

# CALL OPTION PRICE USING PEARSON DIFFUSION PROCESSES

TAPAN KAR, SUPRIO BHAR, BARUN SARKAR, AND SESHA MEKA

**ABSTRACT.** Following the foundational work of the Black–Scholes model, extensive research has been developed to price the option by addressing its underlying assumptions and associated pricing biases. This study introduces a novel framework for pricing European call options by modeling the underlying asset’s return dynamics using Pearson diffusion processes, characterized by a linear drift and a quadratic squared diffusion coefficient. This class of diffusion processes offers a key advantage in its ability to capture the skewness and excess kurtosis of the return distribution, well-documented empirical features of financial returns. We also establish the validity of the risk-neutral measure by verifying the Novikov condition, thereby ensuring that the model does not admit arbitrage opportunities. Further, we study the existence of a unique strong solution of stock prices under the risk-neutral measure. We apply the proposed method to Nifty 50 index option data and conduct a comparative analysis against the classical Black–Scholes and Heston stochastic volatility models. Results indicate that our method shows superior performance compared to the other two benchmark models. These results carry practical implications for market participants, including market makers, hedge funds, and derivative traders.

## 1. INTRODUCTION

Black and Scholes [6] and Merton [23] developed option pricing models based on the assumption that the underlying asset price follows a geometric Brownian motion, implying that its log returns are normally distributed.

---

2020 *Mathematics Subject Classification.* 91B25, 91G20, 60H10, 60H30.

*Key words and phrases.* Call option, Pearson diffusion, Pearson type IV distribution, Novikov condition.

However, empirical evidence shows that log returns deviate from normality, displaying skewness and excess kurtosis. Consequently, the option prices estimated by these models often deviate from observed market prices.

Researchers have proposed various alternative option pricing models to address the limitations of the Black-Scholes framework. An alternative to the Black-Scholes model is the constant elasticity of variance (CEV) approach, which generalizes geometric Brownian motion by allowing the diffusion coefficient to follow a power-law relationship with the stock price. This approach has been adopted by several researchers, including Cox [11, 12], Cox and Ross [13], and Beckers [4], among others. Merton [24] extended the Black-Scholes model by introducing a jump-diffusion framework incorporating a jump component into the asset price dynamics. Heston [17] introduced the stochastic volatility model in which volatility itself is governed by a separate stochastic process. Madan *et al.* [22] proposed the Variance Gamma option pricing model, driven by a stochastic process known as the Variance Gamma process, which is constructed by evaluating Brownian motion with drift at a random time given by a Gamma process. Cervellera and Tucci [8] discussed challenges associated with parameter estimation for the Variance Gamma process. Borland [7] proposed an option pricing model based on a diffusion process described by a stochastic differential equation whose noise was driven by a non-Gaussian distribution, and the resulting distribution of log return process over time can be explicitly characterized by the Tsallis distribution [28]. However, Vellekoop and Nieuwenhuis [29] later argued that Borland's model does not satisfy the Novikov condition, a key requirement for risk-neutral pricing, thereby allowing the possibility of arbitrage. Additionally, Karagozoglu [20] provided a good literature review on option pricing models.

In option pricing, researchers aim to model log returns using distributions that can capture the excess kurtosis and skewness of the return process. The Pearson Type IV distribution is well-suited for this task, as it accommodates

a broad range of kurtosis and skewness depending on its shape parameter values. Nagahara [25] provided a comprehensive analysis of this distribution, deriving its normalizing constant, the first four moments, and illustrating its applicability to modeling stock return distributions. Bhattacharyya *et al.* [5] employed the Pearson Type IV distribution to forecast Value at Risk for leptokurtic equity index returns. Further, Forman and Sørensen [16] investigated the statistical properties of Pearson diffusion processes, a class of diffusion characterized by linear drift and a quadratic squared diffusion coefficient and showed that under specific conditions on the diffusion coefficient, the process follows Pearson Type IV distribution as its invariant distribution.

The main contribution of this paper is introducing a novel method for estimating call option prices by modeling log returns using Pearson diffusion processes, henceforth we refer to this method as the PIV model. In addition, we establish the validity of the Novikov condition within this framework, crucial for risk-neutral pricing, and study the existence of a unique strong solution for the stock price process under the risk-neutral measure. We apply the proposed method to Nifty 50 index option data and perform a comparative analysis against the classical Black–Scholes and Heston stochastic volatility models. The results demonstrate that our approach outperforms both benchmark models.

The rest of the paper is organized in the following way. Section 2 describes the Pearson Type IV distribution. Section 3 presents the Pearson diffusion processes. Section 4 outlines the modeling of log return processes. In Section 5, we derive the risk-neutral measure and present the proposed option pricing model. Section 6 provides empirical evidence, and Section 7 concludes the paper.

## 2. PEARSON TYPE IV DISTRIBUTION

Karl Pearson [27] was the first to develop a generalized family of frequency curves designed to accommodate a wide range of empirical distributions encountered in practice. This comprehensive framework is now recognized as the Pearson system of curves. The Pearson system of curves satisfy the differential equation defined as:

$$(1) \quad \frac{1}{f(x)} \frac{df(x)}{dx} = \frac{x - \alpha}{c_0 + c_1x + c_2x^2}$$

A general solution of the above differential equation is given by

$$(2) \quad f(x) = k \exp \left( \int \frac{x - \alpha}{c_0 + c_1x + c_2x^2} dx \right),$$

where  $k$  is a constant.

The explicit form of the function  $f(x)$  is governed by the integral in the exponent of (2). The evaluation of this integral, in turn, depends on the nature of the roots of the equation  $c_0 + c_1x + c_2x^2 = 0$ . The Pearson Type IV curve is obtained when the roots are imaginary, or when  $c_1^2 < 4c_0c_2$ .

The probability density function (PDF) of Pearson's Type IV distribution is given by

$$(3) \quad f(x) = k \left[ 1 + \left( \frac{x - \lambda}{a} \right)^2 \right]^{-m} \exp \left[ \nu \tan^{-1} \left( \frac{x - \lambda}{a} \right) \right] \quad -\infty < x < \infty,$$

where  $m > \frac{1}{2}$ ,  $a > 0$ , and  $m$ ,  $a$ ,  $\nu$  and  $\lambda$  are real-valued parameters (functions of  $\alpha$ ,  $c_0$ ,  $c_1$ , and  $c_2$ ) and  $k$  is a normalization constant that depends on  $m$ ,  $a$ , and  $\nu$ . Nagahara [25] derived a closed-form expression of the normalization constant  $k$ .

The parameters  $\nu$  and  $m$  control skewness and kurtosis, respectively. Accordingly, the distribution is positively (negatively) skewed as  $\nu$  is negative (positive). Increasing  $m$  reduces the kurtosis. The scale parameter is  $a$ , and

the location parameter is  $\lambda$ . The Pearson Type IV distribution can capture a wide range of shapes depending on its parameter values.

### 3. PEARSON DIFFUSION PROCESSES

Forman and Sørensen [16] studied the statistical properties of the Pearson diffusion processes, which are defined by having a linear drift and a quadratic squared diffusion coefficient.

Let  $\mathbb{P}$  be a probability measure on a sample space  $\Omega$ . Pearson diffusion is a stationary solution to the following stochastic differential equation (SDE):

$$(4) \quad dR_t = -\theta(R_t - \mu) dt + \sqrt{2\theta(aR_t^2 + bR_t + c)} dB_t$$

where  $\theta > 0$  and  $a$ ,  $b$ , and  $c$  are such that the square root term is well defined. Here  $B_t$ ,  $t \geq 0$  is a Brownian motion and  $\mu$  denotes the mean of the invariant distribution.

As a particular case, when  $a > 0$ ,  $b = 0$ , and  $a = c$ , the SDE (4) simplifies to the form:

$$(5) \quad dR_t = -\theta(R_t - \mu) dt + \sqrt{2\theta a(1 + R_t^2)} dB_t$$

Forman and Sørensen [16] stated that when  $\mu = 0$ , the invariant distribution of the SDE (5) is a scaled  $t$ -distribution with  $m_1 = 1 + a^{-1}$  degrees of freedom and a scale parameter  $m_1^{-1/2}$ . In the case where  $\mu \neq 0$ , the invariant distribution corresponds to the Pearson Type IV distribution, with  $\mu$  controlling the skewness and the tails decaying at the same rate as those of a  $t$ -distribution with  $1 + a^{-1}$  degrees of freedom. In both cases, the invariant distribution has moments of order  $k_1$  for  $k_1 < 1 + a^{-1}$ .

### 4. MODEL OF RETURNS: EXISTENCE AND UNIQUENESS OF SOLUTION FOR THE LOG RETURN PROCESS

Let the stock price dynamics be given by,

$$S_t := S_0 e^{R_t}, \quad \forall t \geq 0,$$

with  $S_0$  being the initial stock price and  $R_t, t \geq 0$  be the log return process.

Note that  $R_t = \ln \frac{S_t}{S_0}$  and  $R_0 = 0$ .

Let  $R_t, t \geq 0$  follow Pearson diffusion process:

$$(6) \quad dR_t = -\theta(R_t - \mu) dt + \sigma \sqrt{2\theta a(1 + R_t^2)} dB_t, \quad t \geq 0$$

where  $\sigma$  denotes the volatility of log returns. We use the parameter  $\sigma$  for scaling the invariant distribution.

Next, we study the existence of strong global solution of (6). For that we have to show the drift and diffusion coefficients are globally Lipschitz [26, Theorem 5.2.1]. The drift and diffusion coefficients are given by

$$u(x) := -\theta(x - \mu), \quad v(x) := \sigma \sqrt{2\theta a(1 + x^2)}, \quad \forall x \in \mathbb{R},$$

respectively. Note that,

$$|u(x) - u(y)| = \theta|x - y|, \quad \forall x, y \in \mathbb{R}.$$

Again,

$$\begin{aligned} |v(x) - v(y)| &= \sigma \sqrt{2\theta a} \left| \sqrt{1 + x^2} - \sqrt{1 + y^2} \right| \\ &= \sigma \sqrt{2\theta a} \frac{|x^2 - y^2|}{\left| \sqrt{1 + x^2} + \sqrt{1 + y^2} \right|} \\ &\leq \frac{\sigma \sqrt{2\theta a} |x + y|}{\sqrt{1 + x^2} + \sqrt{1 + y^2}} |x - y| \\ &\leq \sigma \sqrt{2\theta a} |x - y|, \quad \forall x, y \in \mathbb{R}. \end{aligned}$$

The last inequality follows from the estimate

$$|x + y| \leq |x| + |y| \leq \sqrt{1 + x^2} + \sqrt{1 + y^2}.$$

Therefore there exists an unique strong solution of (6). In particular by [26, Theorem 5.2.1],

$$(7) \quad \mathbb{P} \left( \int_0^T R_t^2 dt < \infty \right) = 1.$$

We now derive the SDE governing the stock price using the Itô's formula. Consider the function  $f(x) = e^x, x \in \mathbb{R}$ . Then, we have, a.s.,

$$\begin{aligned} dS_t &= S_0 d(e^{R_t}) \\ &= S_t \{ -\theta(R_t - \mu) + \sigma^2 \theta a (1 + R_t^2) \} dt + S_t \sigma \sqrt{2\theta a (1 + R_t^2)} dB_t. \end{aligned}$$

## 5. EQUIVALENT LOCAL MARTINGALE MEASURE

Consider a European-style call option with strike price  $K$ , time to maturity  $T$ , written on an underlying asset  $S_t$ ,  $0 \leq t \leq T$ . The payoff from the European call option at maturity is  $\max(S_T - K, 0)$ . At a time  $t < T$ , using the information  $I_t$  available up to time  $t$  (more precisely, a filtration  $\{I_t\}_{t \leq T}$ ) the fair market price of the call option denoted by  $C(t)$  is equal to the discounted (under the risk-free interest rate  $r$ ) expected payoff, and is given as:

$$(8) \quad C(t) = e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}[\max(S_T - K, 0) \mid I_t],$$

where  $\mathbb{E}^{\mathbb{Q}}$  denotes expectation under the risk-neutral probability measure  $\mathbb{Q}$ . The market completeness ensures the uniqueness of the risk-neutral measure (see [10]). The equality stated in (8) holds provided that the discounted asset price process  $e^{-rt} S_t$ ,  $t \leq T$  is a martingale under  $\mathbb{Q}$ . In this case,  $\mathbb{Q}$  is called an equivalent martingale measure. The probability measure  $\mathbb{Q}$  is called an equivalent local martingale measure if  $e^{-rt} S_t$ ,  $t \leq T$  becomes a local martingale under  $\mathbb{Q}$  (see [26, Remark, page 165]).

According to the asset pricing theory, to compute the expectation under the risk-neutral measure, we first convert  $e^{-rt} S_t$ ,  $t \leq T$  into a local martingale under the risk-neutral probability measure  $\mathbb{Q}$ . To identify  $\mathbb{Q}$ , we use the celebrated Girsanov's Theorem.

**Theorem 5.1** (Girsanov's Theorem). [26, Theorem 8.6.6] *Let  $Y_t, t \leq T$  be a real valued Itô process of the form*

$$dY_t = \beta(t) dt + \theta(t) dB_t; t \leq T.$$

*Suppose there exist processes  $u_t, t \leq T$  and  $\alpha_t, t \leq T$  such that*

$$\theta_t u_t = \beta_t - \alpha_t$$

*and assume that  $u_t, t \leq T$  satisfies the Novikov's condition*

$$\mathbb{E} \left[ \exp \left( \frac{1}{2} \int_0^T u_s^2 ds \right) \right] < \infty.$$

*Put*

$$M_t := \exp \left( - \int_0^t u_s dB_s - \frac{1}{2} \int_0^t u_s^2 ds \right), t \leq T$$

*and  $d\mathbb{Q}(\omega) = M_T(\omega) d\mathbb{P}(\omega)$  on  $\mathcal{F}_T$ . Then,*

$$\widehat{B}_t := \int_0^t u_s ds + B_t, t \leq T$$

*is a Brownian motion with respect to  $\mathbb{Q}$  and in terms of  $\widehat{B}_t, t \leq T$ , the process  $Y_t, t \leq T$  has the stochastic integral representation*

$$dY_t = \alpha(t) dt + \theta(t) d\widehat{B}_t; t \leq T.$$

**Remark 5.2.** *In Theorem 5.1, we require  $M_t, t \leq T$  to be a  $\mathbb{P}$  martingale, specifically an exponential martingale, see [19, Chapter III, Theorem 5.3].*

We now derive the SDE for the discounted price process  $e^{-rt} S_t, t \leq T$  under the original probability measure  $\mathbb{P}$ . Consider the function  $f(t, x) := e^{-rt} x, t \geq 0, x \in \mathbb{R}$ . Then,

$$\frac{\partial f}{\partial t}(t, x) = -r e^{-rt} x, \quad \frac{\partial f}{\partial x}(t, x) = e^{-rt}, \quad \frac{\partial^2 f}{\partial x^2}(t, x) = 0, \quad \forall t, x.$$

Therefore,

$$(9) \quad \begin{aligned} d(e^{-rt}S_t) &= -re^{-rt}S_t dt + e^{-rt}S_t \{-\theta(R_t - \mu) + \sigma^2\theta a(1 + R_t^2)\} dt \\ &+ e^{-rt}S_t \sigma \sqrt{2\theta a(1 + R_t^2)} dB_t \end{aligned}$$

Girsanov's Theorem 5.1 proposed a new Brownian motion  $\widehat{B}_t, t \leq T$  under  $\mathbb{Q}$  is of the form

$$d\widehat{B}_t = u_t dt + dB_t, t \leq T.$$

Therefore, (9) can be written as:

$$(10) \quad \begin{aligned} d(e^{-rt}S_t) &= -re^{-rt}S_t dt + e^{-rt}S_t \{-\theta(R_t - \mu) + \sigma^2\theta a(1 + R_t^2)\} dt \\ &+ e^{-rt}S_t \sigma \sqrt{2\theta a(1 + R_t^2)} dB_t \\ &= -re^{-rt}S_t dt + e^{-rt}S_t \{-\theta(R_t - \mu) + \sigma^2\theta a(1 + R_t^2)\} dt \\ &+ e^{-rt}S_t \sigma \sqrt{2\theta a(1 + R_t^2)} (d\widehat{B}_t - u_t dt) \end{aligned}$$

To ensure that the discounted price process  $e^{-rt}S_t, t \leq T$  becomes a local martingale under  $\mathbb{Q}$ , we set the drift term of (10) to zero and obtain:

$$(11) \quad u_t = \frac{-r - \theta(R_t - \mu) + \sigma^2\theta a(1 + R_t^2)}{\sigma \sqrt{2\theta a(1 + R_t^2)}}.$$

**Theorem 5.3.** *There exists a equivalent local martingale measure  $\mathbb{Q}$  such that such that under  $\mathbb{Q}$ , the discounted price process  $e^{-rt}S_t, t \leq T$  becomes a local martingale.*

*Proof.* For the proof, see Appendix A. □

**5.1. Black-Scholes type PDE.** Let  $F : [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$  be  $C^{1,2}$  regular. Here,  $C^{1,2}$  denotes the space of two variable functions, once continuously differentiable in the first variable and twice continuously differentiable in the second variable. For any derivative price at time  $t$  given by  $F(t, S_t)$ , by Itô's formula,

$$dF(t, S_t)$$

$$\begin{aligned}
&= \frac{\partial F}{\partial t}(t, S_t) dt + \frac{\partial F}{\partial S}(t, S_t) dS_t + \frac{1}{2} \frac{\partial^2 F}{\partial S^2}(t, S_t) d[S, S]_t \\
&= \frac{\partial F}{\partial t}(t, S_t) dt \\
&\quad + \frac{\partial F}{\partial S}(t, S_t) \left\{ S_t \left\{ -\theta(R_t - \mu) + \sigma^2 \theta a (1 + R_t^2) \right\} dt + S_t \sigma \sqrt{2\theta a(1 + R_t^2)} dB_t \right\} \\
&\quad + \frac{1}{2} \frac{\partial^2 F}{\partial S^2}(t, S_t) \sigma^2 S_t^2 (2\theta a(1 + R_t^2)) dt \\
&= \left\{ \frac{\partial F}{\partial t}(t, S_t) + \frac{\partial F}{\partial S}(t, S_t) [S_t \left\{ -\theta(R_t - \mu) + \sigma^2 \theta a (1 + R_t^2) \right\}] + \frac{1}{2} \frac{\partial^2 F}{\partial S^2}(t, S_t) \sigma^2 S_t^2 (2\theta a(1 + R_t^2)) \right\} dt \\
&\quad + \frac{\partial F}{\partial S}(t, S_t) S_t \sigma \sqrt{2\theta a(1 + R_t^2)} dB_t
\end{aligned}$$

For the discounted derivative price, using  $d\widehat{B}_t$  from Theorem 5.1 and  $u_t$  from (11), we obtain

$$\begin{aligned}
&d(e^{-rt} F(t, S_t)) \\
&= -r e^{-rt} F(t, S_t) dt + e^{-rt} dF(t, S_t) \\
&= e^{-rt} \left\{ -r F(t, S_t) + \frac{\partial F}{\partial t}(t, S_t) + \frac{\partial F}{\partial S}(t, S_t) [S_t \left\{ -\theta(R_t - \mu) + \sigma^2 \theta a (1 + R_t^2) \right\}] \right. \\
&\quad \left. + \frac{1}{2} \frac{\partial^2 F}{\partial S^2}(t, S_t) \sigma^2 S_t^2 (2\theta a(1 + R_t^2)) \right\} dt \\
&\quad + e^{-rt} \frac{\partial F}{\partial S}(t, S_t) S_t \sigma \sqrt{2\theta a(1 + R_t^2)} dB_t \\
&= e^{-rt} \left\{ -r F(t, S_t) + \frac{\partial F}{\partial t}(t, S_t) + \frac{\partial F}{\partial S}(t, S_t) [S_t \left\{ -\theta(R_t - \mu) + \sigma^2 \theta a (1 + R_t^2) \right\}] \right. \\
&\quad \left. + \frac{1}{2} \frac{\partial^2 F}{\partial S^2}(t, S_t) \sigma^2 S_t^2 (2\theta a(1 + R_t^2)) \right\} dt \\
&\quad + e^{-rt} \frac{\partial F}{\partial S}(t, S_t) S_t \sigma \sqrt{2\theta a(1 + R_t^2)} d\widehat{B}_t - e^{-rt} \frac{\partial F}{\partial S}(t, S_t) S_t \sigma \sqrt{2\theta a(1 + R_t^2)} u_t dt, \\
&= e^{-rt} \left\{ -r F(t, S_t) + \frac{\partial F}{\partial t}(t, S_t) + \frac{\partial F}{\partial S}(t, S_t) [S_t \left\{ -\theta(R_t - \mu) + \sigma^2 \theta a (1 + R_t^2) \right\}] \right. \\
&\quad \left. + \frac{1}{2} \frac{\partial^2 F}{\partial S^2}(t, S_t) \sigma^2 S_t^2 (2\theta a(1 + R_t^2)) \right\} dt \\
&\quad + e^{-rt} \frac{\partial F}{\partial S}(t, S_t) S_t \sigma \sqrt{2\theta a(1 + R_t^2)} d\widehat{B}_t \\
&\quad - e^{-rt} \frac{\partial F}{\partial S}(t, S_t) S_t \sigma \sqrt{2\theta a(1 + R_t^2)} \left( \frac{-r - \theta(R_t - \mu) + \sigma^2 \theta a (1 + R_t^2)}{\sigma \sqrt{2\theta a(1 + R_t^2)}} \right) dt
\end{aligned}$$

$$= e^{-rt} \left\{ -rF(t, S_t) + \frac{\partial F}{\partial t}(t, S_t) + rS_t \frac{\partial F}{\partial S}(t, S_t) + \frac{\partial^2 F}{\partial S^2}(t, S_t) \sigma^2 S_t^2 (\theta a(1 + R_t^2)) \right\} dt \\ + e^{-rt} \frac{\partial F}{\partial S}(t, S_t) S_t \sigma \sqrt{2\theta a(1 + R_t^2)} d\widehat{B}_t.$$

Equating the drift term to be zero and converting it to the deterministic condition, we end up with a Black-Scholes type PDE

$$(12) \quad \frac{\partial F}{\partial t}(t, x) + rx \frac{\partial F}{\partial x}(t, x) + \sigma^2 x^2 \left( \theta a \left( 1 + \left( \ln \frac{x}{S_0} \right)^2 \right) \right) \frac{\partial^2 F}{\partial x^2}(t, x) = rF(t, x).$$

The above discussion is formulated in the following proposition.

**Proposition 5.4.** *Assume that the log return follows Pearson diffusion process (6). Then, the Black-Scholes type PDE for derivative pricing is (12).*

**5.2. Existence and Uniqueness of stock price under  $\mathbb{Q}$ .** Under  $\mathbb{Q}$ , the log return process satisfies

$$(13) \quad \begin{aligned} dR_t &= -\theta(R_t - \mu) dt + \sigma \sqrt{2\theta a(1 + R_t^2)} dB_t \\ &= -\theta(R_t - \mu) dt + \sigma \sqrt{2\theta a(1 + R_t^2)} (d\widehat{B}_t - u_t dt) \\ &= -\theta(R_t - \mu) dt + \sigma \sqrt{2\theta a(1 + R_t^2)} d\widehat{B}_t \\ &\quad - \sigma \sqrt{2\theta a(1 + R_t^2)} \left( \frac{-r - \theta(R_t - \mu) + \sigma^2 \theta a (1 + R_t^2)}{\sigma \sqrt{2\theta a(1 + R_t^2)}} \right) dt \\ &= (r - \sigma^2 \theta a (1 + R_t^2)) dt + \sigma \sqrt{2\theta a(1 + R_t^2)} d\widehat{B}_t, \end{aligned}$$

Let  $\Pi(x)$  be the probability density function corresponding to the steady state distribution of (13) under the probability measure  $\mathbb{Q}$ . Then  $\Pi(x)$  satisfies the Fokker-Planck equation defined as:

$$\frac{\partial^2}{\partial x^2} \left\{ \frac{1}{2} \sigma^2 2\theta a (1 + x^2) \Pi(x) \right\} = \frac{\partial}{\partial x} \left\{ (r - \sigma^2 \theta a (1 + x^2)) \Pi(x) \right\},$$

simplifying which we get,

$$(14) \quad (1 + x^2) \Pi''(x) + \left\{ 4x - \frac{r}{\sigma^2 \theta a} + (1 + x^2) \right\} \Pi'(x) + 2(1 + x) \Pi(x) = 0.$$

One can get the steady state distribution of (13) by solving the above differential equation (14). Here, we are unable to obtain any closed form solution of (14).

Next we discuss the stock price dynamics under  $\mathbb{Q}$ ,

(15)

$$\begin{aligned}
dS_t &= S_t \{ -\theta(R_t - \mu) + \sigma^2 \theta a (1 + R_t^2) \} dt + S_t \sigma \sqrt{2\theta a (1 + R_t^2)} dB_t \\
&= S_t \{ -\theta(R_t - \mu) + \sigma^2 \theta a (1 + R_t^2) \} dt + S_t \sigma \sqrt{2\theta a (1 + R_t^2)} (\widehat{dB}_t - u_t dt) \\
&= S_t \{ -\theta(R_t - \mu) + \sigma^2 \theta a (1 + R_t^2) \} dt + S_t \sigma \sqrt{2\theta a (1 + R_t^2)} \widehat{dB}_t \\
&\quad - S_t \sigma \sqrt{2\theta a (1 + R_t^2)} \left( \frac{-r - \theta(R_t - \mu) + \sigma^2 \theta a (1 + R_t^2)}{\sigma \sqrt{2\theta a (1 + R_t^2)}} \right) dt \\
&= r S_t dt + S_t \sigma \sqrt{2\theta a (1 + R_t^2)} \widehat{dB}_t \\
&= r S_t dt + S_t \sigma \sqrt{2\theta a \left( 1 + \left( \ln \frac{S_t}{S_0} \right)^2 \right)} \widehat{dB}_t.
\end{aligned}$$

By Girsanov's Theorem, the SDE (15) under  $\mathbb{Q}$  has a weak solution. In the next theorem, we establish the existence and uniqueness of strong solution of this equation.

**Theorem 5.5.** *The SDE (15) has a unique strong solution.*

*Proof.* Consider the diffusion coefficient -

$$\hat{\sigma}(x) := x \sigma \sqrt{2\theta a \left( 1 + \left( \ln \frac{x}{S_0} \right)^2 \right)}, \quad \forall x \in \mathbb{R}.$$

Then,

$$\hat{\sigma}'(x) = \sigma \sqrt{2\theta a \left( 1 + \left( \ln \frac{x}{S_0} \right)^2 \right)} + x \sigma \frac{1}{2} \frac{1}{\sqrt{2\theta a \left( 1 + \left( \ln \frac{x}{S_0} \right)^2 \right)}} 2 \left( \ln \frac{x}{S_0} \right) \frac{1}{x}$$

$$= \sigma \sqrt{2\theta a \left( 1 + \left( \ln \frac{x}{S_0} \right)^2 \right)} + \frac{\sigma \ln \frac{x}{S_0}}{\sqrt{2\theta a \left( 1 + \left( \ln \frac{x}{S_0} \right)^2 \right)}},$$

and consequently,

$$|\hat{\sigma}'(x)| \leq \sigma \sqrt{2\theta a \left( 1 + \left( \ln \frac{x}{S_0} \right)^2 \right)} + \frac{\sigma}{\sqrt{2\theta a}}.$$

Now, for any large  $L > 0$ , with  $|x|, |y| \leq L$ , by the Mean Value Theorem, we have

$$\begin{aligned} & |\hat{\sigma}(x) - \hat{\sigma}(y)| \\ &= |\hat{\sigma}'(z)| |x - y|, \quad [\text{for } z \in (x, y) \Rightarrow |z| \leq L] \\ &\leq \left( \sigma \sqrt{2\theta a \left( 1 + \left( \ln \frac{L}{S_0} \right)^2 \right)} + \frac{\sigma}{\sqrt{2\theta a}} \right) |x - y|. \end{aligned}$$

Therefore, the diffusion coefficient is locally Lipschitz and the drift coefficient  $\hat{b}(x) := rx, x \in \mathbb{R}$  is clearly globally Lipschitz. Therefore, from [21, Chapter 5, Theorem 2.5], we have the local existence and uniqueness for strong solution of  $S_t$ .

Now, we verify condition (3.13) of [9]. For  $x^2 > e$ ,

$$\begin{aligned} & r\hat{b}(x) - \frac{1}{2}\hat{\sigma}^2(x) \\ &= rx^2 - \frac{1}{2}x^2\sigma^22\theta a \left( 1 + \left( \ln \frac{x}{S_0} \right)^2 \right) \\ &= (r - \sigma^2\theta a)x^2 - x^2\sigma^2\theta a \left( \ln \frac{x}{S_0} \right)^2 \\ &\leq |r - \sigma^2\theta a| x^2 \\ &\leq |r - \sigma^2\theta a| x^2 \ln(x^2), \end{aligned}$$

where we use the fact that  $\ln(x^2) > 1$ . Therefore, we conclude that the solution is always regular i.e. non-explosive. This completes the proof.  $\square$

**Remark 5.6.** *By the existence and uniqueness of the strong solution discussed above, we have the global solution of  $S_t, t \leq T$ . However, in our approach we have not been able to establish the integrability of  $S_t$  under  $\mathbb{Q}$ . As a result,  $e^{-rt}S_t, t \leq T$  becomes a local martingale under  $\mathbb{Q}$  and we have not been able to establish the martingale property.*

**5.3. Estimation of Call option price.** In what follows, we assume that  $e^{-rt}S_t, t \leq T$  is a  $\mathbb{Q}$  martingale. The price of a call option at time  $t$ , with maturity at  $T$  ( $t < T$ ), strike price  $K$ , and risk-free interest rate  $r$  can be estimated as:

$$(16) \quad C(t) = \mathbb{E}^{\mathbb{Q}} \left[ e^{-r(T-t)} \max\{S_T - K, 0\} | I_t \right].$$

In this study, we are not able to get any closed-form solution for the call option price. Therefore, we estimate the option price through a simulation approach. We first simulate the stock price processes  $S_t, t \leq T$  using the stochastic differential equation (15), and then compute the call option price using (16).

**Remark 5.7.** *From Girsanov Theorem 5.1, the call option price at  $t = 0$ ,*

$$\begin{aligned} C(0) &= \int_{\Omega} e^{-rT} \max\{S_T - K, 0\} d\mathbb{Q}(\omega) \\ &= \int_{\Omega} e^{-rT} \max\{S_T - K, 0\} M_T(\omega) d\mathbb{P}(\omega) \\ &= \int_{\Omega} e^{-rT} \max\{S_T - K, 0\} \exp \left( - \int_0^T u_s dB_s - \frac{1}{2} \int_0^T u_s^2 ds \right) d\mathbb{P}. \end{aligned}$$

## 6. EMPIRICAL EVIDENCE

To empirically assess the effectiveness of the proposed method, we use daily data of European call options on Nifty 50 index. We select the Nifty 50 index because it is India's most actively traded derivative instrument on the National Stock Exchange (NSE). Our study covers April 11, 2022, to August 30, 2024, focusing on recent historical data. The data is sourced from the publicly available NSE website. The dataset has information about trade

date, time to maturity, underlying asset price, strike price, trading volume, total turnover, and the open, high, low, and close prices for all the option contracts for the sample period. In our sample, we restrict the maturity time of options to a maximum of three months. More than three-month maturity option contracts on the NSE are less frequently traded and tend to exhibit lower liquidity.

A significant portion of the option contracts in our sample display low or negligible liquidity. To address this, we use turnover as a proxy for liquidity. Turnover is calculated as the product of contracts traded, lot size, and the underlying price, and is measured in lakhs of Indian rupees. We exclude contracts falling below the seventh decile of the turnover. After applying this filter, the final dataset consists of 62442 observations, with the minimum turnover value of 28282.52. We use the yield on 91-day Treasury bills as the risk-free interest rate. We collect the interest rate data from Reserve Bank of India (RBI) website. The data and the R-codes are available on request.

We group the option data based on moneyness and time to maturity categories. Let  $S_t$  be the underlying asset price at time  $t$  and  $K$  be the strike price. A call option is classified as at-the-money (ATM) if  $\frac{S_t}{K} \in (0.97, 1.03)$ , out-of-the-money (OTM) if  $\frac{S_t}{K} \leq 0.97$ , and in-the-money (ITM) if  $\frac{S_t}{K} \geq 1.03$  (see: Bakshi *et al.* [1]).

Based on time to maturity, option contracts are classified into five categories: A (0 day < Time to maturity  $\leq$  7 days); B (7 days < Time to maturity  $\leq$  15 days); C (15 days < Time to maturity  $\leq$  30 days); D (30 days < Time to maturity  $\leq$  60 days); and E (60 days < Time to maturity  $\leq$  90 days).

Tables 1 and 2 report the time-to-maturity distribution and the moneyness distribution of the traded option contracts, respectively.

**Table 1.** Time to maturity distribution of the option contracts.

|                             | A       | B       | C       | D      | E      |
|-----------------------------|---------|---------|---------|--------|--------|
| Number of contracts         | 32142   | 17507   | 9548    | 2965   | 280    |
| Percentage of contracts (%) | 51.4750 | 28.0372 | 15.2910 | 4.7484 | 0.4484 |

**Table 2.** Moneyness distribution of the option contracts.

|                             | ATM     | ITM    | OTM     |
|-----------------------------|---------|--------|---------|
| Number of contracts         | 31617   | 1131   | 29694   |
| Percentage of contracts (%) | 50.6342 | 1.8113 | 47.5545 |

As shown in Table 1, the majority of the traded options (94.8032%) have a time to maturity within categories A to C, while category E contains the smallest proportion of contracts.

Table 2 indicates that trading activity is more concentrated in at-the-money (ATM) and out-of-the-money (OTM) call options compared to in-the-money (ITM) call options.

**6.1. Model comparison using the historical approach.** We evaluate our model's performance against the classical Black–Scholes and Heston models. Under the Pearson diffusion framework, the statistical dynamics of log returns  $R_t$  is governed by (6), an SDE having four parameters:  $(\theta, a, \mu, \sigma)$ . Our first step is to estimate these parameters. In the historical approach, model parameters are estimated from historical log return data and then used to estimate the model option price. For the historical approach, we need to select an appropriate window size. As mentioned in Hull and Basu [18], the estimation window for historical volatility in the Black–Scholes model ranges from 30 to 180 trading days in practice. In this study, we select two rolling window lengths of 90 and 180 trading days. In other words, to estimate the model option price on a given day, we first collect the daily closing prices from the preceding 90 or 180 trading days,

compute the corresponding log returns, estimate the model parameters, and then use these parameters to calculate the model option price. In the simulation study, call option prices are computed using 200000 Monte Carlo replications under the proposed PIV method.

For all three models - PIV, Black–Scholes (BS), and Heston (HS) - parameters are estimated via the maximum likelihood method. In the estimation procedure, we set  $dt = \frac{1}{252}$ , reflecting the use of daily data assuming that the NSE operates 252 trading days per calendar year.

**6.2. Out-of-sample performance of historical approach.** The out-of-sample dataset consists of 53310 and 44680 observations corresponding to the rolling window sizes of 90 and 180 days, respectively. The pricing error of an option is defined as the difference between the model price and the market price. We consider the closing price of a traded option to be the market price. To evaluate model performance, we employ two error metrics: mean absolute error (MAE) and mean squared error (MSE). Tables 3 and 4 display the performance of all models using rolling window sizes of 90 and 180 days, respectively, categorized by moneyness. Tables 5 and 6 show the corresponding results for the same window sizes, for the time to maturity category.

**Table 3.** Model Performance under moneyness category by rolling window approach with size 90.

|            | <b>PIV</b>       | <b>BS</b>  | <b>HS</b> |
|------------|------------------|------------|-----------|
| <b>MAE</b> |                  |            |           |
| ITM        | <b>36.9693</b>   | 54.6638    | 44.7100   |
| OTM        | <b>11.4605</b>   | 34.4065    | 14.9941   |
| ATM        | <b>36.2825</b>   | 81.1552    | 40.6405   |
| ALL        | <b>24.8806</b>   | 59.1877    | 28.9196   |
| <b>MSE</b> |                  |            |           |
| ITM        | <b>3109.4387</b> | 7791.0597  | 4982.0560 |
| OTM        | <b>819.6919</b>  | 6542.1246  | 2399.6011 |
| ATM        | <b>4018.4329</b> | 15329.0959 | 6137.0253 |
| ALL        | <b>2531.3807</b> | 11154.6008 | 4397.8962 |

*Note:* This table reports the mean absolute error (MAE) and mean squared error (MSE) for all models. PIV, BS, and HS denote the Pearson diffusion, Black–Scholes, and Heston models, respectively. ITM, OTM, and ATM refer to in-the-money, out-of-the-money, and at-the-money options. ALL represents the entire out-of-sample dataset. Across the entire out-of-sample dataset, the proportions of ITM, OTM, and ATM options are 1.78%, 45.98%, and 52.24%, respectively. Bold values indicate the lowest MAE and MSE in each row.

**Table 4.** Model Performance under moneyness category by rolling window approach with size 180.

|            | <b>PIV</b>       | <b>BS</b>  | <b>HS</b> |
|------------|------------------|------------|-----------|
| <b>MAE</b> |                  |            |           |
| ITM        | <b>38.6962</b>   | 48.9129    | 47.8571   |
| OTM        | <b>12.0293</b>   | 29.5032    | 16.0753   |
| ATM        | <b>38.2951</b>   | 75.2467    | 43.0964   |
| ALL        | <b>26.4330</b>   | 54.0835    | 30.9744   |
| <b>MSE</b> |                  |            |           |
| ITM        | <b>3367.1278</b> | 5055.1497  | 5630.3482 |
| OTM        | <b>854.5360</b>  | 4231.0854  | 2685.0795 |
| ATM        | <b>3972.1454</b> | 10881.1451 | 6967.7207 |
| ALL        | <b>2551.9969</b> | 7767.2186  | 5007.4213 |

*Note:* This table reports the mean absolute error (MAE) and mean squared error (MSE) for all models. PIV, BS, and HS denote the Pearson diffusion, Black–Scholes, and Heston models, respectively. ITM, OTM, and ATM refer to in-the-money, out-of-the-money, and at-the-money options. ALL represents the entire out-of-sample dataset. Across the entire out-of-sample dataset, the proportions of ITM, OTM, and ATM options are 1.87%, 45.19%, and 52.94%, respectively. Bold values indicate the lowest MAE and MSE in each row.

**Table 5.** Model Performance under time-to-maturity category by rolling window approach with size 90.

|            | <b>PIV</b>        | <b>BS</b>  | <b>HS</b>      |
|------------|-------------------|------------|----------------|
| <b>MAE</b> |                   |            |                |
| A          | <b>11.2749</b>    | 32.0895    | 13.7173        |
| B          | <b>28.5835</b>    | 72.2835    | 35.5088        |
| C          | <b>47.5368</b>    | 97.9993    | 52.2150        |
| D          | 61.6959           | 131.1378   | <b>60.4702</b> |
| E          | <b>128.1301</b>   | 158.6207   | 160.7986       |
| <b>MSE</b> |                   |            |                |
| A          | <b>689.8709</b>   | 3466.8727  | 1180.3141      |
| B          | <b>2464.4858</b>  | 11985.3773 | 4908.5419      |
| C          | <b>5872.2965</b>  | 23112.1791 | 9431.0848      |
| D          | <b>8866.9861</b>  | 40282.0234 | 13405.8085     |
| E          | <b>26497.2694</b> | 86574.5441 | 55562.8110     |

*Note:* This table presents the mean absolute error (MAE) and mean squared error (MSE) for all models. PIV, BS, and HS represent the Pearson diffusion, Black–Scholes, and Heston models, respectively. Categories A through E classify options based on time to expiry as follows: A (7 days or less), B (more than 7 to 15 days), C (more than 15 to 30 days), D (more than 30 to 60 days), and E (more than 60 to 90 days). Across the entire out-of-sample dataset, the proportions of option contracts corresponding to time-to-maturity categories A, B, C, D, and E are 50.78%, 28.18%, 15.72%, 4.79%, and 0.52%, respectively. Bold values highlight the lowest MAE and MSE for each row.

**Table 6.** Model Performance under time-to maturity category by rolling window approach with size 180.

|            | <b>PIV</b>        | <b>BS</b>  | <b>HS</b>      |
|------------|-------------------|------------|----------------|
| <b>MAE</b> |                   |            |                |
| A          | <b>12.0051</b>    | 28.1144    | 14.7708        |
| B          | <b>30.6041</b>    | 65.0758    | 38.5824        |
| C          | <b>49.5831</b>    | 90.8299    | 54.6517        |
| D          | 65.4378           | 127.3506   | <b>62.6735</b> |
| E          | <b>114.3168</b>   | 148.2005   | 155.4543       |
| <b>MSE</b> |                   |            |                |
| A          | <b>729.9839</b>   | 2429.8010  | 1371.3447      |
| B          | <b>2468.0535</b>  | 8254.1154  | 5687.6544      |
| C          | <b>5686.9298</b>  | 15902.2269 | 10627.6116     |
| D          | <b>9194.8577</b>  | 28308.5065 | 14623.8995     |
| E          | <b>23047.8746</b> | 50836.7810 | 53501.0383     |

*Note:* This table presents the mean absolute error (MAE) and mean squared error (MSE) for all models. PIV, BS, and HS represent the Pearson diffusion, Black–Scholes, and Heston models, respectively. Categories A through E classify options based on time to expiry as follows: A (7 days or less), B (more than 7 to 15 days), C (more than 15 to 30 days), D (more than 30 to 60 days), and E (more than 60 to 90 days). Across the entire out-of-sample dataset, the proportions of option contracts corresponding to time-to-maturity categories A, B, C, D, and E are 50.43%, 28.25%, 15.75%, 5.01%, and 0.57%, respectively. Bold values highlight the lowest MAE and MSE for each row.

Tables 3 and 4 show that the proposed method consistently outperforms the classical Black–Scholes model and the Heston stochastic volatility model across all moneyness categories, for both two error metrics and rolling window sizes. Likewise, Tables 5 and 6 indicate that the PIV method delivers superior performance across all time-to-maturity categories, under both error functions and window sizes, except for Category D (options with maturities between 30 and 60 days) under the MAE criterion.

Next, we apply the Diebold and Mariano [14] test to evaluate whether the differences in predictive accuracy between the competing models are statistically significant. The test is conducted using both absolute error (AE) and squared error (SE) as loss functions. The null hypothesis of the test states that the forecasts from the PIV and HS (or BS) models have the same predictive accuracy, while the alternative hypothesis states that the

PIV model forecasts are more accurate than the HS (or BS) model. Tables 7 and 8 report the p-values of the Diebold and Mariano test comparing the predictive performance of the PIV model with the HS and BS models, using rolling window sizes of 90 and 180 days, respectively.

Tables 7 and 8 indicate that the null hypothesis is rejected at the 1% significance level across all moneyness categories, for both error functions and all rolling window sizes, providing statistical evidence that the forecasts produced by the proposed model are more accurate than those of the HS and BS models at 1% significance level.

**Table 7.** p-values of the Diebold-Mariano Test for rolling window size 90.

| <b>PIV vs HS</b> |                      |                      |
|------------------|----------------------|----------------------|
|                  | <b>AE</b>            | <b>SE</b>            |
| ITM              | 4.30E <sup>-10</sup> | 6.46E <sup>-08</sup> |
| OTM              | 2.20E <sup>-16</sup> | 2.20E <sup>-16</sup> |
| ATM              | 2.20E <sup>-16</sup> | 2.20E <sup>-16</sup> |
| ALL              | 2.20E <sup>-16</sup> | 2.20E <sup>-16</sup> |
| <b>PIV vs BS</b> |                      |                      |
|                  | <b>AE</b>            | <b>SE</b>            |
| ITM              | 2.20E <sup>-16</sup> | 1.76E <sup>-11</sup> |
| OTM              | 2.20E <sup>-16</sup> | 2.20E <sup>-16</sup> |
| ATM              | 2.20E <sup>-16</sup> | 2.20E <sup>-16</sup> |
| ALL              | 2.20E <sup>-16</sup> | 2.20E <sup>-16</sup> |

*Note:* This table presents the p-values from the Diebold and Mariano test, used to assess the predictive accuracy of the PIV model relative to the HS (PIV vs HS) and BS (PIV vs BS) models. ITM, OTM, and ATM refer to in-the-money, out-of-the-money, and at-the-money options. ALL represents the entire out-of-sample dataset. The AE and SE columns report the p-values of the test based on the Absolute Error and Squared Error loss functions, respectively.

**Table 8.** p-values of the Diebold-Mariano Test for rolling window size 180.

| PIV vs HS |                      |                      |
|-----------|----------------------|----------------------|
|           | AE                   | SE                   |
| ITM       | 6.77E <sup>-10</sup> | 1.07E <sup>-07</sup> |
| OTM       | 2.20E <sup>-16</sup> | 2.20E <sup>-16</sup> |
| ATM       | 2.20E <sup>-16</sup> | 2.20E <sup>-16</sup> |
| ALL       | 2.20E <sup>-16</sup> | 2.20E <sup>-16</sup> |
| PIV vs BS |                      |                      |
|           | AE                   | SE                   |
| ITM       | 6.88E <sup>-11</sup> | 4.68E <sup>-06</sup> |
| OTM       | 2.20E <sup>-16</sup> | 2.20E <sup>-16</sup> |
| ATM       | 2.20E <sup>-16</sup> | 2.20E <sup>-16</sup> |
| ALL       | 2.20E <sup>-16</sup> | 2.20E <sup>-16</sup> |

*Note:* This table presents the p-values from the Diebold and Mariano test, used to assess the predictive accuracy of the PIV model relative to the HS (PIV vs HS) and BS (PIV vs BS) models. ITM, OTM, and ATM refer to in-the-money, out-of-the-money, and at-the-money options, respectively. ALL represents the entire out-of-sample dataset. The AE and SE columns report the p-values of the test based on the Absolute Error and Squared Error loss functions, respectively.

**6.3. Out-of-sample performance using Implied approach.** The historical approach requires selecting a window size to estimate option model parameters, which can have practical challenges due to its stringent requirements on historical data (Bakshi *et al.* [1]). To address this limitation, practitioners have traditionally opted to use the implied approach, where model parameters are estimated by inverting the option pricing model using recent market option prices. The implied parameters may be interpreted as the instant market estimate of the future price movement of the underlying asset. This method not only significantly reduces data requirements but has also been shown to improve performance (see: Bates [2, 3]).

For the case of the implied approach, we follow Bakshi *et al.* to estimate the model parameters. [1]. To estimate the call option price for the day  $t$ , we first collect all options traded on the preceding day. Let  $C_{market,t-1}$  denote the observed market price and  $C_{model,t-1}$  be model option price both at time

$t - 1$ . Let  $\Theta$  be the vector of model parameters. The implied parameters of each model for day  $t - 1$  are estimated by minimizing the sum of squared pricing errors:

$$\hat{\Theta} = \min_{\Theta} \sum (C_{\text{market},t-1} - C_{\text{model},t-1})^2$$

where the summation is taken over all options traded on day  $t - 1$ . Next we utilize those estimated implied parameters to estimate the option price of day  $t$ . Under the implied approach, the out-of-sample dataset comprises a total of 62103 observations.

Tables 9 and 10 report the out-of-sample forecast performance of all three models across moneyness and time-to-maturity categories, respectively. Table 11 presents the p-values of the Diebold and Mariano test, assessing whether the differences in predictive accuracy among the models are statistically significant.

Table 9 demonstrates that the PIV model outperforms the other models for at-the-money (ATM) options as well as across the entire out-of-sample dataset under both error metrics. In contrast, the HS model delivers better forecasts for in-the-money (ITM) and out-of-the-money (OTM) options. Given that over 50% of the options in both the full sample (Table 2) and the out-of-sample dataset (Table 9) fall within the ATM category, the superior performance of the proposed model in this segment is particularly significant. Table 11 presents statistical evidence confirming that the PIV model's forecasts for ATM options are more accurate than those of the HS and BS models at 1% significance level.

For data grouped by time-to-maturity (Table 10), the PIV method outperforms others in categories A and B under the MAE metric, which together account for over 79% of traded options in both the full sample (Table 1) and the out-of-sample (Table 10) dataset. The proposed PIV model delivers superior performance in the groups with the highest trading activity, highlighting its effectiveness within the implied approach framework as well.

However, under the MSE metric, the proposed method does not exhibit superior performance, the BS model performs better in categories B and C, while the HS model outperforms in the remaining three categories.

**Table 9.** Model Performance under moneyness category by implied approach.

|            | PIV             | BS        | HS               |
|------------|-----------------|-----------|------------------|
| <b>MAE</b> |                 |           |                  |
| ITM        | 27.5377         | 26.9758   | <b>26.1722</b>   |
| OTM        | 5.6576          | 5.4501    | <b>4.9065</b>    |
| ATM        | <b>15.0136</b>  | 16.9099   | 15.8467          |
| ALL        | <b>10.7919</b>  | 11.6430   | 10.8316          |
| <b>MSE</b> |                 |           |                  |
| ITM        | 1628.9130       | 1565.7580 | <b>1517.8892</b> |
| OTM        | 189.8867        | 174.3626  | <b>131.9044</b>  |
| ATM        | <b>583.6524</b> | 634.4850  | 623.1860         |
| ALL        | <b>415.4041</b> | 432.6046  | 405.8215         |

*Note:* This table reports the mean absolute error (MAE) and mean squared error (MSE) for all models. To compute the option price on a given day for all models, we first estimate the implied parameters using all options traded on the previous day, and then use these parameters to calculate the option price. PIV, BS, and HS denote the Pearson diffusion, Black–Scholes, and Heston models, respectively. ITM, OTM, and ATM refer to in-the-money, out-of-the-money, and at-the-money options. ALL represents the entire out-of-sample dataset. Across the entire out-of-sample dataset, the proportions of ITM, OTM, and ATM options are 1.82%, 47.56%, and 50.62%, respectively. Bold values indicate the lowest MAE and MSE in each row.

**Table 10.** Model Performance under time-to-maturity category by implied approach.

|            | <b>PIV</b>     | <b>BS</b>       | <b>HS</b>        |
|------------|----------------|-----------------|------------------|
| <b>MAE</b> |                |                 |                  |
| A          | <b>6.6631</b>  | 8.3303          | 6.9006           |
| B          | <b>12.1654</b> | 12.3763         | 13.4534          |
| C          | 17.1369        | 16.8262         | <b>16.0479</b>   |
| D          | 23.0805        | 22.4659         | <b>18.8199</b>   |
| E          | 50.3616        | 53.0035         | <b>34.0092</b>   |
| <b>MSE</b> |                |                 |                  |
| A          | 204.2751       | 246.4082        | <b>196.3507</b>  |
| B          | 420.3412       | <b>404.1864</b> | 480.5811         |
| C          | 681.9721       | <b>675.7133</b> | 735.1835         |
| D          | 1248.3535      | 1257.3568       | <b>936.6570</b>  |
| E          | 6308.4326      | 6443.3622       | <b>2822.3595</b> |

*Note:* This table presents the mean absolute error (MAE) and mean squared error (MSE) for all models by implied approach. To compute the option price on a given day for all models, we first estimate the implied parameters using all options traded on the previous day, and then use these parameters to calculate the option price. PIV, BS, and HS represent the Pearson diffusion, Black–Scholes, and Heston models, respectively. Categories A through E classify options based on time to expiry as follows: A (7 days or less), B (more than 7 to 15 days), C (more than 15 to 30 days), D (more than 30 to 60 days), and E (more than 60 to 90 days). Across the entire out-of-sample dataset, the proportions of option contracts corresponding to time-to-maturity categories A, B, C, D, and E are 51.40%, 28.06%, 15.33%, 4.76%, and 0.45%, respectively. Bold values highlight the lowest MAE and MSE for each row.

**Table 11.** p-values of the Diebold-Mariano Test for implied approach.

| PIV vs HS |                      |                      |
|-----------|----------------------|----------------------|
|           | AE                   | SE                   |
| ITM       | 0.9993               | 0.9607               |
| OTM       | 1.0000               | 1.0000               |
| ATM       | 2.20E <sup>-16</sup> | 0.0002               |
| ALL       | 0.2003               | 0.8884               |
| PIV vs BS |                      |                      |
|           | AE                   | SE                   |
| ITM       | 0.9988               | 0.9990               |
| OTM       | 1.0000               | 1.0000               |
| ATM       | 2.20E <sup>-16</sup> | 2.20E <sup>-16</sup> |
| ALL       | 2.20E <sup>-16</sup> | 3.88E <sup>-14</sup> |

*Note:* This table presents the p-values from the Diebold and Mariano test, used to assess the predictive accuracy of the PIV model relative to the HS (PIV vs HS) and BS (PIV vs BS) models. ITM, OTM, and ATM refer to in-the-money, out-of-the-money, and at-the-money options, respectively. ALL represents the entire out-of-sample dataset. The AE and SE columns report the p-values of the test based on the Absolute Error and Squared Error loss functions, respectively.

## 7. CONCLUSION

We propose a novel method for estimating call option prices by modeling the log return dynamics using Pearson diffusion processes, characterized by a linear drift and a quadratic squared diffusion coefficient. The advantage of this class of diffusion is its ability to capture the excess kurtosis and skewness observed in the return distribution. We also verify the Novikov condition, a key requirement for risk-neutral pricing, thereby ensuring that the model does not admit arbitrage opportunities. Further, we establish the existence of a unique strong solution for the stock price process under the risk-neutral measure. We apply our method to Nifty 50 index options and compare the results to the classical Black-Scholes and Heston stochastic volatility models using historical and implied approaches.

Under the historical approach, the proposed method consistently outperforms competing models across all rolling window sizes and evaluation

metrics. These results hold across option categories defined by moneyness and time-to-maturity (except 30 to 60 days maturity options).

Under the implied approach, the PIV method demonstrates superior performance for at-the-money options, which account for over 50% of the observations, as well as across the entire out-of-sample dataset. For the time-to-maturity category, the PIV method outperforms other methods for short-term maturity options (one-week and two-weeks maturity options), which together constitute more than 79% of the dataset. The Heston model exhibits better performance for long-term maturity options.

The findings based on historical and implied approaches show the practical relevance of our approach and suggest that it offers a more effective alternative for option pricing.

**Acknowledgment:** We gratefully acknowledge Professor Malay Bhattacharyya for introducing us to the Pearson diffusion processes. Barun Sarkar would like to acknowledge the seed grant of Indian Institute of Technology Madras. Suprio Bhar acknowledges the support of the Matrics grant MTR/2021/000517 from the Science and Engineering Research Board (Department of Science & Technology, Government of India).

## APPENDIX A. APPENDIX

*Proof of Theorem 5.3.* Girsanov's Theorem 5.1 proposed a new Brownian motion  $\widehat{B}_t, t \leq T$  under  $\mathbb{Q}$  of the form

$$d\widehat{B}_t = u_t dt + dB_t, \quad t \leq T.$$

From (10),

$$\begin{aligned} d(e^{-rt} S_t) &= -re^{-rt} S_t dt + e^{-rt} S_t \{-\theta(R_t - \mu) + \sigma^2 \theta a (1 + R_t^2)\} dt \\ &\quad + e^{-rt} S_t \sigma \sqrt{2\theta a (1 + R_t^2)} (d\widehat{B}_t - u_t dt), \end{aligned}$$

and from (11)

$$u_t = \frac{-r - \theta(R_t - \mu) + \sigma^2 \theta a (1 + R_t^2)}{\sigma \sqrt{2\theta a (1 + R_t^2)}}.$$

The proposed probability measure  $\mathbb{Q}$  is given by

$$(17) \quad d\mathbb{Q} = M_T d\mathbb{P},$$

where

$$(18) \quad M_T := \exp \left( - \int_0^T u_s dB_s - \frac{1}{2} \int_0^T u_s^2 ds \right).$$

If we can show that  $M_t$ ,  $t \leq T$  is a  $\mathbb{P}$ -martingale, then  $\mathbb{Q}$  becomes a probability measure. Under  $\mathbb{Q}$ ,

$$(19) \quad d(e^{-rt} S_t) = e^{-rt} S_t \sigma \sqrt{2\theta a (1 + R_t^2)} d\widehat{B}_t$$

is a local martingale. This achieves our target provided we show that  $M_t$ ,  $t \leq T$  is a  $\mathbb{P}$ -martingale. Observe that from (11)

$$\begin{aligned} (20) \quad u_t^2 &= \left| \frac{-r - \theta(R_t - \mu) + \sigma^2 \theta a (1 + R_t^2)}{\sigma \sqrt{2\theta a (1 + R_t^2)}} \right|^2 \\ &= \left| \frac{-r + \mu\theta - \theta R_t + \sigma^2 \theta a (1 + R_t^2)}{\sigma \sqrt{2\theta a (1 + R_t^2)}} \right|^2 \\ &\leq 3 \left\{ \left| \frac{\mu\theta - r}{\sigma \sqrt{2\theta a (1 + R_t^2)}} \right|^2 + \left| \frac{\theta R_t}{\sigma \sqrt{2\theta a (1 + R_t^2)}} \right|^2 + \left| \sigma \sqrt{\frac{a\theta}{2}} \sqrt{1 + R_t^2} \right|^2 \right\} \\ &\leq 3 \left\{ \frac{(\mu\theta - r)^2}{2\sigma^2 \theta a} + \frac{\theta^2}{2\sigma^2 \theta a} + \frac{\sigma^2 a\theta}{2} (1 + R_t^2) \right\} \\ &\leq K(\mu, \theta, r, \sigma, a) (1 + R_t^2), \end{aligned}$$

where  $K(\mu, \theta, r, \sigma, a) > 0$  is a constant depending only on  $\mu, \theta, r, \sigma$  and  $a$ .

Therefore for Novikov's condition, using (20),

$$\mathbb{E} \left[ \exp \left( \frac{1}{2} \int_0^T u_t^2 dt \right) \right]$$

$$\begin{aligned}
&= \mathbb{E} \left[ \exp \left( \frac{1}{2} \int_0^T \left| \frac{-r - \theta(R_t - \mu) + \sigma^2 \theta a (1 + R_t^2)}{\sigma \sqrt{2\theta a(1 + R_t^2)}} \right|^2 dt \right) \right] \\
&\leq \mathbb{E} \left[ \exp \left( \frac{1}{2} K(\mu, \theta, r, \sigma, a) \int_0^T (1 + R_t^2) dt \right) \right].
\end{aligned}$$

From (7) and (20), we know that

$$\mathbb{P} \left( \int_0^T (1 + R_t^2) dt < \infty \right) = 1.$$

Therefore,

$$(21) \quad \mathbb{P} \left( \int_0^T u_t^2 dt < \infty \right) = 1.$$

Following [15, Theorem 2.1], define

$$\chi_N(t) := \begin{cases} 1, & \text{if } \int_0^t u_s^2 ds \leq N \\ 0, & \text{otherwise.} \end{cases}$$

Therefore, from (21)

$$\lim_{N \rightarrow \infty} \mathbb{P} \left( \int_0^T u_t^2 dt \leq N \right) = 1.$$

Then, from definition of  $\chi_N$

$$\int_0^T (\chi_N(t) u_t)^2 dt \leq N, \quad P\text{-a.s.}$$

Now, define

$$M_t^N := \exp \left( - \int_0^t \chi_N(s) u_s dB_s - \frac{1}{2} \int_0^t (\chi_N(s) u_s)^2 ds \right).$$

Here, the Novikov condition is satisfied

$$\mathbb{E} \left[ \exp \left( \frac{1}{2} \int_0^T (\chi_N(s) u_s)^2 ds \right) \right] < \infty,$$

and consequently,  $\{M_t^N\}_{t \in [0, T]}$ , is an  $\mathcal{L}^2$  exponential martingale [19, Chapter III., Theorem 5.3]. Here,  $\mathcal{L}^2$  denotes the space of square integrable random

variables. Define

$$d\mathbb{Q}^N := M_T^N d\mathbb{P}$$

and note that  $\mathbb{E}(M_T^N) = 1$ . Hence,

$$\begin{aligned} 1 &= \mathbb{E}(M_T^N) \\ &= \mathbb{E}(\chi_N(T)M_T^N) + \mathbb{E}((1 - \chi_N(T))M_T^N) \\ &= \mathbb{E}(\chi_N(T)M_T) + \mathbb{P}(\chi_N(T) = 0) \end{aligned}$$

Now, by the Monotone Convergence Theorem,  $\lim_{N \rightarrow \infty} \mathbb{E}(\chi_N(T)M_T) = \mathbb{E}(M_T)$  and by (21),

$$\lim_{N \rightarrow \infty} \mathbb{P}(\chi_N(T) = 0) = \lim_{N \rightarrow \infty} \mathbb{P}\left(\int_0^T u_t^2 dt > N\right) = 0.$$

Therefore, we conclude

$$\lim_{N \rightarrow \infty} \mathbb{E}(M_T^N) = \mathbb{E}(M_T) = 1.$$

Thus,  $M_t$ ,  $t \leq T$  is a  $\mathbb{P}$ -martingale.  $\square$

## REFERENCES

- [1] Gurdip Bakshi, Charles Cao, and Zhiwu Chen. Empirical performance of alternative option pricing models. *The Journal of finance*, 52(5):2003–2049, 1997.
- [2] David S Bates. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, 9(1):69–107, 1996.
- [3] DS Bates. Testing option pricing models. chapter 20 in handbook of statistics and finance 14: Statistical methods in finance, 1996.
- [4] Stan Beckers. The constant elasticity of variance model and its implications for option pricing. *the Journal of Finance*, 35(3):661–673, 1980.
- [5] Malay Bhattacharyya, Nityanand Misra, and Bharat Kodase. Maxvar for non-normal and heteroskedastic returns. *Quantitative Finance*, 9(8):925–935, 2009.
- [6] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *J. Polit. Econ.*, 81(3):637–654, 1973.
- [7] Lisa Borland. A theory of non-gaussian option pricing. *Quantitative Finance*, 2(6):415, 2002.

- [8] Gian P Cervellera and Marco P Tucci. A note on the estimation of a gamma-variance process: Learning from a failure. *Computational Economics*, 49(3):363–385, 2017.
- [9] Pao-Liu Chow and Rafail Khasminskii. Almost sure explosion of solutions to stochastic differential equations. *Stochastic Process. Appl.*, 124(1):639–645, 2014.
- [10] George M Constantinides, Jens Carsten Jackwerth, and Stylianos Perrakis. Option pricing: real and risk-neutral distributions. *Handbooks in Operations Research and Management Science*, 15:565–591, 2007.
- [11] John Cox. Notes on option pricing i: Constant elasticity of variance diffusions. *Unpublished note, Stanford University, Graduate School of Business*, 1975.
- [12] John C Cox. The constant elasticity of variance option pricing model. *Journal of Portfolio Management*, page 15, 1996.
- [13] John C Cox and Stephen A Ross. The valuation of options for alternative stochastic processes. *Journal of financial economics*, 3(1-2):145–166, 1976.
- [14] Francis X Diebold and Robert S Mariano. Comparing predictive accuracy. *Journal of Business & economic statistics*, 20(1):134–144, 2002.
- [15] Benedetta Ferrario. Absolute continuity of laws for semilinear stochastic equations with additive noise. *Commun. Stoch. Anal.*, 2(2):209–227, 2008.
- [16] Julie Lyng Forman and Michael Sørensen. The Pearson diffusions: a class of statistically tractable diffusion processes. *Scand. J. Statist.*, 35(3):438–465, 2008.
- [17] Steven L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud.*, 6(2):327–343, 1993.
- [18] John C Hull and Sankarshan Basu. *Options, futures, and other derivatives*. Pearson Education India, 2016.
- [19] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, second edition, 1989.
- [20] Ahmet K Karagozoglu. Option pricing models: From black-scholes-merton to present. *Journal of Derivatives*, 29(4), 2022.
- [21] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.
- [22] Dilip B Madan, Peter P Carr, and Eric C Chang. The variance gamma process and option pricing. *Review of Finance*, 2(1):79–105, 1998.
- [23] Robert C. Merton. Theory of rational option pricing. *Bell J. Econom. and Management Sci.*, 4:141–183, 1973.

- [24] Robert C Merton. Option pricing when underlying stock returns are discontinuous. *Journal of financial economics*, 3(1-2):125–144, 1976.
- [25] Yuichi Nagahara. The PDF and CF of Pearson type IV distributions and the ML estimation of the parameters. *Statist. Probab. Lett.*, 43(3):251–264, 1999.
- [26] Bernt Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003. An introduction with applications.
- [27] Karl Pearson. Contributions to the mathematical theory of evolution ii. skew variation in homogeneous material. *Philosophical Transactions of the Royal Society of London. Series A*, 186:343–414, 1895.
- [28] Constantino Tsallis and Dirk Jan Bukman. Anomalous diffusion in the presence of external forces: Exact time-dependent solutions and their thermostatistical basis. *Physical Review E*, 54(3):R2197, 1996.
- [29] Michel Vellekoop and Hans Nieuwenhuis. On option pricing models in the presence of heavy tails. *Quantitative finance*, 7(5):563–573, 2007.

TAPAN KAR, INSTITUTE FOR FINANCIAL MANAGEMENT AND RESEARCH - GRADUATE SCHOOL OF BUSINESS, KREA UNIVERSITY, SRI CITY, ANDHRA PRADESH 524401, INDIA.

*Email address:* `tapan.kar@krea.edu.in`

SUPRIO BHAR, DEPARTMENT OF MATHEMATICS AND STATISTICS, INDIAN INSTITUTE OF TECHNOLOGY KANPUR, KALYANPUR, KANPUR - 208016, INDIA.

*Email address:* `suprio@iitk.ac.in`

BARUN SARKAR, DEPARTMENT OF MATHEMATICS, INDIAN INSTITUTE OF TECHNOLOGY MADRAS, CHENNAI - 600036, INDIA.

*Email address:* `barun@iitm.ac.in`

SESHA MEKA, INSTITUTE FOR FINANCIAL MANAGEMENT AND RESEARCH - GRADUATE SCHOOL OF BUSINESS, KREA UNIVERSITY, SRI CITY, ANDHRA PRADESH 524401, INDIA.

*Email address:* `sesha.meka@krea.edu.in`