \\ (\beta_t)_{12} & (\beta_t)_{22} \end{bmatrix}.$$

Define the convex set

$$\Gamma(t,x) = \left\{ (\alpha,\beta) \in \mathbb{R}^2 \times \mathbb{S}_+^2 \ : \ \alpha_1 = x_2 - \frac{1}{2}\beta_{11} \right\}.$$

Define the cost function

$$F(\alpha,\beta) = \begin{cases} ||\alpha - \overline{\alpha}||_2^2 + ||\beta - \overline{\beta}||_{\mathrm{Fro}}^2, & \text{if } (\alpha,\beta) \in \Gamma(t,x), \\ +\infty, & \text{otherwise.} \end{cases}$$

Where $\bar{\alpha}$ and $\bar{\beta}$ correspond to some reference model. We remark that we will calibrate with interest rate derivatives as well.

Joint Calibration HJB Equation

The dual formulation is similar to the sequential calibration, but with a different cost function. Let $G_i(x)$ denote the payoffs of instruments with maturity τ_i and market value u_i.

Joint Calibration Dual Formulation

Maximise

$$V = \sup_{\lambda \in \mathbb{R}^n} \lambda \cdot u - \phi(0, X_0)$$

Subject to

$$\partial_t \phi + \sup_{\alpha_2 \in \mathbb{R}, \beta \in \mathbb{S}^2_+} \Big\{ \left(x_2 - \frac{1}{2}\beta_{11} \right) \partial_{x_1}\phi + \alpha_2 \partial_{x_2}\phi + \frac{1}{2}\beta_{11}\partial^2_{x_1 x_1}\phi$$

$$+ \frac{1}{2}\beta_{22}\partial^2_{x_2 x_2}\phi + \beta_{12}\partial^2_{x_1 x_2}\phi - ||\alpha - \overline{\alpha}||_2^2 - ||\beta - \overline{\beta}||^2_{\mathrm{Fro}} \Big\}$$

$$- x_2\phi + \sum_{i=1}^{n} \lambda_i G_i(x)\delta_{\tau_i} = 0, \qquad \text{for } (t, x) \in [0, T] \times \mathbb{R}^2$$

Joint Calibration Numerical Method

The numerical method is identical to the sequential calibration method, and we can analytically compute the supremum in the HJB equation.

We calibrate using Call options on the stock and Caplets on the interest rate with a fixed notional of \$1,000 at 60 and 120 days.

The reference models are the CEV local volatility model with a Hull-White interest rate and a CIR interest rate. In both cases, the generating model was the same with shifted parameters. The parameters were given as follows:

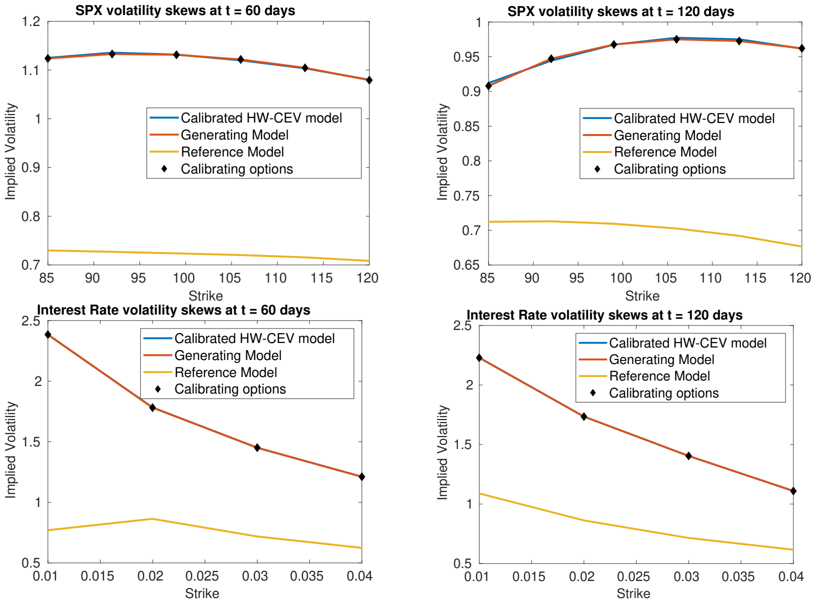
Generating		Reference	
σ	1.50	$\overline{\sigma}$	1.2
γ	0.95	$\overline{\gamma}$	0.89
a	0.05	$\overline{a}$	0.03
σ_r	0.04	$\overline{\sigma}_r$	0.02
ρ	-0.05	$\overline{\rho}$	-0.2

Table: CEV-Hull-White Parameters

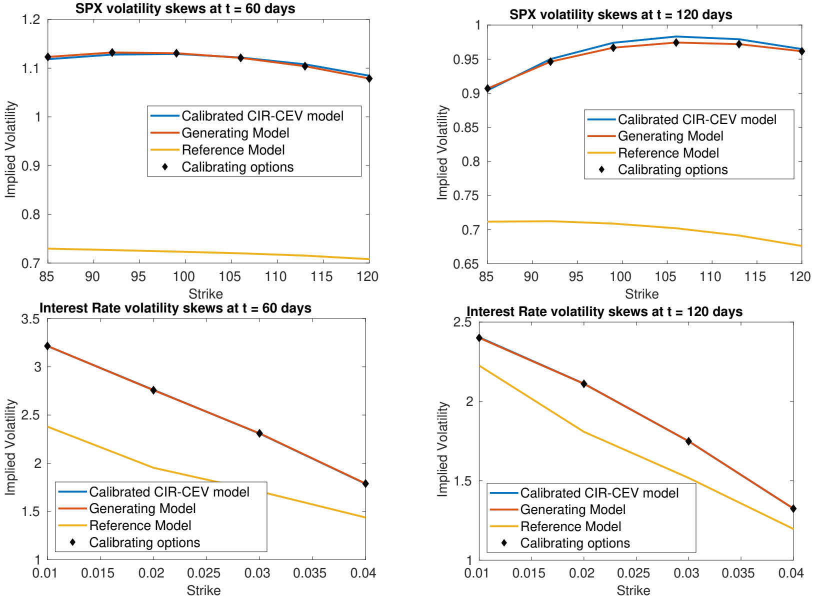
Generating		Reference	
σ	1.5	$\overline{\sigma}$	1.2
γ	0.95	$\overline{\gamma}$	0.89
b	0.03	$\overline{b}$	0.03
a	0.5	$\overline{a}$	0.4
σ_r	0.5	$\overline{\sigma}_r$	0.3
ρ	-0.4	$\overline{\rho}$	-0.2

Table: CEV-CIR Parameters

Simulated Data Example — CEV-HW



Simulated Data Example — CEV-CIR



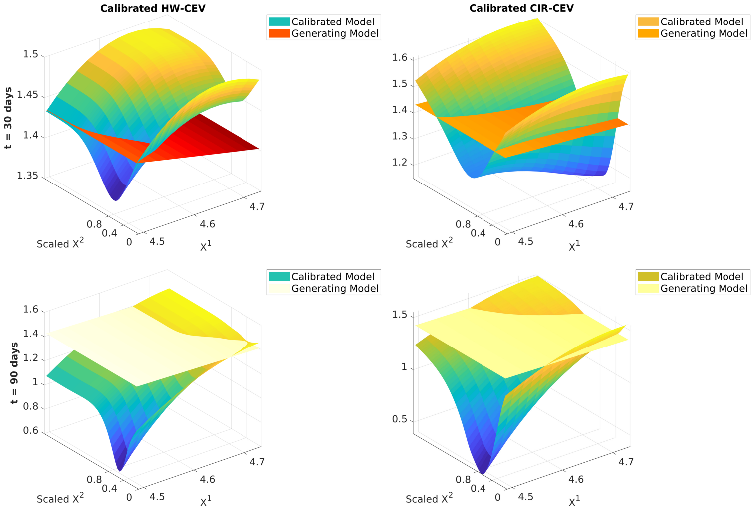


Figure: Compatison of β_{11} for the calibrated and generating model

Simulated Data Example — Plots of Characteristics

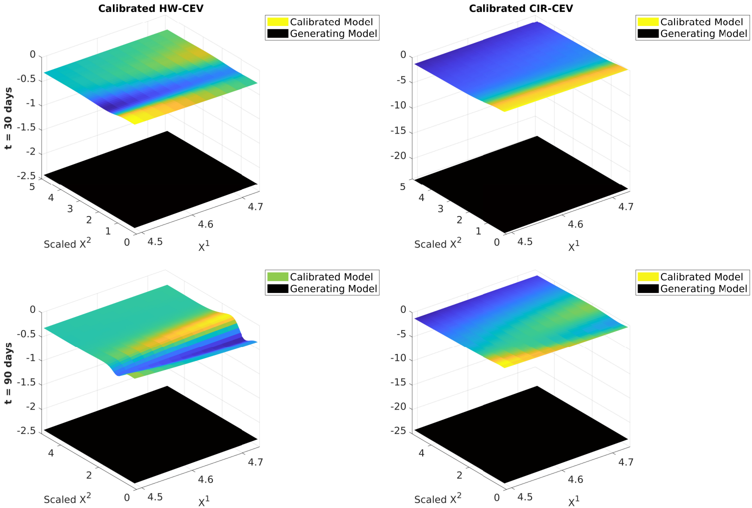


Figure: Comparison of β_{12} for the calibrated and generating model

Simulated Data Example — Plots of Characteristics

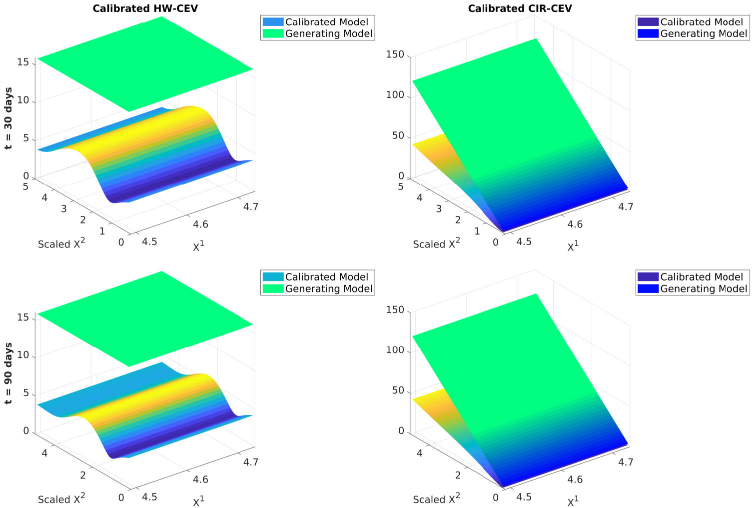


Figure: Comparison of β_{22} for the calibrated and generating model

Simulated Data Example — Plots of Characteristics

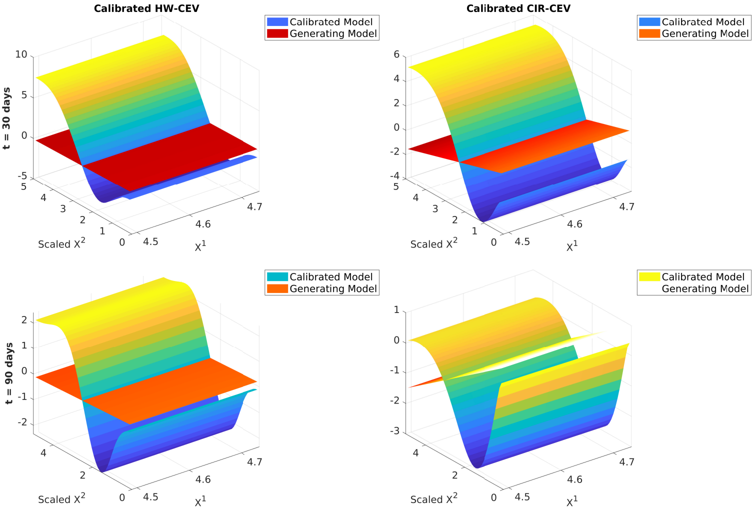


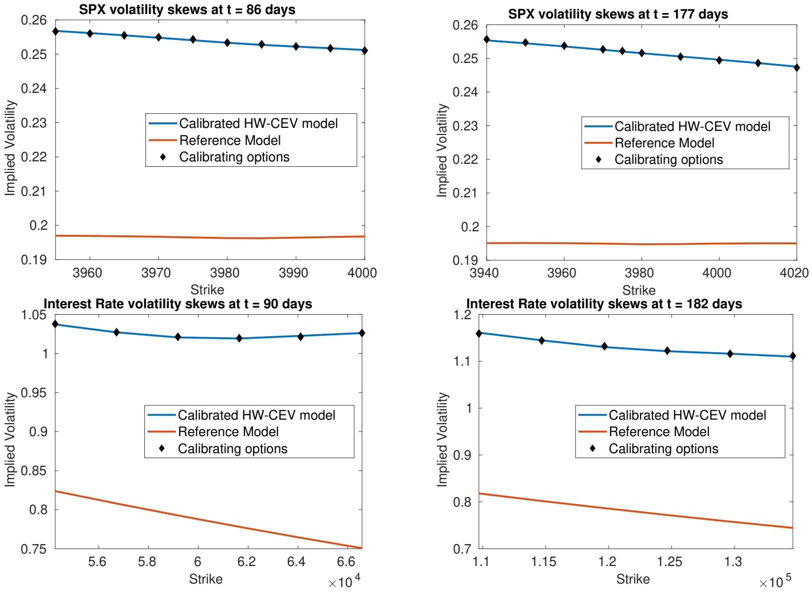
Figure: Comparison of α_2 for the calibrated and generating model

The numerical method is identical to the sequential calibration method, and we can analytically compute the supremum in the HJB equation. We took the SPX as the underlying and the 1M US LIBOR for a proxy of the short rate. We obtained the following data on 23/05/2022 from a Bloomberg terminal:

- Calls on the SPX with expiry 19/08/2022,
- Caps on the one month LIBOR with notional $10,000,000 and expiry 23/08/2022,
- Calls on the SPX with expiry 18/11/2022,
- Caps on the one month LIBOR with notional $10,000,000 and expiry 23/11/2022.

We additionally took a CEV-Hull-White reference model with parameters $(\bar{\sigma}, \gamma, \bar{a}, \bar{\sigma}_r, \bar{\rho}) = (0.3, 0.95, 0.01, 0.02, -0.7)$

Market Data Example — CEV-HW



Market Data Example — Plots of Characteristics

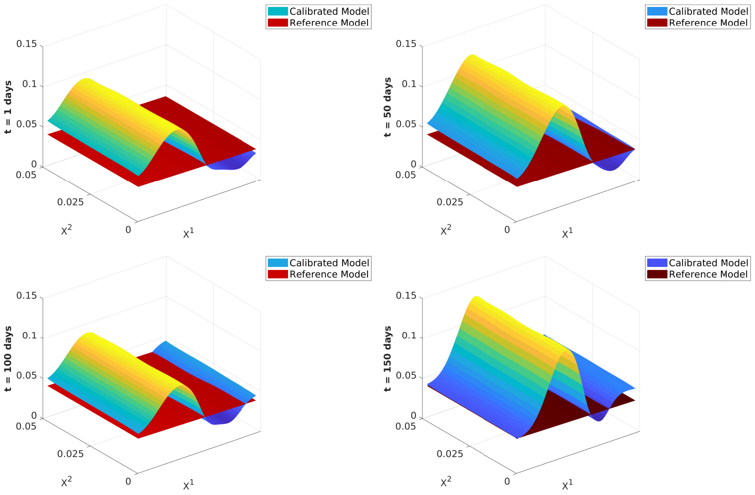


Figure: Compatison of $\beta_{11} = \sigma_X^2$ for the calibrated and generating model

Market Data Example — Plots of Characteristics

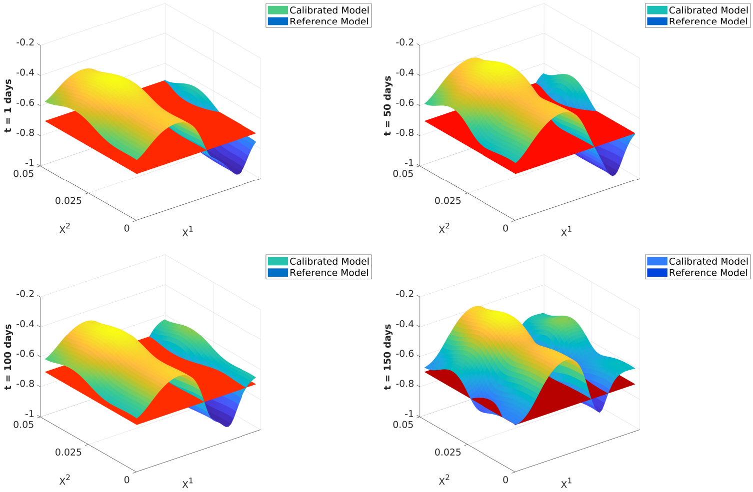


Figure: Comparison of $\beta_{12} = \rho$ for the calibrated and generating model

Market Data Example — Plots of Characteristics

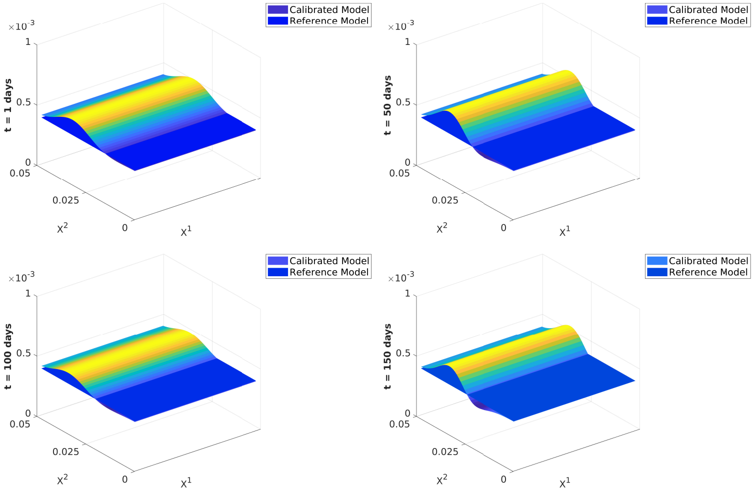


Figure: Comparison of $\beta_{22} = \sigma_r^2$ for the calibrated and generating model

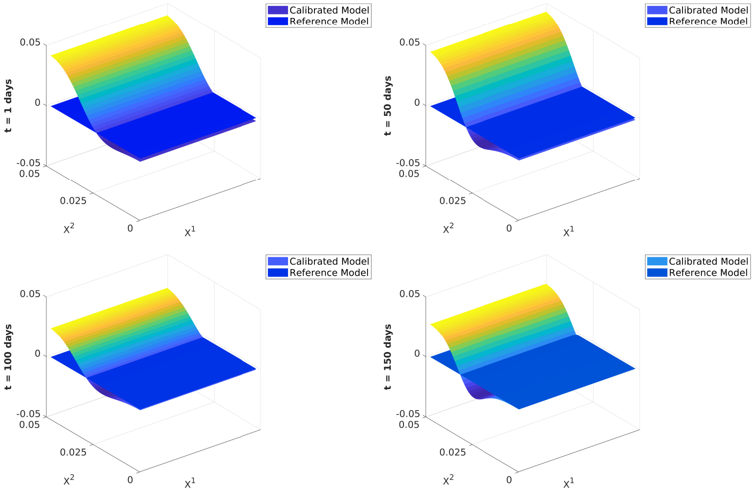


Figure: Comparison of $\alpha_2 = \mu_r$ for the calibrated and generating model

SOT – BB Formulation

Analytic formulae and some magic

based on joint work J. Backhoff, B. Joseph and G. Loeper
see *arXiv: 2406.04016 and 2310.13797*

We want to solve

$$\mathbf{GmBB}_{\mu_0,\mu_1} = \inf_{\substack{S_0\sim\mu_0, S_1\sim\mu_1 \\ S_t = S_0 + \int_0^t \sigma_u S_u dB_u}} \mathbb{E}\left[\int_0^1 (\sigma_t - \bar{\sigma})^2 dt\right], \qquad \text{(G-mBB)}$$

That is, we want to find

- a calibrated model,
- which is the closest to the $\bar{\sigma}$-Black-Scholes model.

(M)OT Motivation

The celebrated Benamou-Brenier reformulation of the classical OT problem is:

$$\inf_{\substack{X_0 \sim \nu_0, X_1 \sim \nu_1 \\ X_t = X_0 + \int_0^t V_s \, ds}} \mathbb{E}\left[\int_0^1 |V_t|^2 dt\right],$$

i.e., we look for a particle with velocity as close as possible to a constant one, with given initial and terminal distributions.

(M)OT Motivation

The celebrated Benamou-Brenier reformulation of the classical OT problem is:

$$\inf_{\substack{X_0\sim\nu_0, X_1\sim\nu_1 \\ X_t = X_0 + \int_0^t V_s\,ds}} \mathbb{E}\left[\int_0^1 |V_t|^2 dt\right],$$

i.e., we look for a particle with velocity as close as possible to a constant one, with given initial and terminal distributions.

More recently, Backhoff et al. '20 and Huesmann & Trevisan '19, considered the martingale analogue of this problem:

$$\mathbf{AmBB}_{\nu_0,\nu_1} = \inf_{\substack{M_0\sim\nu_0, M_1\sim\nu_1 \\ M_t = M_0 + \int_0^t \Sigma_s\,dB_s}} \mathbb{E}\left[\int_0^1 (\Sigma_t - \bar{\Sigma})^2 dt\right], \qquad \text{(A-mBB)}$$

where the optimisation is taken over filtered probability spaces with a Brownian motion $(B_t)_{t\geq 0}$, possibly starting from a non-trivial position B_0.

We rewrite $\mathbf{AmBB}_{\nu_0,\nu_1}$ as

$$\inf_{\substack{M_0\sim\nu_0, M_1\sim\nu_1 \\ M_t = M_0 + \int_0^t \Sigma_s dB_s}} \mathbb{E}\left[\int_0^1 (\Sigma_t - \bar{\Sigma})^2 dt\right] = \bar{\Sigma}^2 + \int x^2 d\nu_1 - \int x^2 d\nu_0 - 2\bar{\Sigma}\mathbf{AP}_{\nu_0,\nu_1},$$

since

$$\mathbb{E}\left[\int_0^1 \Sigma_t^2 dt\right] = \mathbb{E}[\langle M\rangle_1] = \int x^2 d\nu_1 - \int x^2 d\nu_0$$

We rewrite $\mathbf{AmBB}_{\nu_0,\nu_1}$ as

$$\inf_{\substack{M_0\sim\nu_0, M_1\sim\nu_1\\ M_t=M_0+\int_0^t \Sigma_s dB_s}} \mathbb{E}\left[\int_0^1 (\Sigma_t-\bar{\Sigma})^2 dt\right] = \bar{\Sigma}^2 + \int x^2 d\nu_1 - \int x^2 d\nu_0 - 2\bar{\Sigma}\mathbf{AP}_{\nu_0,\nu_1},$$

since

$$\mathbb{E}\left[\int_0^1 \Sigma_t^2 dt\right] = \mathbb{E}[\langle M\rangle_1] = \int x^2 d\nu_1 - \int x^2 d\nu_0$$

and where the new problem is

$$\mathbf{AP}_{\nu_0,\nu_1} = \sup_{\substack{M_0\sim\nu_0, M_1\sim\nu_1\\ M_t=M_0+\int_0^t \Sigma_s dB_s\\ M \text{ martingale}}} \mathbb{E}\left[\int_0^1 \Sigma_t dt\right] = \sup_{\substack{M_0\sim\nu_0, M_1\sim\nu_1\\ M_t=M_0+\int_0^t \Sigma_s dB_s\\ M \text{ martingale}}} \mathbb{E}\left[M_1(B_1-B_0)\right]. \qquad \text{(AP)}$$

We rewrite $\mathbf{AmBB}_{\nu_0,\nu_1}$ as

$$\inf_{\substack{M_0\sim\nu_0, M_1\sim\nu_1\\ M_t=M_0+\int_0^t \Sigma_s dB_s}} \mathbb{E}\left[\int_0^1 (\Sigma_t-\bar{\Sigma})^2 dt\right] = \bar{\Sigma}^2 + \int x^2 d\nu_1 - \int x^2 d\nu_0 - 2\bar{\Sigma}\mathbf{AP}_{\nu_0,\nu_1},$$

since

$$\mathbb{E}\left[\int_0^1 \Sigma_t^2 dt\right] = \mathbb{E}[\langle M\rangle_1] = \int x^2 d\nu_1 - \int x^2 d\nu_0$$

and where the new problem is

$$\mathbf{AP}_{\nu_0,\nu_1} = \sup_{\substack{M_0\sim\nu_0, M_1\sim\nu_1\\ M_t=M_0+\int_0^t \Sigma_s dB_s\\ M \text{ martingale}}} \mathbb{E}\left[\int_0^1 \Sigma_t dt\right] = \sup_{\substack{M_0\sim\nu_0, M_1\sim\nu_1\\ M_t=M_0+\int_0^t \Sigma_s dB_s\\ M \text{ martingale}}} \mathbb{E}\left[M_1(B_1-B_0)\right]. \qquad \text{(AP)}$$

$\implies$ it follows that $M_1 = F^{B_0}(B_1)$ is optimal with F^x increasing.
$\implies$ in fact, we can find $\alpha \sim B_0$, such that $F^x \equiv F$.
$\implies$ and hence $M_t = \mathbb{E}[F(B_1)|\mathcal{F}_t] = (F * \gamma_{1-t})(B_t)$, with $\gamma_t \sim \mathcal{N}(0,t)$.

G-mBB calibration problem

Similarly, in our calibration problem

$$\mathbf{GmBB}_{\mu_0,\mu_1} = \inf_{\substack{S_0\sim\mu_0, S_1\sim\mu_1 \\ S_t = S_0 + \int_0^t \sigma_u S_u dB_u}} \mathbb{E}\left[\int_0^1 (\sigma_t - \bar{\sigma})^2 dt\right],$$

for any such martingale S we have

$$\mathbb{E}\left[\int_0^1 \sigma_t^2 dt\right] = 2\mathbb{E}[\log(S_0/S_1)] = 2\int \log(x) d\mu_0 - 2\int \log(x) d\mu_1$$

and hence $\mathbf{GmBB}_{\mu_0,\mu_1}$ is equivalent to the following problem:

$$\mathbf{GP}_{\mu_0,\mu_1} = \sup_{\substack{S_0\sim\mu_0, S_1\sim\mu_1 \\ S_t = S_0 + \int_0^t \sigma_u S_u dB_u \\ S \text{ martingale}}} \mathbb{E}\left[\int_0^1 \sigma_t dt\right], \tag{GP}$$

where $(B_t)_{t\geq 0}$ is a Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.

From A-mBB to G-mBB and back

It turns out G-mBB can be mapped 1-1 to A-mBB for different marginals!
W.l.o.g., suppose $\int x\mu_0(dx) = \int x\mu_1(dx) = 1$.

From A-mBB to G-mBB and back

It turns out G-mBB can be mapped 1-1 to A-mBB for different marginals!
W.l.o.g., suppose $\int x\mu_0(dx) = \int x\mu_1(dx) = 1$.

The relationship between the problems can be deduced using PDE arguments on the dual side.

It turns out G-mBB can be mapped 1-1 to A-mBB for different marginals!
W.l.o.g., suppose $\int x\mu_0(dx) = \int x\mu_1(dx) = 1$.

The relationship between the problems can be deduced using PDE arguments on the dual side.

On the primal side, it is a change of measure argument, akin to Campi, Laachir and Martini '17; see also Beiglböck, Pammer and Riess '24.

Define $d\tilde{\mathbb{P}} := S_1 d\mathbb{P}$ and let $R_t = 1/S_t$, a $\tilde{\mathbb{P}}$-martingale. Then

$$\mathbb{E}\left[\int_0^1 \sigma_t dt\right] = \tilde{\mathbb{E}}\left[R_1 \int_0^1 \sigma_t dt\right] = \tilde{\mathbb{E}}\left[\int_0^1 R_t \sigma_t dt\right] = \tilde{\mathbb{E}}\left[\int_0^1 \Sigma_t dt\right],$$

where $\Sigma_t := R_t \sigma_t$ and Itô gives $dR_t = \Sigma_t d\tilde{W}_t$, for a $\tilde{\mathbb{P}}$-BM W.

$$\int g d\nu_1 := \tilde{\mathbb{E}}[g(R_1)] = \mathbb{E}\left[\frac{g(R_1)}{R_1}\right] = \mathbb{E}\left[g\left(\frac{1}{S_1}\right) S_1\right] = \int g\left(\frac{1}{y}\right) y\mu_1(dy).$$

From A-mBB to G-mBB and back

For a μ-integrable $f : \mathbb{R}_+ \to \mathbb{R}_+$, we consider the f-reflected measure

$$f_\dagger \mu = \left(y \to \frac{1}{f(y)} \right)_\# \left(\frac{f(y)}{\int f(x)\mu(dx)} \mu(dy) \right).$$

Theorem

Let $\mu_0, \mu_1 \in \mathcal{P}_{-1,1}(\mathbb{R}_+)$ *satisfy* $\mu_0 \preccurlyeq_{cx} \mu_1$. *Let* $\nu_i = Id_\dagger \mu_i$, $i = 0, 1$. *Then*

$$\mathbf{GP}_{\mu_0,\mu_1} = \mathbf{AP}_{\nu_0,\nu_1},$$

and (GP) *admits a unique optimiser in distribution*

From A-mBB to G-mBB and back

For a μ-integrable $f : \mathbb{R}_+ \to \mathbb{R}_+$, we consider the f-reflected measure

$$f_\dagger \mu = \left(y \to \frac{1}{f(y)} \right)_{\#} \left(\frac{f(y)}{\int f(x)\mu(dx)} \mu(dy) \right).$$

Theorem

Let $\mu_0, \mu_1 \in \mathcal{P}_{-1,1}(\mathbb{R}_+)$ *satisfy* $\mu_0 \preccurlyeq_{cx} \mu_1$. *Let* $\nu_i = Id_\dagger \mu_i$, $i = 0, 1$. *Then*

$$\mathbf{GP}_{\mu_0,\mu_1} = \mathbf{AP}_{\nu_0,\nu_1},$$

and (GP) *admits a unique optimiser in distribution characterised by*

$$\mathbb{E}\Big[g\big(\{S_t : t \in [0,1]\}\big)\Big] = \mathbb{E}\Big[g\big(\{1/F(t, B_t) : t \in [0,1]\}\big) \cdot F(1, B_1)\Big],$$

for any measurable functional $g : C([0,1]; \mathbb{R}) \to \mathbb{R}_+$, *where* $(F(t, B_t), t \in [0,1])$ *is an optimiser for* $\mathbf{AP}_{\nu_0,\nu_1}$.

Further properties

- The optimiser in $\mathbf{GP}_{\mu_0,\mu_1}$ solves

$$dS_u = S_u \frac{S_u}{\partial_x F^{-1}(u, \frac{1}{S_u})} dB_u, \quad 0 < u < 1.$$

Further properties

- The optimiser in $\mathbf{GP}_{\mu_0,\mu_1}$ solves

$$dS_u = S_u \frac{S_u}{\partial_x F^{-1}(u, \frac{1}{S_u})} dB_u, \quad 0 < u < 1.$$

- The equivalence extends to a larger class of problems:

$$\mathbb{E}\left[\int_0^1 c(t, S_t, \sigma_t^2) dt\right] = \tilde{\mathbb{E}}\left[\int_0^1 R_t c(t, S_t, \sigma_t^2) dt\right] = \tilde{\mathbb{E}}\left[\int_0^1 R_t c\left(t, \frac{1}{R_t}, \frac{\Sigma_t^2}{R_t^2}\right) dt\right].$$

Further properties

- The optimiser in $\mathbf{GP}_{\mu_0,\mu_1}$ solves

$$dS_u = S_u \frac{S_u}{\partial_x F^{-1}(u, \frac{1}{S_u})} dB_u, \quad 0 < u < 1.$$

- The equivalence extends to a larger class of problems:

$$\mathbb{E}\left[\int_0^1 c(t, S_t, \sigma_t^2) dt\right] = \tilde{\mathbb{E}}\left[\int_0^1 R_t c(t, S_t, \sigma_t^2) dt\right] = \tilde{\mathbb{E}}\left[\int_0^1 R_t c\left(t, \frac{1}{R_t}, \frac{\Sigma_t^2}{R_t^2}\right) dt\right].$$

- The optimiser in (GP) also solves (AP) if and only if

$$\log(S_1/S_0) \sim \mathcal{N}(-\bar{\sigma}^2/2, \bar{\sigma}^2),$$

for some $\bar{\sigma}$.

Further properties

- Consider now projecting a Bass martingale M into $\mathcal{M}(\mu_0, \mu_1)$, with $M_1 = F(B_1)$ and $M_0 \sim \delta_{m_0}$ for simplicity.
  This is equivalent to a weak OT pb:

$$\sup_{S \in \mathcal{M}(\mu_0,\mu_1)} \mathbb{E}[S_1 F(B_1)] \quad \equiv \quad \sup_{\pi} \int MC(\pi^x, q) d\mu_0(x),$$

  which is solved by q-Bass mg of Tschiderer '24, $q \sim M_1 = F(B_1)$. If the optimiser has $S_1 = G(\xi + F(B_1))$, then

$$S_t = \mathbb{E}[S_1 | \mathcal{F}_t] = \int G(\xi + F(B_t + z)) d\gamma_{1-t}(z) := G_t(\xi, B_t).$$

Numerics

Solution to $\mathbf{AmBB}_{\nu_0,\nu_1}$, given by $M_1 = F(B_1)$ with $B_0 \sim \alpha$ is characterised by the Martingale Sinkhorn system:

$$\nu_0 = (\gamma_1 * F)_{\#}\alpha,$$
$$\nu_1 = F_{\#}(\gamma_1 * \alpha),$$

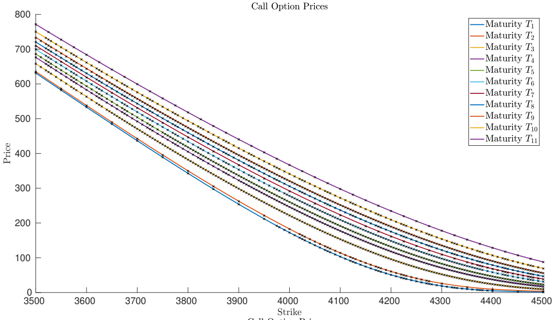
which is another way to write the fixed-point problem of Conze & Henry-Labordère '21, see also Acciaio, Marini and Pammer '23.

The above immediately allows us to solve also $\mathbf{GmBB}_{\mu_0,\mu_1}$.

Furthermore, we can do this across many maturities. Note that the resulting local volatility surface will likely be discontinuous across maturities. We now test and compare A-mBB and G-mBB martingale on market data.

Market Data

- BNP Paribas data on SPX options as of 27/10/2023, with many strikes and maturities: 27/11/2023, 29/12/2023, 19/01/2024, 29/02/2024, 15/03/2024, 28/03/2024, 19/04/2024 and 17/05/2024.
- CDFs built via Breeden Litzenberger formula and interpolated/extrapolated implied vols.
- Rescale variables $S_t \to S_t/S_0$. Com domain $(-0.5, 3) \times (T_k, T_{k+1})$ with 1001 spatial gridpoints and 1t01 time gridpoints.
- Solve heat equation using Crank-Nicolson.
- For G-Bass, we do CDF $\to$ density $\to$ reflected density $\to$ reflected cdf. Reflected density via numerical derivative over 250 gridpoints
- whole numerics took ca 5 min on a laptop.



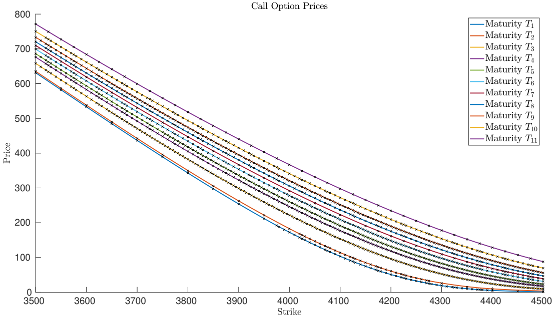


Figure: Call prices: Bass and Geometric Bass models.

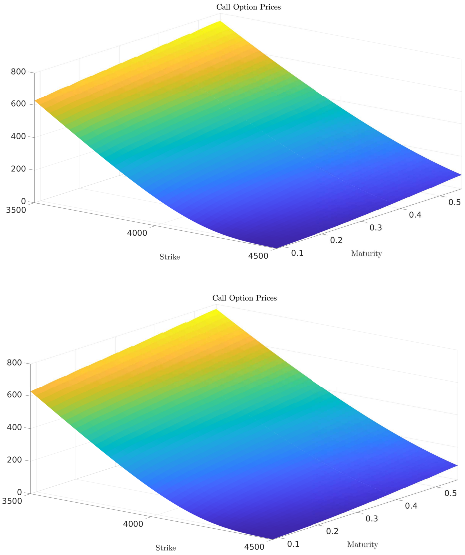


Figure: Call price surface: Bass and Geometric Bass models.

Conclusion and future research

Conclusion:

- Semimartingale OT is a powerful model selection tool
- Arbitrary constraints: European options, path-dependent options and (maybe?) American options
- We develop generic approach to Calibration via OT
- We use it to tackle difficult joint calibration problems: SPX options + VIX futures + VIX options prices; interest rates and SPX options
- Numerical proof-of-concept results

Future research:

- Improving computational efficiency and exploring applications in higher dimensions
  - Deep PDE solvers (see, e.g., Han et al. (2020))
  - Neural SDE (see, e.g., Cuchiero et al. (2020))
- OT Calibration to American options

Thank you!

Selected references I

Benamou, J.-D., and Brenier, Y.
A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem.
Numerische Mathematik 84, 3 (2000), 375–393.

Brunick, G., and Shreve, S.
Mimicking an Itô process by a solution of a stochastic differential equation.
Annals of Applied Probability 23, 4 (2013), 1584–1628.

Cuchiero, C., Khosrawi, W., and Teichmann, J.
A generative adversarial network approach to calibration of local stochastic volatility models.
arXiv preprint arXiv:2005.02505 (2020).

Figalli, A.
Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients.
Journal of Functional Analysis 254, 1 (2008), 109—153.

Fouque, J.-P., and Saporito, Y. F.
Heston stochastic vol-of-vol model for joint calibration of VIX and S&P 500 options.
Quantitative Finance 18, 6 (2018), 1003–1016.

Gatheral, J.
Consistent modeling of SPX and VIX options.
In Bachelier congress (2008), vol. 37, pp. 39–51.

Gatheral, J., Jusselin, P., and Rosenbaum, M.
The quadratic rough Heston model and the joint S&P 500/VIX smile calibration problem.
arXiv preprint arXiv:2001.01789 (2020).

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

Selected references II

Guo, I., and Loeper, G.
Path dependent optimal transport and model calibration on exotic derivatives.
arXiv preprint arXiv:1812.03526 (2018).

Guo, I., Loeper, G., and Wang, S.
Local volatility calibration by optimal transport.
2017 MATRIX Annals (2019).

Guo, I., Loeper, G., and Wang, S.
Calibration of local-stochastic volatility models by optimal transport.
Submitted to Mathematical Finance

Gyöngy, I.
Mimicking the one-dimensional marginal distributions of processes having an Ito differential.
Probability theory and related fields 71, 4 (1986), 501–516

Guyon, J.
The joint S&P 500/VIX smile calibration puzzle solved.
Risk, April (2020).

Han, J., Jentzen, A., et al.
Algorithms for solving high dimensional pdes: From nonlinear monte carlo to machine learning.
arXiv preprint arXiv:2008.13333 (2020).

Tan, X., Touzi, N.
Optimal transportation under controlled stochastic dynamics.
The Annals of Probability 41, 5 (2013), 3201-–3240.

Trevisan, D.
Well-posedness of multidimensional diffusion processes with weakly differen- tiable coefficients.
Electronic Journal of Probability 21 (2016), Paper No. 22, 41.