Risk management of option books with arbitrage-free neural-SDE market models

CHRISTOPH REISINGER

with SAMUEL N. COHEN and VICTOR WANG

supported by CME Group and the Oxford–Man Institute

Soesterberg Winter School 2023

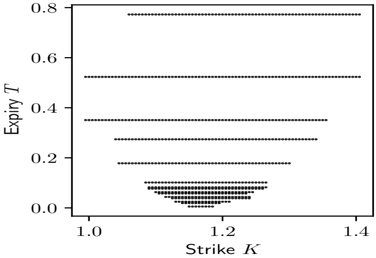
UNIVERSITY OF OXFORD
Mathematical Institute

Oxford Mathematics

UNIVERSITY OF OXFORD
Mathematical Institute

Modelling joint dynamics of liquid vanillas is crucial for arbitrage-free pricing of illiquid derivatives and managing risks of books.

Objective: *Develop a practical, nonparametric model for the European option book respecting underlying financial constraints*.



The quoted strikes and expiries of CME-listed EURUSD calls, 31/05/2018.

Freedom of static (model-independent!) arbitrage

UNIVERSITY OF OXFORD
Mathematical Institute

- The prices of options are heavily constrained and interrelated.
- No arbitrage dictates bounds on option prices in terms of the underlying, e.g.

$$C_t(T, K) \leq e^{-r(T-t)} S_t,$$

and between the options themselves, e.g.

$$C_t(T, K) \leq C_t(T, K') \text{ when } K \geq K'.$$

- In general, absence of model-free arbitrage is characterised by linear constraints: positivity, monotonicity, convexity, etc (Cousot, 2007; Cohen, R., & Wang, 2020)

Freedom of static (model-independent!) arbitrage II

UNIVERSITY OF OXFORD
Mathematical Institute

- Verify discrete static arbitrage conditions verified.
- Use shape preserving interpolation to construct $\check{c}$ in

$$\left\{ s \in C^{1,2}(D) : 0 \le s \le 1, \frac{\partial s}{\partial x} \ge 0, -1 \le \frac{\partial s}{\partial y} \le 0, \frac{\partial^2 s}{\partial y^2} \ge 0 \right\},$$

- By Breeden–Litzenberger, $\exists \ \{\mathbb{Q}_T\}_{T \in [0, T^*)}$, which are NDCO,

$$\mathbb{Q}_{T_1} \ge_{\text{cvx}} \mathbb{Q}_{T_2} \iff \begin{cases} \mathbb{Q}_{T_i} \text{ and } \mathbb{Q}_{T_j} \text{ have equal means;} \\ \int_{\mathbb{R}} (x-k)^+ \, d\mathbb{Q}_{T_1} \ge \int_{\mathbb{R}} (x-k)^+ \, d\mathbb{Q}_{T_2} \quad \forall x \in \mathbb{R}. \end{cases}$$

- By Kellerer's theorem, $\exists$ MMM with these marginals.
- By Carr and Madan, there is no static arbitrage for

$$\check{c}\,(T, k) = \mathbb{E}^{\mathbb{Q}} \left[(M_T - k)^+ \middle| \mathscr{F}_0 \right].$$

Traditional (martingale) models

Mathematical Institute

Martingale approach:

$$C_t(T,K) = D_t(T) \cdot \mathbb{E}^{\mathbb{Q}}[(S_T - K)^+ | \mathscr{F}_t]$$

under some pricing measure $\mathbb{Q}$.

Challenges:

(i) often requiring heavy, model-specific numerical methodology to calibrate these models;

(ii) calibrated model parameters change over time, even though they are assumed constant;

(iii) naturally posed under $\mathbb{Q}$, whereas the historical measure $\mathbb{P}$ is needed for risk management.

Background and contrasting works

Neural networks are ubiquitous these days.

Neural parameter-to-price maps, e.g.:

- Bayer, C., Horvath, B., Muguruza, A., Stemper, B., and Tomas, M. On deep calibration of (rough) stochastic volatility models. arXiv:1908.08806

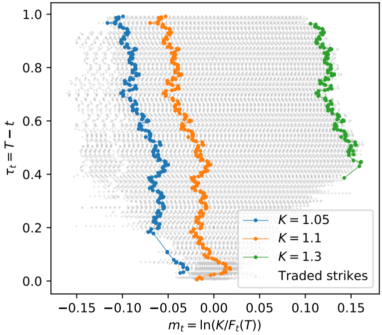
'Neural SDE' martingale models:

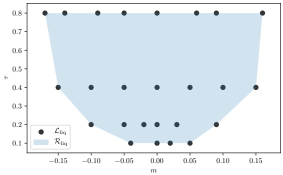
- Cuchiero, C., Khosrawi, W., and Teichmann, J. A generative adversarial network approach to calibration of local stochastic volatility models, Risks, 2020.
- Gierjatowicz, P., Sabate-Vidales, M., Šiška, D., Szpruch, L., and Žurič, Ž. Robust pricing and hedging via neural SDEs, JCF, to appear.

Here: 'Neural SDE' market models; broadly related ideas:

- HJM, BGM (Libor Market Model);
- 'Code book' processes, eg dynamic local vols of Carmona & Nadtochiy.

Normalization





An example of liquid range $\mathcal{R}_{\text{liq}}$ and lattice $\mathcal{L}_{\text{liq}}$.

CME EURUSD options expiring 2020/3/9.

We then model a normalised call price surface

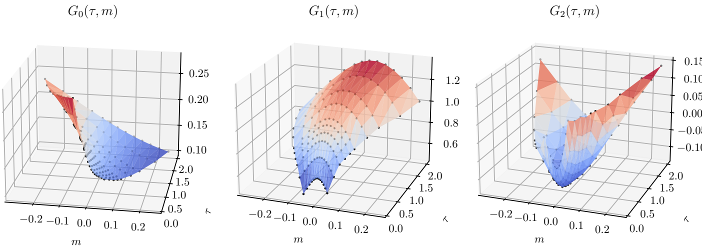
$$\widetilde{c}_t(\tau, m_t) = \frac{C_t(T, K)}{D_t(T) F_t(T)}.$$

UNIVERSITY OF OXFORD
Mathematical Institute

Step 1: Find a factor decomposition, ie factors ξ such that

$$\widetilde{c}_t(\tau, m) \approx G_0(\tau, m) + \sum_{i=1}^{d} G_i(\tau, m)\xi_{it},$$

to minimize statistical errors, dynamic and static arbitrage.



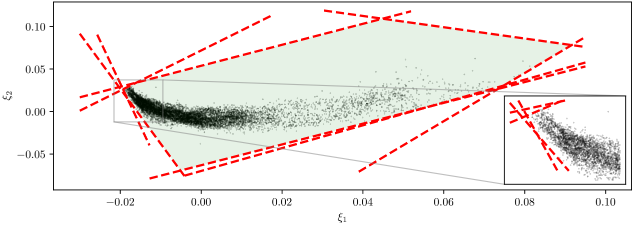
Eurex index options data — see later.

University of Oxford
Mathematical Institute

Step 1: Find a factor decomposition, ie factors ξ such that

$$\widetilde{c}_t(\tau, m) \approx G_0(\tau, m) + \sum_{i=1}^{d} G_i(\tau, m)\xi_{it},$$

to minimize statistical errors, dynamic and static arbitrage.



Step 2: Learn the factor dynamics, ie fit coefficients in a model

$$\begin{cases} \frac{\mathrm{d}S_t}{S_t} &= (\alpha(S_t, \xi_t) - q_t)\ \mathrm{d}t + \gamma(S_t, \xi_t)\,\mathrm{d}W_{0,t}, \\ \mathrm{d}\xi_t &= \mu(S_t, \xi_t)\,\mathrm{d}t + \sigma(S_t, \xi_t)\,\mathrm{d}W_t, \end{cases}$$

to minimize statistical errors, subject to static no-arbitrage.

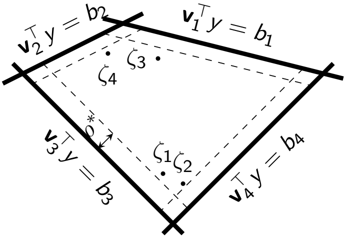
- Note that we work in (τ, m) coordinates, making stationarity more reasonable.
- We later fix α and remove the dependence on S ($\approx$ a scale-invariance property).

Constraining a process $\mathrm{d}Y_t = \mu(Y_t)\,\mathrm{d}t + \sigma(Y_t)\,\mathrm{d}W_t$ (Friedman & Pinsky, 1973)

Mathematical Institute

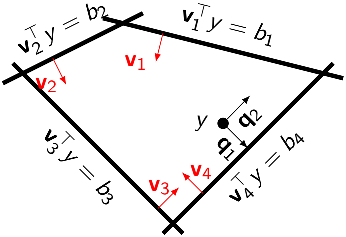
Process stays in interior if on k-th boundary, inward normal $\mathbf{v}_k$:

Drift: $\mathbf{v}_k^\top \mu(y) \geq 0$



$\mu(y) = \hat{\mu}(y) + \sum_k \lambda_k(y)(\zeta_k - y)$,
adding sufficient inwards drift

Diffusion: $\mathbf{v}_k^\top a(y) \mathbf{v}_k = 0$

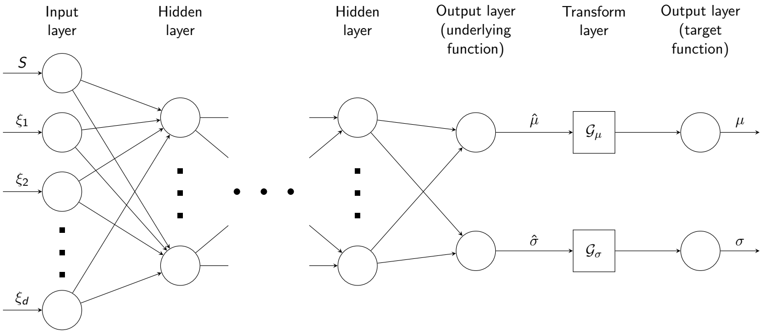


$\sigma(y) = (\mathbf{P}(y))^\top \hat{\sigma}(y)$,
normal component of $\mathbf{P} \to 0$.

Step 2: learning a constrained diffusion

Neural network architecture

UNIVERSITY OF OXFORD
Mathematical Institute



Constrained neural network.

Step 2: learning a constrained diffusion

Objective function

UNIVERSITY OF OXFORD
Mathematical Institute

An Euler–Maruyama approximation leads to the following unconstrained, penalized optimization problem for the MLE:

$$\min_{\hat{\mu},\hat{\sigma}} J[\hat{\mu},\hat{\sigma}] = \sum_{i=0}^{L-1}\left[\ln|a(i)| + \frac{1}{\Delta t}\|y_{t_{i+1}} - y_{t_i}\|^2_{a(i)} + \|\mu(i)\|^2_{a(i)}\Delta t - 2\left(\mu(i), y_{t_{i+1}} - y_{t_i}\right)_{a(i)}\right] + \lambda\mathcal{R}(\hat{\mu},\hat{\sigma}),$$

where $\mu = \mathcal{G}_\mu[\hat{\mu}]$ and $a = \mathcal{G}_\sigma[\hat{\sigma}](\mathcal{G}_\sigma[\hat{\sigma}])^\top$, and

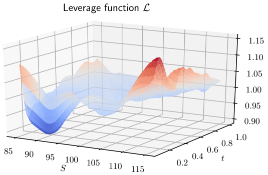
$$\hat{\sigma} = \begin{bmatrix} \exp(\phi_1^\theta) & 0 & \cdots & 0 \\ \phi_2^\theta & \exp(\phi_3^\theta) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \phi^\theta_{\frac{1}{2}d(d-1)+1} & \phi^\theta_{\frac{1}{2}d(d-1)+2} & \cdots & \exp\left(\phi^\theta_{\frac{1}{2}d(d+1)}\right) \end{bmatrix}, \quad \hat{\mu} = \begin{bmatrix} \phi^\theta_{\frac{1}{2}d(d+1)+1} \\ \phi^\theta_{\frac{1}{2}d(d+1)+2} \\ \vdots \\ \phi^\theta_{\frac{1}{2}d(d+3)} \end{bmatrix}.$$

A synthetic market example

Mathematical Institute

We use a calibrated Heston–SLV model

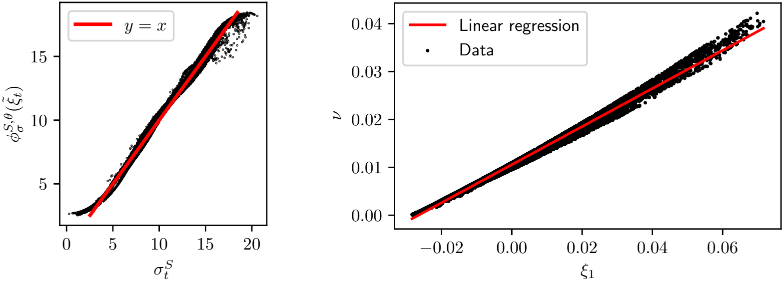
$$\begin{aligned} \mathrm{d}S_u &= \mathcal{L}_t(u, S_u)\sqrt{\nu_t} S_u \,\mathrm{d}W_u^S, \\ \mathrm{d}\nu_u &= \kappa(\theta - \nu_u)\,\mathrm{d}u + \sigma\sqrt{\nu_u}\,\mathrm{d}W_u^\nu, \\ \mathrm{d}\langle W_u^S, W_u^\nu\rangle &= \rho\,\mathrm{d}u, \qquad u \in (t, T^*). \end{aligned}$$



to generate 'market data'.

Heston parameters						Simulation		
S_0	ν_0	θ	κ	σ	ρ	L	Δt	N
100	0.0083	0.0085	8.3	0.32	-0.42	10000	0.0001	46

In-sample comparison

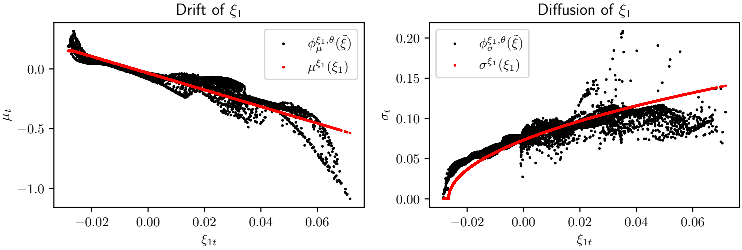


Left: Estimated and ground-truth diffusion coefficient for S.

Right: The (linear) relationship between ξ_1 and ν.

In-sample comparison

UNIVERSITY OF OXFORD
Mathematical Institute

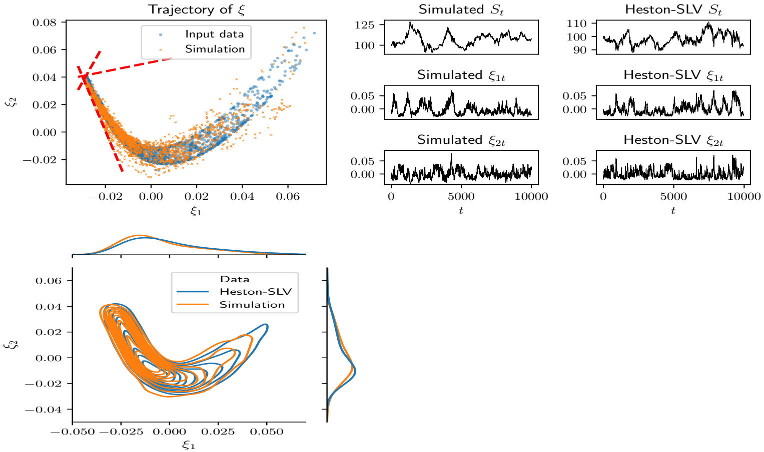


Estimated coefficients and (approximated) ground-truth for ξ_1.

Out-of-sample comparison

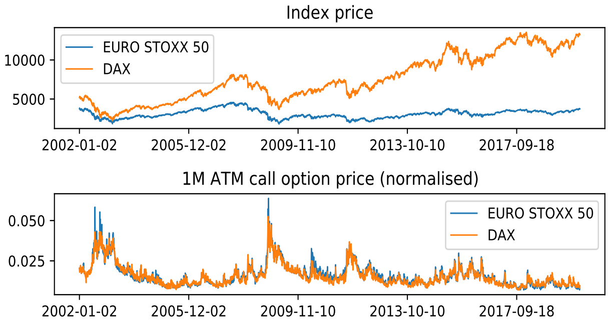
University of Oxford Mathematical Institute

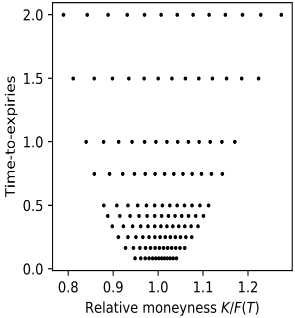
Simulation of S, ξ from learnt model, compared with input data



Application: Eurex index options (data: OptionMetrics)

UNIVERSITY OF OXFORD
Mathematical Institute



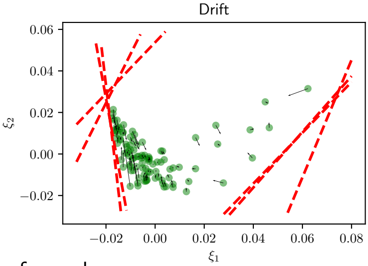


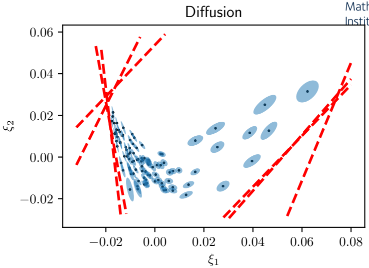
We add 3 'secondary' factors, simple OU processes, which we fit.

- Mean vega-weighted absolute percentage error ≈ 1.33%
- Static arbitrage fraction ≈ 0.05%

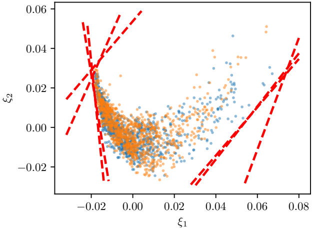
Primary risk factors

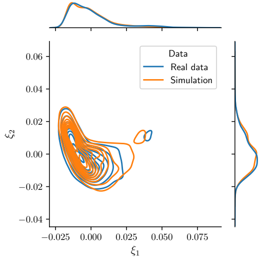
Learnt drift and diffusion:





Out-of-sample:

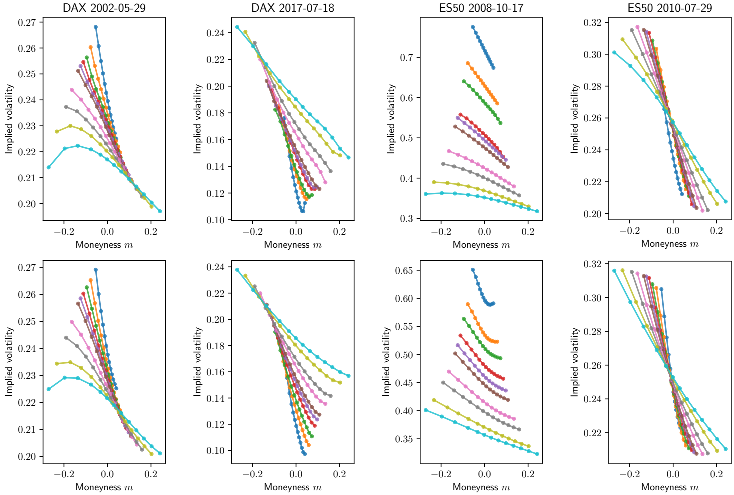




Implied volatility surfaces

Real data

Simulation



Hedging

Let $V = C(T^*, K^*)$ be the option to hedge. Recall

$$C_t(T,K) = S_t \widetilde{c}_t(\tau, m), \qquad \widetilde{c}_t(\tau, m) = G_0(\tau, m) + \sum_{i=1}^{d} G_i(\tau, m)\xi_{it}.$$

- Sensitivity-based hedging: To be delta-neutral, hold:

$$X^S = \frac{\partial V}{\partial S} = \left(\widetilde{c} - \frac{\partial \widetilde{c}}{\partial m}\right)(\tau^*, m^*),\ \tau^* = T^* - t, m^* = \ln(K^*/S).$$

- Hedge ξ-exposure with options, weights $(X^S, X^{C_1}, \dots, X^{C_{d'}})$.
- Independent of the neural-SDE model!

- Minimum variance hedging accounts for dependence.
- For all factors $f = S, f = \xi_j$, for $j = 1, \dots, d'$, $\langle d\Pi, df \rangle = 0$.

Testing methodology

UNIVERSITY OF OXFORD
Mathematical Institute

Number of tested portfolios of various types.

Portfolio	Naive	Outright	Delta spread	Delta butterfly	Strangle	Calendar spread	VIX
Number	1	70	210	30	30	45	1

Error measures:

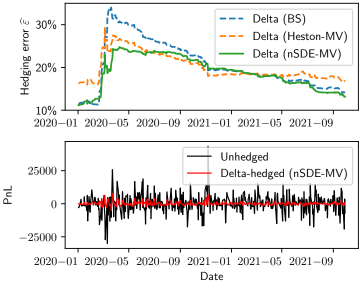
$$\overline{\mathcal{E}}^2(\Pi, \Delta t) = \frac{1}{L-1} \sum_{l=1}^{L} \left(\Pi_{t_l+\Delta t} - \Pi_{t_l}\right)^2,$$

$$\widehat{\mathcal{E}}_{t_l}^2(\Pi, \Delta t, \lambda) = \begin{cases} \left(\Pi_{t_1+\Delta t} - \Pi_{t_1}\right)^2, & \text{if } l = 1, \\ \lambda \widehat{\mathcal{E}}_{t_{l-1}}^2(\Pi, \Delta t, \lambda) + (1-\lambda)\left(\Pi_{t_l+\Delta t} - \Pi_{t_l}\right)^2, & \text{for } l = 2, \ldots, L. \end{cases}$$

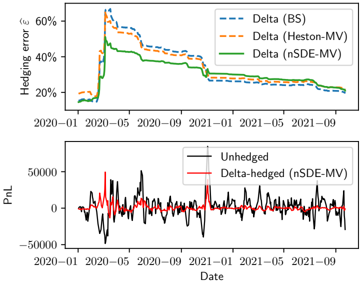
$$\overline{\varepsilon}(\Delta t) = \frac{\overline{\mathcal{E}}(\Pi, \Delta t)}{\overline{\mathcal{E}}(V, \Delta t)} \times 100\%, \quad \widehat{\varepsilon}_{t_l}(\Delta t, \lambda) = \frac{\widehat{\mathcal{E}}_{t_l}(\Pi, \Delta t, \lambda)}{\widehat{\mathcal{E}}_{t_l}(V, \Delta t, \lambda)} \times 100\%.$$

Delta hedging

UNIVERSITY OF OXFORD
Mathematical Institute



Daily rebalancing.

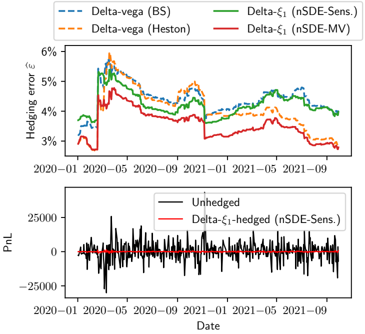


Weekly rebalancing.

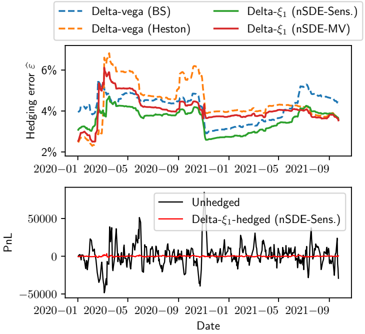
Top – EWMA hedging errors $\widehat{\varepsilon}_t(\Delta t, \lambda = 0.99)$ for the three delta hedging strategies.
Bottom – PnLs for the naive portfolio and for the nSDE-MV delta-hedged portfolio.

Delta-factor hedging

UNIVERSITY OF OXFORD
Mathematical Institute



Daily rebalancing.



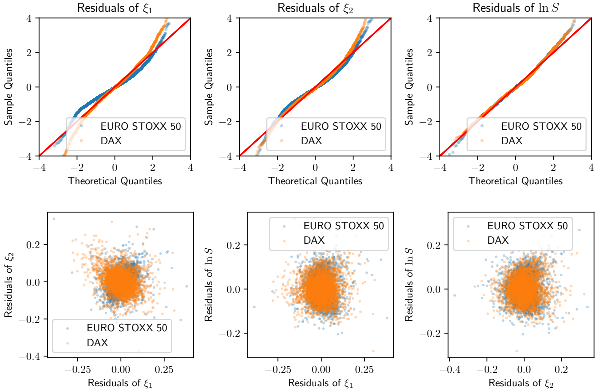
Weekly rebalancing.

Top – EWMA hedging errors $\widehat{\varepsilon}_t(\Delta t, \lambda = 0.99)$ for the four hedging strategies.
Bottom – PnLs for the naive portfolio and for the sensitivity-based delta-ξ_1-hedge.

Value at Risk

UNIVERSITY OF OXFORD
Mathematical Institute

- We can compute risk profiles for option portfolios.
- We use the historical innovations (to allow for unmodelled and higher-order correlation effects) from our training data.



1-day 0.99-Value at Risk

UNIVERSITY OF OXFORD
Mathematical Institute

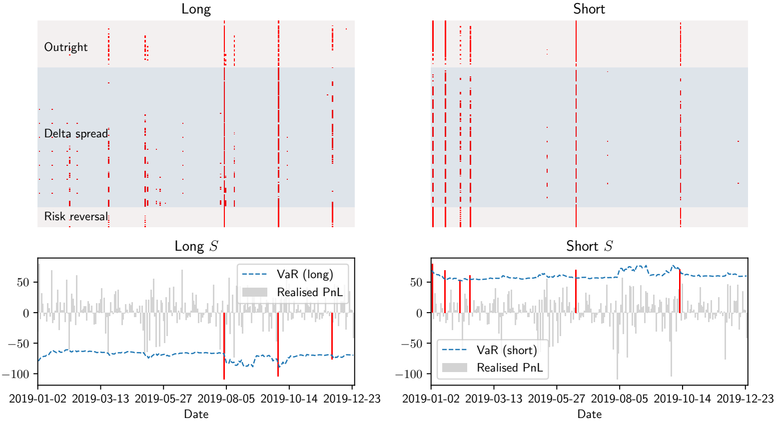
Number of tested portfolios of various types.

Delta-exposed			Delta-hedged				
Outright	Delta spread	Risk reversal	Delta butterfly	Delta-hedged option	Delta-neutral strangle	Calendar spread	VIX
140	420	60	20	60	60	90	2

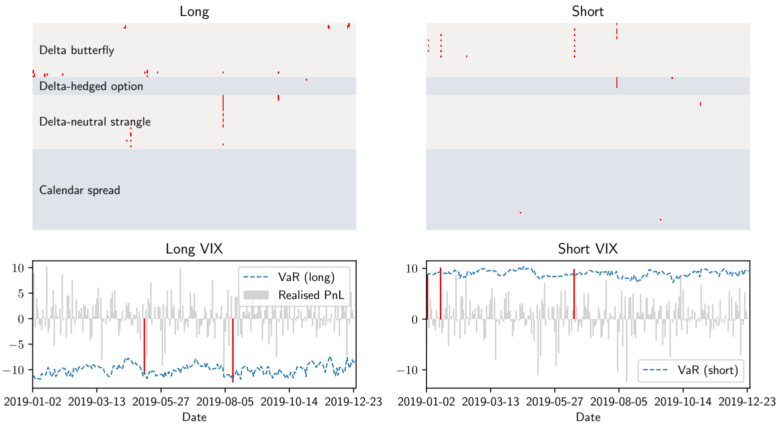
- For comparison, we also use a Filtered Historical Simulation approach, from a time-series model on the Heston parameters.

	Neural-SDE			FHS		
Coverage ratio median	0.9921			0.9881		
Coverage ratio mean	0.9887			0.9742		
Kupiec PF (two-sided)	6.92%			25.23%		
Christoffersen independence	0.70%			11.03%		
Basel committee traffic light	69.1%	29.7%	0.5%	62.4%	25.9%	10.8%

1-day 0.99-Value at Risk



1-day 0.99-Value at Risk



Conclusions

- Combining machine learning with economic modelling (no arbitrage) gives powerful techniques for mathematical finance.
- Market models give significant computational advantages, and can be trained using realistic amounts of historical data.
- American and exotic options are more difficult, due to the lack of good no-arbitrage conditions on option surfaces.
- The hedging performance is comparable to standard model based (stochastic volatility) hedges, at significantly lower cost.
- These models give risk estimates which perform better than traditional filtered historical simulation, at significantly lower computational cost.

References

Mathematical Institute

- *Arbitrage-free neural-SDE market models*, Cohen, R., & Wang, arXiv:2105.11053.
- *Estimating risks of option books using neural-SDE market models*, Cohen, R., & Wang, arXiv:2202.07148.
- *Hedging option books using neural-SDE market models*, Cohen, R., & Wang, arXiv:2205.15991.