

No-Arbitrage and CPT Consistency in Generative Option Pricing Models: A Systematic Literature Investigation

Wen-Chieh Chiu

*Department of English, National Changhua University of Education, Changhua, China
921118sophie@gmail.com*

Abstract. Generative artificial intelligence models are increasingly used for option pricing and market simulation. However, the mathematical consistency of these model outputs and their alignment with real investor behavior remain unclear. Current research on generative pricing models follows two unrelated paths: one verifies mathematical consistency in no-arbitrage, while the other examines whether investor behavior conforms to cumulative prospect theory (CPT). However, no research has yet tested both criteria simultaneously within the same generative framework. This paper, through a systematic literature review, identifies this research gap. First, this paper decomposes the no-arbitrage assumption into four testable criteria: put-call parity, strike price convexity, intertemporal spread consistency, and martingale consistency. Then, it decomposes the consistency of cumulative prospect theory into three testable criteria: the shape of the probability-weighted function, two-tailed pricing bias, and the decomposition of probability-weighted and loss-averse factors. These seven criteria were then applied to eight generative pricing papers covering four model architectures (Generative Adversarial Networks, Neural Stochastic Differential Equations, Diffusion Models, and Variational Autoencoders). A review of the results revealed that convexity and intertemporal spread conditions were tested most frequently, martingale consistency was tested in only three papers, and put-call parity and all three CPT criteria were completely unaddressed in all eight papers. This consistent gap indicates that current generative pricing research narrows arbitrage compliance to a static, curved surface condition without verifying the authenticity of investor behavior, which has practical implications for the reliability of these models in hedging and risk management.

Keywords: Generative artificial intelligence, no-arbitrage consistency, cumulative prospect theory, option pricing, implied volatility surface

1. Introduction

Even if there is no chance for arbitrage, an option pricing model may still deviate from the pricing logic of the real market. The no-arbitrage assumption ensures that there are no loopholes in the market price that can be easily arbitrated, which is a basic principle that option pricing models must follow [1, 2]. However, this principle only answers whether prices are self-consistent, but does not answer whether such prices reflect how real investors view risk and probability. In recent years, generative artificial intelligence technology has developed rapidly, and researchers have begun to

use generative adversarial networks, diffusion models, variational autoencoders and neural stochastic differential equations to construct option pricing models and market simulation tools. With their strong data fitting capabilities, these models have begun to be applied to practical hedging and risk management. Once their pricing logic is out of sync with the behavior of real investors, the prices and hedging positions output by the model may show systematic deviations when the market fluctuates violently, and such deviations are often not noticed until extreme market conditions occur.

The current research on generative pricing models has developed along two separate lines. The first line focuses on whether the model satisfies the no-arbitrage condition. Researchers translate traditional financial theory principles such as call-put parity, convexity condition and martingale consistency into operable test items for the model and embed them into the model architecture or training process [3]. The test criteria of this line are clear and easy to quantify, so it has almost become a standard feature of generative pricing model papers. However, its essence is still a pure mathematical consistency check. Even if the model passes this test, it cannot prove that it reflects the risk perception of real investors. The second line focuses on the field of behavioral finance. It explains empirical phenomena such as option returns and volatility premiums through the probability weighting mechanism of cumulative prospect theory, proving that no-arbitrage restrictions and investor behavioral biases can coexist in a hierarchical structure [4]. This line deeply depicts the systematic distortion of investors' subjective cognition, but it has always remained within the framework of traditional parametric models and has never been used to test the new generation of generative pricing models. Two separate lines of research have accumulated decades of findings within their respective journal communities, yet they have almost never cited each other, forming a peculiar division of labor. Researchers on one side focus solely on price consistency, while those on the other focus on investor rationality; no one simultaneously questions whether these two questions can be answered affirmatively within the same model. Currently, no research systematically examines whether generative pricing models simultaneously meet the criteria of no-arbitrage and consistency with cumulative prospect theory. This means that the credibility of generative pricing tools, considered cutting-edge technology and gradually replacing traditional models, has never actually been verified at the behavioral level.

This paper presents this research gap through a systematic literature review. Methodologically, this paper first organizes the no-arbitrage criterion and the consistency criterion with cumulative prospect theory into actionable test items. Then, it examines existing generative pricing model literature using these two sets of criteria one by one, thereby illustrating the theoretical and practical implications of this gap for the credibility of generative pricing models. The chapters of this paper are arranged as follows: Section 2 summarizes the no-arbitrage criterion; Section 3 summarizes the cumulative prospect theory consistency criterion; Section 4 is the core of this paper, which first briefly describes the basic concepts of generative artificial intelligence, then reviews the existing generative pricing literature in groups according to model type, and cross-examines each of the criteria in Sections 2 and 3, and finally analyzes the current research status, limitations and future directions in the discussion section; Section 5 summarizes the existence of the research gap and its implications, and puts forward suggestions for future research.

2. No-arbitrage criteria

The no-arbitrage assumption is the fundamental principle of option pricing theory. Black and Scholes first proved that if there are no arbitrage opportunities in the market, the option price can be uniquely determined by a replica portfolio constructed from the underlying asset and the risk-free asset [1]. Harrison and Kreps further generalized this result, proving that there is an equivalence

relationship between no-arbitrage and the existence of an equivalent martingale measure, which is the first fundamental theorem of asset pricing [2]. This theorem states that as long as a measure Q can be found such that the discounted price of tradable assets is a martingale under this measure, there are no arbitrage opportunities in the market. The following breaks down this theorem into four directly verifiable conditions, and proposes corresponding auditing methods for each condition, as the criterion for cross-examining generative pricing models in Section 4.

2.1. Call-put parity

Under the assumption of continuous dividends and interest rates, call-put parity requires:

$$C - P = S_0 e^{qT} - K e^{-rT} \quad (1)$$

Where C and P are the call and put prices, respectively, S_0 is the current price of the underlying asset, K is the strike price, q is the dividend rate, r is the risk-free rate, and T is the time to maturity. This relationship is a direct inference of the no-arbitrage assumption between call and put options [1, 2].

To verify this criterion, for each set of call and put prices output by the generative model, calculate the difference between the left and right sides to obtain the parity residual. Take the sample mean and standard deviation of the residuals of all generated samples to construct a confidence interval and check whether the residuals significantly deviate from zero. If they significantly deviate, it means that there is a risk-free arbitrage opportunity in that set of prices, and it is judged as a violation of this criterion.

2.2. Convexity of strike price direction

The option price with respect to the strike price must be a convex function. Equivalently, the implied probability density must be non-negative: $\partial^2 C / \partial K^2 \geq 0$. This result is derived from the fact that the risk-neutral density can be reduced from the second derivative of the option price with respect to the strike price [5]. When expressed in terms of implied volatility $\sigma(K)$, this condition can be written as:

$$\sigma''(K) + 2[\sigma'(K)]^2 / \sigma(K) + 2\sigma'(K) / K + d_1 d_2 / [K \sigma(K)] \geq 0 \quad (2)$$

To verify this criterion, for the implied volatility surface output by the model, estimate $\sigma'(K)$ and $\sigma''(K)$ along the strike price direction at a fixed expiration date using numerical differentiation, and check whether the above formula holds true one by one. Divide the number of execution price points that violate this inequality by the total number of test points to obtain the convexity violation rate; the higher the violation rate, the more often the surface generated by the model implies a negative probability density, that is, there is a butterfly spread arbitrage space.

2.3. Calendar spread conditions across expiry dates

Under a fixed-price condition (moneyness) $m = K/F_t$, the total variance $w(m, T) = \sigma^2(m, T) \cdot T$ must be a non-decreasing function of the expiry date T :

$$w(m, T_2) \geq w(m, T_1), T_2 > T_1 \quad (3)$$

This condition stems from the requirement that volatility surfaces must avoid calendar spread arbitrage [6].

To verify this criterion, for the implied volatility output by the model at different expiry dates, after fixing the price, plot the curve of the total variance against the expiry date, and compare whether a decrease occurs between adjacent expiry dates. Divide the number of comparisons showing a decrease by the total number of comparisons to obtain the calendar spread violation rate, which serves as a quantitative indicator.

2.4. Martingale consistency

The discounted option price must satisfy the martingale property under the equivalent measure Q :

$$D_t(T)C_t(T, K) = E^{Q[D_T(T)C_T(T, K)|\mathcal{F}_t]} \quad (4)$$

This property is directly derived from the definition of the equivalent martingale measure [2].

To verify this criterion, this condition involves the time dynamics of prices and is difficult to verify directly from cross-sectional data at a single point in time, as the previous three conditions can be. The period-by-period increment of the discounted price can be regressed against information lagging one period to check whether the regression coefficient is significantly different from zero; or the Generalized Method of Moments can be used to check whether the conditional expectation of the discounted price increment is zero. If the test significantly rejects the null hypothesis, it indicates that there is a predictable systematic bias in the price dynamics, violating martingale consistency.

The above four conditions and corresponding verification methods together constitute the list of no-arbitrage criteria used in this paper for cross-examining generative pricing models. The common advantage of these criteria is that they can all be written as closed-form equations or inequalities, making it easy to directly check them against the model output one by one. Taking VolGAN as an example, although the arbitrage penalty is not directly embedded in the generator's loss function, the model instead relies on a smoothness penalty during training combined with post-hoc scenario re-weighting to keep the arbitrage penalty of generated surfaces close to that of market data, showing that such criteria can be enforced through mechanisms other than direct loss-function embedding [7].

However, this set of criteria can only confirm whether prices are mathematically consistent internally, without addressing whether prices reflect the subjective perception of real investors when facing probabilities and losses. In other words, a model that passes the no-arbitrage test may still differ from the actual pricing behavior of investors, which is the problem that will be addressed in Section 3.

3. CPT consistency criteria

The core of the CPT consistency criterion lies in examining whether the probability allocation implicit in a pricing model reflects the systematic distortion of objective probability by investors, rather than simply examining their risk aversion [4]. Specifically, it can be summarized into the following three operational test items.

3.1. Probability weighting function shape

CPT assumes that investors modify the objective probability p with the decision weighting function $w(p)$, rather than directly using p itself. A common setting is

$$w(p) = e^{[-(-\ln p)^\gamma]}, \quad 0 < \gamma < 1 \quad (5)$$

When $\gamma < 1$, the function exhibits an inverse S-shape: low probabilities are overestimated and medium-to-high probabilities are underestimated; when $\gamma \rightarrow 1$, the function degenerates into a 45-degree straight line, indicating that the model does not have probability distortion, which is equivalent to the rational expectation hypothesis [8]. This test requires that the value of γ derived from the implicit allocation of the model must be significantly less than 1, and its estimate must be statistically significant, rather than just noise within the estimation error range [8, 9].

3.2. Two-tailed pricing bias

If probability weighting does exist, it should simultaneously lead to an underestimation of the payouts for both out-of-the-money put and call options. This test requires a joint check on the model output payout for call options R_{OTM}^{call} and payout for put options R_{OTM}^{put} :

$$H0: E[R_{OTM}^{call}] \geq 0 \quad \text{or} \quad E[R_{OTM}^{put}] \geq 0 \quad \text{vs.} \quad H1: E[R_{OTM}^{call}] < 0 \quad \text{and} \quad E[R_{OTM}^{put}] < 0 \quad (6)$$

If only one tail is significantly negative, while the other tail is insignificant or even positive, it is insufficient to support the probability-weighted hypothesis, and other sources of explanation must be considered [4].

3.3. Decomposition of probability weighting and loss aversion

The behavioral pricing bias of the model can be decomposed into probability-weighted components and loss aversion components. Using the value:

$$\text{function } v(x) = \begin{cases} x^\alpha, & x \geq 0 \\ -\lambda(-x)^\beta, & x < 0 \end{cases} \quad (7)$$

The decision weights $w(p)$ together determines the prospect value. The test method involves fixing one parameter at a rational baseline ($\gamma=1$ when probability weighting is off; $\lambda=1$ when loss aversion is off) and comparing the explanatory power of the model for volatility premiums or option returns under the two settings [4]. If the model claims to conform to CPT consistency, its explanatory power for abnormal pricing should mainly come from deviations in γ rather than λ ; conversely, if removing probability weighting has almost no impact on the model's explanatory power, then the model may simply be introducing a general risk aversion mechanism, rather than truly conforming to CPT consistency [4]. The above three criteria are all statistical tests at the allocation level. They are not easy to write directly into a single closed-form inequality and embed into the loss function, as is the no-arbitrage criterion. Moreover, their estimation results are more sensitive to the estimation method itself. For example, when using the shape of the pricing kernel as auxiliary evidence, attention should be paid to the consistency of non-parametric estimation methods

to avoid misjudging estimation bias as behavioral signals [10]. These three criteria will serve as the basis for testing when reviewing the Batch B literature one by one in Section 4.

4. Literature review and cross-examination of generative pricing models

4.1. Preliminaries of generative AI

Generative artificial intelligence refers to modeling techniques that learn from the probability distribution of data and generate similar new samples. In recent years, it has been extended to the pricing of derivative products to generate implied volatility surfaces or simulate market paths. This paper reviews models covering four architectures: Generative Adversarial Networks (GANs), Neural Stochastic Differential Equations (Neural SDEs), diffusion models, and Variational Autoencoders (VAEs). These four categories will be reviewed in subsequent subsections.

4.2. Generative adversarial networks (GANs)

VolGAN [7] does not embed the arbitrage penalty into the generator's loss function; the authors report no gain from doing so. Arbitrage control for (II) and (III) instead relies on post-hoc scenario re-weighting using out-of-sample penalty values. However, the authors state that the goal is to make the penalty value lower than or equal to the level of the original market data, rather than to completely eliminate it; during periods of sharp market fluctuations, turning off this penalty term actually improves accuracy, showing that there is a trade-off between arbitrage control and data fitting. Criterion (I) call parity is only used to back-calculate the risk-free rate, which is a preprocessing step, not a post-test. Criterion (IV) only tests the stability of the correlation coefficient between implied volatility and returns, which is different from the definition of martingale properties in this paper. The three CPT terms are not mentioned.

GAN-Enhanced Surface Reconstruction [11] defines two violation rate indicators (corresponding to calendar spread and convexity conditions, respectively), which are independent post-test quantitative measures. GAN-Enhanced Surface Reconstruction [11] keeps calendar violation rates below 0.8% throughout, including March 2020—unlike the SVI baseline it compares against, which rose to 5.8% in that period. Criteria (I) and (IV) were not tested, and the three CPT criteria were not mentioned.

Summary: VolGAN relies on post-hoc re-weighting rather than training-time constraints, while GAN-Enhanced maintains low violation rates even under stress; (I), (IV), and all three CPT criteria remain untested in both.

4.3. Neural SDE class

Arbitrage-Free Neural-SDE Market Models [12] directly embed the absence of arbitrage restrictions into the model architecture in the form of inequalities: the monotonicity of the expiration date corresponds to (III), the convexity of the strike price corresponds to (II), and a drift condition corresponds to the martingale property (IV). All three are built in before training and are tested separately after training. Criterion (I) put-call parity was not addressed. The three CPT criteria were not mentioned. Estimating risks of European option books [13] follows the framework of [12] and provides two additional ex post-hoc quantitative indicators, corresponding to dynamic and static arbitrage violation rates, respectively. The empirical values are all below 3%, which is rare in

literature that provides independent quantitative verification for (II), (III), and (IV) at the same time. Criterion (I) is still not addressed, and the three CPT criteria are not mentioned.

Summary: Both papers incorporate the no-arbitrage constraint into the model design, and [13] further provides actual violation rate figures. However, neither paper addresses the put-call parity, and all three CPT criteria are not mentioned, presenting the same gap as GAN-type literature.

4.4. Diffusion model

Generative Diffusion Model for Risk-Neutral Derivative Pricing [14] This model uses a measure transformation to adjust the inverse denoising process of the trained diffusion model, ensuring that the discounting process of the generated price path satisfies the martingale property, corresponding to criterion (IV). It is verified using a geometric Brownian motion scenario, and the generated European option price closely matches the Black-Scholes theoretical value, making it a rare case where criterion (IV) has been clearly validated post-hoc. Criteria (I), (II), and (III) are not addressed, and the three CPT items are not mentioned.

A diffusion-based generative model for financial time series via GBM [15] This model embeds geometric Brownian motion into the denoising mechanism of the diffusion model, aiming to reproduce statistical characteristics such as heavy tails and volatility clustering, rather than testing the no-arbitrage condition. Although based on Black-Scholes theory, the entire paper does not embed or test criteria (I), (II), (III), and (IV) in any way. The three CPT items are not mentioned.

Summary: In the Diffusion category, only [14] provides explicit post-hoc verification for criterion (IV). The other criteria and the three CPT items are generally not addressed, indicating that this category of literature focuses more on statistical fidelity than financial compliance verification.

4.5. VAE category

Arbitrage-Free Implied Volatility Surface Generation with VAE [16] First, the historical surface is projected onto the parameter space of a known arbitrage-free stochastic model (regime-switching models and Lévy processes). Then, a VAE is trained in this space and sampled to generate the output surface, ensuring arbitrage-free operation through model selection. This is built into the architecture and implicitly covers criteria (II) and (III), but does not provide independent violation rate figures. Criterion (I) is not addressed, and the three CPT items are not mentioned.

Variational Autoencoders for Completing the Volatility Surfaces [17] uses synthetic Heston surfaces to train a VAE and performs probability distribution analysis and testing for calendar spread and butterfly spread conditions. This is an independent post-validation, corresponding to criteria (II) and (III). Criteria (I) and (IV) are not addressed, and the three CPT items are not mentioned.

Summary: Both VAE-related papers address criteria (II) and (III), [16] using built-in guarantees for model selection and [17] providing post-validation procedures, but neither touches upon criteria (I) and (IV), and the three CPT items are not mentioned at all, showing a consistent gap with GAN and Neural SDE-related papers.

4.6. Discussion and analysis

Based on the review of the eight papers in Sections 4.2 to 4.5, the treatment of the four no-arbitrage criteria is not balanced. (II) Convexity and (III) Calendar price difference are the two most frequently treated criteria. Six out of the eight papers address at least one of them through built-in architecture or ex-post quantification. Some papers (GAN-Enhanced, Estimating risks, VAE surface

completion) even provide specific violation rate data. (IV) Martingale consistency appears in three papers: two in the Neural SDE category ([12, 13]) and one in the Diffusion category ([14]). The verification methods are different. Neural SDE is mainly achieved through built-in architecture, [13] which further provides ex-post quantification verification. The Diffusion category directly provides ex-post verification. (I) Call and put parity is a common gap in all eight papers. None of them address this, which shows that existing generative pricing research generally narrows the no-arbitrage constraint to the surface shape condition and ignores the parity relationship across commodity types. The three behavioral criteria of CPT are not mentioned at all in eight articles, which is the core evidence for the gap argument in this paper. The only exception related to the concept of CPT is [18], which touches on the concept of non-monotonicity of pricing core, which is related to probability weighting, but does not transform it into an operational CPT test, remaining only at the conceptual level.

Overall, existing literature focuses on the verification of static conditions (convexity, monotonicity) and deals less with criteria involving time dynamics (martingale consistency, cross-commodity parity), showing that the verification of generative pricing models is still limited to the compliance of surface shape at a single point in time, rather than consistency across time and across commodities. The future audit framework should cover the complete combination of the four no-arbitrage criteria, rather than just convexity and calendar price difference; it should also try to incorporate behavioral criteria such as probability weighting of CPT or two-tailed pricing bias into the training objectives or ex-post verification procedures, rather than just mentioning them at the conceptual level as in [18]. The lack of this dual audit framework is precisely the research gap that this paper hopes to present.

5. Conclusion

This paper, through a systematic literature review, points out a research gap in generative pricing model literature: no-arbitrage consistency and CPT behavioral consistency have never been simultaneously tested in the same model. This paper breaks down the no-arbitrage assumption into four criteria and CPT consistency into three criteria. After cross-examining eight generative pricing papers covering GANs, Neural SDEs, Diffusion, and VAEs, it finds that convexity and calendar spreads are most frequently addressed; martingale consistency is found in three papers; and call-put parity and CPT are completely unaddressed. This gap indicates that existing research narrows "no-arbitrage" to the condition of curved surface shape, ignoring whether investor behavioral biases are internalized by the model, which has practical implications for the credibility of generative tools for hedging and risk management.

References

- [1] Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3), 637–654.
- [2] Harrison, J. M., & Kreps, D. M. (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20(3), 381–408.
- [3] Huang, Y. (2026). From no-arbitrage to foundation models: A review of financial-mathematics-constrained generative AI and deep learning for quantitative finance. SSRN Working Paper No. 6699179. <https://ssrn.com/abstract=6699179>
- [4] Baele, L., Driessen, J., Ebert, S., Londono, J. M., & Spalt, O. G. (2019). Cumulative prospect theory, option returns, and the variance premium. *Review of Financial Studies*, 32(9), 3667–3723.
- [5] Breeden, D. T., & Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices. *Journal of Business*, 51(4), 621–651.

- [6] Gatheral, J., & Jacquier, A. (2014). Arbitrage-free SVI volatility surfaces. *Quantitative Finance*, 14(1), 59–71.
- [7] Vuletić, M., & Cont, R. (2024). VolGAN: A generative model for arbitrage-free implied volatility surfaces. *Applied Mathematical Finance*, 31(4), 203–238.
- [8] Sun, Y.-F., & Peng, W.-Y. (2024). Behavioral option pricing based on prospect theory framework and the Heston model. *Journal of Guangdong University of Technology*, 41(1), 127–134.
- [9] Nardon, M., & Pianca, P. (2019). European option pricing under cumulative prospect theory with constant relative sensitivity probability weighting functions. *Computational Management Science*, 16, 249–274.
- [10] Linn, M., Shive, S., & Shumway, T. (2017). Pricing kernel monotonicity and conditional information. SSRN Working Paper No. 2383527. <https://ssrn.com/abstract=2383527>
- [11] Ge, Y., Wang, Y., Liu, J., & Wang, J. (2025). GAN-enhanced implied volatility surface reconstruction for option pricing error mitigation. *IEEE Access*, 13, 176765–176780.
- [12] Cohen, S. N., Reisinger, C., & Wang, S. (2023). Arbitrage-free neural-SDE market models. *Applied Mathematical Finance*, 30(1), 1–46.
- [13] Cohen, S. N., Reisinger, C., & Wang, S. (2022). Estimating risks of European option books using neural stochastic differential equation market models. *Journal of Computational Finance*, 26(3), 33–72.
- [14] Tiwari, N. (2026). Generative diffusion model for risk-neutral derivative pricing. arXiv:2603.20582.
- [15] Kim, G., Choi, S.-Y., & Kim, Y. (2025). A diffusion-based generative model for financial time series via geometric Brownian motion. arXiv:2507.19003.
- [16] Ning, B. X., Jaimungal, S., Zhang, X., & Bergeron, M. (2023). Arbitrage-free implied volatility surface generation with variational autoencoders. *SIAM Journal on Financial Mathematics*, 14(4), 1004–1027.
- [17] Feugang Nteumagné, B., Azemtsa Donfack, H., & Wafo Soh, C. (2025). Variational autoencoders for completing the volatility surfaces. *Journal of Risk and Financial Management*, 18(5), Article 239.
- [18] Zhang, J., Ma, C., & Schoutens, W. (2026). Forward-looking physical tail risk: A deep learning approach. Working paper, KU Leuven & Northwest University.