Accurate Neural Network Option Pricing Methods with Control Variate Techniques and Data Synthesis/Cleaning with Financial Rationality

Chia-Wei Hsu, Tian-Shyr Dai, Chuan-Ju Wang, Ying-Ping Chen

國立陽明交通大學
NATIONAL YANG MING CHIAO TUNG UNIVERSITY

The 33rd ACM International Conference on Information and Knowledge Management, October 21–25, 2024

中央研究院
ACADEMIA SINICA

01

Introduction

What are options?
Challenges in Pricing Options

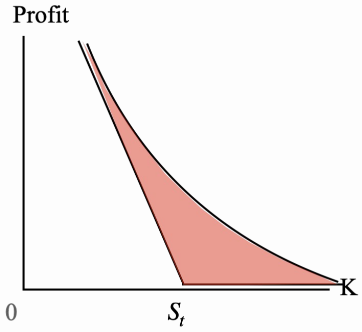
What are options?

Strike Price

- Financial instruments giving the right to buy or sell an underlying assets at a specified price on a maturity date.

- Categories
  - In the Money (ITM): Has intrinsic value and time value.
  - Out of the Money (OTM): No intrinsic value, only time value.
- Components
  - Intrinsic Value: Immediate payoff if exercised.
  - Time Value: Future potential until expiration.

- Take a call option as an example ...



In the case of a call option

S&P 500 INDEX (^spx)

2024-10-17 14:31:34 ET (Delayed)

Underlying Asset Price

Bid: **5845.1699** Ask: **5847.5898** Vol: **0** Last: **5,846.37** Change: **+3.9001 (+0.0668%)**

Options Chain

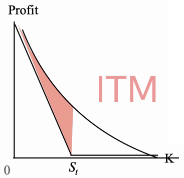
Strike Price ↓

Call Option Price↑

Total Records: 2498

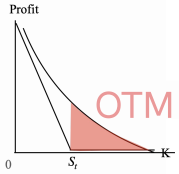
Calls | **Thu Oct 31 2024** | Puts

Last	Net	Bid	Ask	Vol	IV	Delta	Gamma	Int	Strike	Last	Net	Bid	Ask	Vol	IV	Delta	Gamma	Int
4,434.3	+0	4441.4	4455.2	0	2.09	0.9999	0	2	SPXW 1400.000	0.05	+0	0	0.05	0	1.97	-0.0001	0	586
0	+0	4241.2	4255.6	0	0	0.9999	0	1	SPXW 1600.000	0.05	+0	0	0.05	0	1.79	-0.0001	0	861
3,920.59	+0	4042.2	4056	0	1.73	0.9999	0	1	SPXW 1800.000	0.15	+0	0	0.05	0	1.63	-0.0001	0	379
21.17	-0.78	21.1	21.3	89	0.11	0.2836	0.0027	429	SPXW 5930.000	85.4	-11.9	92.8	95.6	1	0.11	-0.7164	0.0027	14
22.48	+2.08	19.6	19.8	38	0.11	0.2693	0.0027	277	SPXW 5935.000	89.3	+0	96.2	99.1	0	0.11	-0.7307	0.0027	4
22.02	+3.07	18.1	18.4	132	0.11	0.2552	0.0026	619	SPXW 5940.000	130.61	+0	99.8	102.6	0	0.11	-0.7448	0.0026	12
14.4	+0	16.7	17	0	0.11	0.2415	0.0026	267	SPXW 5945.000	0	+0	103.4	106.3	0	0.11	-0.7585	0.0026	0
16	-0.35	15.4	15.7	170	0.11	0.2281	0.0025	6978	SPXW 5950.000	95.5	-16.1	107.3	109.8	10	0.11	-0.7719	0.0025	471
18.34	+3.24	14.2	14.5	71	0.1	0.2151	0.0024	240	SPXW 5955.000	100	-15.35	110.8	113.8	1	0.11	-0.7849	0.0024	1
15	+1	13.1	13.3	143	0.1	0.2025	0.0023	1142	SPXW 5960.000	125.54	+0	114.7	117.6	0	0.11	-0.7975	0.0023	11
0.1	+0	0	0.1	0	0.28	0.0004	0	43	SPXW 7000.000	0	+0	1133.6	1148.1	0	0.32	-0.9996	0	0
0.2	+0	0	0.1	0	0.32	0.0002	0	403	SPXW 7200.000	0	+0	1333.2	1347.7	0	0.37	-0.9998	0	0
0.05	+0	0	0.05	0	0.34	0.0001	0	751	SPXW 7400.000	1,632.05	+0	1532.5	1547.3	0	0.38	-0.9999	0	0
0.07	+0.02	0	0.1	7	0.4	0.0001	0	34	SPXW 7600.000	1,729.75	-12.8	1732.4	1746.9	7	0.45	-0.9999	0	4
0.05	+0	0	0.05	0	0.41	0.0001	0	1123	SPXW 7800.000	0	+0	1931.6	1947.2	0	0.51	-0.9999	0	0



ITM

Intrinsic value (=S-K > 0)



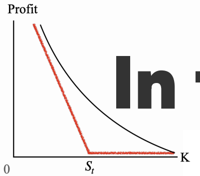
OTM

Intrinsic value (= 0)

Puts: OTM (upper rows), ITM (lower rows)

https://www.cboe.com/delayed_quotes/spx/quote_table

In the case of a call option



Intrinsic Value

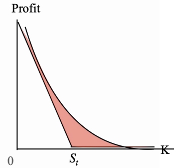
S&P 500 INDEX (^spx)

2024-10-17 14:54:35 ET (Delayed) Bid: **5849.3901** Ask: **5851.3501** Vol: **0** Last: **5,850.29** Change: **+7.82 (+0.1338%)**

Options Chain

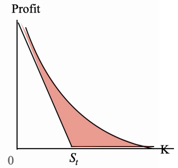
Time to Maturity ↓
Call Option Price↓

Total Records: 32



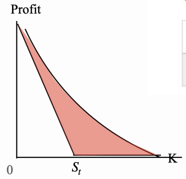
Fri Oct 18 2024

Calls Last	Net	Bid	Ask	Vol	IV	Delta	Gamma	Int	Strike	Puts Last	Net	Bid	Ask	Vol	IV	Delta	Gamma	Int
9.9	-3.05	9.7	10.1	1,339	0.1	0.4687	0.0144	462	SPX 5855.000	13.9	-8.15	12	12.5	1,987	0.1	-0.5295	0.0144	362



Fri Nov 15 2024

Calls Last	Net	Bid	Ask	Vol	IV	Delta	Gamma	Int	Strike	Puts Last	Net	Bid	Ask	Vol	IV	Delta	Gamma	Int
106.98	+0.38	106.1	107	50	0.15	0.5353	0.0016	418	SPX 5855.000	85.7	-9.95	89.5	90.2	8	0.15	-0.4648	0.0016	241



Fri Dec 20 2024

Calls Last	Net	Bid	Ask	Vol	IV	Delta	Gamma	Int	Strike	Puts Last	Net	Bid	Ask	Vol	IV	Delta	Gamma	Int
162.1	-1.45	163.4	164	196	0.15	0.5557	0.0011	3772	SPX 5855.000	123.5	-6.85	124.3	124.9	128	0.15	-0.4451	0.0011	3896

Time Value

https://www.cboe.com/delayed_quotes/spx/quote_table

Challenges in Pricing Options

- Option pricing is crucial for market efficiency.
- Model-based approaches suffers from model misspecification problem.

  Black, F. and Scholes, M. (1973) The Pricing of Options and Corporate Liabilities. Journal of Political Economy, 81, 637-654.
  Kou, S. G. 2002. A jump-diffusion model for option pricing. Management Science 48(8):1086–1101.
- Model-free approaches without rationality may have arbitrage opportunities.

  Liu, S., Oosterlee, C. W., & Bohte, S. M. (2019). Pricing options and computing implied volatilities using neural networks. Risks, 7(1), 16.
  Wei, X., Xie, Z., Cheng, R., & Li, Q. (2020). A cnn based system for predicting the implied volatility and option prices.
  Ge, M., Zhou, S., Luo, S., & Tian, B. (2021, November). 3D Tensor-based Deep Learning Models for Predicting Option Price. In 2021 International Conference on Information Science and Communications
- Model-free approaches with rationality may require unnecessary synthesized data.

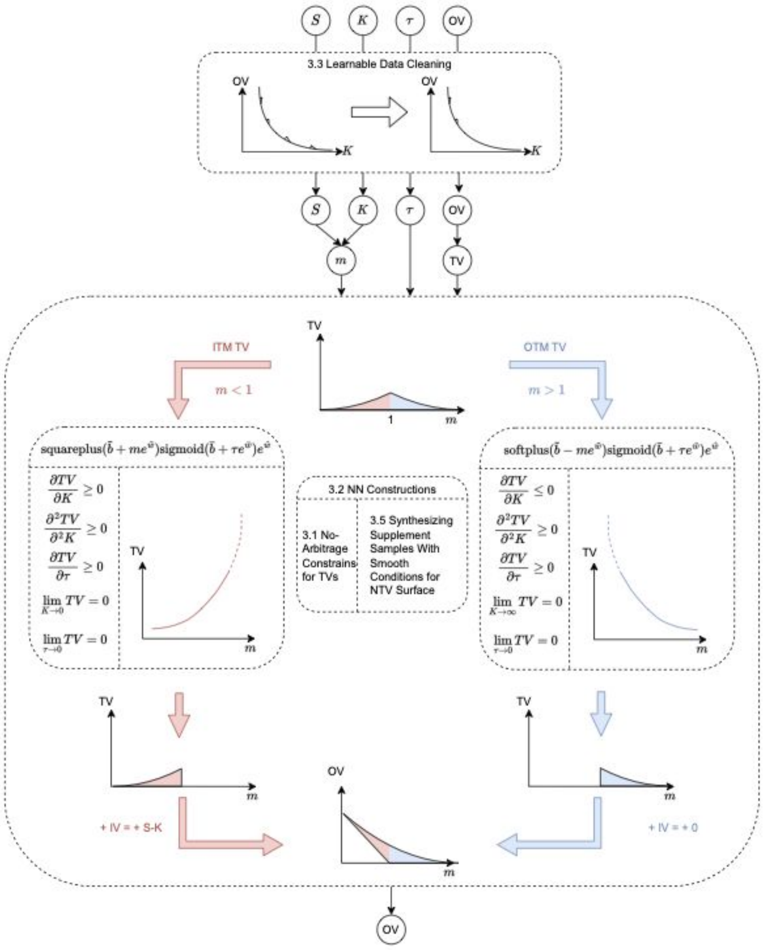
  Yang, Y., Zheng, Y., & Hospedales, T. (2017, February). Gated neural networks for option pricing: Rationality by design. In Proceedings of the AAAI conference on artificial intelligence (Vol. 31, No. 1).
  Ackerer, D., Tagasovska, N., & Vatter, T. (2020). Deep smoothing of the implied volatility surface. Advances in Neural Information Processing Systems, 33, 11552-11563.
  Zheng, Y., Yang, Y., & Chen, B. (2021, August). Incorporating prior financial domain knowledge into neural networks for implied volatility surface prediction. In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining (pp. 3968-3975).
- Stale price in illiquid markets may reduce pricing accuracy.

02

Proposed Method

Control variate approach
- Time value no-arbitrage constraints
- Remove unnecessary synthesized data
- Kink of time value surface

Cleaning data



Control Variates Approach

Option value based no-arbitrage constraints

$$\frac{\partial OV}{\partial K} \leq 0$$

$$\frac{\partial^2 OV}{\partial^2 K} \geq 0$$

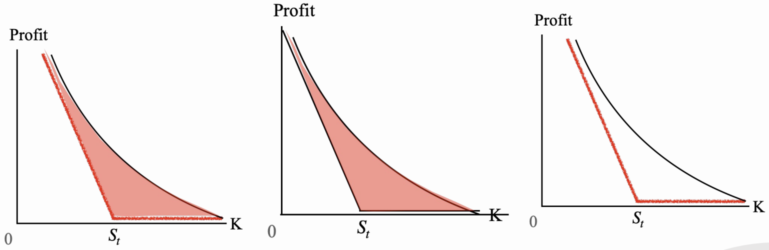
$$\frac{\partial OV}{\partial \tau} \geq 0$$

$$\lim_{k \to 0} OV = S_t \qquad \lim_{k \to \infty} OV = 0$$

$$\lim_{\tau \to 0} OV = \max(0, S_t - K)$$

max(0, S-K)

Option value = time value + intrinsic value



Control Variates Approach

Option value based no-arbitrage constraints

$$\frac{\partial OV}{\partial K} \leq 0$$

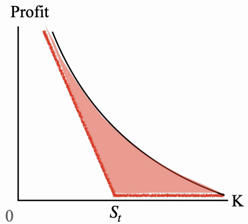
$$\frac{\partial^2 OV}{\partial^2 K} \geq 0$$

$$\frac{\partial OV}{\partial \tau} \geq 0$$

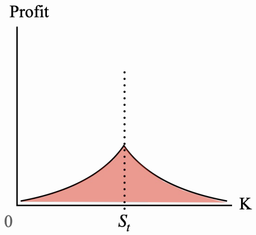
$$\lim_{k \to 0} OV = S_t \qquad \lim_{k \to \infty} OV = 0$$

$$\lim_{\tau \to 0} OV = \max(0, S_t - K)$$

option value



time value



Time value based no-arbitrage constraints

ITM

$$\text{squareplus}(\tilde{b} + me^{\tilde{w}})\text{sigmoid}(\bar{b} + \tau e^{\bar{w}})e^{\hat{w}}$$

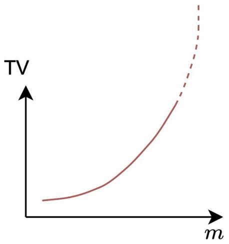
$$\frac{\partial TV}{\partial K} \geq 0$$

$$\frac{\partial^2 TV}{\partial^2 K} \geq 0$$

$$\frac{\partial TV}{\partial \tau} \geq 0$$

$$\lim_{K \to 0} TV = 0$$

$$\lim_{\tau \to 0} TV = 0$$



OTM

$$\text{softplus}(\tilde{b} - me^{\tilde{w}})\text{sigmoid}(\bar{b} + \tau e^{\bar{w}})e^{\hat{w}}$$

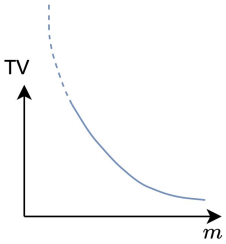
$$\frac{\partial TV}{\partial K} \leq 0$$

$$\frac{\partial^2 TV}{\partial^2 K} \geq 0$$

$$\frac{\partial TV}{\partial \tau} \geq 0$$

$$\lim_{K \to \infty} TV = 0$$

$$\lim_{\tau \to 0} TV = 0$$



ITM OTM

Control Variates Approach

Option value based no-arbitrage constraints

$$\frac{\partial OV}{\partial K} \leq 0$$

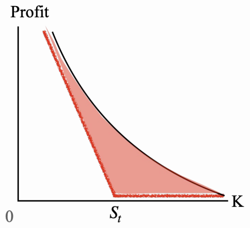
$$\frac{\partial^2 OV}{\partial^2 K} \geq 0$$

$$\frac{\partial OV}{\partial \tau} \geq 0$$

$$\lim_{k \to 0} OV = S_t \qquad \lim_{k \to \infty} OV = 0$$

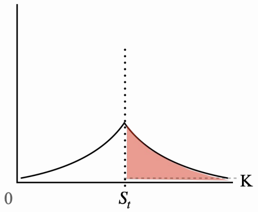
$$\lim_{\tau \to 0} OV = \max(0, S_t - K)$$

option value



Time value based no-arbitrage constraints

time value



Since call OTM option value equals to time value, here we follow Yang et al. 2017.

Yang, Y., Zheng, Y., & Hospedales, T. (2017, February). Gated neural networks for option pricing: Rationality by design. In Proceedings of the AAAI conference on artificial intelligence (Vol. 31, No. 1).

$$\text{softplus}(\tilde{b} - me^{\tilde{w}})\text{sigmoid}(\bar{b} + \tau e^{\bar{w}})e^{\hat{w}}$$

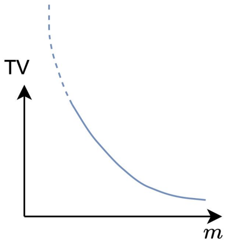
$$\frac{\partial TV}{\partial K} \leq 0$$

$$\frac{\partial^2 TV}{\partial^2 K} \geq 0$$

$$\frac{\partial TV}{\partial \tau} \geq 0$$

$$\lim_{K \to \infty} TV = 0$$

$$\lim_{\tau \to 0} TV = 0$$



OTM

Control Variates Approach

Option value based no-arbitrage constraints

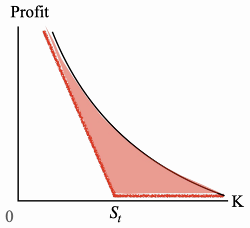
$$\frac{\partial OV}{\partial K} \leq 0$$

$$\frac{\partial^2 OV}{\partial^2 K} \geq 0 \qquad \frac{\partial OV}{\partial \tau} \geq 0$$

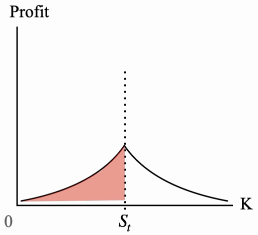
$$\lim_{k \to 0} OV = S_t \qquad \lim_{k \to \infty} OV = 0$$

$$\lim_{\tau \to 0} OV = \max(0, S_t - K)$$

option value



time value



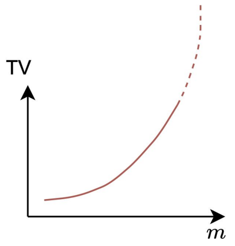
Time value based no-arbitrage constraints

$$\text{squareplus}(\tilde{b} + me^{\tilde{w}})\text{sigmoid}(\bar{b} + \tau e^{\bar{w}})e^{\hat{w}}$$

$$\frac{\partial TV}{\partial K} \geq 0$$

$$\frac{\partial^2 TV}{\partial^2 K} \geq 0 \qquad \frac{\partial TV}{\partial \tau} \geq 0$$

$$\lim_{K \to 0} TV = 0 \qquad \lim_{\tau \to 0} TV = 0$$



$$\text{TV}(K, S_t, \tau) \equiv S_t \times y(m, \tau) \equiv S_t \times \sum_{i=1}^{I} y_i(m, \tau) w_i(m, \tau)$$

$$y_i(m, \tau) = \mathbf{1}_{1 \times J} \sigma_1(\tilde{\boldsymbol{b}} + me^{\tilde{\boldsymbol{w}}}) \sigma_2(\bar{\boldsymbol{b}} + \tau e^{\bar{\boldsymbol{w}}}) e^{\hat{\boldsymbol{w}}}$$

$$w_i(m, \tau) = \text{softmax}\left(\left([m, \tau]\dot{\boldsymbol{W}}^T + \dot{\boldsymbol{b}}^T\right)\ddot{\boldsymbol{W}} + \ddot{\boldsymbol{b}}\right)\mathbf{1}_{I \times 1}$$

$$\sigma_1(x) = \frac{1}{2}\left(x + \sqrt{x^2 + \beta}\right) > 0, \beta > 0$$

$$\sigma_2(x) = \frac{1}{1 + e^{-x}} > 0,$$

ITM

Control Variates Approach

Time value based no-arbitrage conditions

$$\frac{\partial \mathrm{TV}}{\partial K} = \frac{\partial \mathrm{OV}}{\partial K} - \frac{\partial\,(S_t - K)}{\partial K} \geq -1 - (-1) = 0.$$

$$\frac{\partial^2 \mathrm{TV}}{\partial^2 K} = \frac{\partial^2 \mathrm{OV}}{\partial^2 \mathrm{K}} - \frac{\partial^2 (S_t - K)}{\partial^2 K} = f(S_T|S_t, \tau) - 0 \geq 0.$$

$$\frac{\partial \mathrm{TV}}{\partial \tau} = \frac{\partial \mathrm{OV}}{\partial \tau} - \frac{\partial \mathrm{IV}}{\partial \tau} = \frac{\partial \mathrm{OV}}{\partial \tau} - \frac{S_t - K}{\partial \tau} \geq 0.$$

$$\lim_{K\to 0} \mathrm{TV} = \lim_{K\to 0} \mathrm{OV} - \lim_{K\to 0} \mathrm{IV} = S_t - S_t = 0.$$

$$\lim_{\tau\to 0} \mathrm{TV} = \lim_{\tau\to 0} \mathrm{OV} - \lim_{\tau\to 0} \mathrm{IV} = \mathrm{IV} - \mathrm{IV} = 0.$$

Time value based neural network

$$\text{C1)}\ \frac{\partial \mathrm{TV}}{\partial K} = \frac{\partial y_i}{\partial K} = \frac{1}{S} e^{\tilde{\boldsymbol{w}}} \sigma_1'\left(\tilde{\boldsymbol{b}} + \frac{K}{S} e^{\tilde{\boldsymbol{w}}}\right) \sigma_2\left(\bar{\boldsymbol{b}} + \tau e^{\bar{\boldsymbol{w}}}\right) e^{\hat{\boldsymbol{w}}} \geq 0$$

$$\text{C2)}\ \frac{\partial^2 \mathrm{TV}}{\partial^2 K} = \frac{\partial^2 y_i}{\partial^2 K} = \frac{1}{S^2} e^{2\tilde{\boldsymbol{w}}} \sigma_1''\left(\tilde{\boldsymbol{b}} + \frac{K}{S} e^{\tilde{\boldsymbol{w}}}\right) \sigma_2\left(\bar{\boldsymbol{b}} + \tau e^{\bar{\boldsymbol{w}}}\right) e^{\hat{\boldsymbol{w}}} \geq 0$$

$$\text{C3)}\ \frac{\partial \mathrm{TV}}{\partial \tau} = \frac{\partial y_i}{\partial \tau} = e^{\bar{\boldsymbol{w}}} \sigma_1\left(\tilde{\boldsymbol{b}} + \frac{K}{S} e^{\tilde{\boldsymbol{w}}}\right) \sigma_2\left(\bar{\boldsymbol{b}} + \tau e^{\bar{\boldsymbol{w}}}\right) e^{\hat{\boldsymbol{w}}} \geq 0,$$

$$\sigma_1'(x) = \frac{1}{2}\left(1 + \frac{x}{\sqrt{x^2 + \beta}}\right) > 0, \text{ since } \left|\frac{x}{\sqrt{x^2+\beta}}\right| < 1$$

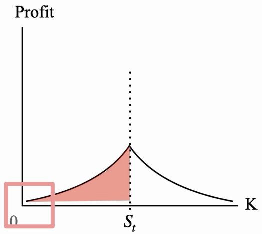
$$\sigma_1''(x) = \frac{1}{2}\left(\frac{\beta}{\sqrt{x^2+\beta}^3}\right) > 0,$$

$$\sigma_2'(x) = \sigma_2(x)(1 - \sigma_2(x)) > 0$$

Control Variates Approach

Time value based approach removes unnecessary synthesized data

time value



Option value based no-arbitrage constraints

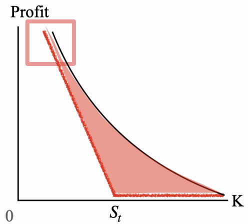
$$\frac{\partial OV}{\partial K} \leq 0$$

$$\frac{\partial^2 OV}{\partial^2 K} \geq 0 \qquad \frac{\partial OV}{\partial \tau} \geq 0$$

$$\lim_{k \to 0} OV = S_t \qquad \lim_{k \to \infty} OV = 0$$

$$\lim_{\tau \to 0} OV = \max(0, S_t - K)$$

option value



Time value based no-arbitrage constraints

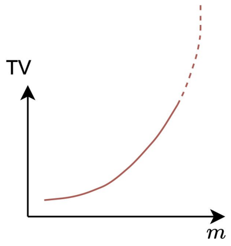
$$\text{squareplus}(\tilde{b} + me^{\tilde{w}})\text{sigmoid}(\bar{b} + \tau e^{\bar{w}})e^{\hat{w}}$$

$$\frac{\partial TV}{\partial K} \geq 0$$

$$\frac{\partial^2 TV}{\partial^2 K} \geq 0 \qquad \frac{\partial TV}{\partial \tau} \geq 0$$

$$\lim_{K \to 0} TV = 0$$

$$\lim_{\tau \to 0} TV = 0$$



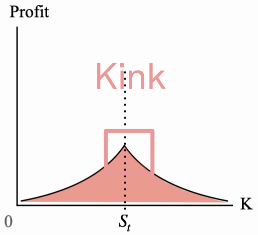
ITM

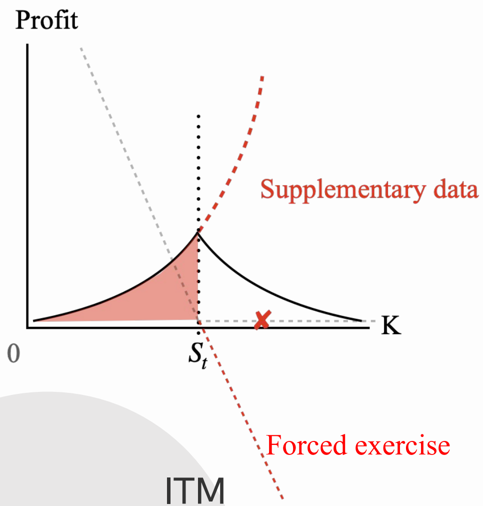
τ	S_t	K	OV
Constraint 6			
5	100	0	100
10	100	0	100
..	..	..	..
5	110	0	110
10	110	0	110
..	..	..	..
Constraint 5			
0	100	10	90
0	100	20	80
..	..	..	..
0	100	90	10
0	110	10	100
0	110	20	90
..	..	..	..

Control Variates Approach

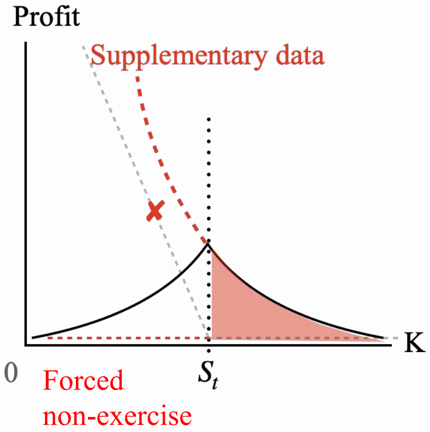
Handle kink of time value surface by supplementary data

time value





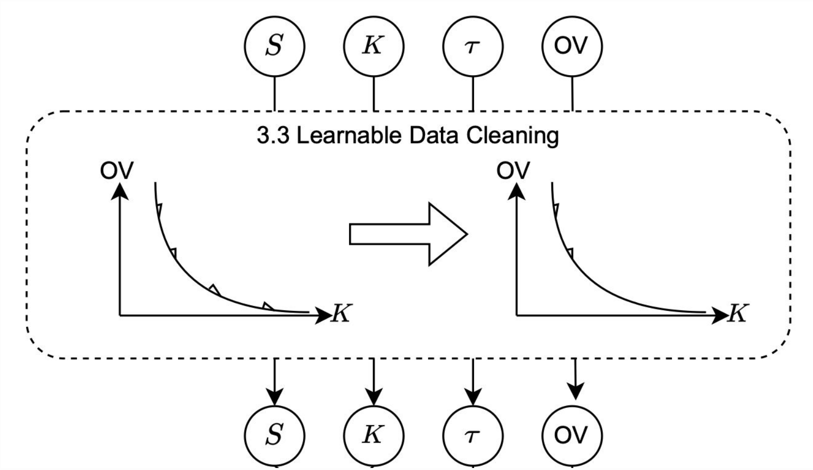
ITM



OTM

Data Cleaning

Trade-off between data quality and quantity



03

Experiments

Datasets
Training Setting
Results on S&P 500 Option
Results on TAIEX Option

Datasets & Environment

TAIEX Index Option

2014 - 2021
Option: TXO from Taiwan Future Exchange
Underlying: TXF from Taiwan Future Exchange
Risk-Free Rate: Bank of Taiwan

S&P 500 Index Option

2010 - 2016
Option: S&P 500 from OptionMetric
Underlying: S&P 500 from Yahoo Finance
Dividend Yield Rate: S&P 500 from OptionMetric
Risk-Free Rate: U.S. Department of the Treasury

Environment

OS: Linux
Memory: 128 GB
CPU: 2 * AMD EPYC 7702P 64-Core Processors

Evaluation Metrics

MSE, MAPE

Results on S&P 500 Option

Model	Error	Training Set Overall	Training Set ITM	Training Set OTM	Testing Set Overall	Testing Set ITM	Testing Set OTM
BS	MSE	75.31	97.21	31.89	78.01	99.86	35.08
[5]	MAPE(%)	48.66	3.60	144.28	50.33	3.67	149.48
Variance Gamma	MSE	17.66	20.03	11.76	19.94	22.19	14.41
[22]	MAPE(%)	15.04	2.71	40.65	16.43	2.84	44.55
Kou's Jump	MSE	14.47	16.34	10.02	16.7	18.49	12.61
[19]	MAPE(%)	12.23	2.18	32.39	14.00	2.29	37.70
Multi without syn.	MSE	113.29	135.15	70.52	117.15	139.07	72.74
[29]	MAPE(%)	10.57	2.74	26.92	13.95	3.05	36.57
Multi with syn.	MSE	8.52	11.27	2.95	11.45	13.88	6.48
[29]	MAPE(%)	5.93	1.10	15.65	9.66	1.44	26.29
CV with supplement	MSE	3.78	3.26	4.93	5.14	4.71	6.18
but without redundant syn.	MAPE(%)	7.66	0.89	21.33	9.33	1.19	25.80
CV with supplement	MSE	5.12	5.67	4.03	7.5	8.68	5.68
and redundant syn.	MAPE(%)	7.25	1.50	18.85	9.81	1.91	25.58
CV with redundant syn.	MSE	9.26	10.97	6.19	117.5	179.34	6.64
but without supplement	MAPE(%)	7.36	2.17	17.81	9.72	2.53	24.26
CV uses one NN	MSE	92.35	115.26	48.25	99.65	121.9	58.55
	MAPE(%)	23.16	5.69	60.05	27.51	5.86	72.80

Table 3: Comparing Pricing Accuracy and Ablation Studies for S&P Options.

Results on S&P 500 Option

Model	Error	Training Set Overall	Training Set ITM	Training Set OTM	Testing Set Overall	Testing Set ITM	Testing Set OTM
BS [5]	MSE	75.31	97.21	31.89	78.01	99.86	35.08
	MAPE(%)	48.66	3.60	144.28	50.33	3.67	149.48
Variance Gamma [22]	MSE	17.66	20.03	11.76	19.94	22.19	14.41
	MAPE(%)	15.04	2.71	40.65	16.43	2.84	44.55
Kou's Jump [19]	MSE	14.47	16.34	10.02	16.7	18.49	12.61
	MAPE(%)	12.23	2.18	32.39	14.00	2.29	37.70
Multi with syn. [29]	MSE	8.52	11.27	2.95	11.45	13.88	6.48
	MAPE(%)	5.93	1.10	15.65	9.66	1.44	26.29
CV with supplement but without redundant syn.	MSE	3.78	3.26	4.93	5.14	4.71	6.18
	MAPE(%)	7.66	0.89	21.33	9.33	1.19	25.80

Table 3: Comparing Pricing Accuracy and Ablation Studies for S&P Options.

Yang, Y., Zheng, Y., & Hospedales, T. (2017, February). Gated neural networks for option pricing: Rationality by design. In Proceedings of the AAAI conference on artificial intelligence (Vol. 31, No. 1).

Results on S&P 500 Option

Model	Error	Training Set Overall	Training Set ITM	Training Set OTM	Testing Set Overall	Testing Set ITM	Testing Set OTM
BS [5]	MSE	75.31	97.21	31.89	78.01	99.86	35.08
	MAPE(%)	48.66	3.60	144.28	50.33	3.67	149.48
Variance Gamma [22]	MSE	17.66	20.03	11.76	19.94	22.19	14.41
	MAPE(%)	15.04	2.71	40.65	16.43	2.84	44.55
Kou's Jump [19]	MSE	14.47	16.34	10.02	16.7	18.49	12.61
	MAPE(%)	12.23	2.18	32.39	14.00	2.29	37.70
Multi with syn. [29]	MSE	8.52	11.27	2.95	11.45	13.88	6.48
	MAPE(%)	5.93	1.10	15.65	9.66	1.44	26.29
CV with supplement but without redundant syn.	MSE	3.78	3.26	4.93	5.14	4.71	6.18
	MAPE(%)	7.66	0.89	21.33	9.33	1.19	25.80

Table 3: Comparing Pricing Accuracy and Ablation Studies for S&P Options.

Results on S&P 500 Option

Model	Error	Training Set Overall	Training Set ITM	Training Set OTM	Testing Set Overall	Testing Set ITM	Testing Set OTM
BS [5]	MSE	75.31	97.21	31.89	78.01	99.86	35.08
	MAPE(%)	48.66	3.60	144.28	50.33	3.67	149.48
Variance Gamma [22]	MSE	17.66	20.03	11.76	19.94	22.19	14.41
	MAPE(%)	15.04	2.71	40.65	16.43	2.84	44.55
Kou's Jump [19]	MSE	14.47	16.34	10.02	16.7	18.49	12.61
	MAPE(%)	12.23	2.18	32.39	14.00	2.29	37.70
Multi with syn. [29]	MSE	8.52	11.27	2.95	11.45	13.88	6.48
	MAPE(%)	5.93	1.10	15.65	9.66	1.44	26.29
CV with supplement but without redundant syn.	MSE	3.78	3.26	4.93	5.14	4.71	6.18
	MAPE(%)	7.66	0.89	21.33	9.33	1.19	25.80

Table 3: Comparing Pricing Accuracy and Ablation Studies for S&P Options.

Results on S&P 500 Option

Model	Error	Training Set			Testing Set		
		Overall	ITM	OTM	Overall	ITM	OTM
Multi without syn. [29]	MSE	113.29	135.15	70.52	117.15	139.07	72.74
	MAPE(%)	10.57	2.74	26.92	13.95	3.05	36.57
Multi with syn. [29]	MSE	8.52	11.27	2.95	11.45	13.88	6.48
	MAPE(%)	5.93	1.10	15.65	9.66	1.44	26.29
CV with supplement but without redundant syn.	MSE	3.78	3.26	4.93	5.14	4.71	6.18
	MAPE(%)	7.66	0.89	21.33	9.33	1.19	25.80
CV with supplement and redundant syn.	MSE	5.12	5.67	4.03	7.5	8.68	5.68
	MAPE(%)	7.25	1.50	18.85	9.81	1.91	25.58

Without synthetic data.

With synthetic data.

degrade

improve

Table 3: Comparing Pricing Accuracy and Ablation Studies for S&P Options.

Results on S&P 500 Option

Ablation Study

Model	Error	Training Set Overall	Training Set ITM	Training Set OTM	Testing Set Overall	Testing Set ITM	Testing Set OTM
CV with supplement but without redundant syn.	MSE	3.78	3.26	4.93	5.14	4.71	6.18
	MAPE(%)	7.66	0.89	21.33	9.33	1.19	25.80
CV with supplement and redundant syn.	MSE	5.12	5.67	4.03	7.5	8.68	5.68
	MAPE(%)	7.25	1.50	18.85	9.81	1.91	25.58
CV with redundant syn. but without supplement	MSE	9.26	10.97	6.19	117.5	179.34	6.64
	MAPE(%)	7.36	2.17	17.81	9.72	2.53	24.26
CV uses one NN	MSE	92.35	115.26	48.25	99.65	121.9	58.55
	MAPE(%)	23.16	5.69	60.05	27.51	5.86	72.80

Table 3: Comparing Pricing Accuracy and Ablation Studies for S&P Options.

Results on TAIEX Option

Model	Error	Before Data Cleaning Overall	Before Data Cleaning ITM	Before Data Cleaning OTM	After Data cleaning Overall	After Data cleaning ITM	After Data cleaning OTM
BS (Black and Scholes [5])	MSE	979.47	1263.32	752.24	994.72	1265.41	767.88
	MAPE(%)	35.52	2.62	82.75	35.09	2.66	81.51
Variance Gamma (Madan et al. [22])	MSE	357.84	416.53	368.51	365.15	418.77	378.33
	MAPE(%)	19.52	1.75	44.44	18.96	1.78	42.98
Kou's Jump (Kou [19])	MSE	319.4	378.26	314.12	323.12	367.51	321.88
	MAPE(%)	17.57	1.63	38.88	17.13	1.65	37.78
Multi (Yang et al. [29])	MSE	8820.45	12401.15	3946.04	7456.88	13713.16	1349.06
	MAPE(%)	65.63	7.69	150.57	37.12	4.1	85.52
Deep Smoothing (Ackerer et al. [1])	MSE	3710.67	5511.63	1191.23	3490.72	5359.1	885.74
	MAPE(%)	51.09	9.7	111.15	49	9.81	105.54
Hybrid gated NN (Cao et al. [7])	MSE	251978.14	423181.76	17719.31	21591.17	21644.39	21580.21
	MAPE(%)	836.46	29.25	1859.18	687.53	10.76	1559.75
Conv-LSTM (Ge et al. [13])	MSE	5205.84	6044	3025	5203	6659	3171
	MAPE(%)	1271.44	477	11641	974.71	397	11600
CV	MSE	537.51	611.56	540.26	301.31	421.35	225.21
	MAPE(%)	13.19	1.48	28.95	11.72	1.27	25.72

Table 4: Comparing Pricing Accuracy for TAIEX Options Before/After Data Cleaning on the Testing Set

04

Conclusions

Conclusions

- Reduces prediction errors by decomposing option values using control variate techniques
- Incorporate no-arbitrage constraints for time value surface into neural network
- Remove those unnecessary data synthesis from past approaches.
- Propose a data cleaning trick to price the option in the illiquid market more accurately

Thanks!

Do you have any questions?

music87.c@nycu.edu.tw