Appendix

A Factor decoding algorithm

In this appendix, we recall the factor decoding algorithm that is introduced as Algorithm 1 in [19]. As formulated in (6), we would like to decode d factors from N call option prices. For some fixed $d < N$, an ideal factor representation should achieve the following three objectives:

(i) **Statistical accuracy**: The reconstruction error $||\boldsymbol{\Upsilon}||_F$ should be as small as possible. Among linear dimension reduction techniques, PCA applied to the observed prices $\mathbf{C}$ gives the optimal reconstruction error.

(ii) **Minimal dynamic arbitrage**: There should be as few violations of the HJM-type drift restrictions (4) as possible. For this purpose, we estimate z_t from our observed prices $\mathbf{C}$, then construct price basis from principal components of z_t, to minimize the discrepancy between z_t and the space spanned by the factors.

(iii) **No static arbitrage**: The reconstructed prices should violate the static arbitrage constraints (5) as seldom as possible.

Suppose we construct d^{st}, d^{da} and d^{sa} factors for the above-mentioned objectives, respectively, where $d^{\mathrm{st}} + d^{\mathrm{da}} + d^{\mathrm{sa}} = d$. Algorithm 1 gives the details of the factor decoding process.

B Data pre-processing

Typical data vendors[15] provide *historical* EOD (end-of-day) option prices with parameters in the format of a fixed set of representative time-to-expiries and Black–Scholes deltas, instead of traded expiries and strikes. This format is popular in industry, as it saves users great efforts in building historical implied volatility surfaces, which otherwise requires the users themselves to reconcile traded option contracts that are changing over time.

This data format is also appealing to our modelling approach, though the raw data need to be pre-processed to (1) get rid of static arbitrage, and (2) interpolate prices to a fixed set of moneyness parameters. We now explain how we pre-process the OptionMetrics dataset as described in Section 3.1.

We use the arbitrage detection method of our previous work [18] to construct static arbitrage constraints in terms of option prices, and check the violation of these constraints by the data over time. In Figure 18, we show the histogram for the number of constructed arbitrage constraints and what percentage of them are violated over time. On average, there are about 3668 constraints per day. Historical option prices are almost free of static arbitrage after October 2008, before which there were about $8 \sim 9\%$ of constraint violations, in the worst cases, in 2002 and 2006.

We use historical *median* moneynesses to define our liquid option lattice $\mathcal{L}_{\mathrm{liq}}$ (as shown in Figure 1a), and interpolate prices to the fixed set of moneyness parameters. In Figure 19, we show boxplots of moneynesses that are calculated over time for the fixed set of 13 deltas and 10 time-to-expiries. The closer to the money a delta is, the less variation we observe in the corresponding moneyness. This implies that less interpolation error is introduced to prices for close-to-the-money options, assuming a C^2 convex price curve over moneyness. We use cubic splines to interpolate the call option price curve over moneyness for each expiry.

After interpolating option prices to the liquid lattice $\mathcal{L}_{\mathrm{liq}}$ for each historical date, we check the amount of static arbitrage in the interpolated prices and present the fraction of arbitrage constraint violation in Figure 20. Compared with Figure 18, the interpolation has introduced extra arbitrage

[15] For example, OptionMetrics (https://wrds-www.wharton.upenn.edu/pages/about/data-vendors/optionmetrics/), IVolatility (https://www.ivolatility.com/home.j), Nasdaq Data Link (https://data.nasdaq.com/data/OWF-optionworks-futures-options/documentation), Bloomberg, etc.

Algorithm 1: Decoding the factor representation

Input : Matrix of price data $\mathbf{C} \in \mathbb{R}^{(L+1)\times N}$; number of factors d; number of statistical accuracy factors d^{st}; number of dynamic arbitrage factors d^{da}; number of static arbitrage factors d^{sa}; data for z_t in (4), denoted as $\mathbf{Z} \in \mathbb{R}^{(L+1)\times N}$; constraint matrices $\mathbf{A}$, $\mathbf{b}$ from (5).

Output: Factor data $\boldsymbol{\Xi} \in \mathbb{R}^{(L+1)\times d}$; price basis vectors $\mathbf{G}_i \in \mathbb{R}^N$ for $i = 0, \ldots, d$.

```
Let $G_{0j} = \frac{1}{L+1}\sum_{l=1}^{L+1} C_{lj}$ for $j = 1, \ldots, N$ ;
/* Construct factors to minimize dynamic arbitrage. */
Compute the residual data $\mathbf{R}_0 = \mathbf{C} - \mathbf{1}_{L+1} \otimes \mathbf{G}_0$;
Compute the principal component decomposition of $\mathbf{Z}$ and assign, $\forall i = 1, \ldots, d^{\mathrm{da}}$ and $\forall l = 1, \ldots, L+1$,

    $$\mathbf{G}_i = i\text{-th principal component of } \mathbf{Z},\ \Xi_{li} = \langle (\mathbf{R}_0)_l, \mathbf{G}_i \rangle,$$

  where $(\mathbf{R}_0)_l$ is the $l$-th row of $\mathbf{R}_0$;
/* Construct factors to maximize statistical accuracy. */
Compute the residual data $\mathbf{R}_{d^{\mathrm{da}}} = \mathbf{R}_0 - \sum_{k=1}^{d^{\mathrm{da}}} \boldsymbol{\Xi}_k \otimes \mathbf{G}_k$, where $\boldsymbol{\Xi}_k = [\Xi_{1k} \cdots \Xi_{L+1,k}]$;
Compute the principal component decomposition of $\mathbf{R}_{d^{\mathrm{da}}}$ and assign, $\forall i = d^{\mathrm{da}} + 1, \ldots, d^{\mathrm{da}} + d^{\mathrm{st}}$ and $\forall l = 1, \ldots, L+1$,

    $$\mathbf{G}_i = (i - d^{\mathrm{da}})\text{-th principal component of } \mathbf{R}_{d^{\mathrm{da}}},\ \Xi_{li} = \langle (\mathbf{R}_{d^{\mathrm{da}}})_l, \mathbf{G}_i \rangle,$$

  where $(\mathbf{R}_{d^{\mathrm{da}}})_l$ is the $l$-th row of $\mathbf{R}_{d^{\mathrm{da}}}$;
/* Construct factors to minimize static arbitrage. */
foreach $i = d^{da} + d^{st}, \ldots, d-1$ do
    Compute the residual data $\mathbf{R}_i = \mathbf{R}_0 - \sum_{k=1}^{i} \boldsymbol{\Xi}_k \otimes \mathbf{G}_k$, where $\boldsymbol{\Xi}_k = [\Xi_{1k} \cdots \Xi_{L+1,k}]$;
    Compute the covariance of the residual data $\boldsymbol{\Sigma}_i = \mathbf{R}_i^\top \mathbf{R}_i$ ;
    Eigen-decompose $\boldsymbol{\Sigma}_i = \mathbf{Q}_i \boldsymbol{\Lambda}_i \mathbf{Q}_i^{-1}$ (where $\mathbf{Q}_i$ is a unitary matrix) ;
    For some $\mathbf{w} \in \mathbb{R}^d$ such that $\|\mathbf{w}\|_2 = 1$, define

        $$\mathbf{q}_i(\mathbf{w}) = \mathbf{Q}_i \mathbf{w},\ \mathbf{s}_i(\mathbf{w}) = \mathbf{R}_i \mathbf{q}_i(\mathbf{w}).$$

    Then solve the following minimization problem:

        $$\mathbf{w}^* = \underset{\|\mathbf{w}\|_2=1}{\arg\min} \Big( - \sum_{l=1}^{L+1} \mathbb{1}_{\left\{\langle \mathbf{A}\mathbf{q}_i(\mathbf{w}), \mathbf{s}_i(\mathbf{w}) \rangle \geq \mathbf{b} - \mathbf{A}\left(\sum_{k=1}^{i-1} \Xi_{lk} \mathbf{G}_k\right)\right\}} + \lambda \left\| \mathbf{R}_i - \mathbf{s}_i(\mathbf{w}) \otimes \mathbf{q}_i(\mathbf{w}) \right\|_2 \Big), \quad (13)$$

    where we require $\lambda < \min\left(1, \frac{1}{\|\mathbf{R}_i\|}\right)$ so that the penalty term is strictly less than 1;
    Let $\mathbf{G}_{i+1} = \mathbf{q}_i(\mathbf{w}^*)$ and $\boldsymbol{\Xi}_{i+1} = \mathbf{s}_i(\mathbf{w}^*)$ ;
end
```

for a few days, but in the worst case the introduced violation percentage is not greater than 2%. Thereafter, we perturb arbitrageable prices using the ℓ^1-repair method [18] to get arbitrage-free prices.

C Alternative model configurations

Apart from the market model discussed in the main text, we present two alternative models that differ in the number of primary factors chosen and the dependence between S and ξ. We also show their corresponding VaR backtesting results in Appendix F, which perform slightly worse than the model presented in the main text.

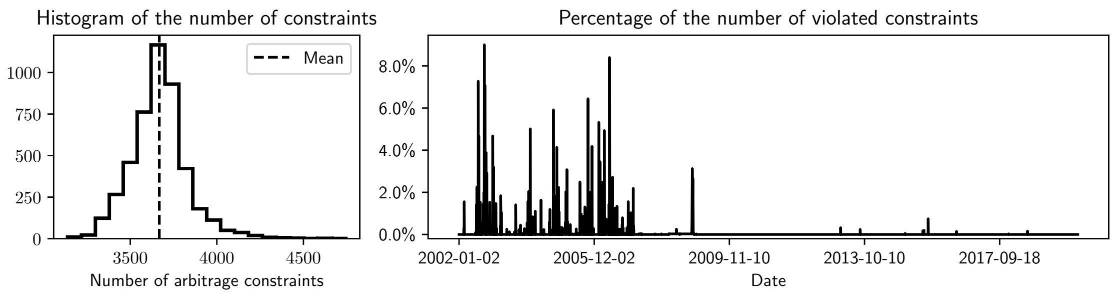

Figure 18: Statistics on the static arbitrage constraints violation for EURO STOXX 50 index option price data.

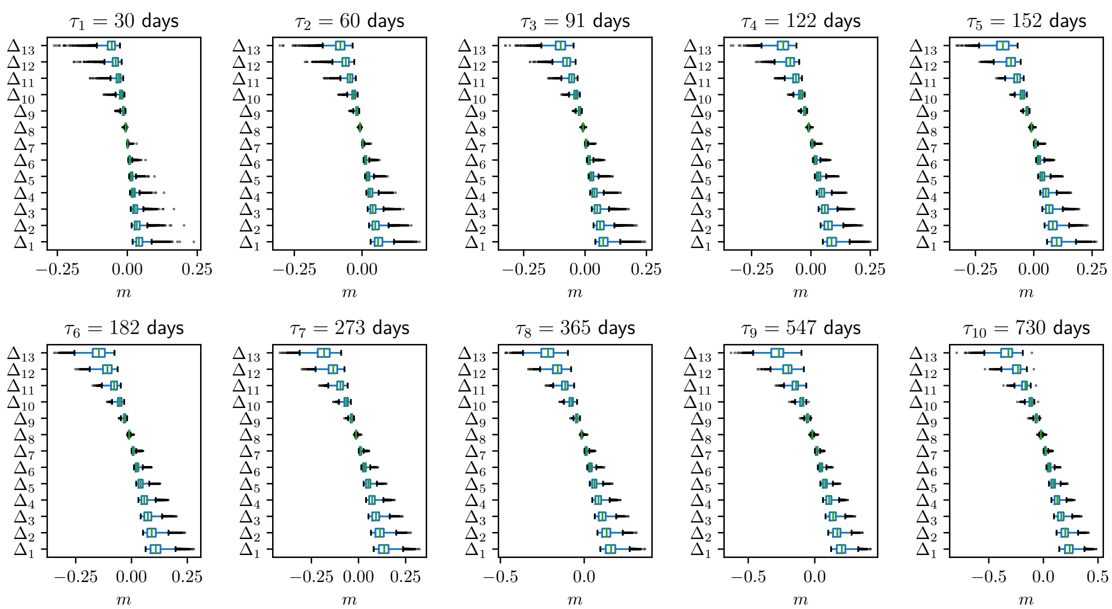

Figure 19: Boxplots of the moneynesses m corresponding to the 13 deltas for the 10 time-to-expiries. $\Delta_1 = 0.2$, $\Delta_{13} = 0.8$, and consecutive deltas have a 0.05 increment.

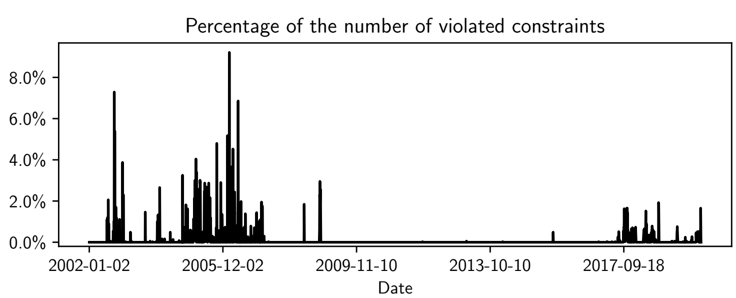

Figure 20: Statistics on the static arbitrage constraints violation for EURO STOXX 50 index option prices interpolated to the liquid lattice $\mathcal{L}_{\text{liq}}$.

C.1 A three-primary-factor market model

In Table 7, we display the types of the three primary factors $\xi = [\xi_1\ \xi_2\ \xi_3]^\top \in \mathbb{R}^3$ we use for representing call option prices, as well as the resultant MAPE, PDA and PSAS. Compared with the two primary factors (last row in Table 2), the additional dynamic arbitrage factor reduces dynamic arbitrage by 6.1% and MAPE in price reconstruction error by 0.56%, but introduces 0.55% more static arbitrage. In Figure 21, we plot the corresponding price basis functions of these three factors.

Factors	MAPE	PDA	PSAS
Dynamic arb. + Statistical acc. + Static arb.	4.26%	1.73%	0.85%

Table 7: MAPE, PDA and PSAS metrics when including three primary factors.

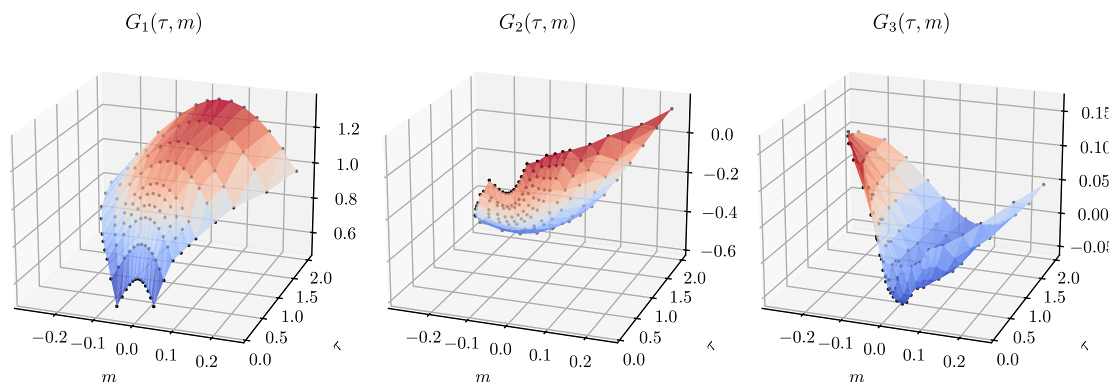

Figure 21: Price basis functions of the normalised call price surface.

In Figure 22, the black dots give trajectories of (ξ_i, ξ_j) for $i, j \in \{1, 2, 3\}$. For each (i, j), we also indicate boundaries of the polygon, that is, the 2D affine projection of the 3D arbitrage-free state space for ξ. Specifically, let $\mathcal{P}$ be the $\mathbb{R}^3$-polytope state space for the factors ξ formed by the static arbitrage constraints, then the affine projection of this polytope onto the (ξ_i, ξ_j)-space is given by

$$\pi_{i,j}(\mathcal{P}) = \left\{ y \in \mathbb{R}^2 : \exists \xi = (\xi_k) \in \mathcal{P}, y = [\xi_i\ \xi_j]^\top \right\}.$$

We plot the boundaries $\partial\pi_{i,j}(\mathcal{P})$ as red dashed lines in Figure 22. These boundaries (which are, in some sense, as wide as possible) suggest how tight the arbitrage bounds are.

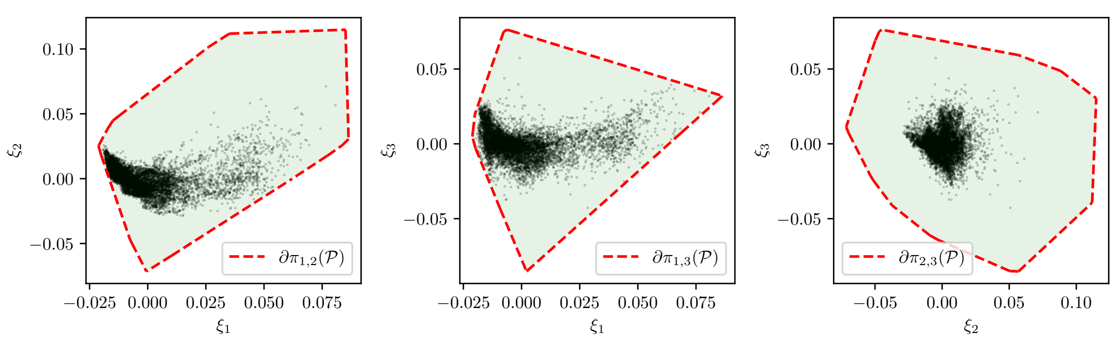

Figure 22: Trajectory (black dots) of the primary $\mathbb{R}^3$ factors and the corresponding static arbitrage constraints (red and orange dashed lines), projected onto $\mathbb{R}^2$ spaces.

We take the three factors and model their dynamics using neural-SDEs. After training the models, we forward-simulate $10,000$ timesteps, and compare the pairwise joint distribution of the simulated factors and that of the real data, as shown in Figure 23. The trained neural-SDE model seems to be capable of simulating long trajectory of the three factors that are similar to the real data, at least in terms of pairwise joint distribution.

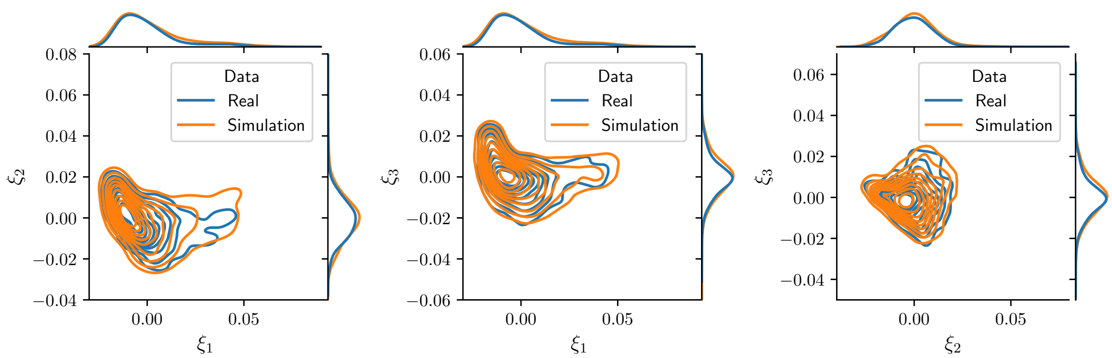

Figure 23: Joint distribution of pairs of the simulated primary factors, compared with that of the real data.

C.2 Modelling and estimating S and ξ jointly

We build and estimate an SDE model for $\tilde{\xi} = (\ln S, \xi)$ with a fully-specified diffusion matrix; in fact, the model (3) is a special case with a block-diagonal diffusion matrix. Given $\xi \in \mathbb{R}^d$, compared with the model (3), the joint model has d more $\mathbb{R}$-valued functions to estimate in the diffusion term, which captures dynamic correlation between $\ln S$ and each of the d factors.

For estimating the joint model, we still use the concatenated data (of EURO STOXX 50 and DAX options) for training the neural nets. We assume that the two stock index processes have the same diffusion function with respect to ξ but different constant drifts estimated by (9). This enables us to use the concatenated stock index data to estimate a universal volatility function for the two indices. In Figure 24, we show the loss history of training the neural nets. Both the training and validation losses decline rapidly over the first 1000 epochs and then converge gradually.

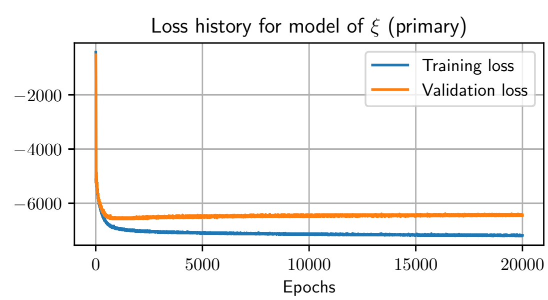

Figure 24: Evolution of training losses and validation losses for the joint model of $\ln S$ and ξ.

To see that the joint model has reasonably captured covariances between the underlying and the the factors, we show in Figure 25 the scatter plots of all pairs of model residuals. Compared with Figure 10, the joint model takes into account the leverage effect and leaves no correlation between residuals of $\ln S$ and ξ_1.

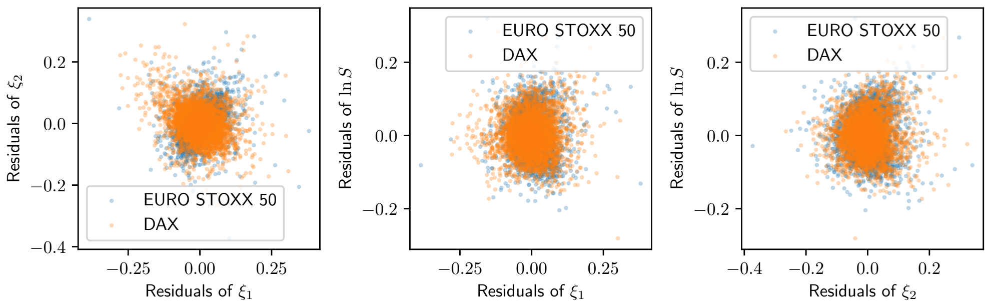

Figure 25: Scattter plots of historical in-sample model residuals.

D Heston model calibration results and historical scenarios

Heston models are constructed from the SDE

$$\mathrm{d}S_t = r_t S_t \,\mathrm{d}t + \sqrt{\nu_t} S_t \,\mathrm{d}W_t^S, \tag{14a}$$

$$\mathrm{d}\nu_t = k(\theta - \nu_t)\,\mathrm{d}t + \sigma\sqrt{\nu_t}\,\mathrm{d}W_t^\nu, \quad \mathrm{d}\langle W_t^S, W_t^\nu\rangle = \rho\,\mathrm{d}t, \tag{14b}$$

We re-write the variance process (14b) as

$$\mathrm{d}\nu_t = k(\theta - \nu_t)\,\mathrm{d}t + \eta\sqrt{\nu_t}\,\mathrm{d}W_t^S + \lambda\sqrt{\nu_t}\,\mathrm{d}W_t^\perp, \quad \mathrm{d}\langle W_t^S, W_t^\perp\rangle = 0,$$

where $\eta = -\sigma\rho$, $\lambda = \sigma\sqrt{1-\rho^2}$. We characterise option price surfaces by the parameters $\Theta = (S_0, \nu_0, \theta, k, \eta, \lambda)$. On each historical date t, we calibrate parameters for a Heston model by solving

$$\widehat{\Theta}_t = \arg\min_{\Theta} \sum_{j=1}^{N} \frac{1}{\mathcal{V}_t(\tau_j, m_j)} \left(c^\Theta(\tau_j, m_j) - \tilde{c}_t(\tau_j, m_j)\right)^2,$$

where $\mathcal{V}$ is the option vega calculated through (7). In Figure 26, we show the time series of the calibrated parameters and calibration errors, measured by MAPE, for historical price data of EURO STOXX 50 index options. Since ν_0, θ, k and λ are positive by construction, we include the positivity constraint for these parameters during the calibration. While η can be positive or non-positive, the calibrated result is always positive, implying negative correlations ρ between the index and volatility over time.

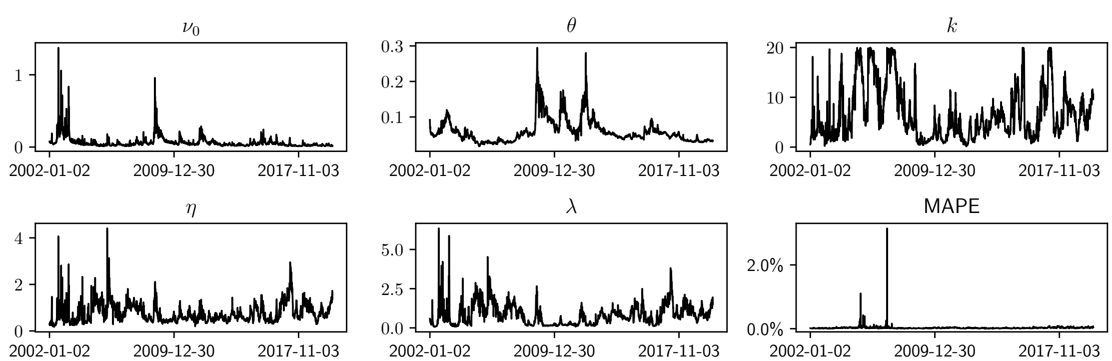

Figure 26: Heston model calibrated parameters and calibration errors (measured as mean absolute percentage error) for historical price data of EURO STOXX 50 index options.

Given that all risk factors are positive, we compute their 1-day log-returns, based on which we estimate the exponentially weighted moving average (EWMA) volatilities with time decay factor 0.95. We plot the EWMA 1-day volatility estimates over the estimation window in Figure 27. The RiskMetrics Technical document [39] suggests a value of 0.94 for the decay factor when estimating daily volatility. Nevertheless, in order to make the FHS-VaR and the nSDE-VaR more comparable, we optimise the choice of the decay factor by minimising the Wasserstein distance between the resulting EWMA vol and the in-sample neural-SDE vol in S (as shown in Figure 8). This yields the decay factor 0.95. Dividing the returns by the volatility estimates, we obtain historical standardised residuals, or *historical shocks*, as shown in Figure 28. The standardised residuals appear to be roughly stationary over time and do not exhibit volatility clustering. In order to generate risk scenarios for each t in the testing window, we multiply these historical shocks by the time-t 1-day EWMA volatility, as described in (12), and apply these filtered returns to the time-t risk factors, as if these historical shocks actually occurred.

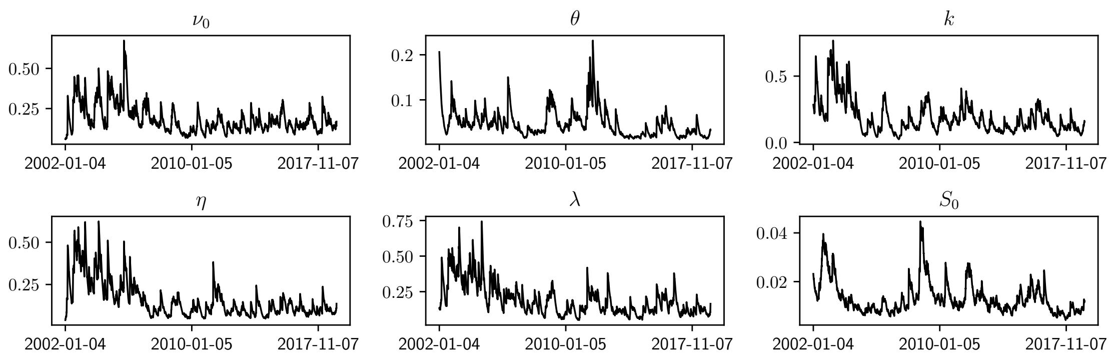

Figure 27: EWMA 1-day volatility estimates (with time decay factor 0.95) for the Heston risk factors for EURO STOXX 50 index options.

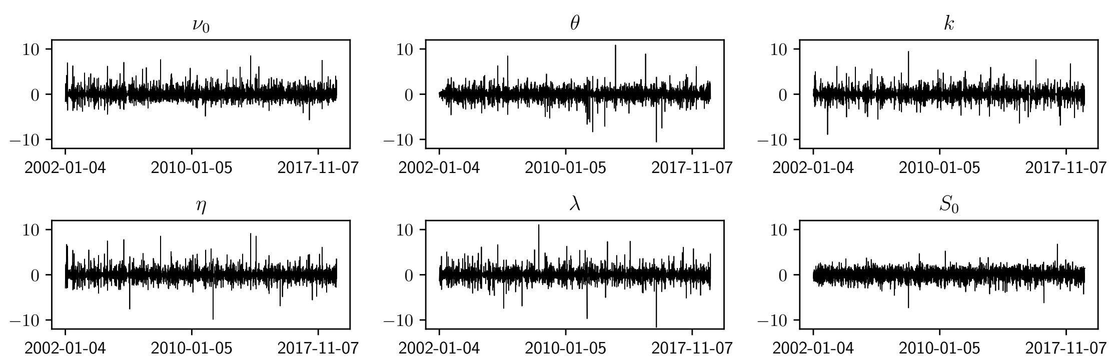

Figure 28: Standardised residuals of the Heston risk factors for EURO STOXX 50 index options.

E Tested option portfolios for risk simulation

To assess performance of a risk model, we backtest its risk evaluation of a variety of representative trading strategies. In particular, defining a set of deltas $\mathcal{D} = \{0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8\}$ and a set of time-to-expiries $\mathcal{T} = \{30, 60, 91, 122, 152, 182, 273, 365, 547, 730\}$ (in calendar days), we consider:

(1) *Outright call* option $C(\tau, \Delta)$ with all deltas $\Delta \in \mathcal{D}$ and all time-to-expiries $\tau \in \mathcal{T}$. There are in total $7 \times 10 = 70$ outright test strategies.

(2) *Delta spread* $C(\tau, \Delta_1) - C(\tau, \Delta_2)$ with all distinguished delta pairs $\Delta_1 > \Delta_2 \in \mathcal{D}$ and all time-to-expiries $\tau \in \mathcal{T}$. There are in total $\binom{7}{2} \times 10 = 210$ delta spread test strategies.

(3) *Delta butterfly* $C(\tau, \Delta) + C(\tau, 1 - \Delta) - 2 \times C(\tau, 0.5)$ with all deltas $\Delta \in \mathcal{D} \setminus \{0.5\}$ and all time-to-expiries $\tau \in \mathcal{T}$. There are in total $3 \times 10 = 30$ delta butterfly spread test strategies.

(4) *Delta-hedged option* $C(\tau, \Delta) - \Delta \times S$ with $\Delta = 0.5$ and all time-to-expiries $\tau \in \mathcal{T}$. There are in total $1 \times 10 = 10$ delta-hedged option test strategies.

(5) *Delta-neural strangle* $C(\tau, \Delta) + C(\tau, 1 - \Delta) - S$ with all deltas $\Delta \in \mathcal{D} \setminus \{0.5\}$ and all time-to-expiries $\tau \in \mathcal{T}$. There are in total $3 \times 10 = 30$ delta-neutral strangle test strategies.
Note that a typical delta-neutral straggle involves simultaneously buying a call and a put, i.e. $C(\tau, \Delta) + P(\tau, -\Delta)$. Here we use put-call parity to express such straddles in terms of only calls and the underlyings, ignoring cash components.

(6) *Risk reversal* $C(\tau, \Delta) - C(\tau, 1 - \Delta) + S$ with all deltas $\Delta \in \{0.2, 0.3, 0.4\}$ and all time-to-expiries $\tau \in \mathcal{T}$. There are in total $3 \times 10 = 30$ risk reversal test strategies.
Note that a typical risk reversal consists of buying an OTM call and selling an OTM put, i.e. $C(\tau, \Delta) - P(\tau, -\Delta)$. Here we use put-call parity to express such straddles in terms of only calls and the underlyings, ignoring cash components.

(7) *Calendar spread* $C(\tau_1, \Delta) - C(\tau_2, \Delta)$ with $\Delta = 0.5$ and all distinguished time-to-expiry pairs $\tau_1 > \tau_2 \in \mathcal{T}$. There are in total $1 \times \binom{10}{2} = 45$ calendar spread test strategies.

(8) *VIX*, which is a linear combination of OTM call and put option prices, and can be further written as a linear combination of call prices only, provided that put-call parity holds under no-arbitrage; see details in [14].

In total, as we consider both long and short positions, there are $= 2 \times (70 + 210 + 30 + 10 + 30 + 30 + 45 + 1) = 852$ tested trading strategies. Among the tested trading strategies, delta butterflies, delta-hedged options, delta-neutral strangles and calendar spreads have low values of Black–Scholes delta.

F More VaR backtesting results

We show backtesting results, using the model in the main text, for 2-day, 5-day and 10-day VaRs for EURO STOXX 50 index option portfolios in Table 8. Similarly, we show backtesting results for 1-day, 2-day, 5-day and 10-day VaRs for DAX option portfolios in Table 9. Aligned with our observation in the main text, the nSDE-VaR model outperforms the FHS-VaR approach through the backtesting analysis in almost all metrics.

	2-day $\text{VaR}_{0.99}$		2-day $\text{VaR}_{0.95}$		5-day $\text{VaR}_{0.99}$		5-day $\text{VaR}_{0.95}$	
	nSDE	FHS	nSDE	FHS	nSDE	FHS	nSDE	FHS
Coverage ratio median	0.9960	0.9920	0.9602	0.9602	0.9960	0.9879	0.9677	0.9637
Coverage ratio mean	0.9950	0.9800	0.9602	0.9402	0.9944	0.9798	0.9618	0.9464
Kupiec PF (two-sided)	0.12%	13.11%	38.29%	35.71%	1.41%	12.76%	45.08%	37.12%
Kupiec PF (one-sided)	0.12%	13.11%	7.26%	15.69%	1.41%	12.76%	7.61%	14.05%

Table 8: Summary statistics on coverage tests for 2-day and 5-day VaRs computed by the neural-SDE (nSDE) market model and the filtered historical simulation (FHS) approach. These results are for EURO STOXX 50 index options.

	1-day $\mathrm{VaR}_{0.99}$		1-day $\mathrm{VaR}_{0.95}$	
	nSDE	FHS	nSDE	FHS
Coverage ratio median	0.9881	0.9841	0.9484	0.9524
Coverage ratio mean	0.9872	0.9726	0.9556	0.9314
Kupiec PF (two-sided)	14.64%	19.56%	33.49%	34.19%
Kupiec PF (one-sided)	14.64%	19.56%	6.21%	18.03%
Christoffersen independence	0.47%	7.61%	4.68%	10.42%
Conditional coverage	8.31%	19.44%	22.60%	33.49 %
Basel committee traffic light	67.3% 31.1% 0.8%	56.9% 28.7% 12.9%	N.A.	N.A.

	2-day $\mathrm{VaR}_{0.99}$		2-day $\mathrm{VaR}_{0.95}$		5-day $\mathrm{VaR}_{0.99}$		5-day $\mathrm{VaR}_{0.95}$	
	nSDE	FHS	nSDE	FHS	nSDE	FHS	nSDE	FHS
Coverage ratio median	0.9960	0.9920	0.9681	0.9482	0.9960	0.9839	0.9758	0.9637
Coverage ratio mean	0.9903	0.9811	0.9602	0.9335	0.9928	0.9778	0.9644	0.9453
Kupiec PF (two-sided)	8.31%	13.93%	54.33%	37.00%	0.00%	14.40%	58.90%	42.62%
Kupiec PF (one-sided)	8.31%	13.93%	10.42%	20.37%	0.00%	14.40%	8.78%	14.64%

	10-day $\mathrm{VaR}_{0.99}$		10-day $\mathrm{VaR}_{0.95}$	
	nSDE	FHS	nSDE	FHS
Coverage ratio median	1.0000	0.9794	0.9671	0.9630
Coverage ratio mean	0.9949	0.9742	0.9659	0.9367
Kupiec PF (two-sided)	0.12%	22.72%	43.44%	37.59%
Kupiec PF (one-sided)	0.12%	22.72%	7.61%	20.37%

Table 9: Summary statistics on coverage tests for 2-day, 5-day and 10-day VaRs computed by the neural-SDE (nSDE) market model and the filtered historical simulation (FHS) approach. These results are for DAX options.

In Figure 29 and 30, we show the time series of 1-day FHS-$\mathrm{VaR}_{0.99}$ breaches for delta-exposed and vega-exposed option portfolios respectively, and compare with those for positions in the underlying index and the VIX portfolio. These are results for EURO STOXX 50 index options, and are supposed to provide a comparison with Figure 15 and 16. The FHS-VaRs tend to have poor coverage performance for delta-hedged option portfolios.

In addition to the coverage and independence test results, in Figure 31 and 32, we show the comparison of more procyclicality results produced by the nSDE-VaRs and the FHS-VaRs for both index options. The nSDE-VaRs are consistently less procyclical than the FHS-VaRs in all scenarios.

Finally, we compare the VaR backtesting results for EURO STOXX 50 index options generated by the neural-SDE models with different configurations. Specifically, we have described the model with three primary factors (referred to as "nSDE-3") in Appendix C.1 and the model where a full diffusion matrix of (S, ξ) is specified (referred to as "nSDE-J") in Appendix C.2. In Table 10, we list the statistics on coverage and independence tests for 1-day, 2-day and 5-day VaRs computed by the two models. Compared with the original model for which the related 1-day VaR backtesting results are shown in Table 4, the 3-factor model slightly outperforms the original 2-factor model for most tests, with 17.3% more portfolios in the green zone. Nevertheless, when compared with the original model for which the related 2-day and 5-day VaR backtesting results are shown in Table 8, neither the 3-factor model nor the original 2-factor model is significantly better than the other.

G In-sample VaR backtesting

In addition to the out-of-sample VaR backtesting analysis, we show how in-sample nSDE-VaR looks in this section. Specifically, since we have used option data from 2002 to 2018 to train the neural-SDE model, the "in-sample" nSDE-VaRs then refer to the VaR time series from 2002 to 2018, computed using the trained neural-SDE model.

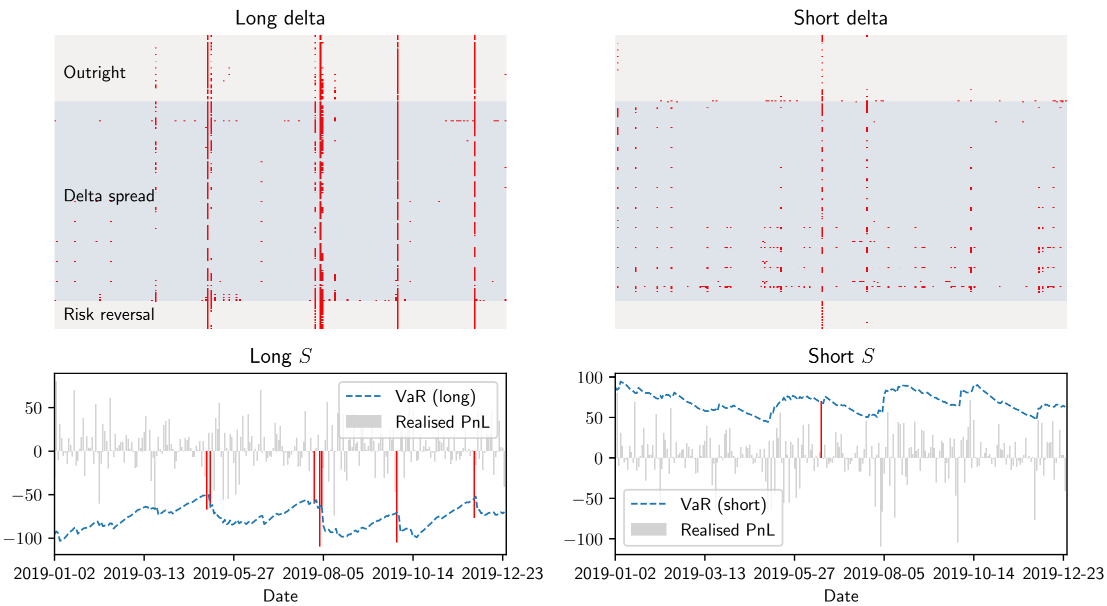

Figure 29: Times series of 1-day FHS-VaR$_{0.99}$ breaches (red dots) for delta-exposed option portfolios, compared with that for positions in the underlying index.

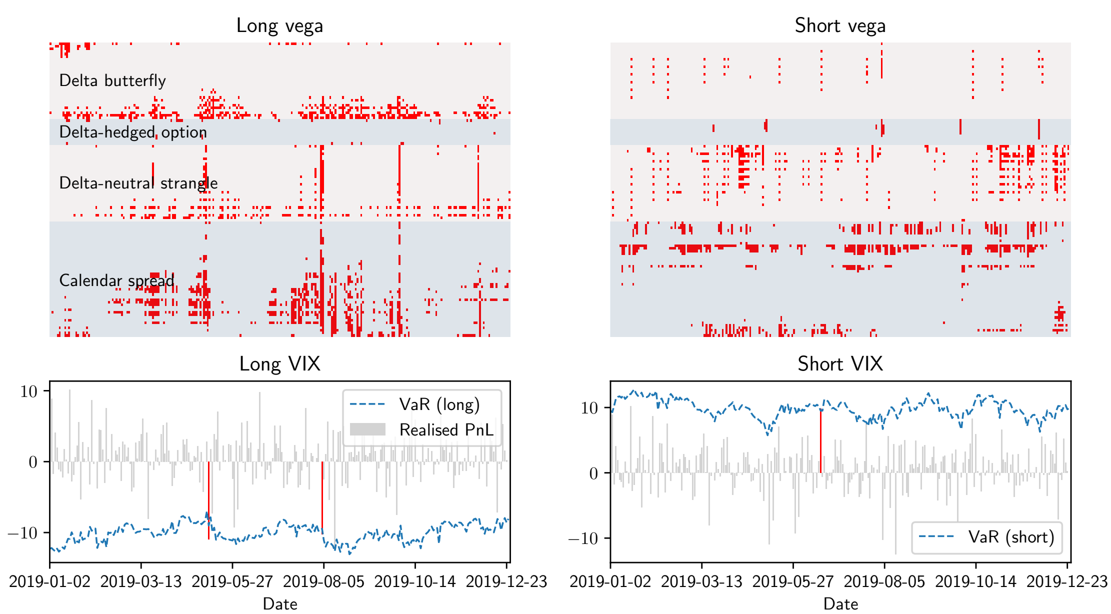

Figure 30: Times series of 1-day FHS-VaR$_{0.99}$ breaches (red dots) for delta-neutral and vega-exposed option portfolios, compared with that for positions in the VIX portfolio.

In Figure 33, we show the time series of 1-day realised PnLs, nSDE-VaR$_{0.99}$'s and the VaR breaches for long and short positions in EURO STOXX 50 index. The market experienced three notable stressed periods since 2002, corresponding to the three volatility clusterings in 2002 (the dot-com bubble bursting), 2008 (the global financial crisis triggered by the US subprime mortgage crisis) and 2015 (the European sovereign debt crisis). Overall, the nSDE-VaRs are reactive to changes in

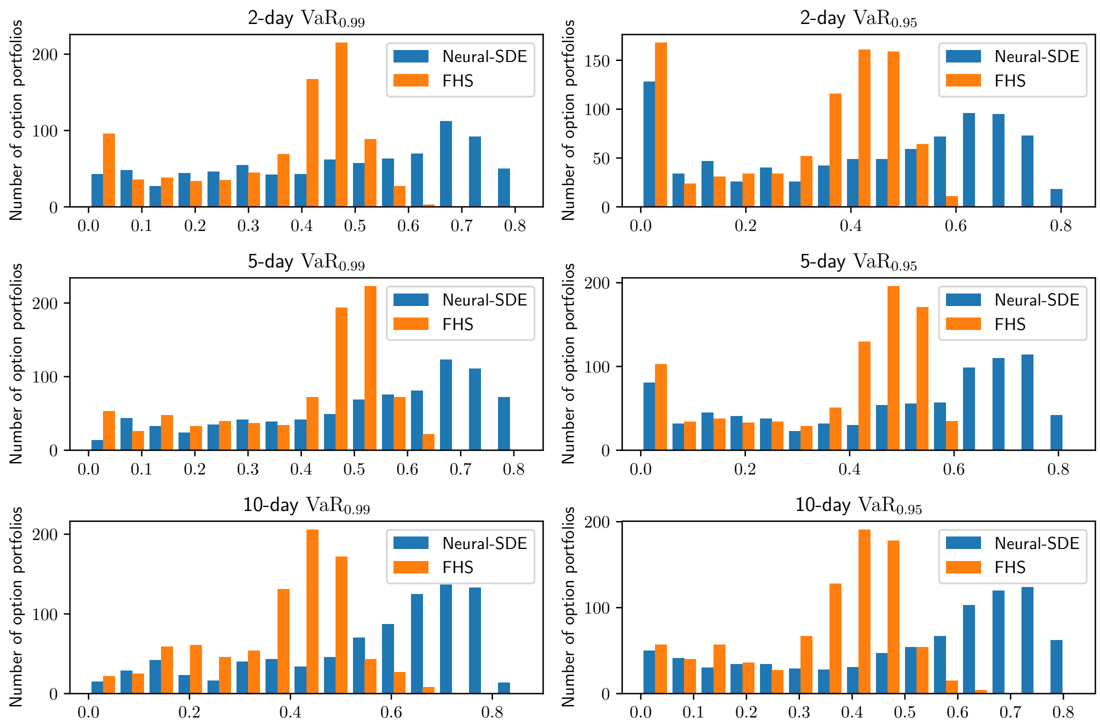

Figure 31: Distribution of trough-to-peak ratios over different tested portfolios of options on the EURO STOXX 50 index.

	1-day $\mathrm{VaR}_{0.99}$		1-day $\mathrm{VaR}_{0.95}$	
	nSDE-3	nSDE-J	nSDE-3	nSDE-J
Coverage ratio median	0.9881	0.9960	0.9643	0.9742
Coverage ratio mean	0.9891	0.9911	0.9603	0.9667
Kupiec PF (two-sided)	4.68%	11.12%	28.81%	51.41%
Kupiec PF (one-sided)	4.68%	11.12%	5.15%	6.32%

	2-day $\mathrm{VaR}_{0.99}$		2-day $\mathrm{VaR}_{0.95}$		5-day $\mathrm{VaR}_{0.99}$		5-day $\mathrm{VaR}_{0.95}$	
	nSDE-3	nSDE-J	nSDE-3	nSDE-J	nSDE-3	nSDE-J	nSDE-3	nSDE-J
Coverage ratio median	0.9960	1.0000	0.9641	0.9880	1.0000	1.0000	0.9637	0.9899
Coverage ratio mean	0.9946	0.9938	0.9634	0.9718	0.9966	0.9946	0.9694	0.9731
Kupiec PF (two-sided)	1.05%	3.04%	33.84%	65.11%	0.00%	3.75%	40.52%	75.76%
Kupiec PF (one-sided)	1.05%	3.04%	4.57%	6.91%	0.00%	3.75%	2.34%	8.08%

Table 10: Summary statistics on coverage and independence tests for 1-day, 2-day and 5-day VaRs computed by the neural-SDE market models of different configurations. These results are for EURO STOXX 50 options.

market conditions, and have produced statistically sufficient risk margins over time.

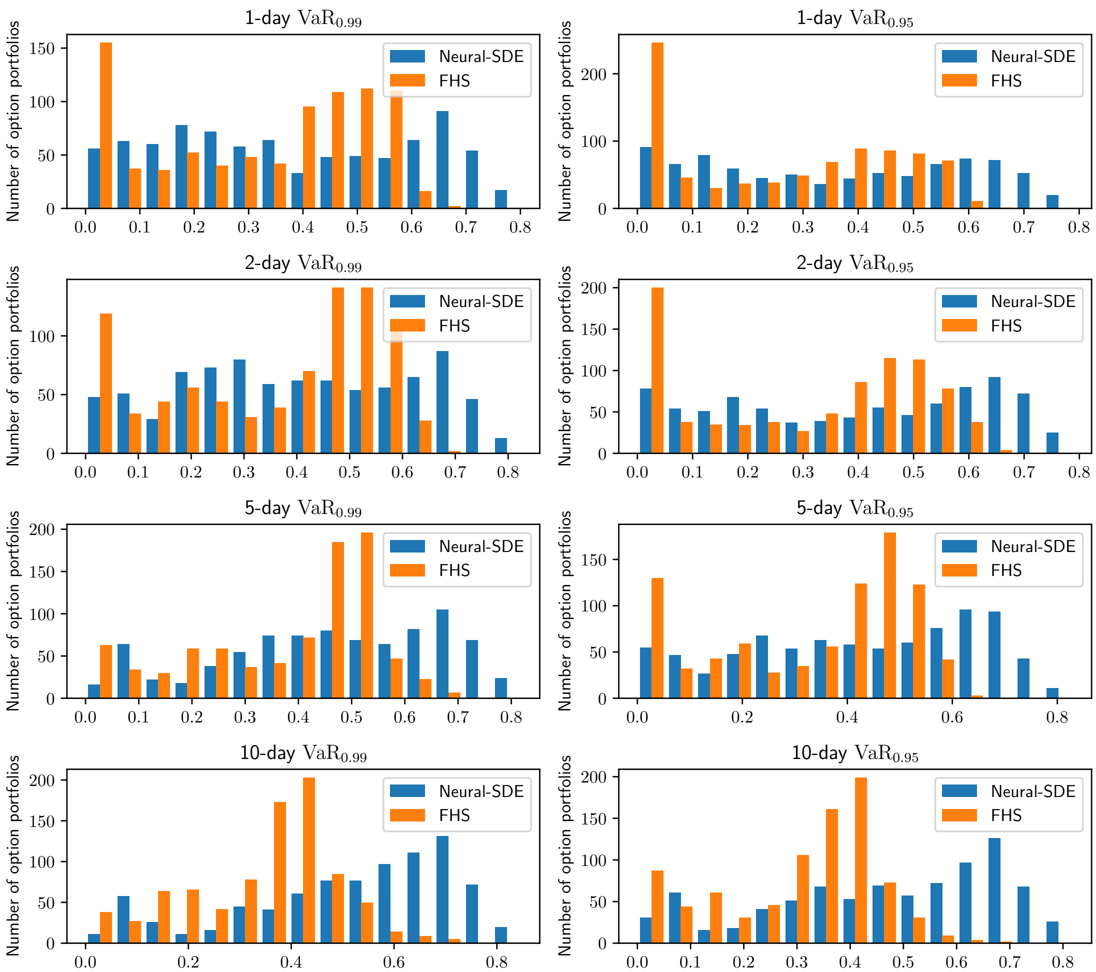

Figure 32: Distribution of trough-to-peak ratios over different tested portfolios of options on the DAX index.

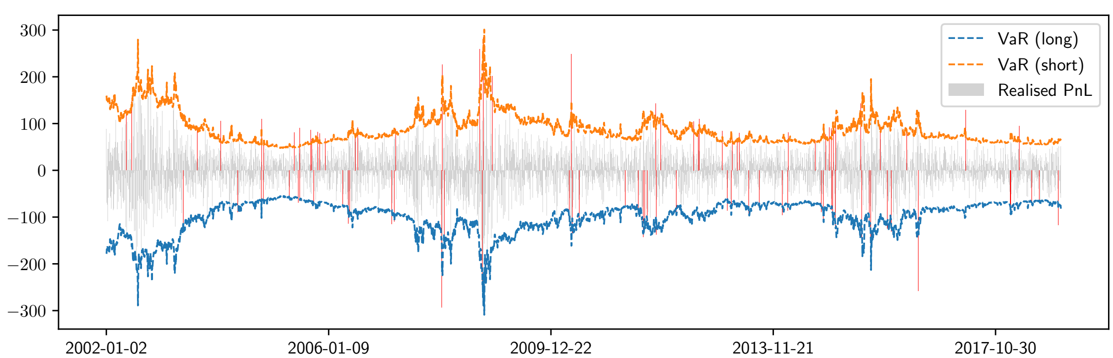

Figure 33: Time series of realised 1-day PnL and nSDE-$\mathrm{VaR}_{0.99}$ breaches (red bars) for long and short positions in the underlying index (EURO STOXX 50).