

POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

A Neural Network Approach to American Option: from Implied Volatility and Dividend Recovery to Non-Gaussian Models

Tesi di Laurea Magistrale in
Mathematical Engineering, Quantitative Finance

Gaia Imperatore, 260554

Abstract: This thesis studies the recovery of implied volatility, dividend yields, and model parameters from American option prices. The analysis combines numerical pricing with neural-network calibration. For the synthetic experiments, American option prices are generated with the Fourier–cosine (COS) method and used to train neural networks that approximate the pricing function. Differential Evolution then searches for the parameters that best reproduce the target prices. The recovered parameters are inserted again into the COS pricer to check the calibration results. In the Black–Scholes case, the Cox–Ross–Rubinstein tree provides an additional numerical benchmark.

Starting from the American Black–Scholes setting considered in the literature, the analysis moves from individual matched put–call pairs to groups of options with several strikes and maturities. Dividends are represented by a curve that can change with maturity while remaining common across strikes. The same procedure is then studied under Black–Scholes, Normal Inverse Gaussian (NIG), R–GBA, and ATS dynamics.

The synthetic experiments show that option prices, implied-volatility smiles, and dividend curves can be recovered accurately when the calibration imposes a clear structure on the unknown quantities. They also show that a small pricing error does not always imply an equally accurate recovery of every model parameter. In particular, when a separate dividend is estimated for every strike and maturity together with all the other model parameters, different parameter combinations can produce similar option prices.

The final application uses Apple Inc. option MID prices (Bloomberg ticker: AAPL). Calibration is performed separately for each expiry using matched put–call observations, and the results are evaluated on strikes that were not used for calibration. NIG gives a lower pricing error for calls than the Black–Scholes benchmark, while the put errors remain similar. In both models, the dividend recovered from option prices differs from the Bloomberg dividend field. Additional comparisons show that using the recovered dividend improves the pricing of the excluded strikes even when the volatility input or the NIG parameters are kept fixed. This difference is interpreted as a possible model-implied adjustment to the effective carry used in option pricing, rather than as evidence of a negative corporate dividend.

Advisor:
Prof. Michele Azzone

Academic year:
2025–2026

Key-words: American options; neural calibration; implied volatility; dividend recovery; COS method; non-Gaussian models

Contents

1 Introduction **2**

2 Models and methodology **3**
2.1 Market setting and models 4
2.2 American-option pricing and numerical checks 6
2.3 Neural calibration and inverse problem 9

3 Synthetic results **13**
3.1 Black–Scholes 13
3.2 Normal Inverse Gaussian (NIG) model 19
3.3 R–GBA and ATS extensions 23

4 Application to market data: AAPL options **26**

5 Conclusions **37**

List of symbols and abbreviations **39**

1. Introduction

American options may be exercised at any admissible time before maturity. If the holder exercises at time $\tau \leq T$, an American put pays $(K - S_\tau)^+$, while an American call pays $(S_\tau - K)^+$. The holder chooses the exercise time that maximizes the discounted expected payoff. This additional stopping decision makes American options more difficult to price and calibrate than European options, whose payoff is evaluated only at maturity. The value of an equity American option depends on volatility, interest rates, dividends, strike, maturity, and the exercise decision. The inverse problem is therefore delicate: different combinations of volatility, dividends, and model-specific parameters can produce similar option prices. For this reason, a small pricing error does not automatically imply that every calibrated parameter has been recovered accurately.

This thesis studies neural networks as fast approximations of American option prices and uses them to recover implied information from option prices. The main quantities considered in the experiments are implied volatility, dividend yields, and the parameters of more general pricing models. The aim is not only to obtain accurate prices, but also to understand which quantities can be recovered reliably when several options are calibrated jointly.

Black and Scholes introduced the classical lognormal model for option pricing, while Merton included continuous dividends in the same framework (Black and Scholes, 1973; Merton, 1973). The Cox–Ross–Rubinstein (CRR) tree provides a discrete-time method for American options: at every node it compares the value of immediate exercise with the value of continuing the contract (Cox et al., 1979). In this thesis, CRR is mainly used as an independent benchmark for the American COS implementation.

The COS method represents the transition density through a Fourier–cosine expansion based on the characteristic function of the return process (Fang and Oosterlee, 2008). Fang and Oosterlee (2009) extend this method to American options by working backwards over a grid of possible exercise dates. COS is useful here for two reasons. First, it can efficiently generate the large number of option prices needed for the experiments. Since these prices are generated by a pricing model rather than observed in the market, we refer to them as synthetic option prices. Second, COS can also be used when the return model is described through its characteristic function.

Neural networks have been used to approximate option-pricing functions and to reduce the cost of repeated pricing calculations (Hutchinson et al., 1994; Liu et al., 2019b; Horvath et al., 2021). The Calibration Neural Network framework of Liu et al. (2019a) combines a neural network that approximates the pricing function with a global optimization method.

Our starting point is the American Black–Scholes procedure proposed by Liu et al. (2021). In their work, American option prices are first generated with the COS method and used to train neural networks. When the dividend yield is known, the network is used to recover the implied volatility from an individual American put price. When both volatility and dividend yield are unknown, the network approximates the prices of a matched put–call pair. When this calibration is repeated independently at different strikes, it does not impose that the recovered dividend yield be the same at every strike.

Differential Evolution (DE) is the optimization method used to recover the unknown parameters (Storn and Price, 1997). It starts from a population, that is a set of possible parameter combinations. For each combination,

the neural network produces option prices and the difference from the target prices is calculated. At each iteration, DE combines and modifies the current parameter values to produce new candidates. A new candidate is retained only when it gives a smaller pricing error.
DE does not require derivatives of the pricing function. This is useful for American options because the decision between continuing and exercising is evaluated at every step of the backward calculation, making price derivatives with respect to the model parameters more difficult to compute. Parameter combinations that are outside the allowed ranges or do not satisfy the conditions of the model are assigned a large error and are not selected.
We use these settings as a benchmark. We then extend the calibration from individual puts or one matched put–call pair to groups of strikes and maturities.
We consider constant dividends and dividend curves that vary with maturity while remaining common across strikes. As a comparison, we also run calibrations in which a separate dividend value is allowed for each strike and maturity. These cases are not interpreted as different company dividend policies. They are controls used to show what happens when the calibration is allowed to fit each option with a separate dividend value. We then study how these choices affect the recovery of volatility, dividends, and non-Gaussian model parameters.
The Normal Inverse Gaussian (NIG) model introduces asymmetry and heavier tails relative to Gaussian returns while retaining a characteristic function suitable for COS pricing (Barndorff-Nielsen, 1997). NIG is used first to generate non-flat volatility smiles in a Black–Scholes experiment and then directly as the pricing model inside the American COS recursion. The R–GBA and ATS models provide two further additive-process specifications (Azzone and Torricelli, 2025; Azzone and Baviera, 2022). They are included to test whether the same pricing and calibration procedure can also be used with more flexible return dynamics.
Previous studies show that American option prices can contain information about volatility, dividends, and carry-related quantities. Zhang and Teo (2015) study the recovery of a local-volatility surface from American option prices, while Kohler et al. (2010) use neural methods for American option valuation. Related work also studies funding-related information extracted from option prices (Azzone and Baviera, 2021). These studies motivate the final application to market data.

The contributions of this thesis are threefold. The first contribution is a replication of the American Black–Scholes setting of Liu et al. (2021). Before extending the analysis to several strikes, maturities, and models, we also compare the COS prices with the CRR tree. This provides an independent check of the numerical prices used to construct the synthetic datasets.
The second contribution is a systematic synthetic study of the inverse problem. We move from individual put options and single matched put–call pairs to smiles and complete strike–maturity surfaces. The experiments introduce structured volatility smiles, forward dividend curves shared across strikes, and control experiments in which a separate dividend is estimated for each strike and maturity. The same procedure is applied under Black–Scholes, Normal Inverse Gaussian (NIG), R–GBA, and ATS dynamics.
In a synthetic experiment, we first choose a set of model parameters and use them to generate option prices. The calibration procedure receives the option prices and contract inputs, but not the parameters used to generate them. After calibration, we can therefore check both whether the recovered parameters reproduce the synthetic prices and whether they are close to the original parameter values. This distinguishes a good price fit from an accurate recovery of the underlying parameters.
Third, the procedure is applied to real AAPL option data. For each maturity, matched call and put prices at several strikes are calibrated jointly, while other strikes are retained for testing. The Black–Scholes and NIG results are compared with Bloomberg MID prices and with Bloomberg volatility and dividend field. The final analysis studies the difference between the Bloomberg dividend and the dividend implied by option prices and examines whether it may help explain the remaining pricing differences.
The rest of the thesis is organized as follows. Section 2 introduces the market setting, pricing models, dividend representation, and American COS and CRR methods. Section 2.2 describes the datasets, neural pricing approximation, calibration procedure, and numerical validation. Section 3 presents the synthetic experiments under Black–Scholes, NIG, R–GBA, and ATS dynamics. Section 4 applies the framework to AAPL market data, while section 5 summarizes the main results and the limitations of parameter recovery.

2. Models and methodology

This chapter introduces the common setting used in the experiments. It first defines the market variables and the pricing models. It then describes the COS and CRR methods used for American options. The construction of the synthetic datasets and the neural calibration procedure are presented in the next section.
Before using COS to generate the synthetic option prices, we check its accuracy in the Black–Scholes case. We consider 496 American options and compute the implied volatility associated with the COS price and with the CRR price. The mean absolute difference between the two implied volatilities is 4.5×10^{-4}, while the RMSE

is 1.79×10^{-3} (figure 2). The MAE corresponds to 0.045 percentage points of volatility. In practical terms, this is a very small difference and is more than acceptable for implied-volatility calculations in option markets. We therefore use COS to generate the datasets, while CRR remains an independent numerical check in the Black–Scholes experiments.

2.1. Market setting and models

Let S_t be the stock price, $r(t)$ the continuously compounded financing rate, and $q(t)$ the continuous dividend yield. For a payoff $g(S_T)$, the American value at time zero is

$$V_0 = \sup_{\tau \in \mathcal{T}_{[0,T]}} \mathbb{E}^{\mathbb{Q}} \left[\exp\left(- \int_0^\tau r(u)\, \mathrm{d}u \right) g(S_\tau) \right], \tag{1}$$

where $\mathcal{T}_{[0,T]}$ is the set of admissible stopping times. Put and call payoffs are $(K-S)^+$ and $(S-K)^+$, respectively. The initial Black–Scholes dataset and the COS–CRR validation use individual options: strike and maturity vary from one observation to another, without a fixed set of strikes. A fixed strike grid is used only in the later multi-strike simulations.
In these simulations, a *synthetic surface* is a collection of put and call prices generated by the same model at several strikes and maturities. The prices are generated first and are then calibrated jointly. For each surface, the spot is normalized to $S_0 = 1$, so that strikes can be expressed as the moneyness ratio K/S_0. At every selected maturity, put and call prices are generated at the seven strikes

$$\frac{K}{S_0} \in \{0.75, 0.85, 0.95, 1.00, 1.05, 1.15, 1.25\}. \tag{2}$$

The grid includes strikes below, close to, and above the spot price. It is used to test whether the same model parameters and dividend specification can explain the prices of several options with different strikes. Normalizing the spot makes this grid independent of the monetary unit used for the stock price.

The pricing models in this thesis represent dividends through a continuous dividend yield. This means that dividends are treated as if they were paid continuously through time, rather than as separate cash payments on specific dates. This is a simplifying assumption: actual equities, including AAPL in the market application, pay discrete cash dividends. The continuous-yield representation is nevertheless convenient because it can be used directly in the pricing models and in the American-option recursion.
The dividend yield may be constant or may change over time. For maturities

$$0 = T_0 < T_1 < \cdots < T_m,$$

we represent a time-dependent dividend curve through one forward dividend yield for each interval:

$$q(t) = q_j, \qquad t \in (T_{j-1}, T_j]. \tag{3}$$

Thus, q_j is the dividend yield used between T_{j-1} and T_j.
American options are priced by backward induction. The calculation starts at maturity, where the option value equals its payoff. It then moves backwards over the possible exercise dates. At each date, the value of continuing the option is calculated from the value at the following date and is compared with the payoff from immediate exercise. The larger of the two values is retained. When the calculation moves through the interval $(T_{j-1}, T_j]$, it uses the corresponding dividend yield q_j. Consequently, an option maturing at T_j depends on $q_1, \ldots, q_j$, but not on dividend values for later intervals.
In the neural-network input, the same curve is sometimes written as

$$(q_1, \Delta q_2, \ldots, \Delta q_m), \qquad \Delta q_j = q_j - q_{j-1}.$$

This is only a different way of recording the same forward dividend curve: the first value is q_1, while each later input describes the change from the previous interval.
Under risk-neutral pricing, the expected growth rate of the stock is determined by $r-q$, where r is the financing rate and q is the continuous dividend yield. Holding the other inputs fixed, a higher dividend yield lowers the expected future stock price. It therefore tends to lower call prices and increase put prices. Higher volatility instead makes future stock prices more uncertain. Since both put and call payoffs benefit from large price movements, their values usually increase when volatility increases.
For American options, dividends also affect the decision of whether to exercise early. Exercising a call gives the holder the stock earlier and can therefore become more attractive when the dividend yield is higher. For puts,

the effect of early exercise also depends on the interest rate and on whether the strike is above or below the current stock price. The dividend curve must therefore be used consistently both in the stock dynamics and in every step of the backward-induction calculation.

These effects are important when parameters are recovered from prices. For very short maturities, options with little time value, or options that react only weakly to dividends, a change in the dividend can be partly offset by a change in volatility. In these cases, different parameter values can produce similar prices. Using calls and puts at several strikes provides more information, but it does not guarantee that every parameter can be recovered uniquely when all of them are left free.

In the main synthetic experiments, the preferred specification uses one dividend curve shared across strikes and allowed to change with maturity. For comparison, some control experiments estimate a separate value $q(K,T)$ for each strike and maturity. These controls show how the results change when the strike-common restriction is removed; they are not interpreted as different corporate dividends for different strikes. In the AAPL application, the dividend recovered from option prices is compared with the Bloomberg dividend field. The interpretation of any difference between the two is discussed later in the market-data chapter.

We now describe the four return models used in the experiments: Black–Scholes, NIG, R–GBA, and ATS. They differ in the assumptions made about stock returns, while the dividend representation above is kept unchanged. Under Black–Scholes (Black and Scholes, 1973; Merton, 1973),

$$\frac{\mathrm{d}S_t}{S_t} = (r(t) - q(t))\,\mathrm{d}t + \sigma\,\mathrm{d}W_t^{\mathbb{Q}}. \tag{4}$$

In the benchmark setting of Liu et al. (2021), the calibration is performed for one option quote, or for one matched call–put pair with fixed K and T. In the joint calibration of volatility and dividend, the same scalar volatility σ is used for the put and call of the matched pair. The procedure does not estimate one volatility smile shared across several strikes and maturities. Our later experiments extend this setting by calibrating groups of quotes and by introducing structured volatility smiles and surfaces.

In our first controlled tests, one scalar volatility is shared by all quotes within a given test. In the later Black–Scholes surface experiments, the volatility is obtained from a smooth function $\sigma_{\mathrm{imp}}(K,T)$. Different strikes and maturities can therefore have different volatility values, producing a synthetic smile or surface. This construction is used to test the recovery of non-flat volatility patterns. The pricing formula remains the Black–Scholes formula, with the corresponding volatility assigned to each (K,T) pair.

The volatility surface is described by a small number of coefficients controlling its level, skew, curvature, and maturity dependence. This structured representation is easier to identify than one independent volatility parameter for every option quote. European Black–Scholes prices provide an analytical reference for validating the Fourier implementation. We next keep the same general pricing framework but replace the Gaussian return distribution with a non-Gaussian model.

The Normal Inverse Gaussian (NIG) model (Barndorff-Nielsen, 1997) is obtained by setting $\alpha = 1/2$ in the Normal Tempered Stable family. It replaces Gaussian log-returns with a pure-jump process. The characteristic function of an NIG log-return increment over an interval of length Δt is denoted by $\phi_Y(u;\Delta t)$ and is given by

$$\phi_Y(u;\Delta t) = \exp\left\{ \frac{\Delta t}{\kappa}\left[1 - \sqrt{1 + 2\kappa\left(\mathrm{i}u\eta + \frac{1}{2}\sigma_{\mathrm{NIG}}^2 u^2\right)}\right]\right\}. \tag{5}$$

The exponential moment required by the risk-neutral model exists when

$$1 + 2\kappa\eta - \kappa\sigma_{\mathrm{NIG}}^2 > 0. \tag{6}$$

This condition guarantees that the martingale correction used in the pricing model is well defined.

For an interval $(t_{j-1}, t_j]$ of length $\Delta t_j = t_j - t_{j-1}$, with interest rate r_j and dividend yield q_j, the deterministic part of the log-price increment is

$$\omega_j = (r_j - q_j) - \frac{1}{\Delta t_j}\log\phi_Y(-\mathrm{i};\Delta t_j). \tag{7}$$

The second term is the standard martingale correction used in exponential Lévy models (Fang and Oosterlee, 2008). It adjusts the NIG increment so that, under the risk-neutral measure, the expected growth of the stock is consistent with $r_j - q_j$. In the constant-rate and constant-dividend experiments, $r_j = r$ and $q_j = q$ in every interval. When the dividend changes with maturity, the corresponding forward dividend yield q_j is used in each interval of the American backward calculation.

The three NIG structural parameters are $(\sigma_{\mathrm{NIG}}, \eta, \kappa)$. The scale parameter σ_{NIG} mainly affects the overall level of the implied-volatility smile, κ controls the variability of the time change and the importance of the tails, while η affects the asymmetry of the return distribution and therefore the skew of the smile. These effects may be

not completely independent over a limited strike and maturity range. Consequently, similar implied-volatility smiles can sometimes be obtained from different combinations of NIG parameters. The experiments therefore compare both the quality of the reconstructed prices and smile and the recovery of the individual structural parameters.

We also test whether the American COS pricing and calibration procedure can be used with additive-process models beyond Black–Scholes and NIG. Following the reduced-form specification discussed by Azzone and Torricelli (2025), we use

$$\alpha(t) = \alpha_0, \qquad \beta(t) = \beta_0 + \sigma(t), \qquad \sigma(t) = \sigma_0 t^{H_0}, \tag{8}$$

where $H_0 = 0.5$ is fixed. The structural parameters calibrated in the experiments are $(\alpha_0, \beta_0, \sigma_0)$.

According to Azzone and Torricelli (2025, Corollary 4.7), the R–GBA implied-volatility skew at short maturities is negative when $\alpha_0 < \beta_0$ and positive when $\alpha_0 > \beta_0$. The case $\alpha_0 = \beta_0$ is treated separately as the boundary between these two cases. Since equity options usually have a negative skew, we impose $\beta_0 > \alpha_0$ both when generating the synthetic prices and during calibration. The parameter ranges are fixed consistently during data generation and calibration.

The ATS specification follows Azzone and Baviera (2022). We fix $\alpha = 0.5$ and describe the maturity dependence through

$$\kappa(T) = \kappa_0 T^{\beta}, \qquad \eta(T) = \eta_0 T^{\delta}. \tag{9}$$

The calibrated ATS vector is $(\sigma, \kappa_0, \eta_0, \beta, \delta)$. The parameters κ_0 and β control the maturity dependence of the tail behaviour, while η_0 and δ determine how the asymmetry changes with maturity.

In both models, the dividend curve is represented by the forward yields in equation (3). The stochastic model therefore changes the return distribution and the resulting smile, but not the economic interpretation of the recovered dividend. Since some parameters affect the distribution in similar ways over a short maturity grid, the experiments report both price errors and parameter errors. A good price fit alone does not guarantee that every structural parameter has been uniquely identified.

The reference paper studies European vanilla prices. In this thesis, the same R–GBA and ATS additive-process dynamics are inserted into the COS backward-induction recursion for American options. This provides the extension from European pricing to the American pricing and calibration experiments considered here.

The parameters of NIG, R–GBA, and ATS have different meanings and cannot be compared directly across models. We therefore use Black–Scholes-equivalent implied volatility as a common way of displaying the option-price pattern generated by each model.

For a European call price C^{Eur}_{model} produced by one of these models, the equivalent implied volatility σ_{IV} is the Black–Scholes volatility that gives the same European call price:

$$C^{Eur}_{\text{BS}}(S_0, K, T, r, q, \sigma_{\text{IV}}) = C^{Eur}_{\text{model}}(S_0, K, T). \tag{10}$$

Here C^{Eur}_{BS} is the European Black–Scholes call price and C^{Eur}_{model} is the European call price generated by NIG, R–GBA, or ATS. We solve this equation for σ_{IV} and plot the result against strike. The resulting curve is the Black–Scholes-equivalent implied-volatility smile used in the non-Gaussian results.

In the COS–CRR validation, implied volatility is instead obtained by matching an American option price with the same American pricing method. The two quantities are labeled separately in the figures and tables.

2.2. American-option pricing and numerical checks

The COS method approximates the transition density of the log-price through a Fourier–cosine expansion (Fang and Oosterlee, 2008, 2009). The transition density describes the probability distribution of the future log-price conditional on its current value. The first Black–Scholes implementation follows the American-option construction used by Liu et al. (2021).

Let x denote the current log-price state and y the log-price state at the following date. On a finite interval $[a, b]$, the transition density is approximated by

$$f(y \mid x) \approx \sum_{k=0}^{N-1}{}' A_k(x) \cos\left(k\pi \frac{y-a}{b-a}\right), \qquad A_k(x) = \frac{2}{b-a} \Re\left[\phi\left(\frac{k\pi}{b-a}\right) e^{\mathrm{i}k\pi(x-a)/(b-a)}\right], \tag{11}$$

where N is the number of cosine terms, ϕ is the characteristic function of the model increment under the selected model and the prime on the summation sign means that the first term in the summation is weighted by one-half. Once the density coefficients have been calculated, they are combined with the payoff coefficients to obtain the option price. The payoff part of the calculation is the same for all models; Black–Scholes, NIG, R–GBA, and ATS differ through their characteristic functions.

The interval $[a, b]$ must contain the relevant range of log-price values. It is selected from numerical measures of the location and spread of the log-return distribution, together with a safety multiplier L. If the interval is too small, relevant extreme values are excluded. If it is too large, more cosine terms are needed. Before choosing the settings used for the datasets, we vary the number of cosine terms N, the interval multiplier L, and the number of exercise dates.

For a European option, COS combines the payoff coefficients with the transition-density coefficients at maturity. For an American option, exercise is theoretically possible at any time before maturity. In the COS implementation used in this thesis, we approximate this contract by a Bermudan option, which can be exercised only on a finite set of exercise dates. With more exercise dates, the Bermudan price approaches the American price and the calculation becomes more expensive.

The Bermudan price is calculated by backward induction over the exercise dates. At date t_j, $g(S)$ is the payoff obtained by exercising immediately, while $C(t_j, S)$ is the continuation value: the discounted expected value of keeping the option until the next exercise date. The option value is the larger of these two quantities,

$$V(t_j, S) = \max\{g(S), C(t_j, S)\}. \tag{12}$$

Following the approach reported in the appendix of Liu et al. (2021), based on Fang and Oosterlee (2009), we use Richardson extrapolation to improve the Bermudan approximation. Let $V_{\text{Ber}}(M)$ denote the Bermudan COS price obtained with M equally spaced exercise dates. The American price is approximated by

$$V_{\text{Am}} \approx \frac{1}{21}\left[64V_{\text{Ber}}(8M) - 56V_{\text{Ber}}(4M) + 14V_{\text{Ber}}(2M) - V_{\text{Ber}}(M)\right]. \tag{13}$$

The number N controls the number of cosine terms used to approximate the transition density. The number M controls the number of exercise dates in the Bermudan approximation. Richardson extrapolation combines the prices obtained with M, $2M$, $4M$, and $8M$ exercise dates to improve the approximation of the American price. To show how the approximation changes, figure 1 plots the number of exercise dates on the horizontal axis and the option price on the vertical axis. The figure shows the Bermudan COS prices, the Richardson estimate, and the corresponding European price for a put with strike equal to the spot price. As the number of exercise dates increases, the Bermudan COS price approaches the Richardson estimate. The European price remains lower because a European option cannot be exercised before maturity and therefore does not include the early-exercise premium.

After checking the convergence of the COS approximation, we use the Cox–Ross–Rubinstein (CRR) tree (Cox et al., 1979) as an independent benchmark for American Black–Scholes prices. With $u = \exp(\sigma\sqrt{\Delta t})$, $d = 1/u$, and a constant dividend yield over one time step, the risk-neutral probability of an upward movement is

$$p = \frac{\exp((r - q)\Delta t) - d}{u - d}. \tag{14}$$

The option value is then computed backwards through the tree. At each node, we compare the value of continuing with the value obtained by exercising immediately, as in equation (12).

American call prices are obtained using the put–call symmetry described by Liu et al. (2021). In practice, a call with inputs (S_0, K, r, q) is calculated using the same CRR calculation used for a put, after exchanging these quantities as (K, S_0, q, r) and leaving maturity and volatility unchanged. This provides a direct check of the symmetry relation without requiring a separate call implementation.

Before generating the large synthetic datasets, we compare COS and CRR on 496 American Black–Scholes options. For each method, the option price is converted into an implied volatility. Figure 2 reports this comparison in three ways. The left panel compares the COS and CRR implied volatilities directly: almost all points lie on the 45-degree line. The middle panel shows their difference as a function of the volatility used to generate the option, with colour indicating maturity. Most differences are close to zero, while the few larger differences occur for short maturities. The right panel shows that the distribution of the differences is centred close to zero.

The figure therefore shows that COS and CRR give very similar American Black–Scholes prices over the cases considered here. The mean absolute difference in implied volatility is 4.5×10^{-4}, or 0.045 percentage points. This supports the use of COS as the main numerical pricer in the following experiments.

The CRR tree becomes computationally expensive when many options must be priced. With n_{tree} time steps, the backward calculation requires approximately

$$\frac{n_{\text{tree}}(n_{\text{tree}} + 1)}{2}$$

node evaluations for each option. The computational cost therefore grows approximately with the square of the number of time steps. COS is therefore used to generate the large datasets and to reprice the final calibrated parameters. CRR is retained as an independent check on selected Black–Scholes cases.

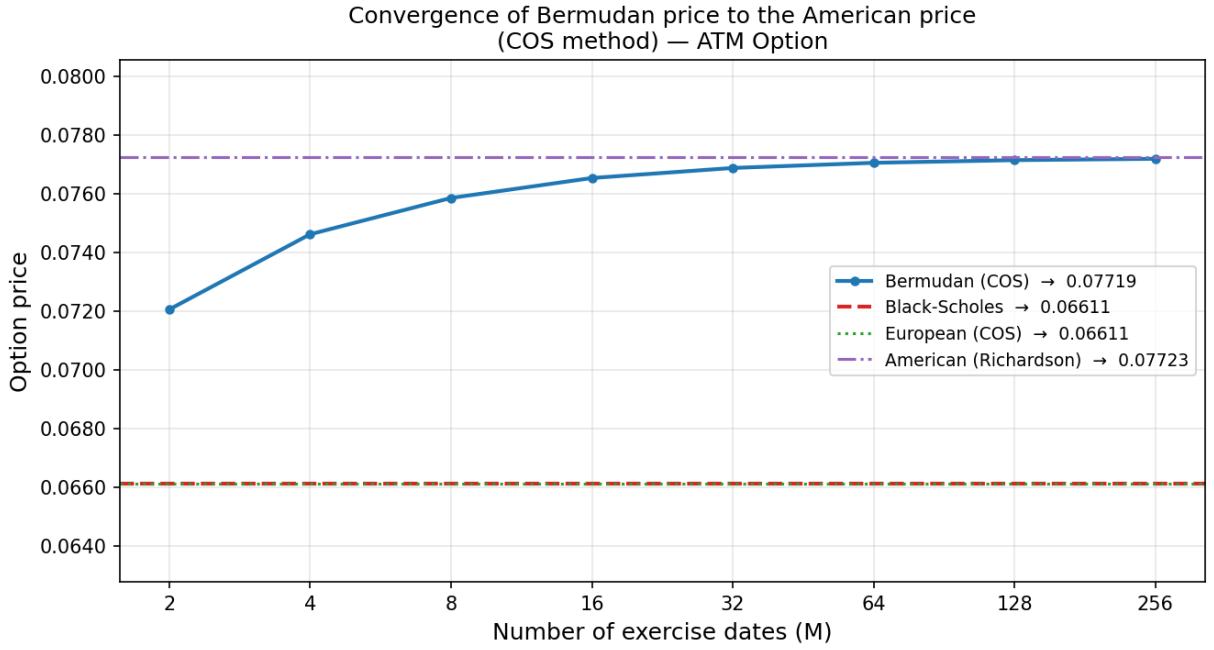


Figure 1: Convergence of the Bermudan COS approximation for a put with $K = S_0$. The horizontal axis reports the number of exercise dates and the vertical axis the option price. As the number of exercise dates increases, the Bermudan price approaches the Richardson estimate of the American price. The European value is lower because early exercise is not allowed.

Before a simulated option price is added to a dataset, we apply numerical and economic checks. The characteristic function must equal one when computed at zero and must return finite numerical values at all the frequencies used by COS. We also check the mathematical conditions required for the pricing formula to be well defined. For NIG, this includes equation (6); analogous conditions are checked for R–GBA and ATS. Parameter combinations that fail these checks are excluded from the dataset.

For constant interest rates and dividend yields, American put and call prices must also satisfy the following bounds:

$$\max\{K - S_0,\ Ke^{-rT} - S_0e^{-qT},\ 0\} \le P_A \le Ke^{\max(-r,0)T}, \tag{15}$$

$$\max\{S_0 - K,\ S_0e^{-qT} - Ke^{-rT},\ 0\} \le C_A \le S_0e^{\max(-q,0)T}. \tag{16}$$

The lower bounds include the value obtained by immediate exercise. Small numerical deviations are accepted only within a fixed numerical tolerance; larger deviations lead to rejection of the simulated price. Prices must also be finite and non-negative, and must satisfy both the put and call bounds in equations (15) and (16).

We check COS convergence by changing three numerical settings: the number of cosine terms N, the interval multiplier L, and the number of Bermudan exercise dates. For each setting, the same test options are priced again and the resulting values are compared. The settings used for the datasets are chosen only after further increases in these quantities do not change the prices at the numerical precision used in the experiments. In the Black–Scholes case, the final COS prices are also compared with analytical European prices and with CRR American prices. For the non-Gaussian models, we check the model conditions, the price bounds, the stability of prices under different COS settings, and the corresponding implied-volatility curves.

These checks are used first to verify the prices that train the neural networks and later to verify the calibrated results. After calibration, the parameters found through the neural network are inserted again into COS. If the network price and the COS price differ substantially at these parameters, the network is not sufficiently accurate in that region. In the Black–Scholes case, agreement between COS and CRR provides an additional check based on a different numerical method. Finally, in synthetic experiments, it is possible that the recovered parameters reproduce the prices accurately but differ from the values used to generate the data. This means that the available option prices do not contain enough information to separate those parameters clearly. This issue is examined in the results chapter.

The COS implementation first reproduces the American Black–Scholes setting of Liu et al. (2021). The same early-exercise procedure is then used with the characteristic functions of the other models. The forward dividend curve in equation (3) is used directly in the American recursion. The convergence test in figure 1 and the COS–CRR comparison in figure 2 support the use of COS prices as targets for the neural-network datasets.

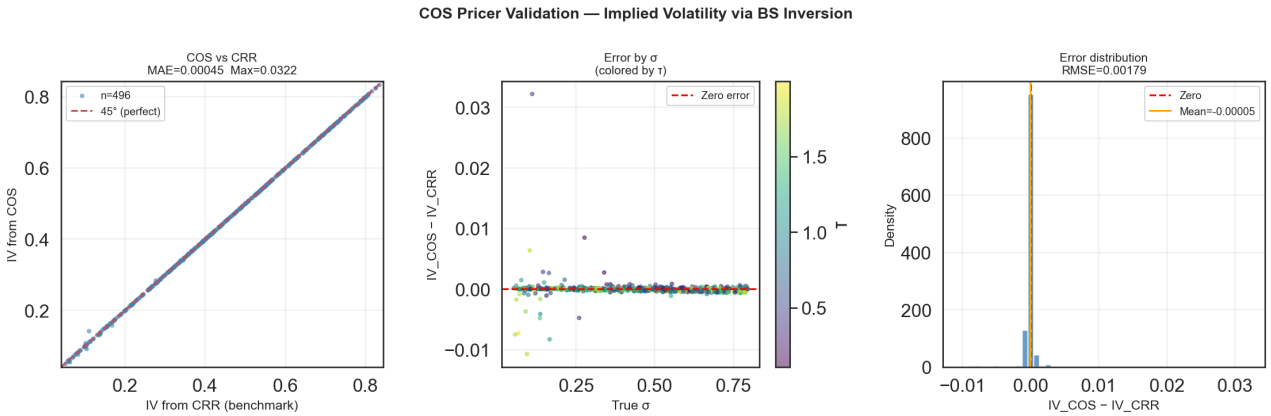


Figure 2: Comparison between COS and CRR implied volatilities for 496 American Black–Scholes options. The left panel compares the two implied volatilities directly, the middle panel reports their difference against the volatility used to generate each option, and the right panel shows the distribution of the differences. The absolute MAE is 4.5×10^{-4} and the RMSE is 1.79×10^{-3} in decimal-volatility units.

Table 1: Synthetic datasets used for the multi-strike and multi-maturity experiments. Each complete surface is assigned to only one of the training, validation, or test partitions. Therefore, prices from the same surface cannot appear in more than one partition.

Model	Surfaces	Strikes	Maturities	Rows
Black–Scholes, time-dependent dividends	5,000	7	6	210,000
NIG, time-dependent dividends	6,000	7	6	252,000
R–GBA	3,000	7	4	84,000
ATS	4,000	7	4	112,000

2.3. Neural calibration and inverse problem

We describe how the numerical pricing map is converted into a fast inverse calibration tool. The procedure has two stages. In the offline stage, we generate complete option surfaces and train a forward neural network to map model inputs to put and call prices. In the online stage, the network weights are frozen and Differential Evolution searches for the parameters that best reproduce a target surface. The final parameters are then checked with the original numerical pricer.

The checks have a different meaning for synthetic and market data. For a synthetic surface, the parameters used to generate the prices are known, so we can measure both pricing and parameter errors. For Apple, the true parameters are unknown; we instead use hold-out prices, Bloomberg fields, and independent COS repricing as external diagnostics.

The first Black–Scholes dataset reproduces the setting of Liu et al. (2021), in which each observation represents one synthetic American put option priced with COS. We set $S_0 = 1$ and sample the strike K, maturity T, interest rate r, dividend yield q, and volatility σ over the ranges

$$K \in [0.60, 1.40], \qquad T \in [0.05, 3], \qquad r, q \in [-0.05, 0.10], \qquad \sigma \in [0.01, 1.05].$$

Approximately one million parameter combinations are generated through Latin hypercube sampling. This method divides the range of each input into intervals and selects combinations that cover the full set of possible inputs more evenly than independent random draws (McKay et al., 1979). Parameter combinations that do not satisfy the model checks are discarded before pricing.

In every surface experiment, the strikes are not sampled separately. Each surface uses the same seven fixed moneyness values in equation (2). At each strike and maturity, the dataset contains the corresponding put and call prices. The maturity grid is also fixed: Black–Scholes and NIG use

$$\left\{\frac{1}{12}, \frac{1}{4}, \frac{1}{2}, 1, 2, 3\right\}$$

years, while R–GBA and ATS use

$$\left\{\frac{1}{2}, 1, 2, 3\right\}$$

years.

What changes from one surface to another is the parameter vector used to generate it. This includes the interest rate, the forward dividend values, and the parameters of the selected pricing model. In the Black–Scholes surface experiments, the coefficients of the volatility smile also change between surfaces. Before a surface is retained, we check that the model conditions are satisfied and that all generated prices are finite, non-negative, and within the American-option bounds described in equations (15) and (16). Surfaces that fail these checks are discarded, so that the neural network is trained only on valid numerical prices.

The inverse-volatility experiment uses a separate dataset of individual options. Its aim is to recover volatility from individual American option prices. For this dataset, we retain only options with positive American time value, namely options whose price is above the value obtained from immediate exercise. When an option is exercised immediately, or is very close to this region, its price is determined mainly by the payoff and changes little when volatility changes. Such observations contain limited information for recovering volatility.

Removing these cases focuses the inverse-volatility experiment on options whose prices contain a meaningful time-value component. The positive time value also allows the logarithmic time-value transformation used as an input to the inverse network.

For individual Black–Scholes options, the inputs are $(\log(S_0/K), T, r, q, \sigma)$; NIG adds $(\sigma_{\mathrm{NIG}}, \eta, \kappa)$. In the surface experiments, the dividend yield is allowed to change across maturity intervals. For brevity, we refer to this time-dependent dividend-yield specification as TDQ, where q is the usual notation for the dividend yield. TDQ surface inputs contain the relevant volatility-smile or model coefficients together with the forward dividend values.

For an option expiring at T_j, only the dividend values up to T_j can affect its price. The dividend values for later intervals are therefore set to zero in the network input. This keeps the input length equal for all maturities, while ensuring that the price at T_j does not use information from future dividend intervals.

Put and call prices are learned jointly. We use the homogeneity of option prices to work with normalized values. A single normalization convention is used within each dataset, and the same saved scaler is applied during calibration. Before any price error is computed, the network output is mapped back to the original price scale. Mixing conventions within one dataset would change the objective and invalidate the recovered parameters.

The Black–Scholes replication also uses the American put–call symmetry described by Liu et al. (2021). This relation allows the price of a call with inputs (S_0, K, r, q) to be obtained from the price of a put after exchanging the spot and strike and also exchanging the interest rate and dividend yield:

$$(S_0, K, r, q) \longrightarrow (K, S_0, q, r).$$

Since the put and call prices are normalized by different strikes, the corresponding normalized relation is

$$\frac{C(S_0, K, r, q, \sigma)}{K} = \frac{S_0}{K}\frac{P(K, S_0, q, r, \sigma)}{S_0}. \tag{17}$$

We check this relation numerically by comparing the resulting call prices with calls priced directly.

When the dataset is divided into training, validation, and test sets, an observation and the corresponding observation obtained by this exchange are always assigned to the same set. Therefore, the transformed version of an option is not used for training while the original version is used for validation or testing.

The experiments use the same general calibration logic, but the validation is adapted to the information available in each case. For synthetic data, prices are generated from known parameters. These values are never given to the optimizer: they are used only after calibration to measure parameter, dividend, implied-volatility, and price-reconstruction errors. In every synthetic experiment, the recovered parameters are inserted again into the corresponding numerical COS pricer.

For Black–Scholes, the CRR tree provides an additional independent check. This comparison is not extended to NIG, R–GBA, or ATS, because a standard CRR tree is a Black–Scholes model and would not price the same non-Gaussian dynamics. For market data, the true parameters are unknown. The relevant validation is therefore the error on strikes not used in calibration, measured against Bloomberg MID prices, together with COS repricing of the recovered parameters.

Table 2 summarizes the main calibration settings. The experiments become progressively more flexible: they move from one matched put–call pair to smiles and complete strike–maturity surfaces. The preferred specification uses one dividend curve shared across strikes. For comparison, we also consider a control in which a different dividend value $q(K, T)$ is estimated for each strike and maturity. These controls are used to assess the effect of removing the strike-common restriction.

The validation procedures are not identical across all rows because the available information is different. Synthetic experiments allow direct comparison with the parameters used to generate the data. In the AAPL application, these parameters are not observed, so the relevant test is the ability to reproduce Bloomberg MID

Table 2: Main inverse problems and validation procedures. The numerical checks depend on the model and on whether the data are synthetic or observed market prices.

Experiment	Calibration unit and unknowns	Dividend specification	Validation
BS benchmark	One synthetic put–call pair; (σ, q)	One scalar q	Known-parameter errors; COS repricing; CRR check
BS smiles and TDQ surfaces	One smile or full (K,T) surface; volatility-surface coefficients and $q_1, \ldots, q_6$	$q(T)$ common across strikes; control with $q(K,T)$ varying by strike and maturity	Known-parameter and IV errors; COS repricing; selected CRR checks
NIG-generated smile with BS pricing	Black–Scholes calibration on a synthetic NIG-shaped smile	Common $q(T)$ and free-$q(K,T)$ diagnostic cases	Black–Scholes COS and CRR repricing; IV and dividend recovery
Direct NIG calibration	One smile or full (K,T) surface; $(\sigma_{\mathrm{NIG}}, \eta, \kappa)$ and dividend nodes	$q(T)$ common across strikes; control with $q(K,T)$ varying by strike and maturity	Known-parameter, dividend, and IV errors; NIG–COS repricing
R–GBA and ATS	Full synthetic surface; model-specific structural vector and four dividend nodes	$q(T)$ common across strikes; control with $q(K,T)$ varying by strike and maturity	Known-parameter, dividend, and equivalent-IV errors; model-specific COS repricing
Apple market Application	One observed smile for each expiry; Black–Scholes or NIG parameters and one dividend common across strikes	One maturity-specific implied dividend	Holdout MID-price errors; COS repricing; comparison with Bloomberg fields

prices on strikes excluded from calibration. In all cases, the recovered parameters are finally checked by pricing again with the original numerical COS model.

For each experiment, a neural network is trained to approximate the American pricing function. For one option quote, its inputs include moneyness, maturity, the interest rate, the dividend information, and the parameters of the corresponding model. Its outputs are the normalized put and call prices. The training targets are generated offline with the COS pricer.

The Black–Scholes benchmark follows the main architecture of Liu et al. (2021): four hidden layers with 200 neurons and Softplus activation functions. For the larger Black–Scholes and NIG surface datasets, we use four hidden layers with 256 neurons. The R–GBA and ATS networks use different hidden-layer sizes chosen during the implementation.

All input and output scalers are fitted on the training partition only. Training minimizes the mean squared error between network prices and COS prices. Training is performed in mini-batches using Adam or AdamW optimizers. The dataset is split into training, validation, and test partitions, keeping every row of the same synthetic surface in the same partition. Therefore, information from one surface cannot appear partially in training and partially in validation.

During training, we save the network state with the lowest validation loss. A learning-rate reduction and early stopping are used when validation performance no longer improves. The training history reports loss and R^2 on both training and validation data.

Figure 3 shows the training history for one representative Black–Scholes network. The figure reports the pricing error and R^2 on the training and validation sets as the training proceeds. Both errors decrease and remain close to each other, while the validation R^2 approaches one. This shows that the network reproduces the COS prices accurately also on data that were not used to update its weights. The network can therefore be used as a fast approximation of the pricing function during calibration.

The network evaluates one option quote at a time. During calibration, DE tests one candidate parameter vector at a time. For each candidate, the network is evaluated on all the quotes in the selected smile or surface. In this way, the calibration uses information from several strikes and maturities while imposing the chosen structure on volatility, dividends, and model parameters.

After calibration, the recovered parameters are inserted again into the numerical COS pricer. This verifies that

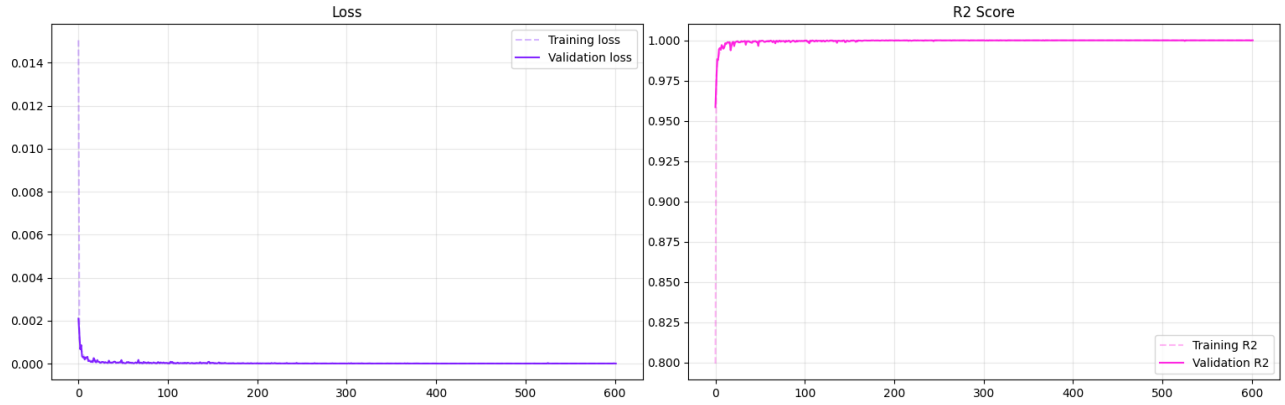


Figure 3: Representative training history for the symmetry-aware Black–Scholes put–call network. Training and validation loss decrease together, and the validation R^2 approaches one.

the parameters selected through the network also reproduce the target prices in the original pricing model.
After training, the network weights are fixed. The neural network is used as a fast approximation of the pricing function, while Differential Evolution (DE) searches for the model parameters that best reproduce the observed option prices. This follows the general CaNN approach of Liu et al. (2019a, 2021).
Let $\boldsymbol{x}_i$ denote the fixed inputs of option i, such as moneyness, maturity, and interest rate. Let $\boldsymbol{p}$ be the vector of parameters to be calibrated. Depending on the experiment, $\boldsymbol{p}$ can contain volatility, dividend nodes, NIG parameters, or the structural parameters of R–GBA and ATS. For a candidate vector $\boldsymbol{p}$, the trained network f_θ returns the predicted put and call prices:

$$(\widehat{y}_{P,i}(\boldsymbol{p}), \widehat{y}_{C,i}(\boldsymbol{p})) = f_\theta(\boldsymbol{x}_i, \boldsymbol{p}), \tag{18}$$

where $y_{P,i}$ and $y_{C,i}$ are the corresponding target put and call prices.
For each candidate vector $\boldsymbol{p}$, DE compares the prices predicted by the network with the target prices. It minimizes the average squared pricing difference across all the options in the selected smile or surface:

$$\widehat{\boldsymbol{p}} = \arg\min_{\boldsymbol{p}\in\mathcal{P}} J(\boldsymbol{p}), \qquad J(\boldsymbol{p}) = \frac{1}{n}\sum_{i=1}^{n}\left[(\widehat{y}_{P,i}(\boldsymbol{p}) - y_{P,i})^2 + (\widehat{y}_{C,i}(\boldsymbol{p}) - y_{C,i})^2\right], \tag{19}$$

where all price terms use the same normalization adopted in the relevant dataset. The optimization is subject to parameter bounds and model admissibility conditions. Put and call prices are calibrated jointly because volatility and dividends affect them in different ways.
The role of DE in a simple calibration is illustrated in figure 4. The figure shows the lowest value of the pricing objective at each generation for one synthetic Black–Scholes put–call pair. The horizontal axis is the generation number, while the vertical axis reports the pricing objective on a logarithmic scale. The error decreases from about 10^{-4} to 10^{-8} after 16 generations.
For this example, the recovered volatility and dividend differ from the values used to generate the prices by at most 4.4×10^{-4}. This shows that, in this simple setting, the neural network and DE recover the two unknown parameters accurately. The later calibrations of smiles and surfaces contain more unknown quantities. They therefore require more candidate parameter combinations and more iterations of the optimization.
The value $J(\widehat{\boldsymbol{p}})$ measures the fit seen by DE through the neural network. It is useful during optimization, but it is not sufficient by itself. For this reason, the recovered parameters are inserted again into the original COS pricer. The resulting COS prices are compared with the target prices using MAE or RMSE. This numerical repricing check verifies that the selected parameters also work in the original pricing model and not only in the neural approximation.
We also compare the network prices with the COS prices evaluated at the recovered parameters. Their difference measures the approximation error of the network at the final calibrated solution. For Black–Scholes, CRR provides a further independent comparison on selected cases. The numerical COS repricing is used for every model, whereas CRR is restricted to Black–Scholes because a standard CRR tree does not represent NIG, R–GBA, or ATS dynamics.
For synthetic data, the parameters used to generate each surface are known. We can therefore also compare recovered volatility, dividend values, and structural parameters with their generating values. For Apple application, the true model parameters are unknown, price errors are measured in USD against Bloomberg MID prices and Bloomberg implied-volatility and dividend fields are used as external benchmarks.

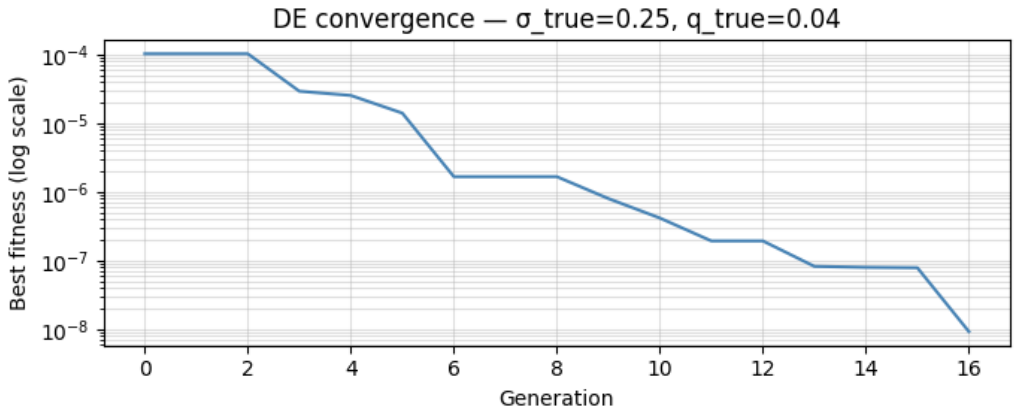


Figure 4: Differential Evolution convergence for one synthetic Black–Scholes put–call pair. The best objective value decreases by more than four orders of magnitude. The recovered values are $\widehat{\sigma} = 0.2501$ and $\widehat{q} = 0.0396$, compared with true values $\sigma = 0.25$ and $q = 0.04$.

The synthetic experiments are designed to study how stable the recovered parameters remain when the inverse problem becomes more flexible. Since the generating parameters are known, we can compare price reconstruction with the recovery of volatility, dividends, and structural model parameters.
The preferred specification uses one forward dividend curve $q(T)$ shared by all strikes. The synthetic option prices are generated with this same strike-common dividend curve. For comparison, we also run a control calibration in which a separate dividend value $q(K, T)$ is estimated for every strike and maturity.
This control does not represent a different dividend policy for each strike. Its purpose is to show whether different dividend estimates appear when the calibration is allowed to use many more free values. Comparing this case with the strike-common specification shows how the additional flexibility affects price reconstruction and parameter recovery.
The comparison between the constrained and unrestricted cases is based on three quantities: price errors after COS repricing, recovery errors relative to the known synthetic parameters, and the dispersion of the recovered dividend across strikes. In the NIG experiments, we also compare a conditional case with fixed NIG parameters with the fully joint calibration of NIG parameters and dividends. This shows separately the effect of freeing the dividend curve and the effect of freeing the non-Gaussian structural parameters.
The results show a consistent pattern. Prices and implied-volatility smiles can remain accurately reconstructed when several parameter combinations are allowed. However, individual dividend values and some structural parameters become less stable, especially at short maturities and in the fully unrestricted cases. The strike-common dividend curve is therefore the main economic specification, while the free $q(K, T)$ cases are retained as diagnostic stress tests.
Summing up, the neural network makes the inverse search substantially faster because DE evaluates candidate parameters through matrix operations rather than repeatedly running the American COS recursion. The financial model itself is unchanged: the COS pricer remains the source of the training prices and the final numerical check. The synthetic experiments additionally allow direct comparison with known generating parameters, while the Apple application uses holdout Bloomberg MID prices as the main out-of-sample validation.
With this framework in place, the next chapter begins with the Black–Scholes experiments and then progressively introduces more flexible volatility, dividend, and return specifications.
The final AAPL application uses the Black–Scholes and NIG versions of this procedure with observed Bloomberg MID prices rather than known synthetic inputs.

3. Synthetic results

3.1. Black–Scholes

We begin with Black–Scholes because it allows us to study the pricing model, the dividend specification, and the calibration procedure in a simple setting. We first reproduce the calibration of individual options used by Liu et al. (2021). We then extend the analysis to smiles and to sets of option prices observed at several strikes and maturities.
This sequence shows how the results change when more option prices are used and when the volatility or dividend specification becomes more flexible.

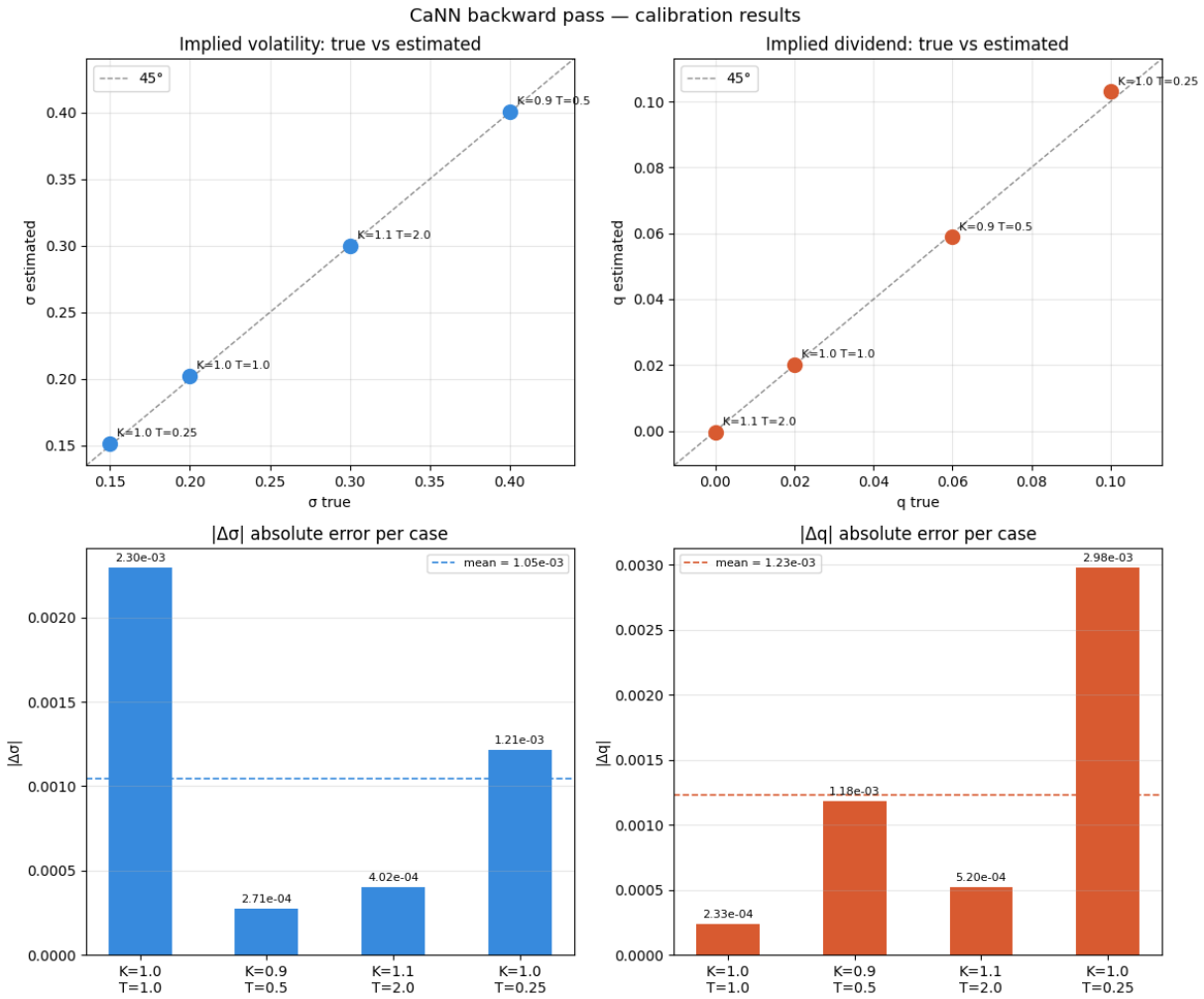


Figure 5: Black–Scholes CaNN–DE recovery for individual synthetic options. The upper panels compare the true and recovered volatility and dividend yield. The lower panels show the absolute errors for four synthetic strike–maturity cases. Put and call prices are calibrated jointly.

The first experiment reproduces the two main Black–Scholes steps of Liu et al. (2021). In the first step, a separate inverse neural network estimates the implied volatility of an American put. The synthetic parameters are sampled with Latin hypercube sampling and the American prices are generated with COS.

Options for which immediate exercise is optimal are excluded from this inverse-volatility dataset. In these cases, the option price is almost entirely explained by the exercise payoff and contains little information about volatility.

In the second step, a forward CaNN receives $(\log(S_0/K), T, r, q, \sigma)$ and returns the American put and call prices. Differential Evolution then recovers σ and q from target option prices. We use matched put and call prices jointly in this calibration. This provides more information than fitting only one option type, because volatility and the dividend yield affect puts and calls differently.

These initial tests also verify the numerical implementation. COS prices are compared with CRR in selected Black–Scholes cases, and the parameters found through CaNN–DE are inserted again into COS.

Figure 5 shows four synthetic examples. The recovered volatility and dividend values are close to the values used to generate the four examples. The figure illustrates calibration carried out on individual synthetic examples. The broader tests later in this section evaluate the method over complete sets of strikes and maturities.

The first Black–Scholes benchmark follows the calibration setting of Liu et al. (2021), in which one option contract, or one matched put–call pair, is calibrated at a time. This approach does not impose that the implied dividend recovered at one strike must be the same as the dividend recovered at another strike. Repeating the calibration separately across strikes can therefore produce different dividend estimates. In this thesis, these separate estimates are used as a benchmark for the joint calibration that follows.

We then calibrate the options at all strikes jointly. For a given maturity, one dividend yield is shared by all strikes, while volatility may vary with strike in order to describe a volatility smile. We next extend the same idea to several maturities: the dividend is represented by a curve $q(T)$ that is common across strikes and allowed

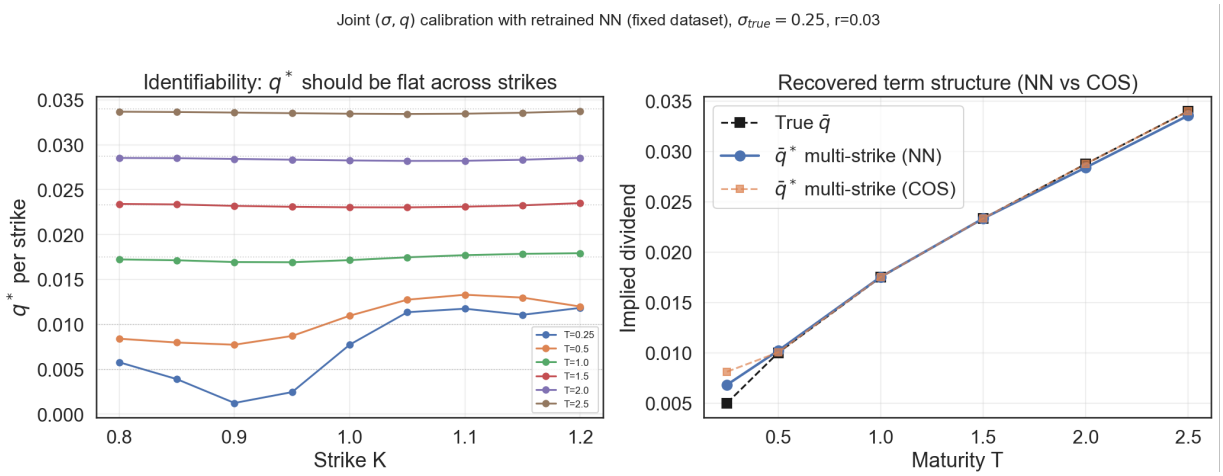


Figure 6: Preliminary Black–Scholes dividend term-structure diagnostic. Left: dividend estimates obtained by calibrating each strike separately. Dispersion is largest at short maturity although the true dividend is common across strikes. Right: maturity-average dividend yields recovered by joint multi-strike calibration with the neural network and direct COS. The later TDQ experiments price directly with forward interval yields.

to change over time.

This restriction has a direct economic meaning. At a given date, one stock does not have a different dividend policy for each strike. Therefore, a dividend estimate that changes with strike is not interpreted as a literal corporate dividend. The cases in which a separate dividend is estimated at each strike are retained as control experiments. They show whether strike-dependent estimates appear when the calibration is given more free parameters, for example because volatility and dividend are difficult to separate from the available option prices.

As a preliminary term-structure test, volatility is fixed at its true value and one maturity-average dividend $\bar{q}(T_j)$ is recovered at each maturity. The corresponding forward interval yield is obtained through

$$\widehat{q}_j = \frac{T_j \widehat{\bar{q}}(T_j) - T_{j-1} \widehat{\bar{q}}(T_{j-1})}{T_j - T_{j-1}}. \tag{20}$$

This bootstrap is used only to inspect the relation between maturity-average and forward yields. The final TDQ experiment below does not rely on this conversion: it prices directly with the forward interval yields at every American exercise step.

Figure 6 compares separate strike-by-strike calibration with joint multi-strike calibration. At the shortest maturity, the range of the separate dividend estimates is 1.059×10^{-2}. It decreases to 3.10×10^{-4} at $T = 2.5$, showing that short maturities contain less information about the dividend. The multi-strike neural calibration recovers the maturity-average dividend with an error of 1.813×10^{-3} at $T = 0.25$ and with errors below 4.42×10^{-4} at the longer maturities.

Table 3 compares two calibrations of the same seven matched put–call pairs. In the first calibration, one dividend value is estimated for every strike. In the second, one common dividend is estimated from all strikes together. Volatility remains free at every strike in both cases.

The specification with one dividend per strike has a smaller neural price MSE because it has more parameters. Here, the price MSE is the mean squared difference between the prices returned by the neural network and the target prices. However, the dividend MAE is not lower and the recovered dividends differ across strikes. Since the underlying stock has one dividend policy, the common-dividend specification is retained in the later experiments.

The preceding tests use a constant true volatility in order to isolate the recovery of the dividend. We next introduce a smooth volatility smile, so that volatility can change across strikes. This is important because an incorrect volatility specification may otherwise be absorbed by the dividend estimate.

When the calibration includes the volatility smile, one dividend curve common across strikes can still be used. If volatility is constrained too strongly, allowing a separate dividend for every strike can partly compensate for the missing smile. The cases with free $q(K, T)$ are therefore used only to study this effect; they are not the preferred dividend specification.

The final Black–Scholes experiment uses the BS–TDQ specification, in which volatility is described by a smooth surface and the dividend is represented by forward yields that can change across maturities. Each of the 5,000 synthetic surfaces contains seven strikes and six maturities, for a total of 210,000 option rows. Volatility is

Table 3: Controlled Black–Scholes comparison at $T = 1$ with true $q = 0.02$. The free-$q(K)$ case is a benchmark in which each strike has its own dividend value; the common-q case uses the economically preferred restriction that one dividend is shared by all strikes. Both specifications use the same seven matched put–call pairs and one volatility value at each strike.

Dividend specification	q values	q MAE	Standard deviation across strikes	Price MSE
Free $q(K)$	7	4.52×10^{-4}	3.80×10^{-4}	9.11×10^{-16}
Common q	1	4.25×10^{-4}	0	5.38×10^{-8}

Table 4: Representative BS–TDQ surface recovery. Dividend errors are in decimal yield units; for example, 10^{-3} corresponds to 0.1 percentage points.

Quantity	True	Recovered	Absolute error
ATM volatility level	0.274193	0.274423	2.30×10^{-4}
Smile skew	0.102006	0.102802	7.96×10^{-4}
Smile curvature	0.191425	0.197143	5.72×10^{-3}
Term slope	−0.039641	−0.039845	2.04×10^{-4}
Forward dividend q_1	0.028269	0.027344	9.25×10^{-4}
Forward dividend q_2	0.036948	0.037269	3.22×10^{-4}
Forward dividend q_3	0.027097	0.027594	4.97×10^{-4}
Forward dividend q_4	0.017979	0.017626	3.54×10^{-4}
Forward dividend q_5	0.017850	0.017549	3.01×10^{-4}
Forward dividend q_6	0.021299	0.021339	3.96×10^{-5}
Mean absolute dividend error	Six dividend values		4.06×10^{-4}

generated from a surface described by four coefficients: its level, skew, curvature, and maturity dependence. The dividend is generated as six forward interval yields $(q_1, \ldots, q_6)$. At maturity T_j, the price uses only the dividend information available up to that maturity.

The calibration estimates the volatility-surface coefficients and the forward dividend curve jointly. Table 4 reports the results for one representative set of option prices. In this example, the mean absolute error across the six forward dividend yields is 4.06×10^{-4}, while the largest individual error is 9.25×10^{-4}.

This single example illustrates how the calibration works when volatility and the dividend curve are both unknown. It is not sufficient, however, to assess the method over the full range of simulated parameter values. For this reason, we next repeat the same calibration on every surface in the held-out test set.

The full test set contains 500 surfaces that were not used during network training or validation. Each surface has its own volatility surface and its own six-interval forward dividend curve. The dividend is common across strikes, but it is allowed to change across maturities.

Across the 500 test surfaces, the mean absolute error of the recovered forward dividend yields is 9.33×10^{-4}. The mean absolute error of the reconstructed volatility surface is 5.46×10^{-4}. When the recovered parameters are inserted again into the COS pricer, the mean normalized repricing MAE is 1.85×10^{-4} for puts and 1.86×10^{-4} for calls.

Figure 7 summarizes these results over the full test set. It does not show one selected calibration: every point and every error summary is computed from the 500 test surfaces.

Table 5: BS–TDQ calibration on the held-out synthetic test set. The dividend is allowed to change across the six maturity intervals and is common across strikes. Dividend and volatility errors are reported in decimal units. Repricing errors use prices normalized by the strike. COS repricing uses the parameters recovered by CaNN–DE.

Quantity	Value
Test surfaces calibrated	500
Forward dividend MAE	9.33×10^{-4}
Forward dividend RMSE	1.64×10^{-3}
Median forward dividend MAE	7.41×10^{-4}
Volatility-surface MAE	5.46×10^{-4}
COS put repricing MAE	1.86×10^{-4}
COS call repricing MAE	1.81×10^{-4}
Normalized COS put repricing MAE	1.85×10^{-4}
Normalized COS call repricing MAE	1.86×10^{-4}

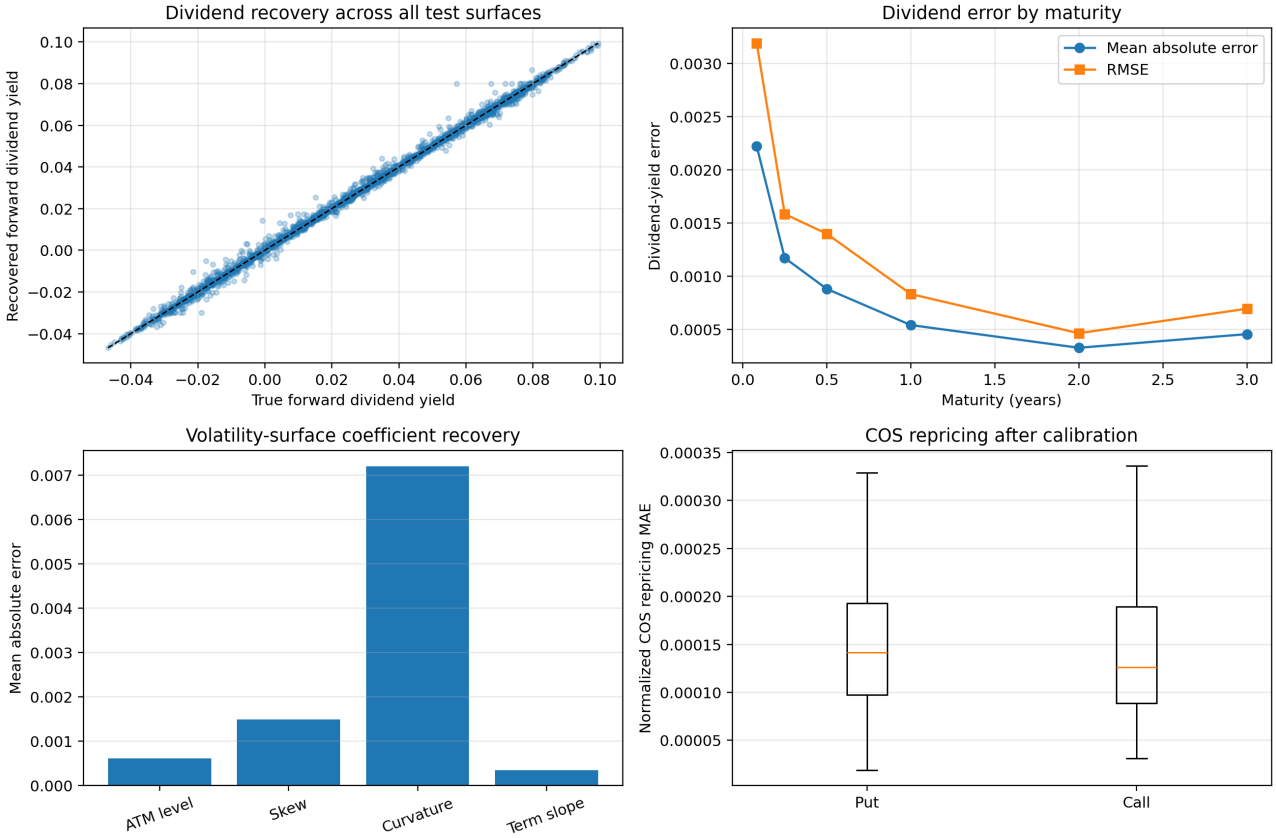


Figure 7: BS–TDQ recovery on all 500 held-out synthetic test surfaces. The upper-left panel compares the 3,000 recovered forward dividend yields with the values used to generate the surfaces; the dashed line represents exact recovery. The upper-right panel reports the dividend-yield error at each maturity. The lower-left panel reports the mean absolute errors of the four volatility-surface coefficients. The lower-right panel shows the distribution across surfaces of the normalized COS repricing MAE for puts and calls after calibration.

The upper-left panel shows that the recovered forward dividend yields are generally close to their generating values. The upper-right panel shows that the recovery is less precise at the shortest maturity. The mean absolute error is 2.22×10^{-3} at one month and falls to 3.27×10^{-4} at two years. Thus, the full test-set results support the recovery of a maturity-dependent dividend curve, while also showing that its shortest end is less stable.

The lower panels provide two additional checks. The volatility-surface coefficients have different scales and roles in the surface specification, so their errors are not directly comparable one by one. The resulting volatility-

Table 6: Total calibration time and parameter recovery for the Black–Scholes multi-strike experiment. Each method calibrates six maturities, using seven matched put–call pairs at each maturity. Errors are absolute deviations from the known synthetic volatility and dividend yield.

Method	σ MAE	q MAE	Total time (s)	Relative time
CaNN–DE	9.37×10^{-4}	3.71×10^{-4}	101.37	$\times 1.0$
CRR–DE ($M = 300$)	5.37×10^{-5}	1.38×10^{-5}	265.57	$\times 2.62$
Direct COS–DE	0	0	9959.61	$\times 98.25$

surface MAE reported in table 5 is the more relevant overall measure. Finally, the lower-right panel shows that the parameters found through CaNN–DE also reproduce the option prices when they are evaluated again with COS.

The final Black–Scholes comparison considers the same multi-strike synthetic smiles with three calibration methods: CaNN–DE, direct COS–DE, and CRR–DE with 300 tree steps. Each maturity is calibrated separately using seven matched put–call pairs. The reported time is the total time required to calibrate the six maturities. At each maturity, the calibration uses the same seven matched put–call pairs. The timing begins when the parameter search starts and therefore excludes the earlier generation of the COS dataset and the training of the neural network.

Table 6 reports the mean parameter errors over the six maturities and the total calibration time. In this implementation, CaNN–DE is about 2.6 times faster than CRR–DE and about 98 times faster than direct COS–DE. Its parameter errors are larger than those of direct numerical calibration, but remain of order 10^{-4}–10^{-3}.

The displayed zero error for direct COS–DE means that the synthetic prices are generated and calibrated with the same COS model, up to the numerical precision shown in the table. CRR–DE is also very accurate, but it is slower than CaNN–DE. The neural network is therefore used to reduce the repeated pricing cost during calibration, while COS remains the model used to generate the synthetic prices and to check the recovered parameters.

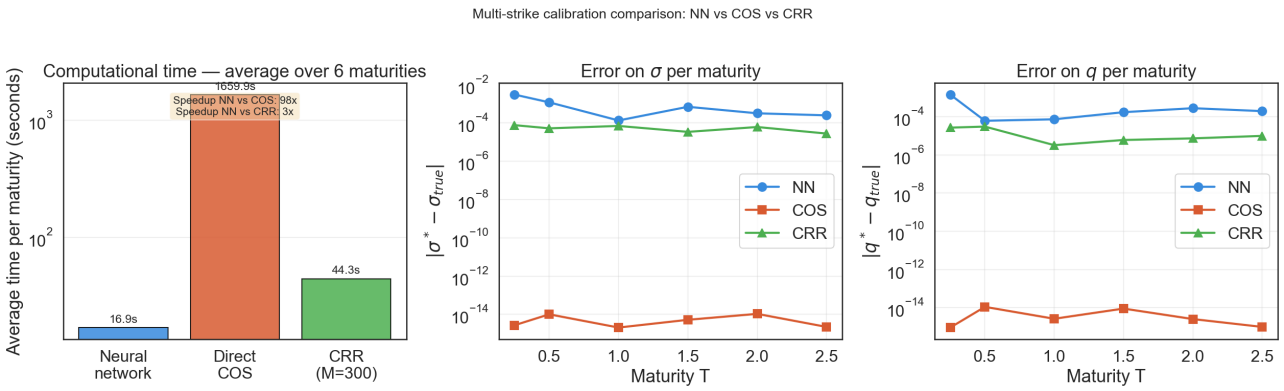


Figure 8: Speed–accuracy comparison for multi-strike Black–Scholes calibration. The left panel reports the total calibration time for the six maturities. The remaining panels show absolute volatility and dividend errors relative to their known synthetic values.

Figure 8 compares the three methods. The left panel shows the total time needed to calibrate all six maturities. The two right panels show the volatility and dividend errors at each maturity. Direct COS–DE has errors close to numerical precision, CRR–DE is also highly accurate, and CaNN–DE obtains a shorter calibration time at the cost of slightly larger parameter errors.

CRR is used only in the Black–Scholes comparison because it is an independent tree approximation of Black–Scholes dynamics. It does not represent the non-Gaussian transition laws of NIG, R–GBA, or ATS. For these models, the corresponding COS pricer is used for the final numerical repricing check.

In conclusion, the Black–Scholes experiments show that calibrating matched puts and calls jointly across several strikes improves the recovery of the dividend. A dividend curve that is common across strikes remains the preferred specification because a stock does not have a different dividend policy for different strikes. The BS–TDQ test over 500 previously unseen surfaces shows that the same procedure can recover both a structured volatility surface and a maturity-dependent forward dividend curve. The recovery is less precise at the shortest maturity, which is consistent with the earlier preliminary tests. The next section replaces Black–Scholes dynamics with

the NIG model and examines whether these findings remain valid with non-Gaussian returns.

3.2. Normal Inverse Gaussian (NIG) model

The NIG model extends the Black–Scholes experiments by allowing non-Gaussian returns with asymmetry and heavier tails. It is used in two different ways. First, it generates a non-flat volatility smile while American option prices are still computed with Black–Scholes. Second, its characteristic function is inserted directly into the American COS pricer. Only the second part is an American pricing and calibration experiment under NIG dynamics.

To display the effect of the NIG parameters on option prices, we use the Black–Scholes-equivalent implied volatility. The relation is

$$(\sigma_{\mathrm{NIG}}, \eta, \kappa) \longrightarrow \phi_{\mathrm{NIG}} \longrightarrow C_{\mathrm{NIG}}(K,T) \longrightarrow \sigma_{\mathrm{IV}}(K,T), \tag{21}$$

where ϕ_{NIG} is the NIG characteristic function, $C_{\mathrm{NIG}}(K,T)$ is the corresponding European NIG call price, and $\sigma_{\mathrm{IV}}(K,T)$ is obtained by inverting the European Black–Scholes price. The equivalent implied-volatility smile is a convenient way to compare price shapes. It is not itself a NIG parameter: the structural quantities to be recovered remain $(\sigma_{\mathrm{NIG}}, \eta, \kappa)$ together with the dividend curve.

The first NIG experiments use European NIG prices only to construct a non-flat implied-volatility smile. This smile is then used as input to the American Black–Scholes COS pricer. The purpose is to understand whether a more realistic strike-dependent volatility affects dividend recovery before changing the underlying pricing dynamics.

When the Black–Scholes calibration is allowed to represent the NIG-generated smile, one dividend common across strikes can be recovered accurately. When volatility is restricted too strongly, the dividend can compensate for the missing part of the smile and may vary across strikes. This confirms the Black–Scholes result: a strike-dependent recovered dividend is not evidence that the company dividend depends on strike. It is a signal that the volatility and dividend information cannot be cleanly separated under that specification. This first stage does not estimate NIG parameters; it remains a Black–Scholes pricing experiment with an NIG-shaped volatility.

We next use the NIG characteristic function directly inside the American COS recursion. The characteristic function, martingale correction, moment condition in equation (6), price bounds, and COS convergence are checked before generating the dataset. The NIG forward network is then trained on synthetic American put and call prices.

The final NIG–TDQ dataset contains 6,000 synthetic surfaces and 252,000 option rows. Each surface has seven strikes and six maturities. One NIG vector $(\sigma_{\mathrm{NIG}}, \eta, \kappa)$ is shared by the entire surface, while the dividend is represented by six forward interval yields. At maturity T_j, only the dividend information available up to that maturity enters the price. The network takes log-moneyness, maturity, rate, the NIG parameters, and the forward dividend values available up to the option maturity. It returns normalized put and call prices.

Figure 9 shows the forward validation on one held-out surface. The CaNN prices remain close to the COS targets over the full strike and maturity grid. This confirms that the network can reproduce the NIG American pricing map before it is used inside Differential Evolution.

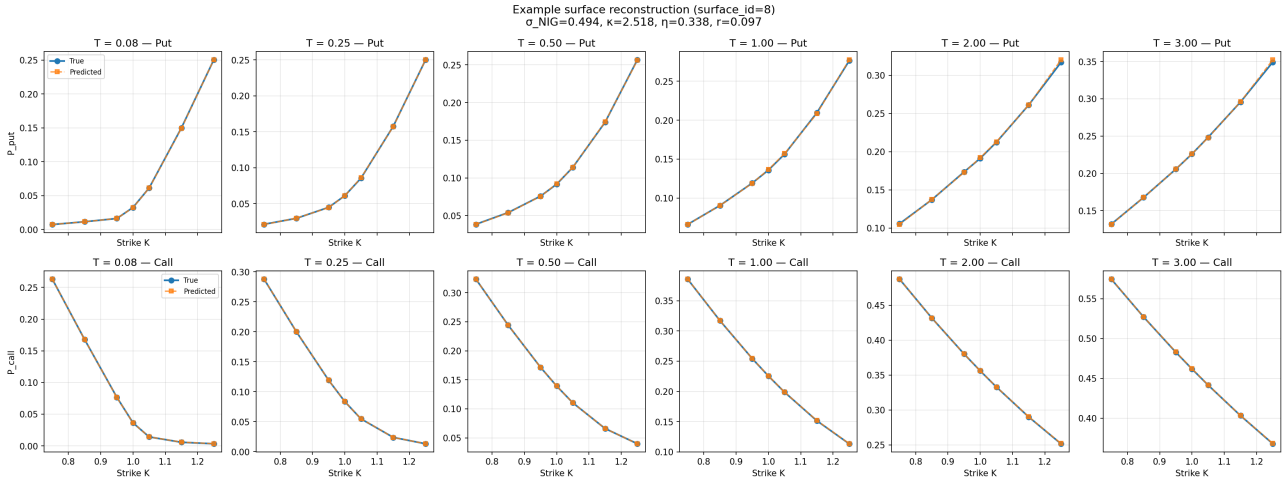


Figure 9: Forward NIG–TDQ surface validation with piecewise-forward dividends. COS-generated synthetic put and call prices are compared with CaNN predictions over all seven strikes and six maturities of one held-out surface. The close agreement preserves both the strike shape and the maturity evolution of the price surface.

The main calibration uses one shared set of NIG parameters and one forward dividend curve for each surface. Table 7 compares this CaNN–DE calibration with two deliberately smaller optimization budgets. For the same seven surfaces, we also compare the equivalent implied-volatility smile generated by the recovered and true NIG parameters. In this comparison, the dividend curve is kept equal to the curve used to generate the surface. Therefore, the IV MAE measures the effect of errors in the NIG parameters, while the separate q MAE column measures the recovery of the dividend curve.

The large neural search recovers σ_{NIG}, η, and κ with mean relative errors of 0.793%, 1.124%, and 2.091%, respectively. The mean absolute error of the six dividend nodes is 1.268×10^{-3}, the equivalent IV MAE is 4.80×10^{-4}, and the COS-repriced put MAE is 2.10×10^{-4}.

The direct-COS row is not an equally converged competitor. It uses only a population of 16, 12 generations, and 64 COS coefficients in order to keep the calculation feasible. Its larger errors therefore show the difficulty of searching directly over a high-dimensional American NIG problem with a small budget; they do not imply that the exact COS objective has a worse optimum.

Table 7: NIG–TDQ calibration over seven held-out synthetic surfaces. All entries are averages over the seven calibrations. NIG parameter errors are relative to the known generating values. Dividend and IV MAEs are reported in decimal units. The IV MAE compares the Black–Scholes-equivalent IV smile generated by the recovered and true NIG parameters, using the true dividend curve in both calculations. Put MAE compares COS-repriced normalized prices at the recovered parameters with the synthetic COS targets.

Method	σ_{NIG} err.	κ err.	η err.	q MAE	IV MAE	Put MAE	Mean time per surface (s)
CaNN–DE (large)	0.793%	2.091%	1.124%	1.268×10^{-3}	4.80×10^{-4}	2.10×10^{-4}	6.67
CaNN–DE (reduced)	10.47%	23.88%	20.54%	5.159×10^{-3}	6.03×10^{-3}	1.937×10^{-3}	0.108
COS–DE (reduced)	28.12%	45.03%	297.10%	7.714×10^{-3}	1.92×10^{-2}	2.794×10^{-3}	1093.23

The comparison illustrates the practical role of the neural network. Once trained, the network allows Differential Evolution to evaluate a substantially larger number of NIG parameter vectors in a short time. COS remains essential: it generates the synthetic target prices and is used again to check the prices obtained from the recovered parameters. The advantage of the neural network is therefore computational: repeatedly calling the American COS pricer during a calibration with many unknown parameters is considerably more expensive.

The final experiments test the effect of relaxing the strike restriction on the dividend. In the preferred specification, the forward dividend curve $q(T)$ is common across strikes. In the unrestricted case, the calibration uses one dividend value for every strike and maturity, $q(K, T)$. Together with the three NIG parameters, this gives 45 unknown quantities on a surface with 84 put and call prices.

We first fix the NIG parameters at their known synthetic values and estimate only the unrestricted dividend. This separates dividend recovery from the uncertainty in the NIG vector. Across 40 surfaces, dividend MAE is

1.228×10^{-3}. The mean cross-strike standard deviation decreases from 5.276×10^{-3} at the shortest maturity to 8.02×10^{-4} at the longest maturity. Therefore, the recovered dividend becomes approximately constant across strikes when maturity is long enough, even though this restriction is not imposed.

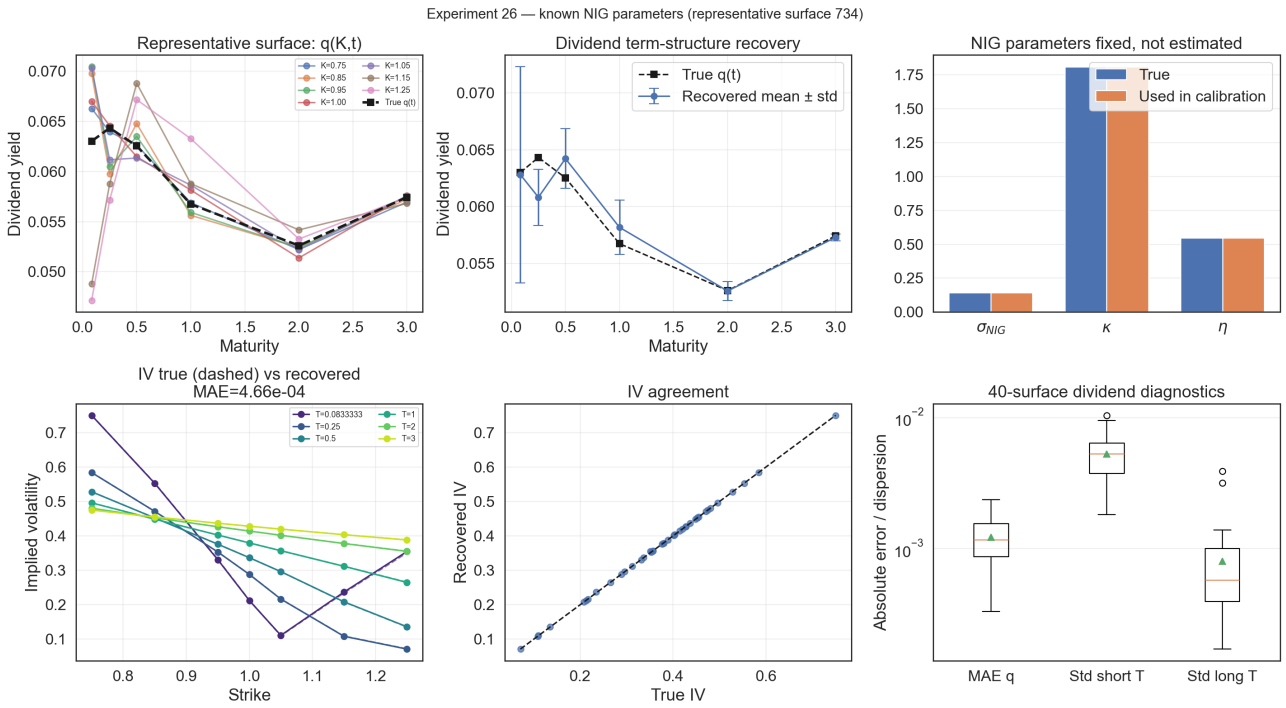


Figure 10: Free $q(K,T)$ calibration with NIG parameters fixed at their known synthetic values. The upper panels show the strike-by-strike recovered dividends, their average term structure, and the fact that the NIG parameters are fixed in this experiment. The lower panels show the NIG-equivalent implied-volatility recovery and the dividend diagnostics over 40 synthetic surfaces. The recovered dividend becomes nearly constant across strikes at medium and long maturities, while the short end remains more dispersed. The IV MAE on the displayed surface is 4.66×10^{-4} in decimal-volatility units.

We then estimate the NIG vector and all dividend values jointly. Across eight surfaces, the dividend MAE increases to 3.584×10^{-3} and the cross-strike standard deviation is 5.062×10^{-3}. The mean relative errors are 5.707% for σ_{NIG}, 18.643% for κ, and 28.922% for η. At the same time, normalized COS-repricing MAEs remain 8.33×10^{-4} for puts and 1.10×10^{-3} for calls.

On the representative surface shown in figure 11, the recovered equivalent implied-volatility smile remains close to the true smile, with MAE 3.66×10^{-3}.

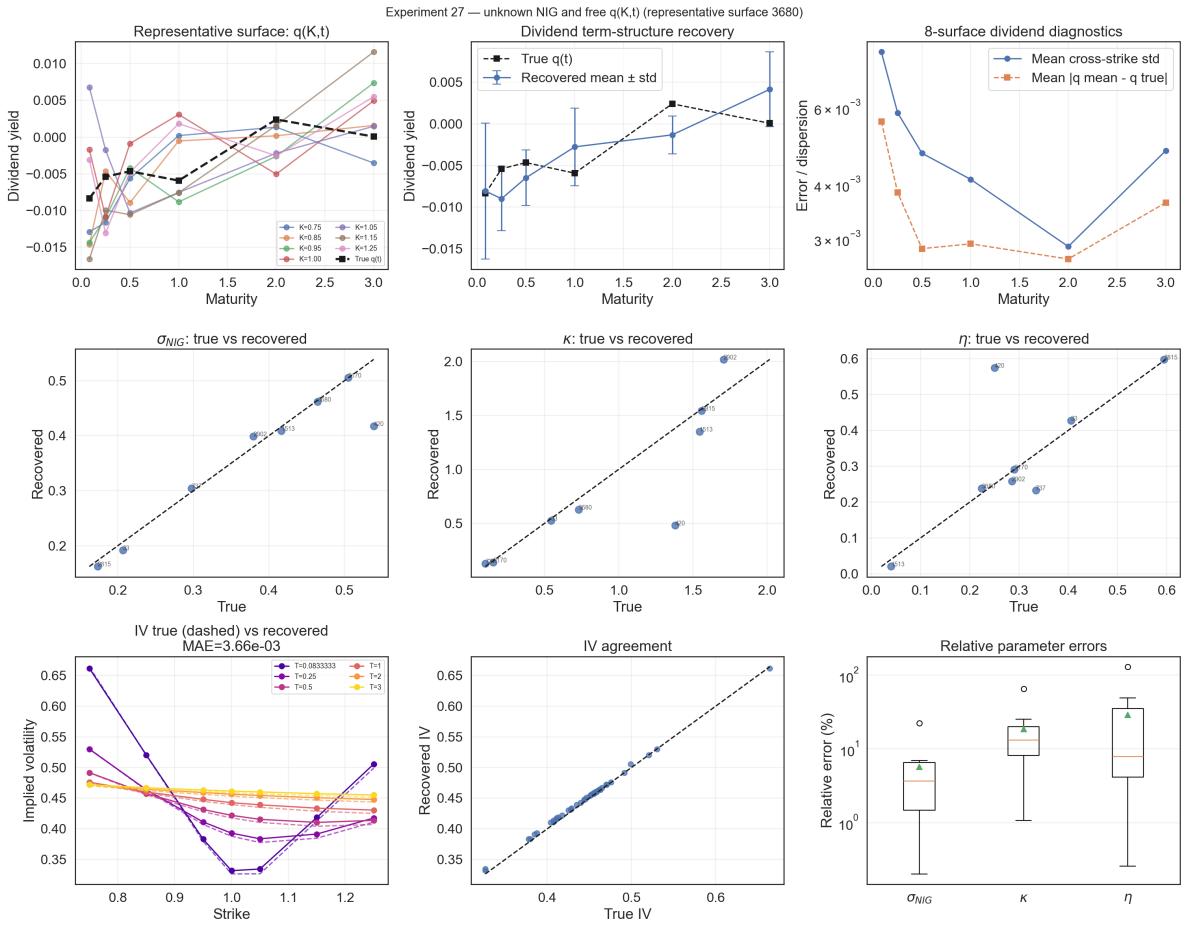


Figure 11: Joint calibration of the NIG vector and the unrestricted dividend $q(K,T)$ on one representative synthetic surface. The upper panels show the dividend estimates across strikes and maturities. The middle panels compare the true and recovered NIG parameters. The lower panels compare the true and recovered NIG-equivalent implied-volatility smiles. Prices and the smile remain close to the targets, while the individual NIG parameters and the strike-dependent dividend values are less stable.

Table 8: Effect of freeing the NIG parameters and the dividend curve. The known-NIG case isolates dividend recovery, whereas the fully joint case estimates the NIG vector and all dividend values simultaneously. Dividend and price entries are averages over the successful synthetic surfaces. Price MAEs are normalized COS-repricing errors for puts and calls.

Specification	Surfaces	Dividend recovery (q MAE and cross-strike dispersion)	COS-repricing MAE (1. put / 2. call)
Known NIG parameters	40	q MAE: 1.228×10^{-3} Dispersion: from 5.276×10^{-3} at short maturity to 8.02×10^{-4} at long maturity	1. 1.111×10^{-3} 2. 2.125×10^{-4}
Jointly free NIG and $q(K,T)$	8	q MAE: 3.584×10^{-3} Dispersion: 5.062×10^{-3}	1. 8.33×10^{-4} 2. 1.10×10^{-3}

The first NIG experiments show that a strike-dependent dividend can be created artificially when a Black–Scholes calibration cannot represent a non-flat smile. Direct American NIG pricing removes this inconsistency by using the NIG characteristic function in the COS recursion.

For the main NIG–TDQ calibration, one NIG vector shared across the surface and one dividend curve shared across strikes provide accurate price reconstruction, a close equivalent implied-volatility smile, and useful recovery of the synthetic dividend curve. When all NIG parameters and all dividend values are free, prices and implied volatilities remain accurate but the individual quantities become less stable.

This supports the use of a forward dividend curve shared across strikes and a limited number of NIG parameters. We next test whether the same conclusion also holds for the R–GBA and ATS models.

3.3. R–GBA and ATS extensions

This section studies whether the American pricing and calibration framework can also be used with the R–GBA and ATS additive-process models. These models have richer return distributions than Black–Scholes and different parameters from NIG. The purpose is therefore not only to obtain small price errors, but also to understand which quantities can be recovered reliably when the model becomes more flexible.
For both models, each synthetic surface contains seven strikes and four maturities. The American COS pricer generates matched put and call prices. We use four forward dividend values, one for each maturity interval. Each complete surface is assigned to only one of the training, validation, or test sets. The recovered parameters are then checked by COS repricing.
The true synthetic dividend is common across strikes. We consider three dividend specifications. In the first and third tests, the calibration is allowed to assign different dividend values to different strikes. These are diagnostic tests: they show whether allowing more parameters produces a strike-dependent recovered dividend even though the dividend used to generate the data is common across strikes. The preferred specification is a dividend curve $q(T)$ shared by all strikes and allowed to vary with maturity.
Following the reduced-form specification discussed by Azzone and Torricelli (2025), the R–GBA term functions are defined in equation (8). The source reference studies European additive-process pricing. Here, the same dynamics are used in the American COS backward induction and in the neural calibration procedure.
The structural parameters are sampled from

$$\alpha_0 \in [1.2, 3.5], \qquad \beta_0 \in [1.6, 7], \qquad \sigma_0 \in [0.08, 0.35], \tag{22}$$

with $H_0 = 0.5$ and forward dividend yields in $[-0.01, 0.07]$. We impose $\beta_0 > \alpha_0$. As discussed in Azzone and Torricelli (2025), the equality $\alpha_0 = \beta_0$ corresponds to a zero-skew boundary, whereas $\beta_0 > \alpha_0$ describes the negative-skew region relevant for equity options. The strict inequality is therefore used to focus the synthetic experiment on this region and to avoid evaluating the boundary itself.
The R–GBA dataset contains 3,000 surfaces, corresponding to 84,000 option rows. Each input includes log-moneyness, maturity, interest rate, the four forward dividend values, and the three R–GBA parameters. The targets are normalized put and call prices.
Table 9 reports the results of the three dividend experiments. All three specifications give small errors when the recovered parameters are inserted into the COS pricer. The specification with one value of $q(T)$ for each maturity is the fastest and, by definition, uses the same dividend for all strikes. When a different dividend is allowed for each strike, the dividend MAE is slightly lower in this small sample, but the calibration requires more parameters. In the last experiment, although the dividend is free to vary across both strikes and maturities, the recovered values remain similar across strikes with the same maturity. This result is consistent with the fact that the dividend belongs to the underlying stock and should not depend on the option strike. The more flexible calibration, however, takes longer.

Table 9: R–GBA calibration results. The q MAE compares recovered and generating synthetic dividends in decimal units. The standard deviation across strikes measures the dispersion of the recovered dividend estimates. COS put and call errors are normalized MAEs between COS repricing at the recovered parameters and the COS prices generated from the true parameters. The final column reports the mean calibration time per surface.

Dividend specification	Surfaces	q MAE	Std. across strikes	COS put MAE	COS call MAE	Mean time per surface (s)
Constant true, free $q(K)$	10	1.78×10^{-4}	1.79×10^{-4}	9.54×10^{-5}	1.23×10^{-4}	4.46
Common $q(T)$	10	4.25×10^{-4}	≈ 0	1.27×10^{-4}	2.09×10^{-4}	2.28
Common true $q(T)$, free $q(K,T)$	5	3.13×10^{-4}	5.30×10^{-5}	6.64×10^{-5}	1.17×10^{-4}	5.84

For the preferred specification with one dividend curve $q(T)$ shared across strikes, the mean absolute errors of the R–GBA parameters are

$$\Big(\mathrm{MAE}(\widehat{\alpha}_0),\ \mathrm{MAE}(\widehat{\beta}_0),\ \mathrm{MAE}(\widehat{\sigma}_0)\Big) = (0.231,\ 0.702,\ 0.0125). \tag{23}$$

The three parameters have different ranges in the synthetic dataset, so their absolute errors should not be compared directly. Relative to the corresponding sampling intervals, these errors are approximately 10.0%, 13.0%, and 4.6%, respectively. Thus, σ_0 is recovered more stably than α_0 and β_0, while β_0 is the least stable parameter. At the same time, the COS-repricing errors in table 9 remain small. This shows that accurate option prices do not necessarily imply equally accurate recovery of every structural parameter.

Figure 12 shows one representative calibration with a dividend curve shared across strikes. The CaNN predictions and COS repricing remain close to the synthetic targets, while the structural panels make the weaker recovery of β_0 visible.

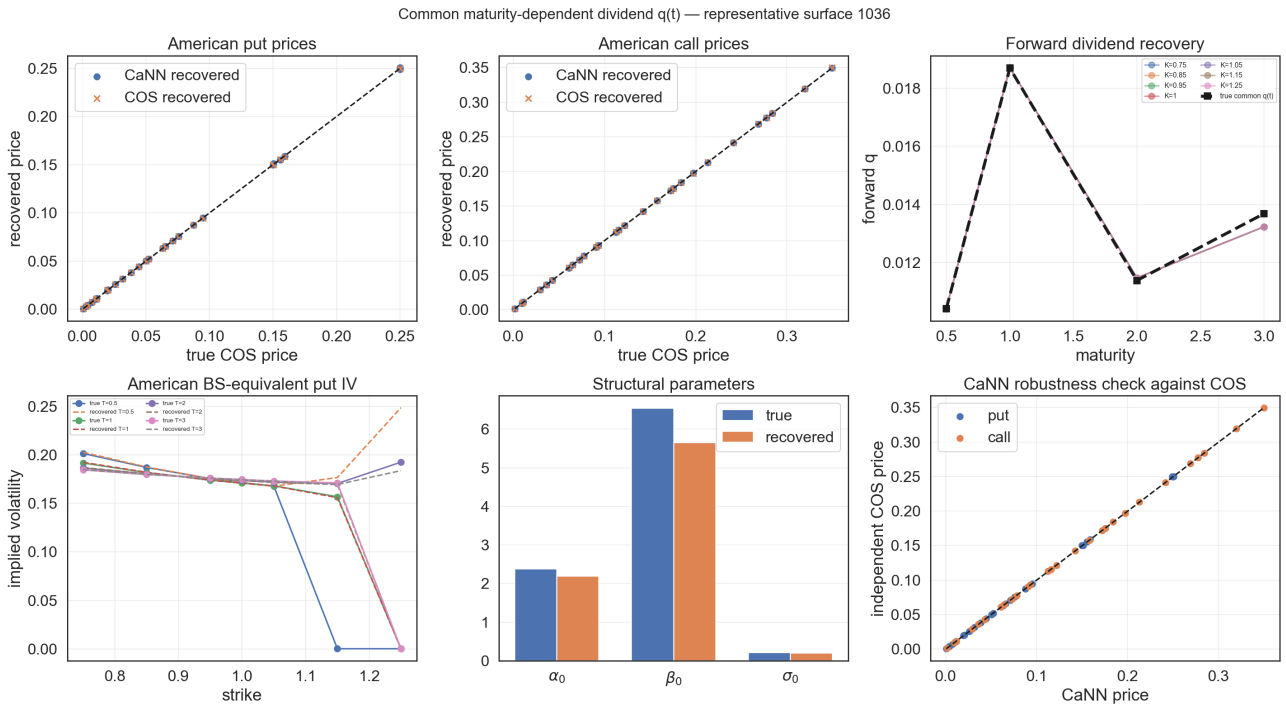


Figure 12: R–GBA calibration with a maturity-dependent dividend shared across strikes. The figure compares the true and recovered dividend curve, the structural parameters, the CaNN prices, and COS prices obtained by repricing the recovered parameters. The price points remain close to the diagonal, whereas β_0 is less accurately recovered than the other structural quantities.

For ATS, the sampled parameters are

$$\sigma \in [0.10, 0.35], \qquad \kappa_0 \in [0.2, 1.4], \qquad \eta_0 \in [0.05, 0.4], \tag{24}$$

$$\beta \in [0.05, 0.75], \qquad \delta \in [-0.35, -0.05], \qquad q_j \in [-0.02, 0.08]. \tag{25}$$

The dataset contains 4,000 surfaces, or 112,000 option rows. The input contains the five ATS parameters, the four forward dividend values, and the contract variables. The forward network has a very accurate held-out fit; its put and call R^2 values are reported together with the other models in table 11.
The same three dividend specifications are used for ATS. The results in table 10 again show small repricing errors in every case. The strike-common curve is the fastest specification. The free cases recover almost strike-constant dividends, but their additional flexibility does not provide a clearer economic interpretation.

Table 10: ATS calibration summary. The q MAE compares recovered and generating synthetic yields in decimal units. The standard deviation across strikes measures the dispersion of the recovered dividend estimates. COS put and call entries are normalized MAEs between ATS–COS repricing at recovered parameters and ATS–COS prices generated from the true parameters. The final column reports the mean calibration time per surface.

Dividend Specification	Surfaces	q MAE	Std. across strikes	COS put MAE	COS call MAE	Mean time per surface (s)
Constant true, free $q(K)$	10	4.79×10^{-4}	5.59×10^{-4}	1.18×10^{-4}	3.06×10^{-4}	9.81
Common $q(T)$	10	1.316×10^{-3}	≈ 0	2.34×10^{-4}	3.91×10^{-4}	6.30
Common true $q(T)$, free $q(K,T)$	5	8.73×10^{-4}	9.0×10^{-5}	1.59×10^{-4}	2.59×10^{-4}	16.12

Figure 13 gives the corresponding ATS diagnostics. The figure confirms that the recovered prices and implied-volatility shapes remain close to the synthetic targets under all three dividend specifications.

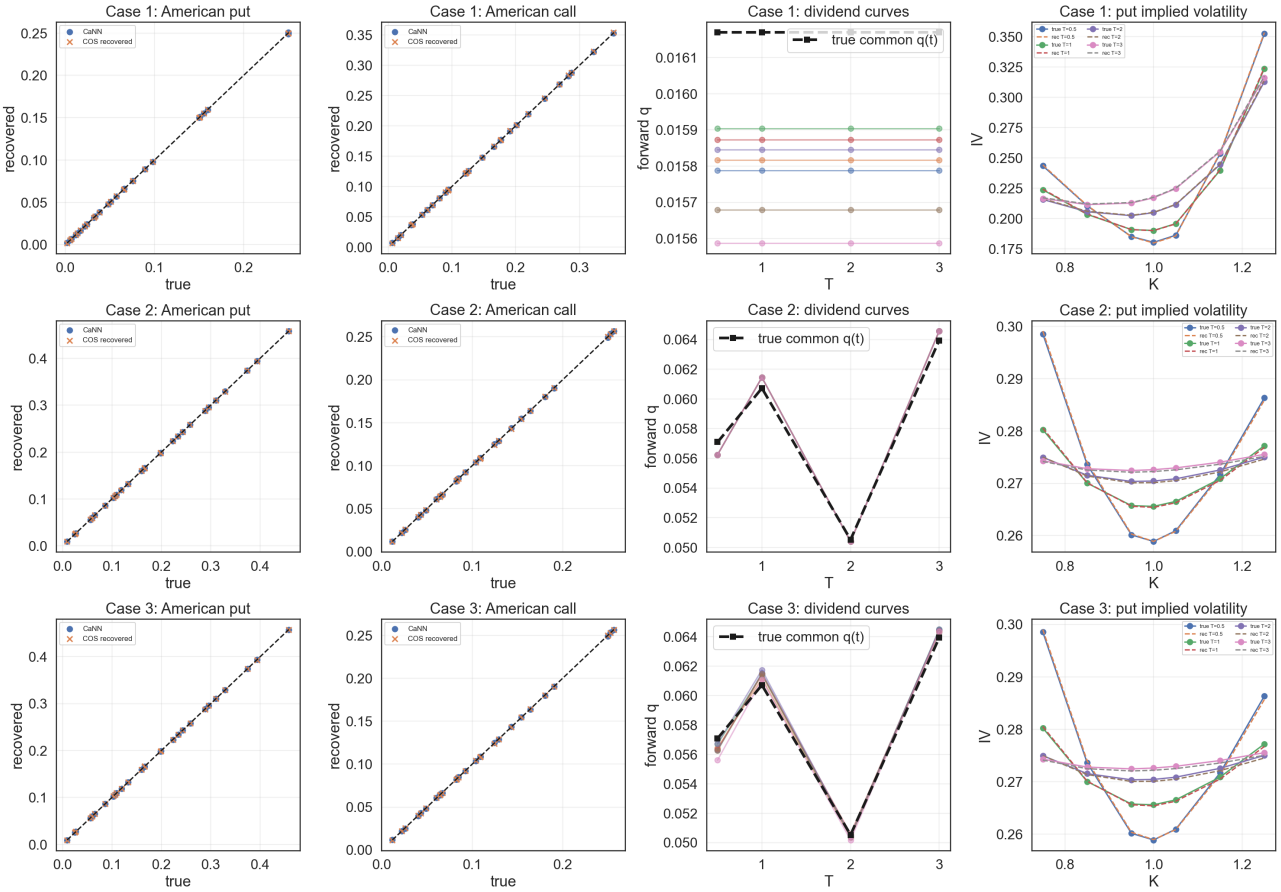


Figure 13: ATS recovery diagnostics for the three dividend specifications. From top to bottom, the rows show a constant true yield recovered freely by strike, a maturity-dependent dividend curve shared by all strikes, and a common true curve recovered with unrestricted strike dependence. Price panels compare ATS–COS targets with CaNN predictions and COS repricing. The remaining panels compare recovered dividend curves and put implied volatilities with their synthetic targets.

The final comparison below reports the parameter-recovery error for ATS together with the other models. It shows that the option prices and the implied-volatility shapes can be recovered accurately even when the five ATS parameters are less precisely recovered than the smaller Black–Scholes and NIG parameter sets.

The final comparison uses 300 previously unseen synthetic surfaces for each model. Each surface is generated and calibrated with the same model family: Black–Scholes surfaces are evaluated with Black–Scholes, NIG surfaces with NIG, and similarly for R–GBA and ATS. The comparison therefore does not test which model fits a surface generated by another model. It shows how accurately the same calibration procedure recovers the quantities of each correctly specified synthetic model.

The forward fit is measured on all held-out option rows. The inverse errors are then averaged over the 300 calibrated surfaces. Prices in table 11 are reported in the original synthetic scale, with $S_0 = 1$; the neural networks themselves are trained on normalized prices. Dividend MAE is reported in basis points and equivalent-IV MAE in percentage points.

The four models have parameters with different units and ranges. To compare their overall parameter recovery in one column, we use the mean normalized parameter error. For each non-dividend parameter p sampled in $[p_{\min}, p_{\max}]$, we calculate

$$\frac{|\widehat{p} - p|}{p_{\max} - p_{\min}}. \tag{26}$$

We then average this quantity over the model parameters and over the 300 surfaces. A value of 10%, for example, means that the average error is one tenth of the interval explored for that parameter in the synthetic dataset. This normalization is only used for the common comparison; the previous tables and figures retain the individual parameter errors.

Table 11: Mean results over 300 unseen synthetic surfaces for each model. The forward R^2 is the mean of the held-out put and call R^2 values. Dividend and implied-volatility errors are in basis points and percentage points, respectively. CaNN put and call MAEs are evaluated at the calibrated parameters and are reported in the original synthetic price scale with $S_0 = 1$.

Model	Mean forward R^2	q MAE (bp)	IV MAE (pp)	CaNN put MAE (price units)	CaNN call MAE (price units)	Mean parameter error (% of sampling range)	Time/surface (s)
Black–Scholes	0.999969	25.39	0.642	3.75×10^{-4}	3.74×10^{-4}	1.55	1.03
NIG	0.999963	23.18	0.314	5.09×10^{-4}	4.93×10^{-4}	2.25	0.84
R–GBA	0.999987	3.57	0.121	1.32×10^{-4}	1.49×10^{-4}	8.87	0.97
ATS	0.999975	5.10	0.081	1.68×10^{-4}	1.86×10^{-4}	8.50	0.76

All four models reconstruct the synthetic prices accurately. R–GBA has the lowest mean dividend error in this comparison, followed by ATS. ATS has the smallest equivalent-IV error. Black–Scholes and NIG have lower normalized structural errors, while the richer R–GBA and ATS vectors are harder to separate into individual parameters.

As a final numerical consistency check, we select one fixed synthetic surface for each model. The parameters recovered by CaNN–DE are inserted directly into the corresponding COS pricer, and the resulting prices are compared with the synthetic COS targets. The results are reported in table 12. These checks are separate from the 300-surface benchmark: they show directly that the calibrated parameters remain valid when they are evaluated again by the numerical pricer.

Table 12: Direct COS-repricing check on one fixed synthetic surface for each model. Put and call MAEs compare COS prices at the recovered parameters with the COS targets generated from the true parameters. Prices are in the original synthetic scale, with $S_0 = 1$.

Model	COS put MAE	COS call MAE
Black–Scholes	3.01×10^{-4}	3.77×10^{-4}
NIG	1.59×10^{-4}	2.90×10^{-4}
R–GBA	9.4×10^{-5}	1.43×10^{-4}
ATS	6.5×10^{-5}	1.47×10^{-4}

R–GBA and ATS confirm that the framework can be used beyond Black–Scholes and NIG. The trained neural network makes the calibration faster, while COS repricing verifies that the recovered parameters also work in the original American pricing model. The results also repeat an important pattern of the earlier experiments: option prices and implied-volatility shapes can be reconstructed accurately even when some internal parameters are less stable. For this reason, the strike-common maturity-dependent dividend curve remains the preferred specification, while fully unrestricted dividends are used to study the limits of the inverse problem.
The synthetic results therefore show how the procedure behaves when the true parameters are known and when the pricing model becomes more flexible. The next chapter applies the Black–Scholes and NIG models to AAPL market prices. There, the true parameters are unknown, so the main check is the pricing error on strikes that were not used in calibration.

4. Application to market data: AAPL options

This section applies the calibration framework to AAPL option data from Bloomberg. The aim is to examine how well the models reproduce observed option prices and to study the dividend values obtained from the calibration.
We first construct matched put–call pairs and compare the prices obtained from the Bloomberg reference fields with Bloomberg MID prices. We then test the calibration using prices generated directly by the pricing model from known parameters. Since these parameters are known, we can check whether the calibration recovers their original values. Finally, for each maturity, we calibrate the model using Bloomberg MID prices at a subset of strikes and evaluate the recovered parameters on strikes that were not used in calibration. This final test shows whether the dividend recovered from the calibration also improves the pricing of other options with the same maturity and helps us interpret its difference from the Bloomberg dividend field.
In the market application, spot, strike, maturity, MID price, financing rate, Bloomberg dividend, and Bloomberg implied volatility are observed data fields. Prices computed using the Bloomberg fields are therefore used as a

benchmark, not as prices generated from known model parameters.
The sample contains AAPL option quotes observed on 17 July 2026. The spot is $S_0 = 333.26$ USD and the dataset contains 14 expiries between 21 August 2026 and 15 December 2028. The calibration target is the Bloomberg MID price. The rate input is Bloomberg `OPT_FINANCE_RT`, used consistently in all the experiments of this section.
Bloomberg dividend yields and financing rates are reported as annual percentages and are divided by 100 before pricing. The dividend field ranges from 0.325% to 0.845%, while the financing rate ranges from 4.028% to 4.281%. In the pricing models, the Bloomberg dividend field is used as a continuous annual yield. This is a modelling approximation, since AAPL pays discrete cash dividends at specific ex-dividend dates.

Table 13: AAPL data flow. A pair is formed by one put and one call with the same strike and expiry. Pairing, elementary price checks, and the saved CaNN domain reduce 2,304 raw Bloomberg rows to 482 comparable put–call pairs.

Stage	Observations
Raw Bloomberg option rows	2,304
Rows with complete required fields	1,976
Complete put–call pairs	834
Pairs passing elementary price bounds	799
Pairs inside the saved CaNN domain	482
Expiries	14

Thirty-five pairs fail the elementary price checks, mainly because the put MID is below its intrinsic value. After these checks, 799 matched pairs remain. Restricting the sample to the domain of the neural networks retains 482 pairs, so that 317 pairs are excluded because at least one required input is outside the training domain, including $K/S_0 \in [0.6, 1.4]$ and $T \in [0.05, 3]$ years. The final sample contains 361 calibration pairs and 121 holdout pairs, spread over 14 maturities. It therefore retains several matched put–call quotes for each maturity while avoiding extrapolation outside the neural-network domain.
Figure 14 provides an overview of the observations retained after these checks. It shows the available strikes and maturities together with the MID prices and the Bloomberg volatility, financing-rate, and dividend fields used in the following comparisons.

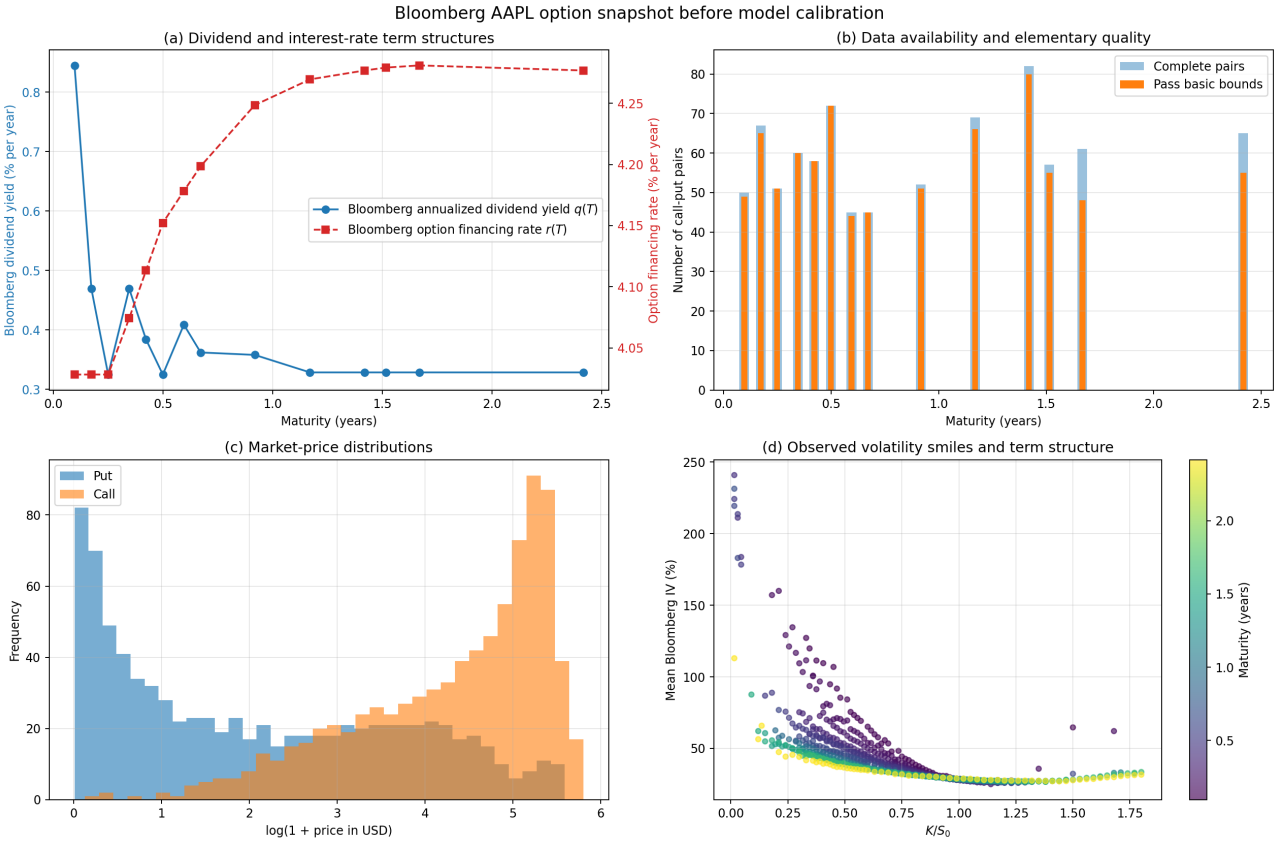


Figure 14: AAPL market data after preprocessing. The panels show strike coverage, MID prices, Bloomberg implied volatility, financing rates, and dividend yields. Rates and dividend yields are annual percentages.

The holdout is defined by strike rather than by individual option row. One observation of the inverse problem is a matched call–put pair. Holding out only the call while retaining the put with the same strike would reuse information from the same point of the smile. We therefore remove both options at every fourth sorted strike. This gives 361 calibration pairs and 121 holdout pairs, corresponding to 242 holdout option prices.
Holding out an entire maturity would answer a different question. The model is calibrated separately for each expiry, so an excluded maturity would have no local dividend or NIG parameters to estimate.
We first price the options by using the Bloomberg implied-volatility, dividend, and financing-rate fields as inputs. This is a reference-field benchmark: the resulting prices are compared with Bloomberg MID prices, but they are not generated from known true model parameters. Table 14 reports the errors on all 482 in-domain matched put–call pairs.

Table 14: Pricing errors obtained by inserting Bloomberg fields into American Black–Scholes pricers. MAEs are in USD and are measured against Bloomberg MID prices on the 482 in-domain matched put–call pairs.

Pricing method	Put MAE	Call MAE
COS	0.3715	0.9686
CRR	0.3666	0.9666
CaNN	0.4216	1.1082

COS and CRR give almost the same errors: their MAEs differ by less than one cent for both puts and calls. The remaining difference from market MID prices is therefore not mainly caused by the American numerical solver. On the holdout sample, the mean MID price is 45.21 USD, with values ranging from 0.105 USD to 145.875 USD. A one-dollar MAE corresponds to approximately 2.21% of the mean MID price. Thus, the initial discrepancies are moderate in price terms, but they can still affect the parameters obtained from calibration.
The calibration of each matched put–call pair separately makes this distinction clear. When a separate volatility and dividend are allowed for each matched quote, CaNN calibration almost interpolates the observed prices:

the put and call MAEs are 0.00295 USD and 0.00020 USD. This does not imply that the recovered inputs are equally reliable. Relative to the Bloomberg fields, the volatility MAE is 1.734 percentage points and the dividend MAE is 0.715 percentage points. The latter is large compared with the Bloomberg dividend-yield range of 0.325%–0.845%.

When the COS pricer is calibrated directly to each matched put–call pair, the conclusion is unchanged. The put and call MAEs are 0.00560 USD and 0.00411 USD, while the volatility and dividend MAEs are still 0.673 and 0.662 percentage points, respectively. Therefore, the difference in the recovered dividend is not mainly caused by the neural-network approximation. Over a limited set of option prices, different combinations of volatility and dividend can produce very similar prices.

We next consider a more restricted Black–Scholes specification. This preliminary exercise uses all 482 in-domain pairs and is intended to examine the pattern of the residuals; the holdout comparisons are introduced later. For each expiry, volatility is estimated at the observed strikes and is penalised when it changes too irregularly across strike. The dividend is described by one non-negative curve shared across all strikes,

$$q(T) = q_\infty + a\exp(-bT). \tag{27}$$

The three parameters (q_∞, a, b) are estimated from the MID prices without using the Bloomberg dividend as a prior.

The lowest calibration loss is obtained with $q_\infty = 0$ and $a = 0$, which are the lower bounds imposed on these two coefficients. Therefore, $q(T) = 0$ at every expiry in this particular specification; once $q_\infty = a = 0$, the value of b no longer affects the curve. This result does not mean that AAPL pays no dividends. It only means that, within this restricted Black–Scholes fit, the smallest pricing loss is obtained at the lowest admissible continuous dividend yield. The resulting put and call MAEs are 0.296 USD and 0.511 USD, respectively, and the mean call residual is −0.493 USD.

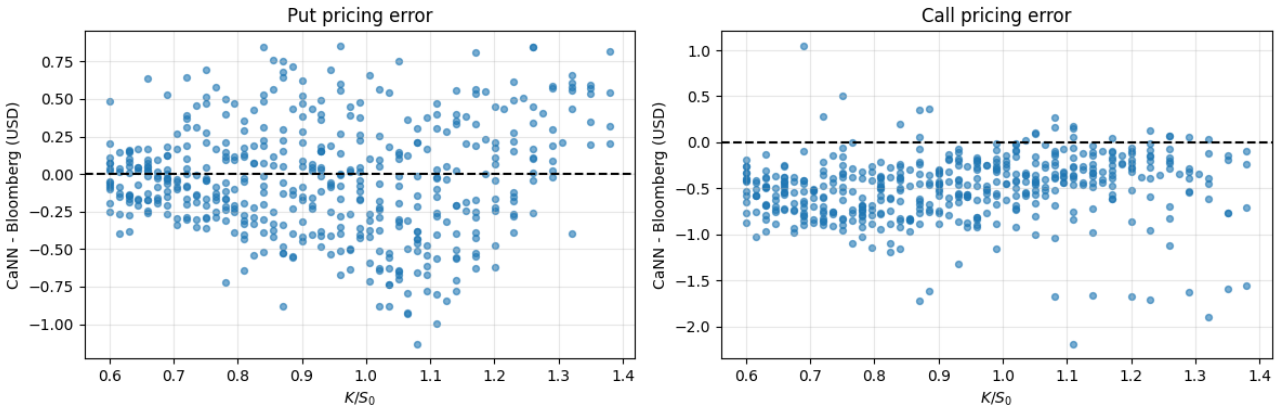


Figure 15: Residuals of the structured Black–Scholes calibration without a Bloomberg dividend prior. The horizontal axis is moneyness K/S_0 and the vertical axis is CaNN price minus Bloomberg MID, in USD. Volatility is smooth in strike and the dividend is one non-negative curve $q(T)$ shared across strikes. The predominantly negative call residuals reveal a systematic pattern that is not visible from the average error alone.

Using one dividend for all strikes removes the economically implausible strike-by-strike variation produced when each matched pair is fitted separately, but it does not force the recovered dividend to coincide with the Bloomberg field. A penalty centred on Bloomberg dividend values would make the estimate closer to those values, but it would introduce external information rather than recover the dividend from option prices alone. The contours in figure 16 show that volatility and dividend can partly compensate for one another.

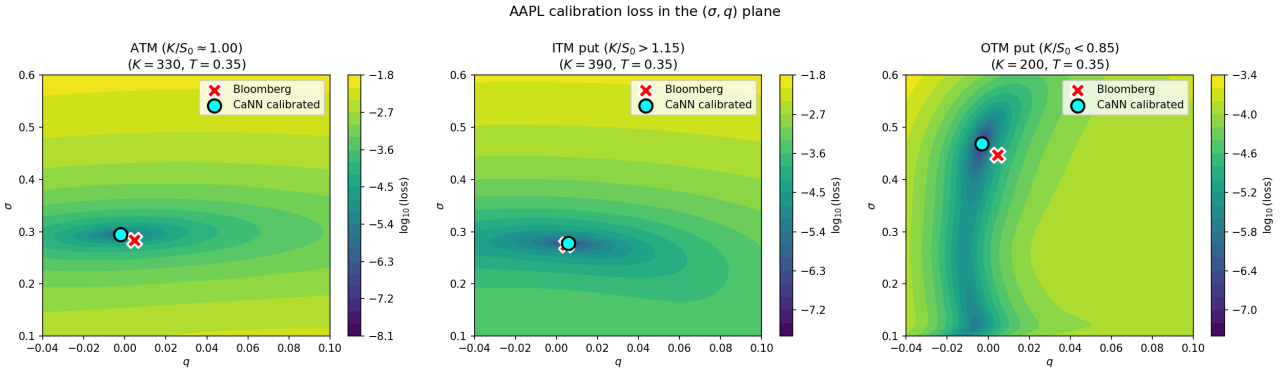


Figure 16: Black–Scholes loss surfaces in the volatility–dividend plane for three representative matched AAPL put–call pairs with $T = 0.35$: an at-the-money pair ($K = 330$), an in-the-money put ($K = 390$), and an out-of-the-money put ($K = 200$). Colours report the logarithm of the joint put–call pricing loss. The red cross is the Bloomberg reference point and the cyan circle is the CaNN-calibrated point. The broad low-loss regions show that, for an individual option pair, different volatility–dividend combinations can give similarly accurate prices.

The following experiments examine the dividend result in three steps. We first verify the inverse calibration using prices generated from known inputs. We then keep the volatility or NIG parameters fixed in order to isolate the effect of the dividend. Finally, we evaluate the recovered parameters on strikes that were not used in calibration. This sequence shows whether the dividend difference produces a pricing improvement beyond the observations used to estimate it.

We first verify that the inverse calibration can recover known inputs. For each maturity, the trained CaNN generates prices from a known dividend and a known set of model parameters. These inputs are then hidden from Differential Evolution and recovered using only the generated put and call prices. Since the target prices are produced by the same forward map used during calibration, this experiment isolates the accuracy of the inverse search.

In the Black–Scholes control, the generating dividend is the Bloomberg dividend and the remaining inputs are the Bloomberg implied-volatility fields. In the NIG control, the generating dividend is again the Bloomberg dividend, while the NIG vector is fixed at a calibrated reference value.

For Black–Scholes, the mean dividend error is below 10^{-3} basis points and the repricing MAE is below 10^{-5} USD. For NIG, 13 of the 14 maturities recover the generating dividend and NIG vector almost exactly. At the remaining maturity, the dividend error is approximately 1.5 basis points and the repricing MAE is 0.144 USD. These results confirm that the CaNN–Differential Evolution procedure can recover the generating inputs accurately in a controlled experiment. Together with the COS repricing checks, they support its use in the following market calibrations. The larger dividend difference obtained from Bloomberg MID prices is therefore not caused by an inability of the inverse procedure to recover known inputs.

Having checked the procedure on known inputs, we now use NIG directly to study whether a non-Gaussian return distribution improves the fit of the AAPL option prices.

For every expiry T_j, the NIG calibration estimates one vector

$$(\sigma_{\mathrm{NIG},j}, \eta_j, \kappa_j, q_j)\,, \tag{28}$$

where q_j is one dividend value shared by all strikes of that expiry. Both the NIG vector and the dividend are allowed to change across maturities. The financing rate remains the expiry-specific Bloomberg option financing rate.

The neural network evaluates one strike at a time: one input row returns the put and call prices for that strike. The calibration is nevertheless performed on the complete smile. At each expiry, Differential Evolution evaluates all calibration strikes together, using one common dividend and one common NIG vector. Thus, the procedure calibrates a maturity-by-maturity term structure without requiring the network itself to output a complete surface. Table 15 reports six representative maturities. The NIG parameters change when the dividend is estimated together with them, showing that the dividend and the shape of the return distribution affect the fit jointly. Since the true parameters are unknown for market data, the table is not used to measure parameter-recovery accuracy. It reports the estimates that produce the price comparisons presented next.

Table 15: Representative NIG estimates obtained from Bloomberg MID prices. The parameters in the middle block are calibrated with Bloomberg dividend fixed. The parameters in the right block are calibrated jointly with one implied dividend common to all strikes of the expiry. These are model-implied NIG estimates, not Bloomberg NIG parameters. The complete implied-dividend term structure is reported in table 19.

Expiry	q_{impl} (%)	$\sigma \mid q_{\text{BBG}}$	$\eta \mid q_{\text{BBG}}$	$\kappa \mid q_{\text{BBG}}$	$\sigma \mid q_{\text{impl}}$	$\eta \mid q_{\text{impl}}$	$\kappa \mid q_{\text{impl}}$
21 Aug 2026	−1.046	0.413	0.100	0.877	0.396	0.051	0.821
20 Nov 2026	−0.338	0.323	0.165	0.541	0.289	0.314	0.176
19 Mar 2027	−0.122	0.309	0.140	0.935	0.292	0.183	0.537
17 Sep 2027	−0.185	0.320	0.111	1.641	0.290	0.171	0.689
17 Mar 2028	−0.113	0.324	0.096	2.193	0.302	0.119	1.118
15 Dec 2028	−0.044	0.328	0.084	2.849	0.305	0.109	1.547

The initial holdout comparison includes the CaNN calibration, COS repricing of the neural solution, and direct NIG–COS calibration. Every error in table 16 is measured against Bloomberg MID on the 121 holdout strikes.

Table 16: AAPL deterministic strike-holdout MAE in USD. Direct NIG–COS reduces the call error substantially relative to the Black–Scholes reference benchmark, while the put error remains similar.

Method	Put MAE	Call MAE
Black–Scholes: Bloomberg IV and q	0.245845	1.225897
NIG CaNN: Bloomberg q fixed	0.470464	0.672056
NIG CaNN: free q calibrated	0.304551	0.336160
COS repricing of CaNN calibrated q result	0.289265	0.320353
Direct NIG–COS: free q calibrated	0.240505	0.268574

The agreement between the CaNN result and COS repricing shows that the neural calibration is also meaningful in the numerical NIG pricer. To compare the fitted distribution with market volatility, the NIG call curve is converted into a Black–Scholes-equivalent implied-volatility smile through equation (21). Direct NIG–COS has a holdout IV MAE of 0.938 percentage points, compared with 1.265 percentage points for CaNN–DE. The largest differences occur at the shortest maturity.
Figure 17 shows this comparison at representative maturities. The fitted NIG curves reproduce the main shape of the market smile, while the largest difference is visible at the shortest maturity, consistently with the larger IV error reported above.

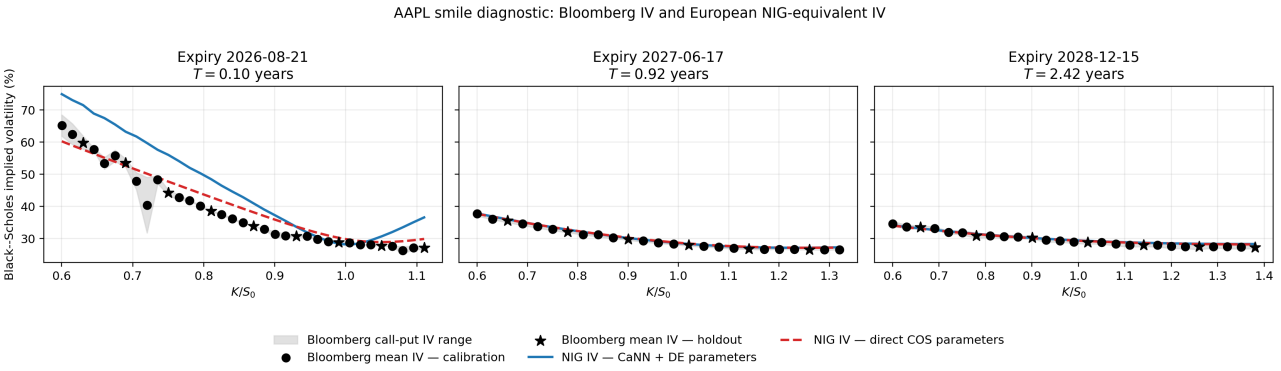


Figure 17: AAPL NIG-equivalent implied-volatility smiles at representative maturities. Black markers are the average of Bloomberg call and put implied volatilities, the grey band is their range, and the curves use CaNN–DE and direct NIG–COS estimates.

The previous results show that the dividend obtained from MID prices can differ from the Bloomberg dividend field. We now test whether this difference improves the pricing of strikes that were not used in calibration.
All calibrations in this part are carried out separately for each of the 14 expiries. Within one expiry, the model is fitted using only the calibration strikes. Every fourth strike is excluded together with its put and call, and these excluded pairs are used only for the final evaluation. The holdout test checks whether the parameters

estimated from some strikes can also reproduce the prices at the other strikes with the same expiry. All errors below are therefore measured against Bloomberg MID prices on the same 242 holdout option prices.
The first Black–Scholes test estimates one dividend yield for each expiry, while the Bloomberg put and call implied volatilities are kept fixed. Thus, the dividend is the only quantity changed between the Bloomberg-dividend baseline and the implied-dividend alternative. In the second Black–Scholes test, one volatility value is estimated at each calibration strike and is shared by the matched put and call. The values obtained at the different strikes are interpolated to form one smooth volatility smile for each maturity. The smile is estimated separately under the Bloomberg dividend and under the implied dividend.
For NIG, one vector $(\sigma_{\mathrm{NIG}}, \eta, \kappa)$ is fitted to all calibration strikes of an expiry. In the first comparison, this vector is kept fixed while only the dividend is changed. In the second comparison, the dividend and the NIG vector are estimated together. These tests distinguish the effect of changing the dividend from the effect of changing the volatility smile or the NIG parameters.

Table 17: Controlled AAPL price comparisons on 242 holdout option prices. For each expiry, all parameters are estimated using calibration strikes only. The mean Bloomberg MID is 45.21 USD. Relative MAE is the absolute MAE divided by this mean price. Confidence intervals are obtained by resampling the 14 expiries as whole groups.

Test	Bloomberg-q baseline		Implied-q alternative		MAE reduction, 95% CI
	MAE (USD)	Rel. (%)	MAE (USD)	Rel. (%)	USD
BS, IV fixed	0.773	1.71	0.503	1.11	0.269 [0.202, 0.343]
BS, q and smile calibrated	0.647	1.43	0.232	0.51	0.414 [0.262, 0.578]
NIG, vector fixed	0.631	1.40	0.245	0.54	0.386 [0.268, 0.508]
NIG, joint calibration	0.606	1.34	0.245	0.54	0.361 [0.249, 0.490]

The first and third rows are the most direct tests.
In Black–Scholes, changing only the dividend while keeping Bloomberg IV fixed reduces holdout MAE by 34.8%. In NIG, changing only the dividend while retaining the same NIG vector reduces MAE by 61.2%. Therefore, the improvement cannot be explained only by a new volatility or NIG-parameter calibration.
The common Black–Scholes smile gives a further improvement. When the dividend is fixed at the Bloomberg value, the combined MAE is 0.647 USD. When the implied dividend and the common smile are calibrated together, it falls to 0.232 USD, equal to approximately 0.51% of the mean MID price. The improvement is also stable across maturities. In all four comparisons in table 17, the lower endpoint of the 95% confidence interval for the MAE reduction is positive. Therefore, the result remains present when the 14 maturities are resampled and is not determined by only one maturity.

Table 18: Controlled combinations of dividend, volatility, and NIG inputs on the 242 holdout option prices. All MAEs are in USD against Bloomberg MID. In Black–Scholes, the calibrated smile is common to put and call prices. In NIG, the structural vector plays the role of the volatility specification.

Model	Dividend	Volatility or NIG-vector input	Combined MAE (USD)	Relative MAE (%)
BS	q_{BBG}	Separate Bloomberg put/call IV	0.773	1.71
BS	q_{BBG}	Common smile calibrated with q_{BBG}	0.647	1.43
BS	q_{impl}	Separate Bloomberg put/call IV	0.503	1.11
BS	q_{impl}	Common smile calibrated with q_{impl}	0.232	0.51
NIG	q_{BBG}	NIG vector calibrated with q_{BBG}	0.606	1.34
NIG	q_{BBG}	NIG vector calibrated with q_{impl}	0.631	1.40
NIG	q_{impl}	NIG vector calibrated with q_{BBG}	0.556	1.23
NIG	q_{impl}	NIG vector calibrated with q_{impl}	0.245	0.54

Table 18 shows that both components matter. In Black–Scholes, replacing the separate Bloomberg IV fields with a common calibrated smile improves the MAE at both dividend values. Replacing only the dividend also improves the error even when Bloomberg IV fields are unchanged.
The NIG comparison gives the same message in a different form. At Bloomberg dividend, changing the NIG vector alone gives similar errors, 0.606 and 0.631 USD. At implied dividend, the vector calibrated jointly with that dividend reduces MAE from 0.556 to 0.245 USD. Most importantly, keeping the NIG vector obtained with the implied dividend unchanged and changing only the dividend reduces MAE from 0.631 to 0.245 USD.
Taken together, these experiments provide the main market result of this section. The dividend recovered from the calibration improves the pricing of strikes that were not used to estimate it, in both Black–Scholes and NIG.

The improvement remains present when the other model inputs are kept fixed, showing that it is not produced only by a simultaneous change in volatility or in the NIG parameters.
Figure 18 shows the individual holdout errors behind the average values in the tables. With the implied dividend, the errors are more concentrated around zero, especially when the volatility smile or the NIG vector is estimated together with the dividend.

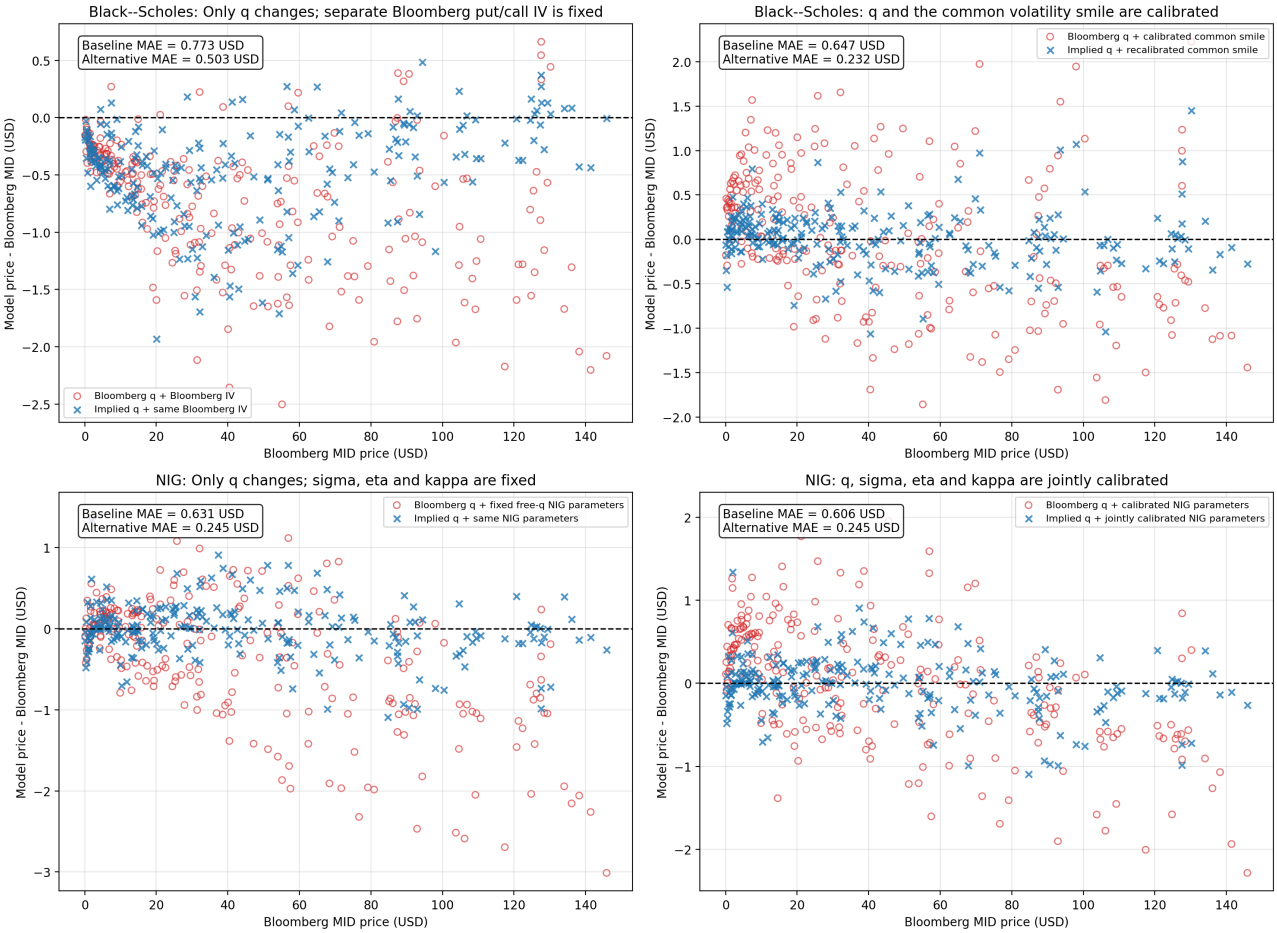


Figure 18: Signed holdout pricing errors for the controlled comparisons. The horizontal axis is Bloomberg MID and the vertical axis is model price minus MID. Blue crosses use the implied dividend and are more concentrated around zero than the Bloomberg-dividend baseline, especially when the volatility smile or NIG vector is calibrated jointly.

Figure 19 reports the Black–Scholes errors against moneyness after estimating one dividend and one common volatility smile for each maturity. It shows how the remaining errors are distributed across strikes rather than reporting only their average.

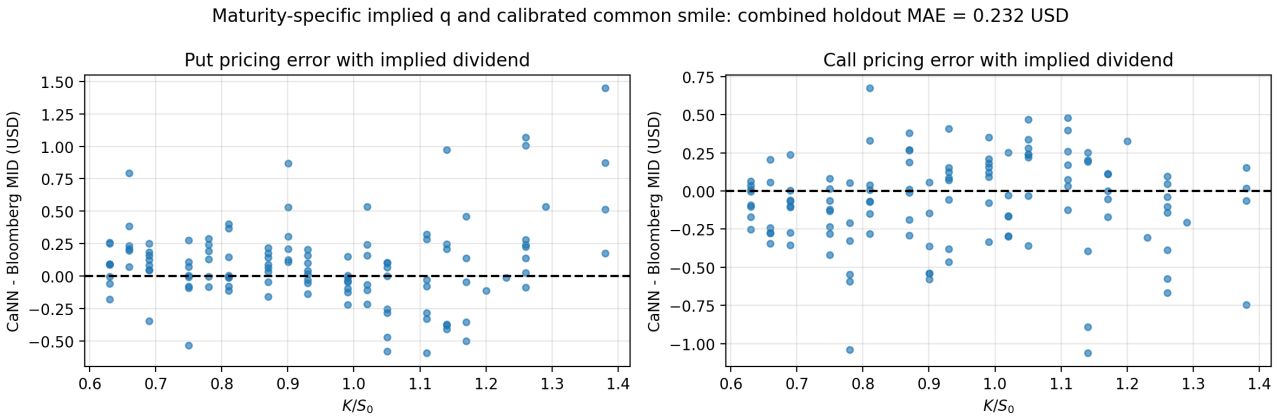


Figure 19: Black–Scholes holdout residuals by moneyness after introducing one implied dividend for each maturity and recalibrating the common volatility smile. The vertical axis is CaNN price minus Bloomberg MID.

The controlled comparisons show that replacing the Bloomberg dividend with the dividend recovered from option prices improves the holdout pricing errors. To describe this difference, we define

$$f(T) = q_{\mathrm{BBG}}(T) - q_{\mathrm{impl}}(T), \tag{29}$$

where $q_{\mathrm{BBG}}(T)$ is the Bloomberg dividend yield and $q_{\mathrm{impl}}(T)$ is the dividend obtained from calibration.
In the pricing models, the financing rate and the dividend yield enter the stock dynamics through the difference $r(T) - q(T)$, which we call the stock carry. Since the Bloomberg financing rate is kept fixed during calibration, the difference between $q_{\mathrm{BBG}}(T)$ and $q_{\mathrm{impl}}(T)$ also represents a difference in the carry required to reproduce the option prices.
Using the definition of $f(T)$, the carry obtained with the calibrated dividend can also be written as

$$r(T) - q_{\mathrm{impl}}(T) = r(T) + f(T) - q_{\mathrm{BBG}}(T).$$

This equivalent expression explains the possible funding interpretation of $f(T)$. If the Bloomberg dividend is retained, the same stock carry can be obtained by adding $f(T)$ to the financing component. A positive value of $f(T)$ means that the dividend recovered from option prices is lower than the Bloomberg dividend.
The numerical pricers do not actually replace $r(T)$ with $r(T) + f(T)$. They continue to use the Bloomberg financing rate and the calibrated dividend. The equality above is only a different way of expressing the same value of $r - q$.

Table 19: Model-implied AAPL carry adjustment by maturity. All entries are annual percentage points. The adjustment is $f = q_{\mathrm{BBG}} - q_{\mathrm{impl}}$. It expresses the difference between the two dividend yields as a possible adjustment to the financing component of the stock carry. It is not a directly observed funding cost.

Maturity	r_{BBG}	q_{BBG}	q_{BS}	f_{BS}	q_{NIG}	f_{NIG}
21 Aug 2026	4.028	0.845	-1.906	2.751	-1.046	1.891
18 Sep 2026	4.028	0.469	-0.474	0.943	-1.120	1.589
16 Oct 2026	4.028	0.325	-0.437	0.762	-0.802	1.127
20 Nov 2026	4.074	0.469	-0.361	0.830	-0.338	0.807
18 Dec 2026	4.113	0.384	-0.313	0.697	-0.315	0.699
15 Jan 2027	4.152	0.325	-0.134	0.459	-0.320	0.645
19 Feb 2027	4.178	0.409	-0.086	0.494	-0.091	0.500
19 Mar 2027	4.199	0.362	-0.073	0.435	-0.122	0.484
17 Jun 2027	4.249	0.358	-0.129	0.487	-0.135	0.493
17 Sep 2027	4.270	0.329	-0.122	0.450	-0.185	0.514
17 Dec 2027	4.277	0.329	-0.083	0.411	-0.153	0.481
21 Jan 2028	4.279	0.329	-0.060	0.388	-0.170	0.498
17 Mar 2028	4.281	0.329	0.045	0.284	-0.113	0.442
15 Dec 2028	4.277	0.329	0.034	0.295	-0.044	0.373

Across all maturities, the mean adjustment is 0.692 percentage points in Black–Scholes and 0.753 percentage points in NIG, while the corresponding medians are 0.473 and 0.507 percentage points. The larger means are mainly due to the annualized values at the shortest maturities, which are examined separately below.

For maturities of at least 0.25 years, the average adjustment is 0.475 percentage points in Black–Scholes and 0.540 percentage points in NIG. For maturities of at least 0.50 years, the corresponding values are 0.406 and 0.473 percentage points. The two models therefore give a consistent central result: beyond the shortest maturities, the implied adjustment is generally around 0.4–0.5 annual percentage points.

The shortest maturity requires a separate interpretation because all quantities in the table are annualized. Bloomberg reports an annual dividend yield, while AAPL pays a fixed cash amount on each dividend date. To compare the two quantities, we use $q_{\mathrm{BBG}} S_0 T$. For a small dividend yield, this provides an approximation of the total cash dividend expected before the option matures. In the present data, it is approximately 0.27 USD for the first three maturities, 0.54 USD for the following three, and 0.81 USD for the next two. These values correspond to one, two, and three quarterly dividends of 0.27 USD, respectively (Apple Inc., 2026).

These amounts are consistent with the Bloomberg field representing expected discrete dividends as an annualized yield, although its exact methodology is not observed.

At the first maturity, the implied adjustments are 2.75 percentage points in Black–Scholes and 1.89 percentage points in NIG. These annualized values look large because the maturity is only $T = 0.0959$ years. Over the actual option horizon, the corresponding adjustments fT are only 26.4 and 18.1 basis points. The first observation should therefore not be read as an annual borrowing cost. It is more sensitive to dividend timing, early exercise, and the limited information available in short-dated options. Figure 20 compares the annualized adjustments with the corresponding values over the actual option horizon. The right panel shows that the large first annualized values correspond to much smaller adjustments when they are measured only until maturity.

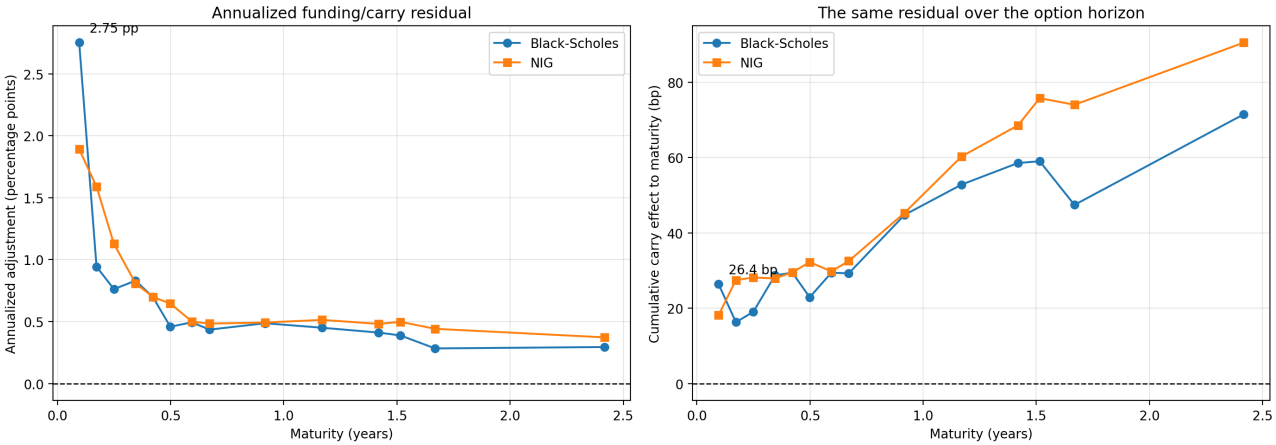


Figure 20: Short-end carry adjustment in annual and cumulative units. The left panel reports annualized values. The right panel reports the same adjustment accumulated only until expiry. The first BS and NIG observations correspond to 26.4 and 18.1 basis points, respectively, over the actual short horizon.

Figure 21 brings together the evidence behind this interpretation. The upper panels show the change in holdout pricing errors when only the dividend is replaced. The lower panels show the decomposition $q_{\mathrm{BBG}} = q_{\mathrm{impl}} + f$. The pricing improvement is the empirical result, while the decomposition follows directly from the definition of f.

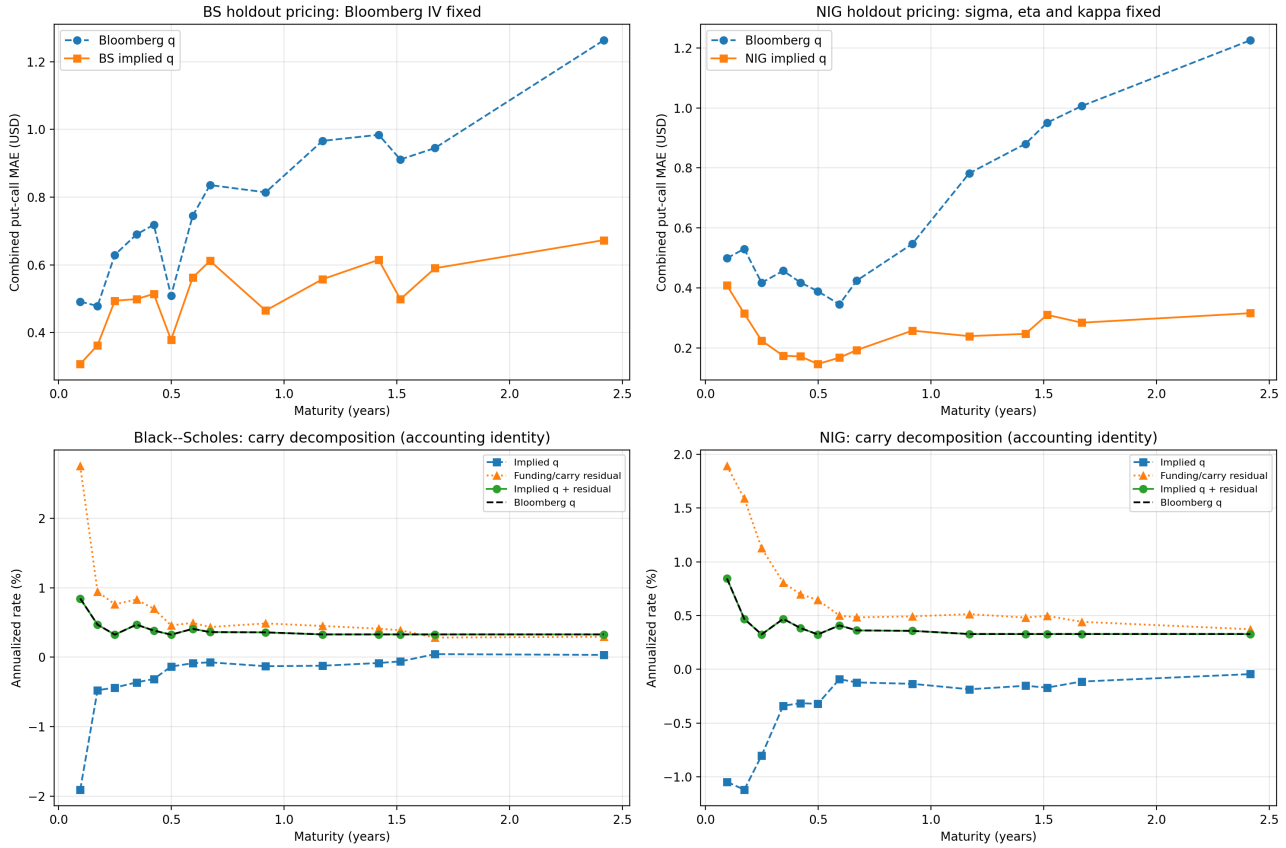


Figure 21: Pricing evidence and model-implied carry decomposition. The upper panels show holdout MAE by maturity when only the dividend is changed. The implied dividend improves the error in every BS and NIG maturity. The lower panels show q_{impl}, the carry adjustment f, and their sum. The equality $q_{\text{impl}} + f = q_{\text{BBG}}$ is an accounting identity; the independent evidence is the holdout-price improvement shown above and in table 17.

The controlled tests give a clear result. Using the Bloomberg dividend and financing fields directly does not fully reproduce the AAPL MID prices in either Black–Scholes or NIG. When the Bloomberg dividend is replaced by the dividend implied by the option prices, the errors on holdout strikes decrease. This also happens when the volatility input is kept fixed. The difference is therefore not explained only by changing volatility during calibration.

Because the Bloomberg financing rate is fixed during calibration, the difference in the carry required by the option prices is represented in the model by the recovered dividend. The equivalent expression above shows that the same difference can be written as an adjustment $f(T)$ to the financing component. For this reason, $f(T)$ is reported as a model-implied carry adjustment and may be interpreted as a possible cost-of-funding adjustment. Related studies also use option prices to examine funding-related quantities (Azzone and Baviera, 2021).

This interpretation does not mean that $f(T)$ is the true funding cost of Apple. Option prices provide information about the combined quantity $r - q$, but they do not identify separately the economic source of every difference. The result may also depend on the use of a continuous dividend yield and on the pricing model. Moreover, Black–Scholes and NIG give similar central values, but not identical estimates. We therefore interpret $f(T)$ as a possible model-implied cost-of-funding adjustment, rather than as a directly observed funding cost.

For the same reason, the negative values of $q_{\text{impl}}(T)$ in table 19 do not mean that Apple pays a negative cash dividend. They are the values of the continuous dividend input that allow each model to reproduce the carry implied by the option prices while the Bloomberg financing rate is kept fixed.

A further limitation concerns the implied-volatility fields reported by Bloomberg, which are calculated separately for calls and puts. In this thesis, a matched put and call with the same strike and maturity are instead calibrated jointly using one volatility smile and one dividend value. The Bloomberg implied volatilities are therefore used as market references, but not as two exact volatility inputs for the joint calibration.

This treatment is consistent with the evidence reported by Wallmeier (2024) for a different option database. The paper shows that separately reported call and put implied volatilities can differ because of the way in which they are calculated, even when the underlying prices are consistent with the relevant parity relation. This suggests that a difference between reported call and put implied volatilities does not necessarily imply that

two separate volatility smiles are required.
The AAPL application gives three main results.
First, direct NIG–COS calibration reduces the holdout call MAE from 1.226 USD in the Black–Scholes benchmark to 0.269 USD, while producing a similar put error. Its Black–Scholes-equivalent implied-volatility smile has a holdout MAE of 0.938 percentage points.
Second, both neural and direct numerical calibration show that small pricing errors can coexist with a dividend estimate that differs substantially from the Bloomberg dividend field.
Third, replacing the Bloomberg dividend with the implied dividend reduces the holdout pricing error in all the controlled Black–Scholes and NIG comparisons. This improvement remains present when the volatility input or the NIG vector is kept fixed.
Beyond the shortest maturity, Black–Scholes and NIG both place the implied adjustment at approximately 0.4–0.5 annual percentage points. The result does not mean that AAPL pays a negative cash dividend. Rather, within both American pricing models, the matched put–call MID prices imply a stock carry different from the one obtained by directly combining the Bloomberg financing rate and dividend field. The agreement between the two models, together with the pricing improvement on holdout strikes, supports the interpretation of this difference as a possible model-implied cost-of-funding adjustment. In this sense, the experiments show that matched American put and call prices can be used to recover an implied carry term structure.

5. Conclusions

The results of the thesis can be summarized in three main contributions.

1. *A validated and fast calibration procedure for American options.*
   Starting from the American Black–Scholes framework of Liu et al. (2021), we combined American COS pricing, a CRR benchmark, forward CaNNs, and Differential Evolution. The neural network is used only as a fast approximation of the pricing map during calibration; the financial model remains the COS pricer. The COS and CRR validation gives an implied-volatility MAE of 4.5×10^{-4} in decimal-volatility units (figure 2). For the main inverse experiments, the parameters obtained through CaNN–DE are inserted again into COS. This numerical repricing checks that the recovered parameters work in the original pricing model and not only inside the neural network.
   The NIG experiment with time-dependent dividends provides the clearest computational comparison. With the large CaNN–DE budget, the mean relative errors of the three NIG parameters are below 2.1%, the dividend MAE is 1.268×10^{-3}, and the average calibration time is 6.67 seconds per surface. The direct COS calibration in table 7 uses a deliberately smaller feasible budget. It requires 1093.23 seconds and gives larger recovery errors. This does not mean that COS is inaccurate. It shows that, after the offline dataset generation and training stage, the trained neural network makes the high-dimensional search computationally practical.
2. *A forward dividend curve shared across strikes.* The initial Black–Scholes benchmark follows the calibration setting of Liu et al. (2021), in which one option contract, or one matched put–call pair, is calibrated at a time. The implied dividend is therefore not imposed to be the same at different strikes. If options are calibrated separately, the procedure can therefore return different dividend estimates at different strikes. In this thesis, these strike-by-strike estimates are used only as a benchmark: they are not interpreted as different company dividends.
   One of the extensions developed in this thesis is the joint calibration of several strikes using one forward dividend curve $q(T)$ shared across strikes. This reflects the fact that one stock has one dividend policy, while still allowing the dividend yield to change across maturity intervals. The tests with a separate value $q(K, T)$ at every strike and maturity are retained as diagnostic benchmarks. They show what happens when the strike-common restriction is removed.
   Under Black–Scholes, the joint recovery produces dividend MAE 4.1×10^{-4} and accurate smile coefficients (table 4). The NIG tests show that the free dividend estimates are close across strikes when the NIG parameters are fixed (figure 10), but become less stable when the NIG parameters are also calibrated (figure 11). The R–GBA and ATS results in tables 9 and 10 lead to the same conclusion: allowing a separate dividend at every strike increases the number of unknowns without necessarily improving the recovery of the underlying dividend curve.
   The common comparison over 300 unseen surfaces per model gives dividend MAEs of 25.39, 23.18, 3.57, and 5.10 basis points for Black–Scholes, NIG, R–GBA, and ATS, respectively (table 11). Overall, the synthetic results support the use of a forward dividend curve shared across strikes as the preferred specification.
   The additional control experiments, in which a separate dividend is estimated at each strike and maturity, remain useful because they show directly how the results change when the restriction of one dividend curve shared across strikes is removed.

3. *A maturity-by-maturity NIG application to AAPL option prices.*
   For the AAPL dataset, we calibrate one NIG vector and one dividend common across strikes for each of 14 expiries. Representative NIG estimates obtained with Bloomberg dividend fixed and with free implied dividend are reported in table 15.
   On the deterministic strike holdout, direct NIG–COS calibration gives put and call MAEs of 0.2405 USD and 0.2686 USD, respectively. The corresponding Black–Scholes benchmark errors are 0.2458 USD and 1.2259 USD (table 16). The NIG model therefore improves the call fit materially. Its NIG-equivalent implied-volatility smile has holdout MAE 0.938 percentage points (figure 17).
   The recovered dividend differs substantially from the Bloomberg dividend field. This difference is economically relevant for pricing. When Bloomberg IV is kept fixed, replacing only the Bloomberg dividend by the implied dividend reduces Black–Scholes holdout MAE from 0.773 USD to 0.503 USD. When the NIG vector is kept fixed, the same dividend replacement reduces MAE from 0.631 USD to 0.245 USD (table 17). The improvement is also visible in the signed residuals in figure 18.
   We interpret the difference between Bloomberg and implied dividend as a model-implied carry adjustment. Over the more stable maturities, this adjustment is approximately 0.4–0.5 percentage points in both Black–Scholes and NIG (table 19 and figure 21). The large annualized value at the first expiry corresponds to only 26.4 bp in Black–Scholes and 18.1 bp in NIG over the actual option horizon (figure 20). The result does not mean that AAPL pays a negative dividend or that a unique corporate funding cost has been observed. It shows that, within the models considered here, the matched American put–call cross-section supports an effective carry different from the direct combination of Bloomberg financing and continuous-dividend fields.

The main limitation of the market application is that AAPL pays discrete cash dividends at known ex-dividend dates, whereas the pricing models use a continuous dividend yield. Market MID prices may also be affected by bid–ask spreads and by call, put, and stock quotes recorded at slightly different times. In addition, the implied volatilities reported by Bloomberg depend on the assumptions and numerical method used to calculate them from option prices.

Future work could include the known cash-dividend dates and amounts directly in the American pricing model, rather than approximating dividends through a continuous yield. It could also account for bid–ask spreads when comparing model prices with market quotes. Finally, repeating the calibration from different starting populations would help assess how stable the recovered parameters are.

List of symbols and abbreviations

Principal mathematical symbols used throughout the thesis.

Symbol	Meaning
S_t, S_0	Underlying price at time t and at the valuation date
K	Option strike
T, T_j	Time to maturity and maturity node, in years
$r(t)$, r_j	Continuously compounded financing rate and its value on interval j
$q(t)$, q_j	Continuous dividend yield and its forward value on interval j
$\bar{q}(T)$	Cumulative average dividend yield up to maturity T
q_{BBG}, q_{impl}	Bloomberg dividend field and dividend implied by option prices
$f(T)$	Model-implied carry adjustment, $f = q_{\mathrm{BBG}} - q_{\mathrm{impl}}$
σ	Black–Scholes volatility
σ_{IV}	Black–Scholes-equivalent implied volatility
$\sigma_{\mathrm{NIG}}, \eta, \kappa$	NIG scale, asymmetry, and activity parameters
$\alpha_0, \beta_0, \sigma_0, H_0$	R–GBA structural and maturity-scaling parameters
$\kappa_0, \eta_0, \beta, \delta$	ATS structural and maturity-scaling parameters
$\phi_Y(u;t)$	Characteristic function of a log-return increment
$g(S)$	Put or call payoff at stock price S
$C(t,S)$	Continuation value at time t and stock price S
V_M	Bermudan approximation with M exercise dates
N	Number of COS expansion terms
$\boldsymbol{x}$	Fixed option inputs, such as moneyness, maturity, and rate
$\boldsymbol{p}$	Parameter vector optimized during calibration
$\boldsymbol{y}$	Synthetic or observed target put–call prices
f_θ	Trained neural pricing approximation with weights θ
$J(\boldsymbol{p})$	Calibration objective minimized by Differential Evolution
$\mathbb{E}^{\mathbb{Q}}$	Expectation under the risk-neutral measure

Abbreviations: AAPL: Apple Inc. equity ticker;
ATS: Additive Normal Tempered Stable model;
CaNN: Calibration Neural Network;
COS: Fourier–cosine pricing method;
CRR: Cox–Ross–Rubinstein binomial tree;
DE: Differential Evolution;
IV: implied volatility;
LHS: Latin Hypercube Sampling;
MAE/RMSE: mean absolute error/root mean squared error;
MID: Bloomberg mid option price;
NIG: Normal Inverse Gaussian model;
R–GBA: reduced-form Generalized Beta Additive model;
TDQ: causal time-dependent dividend-yield representation.

References

Apple Inc. Dividend history. Apple Investor Relations, 2026. URL `https://investor.apple.com/dividend-history/default.aspx`. Accessed 16 August 2026.

Michele Azzone and Roberto Baviera. Synthetic forwards and cost of funding in the equity derivative market. *Finance Research Letters*, 41:101841, 2021. doi: 10.1016/j.frl.2020.101841.

Michele Azzone and Roberto Baviera. Additive Normal Tempered Stable processes for equity derivatives and power-law scaling. *Quantitative Finance*, 22(3):501–518, 2022. doi: 10.1080/14697688.2021.1983200.

Michele Azzone and Lorenzo Torricelli. Explicit option pricing with additive processes. *SIAM Journal on Financial Mathematics*, 16(3):747–802, 2025. doi: 10.1137/23M1621241.

Ole E. Barndorff-Nielsen. Normal Inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics*, 24(1):1–13, 1997. doi: 10.1111/1467-9469.00045.

Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973. doi: 10.1086/260062.

John C. Cox, Stephen A. Ross, and Mark Rubinstein. Option pricing: A simplified approach. *Journal of Financial Economics*, 7(3):229–263, 1979. doi: 10.1016/0304-405X(79)90015-1.

Fang Fang and Cornelis W. Oosterlee. A novel pricing method for European options based on Fourier–Cosine series expansions. *SIAM Journal on Scientific Computing*, 31(2):826–848, 2008. doi: 10.1137/080718061.

Fang Fang and Cornelis W. Oosterlee. Pricing early-exercise and discrete barrier options by Fourier–Cosine series expansions. *Numerische Mathematik*, 114(1):27–62, 2009. doi: 10.1007/s00211-009-0252-4.

Blanka Horvath, Aitor Muguruza, and Mehdi Tomas. Deep learning volatility: A deep neural network perspective on pricing and calibration in (rough) volatility models. *Quantitative Finance*, 21(1):11–27, 2021. doi: 10.1080/14697688.2020.1817974.

James M. Hutchinson, Andrew W. Lo, and Tomaso Poggio. A nonparametric approach to pricing and hedging derivative securities via learning networks. *The Journal of Finance*, 49(3):851–889, 1994. doi: 10.1111/j.1540-6261.1994.tb00081.x.

Michael Kohler, Adam Krzyzak, and Nebojsa Todorovic. Pricing of high-dimensional American options by neural networks. *Mathematical Finance*, 20(3):383–410, 2010. doi: 10.1111/j.1467-9965.2010.00404.x.

Shuaiqiang Liu, Anastasia Borovykh, Lech A. Grzelak, and Cornelis W. Oosterlee. A neural network-based framework for financial model calibration. *Journal of Mathematics in Industry*, 9:9, 2019a. doi: 10.1186/s13362-019-0066-7. Article 9.

Shuaiqiang Liu, Cornelis W. Oosterlee, and Sander M. Bohte. Pricing options and computing implied volatilities using neural networks. *Risks*, 7(1):16, 2019b. doi: 10.3390/risks7010016. Article 16.

Shuaiqiang Liu, Alvaro Leitao, Anastasia Borovykh, and Cornelis W. Oosterlee. On a neural network to extract implied information from American options. *Applied Mathematical Finance*, 28(5):449–475, 2021. doi: 10.1080/1350486X.2022.2097099. Published online in 2022.

Michael D. McKay, Richard J. Beckman, and William J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979. doi: 10.1080/00401706.1979.10489755.

Robert C. Merton. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1):141–183, 1973. doi: 10.2307/3003143.

Rainer Storn and Kenneth Price. Differential Evolution—a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11:341–359, 1997. doi: 10.1023/A:1008202821328.

Martin Wallmeier. Quality issues of implied volatilities of index and stock options in the OptionMetrics IvyDB database. *Journal of Futures Markets*, 44(5):854–875, 2024. doi: 10.1002/fut.22495.

Kai Zhang and Kok Lay Teo. A penalty-based method for reconstructing a smooth local volatility surface from American options. *Journal of Industrial and Management Optimization*, 11(2):631–644, 2015. doi: 10.3934/jimo.2015.11.631.