Bridging Theory and Data: Finance-Informed Neural Networks

Heqing Shi
MSBE Group, University of Edinburgh Business School
https://shiheqing.github.io
Heqing.Shi@ed.ac.uk

Joint work with
Yi Cao, Xi'an Jiaotong-Liverpool University
Zexun Chen, University of Edinburgh Business School

May 16, 2025

Motivation

Structural pricing models are still dominant in the option pricing context since good generalizability. But in the financial “big-data” era:

- their capacity of learning from data is limited.
- one structural specification does not adapt to all market conditions.
- the structural parameters can change rapidly.

Since Malliaris and Salchenberger (1993), Hutchinson et al. (1994), there were attempts to price options using nonparametric approaches.

But they have no economic awareness $\rightarrow$ can’t generalize well!

Motivation

There were also attempts to learn from data and economic theories.
$\hookrightarrow$ *e.g.*"*Teaching Economics to the Machines*" by Chen et al. (2023).

Chen et al. (2023) use transfer learning which:

- separates the learning of theory from empirical calibration,
- might lead to a mismatch when the theory diverges from the data,
- the learned theory can be forgotten once transferred to the data, akin to "catastrophic forgetting".

Motivation

Research Question: Can a nonparametric pricing model be encoded with parametric knowledge, such that economic theories and market data can be jointly learned? $\rightarrow$ Data-driven model + Structural specification.

- The data-driven part:
  - ✓ no model assumption
  - ✓ strong expressive power
  - ✗ prone to overfitting
  - ✗ violate economic principles
- The structural (model-driven) part:
  - ✓ parsimonious & analytical
  - ✓ less prone to overfitting
  - ✗ strong assumption
  - ✗ mis-specification/ identification

Literature

Nonparametric model only:

$\hookrightarrow$ pure data-drive, no economic awareness

- MLP [Malliaris and Salchenberger (1993)]
- RBF, MLP, PPR [Hutchinson et al. (1994)]
- SVM [Liang et al. (2009)]

Literature

With model-free constraints:

↪ only shape prior knowledge, no structural specification

- NN [Ackerer et al. (2020)]
- GP [Chataigner et al. (2021)]
- Sparsely-connected NN [Chataigner et al. (2020)]→special topology

With structural models:

- Transfer learning [Chen et al. (2023)]
  ↪ fixed structural model + theory overriding
- Residual learning [Almeida et al. (2023)]
  ↪ need to estimate a structural model first
- PINN [Aboussalah et al. (2024)]
  ↪ gradient pathologies, limited structural specifications applicable

This Paper

Contribution:

- We propose a novel Finance-Informed Neural Network setup which:
  - allows a battery of structural option pricing models to be encoded
  - learn the NN parameters and the hidden economic states jointly
  - work efficiently with standard NNs
- FINNs exhibit improved option pricing and hedging performance
  - panels of S&P 500 index options
  - 6.81%-10.63% (6.00%-9.86%) pricing error reduction, 3.69%-6.11% (1.21%-1.41%) hedging error reduction in average
  - robust to volatile market conditions

PINN in Finance

Standard PINN

Given a neural network $f_\theta(t, x)$, the nonlinear partial differential operation:

$$\partial_t f_\theta(t, x) + \mathfrak{N}_x [f_\theta(t, x); \phi], \tag{1}$$

can be computed under the f_θ representation using auto-differentiation.

Setting (1) to zero formulates a PDE parameterized by ϕ when f_θ is the solution. Raissi et al. (2019) defines the PDE residual as:

$$r(\theta; t, x) = \partial_t f_\theta(t, x) + \mathfrak{N}_x [f_\theta(t, x); \phi], \tag{2}$$

A standard PINN attempts to approximate such solution by minimizing a PDE-loss:

$$\mathcal{L}_{PDE}(\theta; t, x) = \|r(\theta; t, x)\|_{l_2}. \tag{3}$$

Standard PINN—Limitations

Consider the Black and Scholes (1973) PDE:

$$-\partial_\tau V + rS\partial_S V + \frac{1}{2}\sigma^2 S^2 \partial_{SS} V - rV = 0, \tag{4}$$

A PINN $f_\theta(X)$ is trained such that $f_\theta(X) = V$ satisfies (4).

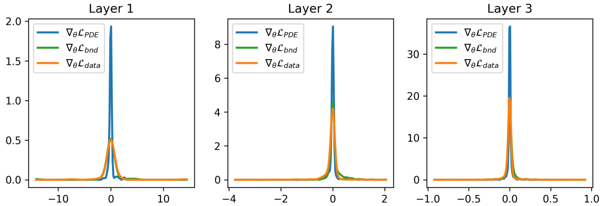
- Condition 1: X has to include the partial derivative variables.
  $\hookrightarrow$ Not applicable when those variables are unobservable, e.g.

$$-\partial_\tau V + \frac{1}{2} vS^2 \partial_{SS} V + \rho\sigma vS\partial_{Sv} V + \frac{1}{2}\sigma^2 v\partial_{vv} V + rS\partial_S V + \left[\kappa\left(\theta - v(t)\right) - \lambda\left(S, v, t\right)\right]\partial_v V - rV = 0$$

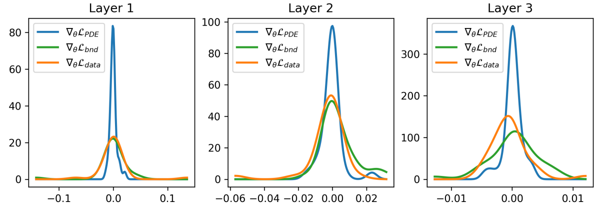
- Condition 2: the partial derivatives have to be learned accurately.
  $\hookrightarrow$ Deep chain rule suffers from vanishing gradients.

Standard PINN—Limitations

Figure 1: The vanishing gradient pathologies of a standard PINN. The loss gradients are obtained from the 5,000th iteration of the model training. The p.d.f. of $\nabla_\theta \mathcal{L}_{PDE}$ is strongly discrepant with that of $\nabla_\theta \mathcal{L}_{data}$, and is sharply centred around zero.



(a) Gradients w.r.t. neural network weights



(b) Gradients w.r.t. neural network biases

Our Model

FINN Setup

Instead of using the PDE formulation of a structural model, a FINN uses its analytical solution $g\left(X_{chara};\phi^*\right)$.

X_{chara} denotes option and market characteristics $\{S_t, r_t, K, \tau\}$ that are observable. ϕ^* denotes the hidden economic states that are unobservable.

$$V = g\left(X_{chara};\phi^*\right) \tag{5}$$

Hidden economic states refer to the unknown parameters in a structural option pricing model, e.g. the implied volatility in the Black-Scholes.

(5) determines the no-arbitrage option prices exactly given ϕ^*.

FINN Setup

A FINN can learn with arbitrary features $X \in \mathbb{R}^d$ as long as $\{K, \tau\} \subseteq X$.
$\hookrightarrow$ $f_\theta(X)$ and $g(X_{chara}; \phi^*)$ can be evaluated at the same $\{K, \tau\} \subseteq X$.

The parametric knowledge from a structural $g(X_{chara}; \phi^*)$ is encoded to a data-driven neural network $f_\theta(X)$ through a "fin(ance)-loss":

$$\mathcal{L}_{fin}(\theta, \phi^*; X) := \lambda_{fin} \| f_\theta(X) - g(X_{chara}; \phi^*) \|_{l_2}. \tag{6}$$

The data-driven knowledge is learned through a "data-loss":

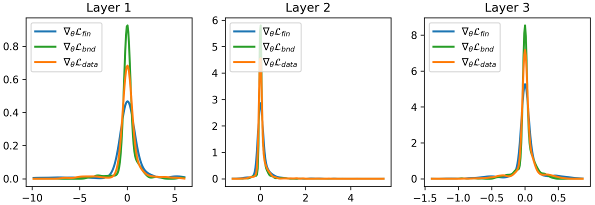
$$\mathcal{L}_{data}(\theta; X) := \lambda_{data} \| f_\theta(X) - V^{obs} \|_{l_2}. \tag{7}$$

The joint search of $\theta_{total} = \{\theta, \phi^*\}$ is solved by one optimization:

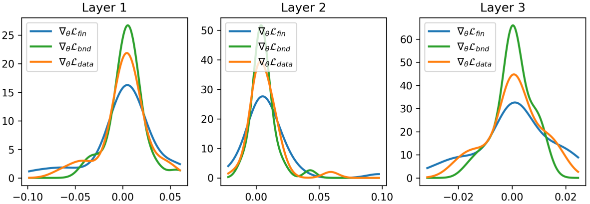
$$\operatorname*{arg\,min}_{\{\theta, \phi^*\}} \; \mathcal{L}_{data}(\theta; X) + \mathcal{L}_{fin}(\theta, \phi^*; X_{chara}). \tag{8}$$

FINN Setup

Figure 3: The p.d.f. of $\nabla_\theta \mathcal{L}_{fin}$ is close to that of $\nabla_\theta \mathcal{L}_{data}$ in a FINN.



(a) Gradients w.r.t. neural network weights



(b) Gradients w.r.t. neural network biases

FINN Setup

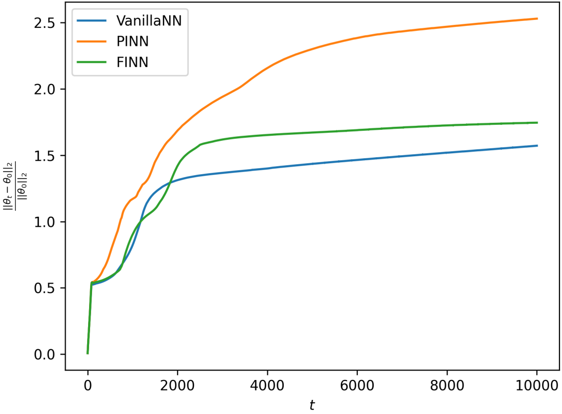


Figure 5: The relative change of the neural network parameters at the t-th iteration of model training compared to the parameters at the initialization. The relative change of parameters is expected to be stable after sufficient iterations of training.

$g\left(\phi^*\right)$ Specifications

Learnable Hidden Economic States

Structural models with closed-form or semi-closed-form solutions widely exist in the option pricing context. $\rightarrow$ A battery of $g\left(X_{chara};\phi^*\right)$ choices.

<u>Black-Scholes Model</u>: (FINN-BS)

$$\frac{dS_t}{S_t} = r\,dt + \sigma\,dW_t, \tag{9}$$

$$g^{BS}\left(X_{chara};\phi^* = \{\sigma\}\right) = V\left(S_t, r, \sigma, K, \tau\right), \tag{10}$$

$$= S_t\Phi\left(d_1\left(\phi^*\right)\right) - Ke^{-r\tau}\Phi\left(d_2\left(\phi^*\right)\right). \tag{11}$$

$\hookrightarrow \mathcal{L}_{fin}\left(t,\theta,\phi^*\right) = \lambda_{fin}\|f_t\left(\theta;X\right) - g_t^{BS}\left(\phi^*;X_{chara}\right)\|_{l_2}.$

Learnable Hidden Economic States

<u>**Ad-hoc Black-Scholes Model**</u>**:** (FINN-AHBS)

- Black-Scholes + smiled implied volatilities.
- Quadratic regression of the implied volatilities.

$$\sigma_{i,t} = \alpha_{0,t} + \alpha_{1,t} m_{i,t} + \alpha_{2,t} m_{i,t}^2 + \alpha_{3,t} \tau_{i,t} + \alpha_{4,t} \tau_{i,t}^2 + \alpha_{5,t} m_{i,t} \tau_{i,t} + \epsilon_{i,t}, \forall i \leq 1 \leq n, \tag{12}$$

where $\{(m_{i,t}, \tau_{i,t})\}_{i=1}^{n}$ is a cross-section of option characteristics at time t.

$$g^{AHBS}\left(X_{chara}; \phi^* = \{\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5\}\right) = S_t \Phi\left(d_1\left(\phi^*\right)\right) - Ke^{-r\tau} \Phi\left(d_2\left(\phi^*\right)\right). \tag{13}$$

$$\hookrightarrow \mathcal{L}_{fin}\left(t, \theta, \phi^*\right) = \lambda_{fin} \| f_t\left(\theta; X\right) - g_t^{AHBS}\left(\phi^*; X_{chara}\right) \|_{l_2}.$$

Learnable Hidden Economic States

Heston Stochastic Volatility Model: (FINN-HSV)

- The volatility is modelled as a separate OU process.
- Semi-closed-form solution with its characteristic function.
  $\hookrightarrow$Fourier-cosine series expansion of Fang and Oosterlee (2009).

$$dS_t = rS_t \, dt + \sqrt{v_t} S_t \, dW_t^1, \tag{14}$$

$$dv_t = \kappa \left(\theta - v_t\right) \, dt + \sigma_v \sqrt{v_t} \, dW_t^2, \tag{15}$$

$$\langle W^1, W^2 \rangle_t = \rho t. \tag{16}$$

$\rightarrow$The learnable hidden economic states $\phi^* = \{v_0, \sigma_v, \kappa, \theta, \rho\}$ in this case.

Learnable Hidden Economic States

The characteristic function of $\log(S_t)$ is written:

$$\psi_\tau(u;\phi^*) = \exp\left[C_\tau\left(u;\phi^*\right)\theta + D_\tau\left(u;\phi^*\right)v_0 + iu\log\left(S_t e^{r\tau}\right)\right], \tag{17}$$

where C_τ, D_τ are functions differentiable w.r.t. ϕ^*.

According to the COS method of Fang and Oosterlee (2009):

$$g^{HSV}\left(X;\phi^* = \{\theta, v_0, \sigma_v, \rho, \kappa\}\right) = e^{-r\tau}\sum_{k=0}^{N-1}{}' Re\left\{\psi_\tau\left(\frac{k\pi}{b-a}; S_t, r, \phi^*\right)e^{-ik\pi\frac{a}{b-a}}\right\}V_k\left(K\right). \tag{18}$$

$$\hookrightarrow \mathcal{L}_{fin}\left(t,\theta,\phi^*\right) = \lambda_{fin}\|f_t\left(\theta;X\right) - g_t^{HSV}\left(\phi^*;X_{chara}\right)\|_{l_2}.$$

Learnable Hidden Economic States

Martingale Option Pricing Approach: (FINN-MOPA)

- No asset price dynamics.
- Learnable hidden economic states inflation, e.g. 2,000.

The first fundamental theorem of asset pricing (FFTAP) tells that:

$$V_t(K,\tau) = D_t(T)\,\mathbb{E}^{\mathbb{Q}}\left[(S_T - K)^+ \,|\mathcal{F}_t\right] \tag{19}$$

$$\simeq D_t(T)\sum_{i=1}^{q}\left(S_T^{(i)} - K\right)^+ \pi^{\mathbb{Q}}\left(S_T^{(i)}\right), \tag{20}$$

assuming $S_T \in \left[S_T^{(1)}, S_T^{(2)}, \cdots, S_T^{(q)}\right]$ spans all possible terminal asset prices, and we know the probability distribution $\pi^{\mathbb{Q}}\left(S_T^{(i)}\middle| S_t\right) \forall 1 \le i \le q$.

Learnable Hidden Economic States

For each τ, q risk-neutral probabilities need to be estimated:

$$\pi^{\mathbb{Q}}_{s\times q} = \begin{bmatrix} \pi_{\tau_1,1} & \pi_{\tau_1,2} & \cdots & \pi_{\tau_1,q} \\ \pi_{\tau_2,1} & \pi_{\tau_2,2} & \cdots & \pi_{\tau_2,q} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{\tau_s,1} & \pi_{\tau_s,2} & \cdots & \pi_{\tau_s,q}. \end{bmatrix}. \tag{21}$$

$\forall T_k, 1 \le k \le s$, define $S_{T_k} \in \left[S^{(1)}_{T_k}, S^{(2)}_{T_k}, \cdots, S^{(q)}_{T_k}\right]$, which gives a matrix of terminal payoffs:

$$V_{T_k} = \begin{bmatrix} \left(S^{(1)}_{T_k} - K_1\right)^+ & \left(S^{(2)}_{T_k} - K_1\right)^+ & \cdots & \left(S^{(q)}_{T_k} - K_1\right)^+ \\ \left(S^{(1)}_{T_k} - K_2\right)^+ & \left(S^{(2)}_{T_k} - K_2\right)^+ & \cdots & \left(S^{(q)}_{T_k} - K_3\right)^+ \\ \vdots & \vdots & \ddots & \vdots \\ \left(S^{(1)}_{T_k} - K_h\right)^+ & \left(S^{(2)}_{T_k} - K_h\right)^+ & \cdots & \left(S^{(q)}_{T_k} - K_h\right)^+ \end{bmatrix}. \tag{22}$$

Learnable Hidden Economic States

$\rightarrow$ A $s \cdot h \times s \cdot q$ block-diagonal payoff matrix for all time-to-maturities:

$$V_T = \begin{bmatrix} V_{T_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & V_{T_2} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & V_{T_s} \end{bmatrix}. \tag{23}$$

By FFTAP, the no-arbitrage option prices can be calculated by:

$$g^{MOPA}\left(X_{chara}; \phi^* = \{\pi_{\tau_k, j}, 1 \leq k \leq s, 1 \leq j \leq q\}\right) = D \cdot \left(V_T\right)_{sh \times sq} \cdot \phi^*_{sq \times 1} \tag{24}$$

$$\hookrightarrow \mathcal{L}_{fin}\left(t, \theta, \phi^*\right) = \lambda_{fin} \| f_t\left(\theta; X\right) - g_t^{MOPA}\left(\phi^*; X_{chara}\right) \|_{l_2}.$$

Learnable Hidden Economic States

To ensure the regularity conditions of the learned probabilities:

- $\pi_{\tau_k,j} \geq 0, \forall 1 \leq k \leq s, q \leq j \leq q,$
- $\sum_{j=1}^{q} \pi_{\tau_k,j} = 1, \forall 1 \leq k \leq s,$

at each iteration, we apply a row-wise softmax transformation for $\pi^{\mathbb{Q}}_{s\times q}$:

$$\pi^{\mathbb{Q}}_{s\times q} = \begin{bmatrix} e^{\pi_{\tau_1,1}} \Big/ \sum_{i=1}^{q} e^{\pi_{\tau_1,i}} & e^{\pi_{\tau_1,2}} \Big/ \sum_{i=1}^{q} e^{\pi_{\tau_1,i}} & \cdots & e^{\pi_{\tau_1,q}} \Big/ \sum_{i=1}^{q} e^{\pi_{\tau_1,i}} \\ e^{\pi_{\tau_2,1}} \Big/ \sum_{i=1}^{q} e^{\pi_{\tau_2,i}} & e^{\pi_{\tau_2,2}} \Big/ \sum_{i=1}^{q} e^{\pi_{\tau_2,i}} & \cdots & e^{\pi_{\tau_2,q}} \Big/ \sum_{i=1}^{q} e^{\pi_{\tau_2,i}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{\pi_{\tau_s,1}} \Big/ \sum_{i=1}^{q} e^{\pi_{\tau_s,i}} & e^{\pi_{\tau_s,2}} \Big/ \sum_{i=1}^{q} e^{\pi_{\tau_s,i}} & \cdots & e^{\pi_{\tau_s,q}} \Big/ \sum_{i=1}^{q} e^{\pi_{\tau_s,i}} \end{bmatrix}. \tag{25}$$

FINN Setup

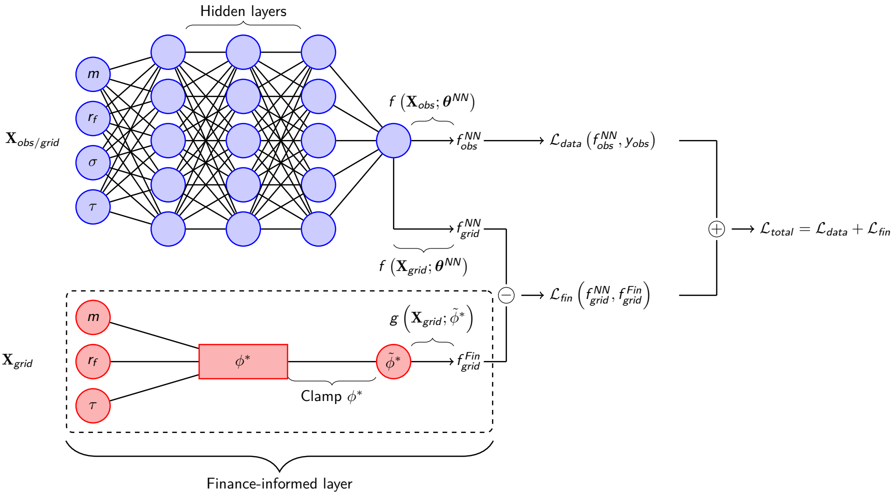


Figure 6: An overview of the FINN setup.

FINN Setup

Algorithm FINN training pseudo code.

Input: X, y^{obs}, initial θ_0^{NN}, ϕ_0^*

for epoch $= 1$ to N_{epochs} **do**

  for batch $= 1$ to N_{batch} **do**

    $f^{NN}(X_{\text{batch}}; \theta_{epoch,batch}) = f_L \circ h_{L-1} \circ f_{L-1} \circ \cdots \circ h_1 f_1 \left(X_{\text{batch}}; \theta_{epoch,batch}\right)$

    $f^{Fin}\left(X_{batch}; \phi^*_{epoch,batch}\right) = g\left(\{S, r, K, \tau\}_{batch}; \phi^*_{epoch,batch}\right)$

    $\mathcal{L}_{data} = \frac{1}{\sqrt{N_{\text{batch}}}} \|f^{NN} - y^{obs}\|_2$

    $\mathcal{L}_{fin} = \frac{1}{\sqrt{M}} \|f^{NN} - f^{Fin}\|_2$

    $\mathcal{L}_{total}\left(\{\theta, \phi^*\}_{epoch,batch}; X_{batch}\right) = \mathcal{L}_{data} + \mathcal{L}_{fin}$

    $\theta_{epoch,batch+1} = \theta_{epoch,batch} - \eta \nabla_\theta \mathcal{L}_{data} - \eta \nabla_\theta \mathcal{L}_{fin}$

    $\phi^*_{epoch,batch+1} = \phi^*_{epoch,batch} - \eta \nabla_{\phi^*} \mathcal{L}_{fin}$

    Clamp $\phi^*_{epoch,batch+1}$ by mathematical and economic regularity conditions.

  end for

end for

Simulation

Learn A Cross-Section

- Simulated economy according to the Black-Scholes.
- $S_t = 100$, $r = 15\%$, $\sigma = 20\%$.
- 10 equidistant strike prices from $[20, 180]$ and 10 equidistant expiries from $[0, 1]$. $\rightarrow$ 100 different option characteristics.
- Price noises $\sim \mathcal{N}\left(0, 10^2\right)$.

Learn A Cross-Section

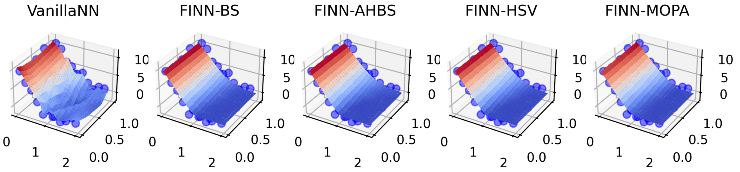


Figure 7: All neural networks have 3 hidden layers with 32 neurons each layer. The activation function is the rectified linear unit. Models are trained for 5,000 epochs with a learning rate of 1e-3, using Adam as the optimizer. The hidden economic states ϕ^* are $\{\sigma = 0.2603\}$, $\{\alpha_0 = 0.0523, \alpha_1 = 0, \alpha_2 = 0, \alpha_3 = 0.3025, \alpha_4 = 0.02, \alpha_5 = 0\}$, $\{\theta = 0.3482, v_0 = 0.0265, \sigma_v = 0.2819, \rho = -1, \kappa = 0.6895\}$, learned respectively by FINN-BS, FINN-AHBS and FINN-HSV.

Learn A Cross-Section

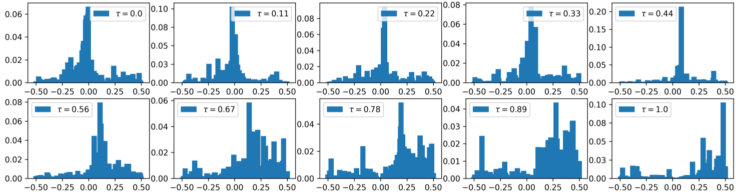


Figure 8: The hidden economic states ϕ^* (interpreted as the risk-neutral probabilities) learned by FINN-MOPA from the simulated Black-Scholes economy.

Empirical Analysis

Data

- Daily S&P 500 index call options sourced from OptionMetrics.
- Covering the period from 4 Jan 1996 to 30 Sep 2022.
- Options expiry in 7 days to 1 year.
- Train with 3-month option panels, 1-month rolling window.

Training Configuration

- Architecture: fully-connected
- Hidden layers: 3
- Neurons per layer: 32
- Parameters: 2,305
- Activation function: ReLU
- Learning rate: 1e-3
- Epochs: 500
- Optimizer: Adam
- Batch method: daily cross-section as a mini-batch

Benchmark Models

Benchmark Models

With Model-free Constraints:

- Reg(ularized) Der(ivatives) NN, Ackerer et al. (2020)
  - ▷ monotonicity, convexity constraints
  - ▷ auto-diff
- Ineq(uality)-Cons(trained) NN
  - ▷ discretized monotonicity, convexity constraints
  - ▷ quadratic programming, Cohen et al. (2020) $\rightarrow$ no need to auto-diff
  - ▷ option boundary conditions $\rightarrow$ Bounded-Ineq(uality)-Cons(trained) NN

Benchmark Models

With Structural Models:

- NN + Black-Scholes PDE $\rightarrow$ forward-problem PINN
  - ▷ structural PDE formulation
  - ▷ historical volatility, hVol (PDE-hVol)
- NN + Black-Scholes formula $\rightarrow$ forward-problem FINN
  - ▷ structural analytical formulation
  - ▷ historical volatility, hVol (Analy-hVol)
  - ▷ structural volatility implied in the Black-Scholes, sVol (Analy-sVol)

Structural Model:

- The Black-Scholes model

Fine-tuning

If one has different weights on the data and the structural information:

$$\mathcal{L}_{FINN}(\theta, \phi^*; X) := \gamma_{data}\mathcal{L}_{data}\left(\theta; X^{obs}\right) + \gamma_{fin}\mathcal{L}_{fin}\left(\theta, \phi^*; X'_{chara}\right). \quad (26)$$

$\rightarrow \gamma_{data} = 1, \gamma_{fin} \in \{1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024\}$

- Train models with different γ_{fin} on a 3-month training option dataset.
- Select the best model on a 1-month validation option dataset.
- Test the selected model on a further 1-month test option dataset.

Fine-tuning

Fine-tuning

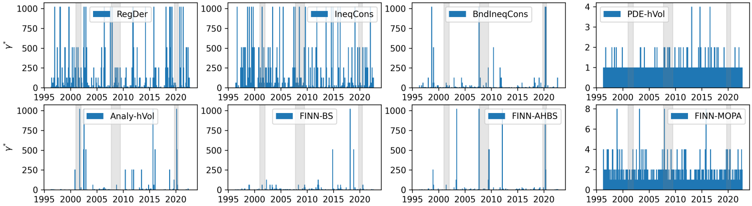


Figure 9: The fine-tuned γ_{fin} on validation option datasets according to different models. The shaded areas indicate NBER major recession periods.

Fine-tuning

	RegDer	IneqCons	BndIneqCons	PDE-hVol	Analy-hVol	FINN-BS	FINN-AHBS	FINN-MOPA
VIX	3.09*	2.03	3.93***	-0.00	3.96***	-0.08	2.31***	0.02**
	(1.85)	(1.88)	(0.82)	(0.00)	(0.88)	(0.49)	(0.84)	(0.01)
Constant	71.67*	106.57**	-52.74***	1.18***	-53.06***	15.84	-24.90	1.43***
	(41.05)	(41.64)	(18.21)	(0.06)	(19.42)	(10.93)	(18.51)	(0.17)
Obs	316	316	316	316	316	316	316	316
Adj. R^2	0.01	0.00	0.06	-0.00	0.06	-0.00	0.02	0.01
F-stat	2.78	1.17	22.84	0.59	20.43	0.03	7.63	4.16

Table 1: We estimate the regression $\gamma_{fin}(t) = \beta_0 + \beta_1 VIX(t) + \epsilon(t)$ for each model. $VIX(t)$ and $\gamma_{fin}(t)$ are from the same validation period. The number in the bracket is the standard error of the estimated coefficient.

Option Pricing

Pricing Error

- NNs with model-free constraints are sensitive to fine-tuning.
- FINNs are instead less sensitive to fine-tuning.
  $\hookrightarrow$ set $\gamma_{data} = 1, \gamma_{fin} = 1$
  $\hookrightarrow$ 3-month training, 2-month testing

Pricing Error

Table 2: **All market periods**. All models are tested using the options in the **first** month after the 3-month training option dataset, using the RMSE as the pricing error metric. We use the Wilcoxon (1947) sign-rank test to evaluate the predictive performance of two models. The table reports the p-values of the tests. Stars on the left indicate that the model in the row index statistically outperforms the model in the column index and vice versa.

		Model-free			**Forward Problem**			**FINN**			
Model	Vanilla	RegDer	IneqCons	BndIneqCons	PDE-hVol	Analy-hVol	Analy-sVol	BS	AHBS	MOPA	Structural-BS
Vanilla	–	*0.92	**0.96	0.45	***1.00	***1.00	**0.98	0.83	0.72	0.00***	***1.00
RegDer	–	–	0.87	0.19	***1.00	***1.00	*0.93	0.56	0.58	0.00***	***1.00
IneqCons	–	–	–	0.03**	***1.00	***1.00	0.82	0.20	0.23	0.00***	***1.00
BndIneqCons	–	–	–	–	***1.00	***1.00	***1.00	*0.91	0.83	0.00***	***1.00
PDE-hVol	–	–	–	–	–	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***
Analy-hVol	–	–	–	–	–	–	0.00***	0.00***	0.00***	0.00***	0.33
Analy-sVol	–	–	–	–	–	–	–	0.00***	0.01***	0.00***	***1.00
BS	–	–	–	–	–	–	–	–	0.53	0.00***	***1.00
AHBS	–	–	–	–	–	–	–	–	–	0.00***	***1.00
MOPA	–	–	–	–	–	–	–	–	–	–	***1.00
Structural-BS	–	–	–	–	–	–	–	–	–	–	–

Pricing Error

Table 3: **Low volatility market periods**. All models are tested using the options in the **first** month after the 3-month training option dataset, using the RMSE as the pricing error metric. We use the Wilcoxon (1947) sign-rank test to evaluate the predictive performance of two models. The table reports the p-values of the tests. Stars on the left indicate that the model in the row index statistically outperforms the model in the column index and vice versa.

		Model-free			Forward Problem			FINN			
Model	Vanilla	RegDer	IneqCons	BndIneqCons	PDE-hVol	Analy-hVol	Analy-sVol	BS	AHBS	MOPA	Structural-BS
Vanilla	–	0.68	0.81	*0.90	***1.00	**0.99	**0.98	*0.91	*0.91	0.03**	***1.00
RegDer	–	–	0.57	0.87	***1.00	**0.98	*0.91	0.88	0.70	0.03**	***1.00
IneqCons	–	–	–	0.67	***1.00	0.90	0.83	0.45	0.47	0.02**	**0.99
BndIneqCons	–	–	–	–	***1.00	*0.93	0.77	0.80	0.40	0.00***	***0.99
PDE-hVol	–	–	–	–	–	0.00***	0.00***	0.00***	0.00***	0.00***	0.06*
Analy-hVol	–	–	–	–	–	–	0.14	0.09*	0.01**	0.00***	*0.93
Analy-sVol	–	–	–	–	–	–	–	0.29	0.10	0.00***	**0.98
BS	–	–	–	–	–	–	–	–	0.22	0.00***	***1.00
AHBS	–	–	–	–	–	–	–	–	–	0.00***	***1.00
MOPA	–	–	–	–	–	–	–	–	–	–	***1.00
Structural-BS	–	–	–	–	–	–	–	–	–	–	–

Pricing Error

Table 4: **Median volatility market periods**. All models are tested using the options in the **first** month after the 3-month training option dataset, using the RMSE as the pricing error metric. We use the Wilcoxon (1947) sign-rank test to evaluate the predictive performance of two models. The table reports the p-values of the tests. Stars on the left indicate that the model in the row index statistically outperforms the model in the column index and vice versa.

		Model-free			Forward Problem			FINN			
Model	Vanilla	RegDer	IneqCons	BndIneqCons	PDE-hVol	Analy-hVol	Analy-sVol	BS	AHBS	MOPA	Structural-BS
Vanilla	–	*0.94	*0.93	0.38	***1.00	***1.00	*0.94	0.87	0.77	0.01***	***1.00
RegDer	–	–	0.86	0.16	***1.00	***1.00	0.85	0.55	0.64	0.00***	***1.00
IneqCons	–	–	–	0.04**	***1.00	***1.00	0.66	0.34	0.35	0.00***	***1.00
BndIneqCons	–	–	–	–	***1.00	***1.00	***0.99	0.87	0.89	0.00***	***1.00
PDE-hVol	–	–	–	–	–	0.04**	0.00***	0.00***	0.00***	0.00***	0.00***
Analy-hVol	–	–	–	–	–	–	0.00***	0.00***	0.00***	0.00***	0.05**
Analy-sVol	–	–	–	–	–	–	–	0.00***	0.11	0.00***	***1.00
BS	–	–	–	–	–	–	–	–	0.86	0.00***	***1.00
AHBS	–	–	–	–	–	–	–	–	–	0.00***	***1.00
MOPA	–	–	–	–	–	–	–	–	–	–	***1.00
Structural-BS	–	–	–	–	–	–	–	–	–	–	–

Pricing Error

Table 5: **High volatility market periods**. All models are tested using the options in the **first** month after the 3-month training option dataset, using the RMSE as the pricing error metric. We use the Wilcoxon (1947) sign-rank test to evaluate the predictive performance of two models. The table reports the p-values of the tests. Stars on the left indicate that the model in the row index statistically outperforms the model in the column index and vice versa.

		Model-free			Forward Problem			FINN			
Model	Vanilla	RegDer	IneqCons	BndIneqCons	PDE-hVol	Analy-hVol	Analy-sVol	BS	AHBS	MOPA	Structural-BS
Vanilla	–	0.50	0.68	0.18	***1.00	*0.94	0.61	0.27	0.17	0.06*	***0.99
RegDer	–	–	0.57	0.09*	***1.00	*0.93	0.62	0.24	0.29	0.08*	***0.99
IneqCons	–	–	–	0.09*	***1.00	*0.91	0.64	0.19	0.20	0.05*	***0.99
BndIneqCons	–	–	–	–	***1.00	**0.98	*0.93	0.59	0.55	0.04**	***1.00
PDE-hVol	–	–	–	–	–	0.07*	0.00***	0.00***	0.00***	0.00***	0.34
Analy-hVol	–	–	–	–	–	–	0.06*	0.01***	0.01***	0.00***	0.74
Analy-sVol	–	–	–	–	–	–	–	0.05**	0.03**	0.00***	***1.00
BS	–	–	–	–	–	–	–	–	0.22	0.09*	***1.00
AHBS	–	–	–	–	–	–	–	–	–	0.08*	***1.00
MOPA	–	–	–	–	–	–	–	–	–	–	***1.00
Structural-BS	–	–	–	–	–	–	–	–	–	–	–

Pricing Error

Table 6: **All market periods**. All models are tested using the options in the **second** month after the 3-month training option dataset, using the RMSE as the pricing error metric. We use the Wilcoxon (1947) sign-rank test to evaluate the predictive performance of two models. The table reports the p-values of the tests. Stars on the left indicate that the model in the row index statistically outperforms the model in the column index and vice versa.

		Model-free			Forward Problem			FINN			
Model	Vanilla	RegDer	IneqCons	BndIneqCons	PDE-hVol	Analy-hVol	Analy-sVol	BS	AHBS	MOPA	Structural-BS
Vanilla	–	0.49	*0.94	0.08*	***1.00	***0.99	0.02**	0.00***	0.00***	0.00***	0.19
RegDer	–	–	**0.96	0.03**	***1.00	**0.99	0.01**	0.00***	0.00***	0.00***	0.29
IneqCons	–	–	–	0.00***	***1.00	*0.93	0.00***	0.00***	0.00***	0.00***	0.07*
BndIneqCons	–	–	–	–	***1.00	***1.00	0.34	0.05*	0.04**	0.00***	0.73
PDE-hVol	–	–	–	–	–	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***
Analy-hVol	–	–	–	–	–	–	0.00***	0.00***	0.00***	0.00***	0.00***
Analy-sVol	–	–	–	–	–	–	–	0.00***	0.01**	0.00***	0.83
BS	–	–	–	–	–	–	–	–	0.30	0.00***	**0.96
AHBS	–	–	–	–	–	–	–	–	–	0.00***	***0.99
MOPA	–	–	–	–	–	–	–	–	–	–	***1.00
Structural-BS	–	–	–	–	–	–	–	–	–	–	–

Pricing Error

Table 7: **Low volatility market periods**. All models are tested using the options in the **second** month after the 3-month training option dataset, using the RMSE as the pricing error metric. We use the Wilcoxon (1947) sign-rank test to evaluate the predictive performance of two models. The table reports the p-values of the tests. Stars on the left indicate that the model in the row index statistically outperforms the model in the column index and vice versa.

		Model-free			Forward Problem			FINN			
Model	Vanilla	RegDer	IneqCons	BndIneqCons	PDE-hVol	Analy-hVol	Analy-sVol	BS	AHBS	MOPA	Structural-BS
Vanilla	–	0.32	0.64	0.24	***1.00	0.57	0.36	0.31	0.45	0.01**	0.63
RegDer	–	–	0.78	0.49	***1.00	0.83	0.51	0.62	0.61	0.09*	0.76
IneqCons	–	–	–	0.28	***1.00	0.34	0.27	0.19	0.36	0.02**	0.53
BndIneqCons	–	–	–	–	***1.00	0.88	0.50	0.78	0.63	0.02**	0.86
PDE-hVol	–	–	–	–	–	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***
Analy-hVol	–	–	–	–	–	–	0.06*	0.09*	0.13	0.00***	0.58
Analy-sVol	–	–	–	–	–	–	–	0.48	0.69	0.01***	0.89
BS	–	–	–	–	–	–	–	–	0.48	0.05**	0.78
AHBS	–	–	–	–	–	–	–	–	–	0.01***	0.83
MOPA	–	–	–	–	–	–	–	–	–	–	**0.99
Structural-BS	–	–	–	–	–	–	–	–	–	–	–

Pricing Error

Table 8: **Median volatility market periods**. All models are tested using the options in the **second** month after the 3-month training option dataset, using the RMSE as the pricing error metric. We use the Wilcoxon (1947) sign-rank test to evaluate the predictive performance of two models. The table reports the p-values of the tests. Stars on the left indicate that the model in the row index statistically outperforms the model in the column index and vice versa.

		Model-free			Forward Problem			FINN			
Model	Vanilla	RegDer	IneqCons	BndIneqCons	PDE-hVol	Analy-hVol	Analy-sVol	BS	AHBS	MOPA	Structural-BS
Vanilla	–	0.75	0.90	0.23	***1.00	***1.00	0.14	0.01**	0.00***	0.00***	0.11
RegDer	–	–	0.76	0.05**	***1.00	***1.00	0.03**	0.00***	0.00***	0.00***	0.07*
IneqCons	–	–	–	0.05**	***1.00	***1.00	0.02**	0.00***	0.00***	0.00***	0.06*
BndIneqCons	–	–	–	–	***1.00	***1.00	0.44	0.04**	0.04**	0.00***	0.31
PDE-hVol	–	–	–	–	–	0.14	0.00***	0.00***	0.00***	0.00***	0.00***
Analy-hVol	–	–	–	–	–	–	0.00***	0.00***	0.00***	0.00***	0.00***
Analy-sVol	–	–	–	–	–	–	–	0.01***	0.01**	0.00***	0.21
BS	–	–	–	–	–	–	–	–	0.37	0.00***	0.57
AHBS	–	–	–	–	–	–	–	–	–	0.00***	0.79
MOPA	–	–	–	–	–	–	–	–	–	–	**0.99
Structural-BS	–	–	–	–	–	–	–	–	–	–	–

Pricing Error

Table 9: **High volatility market periods**. All models are tested using the options in the **second** month after the 3-month training option dataset, using the RMSE as the pricing error metric. We use the Wilcoxon (1947) sign-rank test to evaluate the predictive performance of two models. The table reports the p-values of the tests. Stars on the left indicate that the model in the row index statistically outperforms the model in the column index and vice versa.

		Model-free			Forward Problem			FINN			
Model	Vanilla	RegDer	IneqCons	BndIneqCons	PDE-hVol	Analy-hVol	Analy-sVol	BS	AHBS	MOPA	Structural-BS
Vanilla	–	0.29	0.84	0.13	***1.00	0.41	0.02**	0.01**	0.01***	0.00***	0.45
RegDer	–	–	**0.96	0.12	***1.00	0.35	0.06*	0.01**	0.01**	0.00***	0.67
IneqCons	–	–	–	0.02**	***1.00	0.21	0.01**	0.00***	0.00***	0.00***	0.36
BndIneqCons	–	–	–	–	***1.00	0.49	0.26	0.09*	0.06*	0.02**	0.85
PDE-hVol	–	–	–	–	–	0.01***	0.00***	0.00***	0.00***	0.00***	0.05*
Analy-hVol	–	–	–	–	–	–	0.06*	0.02**	0.01**	0.00***	0.56
Analy-sVol	–	–	–	–	–	–	–	0.02**	0.05**	0.01***	**0.97
BS	–	–	–	–	–	–	–	–	0.32	0.24	***0.99
AHBS	–	–	–	–	–	–	–	–	–	0.21	***0.99
MOPA	–	–	–	–	–	–	–	–	–	–	**0.99
Structural-BS	–	–	–	–	–	–	–	–	–	–	–

Pricing Error

- All FINNs statistically outperform the benchmark models when tested for longer periods (the second month).
- The outperformance is mainly attributed to the lower pricing error during the median-high volatility periods.

Pricing Error

Table 10: We estimate the regression $\Delta RMSE(t) = \beta_0 + \beta_1 AvgVIX(t) + \epsilon(t)$ for each model, where $\Delta RMSE(t)$ denotes the model's RMSE minus the RMSE of a plain vanilla neural network at the period t and $AvgVIX(t)$ denotes the average VIX during the period. A statistically significant negative β_1 implies that the model has lower pricing error compared with a plain vanilla neural network when the market is more volatile. This table uses the models' performance in **all market periods**, when considering the options in the **second** month after the 3-month training option dataset.

	Model-free			**Forward Problem**			**FINN**			
	RegDer	IneqCons	BndIneqCons	PDE-hVol	Analy-hVol	Analy-sVol	BS	AHBS	MOPA	Structural-BS
AvgVIX	0.0041	0.0088	-0.0013	0.0098	-0.0451**	-0.0757***	-0.0870***	-0.0840***	-0.0932***	-0.0317
	(0.0070)	(0.0073)	(0.0099)	(0.0128)	(0.0177)	(0.0137)	(0.0140)	(0.0137)	(0.0135)	(0.0214)
Constant	-0.0010	-0.0010	-0.0016	0.0032	0.0111***	0.0111***	0.0125***	0.0119***	0.0123***	0.0019
	(0.0015)	(0.0016)	(0.0022)	(0.0028)	(0.0039)	(0.0030)	(0.0031)	(0.0030)	(0.0030)	(0.0047)
Obs	317	317	317	317	317	317	317	317	317	317
Adj. R^2	-0.0021	0.0015	-0.0031	-0.0013	0.0172	0.0853	0.1066	0.1030	0.1280	0.0038
F-stat	0.3462	1.4692	0.0172	0.5908	6.5188	30.4658	38.7086	37.3012	47.3977	2.1941

Pricing Error

Table 11: We estimate the regression $\Delta RMSE(t) = \beta_0 + \beta_1 AvgVIX(t) + \epsilon(t)$ for each model, where $\Delta RMSE(t)$ denotes the model's RMSE minus the RMSE of a plain vanilla neural network at the period t and $AvgVIX(t)$ denotes the average VIX during the period. A statistically significant negative β_1 implies that the model has lower pricing error compared with a plain vanilla neural network when the market is more volatile. This table uses the models' performance in **low volatility market periods**, when considering the options in the **second** month after the 3-month training option dataset.

	Model-free			Forward Problem			FINN			
	RegDer	IneqCons	BndIneqCons	PDE-hVol	Analy-hVol	Analy-sVol	BS	AHBS	MOPA	Structural-BS
AvgVIX	-0.0078	-0.0605	-0.0863	-0.2260*	-0.0374	-0.0717	-0.1015	-0.1107	-0.1249	-0.1979
	(0.0969)	(0.1167)	(0.1155)	(0.1239)	(0.1325)	(0.1228)	(0.1301)	(0.1177)	(0.1198)	(0.1657)
Constant	0.0001	0.0074	0.0089	0.0323**	0.0033	0.0067	0.0105	0.0120	0.0118	0.0230
	(0.0120)	(0.0144)	(0.0143)	(0.0153)	(0.0164)	(0.0152)	(0.0161)	(0.0146)	(0.0148)	(0.0205)
Obs	64	64	64	64	64	64	64	64	64	64
Adj. R^2	-0.0160	-0.0117	-0.0071	0.0356	-0.0148	-0.0106	-0.0062	-0.0018	0.0014	0.0067
F-stat	0.0064	0.2687	0.5576	3.3278	0.0798	0.3411	0.6087	0.8855	1.0876	1.4274

Pricing Error

Table 12: We estimate the regression $\Delta RMSE(t) = \beta_0 + \beta_1 AvgVIX(t) + \epsilon(t)$ for each model, where $\Delta RMSE(t)$ denotes the model's RMSE minus the RMSE of a plain vanilla neural network at the period t and $AvgVIX(t)$ denotes the average VIX during the period. A statistically significant negative β_1 implies that the model has lower pricing error compared with a plain vanilla neural network when the market is more volatile. This table uses the models' performance in **median volatility market periods**, when considering the options in the **second** month after the 3-month training option dataset.

	Model-free			**Forward Problem**			**FINN**			
	RegDer	IneqCons	BndIneqCons	PDE-hVol	Analy-hVol	Analy-sVol	BS	AHBS	MOPA	Structural-BS
AvgVIX	-0.0226	0.0002	-0.0807***	0.0311	0.0246	-0.0841**	-0.0806**	-0.0891***	-0.0814**	-0.0708
	(0.0168)	(0.0187)	(0.0288)	(0.0384)	(0.0500)	(0.0346)	(0.0353)	(0.0334)	(0.0326)	(0.0525)
Constant	0.0044	0.0006	0.0136**	-0.0013	-0.0005	0.0129*	0.0115*	0.0130**	0.0102	0.0086
	(0.0033)	(0.0036)	(0.0056)	(0.0075)	(0.0098)	(0.0067)	(0.0069)	(0.0065)	(0.0064)	(0.0102)
Obs	189	189	189	189	189	189	189	189	189	189
Adj. R^2	0.0042	-0.0053	0.0353	-0.0018	-0.0041	0.0255	0.0220	0.0315	0.0271	0.0043
F-stat	1.7951	0.0002	7.8710	0.6544	0.2414	5.9238	5.2228	7.1209	6.2424	1.8196

Pricing Error

Table 13: We estimate the regression $\Delta RMSE(t) = \beta_0 + \beta_1 AvgVIX(t) + \epsilon(t)$ for each model, where $\Delta RMSE(t)$ denotes the model's RMSE minus the RMSE of a plain vanilla neural network at the period t and $AvgVIX(t)$ denotes the average VIX during the period. A statistically significant negative β_1 implies that the model has lower pricing error compared with a plain vanilla neural network when the market is more volatile. This table uses the models' performance in **high volatility market periods**, when considering the options in the **second** month after the 3-month training option dataset.

	Model-free			**Forward Problem**			**FINN**			
	RegDer	IneqCons	BndIneqCons	PDE-hVol	Analy-hVol	Analy-sVol	BS	AHBS	MOPA	Structural-BS
AvgVIX	0.0194	0.0054	0.0355	-0.0235	-0.1675***	-0.1675***	-0.1855***	-0.1639***	-0.2143***	-0.0364
	(0.0234)	(0.0225)	(0.0282)	(0.0375)	(0.0547)	(0.0468)	(0.0474)	(0.0489)	(0.0464)	(0.0783)
Constant	-0.0063	0.0005	-0.0135	0.0155	0.0521***	0.0440***	0.0477***	0.0406**	0.0556***	0.0051
	(0.0078)	(0.0075)	(0.0095)	(0.0126)	(0.0183)	(0.0157)	(0.0159)	(0.0164)	(0.0155)	(0.0262)
Obs	64	64	64	64	64	64	64	64	64	64
Adj. R^2	-0.0049	-0.0152	0.0092	-0.0097	0.1173	0.1576	0.1852	0.1395	0.2442	-0.0126
F-stat	0.6899	0.0581	1.5851	0.3941	9.3755	12.7900	15.3210	11.2097	21.3601	0.2165

Option Hedging

Hedging Error

Consider a hedging portfolio according to Hutchinson et al. (1994):

$$H(t) = V_{Spot}(t) + V_{Bond}(t) + V_{Call}(t). \tag{27}$$

We initialize a delta-hedged portfolio at t by setting:

$$V_{Spot}(t) = S_t \Delta^{model}(t), \tag{28}$$

$$V_{Call}(t) = -V^{obs}(t), \tag{29}$$

$$V_{Bond}(t) = -\left(V_{Spot}(t) + V_{Call}(t)\right). \tag{30}$$

Hedging Error

We calculate the delta-hedging portfolio value at $t+1$ by:

$$V_{Spot}(t+1) = S_{t+1}\Delta^{model}(t), \tag{31}$$

$$V_{Call}(t+1) = -V^{obs}(t+1), \tag{32}$$

$$V_{Bond}(t+1) = e^{r\times\frac{1}{252}}V_{Bond}(t). \tag{33}$$

The hedging capability of a model is measured by the mean hedging error:

$$MHE\left(t,T\right) = \frac{1}{T-t}\sum_{i=t}^{T-1}\left|H\left(t+1\right)\right|, \tag{34}$$

and the standard deviation of the hedging error:

$$SHE\left(t,T\right) = \sqrt{\mathbb{V}ar\left(\left|H\left(t+1\right)\right|\right)}, t \leq 0 \leq T-1. \tag{35}$$

Hedging Error

Table 14: **All market periods**. All models are tested using the options in the **first** month after the 3-month training option dataset, using MHE and SHE as the hedging error metrics. We use the Wilcoxon (1947) sign-rank test to evaluate the predictive performance of two models. The table reports the p-values of the tests. Stars on the left indicate that the model in the row index statistically outperforms the model in the column index and vice versa.

(a) MHE

Model	Vanilla	FINN-BS	FINN-AHBS	FINN-MOPA	BndIneqCons	Structural-BS	hVol-BS	implVol-BS
Vanilla	–	0.00***	0.00***	0.11	0.27	***0.99	***1.00	**0.98
FINN-BS	–	–	0.31	***1.00	***1.00	***1.00	***1.00	***1.00
FINN-AHBS	–	–	–	***1.00	***1.00	***1.00	***1.00	***1.00
FINN-MOPA	–	–	–	–	0.86	***1.00	***1.00	***1.00
BndIneqCons	–	–	–	–	–	***1.00	***1.00	***1.00
Structural-BS	–	–	–	–	–	–	***1.00	0.06*
hVol-BS	–	–	–	–	–	–	–	0.00***
implVol-BS	–	–	–	–	–	–	–	–

(b) SHE

Model	Vanilla	FINN-BS	FINN-AHBS	FINN-MOPA	BndIneqCons	Structural-BS	hVol-BS	implVol-BS
Vanilla	–	0.00***	0.00***	0.11	0.27	***0.99	***1.00	**0.98
FINN-BS	–	–	0.31	***1.00	***1.00	***1.00	***1.00	***1.00
FINN-AHBS	–	–	–	***1.00	***1.00	***1.00	***1.00	***1.00
FINN-MOPA	–	–	–	–	0.86	***1.00	***1.00	***1.00
BndIneqCons	–	–	–	–	–	***1.00	***1.00	***1.00
Structural-BS	–	–	–	–	–	–	***1.00	0.06*
hVol-BS	–	–	–	–	–	–	–	0.00***
implVol-BS	–	–	–	–	–	–	–	–

Hedging Error

Table 16: **Low volatility market periods**. All models are tested using the options in the **first** month after the 3-month training option dataset, using MHE and SHE as the hedging error metrics. We use the Wilcoxon (1947) sign-rank test to evaluate the predictive performance of two models. The table reports the p-values of the tests. Stars on the left indicate that the model in the row index statistically outperforms the model in the column index and vice versa.

(a) MHE

Model	Vanilla	FINN-BS	FINN-AHBS	FINN-MOPA	BndIneqCons	Structural-BS	hVol-BS	implVol-BS
Vanilla	–	0.19	0.18	0.62	0.63	***1.00	***1.00	**0.98
FINN-BS	–	–	0.59	**0.97	*0.91	***1.00	***1.00	***1.00
FINN-AHBS	–	–	–	*0.95	0.81	***1.00	***1.00	***1.00
FINN-MOPA	–	–	–	–	0.36	***1.00	***1.00	**0.98
BndIneqCons	–	–	–	–	–	***1.00	***1.00	**0.99
Structural-BS	–	–	–	–	–	–	0.71	0.01***
hVol-BS	–	–	–	–	–	–	–	0.00***
implVol-BS	–	–	–	–	–	–	–	–

(b) SHE

Model	Vanilla	FINN-BS	FINN-AHBS	FINN-MOPA	BndIneqCons	Structural-BS	hVol-BS	implVol-BS
Vanilla	–	0.23	0.13	0.56	0.60	***0.99	***1.00	**0.98
FINN-BS	–	–	0.61	*0.94	*0.91	***1.00	***1.00	***1.00
FINN-AHBS	–	–	–	*0.95	0.85	***1.00	***1.00	***1.00
FINN-MOPA	–	–	–	–	0.35	***1.00	***1.00	**0.99
BndIneqCons	–	–	–	–	–	***1.00	***1.00	***0.99
Structural-BS	–	–	–	–	–	–	0.77	0.02**
hVol-BS	–	–	–	–	–	–	–	0.00***
implVol-BS	–	–	–	–	–	–	–	–

Hedging Error

Table 18: **Median volatility market periods**. All models are tested using the options in the **first** month after the 3-month training option dataset, using MHE and SHE as the hedging error metrics. We use the Wilcoxon (1947) sign-rank test to evaluate the predictive performance of two models. The table reports the p-values of the tests. Stars on the left indicate that the model in the row index statistically outperforms the model in the column index and vice versa.

(a) MHE

Model	Vanilla	FINN-BS	FINN-AHBS	FINN-MOPA	BndIneqCons	Structural-BS	hVol-BS	implVol-BS
Vanilla	–	0.07*	0.01**	0.46	0.65	***1.00	***1.00	**0.98
FINN-BS	–	–	0.32	***1.00	***1.00	***1.00	***1.00	***1.00
FINN-AHBS	–	–	–	***1.00	***1.00	***1.00	***1.00	***1.00
FINN-MOPA	–	–	–	–	0.73	***1.00	***1.00	***1.00
BndIneqCons	–	–	–	–	–	***1.00	***1.00	**0.99
Structural-BS	–	–	–	–	–	–	***1.00	0.01***
hVol-BS	–	–	–	–	–	–	–	0.00***
implVol-BS	–	–	–	–	–	–	–	–

(b) SHE

Model	Vanilla	FINN-BS	FINN-AHBS	FINN-MOPA	BndIneqCons	Structural-BS	hVol-BS	implVol-BS
Vanilla	–	0.12	0.02**	0.69	0.74	***1.00	***1.00	**0.99
FINN-BS	–	–	0.34	***1.00	***1.00	***1.00	***1.00	***1.00
FINN-AHBS	–	–	–	***1.00	***1.00	***1.00	***1.00	***1.00
FINN-MOPA	–	–	–	–	0.60	***1.00	***1.00	***1.00
BndIneqCons	–	–	–	–	–	***1.00	***1.00	**0.99
Structural-BS	–	–	–	–	–	–	***1.00	0.02**
hVol-BS	–	–	–	–	–	–	–	0.00***
implVol-BS	–	–	–	–	–	–	–	–

Hedging Error

Table 20: **High volatility market periods**. All models are tested using the options in the **first** month after the 3-month training option dataset, using MHE and SHE as the hedging error metrics. We use the Wilcoxon (1947) sign-rank test to evaluate the predictive performance of two models. The table reports the p-values of the tests. Stars on the left indicate that the model in the row index statistically outperforms the model in the column index and vice versa.

(a) MHE

Model	Vanilla	FINN-BS	FINN-AHBS	FINN-MOPA	BndIneqCons	Structural-BS	hVol-BS	implVol-BS
Vanilla	–	0.00***	0.00***	0.01***	0.01**	0.05**	0.23	0.21
FINN-BS	–	–	0.30	***1.00	***1.00	*0.94	**0.98	***1.00
FINN-AHBS	–	–	–	***1.00	***1.00	**0.96	**0.97	***1.00
FINN-MOPA	–	–	–	–	*0.94	0.39	0.77	0.83
BndIneqCons	–	–	–	–	–	0.23	0.53	0.49
Structural-BS	–	–	–	–	–	–	***1.00	**0.99
hVol-BS	–	–	–	–	–	–	–	0.75
implVol-BS	–	–	–	–	–	–	–	–

(b) SHE

Model	Vanilla	FINN-BS	FINN-AHBS	FINN-MOPA	BndIneqCons	Structural-BS	hVol-BS	implVol-BS
Vanilla	–	0.00***	0.00***	0.03**	0.09*	0.32	0.32	0.22
FINN-BS	–	–	0.72	***1.00	***1.00	***1.00	***1.00	***1.00
FINN-AHBS	–	–	–	***1.00	***1.00	***1.00	***0.99	***1.00
FINN-MOPA	–	–	–	–	0.86	0.73	*0.91	0.87
BndIneqCons	–	–	–	–	–	0.69	0.64	0.62
Structural-BS	–	–	–	–	–	–	0.89	0.65
hVol-BS	–	–	–	–	–	–	–	0.76
implVol-BS	–	–	–	–	–	–	–	–

Hedging Error

Table 22: **All market periods**. All models are tested using the options in the **second** month after the 3-month training option dataset, using MHE and SHE as the hedging error metrics. We use the Wilcoxon (1947) sign-rank test to evaluate the predictive performance of two models. The table reports the p-values of the tests. Stars on the left indicate that the model in the row index statistically outperforms the model in the column index and vice versa.

(a) MHE

Model	Vanilla	FINN-BS	FINN-AHBS	FINN-MOPA	BndIneqCons	Structural-BS	hVol-BS	implVol-BS
Vanilla	–	0.00***	0.00***	0.13	0.15	**0.98	***1.00	*0.94
FINN-BS	–	–	0.55	***1.00	***1.00	***1.00	***1.00	***1.00
FINN-AHBS	–	–	–	***1.00	***1.00	***1.00	***1.00	***1.00
FINN-MOPA	–	–	–	–	0.80	***1.00	***1.00	**0.99
BndIneqCons	–	–	–	–	–	***1.00	***1.00	*0.94
Structural-BS	–	–	–	–	–	–	***1.00	0.03**
hVol-BS	–	–	–	–	–	–	–	0.00***
implVol-BS	–	–	–	–	–	–	–	–

(b) SHE

Model	Vanilla	FINN-BS	FINN-AHBS	FINN-MOPA	BndIneqCons	Structural-BS	hVol-BS	implVol-BS
Vanilla	–	0.00***	0.00***	0.09*	0.51	**0.99	***1.00	0.85
FINN-BS	–	–	0.51	***1.00	***1.00	***1.00	***1.00	***1.00
FINN-AHBS	–	–	–	***1.00	***1.00	***1.00	***1.00	***1.00
FINN-MOPA	–	–	–	–	**0.98	***1.00	***1.00	*0.93
BndIneqCons	–	–	–	–	–	***0.99	***1.00	0.80
Structural-BS	–	–	–	–	–	–	***1.00	0.01***
hVol-BS	–	–	–	–	–	–	–	0.00***
implVol-BS	–	–	–	–	–	–	–	–

Hedging Error

Table 24: **Low volatility market periods**. All models are tested using the options in the **second** month after the 3-month training option dataset, using MHE and SHE as the hedging error metrics. We use the Wilcoxon (1947) sign-rank test to evaluate the predictive performance of two models. The table reports the p-values of the tests. Stars on the left indicate that the model in the row index statistically outperforms the model in the column index and vice versa.

(a) MHE

Model	Vanilla	FINN-BS	FINN-AHBS	FINN-MOPA	BndIneqCons	Structural-BS	hVol-BS	implVol-BS
Vanilla	–	0.49	0.43	0.76	0.71	***1.00	***1.00	**0.98
FINN-BS	–	–	0.86	**0.96	0.79	***1.00	***1.00	*0.95
FINN-AHBS	–	–	–	0.76	0.69	***1.00	***1.00	*0.93
FINN-MOPA	–	–	–	–	0.61	***1.00	***1.00	0.81
BndIneqCons	–	–	–	–	–	***1.00	***1.00	0.85
Structural-BS	–	–	–	–	–	–	0.35	0.00***
hVol-BS	–	–	–	–	–	–	–	0.00***
implVol-BS	–	–	–	–	–	–	–	–

(b) SHE

Model	Vanilla	FINN-BS	FINN-AHBS	FINN-MOPA	BndIneqCons	Structural-BS	hVol-BS	implVol-BS
Vanilla	–	0.36	0.42	0.77	0.77	***0.99	***1.00	*0.92
FINN-BS	–	–	0.85	***0.99	*0.94	***1.00	***1.00	**0.96
FINN-AHBS	–	–	–	*0.94	*0.91	***1.00	***1.00	*0.95
FINN-MOPA	–	–	–	–	0.50	***1.00	***1.00	0.85
BndIneqCons	–	–	–	–	–	***0.99	***0.99	0.82
Structural-BS	–	–	–	–	–	–	0.14	0.00***
hVol-BS	–	–	–	–	–	–	–	0.00***
implVol-BS	–	–	–	–	–	–	–	–

Hedging Error

Table 26: **Median volatility market periods**. All models are tested using the options in the **second** month after the 3-month training option dataset, using MHE and SHE as the hedging error metrics. We use the Wilcoxon (1947) sign-rank test to evaluate the predictive performance of two models. The table reports the p-values of the tests. Stars on the left indicate that the model in the row index statistically outperforms the model in the column index and vice versa.

(a) MHE

Model	Vanilla	FINN-BS	FINN-AHBS	FINN-MOPA	BndIneqCons	Structural-BS	hVol-BS	implVol-BS
Vanilla	–	0.06*	0.01***	0.33	0.48	***0.99	***1.00	*0.95
FINN-BS	–	–	0.69	***1.00	***1.00	***1.00	***1.00	***1.00
FINN-AHBS	–	–	–	***1.00	***1.00	***1.00	***1.00	***1.00
FINN-MOPA	–	–	–	–	0.63	***1.00	***1.00	*0.91
BndIneqCons	–	–	–	–	–	***1.00	***1.00	*0.93
Structural-BS	–	–	–	–	–	–	***1.00	0.01***
hVol-BS	–	–	–	–	–	–	–	0.00***
implVol-BS	–	–	–	–	–	–	–	–

(b) SHE

Model	Vanilla	FINN-BS	FINN-AHBS	FINN-MOPA	BndIneqCons	Structural-BS	hVol-BS	implVol-BS
Vanilla	–	0.04**	0.00***	0.29	0.80	**0.98	***1.00	0.87
FINN-BS	–	–	0.59	***1.00	***1.00	***1.00	***1.00	***1.00
FINN-AHBS	–	–	–	***1.00	***1.00	***1.00	***1.00	***1.00
FINN-MOPA	–	–	–	–	*0.94	***0.99	***1.00	0.90
BndIneqCons	–	–	–	–	–	**0.98	***1.00	0.83
Structural-BS	–	–	–	–	–	–	***1.00	0.00***
hVol-BS	–	–	–	–	–	–	–	0.00***
implVol-BS	–	–	–	–	–	–	–	–

Hedging Error

Table 28: **High volatility market periods**. All models are tested using the options in the **second** month after the 3-month training option dataset, using MHE and SHE as the hedging error metrics. We use the Wilcoxon (1947) sign-rank test to evaluate the predictive performance of two models. The table reports the p-values of the tests. Stars on the left indicate that the model in the row index statistically outperforms the model in the column index and vice versa.

(a) MHE

Model	Vanilla	FINN-BS	FINN-AHBS	FINN-MOPA	BndIneqCons	Structural-BS	hVol-BS	implVol-BS
Vanilla	–	0.00***	0.00***	0.02**	0.01***	0.03**	0.11	0.19
FINN-BS	–	–	0.08*	***1.00	***1.00	0.87	***0.99	***1.00
FINN-AHBS	–	–	–	***1.00	***1.00	*0.93	***0.99	***1.00
FINN-MOPA	–	–	–	–	0.86	0.36	0.79	*0.94
BndIneqCons	–	–	–	–	–	0.17	0.45	0.44
Structural-BS	–	–	–	–	–	–	***1.00	***1.00
hVol-BS	–	–	–	–	–	–	–	0.82
implVol-BS	–	–	–	–	–	–	–	–

(b) SHE

Model	Vanilla	FINN-BS	FINN-AHBS	FINN-MOPA	BndIneqCons	Structural-BS	hVol-BS	implVol-BS
Vanilla	–	0.00***	0.00***	0.01***	0.03**	0.20	0.26	0.21
FINN-BS	–	–	0.12	***1.00	***1.00	*0.92	**0.96	**0.98
FINN-AHBS	–	–	–	***1.00	***1.00	**0.97	**0.98	**0.99
FINN-MOPA	–	–	–	–	**0.96	0.53	0.73	0.67
BndIneqCons	–	–	–	–	–	0.47	0.53	0.38
Structural-BS	–	–	–	–	–	–	**0.97	**0.97
hVol-BS	–	–	–	–	–	–	–	0.79
implVol-BS	–	–	–	–	–	–	–	–

Concluding Remarks

Concluding Remarks

- We propose a FINN setup to encode structural option pricing models into a data-driven neural network, which
  - ▷ incorporates parametric knowledge into data knowledge,
  - ▷ admits a battery of structural models,
  - ▷ identifies hidden economic states jointly,
  - ▷ is computationally and economically more feasible than PINN.
- We systematically assess the FINNs pricing and hedging capability
  - ▷ using real S&P 500 index call options, covering a long span of time,
  - ▷ against a wide range of benchmark models from the related literature,
    ↪ FINNs outperform others over a longer prediction horizon.
    ↪ the outperformance is more evident during volatile market periods.

Thank you!

References I

Aboussalah, A.M., Li, X., Chi, C., Patel, R., 2024. The AI Black-Scholes: Finance-informed neural network. arXiv preprint arXiv:2412.12213 .

Ackerer, D., Tagasovska, N., Vatter, T., 2020. Deep smoothing of the implied volatility surface. Advances in Neural Information Processing Systems 33, 11552–11563.

Almeida, C., Fan, J., Freire, G., Tang, F., 2023. Can a machine correct option pricing models? Journal of Business & Economic Statistics 41, 995–1009.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of political economy 81, 637–654.

Chataigner, M., Cousin, A., Crépey, S., Dixon, M., Gueye, D., 2021. Beyond surrogate modeling: Learning the local volatility via shape constraints. SIAM Journal on Financial Mathematics 12, SC58–SC69.

Chataigner, M., Crépey, S., Dixon, M., 2020. Deep local volatility. Risks 8, 82.

References II

Chen, H., Cheng, Y., Liu, Y., Tang, K., 2023. Teaching economics to the machines. Available at SSRN 4642167.

Cohen, S.N., Reisinger, C., Wang, S., 2020. Detecting and repairing arbitrage in traded option prices. Applied Mathematical Finance 27, 345–373.

Fang, F., Oosterlee, C.W., 2009. A Novel Pricing Method for European Options Based on Fourier-Cosine Series Expansions. SIAM Journal on Scientific Computing 31, 826–848. doi:10.1137/080718061.

Hutchinson, J.M., Lo, A.W., Poggio, T., 1994. A nonparametric approach to pricing and hedging derivative securities via learning networks. The journal of Finance 49, 851–889.

Liang, X., Zhang, H., Xiao, J., Chen, Y., 2009. Improving option price forecasts with neural networks and support vector regressions. Neurocomputing 72, 3055–3065.

References III

Malliaris, M., Salchenberger, L., 1993. A neural network model for estimating option prices. Applied Intelligence 3, 193–206.

Raissi, M., Perdikaris, P., Karniadakis, G.E., 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational physics 378, 686–707.

Wilcoxon, F., 1947. Probability tables for individual comparisons by ranking methods. Biometrics 3, 119–122.