Option Pricing with Dynamic Conditional Skewness*

Fang Liang† Lingshan Du‡

August 29, 2023

Efficiently exploiting information contained in price variations and accurately modeling the skewness of the underlying asset is critical for pricing options and other derivatives. In this paper, we propose a new and flexible option-pricing model that explicitly incorporates the dynamics of skewness and realized volatility. By the inverse Fourier transform, we derive closed-form option valuation formulas. Empirically, the model improves significantly upon benchmarks using S&P 500 index options. Overall, the joint modeling of skewness and realized volatility leads to an out-of-sample gain of 16.80% in pricing accuracy. The improvements are more pronounced for deep in-the-money calls, options with shorter maturities, and during highly volatile periods.

Keywords: Skewness, Inverse Gaussian Distribution, Option Pricing

JEL Classification: G13, C58, C51

*Fang Liang would like to acknowledge funding support provided by the National Natural Science Foundation of China (72101278).

†International School of Business & Finance, Sun Yat-sen University, China. Email: liangfang@mail.sysu.edu.cn

‡Guanghua School of Management, Peking University, China.

1 Introduction

The accurate specification of underlying asset distribution is crucial for option valuation frameworks. In recent years, there has been a growing body of literature advocating for the incorporation of heteroskedasticity of skewness in volatility models. These models utilize analytically approximation of the innovations or non-Gaussian distribution when modelling the return and volatility dynamics. This approach allows for more precise measurement and forecasting of asset volatility. The theoretical and empirical justifications for constructing reliable approximation such as Gram-Charlier expansion have been provided by Jondeau and Rockinger (2001), Wang et al. (2022), Corrado (2007), Schlögl (2013), Christoffersen et al. (2010), and other researchers consider a Cornish-Fisher expansion (see, for example, Aboura and Maillard (2016)) or an Edgeworth expansion (Zhang et al. (2011)). Notably, recent studies by Christoffersen et al. (2006), Barone-Adesi et al. (2008), and Stentoft (2008) have provided option pricing framework considering non-Gaussian distribution under discrete time GARCH-type models.

The aforementioned studies focus exclusively on applying mathematical tools such as a new distribution to capture skewness in the underlying asset and do not consider incorporating the empirical grounded properties simultaneously. Although dynamic volatility models with nonzero third moment innovations are conceptually and theoretically attractive, current discrete-time models may lack the necessary flexibility to fully capture observed option biases, even when leverage parameters are incorporated.

For instance, previous literature (Hull and White (1987), Christoffersen et al. (2006), Chernov and Ghysels (2000),Christoffersen and Jacobs (2004), and so on) agree with the facts that 1) Out-of-the-money put prices (and in-the-money call prices) are relatively high compared to the Black–Scholes price; 2) Options with shorter maturity tend to exhibit stronger bias based on standard models. These options are relatively more sensitive to the market conditions. Since trading occurs at every tick, and traders price in their expectations to option prices instantaneously. It is necessary to incorporate realized measures and skewness in the volatility modeling for deep in-the-money (ITM) calls, deep out-of-the-money (OTM) puts, options with short maturities, and options traded during periods of high market volatility. By introducing realized measures based on consideration of non-zero third moments, we expect to filter more accurate information and subsequently derive more precise option prices.

Figure 1 presents a time series plot comparing the CBOE Skew Index and Realized Volatility. Being similar to the CBOE VIX index, the CBOE SKEW Index (SKEW)[1] is calculated from the S&P 500 index options, reflecting the market expectation of distribution of the underlying assets. Figure 1 reveals a notable negative correlation between the two variables. During periods of market turbulence, such as the 2008 financial crisis and the 2020 stock market crash, Realized Volatility reaches high levels, while the CBOE Skew Index stay at lower levels. Conversely, during more stable market conditions, Realized Volatility tends to be lower, while the CBOE Skew Index remains elevated. This indicates that options traders incorporate the probability of black swan events into option prices, as captured by the Skew Index, during relatively calm market periods. Furthermore, as depicted in Figure 1, the SKEW index has exhibited increasing volatility in recent years, reaching its highest point in 2021, surpassing 170.

[1] CBOE SKEW Index White Paper: https://cdn.cboe.com/resources/indices/documents/SKEWwhitepaperjan2011.pdf

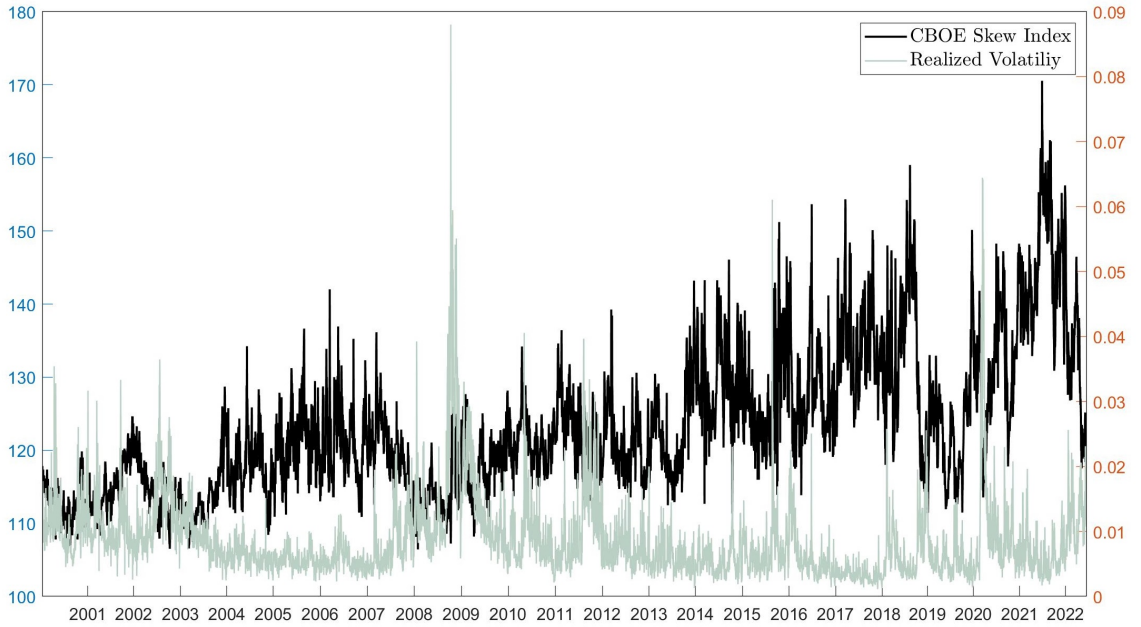


Figure 1: CBOE Skew Index and Realized Volatility Times Series

Notes: Figure 1 plot the series of the CBOE Skew Index (SKEW, left axis) and realized volatility (RV, right axis) constructed by 5-minute close prices of S&P 500 index. The sample covers from January 3, 2000 to June 28, 2022. There are 5592 observations. Daily Data of CBOE Skew Index is from https://www.cboe.com/us/indices/dashboard/skew/. Realized volatilities constructed by 5-minute close prices are from the Realized Library.

From another perspective, Figure 2 illustrates a scatter plot depicting the relationship between Realized Volatility (horizontal axis) and the CBOE Skew Index (vertical axis). The majority of data points are concentrated in the lower-left region of the plot, indicating a significant negative correlation between the two variables. This negative correlation implies that as market panic intensifies, traders incorporate their bullish expectations into option prices, leading to a decrease in the CBOE Skew Index.

Notably, the farthest data point on the right side of the plot corresponds to October 10, 2008, which coincides with the highest level of Realized Volatility observed in Figure 1. This specific point captures a period of heightened market volatility and uncertainty during the 2008 financial crisis. The scatter plot reaffirms the negative relationship between Realized Volatility and the CBOE Skew

Index, emphasizing the tendency for traders to incorporate their bullish outlook into option prices when market conditions deteriorate.

The scatter plot provides empirical evidence of the dynamic interplay between Realized Volatility and the CBOE Skew Index, highlighting the significance of market sentiment and the pricing of options. The occurrence of high SKEW values is observed in connection with both low and high values of VIX. It is worth noting that as RV rises to extreme values above 0.04, the upper limit of SKEW values decreases. This phenomenon can be attributed to the fact that during periods of plummeting stock prices, when the likelihood of a repeat crash may not be perceived as very high, RV experiences a surge. This figure reinforces the importance of considering both volatility and skewness in option pricing models to accurately capture market expectations and adequately price options under different market conditions.

The accurate modeling of volatility dynamics in the underlying asset based on the realized volatility is a crucial aspect of option valuation frameworks. A growing body of literature advocates for incorporating empirically grounded properties into option pricing models. These studies utilize observed (realized) quantities to update volatility, eliminating its latent nature. This modeling approach enables more precise measurement and forecasting of asset volatility.

Researchers, including Andersen et al. (2001a), Andersen et al. (2001b), Barndorff-Nielsen and Shephard (2002), and Corsi et al. (2013), among others, have provided both theoretical and empirical support for the development of dependable measures of realized variance derived from high-frequency intraday observations. Furthermore, recent studies by Christoffersen et al. (2015), Christoffersen et al. (2014), Huang et al. (2017), Majewski et al. (2015), and Tong and Huang (2021) illustrate that incorporating the dynamics of returns and realized variances together in option pricing models outperforms models that solely focus on optimizing returns.

Based on these insights, this study proposes an option pricing model that incorporates both realized measures and nonzero skewness. Our modeling approach allows both dynamic volatilities and skewness, and the inclusive model nests several previous discrete-time affine GARCH family models with Gaussian distributed innovations. To stress the importance of joint information of high-frequency observations and nonzero three moment distribution, the dynamic of skewness can be indirectly upgraded by the realized measures.

We observe that both realized measures and skewness play a significant role in option pricing for the SPX market across different time periods. Our newly proposed model, the Inverse Gaussian

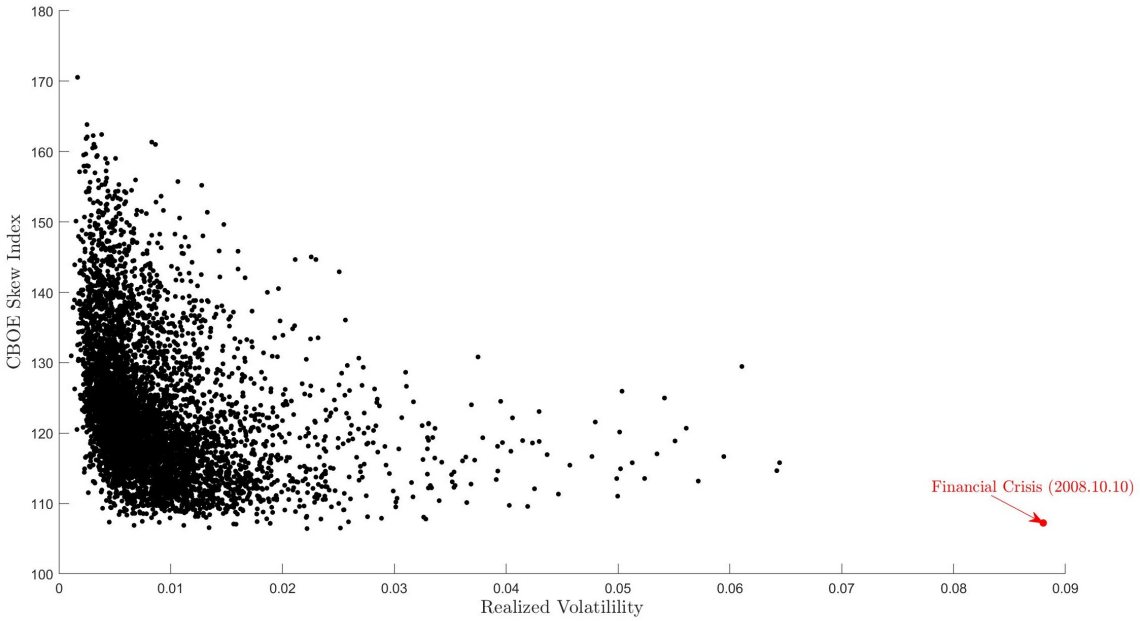


Figure 2: Skew-Volatility Negative Correlation Scatterplot

Notes: Figure 2 presents a scatter plot of realized volatility (RV) against skewness (SKEW) for the observed data. The horizontal axis represents the RV values, while the vertical axis represents the corresponding SKEW values. The scatter plot illustrates the relationship between RV and SKEW, showcasing the distribution of data points. The highlighted red-colored data point in the bottom-right corner corresponds to the highest RV value observed on October 10th, 2008, during the financial crisis. This data point serves as an exemplification of the significant RV and comparatively low SKEW during periods of heightened market distress. The dataset used in this analysis spans from January 3, 2000, to June 28, 2022, encompassing a total of 5592 observations. The daily data for the CBOE Skew Index was obtained from the official website of the Chicago Board Options Exchange (CBOE) at https://www.cboe.com/us/indices/dashboard/skew/. The realized volatilities, calculated based on 5-minute closing prices, were constructed using the realized library.

Affine Realized Volatility model, effectively captures the historical and risk-neutral distributions of S&P 500 index returns. By optimizing the IG-ARV model using a dataset comprising S&P 500 index options, we achieve superior performance compared to various popular specifications. Notably, incorporating the skewness leads to a notable improvement of 16.80% in pricing accuracy during out-of-sample testing, surpassing the modeling of Gaussian distribution with realized measures, while combining the realized measures reduces the pricing error by 30.60%, compared with only modeling the skewed distribution of daily return and volatility. These findings highlight the importance of considering the high-frequency information and the volatility smirk for effective risk hedging during market turbulence and option trading.

This paper introduces a novel option pricing model that offers flexibility in accommodating distinct overnight and intraday variance dynamics in the underlying asset price process. A key aspect of this model is its reliance on nonparametric empirical proxies to capture the dynamics of overnight and intraday variances. By incorporating these proxies, our model contributes to the discrete-time framework. Furthermore, we leverage the multivariate Edgeworth-Sargan density to derive analytical approximations for option valuation formulas, encompassing multiple nested option-pricing specifications. This analytical approach simplifies the estimation process, enables direct comparisons between nested models, and eliminates the need for simulation techniques.

The organization of the paper is as follows. In Section 2, we present the theoretical and empirical arguments underpinning the use of realized volatility and inverse Gaussian distribution when describing skewness. Section 3 introduces a novel option pricing model that is general enough to variance dynamics and skewness in the underlying return process while drawing relevant information from empirical quantities. In Section 4, we describe the physical estimation strategy and discuss the different specifications that our option pricing framework encompasses. We also discuss the estimation findings based on historical observations. Section 5 derives the option pricing formula and investigates the empirical ability to fit the risk-neutral distribution embedded in option contracts. We implement a risk-neutral optimization procedure and analyze the pricing performance of proposed model in Section 6. Section 7 concludes.

2 Realized Volatility and Skewness

2.1 Empirical Dynamics of Returns, Realized Volatilites, and Skewness

Figure 3 plots the daily time series of the S&P 500 index returns (Graph A) from January 3, 2000, to June 28, 2022. Alongside, it presents the square root of the daily realized variances (Graph B), and the skewness constructed by dividing the realized volatility of daily returns after subtracting mean return (Graph C). The graph clearly depicts periods of market instability characterized by significant fluctuations in returns and a heightened level of volatility. Notable turbulent periods include the financial market crisis of 2008-2009, and 2020. We find that the adjusted skewness of return tend to be more negative when encountering dramatic financial crisis, simutaneously to the rise of realized volatilities plotted in Graph B. Note that the adjusted skewness of return present a more volatile dynamic in recent times, especially after 2020, indicating that there exist some information that cannot be captured by the first and second moments. It's of great necessity to specify a new dynamic model concerning heteroskedasticity in skewness.

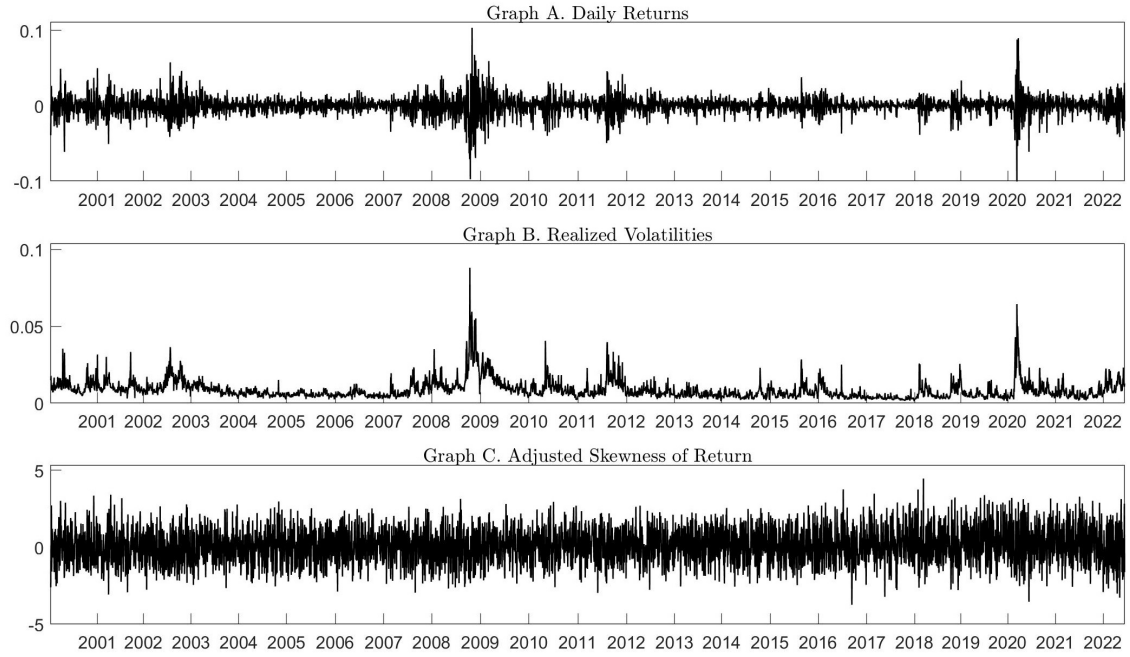


Figure 3: Time Series of Return, Realized Volatility, and Skew

Notes: Figure 3 presents the daily returns on the Standard & Poor's (S&P) 500 index, R_t (Graph A); the daily realized volatilities, $\sqrt{RV_t}$ (Graph B); and the daily RV-adjusted return skewness $(R_t - \mathbb{E}[R_t])/\sqrt{RV_t}$ (Graph C). Daily realized volatility quantities are computed from 5-minute squared returns, sourced from the Realized Library. The sample starts on January 3, 2000, and ends on June. 28, 2022.

Table 1 presents the summary statistics of daily returns and realized volatilities series. We report the annualized percentage values of mean, median, and standard deviation. Table 1 shows that there exist negative skewness (-0.1957) in the return series, while existing positive skewness in the realized volatility series (3.3007). The reported values are similar to previous literature such as Feunou and Okou (2019).

Table 1: Summary Statistics of Returns and Realized Volatilities

Summary Statistics of Returns and Realized Volatilities

	Mean (%)	Median (%)	Std. Dev. (%)	Skewness	Kurtosis
Return	6.3468	17.7244	19.1222	-0.1957	11.0480
Realized Volatility	13.6637	11.1262	9.6992	3.3007	22.6621

Notes: Table 1 presents the summary statistics for the studied series. Mean, median, and standard deviation values are annualized and in percentages. The sample starts on January 3, 2000, and ends on June. 28, 2022.

So far, we have shown the skewed property in the return series, we then further investigate the properties of the distribution of realized volatility.

2.2 Right-Skewed Distribution of Realized Volatility

Many studies have revealed that the Inverse Gaussian distribution (Johnson et al. (1995), Barndorff-Nielsen (1997b), Tweedie (1957), Barndorff-Nielsen (1997a), and so on) can effectively approximate the properties of realized volatility. Forsberg and Bollerslev (2002) find that the distribution of realized volatility, conditioned on past squared daily returns, as well as the unconditional distribution of realized volatility, can be effectively approximated by an Inverse Gaussian (IG) distribution. Similarly, Barndorff-Nielsen and Shephard (2002) show that their model-based estimation of actual volatility demonstrates a good fit, not only with the log-normal distribution but also with the inverse Gaussian distribution. Other studies, such as Barndorff-Nielsen and Shephard (2001), have verified the use of inverse Gaussian models for volatility.

The frequency distribution of realized volatility values is further examined through a histogram analysis, as depicted in Figure 4. The histogram, constructed with 70 bins, provides a comprehensive visualization of the distribution characteristics within the sample period from January 3, 2000, to June 28, 2022. The histogram showcases the varying frequencies of RV values across the observed range. Notably, the distribution exhibits a distinct skewness, indicating the presence of asymmetry in the realized volatility values. To assess the goodness of fit, an inverse Gaussian distribution is fitted to the observed data, and the resulting black curve closely aligns with the histogram's distribution. This demonstrates the efficacy of the inverse Gaussian density function in accurately capturing the distribution's shape and properties, further confirming its suitability as a modeling

framework for the RV dynamics.

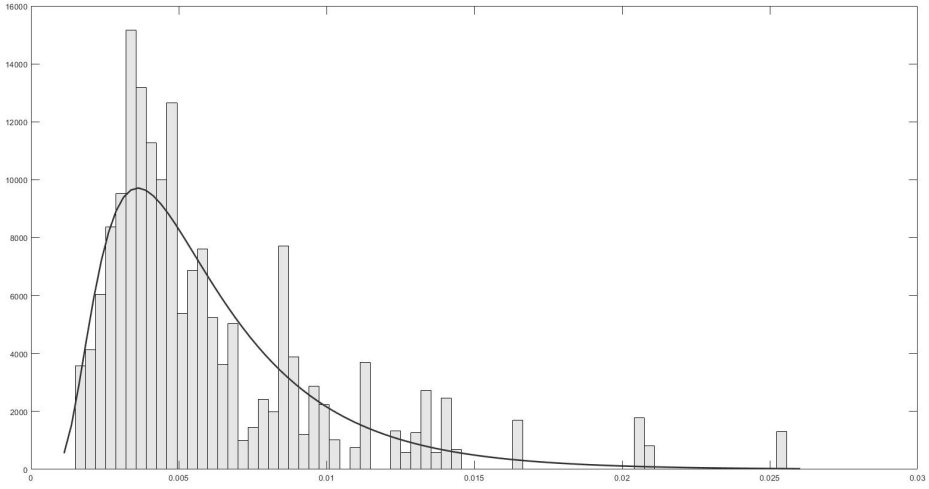


Figure 4: Inverse Gaussian Fitted Realized Volatility

Notes: Figure 4 shows the frequency histogram illustrating the distribution of realized volatility values. The horizontal axis represents the range of realized volatility values, while the vertical axis represents the frequency of occurrence. The histogram consists of 70 bins and pertains to the sample period spanning from January 3, 2000, to June 28, 2022. The bars of the histogram are depicted in a gray color scheme, visually representing the distribution's frequency characteristics. In addition, a fitted inverse Gaussian distribution curve is overlaid in black, providing a representation of the theoretical distribution that closely aligns with the observed realized volatility values. Remarkably, the histogram displays a pronounced skewness in the distribution of realized volatility values, with the inverse Gaussian density function effectively capturing and approximating this distribution.

We have observed significant skewness in both returns and realized volatility. Therefore, it is crucial to incorporate skewness into our modeling framework. In particular, the conditional skewness necessitates modeling along with the realized measures. Our subsequent objective is to develop a dynamic return model that encompasses these distinctive features.

3 The Model

This section develops an option pricing model that allows dynamic conditional asymmetries in returns. Compared with existing models in which conditional skewness is deterministically related

to conditional volatility and both undergo the same return shocks, our proposed model features a separate dynamic equation and innovations for conditional skewness to disentangle the dynamics of conditional volatility and conditional skewness. Moreover, latent variables are updated using their model-free realized measures constructed from high-frequency historical returns on the underlying asset, therefore, the model incorporates the information contained in returns (R_t), realized variances (RV_t), and realized skewness (RS_t). The flexibility that the model offers for conditional skewness as well as high-frequency information from the underlying asset may contribute to better volatility and skewness forecasts of the underlying asset, and thus to more accurate option prices.

3.1 New Dynamic Conditional Skewness Model

The objective of our model is to explicitly include dynamic conditional skewness in the option pricing framework. In traditional models, investors express preferences in terms of the expected returns and the associated variance of their portfolio holdings. However, extensive empirical research suggests that the distribution of asset returns is inadequately described solely by mean and variance parameters. This necessitates the incorporation of the third moment, conditional skewness, into the option pricing framework (see, for example, Harvey and Siddique (2000)). Holding all other variables constant, investors would have a preference for portfolios with positive skewness over those with negative skewness. This paper disentangles the conditional variance and conditional skewness of asset return. We then characterize how conditional skewness enters asset return.

3.1.1 Asset Return Process

We impose that the return on day t+1, namely that R_{t+1}, incorporates effect of conditional skewness by: 1) including rewards for accepting the risk of conditional skewness of the underlying asset; 2) innovating R_{t+1} by a skewed shock other than Gaussian distributed innovation. Specifically,

$$R_{t+1} = r + \left(\lambda_z - \frac{1}{2}\right) h_t + (\lambda_y - \xi)\, s_t + \sqrt{h_t} z_{1,t+1} + \eta_1 y_{1,t+1}, \tag{1}$$

where r is the risk-free rate, h_t and s_t are state variables that depict conditional variance and conditional skewness. The first innovation, z_{t+1}, is a standard normal distributed shock, conditional on the information set on day t. We opt for the inverse Gaussian conditional distribution as the second innovation, y_{t+1}. We consider the one-parameter inverse Gaussian distribution, $y_{t+1}|\mathcal{F}_t \sim IG(s_t)$, where s_t is time-varying degree of freedom parameter. Then the conditional expectation of

R_{t+1},

$$\mathbb{E}_t\left(R_{t+1}\right) = r + \left(\lambda_z - \frac{1}{2}\right) h_t + \left(\lambda_y - \xi + \eta_1\right) s_t, \tag{2}$$

is a linear combination of state variables h_t and s_t. The conditionally expected total return is

$$\mathbb{E}_t\left(\exp\left(R_{t+1}\right)\right) = \exp\left(r + \lambda_z h_t + \lambda_y s_t\right), \tag{3}$$

which in turn ensures that λ_z and λ_y can be viewed as compensation for diffusive volatility and skewness exposure, respectively. The skewness compensator parameter, ξ, in our model is itself a particular function of other parameters:

$$\xi = \sqrt{1 - 2\eta_1} - 1. \tag{4}$$

3.1.2 Conditional Variance and Conditional Skewness

From the model above, it is relatively straightforward to derive the conditional moments based on properties of normal distribution and inverse Gaussian distribution. We have

$$Skew_t\left(R_{t+1}\right) = \frac{3\eta_1^3 s_t}{\sqrt{Var_t\left(R_{t+1}\right)}}, \tag{5}$$

implying that the conditional skewness is entirely determined by s_t. Therefore, the dynamic of s_t fully captures conditional skewness.

Additionally, conditional variance of asset returns is comprised of two components, originating respectively from shocks following a normal distribution and shocks following an inverse Gaussian distribution. These components adhere to the equation depicted as follows:

$$Var_t\left(R_{t+1}\right) = h_t + \eta_1^2 s_t. \tag{6}$$

In practice, it is clear that h_t and s_t are related, but not equal to the conditional variances and skewness of R_{t+1}. Nevertheless, we can still formulate the dynamics of these two state variables with corresponding realized measures.

3.1.3 The Gaussian Shocked Volatility Dynamic

We are prepared to outline the dynamics of Gaussian part of conditional variances. Following previous studies (see, for example, Feunou and Okou (2019), Christoffersen et al. (2014)), we postulate the subsequent recursive dynamics under GARCH framework:

$$h_{t+1} = \omega + \beta h_t + \alpha\left(z_{2,t+1} - \gamma\sqrt{h_t}\right)^2, \tag{7}$$

where innovation term $z_{2,t+1}$ is standard normal distributed, satisfying $Corr\left(z_{1,t+1}, z_{2,t+1}\right) = \rho$. $\alpha\left(z_{2,t+1} - \gamma\sqrt{h_t}\right)^2$ is a non-central Gamma distributed innovation term, capturing the asymmetric shock on variance. Specifically, γ is the parameter that measures the leverage effect. From equation (7), we obtain unconditional conditional gaussian-distributed variance, and volatility persistence:

$$\mathbb{E}_t\left(\bar{h}\right) = \frac{\omega + \alpha}{1 - \text{Persistance}_h}, \tag{8}$$

where $\text{Persistance}_h = \beta + \alpha\gamma^2$. The observed realized measures, $\widetilde{RV_{t+1}}$, and h_t are supposed to be linked as

$$\widetilde{RV_{t+1}} = h_t + \sigma\left[\left(z_{2,t+1} - \gamma\sqrt{h_t}\right)^2 - \left(1 + \gamma^2 h_t\right)\right], \tag{9}$$

where $\mathbb{E}_t\left(\widetilde{RV_{t+1}}\right) = h_t$. The given specification of equation (7) and equation (9) indicates that h_t follow univariate first-order autoregressive (AR(1)) processes.

The subsequent moments implied by the dynamic model will be crucial for comprehending the models' effectiveness in fitting both returns, realized variances and options. To begin with, consider the expectation of variance one day ahead, represented as follows:

$$\mathbb{E}_t\left(h_{t+1}\right) = (\omega + \alpha) + \left(\beta + \alpha\gamma^2\right) h_t. \tag{10}$$

Moving on, the conditional variance of Gaussian shocked variance implied by the model can be derived as follows:

$$Var_t\left(h_{t+1}\right) = 2\alpha^2\left(1 + 2\gamma^2 h_t\right), \tag{11}$$

and

$$Var\left(\widetilde{RV_{t+1}}\right) = 2\sigma^2\left(1 + 2\gamma^2 h_t\right) \tag{12}$$

with a similar structure.

Lastly, the conditional covariances between returns, Gaussian shocked conditional variances, scaled realized variances can be expressed as:

$$Cov_t\left(R_{t+1}, h_{t+1}\right) = -2\alpha\gamma\rho h_t, \tag{13}$$

$$Cov_t\left(RV_{t+1}, R_{t+1}\right) = -2\rho\gamma\sigma h_t. \tag{14}$$

3.1.4 The Inverse Gaussian Shocked Skewness Dynamic

We then model conditional skewness under the GARCH framework following the innovation structure of Christoffersen et al. (2006) and Feunou and Tédongap (2012). The dynamic of s_t

$$s_{t+1} = w + bs_t + cy_{2,t+1} + a\frac{h_t^2}{y_{2,t+1}}, \tag{15}$$

where $y_{2,t+1}$ has an inverse Gaussian conditional distribution with degrees of freedom parameter $\delta_t = \frac{h_t^2}{\eta_2}$, conditional on the available information set on day t. The one-day-ahead conditional expectation can be derived as

$$\mathbb{E}_t\left(s_{t+1}\right) = \left(w + a\eta_2^4\right) + bs_t + \left(\frac{c}{\eta_2^2} + a\eta_2^2\right) h_t, \tag{16}$$

and the unconditional expectation is

$$\mathbb{E}\left(\bar{s}\right) = \frac{\left(w + a\eta_2^4\right) + \left(\frac{c}{\eta_2^2} + a\eta_2^4\right)\mathbb{E}\left(\bar{h}\right)}{1-b}, \tag{17}$$

implying that the unconditional expectation of s is a linear transformation of that of h, where $\mathbb{E}\left(\bar{h}\right)$ is given by equation (8).

In order to keep the consistency of the framework, we impose $\mathbb{E}_t\left(\widetilde{RS_{t+1}}\right) = s_t$. The measurement equation then helps us build the relationship between realized measures and latent variable s_ts:

$$\widetilde{RS_{t+1}} = s_t + d\left[cy_{2,t+1} + a\frac{h_t^2}{y_{2,t+1}} - \left(\left(\frac{c}{\eta_2^2} + a\eta_2^2\right) h_t + a\eta_2^4\right)\right]. \tag{18}$$

We refer to this general specification as the Dynamic Conditional Skewness (DCS) model. It is noteworthy that the inclusion of the realized variance (RV) and realized skewness (RS) as factors in the variance and skewness dynamic yields notable benefits in terms of both enhanced modeling accuracy for current spot volatility and volatility term structure. Furthermore, this approach enables the adoption of more flexible functional forms that effectively capture the skewness of the underlying assets, which are pivotal elements in option valuation.

3.2 Nested Models

3.2.1 The Inverse Gaussian Affine Realized Volatility (IG-ARV) Model

Suppose there is only an inverse Gaussian distributed innovation in the return equation, we have

$$R_{t+1} = r + vh_t^{RV} + \eta y_{1,t+1}, \tag{19}$$

where r is the risk-free rate, and h_t^{RV} is the conditional variance of return. Parameters v and η describe the distribution of R_{t+1}, conditional on the information at time t. The parameter η controls the spread or dispersion of the distribution. A negative η indicates that the distribution is skewed towards the right. The inverse Gaussian distribution is used to model the daily returns, where $y_{1,t+1}$s are independent IG-distributed shocks with degrees of freedom parameter δ_{t+1}, which is determined by the function h_{t+1}^{RV}/η^2.

Then we specify the dynamic of h_t^{RV}. Inverse Gaussian models for volatility were suggested by Barndorff-Nielsen and Shephard (2001). We consider the following GARCH-type dynamic for conditional variance.

$$h_{t+1}^{RV} = w + bh_t^{RV} + cy_{2,t+1} + a\frac{\left(h_t^{RV}\right)^2}{y_{2,t+1}}, \tag{20}$$

where $y_{2,t+1}$ satisfies inverse Gaussian distribution with degree of freedom δ_{t+1}, and $Cov_t(y_{1,t+1}, y_{2,t+1}) = 0$ for all t. Taking conditional expectation, we have

$$\mathbb{E}_t\left[h_{t+1}^{RV}\right] = w + \eta^4 a + \left(b + \frac{c}{\eta^2} + a\eta^2\right) h_t^{RV}, \tag{21}$$

and the unconditional variance can be derived as

$$\mathbb{E}\left[\bar{h}\right] = \frac{w + \eta^4 a}{1 - \text{Persistance}}, \tag{22}$$

where

$$\text{Persistence} = b + \frac{c}{\eta^2} + a\eta^2 \tag{23}$$

denotes the volatility persistence set in the IG-ARV model. Additionally, the conditional variance of variance is given by

$$Var_t[h_{t+1}^{RV}] = \left(\frac{c^2}{\eta^2} - 2ca\eta^2 + a\eta^6\right) h_t^{RV} + 2a^2\eta^8, \tag{24}$$

which also satisfies an affine structure. Finally, the observed RV_{t+1} and its conditional expectation h_t^{RV} are assumed to be linked as follows:

$$RV_{t+1} = h_t^{RV} + d\left[cy_{2,t+1} + a\frac{\left(h_t^{RV}\right)^2}{y_{2,t+1}} - \left(\left(\frac{c}{\eta^2} + a\eta^2\right) h_t^{RV} + a\eta^4\right)\right]. \tag{25}$$

Note that we hold the property that $\mathbb{E}_t\left[RV_{t+1}\right] = h_t^{RV}$, and the conditional variance of RV_{t+1} can be derived as

$$Var_t[RV_{t+1}] = d^2\left(\left(\frac{c^2}{\eta^2} - 2ca\eta^2 + a\eta^6\right) h_t^{RV} + 2a^2\eta^8\right). \tag{26}$$

3.2.2 The Affine Realized Volatility (ARV) Model

Wald (1947) has shown that y will converge to an asymptotic normal distribution as $\delta \to \infty$. Therefore, for a finite h, we can let $\eta \to 0$ to obtain a normal distribution. Therefore, the IG-ARV model exhibits a close relationship with conventional GARCH processes. This connection becomes apparent by considering the parameterization and taking the limit as η approaches zero:

$$\begin{aligned}
&v + \frac{1}{\eta} = \lambda, \quad w + a\eta^4 = \omega + \alpha, \\
&b + \frac{c}{\eta^2} + a\eta^2 = \beta + \alpha\gamma^2, \quad a\eta^4 = \alpha, \\
&\frac{c}{\eta} - \eta^3 a = -2\alpha\gamma, \quad d^2\left(\frac{c}{\eta} - \eta^3 a\right)^2 = 4\sigma^2\gamma_2^2, \\
&d^2 \cdot 2a^2\eta^8 = 2\sigma^2.
\end{aligned} \tag{27}$$

then the IG-ARV model converges to the Affine Realized Volatility (ARV) model in Christoffersen et al. (2014). The ARV model considers normal distribution shocks $\varepsilon_{1,t+1}$ and $\varepsilon_{2,t+1}$:

$$\begin{aligned}
R_{t+1} &= r + \lambda h_t^{\mathrm{RV}} - \frac{1}{2}h_t^{\mathrm{RV}} + \sqrt{h_t^{\mathrm{RV}}}\varepsilon_{1,t+1}, \\
\mathrm{RV}_{t+1} &= h_t^{\mathrm{RV}} + \sigma\left[\left(\varepsilon_{2,t+1} - \gamma_2\sqrt{h_t^{\mathrm{RV}}}\right)^2 - \left(1 + \gamma_2^2 h_t^{\mathrm{RV}}\right)\right], \\
h_{t+1}^{\mathrm{RV}} &= \omega_2 + \beta_2 h_t^{\mathrm{RV}} + \alpha_2\left(\varepsilon_{2,t+1} - \gamma_2\sqrt{h_t^{\mathrm{RV}}}\right)^2.
\end{aligned} \tag{28}$$

It can be shown that the parameterization presented in Equation (27) ensures that the first two conditional moments of the IG-ARV model align with those of the ARV model.

Proof: See Appendix A. ■

3.2.3 The Inverse Gaussian GARCH (IG GARCH) Model

If we close the dynamic of realized measures, and keep only one IG-distributed shock, $y_{1,t+1}$, then we obtain the Inverse Gaussian GARCH (IG GARCH) Model (Christoffersen et al. (2006)):

$$r_{t+1} = r + vh_{t+1} + \eta y_{t+1}, \tag{29}$$

$$h_{t+1} = w + bh_t + cy_t + a\frac{h_t^2}{y_t}, \tag{30}$$

where the degree of freedom of y_{t+1} is $\delta_{t+1} = \frac{h_{t+1}}{\eta^2}$. The IG GARCH model involves integrating an inverse Gaussian distribution with a volatility dynamic that follows a GARCH(1,1) model.

3.2.4 The Heston-Nandi GARCH (HNG) Model

As the degrees of freedom parameter tends towards infinity, the standardized inverse Gaussian distribution approaches the standard Gaussian distribution. By closing the measurement equation of the ARV model, we can obtain the Heston-Nandi GARCH (HNG) model (Heston and Nandi (2000)):

$$\begin{aligned} R_{t+1} &= r + \lambda h_t - \frac{1}{2} h_t + \sqrt{h_t}\varepsilon_{t+1}, \\ h_{t+1} &= \omega_2 + \beta_2 h_t + \alpha_2 \left(\varepsilon_{t+1} - \gamma\sqrt{h_t}\right)^2. \end{aligned} \tag{31}$$

The relationship between the HNG model and the IG GARCH model is parallel to the relationship between the ARV model and the IG-ARV model. As expected, The parameterization (except the last equation) presented in equation (27) establishes a congruence between the first two conditional moments of the IG GARCH model, and those of the HNG model.

4 Physical Measure Estimation

4.1 Optimization Method

In the computation of the conditional quasi-likelihood function, the observation vector for day $t+1$ is factored into its constituent components by leveraging both marginal densities and a Gaussian copula function. Formally expressed, the likelihood contribution of the observation vector for day $t+1$ is:

$$\begin{aligned} f_t\left(R_{t+1}, \widetilde{\mathrm{RV}_{t+1}}, \widetilde{\mathrm{RS}_{t+1}}\right) =& f_{r,t}\left(R_{t+1}\right) f_{h,t}\left(\widetilde{\mathrm{RV}_{t+1}}\right) f_{s,t}\left(\widetilde{\mathrm{RS}_{t+1}}\right) \\ & \times c_t\left(F_{r,t}\left(R_{t+1}\right), F_{h,t}\left(\widetilde{\mathrm{RV}_{t+1}}\right), F_{s,t}\left(\widetilde{\mathrm{RS}_{t+1}}\right)\right), \end{aligned} \tag{32}$$

where $f_{r,t}\left(R_{t+1}\right), f_{h,t}\left(\widetilde{\mathrm{RV}_{t+1}}\right)$, and $f_{s,t}\left(\widetilde{\mathrm{RS}_{t+1}}\right)$ are the conditional marginal densities associated with returns, realized variance, and realized skewness, respectively. Correspondingly, $F_{r,t}\left(R_{t+1}\right), F_{h,t}\left(\widetilde{\mathrm{RV}_{t+1}}\right)$, and $F_{s,t}\left(\widetilde{\mathrm{RS}_{t+1}}\right)$ are the conditional marginal cumulative distribution functions (CDFs) related to these observations. The function $c_t\left(v, v_h, v_s\right)$ specifies the density associated with a Gaussian copula. Note that the innovation term in the measurement equation of

h_{t+1} is Gaussian distributed, we have

$$f_{h,t}\left(\widetilde{\mathrm{RV}_{t+1}}\right)=\frac{\exp\left(-\frac{\left(\widetilde{\mathrm{RV}_{t+1}}-h_t\right)^2}{2\,\mathrm{var}_t\left[\widetilde{\mathrm{RV}_{t+1}}\right]}\right)}{\sqrt{2\pi\,\mathrm{var}_t\left[\widetilde{\mathrm{RV}_{t+1}}\right]}}, \tag{33}$$

$$F_{h,t}\left(\widetilde{\mathrm{RV}_{t+1}}\right)=\Phi\left(\frac{\widetilde{\mathrm{RV}_{t+1}}-h_t}{\sqrt{\mathrm{var}_t\left[\widetilde{\mathrm{RV}_{t+1}}\right]}}\right), \tag{34}$$

where $\Phi(\cdot)$ is the Cumulative Distribution Function (CDF) of a standard normal distribution, while the variance values $\mathrm{var}_t\left[\widetilde{\mathrm{RV}_{t+1}}\right]$ are elaborated in equation (12). We also obtain

$$\begin{aligned}
f_{s,t}\left(\widetilde{RS_{t+1}}\right)&=f_{s,t}\left(y_{2,t+1}\right)\left|\frac{dy_{2,t+1}}{d\widetilde{RS_{t+1}}}\right|\\
&=\frac{\delta_t}{\sqrt{2\pi\left(\frac{-B_1+\sqrt{\Delta}}{2A_1}\right)^3}}\exp\left\{-\frac{1}{2}\left(\sqrt{\frac{-B_1+\sqrt{\Delta}}{2A_1}}-\frac{\delta_t}{\sqrt{\frac{-B_1+\Delta}{2A_1}}}\right)^2\right\}\cdot\left|\frac{1}{2A_1}\left(1-\frac{B_1}{\sqrt{\Delta}}\right)\right|\\
&=\frac{\delta_t}{\sqrt{\Delta\pi\left(\frac{-B_1+\sqrt{\Delta}}{A_1}\right)}}\exp\left\{-\frac{1}{2}\left(\sqrt{\frac{-B_1+\sqrt{\Delta}}{2A_1}}-\frac{\delta_t}{\sqrt{\frac{-B_1+\Delta}{2A_1}}}\right)^2\right\},
\end{aligned} \tag{35}$$

and the conditional CDF of $\widetilde{RS_{t+1}}$ is

$$F_{s,t}\left(\widetilde{RS_{t+1}}\right)=\Phi\left(\frac{-\delta_t}{\sqrt{\frac{-B_1+\sqrt{\Delta}}{2A_1}}}+\sqrt{\frac{-B_1+\sqrt{\Delta}}{2A_1}}\right)+e^{2\delta_t}\Phi\left(\frac{-\delta_t}{\sqrt{\frac{-B_1+\sqrt{\Delta}}{2A_1}}}-\sqrt{\frac{-B_1+\sqrt{\Delta}}{2A_1}}\right), \tag{36}$$

where

$$A_1=dc,\quad B_1=s_t-d\left(\left(\frac{c}{\eta_2^2}+a\eta_2^2\right)h_t+a\eta_2^4\right)-\widetilde{RS_{t+1}},\quad C_1=dah_t^2,\quad \Delta=B_1^2-4A_1C_1. \tag{37}$$

We impose that the asset return is innovated by both normal distribution and a skewed distribution (IG distribution), respectively. It is important to recognize that the exact conditional marginal density of returns, denoted as $f_{r,t}\left(R_{t+1}\right)$, manifests as a convolution of a normal density and an inverse Gaussian density. This function is lacking a closedform representation.

However, the conditional characteristic function $\varphi_{r,t}(v)$ for R_{t+1} is obtainable in a closed-form expression. Explicitly, this function is given as:

$$\begin{aligned}\varphi_{r,t}(v) &\equiv \mathrm{E}_t\left[e^{ivR_{t+1}}\right] \\ &= \exp\left(ivr + iv\lambda_z + \left(iv\left(\lambda_y - \xi\right) + 1 - \sqrt{1 - 2\eta_1 iv}\right)s_t\right),\end{aligned} \tag{38}$$

where i stands for the imaginary unit. This characteristic function provides a comprehensive mathematical framework to capture the dynamics of conditional returns, especially in scenarios where the marginal density fails to offer a tractable form. Thus, we exploit Fourier inversion formulas to compute the quantities of interest:

$$F_{r,t}\left(R_{t+1}\right) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}\left[e^{-i\nu R_{t+1}}\varphi_{r,t}(v)\right]}{\nu}d\nu, \tag{39}$$

$$f_{r,t}\left(R_{t+1}\right) = \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[e^{-i\nu R_{t+1}}\varphi_{r,t}(\nu)\right]d\nu. \tag{40}$$

Additionally, the Gaussian copula function is formally specified as

$$c_t\left(v, v_h, v_s\right) = \frac{1}{\sqrt{|\mathrm{CM}_t|}}\exp\left(-\frac{1}{2}\left(v, v_h, v_s\right)\left(\mathrm{CM}_t^{-1} - I_3\right)\begin{pmatrix} v \\ v_h \\ v_s \end{pmatrix}\right),$$

where CM_t denotes the conditional correlation matrix for $\left(R_{t+1}, \widetilde{\mathrm{RV}_{t+1}}, \widetilde{\mathrm{RS}_{t+1}}\right)$, represented as

$$\mathrm{CM}_t = \begin{bmatrix} 1 & \rho_t & 0 \\ \rho_t & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Here, ρ_t is defined as the conditional correlation between R_{t+1} and $\widetilde{\mathrm{RV}_{t+1}}$ and is expressed as

$$\rho_t = \frac{\mathrm{cov}_t\left(R_{t+1}, \widetilde{\mathrm{RV}_{t+1}}\right)}{\sqrt{\mathrm{var}_t\left[R_{t+1}\right]\mathrm{var}_t\left[\widetilde{\mathrm{RV}_{t+1}}\right]}},$$

where $\mathrm{cov}_t\left(R_{t+1}, \widetilde{\mathrm{RV}_{t+1}}\right)$, $\mathrm{var}_t\left[R_{t+1}\right]$, and $\mathrm{var}_t\left[\widetilde{\mathrm{RV}_{t+1}}\right]$ are elaborated in equation (12). Subsequently, the log-likelihood function is computed via

$$\ln L^P = \sum_{t=1}^{T-1}\ln\left(f_t\left(R_{t+1}, \widetilde{\mathrm{RV}_{t+1}}, \widetilde{\mathrm{RS}_{t+1}}\right)\right). \tag{41}$$

This log-likelihood metric serves as an aggregate measure of the fit between the observed data and the model, encapsulating the combined contributions from returns and both realized variances and realized skewness.

4.2 Estimation Results

Table 2 presents the outcomes obtained through maximum likelihood estimation based on historical data using the physical distribution. The data covers the period from January 3, 2000, to December 31, 2019. The parameters ω and w are estimated by aiming for the unconditional sample variance, which can be implied from the unconditional variance equation of each model.

In all considered models, the market price of risk is ascertained to be positive, a characteristic that can be directly inferred from the parameters λ in the ARV model and the HNG model. Similarly, in the IG model and the IG-ARV model, the equation $v + \frac{1}{\eta}$ yields equivalent outcomes, further indicating a positive market price of risk.

The results imply properties of skewness and volatility. The estimates of η in both the IG-ARV model and the IG model are negative and statistically significant, which implies that the inverse Gaussian distribution to display negative skewness. Negative ηs filtered by these models also explain the negative relationship between skewness and volatility, as in equation (**??**). When considering an offsetting positive value of vh_{t+1} in the return equation, the dynamic nature of the system can produce both positive and negative returns that are centered around a small positive value. Moreover, the positive and statistically significant γ in the HNG model and the ARV model indicate the leverage effect. The parameter d in the IG-ARV model is also significant and statistically positive, indicating the necessity and utility of the modelling of measurement equation.

The likelihood values provide us with the means to evaluate the effectiveness of the different specifications being considered. The loglikelihood value of the IG-ARV model is 67428, exceeding the other models. Therefore, the IG-ARV model exhibits better ability while fitting the observed data. Note that loglikelihood values of models considering realized measures (the IG-ARV model and the ARV model) are generally larger than the models only considering daily close-to-close return dynamics. Additionally, the second-to-last row of Table 2 reports the volatility persistence of each model. All models exhibit variance persistence levels above 0.96.

Table 2: Estimation on Historical Returns and Realized Variance

Estimation on Historical Returns and Realized Variance

Parameters	IG-ARV	IG	ARV	HNG
v	796.21651	1047.7570		
	(104.9118)	(24.6388)		
λ			0.7330	1.49E-06
			(0.0146)	(0.0040)
a	2.35E+04	5.30E+06		
	(262.9164)	(7.15E+05)		
α			4.82E-06	4.47E-06
			(2.40E-08)	(4.72E-09)
b	-0.1964	-10.7152		
	(0.0285)	(0.9741)		
β			3.96E-06	0.7718
			(1.66E-07)	(0.0022)
c	1.99E-06	6.25E-06		
	(1.38E-08)	(2.44E-07)		
γ			454.2649	211.3512
			(1.2055)	(0.8895)
d	1.1363			
	(0.0751)			
σ			9.15E-06	
			(1.43E-07)	
η	-0.0013	-9.54E-04		
	(5.93E-05)	(2.24E-05)		
$\pi^{\mathbb{P}}$	0.9602	0.9705	0.9948	0.9715
$\mathcal{L}$	67428	16447	57616	16384

Notes: The estimation results for the five models using maximum likelihood estimation are presented in Table 2. The dataset comprises daily returns and realized measures of the Standard & Poor's (S&P) 500 index, covering the period from January 3, 2000, to December 31, 2019. The Inverse Gaussian Affine Realized Volatility (IG-ARV) model and the Affine Realized Volatility (ARV) model are estimated using close-to-close returns and realized variances. The Inverse Gaussian GARCH (IG) model and the Heston-Nandi GARCH (HNG) model are estimated using close-to-close returns. The estimated parameters are reported, with their respective standard errors denoted in parentheses. In order to facilitate model comparison, the last row displays the log likelihood value, while the second-to-last row presents the measure of volatility persistence, calculated based on the specific characteristics of each model.

5 Option Pricing Procedures

5.1 Moment Generating Function (MGF)

In order to obtain the characteristic function of the process, we follow previous literature concerning closed-form option pricing formula (see, for example, Christoffersen et al. (2006); Christoffersen et al. (2014), and so on) and derive the Moment Generating Function (MGF). We show that the moment generating function is affine, which satisfites

$$\begin{aligned}\Psi_{t,t+M}(u) &\equiv \mathbb{E}_t\left[\exp\left(u\sum_{j=1}^{M+1}R_{t+j}\right)\right] = \mathbb{E}_t\left[\mathbb{E}_{t+1}\left[\exp\left(u\sum_{j=1}^{M+1}R_{t+j}\right)\right]\right] \\ &= \exp\left(A(u,D(u,M),F(u,M))h_{t+1} + B(u,D(u,M),F(u,M))s_t + C(u,D(u,M),F(u,M)) + G(u,M)\right),\end{aligned} \tag{42}$$

the recursive rule can be written as

$$\begin{aligned}D(u,M+1) &= A(u,D(u,M),F(u,M)),\\ F(u,M+1) &= B(u,D(u,M),F(u,M)),\\ G(u,M+1) &= C(u,D(u,M),F(u,M)) + G(u,M).\end{aligned} \tag{43}$$

where

$$\begin{aligned}D(u,1) &= A(u,0,0)\\ F(u,1) &= B(u,0,0),\\ G(u,1) &= C(u,0,0).\end{aligned} \tag{44}$$

See Appendix B for the explicit expression for the terms

Proof: See Appendix B. ■

5.2 Risk Neutralization

The dynamic conditional skewness (DCS) model under $\mathbb{Q}$ measure is given by

$$R_{t+1} = r - \frac{1}{2}h_t - \left(\sqrt{1-2\eta_1^* C^2}-1\right)s_t^* + \sqrt{h_t}z_{1,t+1}^* + \eta_1^* y_{1,t+1}^*, \tag{45}$$

$$h_{t+1} = \omega + \beta h_t + \alpha\left(z_{2,t+1}^* - \gamma^*\sqrt{h_t}\right)^2, \tag{46}$$

$$s_{t+1}^* = w^* + bs_t^* + c^* y_{2,t+1}^* + a^*\frac{(h_t^*)^2}{y_{2,t+1}^*}, \tag{47}$$

$$\widetilde{RV_{t+1}} = h_t^* + \sigma\left[\left(z_{2,t+1}^* - \gamma^*\sqrt{h_t}\right)^2 - \left(1 + \gamma^{*2}h_t^{*2}\right)\right], \tag{48}$$

$$\widetilde{RS_{t+1}} = s_t + d\left[cy_{2,t+1} + a\frac{h_t^2}{y_{2,t+1}} - \left(\left(\frac{c}{\eta_2^2} + a\eta_2^2\right)h_t + a\eta_2^4\right)\right]. \tag{49}$$

where

$$\eta_1^* = \eta_1 C, \quad w^* = \frac{w}{\sqrt{C}}, \quad c^* = \frac{cD}{\sqrt{C}}, \quad a^* = \frac{a}{D\sqrt{C}\left(1 + \sigma^2\left(\gamma^{*2} - \gamma^2\right)\right)^2}.$$

The mapping relationship of state variables and innovations between $\mathbb{P}$ measure and $\mathbb{Q}$ measure is shown in Appendix C.

Proof: See Appendix C. ■

5.3 Option Pricing Formula

By applying Fourier inversion on the risk-neutral conditional characteristic function, we obtain the following proposition.

Proposition 1 *At time t, the price of a European call option matures at time $t + M$ is:*

$$C(t, M) = S_t P_1(t, M) - \exp(-rM)XP_2(t, M), \tag{50}$$

where

$$\begin{aligned} P_1(t, M) &= \frac{1}{2} + \int_0^{+\infty} \Re\left[\frac{\Psi_{t,t+M}^{Q}(1 + iu)\exp\left(-rM - iu\log\left(\frac{X}{S_t}\right)\right)}{\pi iu}\right] du, \\ P_2(t, M) &= \frac{1}{2} + \int_0^{+\infty} \Re\left[\frac{\Psi_{t,t+M}^{Q}(iu)\exp\left(-iu\log\left(\frac{X}{S_t}\right)\right)}{\pi iu}\right] du, \end{aligned} \tag{51}$$

where $\Psi_{t,t+M}^{Q}$ can by replacing parameters in $\Psi_{t,t+M}$ based on results in subsection 5.2.

5.4 Option Data

We utilize European options that are based on the S&P 500 index. The data we analyze covers the time frame from January 6, 2016, to December 18, 2019. In line with literature for researching the option skewness, we apply a number of standard filters to the data. We only keep European call options sample, and keep observations on Wednesdays. The option dataset consist of option contracts on 205 Wednesdays. Following Christoffersen et al. (2006) and Bakshi et al. (1997), we

restrict attention to option contracts with maturities between 7 and 180 days, and price quotes lower than \$3/8 are not included. Quotes not satisfying the arbitrage restriction

$$C(t,\tau) \geq \max(0, S(t) - K, S(t) - \bar{D}(t,\tau) - KB(t,\tau))$$

are taken out of the sample as well. Finally, our option dataset contain 173158 observations.

Table 3 gives an overview of the full-sample data, sorting by moneyness (S/K) and days-to-maturity (DTM). Options are divided into out-of-the-money ones ($S/K < 0.97$), at-the-money ones ($0.97 \leqslant S/K < 1.03$), and in-the-money ones ($S/K \geqslant 1.06$) by moneyness. The proposed moneyness and maturity classifications produce 36 categories for which the empirical results will be reported. Option prices are reported with notation \$, bid-ask spread are reported in the parentheses, and the number of contracts are reported in the braces. The average call option price for the full sample is \$43.62 and the average bid-ask spread is 0.93. Nearly 41.18% contracts lies in the group with DTM < 30, with a short day-to-maturity. With the DTM increases, the number of contracts decreases, while the average price increases. Bid-ask spread is the smallest for options with 30-60 days to maturity From the perspective of moneyness, 55.74% observations are ATM options. Average price and bid-ask spread increase with moneyness, implying that the intrinsic value of ITM calls and the liquidity of OTM calls.

Our data is divided into two periods: an in-sample period and an out-of-sample period. We apply the in-sample period to estimate the models while applying the results to explore the pricing error in the out-of-sample period. The in-sample period spans from January 6, 2016 to December 26, 2018. The out-of-sample period spans from January 2, 2019 to December 18, 2019. The recognizable volatility smirk can be observed in Table 4. When examining deep in-the-money call options with less than 30 days until maturity, the average implied volatility is 28.77%, whereas the corresponding out-of-the-money call options have implied volatilities of less than 17.53%. However, as maturity increases, the smirk tends to flatten out. For options with more than 150 days until maturity, deep in-the-money calls have average implied volatility of around 20%, while out-of-the-money calls have average implied volatility nearly 11%.

The observed characteristics can also be shown in Figure 5. Figure 5 is generated with four subplot. Subplot (a) represents the implied volatility (IV) smirk for in-sample options with a maturity period (T) of less than 30 days. Subplot (b) displays the IV smirk for in-sample options with a maturity period of 30 days or more. Subplot (c) presents the IV smirk for out-of-sample

Table 3: Sample Properties of S&P 500 Index Options

	Moneyness	Days-to-Maturity						
	S/K	<30	30-60	60-90	90-120	120-150	>150	Subtotal
OTM	<0.94	$1.0140	$1.8322	$3.9381	$5.6319	$7.1442	$10.2967	$3.5860
		(0.2583)	(0.3002)	(0.3923)	(0.4572)	(0.5092)	(0.6382)	(0.3712)
		{2175}	{8062}	{3873}	{2478}	{1622}	{1122}	{19332}
	0.94-0.97	$2.5001	$5.5174	$12.7058	$20.1375	$29.8769	$40.1327	$7.6388
		(0.2583)	(0.3368)	(0.5117)	(0.6459)	(0.7677)	(0.9850)	(0.3765)
		{11357}	{16468}	{4779}	{2290}	{970}	{581}	{36445}
ATM	0.97-1.00	$10.9120	$22.6246	$38.4705	$53.4278	$66.0420	$76.0613	$20.6123
		(0.4000)	(0.5657)	(0.7622)	(0.9278)	(1.0203)	(1.1877)	(0.5278)
		{29137}	{19311}	{5115}	{2379}	{962}	{553}	{57457}
	1.00-1.03	$48.0328	$63.5765	$80.0355	$53.4278	$66.0420	$76.0613	$20.6123
		(1.2301)	(1.0227)	(1.1039)	(1.2554)	(1.2652)	(1.4881)	(1.1556)
		{20769}	{12672}	{3317}	{1396}	{583}	{320}	{39057}
ITM	1.03-1.06	$107.2964	$118.3561	$135.3762	$149.4875	$160.9370	$171.8856	$118.4182
		(3.2344)	(1.9713)	(1.5388)	(1.7240)	(1.9026)	(2.0535)	(2.4648)
		{5679}	{4773}	{1482}	{566}	274	{170}	{12944}
	>=1.06	$234.9214	$258.9795	$291.7120	$341.1983	$354.8294	$367.7687	$275.6942
		(5.3573)	(3.9288)	(3.3761)	(3.4247)	(3.5783)	(3.5153)	(4.1489)
		{2188}	{2911}	{1262}	{754}	{474}	{334}	{7923}
Subtotal		$34.6327	$41.5437	$55.8299	$67.0211	$354.8294	$367.7687	$43.6152
		(0.9928)	(0.8209)	(0.9111)	(1.0270)	(1.1274)	(1.2807)	(0.9306)
		{71305}	{64197}	{19828}	{9863}	{4885}	{3080}	{173158}

Notes: Table 3 reports average option price, bid-ask spread and number of contracts. It consist of the average quoted bid-ask mid-point price and the average effective bid-ask spread (obtained by subtracting the ask price from the bid-ask mid-point). These values are presented in parentheses. Additionally, the total number of observations for each moneyness-maturity category is shown in braces. The data used for analysis spans January 6, 2016, to December 18, 2019, resulting in a total of 173,158 call options. The summary statistics are derived from the daily information of the last quote for each option contract, obtained prior to 3:00 p.m. CST. In this context, S represents the spot S&P 500 index level, and K denotes the exercise price. The terms "OTM", "ATM", and "ITM" indicate out-of-the-money, at-the-money, and in-the-money options, respectively.

Table 4: Implied Volatility, In-Sample and Out-of-Sample

Moneyness	Days-to-Maturity					
S/K	<30	30-60	60-90	90-120	120-150	>150
Panel A. In-Sample, 2016-2018						
<0.94	0.1753	0.1337	0.1204	0.1106	0.1084	0.1116
	(1,801)	(5,880)	(2,640)	(1,622)	(1,052)	(716)
0.94-0.97	0.1259	0.1004	0.1018	0.1044	0.1085	0.1169
	(7,432)	(11,402)	(3,295)	(1,490)	(593)	(350)
0.97-1.00	0.1005	0.1020	0.1117	0.1201	0.1227	0.1275
	(20,011)	(13,775)	(3,610)	(1,544)	(575)	(350)
1.00-1.03	0.1233	0.1236	0.1314	0.1383	0.1394	0.1441
	(14,472)	(9,180)	(2,440)	(921)	(337)	(206)
1.03-1.06	0.1717	0.1511	0.1498	0.1555	0.1545	0.1599
	(4,105)	(3,753)	(1,099)	(385)	(185)	(118)
>=1.06	0.2877	0.2172	0.2058	0.2113	0.2033	0.2008
	(1,570)	(2,327)	(962)	(542)	(324)	(223)
Panel B. Out-of-Sample, 2019						
<0.94	0.1377	0.1199	0.1164	0.1138	0.1156	0.1174
	(374)	(2,182)	(1,233)	(856)	(570)	(406)
0.94-0.97	0.1146	0.1134	0.1134	0.1180	0.1235	0.1263
	(3,925)	(5,066)	(1,484)	(800)	(377)	(231)
0.97-1.00	0.1147	0.1210	0.1292	0.1355	0.1406	0.1439
	(9,126)	(5,536)	(1,505)	(835)	(387)	(203)
1.00-1.03	0.1437	0.1455	0.1488	0.1522	0.1547	0.1597
	(6,297)	(3,492)	(877)	(475)	(246)	(114)
1.03-1.06	0.1808	0.1725	0.1694	0.1732	0.1725	0.1758
	(1,574)	(1,020)	(383)	(181)	(89)	(52)
>=1.06	0.2592	0.2230	0.2189	0.2222	0.2199	0.2148
	(618)	(584)	(300)	(212)	(150)	(111)

Notes: Table 4 reports the implied volatility of S&P 500 index options. The implied volatilities of individual calls are then averaged within each moneyness-maturity category. Moneyness is determined by comparing the spot S&P 500 index level (denoted as S) with the exercise price (denoted as K). The number of contracts are reported in the parentheses. The period of analysis for the in-sample data ranges from January 6, 2016, to December 26, 2018. On the other hand, the out-of-sample data covers the period from January 2, 2019, to December 18, 2019.

options with a maturity period of less than 30 days, while subplot (d) shows the IV smirk for out-of-sample options with a maturity period of 30 days or more.

Figure 5 shows that as the maturity period (T) increases, the IV smirk becomes flatter. This observation aligns with the conclusions derived from Table 4. Specifically, when considering in-sample options, those with shorter maturities exhibit a more pronounced IV smirk, with larger differences in implied volatility between deep in-the-money and out-of-the-money options. However, as the maturity period increases, the differences in implied volatility between these options become less pronounced, resulting in a flatter IV smirk.

5.5 Fitting Options

We evaluate the performance of different models by utilizing the implied volatility rootmean-squared error (IVRMSE) metric, which has been advocated by Renault (1997) as a suitable tool for comparing model performance in option pricing. The IVRMSE quantifies the discrepancy between model-based and market-based implied volatilities. To compute the IVRMSE, we employ the Black and Scholes (1973) formula (BS) to invert the model-based option price, denoted as $\mathrm{C}_j^{\mathrm{Mod}}$, for each contract $\$j\$$. Consequently, the model-based implied volatility, denoted as $\mathrm{IV}_j^{\mathrm{Mod}}$, can be formally derived as:

$$\mathrm{IV}_j^{\mathrm{Mod}} = \mathrm{BS}^{-1}\left(C_j^{\mathrm{Mod}}\right).$$

a similar procedure to obtain the market-based implied volatilities, denoted as $\mathrm{IV}_j^{\mathrm{Mkt}}$:

$$\mathrm{IV}_j^{\mathrm{Mkt}} = \mathrm{BS}^{-1}\left(C_j^{\mathrm{Mkt}}\right).$$

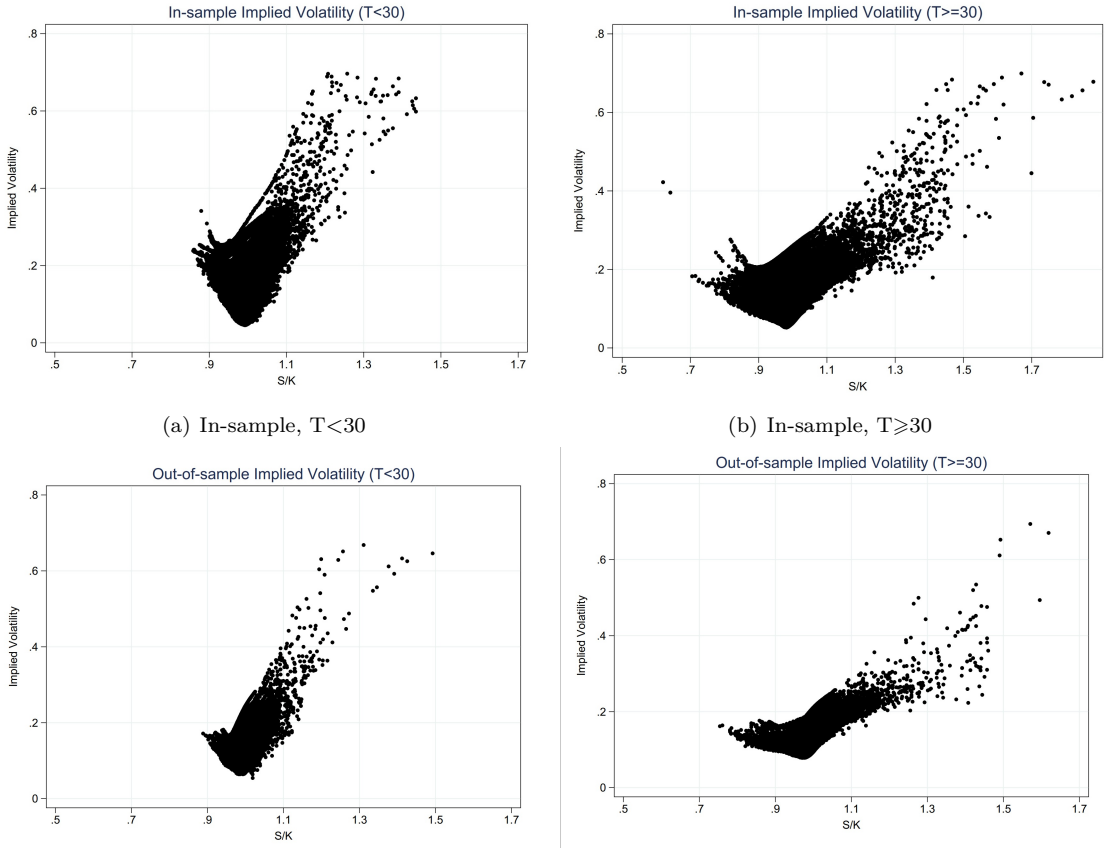
Consequently, the implied volatility error, denoted as $\$e\_\mathrm{j}\$$, is computed as the difference between the market-based and model-based implied volatilities:

$$e_j = \mathrm{IV}_j^{\mathrm{Mkt}} - \mathrm{IV}_j^{\mathrm{Mod}}$$

The IVRMSE is then obtained as the square root of the mean of the squared errors:

$$\mathrm{IVRMSE} \equiv \sqrt{\frac{1}{N}\sum_{j=1}^{N} e_j^2},$$

where N represents the size of the option sample. The IVRMSE serves as a comprehensive measure of the overall discrepancy between the model's implied volatilities and the market's implied volatilities, allowing for a robust comparison of model performance in option pricing.



(a) In-sample, T<30

(b) In-sample, T⩾30

(c) Out-of-Sample, T<30

(d) Out-of-Sample, T⩾30

Figure 5: In-sample and Out-of-Sample Implied Volatility

Notes: The graph illustrates the implied volatility (IV) smirk for different subsets of S&P 500 index options based on their maturity period (T) and whether they are in-sample or out-of-sample. Subplot (a) represents the IV smirk for in-sample options with a maturity period of less than 30 days, consisting of a total of 49,391 observations. Subplot (b) displays the IV smirk for in-sample options with a maturity period of 30 days or more, comprising a larger dataset of 71,896 observations. Subplot (c) shows the IV smirk for out-of-sample options with a maturity period of less than 30 days, consisting of 21,914 observations. Finally, subplot (d) presents the IV smirk for out-of-sample options with a maturity period of 30 days or more, which includes 29,957 observations.

The estimation of model parameters through the minimization of IVRMSE involves significant computational intensity since the Black-Scholes Model (BSM) inversion needs to be performed for each set of model option prices attempted by the optimizer. To overcome this, we adopt the approach proposed by Trolle and Schwartz (2009), who minimize the vega-weighted root-mean-squared error (VWRMSE) instead:

$$\text{VWRMSE} \equiv \sqrt{\frac{1}{N}\sum_{j=1}^{N} e_j^2} \equiv \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(\frac{C_j^{\text{MKT}} - C_j^{\text{MOD}}}{\text{BSV}_j^{\text{MKT}}}\right)^2}$$

where $\text{BSV}_j^{\text{MKT}}$ represents the Black-Scholes vega option (the derivative with respect to volatility) computed using the market-implied volatility level. IVRMSE and VWRMSE generally yield similar values. In the subsequent tables, we will report both IVRMSE and VWRMSE, but our analysis will primarily focus on the IVRMSE values. Instead of directly minimizing VWRMSE, we estimate the risk-neutral parameters by maximizing the likelihood function based on Gaussian vega-weighted option errors:

$$\ln L^O \propto -\frac{1}{2}\sum_{j=1}^{N}\left\{\ln\left(\text{VWRMSE}^2\right) + \frac{e_j^2}{\text{VWRMSE}^2}\right\}$$

where e_j represents the individual option errors.

Table 5 presents the results of the option-based estimation. It is evident that our option-fitting approach produces accurate parameter estimates, as indicated by relatively small standard errors and substantial model likelihoods. Since we solely fit the model on options, the resulting estimates correspond to risk-neutral parameters. Consequently, we do not estimate the market prices of risk, that is, parameters v and λ. It is important to note that, to ensure model consistency during the estimation process, we filter volatility on returns and realized variance (RV) while fitting the option implied volatilities (IVs). Following the historical estimation approach, we apply variance targeting method to and the theoretical unconditional risk-neutral variance formula to match sample variance.

Table 5: In-Sample Estimation Results on Options: 2016-2018

In-Sample Estimation Results on Options: 2016-2018

Parameters	IG-ARV	IG	ARV	HNG
a	105.3807	-465.1278		
	(248.3539)	(3.4471)		
α			3.79E-10	3.42E-06
			(1.02E-08)	(4.25E-08)
b	1.0101	0.8996		
	(0.0338)	(0.0003)		
β			0.9837	0.8266
			(0.3775)	(0.1100)
c	-4.72E-06	5.19E-06		
	(6.56E-07)	(4.18E-09)		
γ			229.5147	213.1854
			(94.0646)	(75.8540)
d	5.9185			
	(0.6588)			
σ			3.18E-06	
			(1.80E-07)	
η	-0.0065	-0.0066		
	(2.52E-06)	(8.72E-06)		
$\pi^{\mathbb{Q}}$	0.9016	0.9998	0.9837	0.9821
$\mathcal{L}$	222188	87334	224858	67428

Notes: Table 5 presents the risk-neutral estimation results for the inverse Gaussian Affine Realized Volatility (IG-ARV) model, the inverse Gaussian GARCH (IG) model, the Affine Realized Volatility (ARV) model, and the Heston-Nandi GARCH (HNG) model. The data used for estimation comprises Wednesday European call contracts sourced from OptionMetrics. Several filters are applied to clean the option data, as illustrated in the data-cleaning subsection. The observed period spans from January 6, 2016, to December 26, 2018, corresponding to the in-sample data in Table 4. The estimated parameters and their corresponding standard errors are reported. Standard errors are reported in the parentheses. The final row presents the loglikelihood values for each model. The second-to-last row reports the volatility persistences under $\mathbb{Q}$ measure, defined by the dynamics of each model, respectively.

We observe that the estimates of the IG-ARV model and the IG model exhibit a high degree of similarity, while the estimates of the ARV model and the HNG model also show considerable similarity. The estimated coefficients for η remain negative and statistically significant, supporting the negative relationship mentioned in the interpretation of Table 2. Additionally, the estimated coefficients for γ remain positive and statistically significant, indicating the robust presence of the leverage effect. The volatility persistences under $\mathbb{Q}$ measure exceed 0.90 for each of the models. The log-likelihood values for the IG-ARV model and the ARV model are 222,188 and 224,858, respectively, as both models incorporate the $\mathcal{L}_t(RV_{t+1})$ component.

6 Empirical Pricing Performance

6.1 In-Sample Pricing Performance

We proceed to analyze the IVRMSE results based on the parameters presented in Table 5 by categorizing the data based on moneyness, maturity, and VIX levels. Moneyness and maturity utilize the bins defined in Table 3. The decomposed IVRMSE results are reported in Table 6.

Table 6 presents the in-sample IVRMSE values for different models. The IG-ARV model demonstrates an IVRMSE of 5.2985, indicating a relatively low level of discrepancy between the model-based implied volatilities and the observed market volatilities. In comparison, the IG model, which also incorporates the inverse Gaussian distribution to model returns and volatility but lacks high-frequency information, exhibits a higher IVRMSE of 6.9496. Similarly, the ARV model, which incorporates high-frequency information but models the dynamic of realized variances using a normal distribution, shows a higher IVRMSE of 5.3981. The HNG model, which solely considers daily return and volatility dynamics with a normal distribution, displays the highest IVRMSE of 10.2170.

By incorporating skewness, the IG-ARV model reduces the IVRMSE by 1.85% compared to the IG model. Furthermore, by incorporating high-frequency information and modeling the dynamics of realized variances, the IG-ARV model reduces the IVRMSE by 23.76% compared to the ARV model. In terms of pricing performance, when compared to the HNG model that only models daily return and volatility, the IG-ARV model exhibits a significant improvement of 48.14%.

Furthermore, Table 6 reveals that the overall improvement in option fit achieved by the IG-ARV model is not driven by any specific subset of the data. However, it is noteworthy that the IG-ARV

model demonstrates particularly strong performance when pricing options with a shorter time to maturity (DTM) of less than 30 days. This subset, as shown in Table 3, constitutes approximately 41.18% of our sample. Options with a DTM of less than 30 days exhibit a distinct volatility smirk, suggesting that the combination of the inverse Gaussian distribution and the inclusion of high-frequency information contributes to enhanced option valuation performance in this category.

Consider now the Panel C of Table 6, which reports the IVRMSE across VIX level categories. We see that the IG-ARV model performs the best in the highest VIX level category, where all the models displays the highest IVRMSE. The IG-ARV model reduces IVRMSE by 9.51% than the second best performing model. Containing empirical quantities improves the pricing performance by 28.95% given inverse Gaussian distributed innovation (comparing to the IG model), while considering skewness help to reduce the IVRMSE by 25.21% given upgrading the forecasting of volatiltity using the realized measures.

6.2 Out-of-Sample Pricing Performance

We further analyze the overall IVRMSE results based on the estimates in Table 5 by disaggregating the data according to moneyness, maturity, and VIX levels. This segmentation, based on the bins provided in Table 3, allows us to examine the individual components of the IVRMSE. Table 7 reports the detailed breakdown of these results.

The first column of Table 7 reports the out-of-sample total IVRMSE of the IG-ARV model, the IG model, the ARV model, and the HNG model. We find, that the IVRMSEs are relatively smaller than those in Table 6, and the IG-ARV model still performs the best among the four models. Total IVRMSE of the IG-ARV model is 5.2985, reducing the pricing error of 16.80% by considering skewness under the framework (in comparison with the ARV model). Besides, taking into account the empirical quantities help to reduce IVRMSE by 30.60% (in comparison with the IG model).

When coming to the different categories by moneyness, maturity, and VIX level, Table 7 shows that the IVRMSE the IG-ARV model, which has the lowest overall IVRMSE in Table 7, has the lowest IVRMSE in each of the six moneyness categories considered. As is shown by Christoffersen et al. (2006), models consider skewness (the IG-ARV model and the IG model) exhibit better performances when pricing the deep ITM call options, that is, when $S/K \geqslant 1.06$. Here the IVRMSE of the IG-ARV model and the IG model are 12.1286 and 12.2141, respectively, indicating the better fitness of inverse Gaussian based modelling.

Table 6: In-Sample IVRMSE Option Error by Moneyness, Maturity, and VIX: 2016-2018

In-Sample IVRMSE Option Error by Moneyness, Maturity and VIX

Model	Total	S/K < 0.94	0.94⩽S/K<0.97	0.97⩽S/K<1.00	1.00⩽S/K<1.03	1.03⩽S/K<1.06	S/K⩾1.06
Panel A. Partitioned by Moneyness							
IG-ARV	5.2985	3.5251	4.7216	5.0063	4.3604	4.2306	12.8750
IG	6.9496	2.8288	4.1337	6.9073	7.0915	7.7373	15.6413
ARV	5.3981	4.5843	3.7325	3.8820	4.6590	6.9478	14.3931
HNG	10.2170	10.0187	11.3259	11.0412	9.0239	7.0315	9.6514
Model	Total	T<30	30⩽T<60	60⩽T<90	90⩽T<120	120⩽T<150	T⩾150
Panel B. Partitioned by Maturity							
IG-ARV	5.2985	5.7235	5.0877	4.8672	4.7640	4.6825	4.5392
IG	6.9496	8.4675	6.2353	4.8376	3.9557	3.9205	4.3740
ARV	5.3981	6.3555	4.8063	4.4043	4.2171	3.9706	4.0697
HNG	10.2170	8.9401	10.9136	11.1979	11.1967	11.2305	10.9079
Model	Total	VIX<15	15⩽VIX<20	20⩽VIX<25	25⩽VIX<30	VIX⩽30	
Panel C. Partitioned by VIX Level							
IG-ARV	5.2985	5.0669	4.2095	5.7276	7.4121	10.7804	
IG	6.9496	4.9291	8.0286	10.3139	11.3134	15.1727	
ARV	5.3981	3.5886	5.4219	8.0778	10.3626	14.4146	
HNG	10.2170	6.0213	6.1931	6.8937	8.9406	11.9140	

Notes: Table 6 presents the IVRMSE for the IG-ARV, IG, ARV, and HNG models, categorized by moneyness, maturity, and VIX level. The sample consists of European call options on the S&P 500 index from OptionMetrics, with parameter estimates obtained from Table 5. The in-sample period spans from January 6, 2016, to December 26, 2018. Panel A displays IVRMSEs based on moneyness (measured by S/K), Panel B presents IVRMSEs based on days to maturity (DTM), and Panel C shows IVRMSEs based on the VIX level. All IVRMSE values are expressed as percentages, representing the discrepancy between model-based implied volatilities and observed market volatilities.

Table 7: Out-of-Sample IVRMSE Option Error by Moneyness, Maturity, and VIX: 2019

Out-of-Sample IVRMSE Option Error by Moneyness, Maturity and VIX

Model	Total	S/K < 0.94	0.94⩽S/K<0.97	0.97⩽S/K<1.00	1.00⩽S/K<1.03	1.03⩽S/K<1.06	S/K⩾1.06
Panel A. Partitioned by Moneyness							
IG-ARV	3.4988	1.6536	2.0662	2.4188	3.1178	4.5753	12.1286
IG	5.0413	6.2313	3.7478	3.2014	4.7909	7.7735	12.2141
ARV	4.2055	1.7735	2.1061	2.8486	4.3343	6.8437	13.1212
HNG	4.3865	1.9231	1.9801	3.0704	4.9489	7.2511	12.6400
Model	Total	T<30	30⩽T<60	60⩽T<90	90⩽T<120	120⩽T<150	T⩾150
Panel B. Partitioned by Maturity							
IG-ARV	3.4988	3.8033	3.1218	3.3239	3.4656	3.6106	3.7650
IG	5.0413	5.1952	4.7622	4.9466	5.1260	5.5206	5.6910
ARV	4.2055	4.6768	3.7089	3.8670	3.9883	4.1601	4.3164
HNG	4.3865	4.5882	3.9921	4.3688	4.5902	4.8256	5.0335
Model	Total	VIX<15	15⩽VIX<20	VIX⩾20			
Panel C. Partitioned by VIX Level							
IG-ARV	3.4988	3.0616	3.4876	5.5606			
IG	5.0413	4.3102	5.3459	7.2013			
ARV	4.2055	3.4012	4.4044	6.8833			
HNG	4.3865	3.3363	4.8452	6.9286			

Notes: Table 7 presents the IVRMSE for the IG-ARV, IG, ARV, and HNG models, categorized by moneyness, maturity, and VIX level. The sample consists of European call options on the S&P 500 index from OptionMetrics, with parameter estimates obtained from Table 5. The out-of-sample period spans from January 2, 2019, to December 18, 2019. Panel A displays IVRMSEs based on moneyness (measured by S/K), Panel B presents IVRMSEs based on days to maturity (DTM), and Panel C shows IVRMSEs based on the VIX level. All IVRMSE values are expressed as percentages, representing the discrepancy between model-based implied volatilities and observed market volatilities.

Robust improvements given by considering both skewness and high-frequency information are shown in Panel B and Panel C. The advantages provided by the IG-ARV model are not limited to any specific range of maturity or VIX level. In terms of shortest maturity, the IG-ARV model exhibits improvement of 17.11%, while 19.22% reduction of IVRMSE in the hightest VIX level category. As expected, all models face the greatest challenge in accurately fitting options during periods of high market volatility. We also find that considering realized measures help to reduce pricing error, since the IG-ARV model and the ARV model perform better in the highly-volatile times.

7 Conclusion

This study develops a new and flexible option-pricing model that explicitly incorporates the dynamics of skewness and realized volatility. An important feature of the model is that the dynamics of variances are governed by their empirical proxies. Given that these proxies are constructed in discrete time, our model contributes to the discrete-time family.

From a theoretical viewpoint, the structure of the new model and the application of the inverse Fourier transform enable us to derive closed-form option valuation formulas that nest several option-pricing specifications. This feature facilitates the estimation procedure, allows for a direct comparison of nested models, and avoids the need to resort to simulation techniques.

From an empirical viewpoint, our new proposed option pricing model performs well. In terms of pricing accuracy, the model improves significantly upon popular specifications when optimized on a dataset of S&P 500 index options, realized variances, and returns. In particular, the joint modeling of skewness and realized volatility accounts for an out-of-sample gain of 16.80% in pricing accuracy.

Appendix

Appendix A. Proof of Moment Matching

Following Christoffersen et al. (2006), the random shocks in the return equation and GARCH equation follow Inverse-Gaussian distribution:

$$\begin{aligned} R_{t+1} &= r + v h_t^{RV} + \eta y_{1,t+1}, \\ h_{t+1}^{RV} &= w + b h_t^{RV} + c y_{2,t+1} + a \frac{\left(h_t^{RV}\right)^2}{y_{2,t+1}}, \end{aligned} \tag{A-1}$$

where $Cov\left(y_{1,t+1}, y_{2,t+1}\right) = 0$, and $y_{1,t+1}, y_{2,t+1} \overset{i.i.d}{\sim} IG(\delta)$.[2]

Applying the same technique as IGG model, we can derive the first two moments of equation A-1.

Note that here we need to specify the measurement equation, in which the basic idea is $E_t[RV_{t+1}] = h_t^{RV}$. We follow the specification of the measurement equation of ARV model, that is,

$$RV_{t+1} = h_t^{RV} + d \underbrace{\left[c y_{2,t+1} + a \frac{\left(h_t^{RV}\right)^2}{y_{2,t+1}} - (\cdots) \right]}_{\text{whose conditional expectation equals to 0}} . \tag{A-2}$$

We first explicitly write down the terms in the square bracket.

$$\begin{aligned} E_t \left[c y_{2,t+1} + a \frac{\left(h_t^{RV}\right)^2}{y_{2,t+1}} \right] &= c \cdot \delta_{t+1} + a \cdot \left(h_t^{RV}\right)^2 \cdot \left(\frac{1}{\delta_{t+1}} + \frac{1}{\delta_{t+1}^2} \right) \\ &= c \cdot \frac{h_t^{RV}}{\eta^2} + a \left(h_t^{RV}\right)^2 \cdot \left(\frac{\eta^2}{h_t^{RV}} + \frac{\eta^4}{\left(h_t^{RV}\right)^2} \right) \\ &= \left(\frac{c}{\eta^2} + a\eta^2 \right) h_t^{RV} + a\eta^4, \end{aligned} \tag{A-3}$$

thus the measurement equation can be written into

$$RV_{t+1} = h_t^{RV} + d \left[c y_{2,t+1} + a \frac{\left(h_t^{RV}\right)^2}{y_{2,t+1}} - \left(\left(\frac{c}{\eta^2} + a\eta^2 \right) h_t^{RV} + a\eta^4 \right) \right]. \tag{A-4}$$

The additional parameter d is born to have something to do with parameter σ, which will be pinned down later in the moment-matching equations.

[2] See Section 1.1 in Christoffersen et al. (2006).

Then we derive the first two moments of IG-ARV model. Note that there's little difference in return equation and GARCH equation in the IG model and the IG-ARV model, we have

$$\begin{aligned} E_t[R_{t+1}] &= r + vh_t^{RV} + \eta \cdot \delta_{t+1} \\ &= r + vh_t^{RV} + \eta \cdot \frac{h_t^{RV}}{\eta^2} \\ &= r + \left(v + \frac{1}{\eta}\right) h_t^{RV}. \end{aligned} \tag{A-5}$$

$$Var_t[R_{t+1}] = \eta^2 Var_t[y_{1,t+1}] = h_t^{RV}. \tag{A-6}$$

In terms of RV and h, we have

$$E_t[RV_{t+1}] = h_t^{RV}, \tag{A-7}$$

$$\begin{aligned} Var_t[RV_{t+1}] &= d^2 \left(c^2 Var_t[y_{2,t+1}] + 2ca \left(h_t^{RV}\right)^2 Cov_t \left(y_{2,t+1}, \frac{1}{y_{2,t+1}} \right) \right) \\ &= d^2 \left(\left(\frac{c^2}{\eta^2} - 2ca\eta^2 + a\eta^6 \right) h_t^{RV} + 2a^2\eta^8 \right), \end{aligned} \tag{A-8}$$

$$E_t[h_{t+1}^{RV}] = \left(w + a\eta^4\right) + \left(b + \frac{c}{\eta^2} + a\eta^2 \right) h_t^{RV}, \tag{A-9}$$

$$Var_t[h_{t+1}^{RV}] = \left(\frac{c^2}{\eta^2} - 2ca\eta^2 + a\eta^6 \right) h_t^{RV} + 2a^2\eta^8. \tag{A-10}$$

To sum up, we have

$$\begin{aligned} v + \frac{1}{\eta} &= \lambda, \\ w + a\eta^4 &= \omega + \alpha, \\ b + \frac{c}{\eta^2} + a\eta^2 &= \beta + \alpha\gamma^2, \\ a\eta^4 &= \alpha, \\ \frac{c}{\eta} - \eta^3 a &= -2\alpha\gamma, \\ d^2 \left(\frac{c}{\eta} - \eta^3 a \right)^2 &= 4\sigma^2\gamma_2^2, \\ d^2 \cdot 2a^2\eta^8 &= 2\sigma^2. \end{aligned} \tag{A-11}$$

Thus we obtain all the parameters of IG-ARV model, where $d = \frac{\sigma}{\alpha}$ and the others are the same with IG model.

Appendix B. Proof of Moment Generating Function

To begin with, consider the one-step-ahead MGF:

$$\begin{aligned}
&\mathbb{E}_t\left[\exp\left(\nu_1 R_{t+1}+\nu_2 h_{t+1}+\nu_3 s_{t+1}\right)\right] \\
&=\mathbb{E}_t\left[\exp\left(\begin{array}{c}\nu_1\left(r+\left(\lambda_z-\frac{1}{2}\right)h_t+\left(\lambda_y-\xi\right)s_t+\sqrt{h_t}z_{1,t+1}+\eta_1 y_{1,t+1}\right)\\ +\nu_2\left(\omega+\beta h_t+\alpha\left(z_{2,t+1}-\gamma\sqrt{h_t}\right)^2\right)\\ +\nu_3\left(w+bs_t+cy_{2,t+1}+a\frac{(h_t)^2}{y_{2,t+1}}\right)\end{array}\right)\right] \\
&=\exp\left(\nu_1 r+\nu_2\omega+\nu_3 w+\left(\nu_1\left(\lambda_z-\frac{1}{2}\right)+\nu_2\beta\right)h_t+\nu_3 b s_t\right) \\
&\quad\cdot\mathbb{E}_t\left[\exp\left(\nu_1\sqrt{h_t}z_{1,t+1}+\nu_2\alpha\left(z_{2,t+1}-\gamma\sqrt{h_t}\right)^2\right)\right]\cdot\mathbb{E}_t\left[\left(\nu_1\eta_1 y_{1,t+1}\right)\right]\cdot\mathbb{E}_t\left[\exp\left(\nu_3\left(cy_{2,t+1}+a\frac{h_t^2}{y_{2,t+1}}\right)\right)\right]
\end{aligned} \tag{B-1}$$

based on the independence of $y_{1,t+1}$ and $y_{2,t+1}$.

Using the Cholesky representation, we can rewrite

$$z_{1,t+1}=\rho z_{2,t+1}+\sqrt{1-\rho^2}z_{3,t+1},$$

where $Corr\left[z_{2,t+1},z_{3,t+1}\right]=0$. Moreover, by exploiting the identity

$$\log\mathbb{E}\left[\exp\left(a(z+b)^2\right)\right]=\frac{ab^2}{1-2a}-\frac{1}{2}\log(1-2a),$$

for z distributed as a standard normal random variable. One can check that

$$\mathbb{E}_t\left[\exp\left(\nu_1\sqrt{h_t}z_{1,t+1}+\nu_2\alpha\left(z_{2,t+1}-\gamma\sqrt{h_t}\right)^2\right)\right]=\exp\left(\left(\nu_2\alpha\gamma^2+\frac{2\nu_2\alpha}{1-2\nu_2\alpha}\left(\nu_1\rho-2\nu_2\alpha\gamma\right)^2\right)h_t-\frac{1}{2}\log\left(1-2\nu_2\alpha\right)\right), \tag{B-2}$$

$$\mathbb{E}_t\left[\exp\left(\nu_1\eta_1 y_{1,t+1}\right)\right]=\exp\left(s_t-s_t\sqrt{1-2\nu_1\eta_1}\right), \tag{B-3}$$

$$\mathbb{E}_t\left[\exp\left(\nu_3\left(cy_{2,t+1}+a\frac{h_t^2}{y_{2,t+1}}\right)\right)\right]=\frac{\delta_{t+1}}{\sqrt{\delta_{t+1}^2-2\nu_3 c h_t^2}}\exp\left(\delta_{t+1}-\sqrt{\left(\delta_{t+1}^2-2\nu_3 c h_t^2\right)\left(1-2\nu_3 c\right)}\right). \tag{B-4}$$

Note that $\delta_{t+1}=\frac{h_t}{\eta_2^2}$, multiplication of the three terms yields

$$\exp\left(\begin{array}{c}\left(\nu_2\alpha\gamma^2+\frac{2\nu_2\alpha}{1-2\nu_2\alpha}\left(\nu_1\rho-2\nu_2\alpha\gamma\right)\right)h_t-\frac{1}{2}\log\left(1-2\nu_2\alpha\right)\\ +s_t-s_t\sqrt{1-2\nu_1\eta_1}\\ -\frac{1}{2}\log\left(1-2\nu_3 c\eta_2^4\right)+\left(\frac{1}{\eta_2^2}-\frac{1}{\eta_2^2}\sqrt{\left(1-2\nu_3 c\eta_2^4\right)\left(1-2\nu_3 c\right)}\right)h_t\end{array}\right). \tag{B-5}$$

Therefore, we obtain

$$\begin{aligned}
&\mathbb{E}_t\left[\exp\left(\nu_1 R_{t+1}+\nu_2 h_{t+1}+\nu_3 s_{t+1}\right)\right] \\
&=\exp\left(\nu_1 r+\nu_2\omega+\nu_3 w+\left(\nu_1\left(\lambda_z-\frac{1}{2}\right)+\nu_2\beta\right)h_t+\nu_3 b s_t\right) \\
&\cdot\exp\begin{pmatrix}\left(\nu_2\alpha\gamma^2+\frac{2\nu_2\alpha}{1-2\nu_2\alpha}\left(\nu_1\rho-2\nu_2\alpha\gamma\right)\right)h_t-\frac{1}{2}\log\left(1-2\nu_2\alpha\right)\\ +s_t-s_t\sqrt{1-2\nu_1\eta_1}\\ -\frac{1}{2}\log\left(1-2\nu_3 c\eta_2^4\right)+\left(\frac{1}{\eta_2^2}-\frac{1}{\eta_2^2}\sqrt{\left(1-2\nu_3 c\eta_2^4\right)\left(1-2\nu_3 c\right)}\right)h_t\end{pmatrix}. \\
&=\exp\left(A(\nu_1,\nu_2,\nu_3)h_t+B(\nu_1,\nu_2,\nu_3)s_t+C(\nu_1,\nu_2,\nu_3)\right),
\end{aligned} \tag{B-6}$$

where

$$\begin{aligned}
A(\nu_1,\nu_2,\nu_3)&=\nu_1\left(\lambda_z-\frac{1}{2}\right)+\nu_2\beta+\nu_2\alpha\gamma^2+\frac{2\nu_2\alpha}{1-2\nu_2\alpha}\left(\nu_1\rho-2\nu_2\alpha\gamma\right)+\frac{1}{\eta_2^2}-\frac{1}{\eta_2^2}\sqrt{\left(1-2\nu_3 c\eta_2^4\right)\left(1-2\nu_3 c\right)},\\
B(\nu_1,\nu_2,\nu_3)&=\nu_3 b+1-\sqrt{1-2\nu_1\eta_1},\\
C(\nu_1,\nu_2,\nu_3)&=\nu_1 r+\nu_2\omega+\nu_3 w-\frac{1}{2}\log\left(1-2\nu_2\alpha\right)-\frac{1}{2}\log\left(1-2\nu_3 c\eta_2^4\right).
\end{aligned}$$

For the multi-period one, we denote Ψ as the corresponding MGF, and denote M as day-to-maturity. We expect that the multi-period MGF keep exponentially affine, that is,

$$\Psi_{t,t+M}(u)\equiv\mathbb{E}_t\left[\exp\left(u\sum_{j=1}^{M}R_{t+j}\right)\right]=\exp\left(D(u,M)h_t+F(u,M)s_t+G(u,M)\right).$$

By applying the law of conditional expectation iteration,

$$\begin{aligned}
\Psi_{t,t+M}(u)&\equiv\mathbb{E}_t\left[\exp\left(u\sum_{j=1}^{M+1}R_{t+j}\right)\right]=\mathbb{E}_t\left[\mathbb{E}_{t+1}\left[\exp\left(u\sum_{j=1}^{M+1}R_{t+j}\right)\right]\right]\\
&=\mathbb{E}_t\left[\exp\left(uR_{t+1}\right)\cdot\mathbb{E}_{t+1}\left[\exp\left(u\sum_{j=2}^{M+1}R_{t+j}\right)\right]\right]\\
&=\mathbb{E}_t\left[\exp\left(uR_{t+1}\right)\cdot\mathbb{E}_{t+1}\left[\exp\left(u\sum_{k=1}^{M}R_{t+1+k}\right)\right]\right]\\
&=\mathbb{E}_t\left[uR_{t+1}+D(u,M)h_{t+1}+F(u,M)s_{t+1}+G(u,M)\right]\\
&=\exp\left(A(u,D(u,M),F(u,M))h_{t+1}+B(u,D(u,M),F(u,M))s_t+C(u,D(u,M),F(u,M))+G(u,M)\right),
\end{aligned} \tag{B-7}$$

the recursive rule can be written as

$$\begin{aligned} D(u, M+1) &= A(u, D(u, M), F(u, M)), \\ F(u, M+1) &= B(u, D(u, M), F(u, M)), \\ G(u, M+1) &= C(u, D(u, M), F(u, M)) + G(u, M). \end{aligned} \tag{B-8}$$

where

$$\begin{aligned} D(u, 1) &= A(u, 0, 0) \\ F(u, 1) &= B(u, 0, 0), \\ G(u, 1) &= C(u, 0, 0). \end{aligned} \tag{B-9}$$

Appendix C. Proof of Risk-Neutralization

Note that the model contains three random shocks, we follow Christoffersen et al. (2010) to apply the following pricing kernel:

$$Z_{t+1} = \frac{\exp\left(\nu_{1,t} z_{1,t+1} + \nu_{2,t} z_{2,t+1} + \nu_{3,t} y_{1,t+1} + \nu_{4,t} y_{2,t+1}\right)}{\mathbb{E}_t\left[\exp\left(\nu_{1,t} z_{1,t+1} + \nu_{2,t} z_{2,t+1} + \nu_{3,t} y_{1,t+1} + \nu_{4,t} y_{2,t+1}\right)\right]}. \tag{C-1}$$

As is illustrated in the model, $Corr_t(y_{1,t+1}, y_{2,t+1}) = 0, Corr_t(z_{1,t+1}, z_{2,t+1}) = \rho$. Then we obtain

$$Z_{t+1} = \frac{\exp\left(\nu_{1,t} z_{1,t+1} + \nu_{2,t} z_{2,t+1} + \nu_{3,t} y_{1,t+1} + \nu_{4,t} y_{2,t+1}\right)}{\exp\left(\frac{1}{2}\nu_{1,t}^2 + \rho\nu_{1,t}\nu_{2,t} + \frac{1}{2}\nu_{2,t}^2 + s_t - s_t\sqrt{1-2\nu_{3,t}} + \delta_t - \delta_t\sqrt{1-2\nu_{4,t}}\right)} \tag{C-2}$$

We need to impose that

$$\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(R_{t+1}\right)\right] = \exp(r), \tag{C-3}$$

where

$$\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(R_{t+1}\right)\right] = \mathbb{E}_t\left[Z_{t+1}\exp\left(R_{t+1}\right)\right], \tag{C-4}$$

that is,

$$\mathbb{E}_t\left[\exp\left(\begin{array}{c} \nu_{1,t} z_{1,t+1} + \nu_{2,t} z_{2,t+1} + \nu_{3,t} y_{1,t+1} + \nu_{4,t} y_{2,t+1} \\ -\frac{1}{2}\nu_{1,t}^2 - \rho\nu_{1,t}\nu_{2,t} - \frac{1}{2}\nu_{2,t}^2 - s_t + s_t\sqrt{1-2\nu_{3,t}} - \delta_t + \delta_t\sqrt{1-2\nu_{4,t}} \\ +r + \left(\lambda_z - \frac{1}{2}\right) h_t + \left(\lambda_y - \xi\right) s_t + \sqrt{h_t} z_{1,t+1} + \eta_1 y_{1,t+1} \end{array}\right)\right] = \exp(r), \tag{C-5}$$

therefore,

$$\nu_{1,t}\sqrt{h_t}+\rho\nu_{2,t}\sqrt{h_t}+\lambda_z h_t-s_t\sqrt{1-2\left(\nu_{3,t}+\eta_1\right)}+s_t\sqrt{1-2\nu_{3,t}}+\left(\lambda_y-\xi\right)s_t=0. \tag{C-6}$$

For the shocks, we have the risk-neutral expectation

$$\begin{aligned}
&\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(u_1 z_{1,t+1}+u_2 z_{2,t+1}+u_3 y_{1,t+1}+u_4 y_{2,t+1}\right)\right]\\
&=\mathbb{E}_t\left[Z_{t+1}\exp\left(u_1 z_{1,t+1}+u_2 z_{2,t+1}+u_3 y_{1,t+1}+u_4 y_{2,t+1}\right)\right]\\
&=\mathbb{E}_t\left[\exp\left(\begin{array}{c}\nu_{1,t}z_{1,t+1}+\nu_{2,t}z_{2,t+1}+\nu_{3,t}y_{1,t+1}+\nu_{4,t}y_{2,t+1}\\ -\frac{1}{2}\nu_{1,t}^2-\rho\nu_{1,t}\nu_{2,t}-s_t+s_t\sqrt{1-2\nu_{3,t}}-\delta_t+\delta_t\sqrt{1-2\nu_{4,t}}\\ +u_1 z_{1,t+1}+u_2 z_{2,t+1}+u_3 y_{1,t+1}+u_4 y_{2,t+1}\end{array}\right)\right]\\
&=\exp\left(\begin{array}{c}\frac{1}{2}u_1^2+\left(u_1+\rho u_2\right)\nu_{1,t}+\frac{1}{2}u_2^2+\left(u_2+\rho u_1\right)\nu_{2,t}+\rho u_1 u_2\\ -s_t\sqrt{1-2\left(\nu_{3,t}+u_3\right)}+s_t\sqrt{1-2\nu_{3,t}}-\delta_t\sqrt{1-2\left(\nu_{4,t}+u_4\right)}+\delta_t\sqrt{1-2\nu_{4,t}}\end{array}\right).
\end{aligned} \tag{C-7}$$

Consider the following transformation:

$$\left(z_{1,t+1},z_{2,t+1},y_{1,t+1},y_{2,t+1}\right)'=\left(z_{1,t+1}^*+A,z_{2,t+1}^*+B,Cy_{1,t+1}^*,Dy_{2,t+1}^*\right)', \tag{C-8}$$

therefore,

$$\begin{aligned}
&\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(u_1 z_{1,t+1}+u_2 z_{2,t+1}+u_3 y_{1,t+1}+u_4 y_{2,+1}\right)\right]\\
&=\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(u_1\left(z_{1,t+1}^*+A\right)+u_2\left(z_{2,t+1}^*+B\right)+u_3 C y_{1,t+1}^*+u_4 D y_{2,t+1}^*\right)\right]\\
&=\exp\left(\frac{1}{2}u_1^2+u_1 A+\frac{1}{2}u_2^2+u_2 B+\rho u_1 u_2+s_t^*-s_t^*\sqrt{1-2u_3 C}+\delta_t^*-\delta_t^*\sqrt{1-2u_4 D}\right),
\end{aligned} \tag{C-9}$$

which implies that

$$z_{1,t+1}=z_{1,t+1}^*+\left(\nu_{1,t}+\rho\nu_{2,t}\right), \tag{C-10}$$

$$z_{2,t+1}=z_{2,t+1}^*+\left(\rho\nu_{1,t}+\nu_{2,t}\right). \tag{C-11}$$

Additionally,

$$-s_t\sqrt{1-2\left(\nu_{3,t}+u_3\right)}+s_t\sqrt{1-2\nu_{3,t}}=s_t^*-s_t^*\sqrt{1-2u_3 C}, \tag{C-12}$$

$$-\delta_t\sqrt{1-2\left(\nu_{4,t}+u_4\right)}+\delta_t\sqrt{1-2\nu_{4,t}}=\delta_t^*-\delta_t^*\sqrt{1-2u_4 D}. \tag{C-13}$$

To write the equation above in a tighter way, we take s_t as an example:

$$\begin{aligned}
s_t\cdot\frac{2u_3}{\sqrt{1-2\nu_{3,t}}+\sqrt{1-2\left(\nu_{3,t}+u_3\right)}}&=s_t^*\frac{2u_3 C}{1+\sqrt{1-2u_3 C}},\\
s_t\cdot\frac{\frac{1}{\sqrt{1-2\nu_{3,t}}}}{1+\sqrt{1-2u_3\cdot\frac{1}{1-2\nu_{3,t}}}}&=s_t^*\frac{C}{1+\sqrt{1-2u_3 C}}.
\end{aligned} \tag{C-14}$$

Then we have

$$C = \frac{1}{1-2\nu_{3,t}}, \quad s_t \cdot \sqrt{C} = s_t^* C \Rightarrow s_t = s_t^* \sqrt{C}. \tag{C-15}$$

Similarly,

$$D = \frac{1}{1-2\nu_{4,t}}, \quad \delta_t = \delta_t^* \sqrt{D}. \tag{C-16}$$

Consequently, we can derive the the dynamic conditional skewness (DCS) model under $\mathbb{Q}$ measure:

$$\begin{aligned} R_{t+1} &= r + \left(\lambda_z - \frac{1}{2}\right) h_t + (\lambda_y - \xi)\, s_t + \sqrt{h_t} z_{1,t+1} + \eta_1 y_{1,t+1} \\ &= r + \left(\lambda_z - \frac{1}{2}\right) h_t + (\lambda_y - \xi)\, s_t + \sqrt{h_t}\left(z_{1,t+1}^* + \nu_{1,t} + \rho\nu_{2,t}\right) + \eta_1 C y_{1,t+1}^* \\ &= r - \frac{1}{2} h_t - \left(\sqrt{1-2\nu_{3,t}} - \sqrt{1-2\left(\nu_{3,t} + \eta_1^* C\right)}\right) s_t + \sqrt{h_t} z_{1,t+1}^* + \eta_1^* y_{1,t+1}^* \\ &= r - \frac{1}{2} h_t - \left(\sqrt{1-2\eta_1^* C^2} - 1\right) s_t^* + \sqrt{h_t} z_{1,t+1}^* + \eta_1^* y_{1,t+1}^*. \end{aligned} \tag{C-17}$$

Based on the mapping relationship of shocks under $\mathbb{P}$ measure and $\mathbb{Q}$ measure, the dynamic of h_t under $\mathbb{Q}$ can be written as

$$\begin{aligned} h_{t+1} &= \omega + \beta h_t + \alpha \left(z_{2,t+1}^* + \rho\nu_{1,t} + \nu_{2,t} - \gamma\sqrt{h_t}\right)^2 \\ &= \omega + \beta h_t + \alpha \left(z_{2,t+1}^* - \gamma_t^* \sqrt{h_t}\right)^2, \end{aligned} \tag{C-18}$$

where $\gamma_t^* = \gamma - \frac{\rho\nu_{1,t} + \nu_{2,t}}{\sqrt{h_t}}$. In order to keep the model affine, we set $\rho\nu_{1,t} + \nu_{2,t} = \chi\sqrt{h_t}$, then

$$\gamma_t^* = \gamma^* = \gamma - \chi, \tag{C-19}$$

where χ is a constant premium to estimate. Therefore, the measurement equation concerning h_t under $\mathbb{Q}$ measure is

$$\begin{aligned} \widetilde{RV_{t+1}} &= h_t + \sigma\left[\left(z_{2,t+1}^* - \gamma^*\sqrt{h_t}\right)^2 - \left(1 + \gamma^2 h_t\right)\right] \\ &= h_t^* + \sigma\left[\left(z_{2,t+1}^* - \gamma^*\sqrt{h_t}\right)^2 - \left(1 + \gamma^2 h_t^*\right)\right], \end{aligned} \tag{C-20}$$

where

$$h_t^* = h_t + \sigma\left(\gamma^{*2} - \gamma^2\right) h_t. \tag{C-21}$$

Then, the dynamic of conditional skewness is equivalent to

$$s_{t+1}^{*}\sqrt{C} = w + bs_t^{*}\sqrt{C} + cD \cdot y_{2,t+1}^{*} + a\frac{\frac{h_t^{*2}}{\left(1+\sigma(\gamma^{*2}-\gamma^2)^2\right)}}{y_{2,t+1}^{*}D}, \tag{C-22}$$

which can be reduced to

$$s_{t+1}^{*} = w^{*} + bs_t^{*} + c^{*}y_{2,t+1}^{*} + a^{*}\frac{h_t^{*2}}{y_{2,t+1}^{*}}, \tag{C-23}$$

where

$$w^{*} = \frac{w}{\sqrt{C}}, \quad c^{*} = \frac{cD}{\sqrt{C}}, \quad a^{*} = \frac{a}{D\sqrt{C}\left(1+\sigma\left(\gamma^{*2}-\gamma^2\right)\right)^2}. \tag{C-24}$$

Appendix D: Optimization Method of the IG-ARV model

We apply the Quasi Maximum Likelihood Estimation (QMLE) when estimating the models. The joint quasi-log-likelihood of returns and realized variance is determined by taking the logarithm of the bivariate inverse Gaussian distribution (BVIGD). We refer to as $\mathcal{L}_t\left(R_{t+1}, RV_{t+1}\right)$, which combines two components, the return part and the RV part.

The quasi-log-likelihood of returns at time $t+1$ conditional on information known at time t is defined using the inverse Gaussian distribution

$$\begin{aligned}\mathcal{L}_t(R_{t+1}) =& \log(\delta_{t+1}) - \frac{1}{2}\log(2\pi) - \frac{3}{2}\log\left(\frac{R_{t+1}-(r_f+vh_{t+1})}{\eta}\right) \\ &- \frac{1}{2}\left(\sqrt{\frac{R_{t+1}-(r_f+vh_{t+1})}{\eta}} - \frac{\delta_{t+1}}{\sqrt{\frac{R_{t+1}-(r_f+vh_{t+1})}{\eta}}}\right)^2 - \log\left(-\eta\right).\end{aligned} \tag{D-1}$$

The quasi-log-likelihood of realized variance at time $t+1$ conditional on information known at time t is similarly given by:

$$\begin{aligned}\mathcal{L}_t(RV_{t+1}) =& \log(\delta_{t+1}) - \frac{1}{2}\log(2\pi) - \frac{3}{2}\log\left(\frac{-\left(C-RV_{t+1}\right)+\sqrt{\Delta}}{2A}\right) \\ &- \frac{1}{2}\left(\sqrt{\frac{-(C-RV_{t+1})+\sqrt{\Delta}}{2A}} - \frac{\delta_{t+1}}{\sqrt{\frac{-(C-RV_{t+1})+\sqrt{\Delta}}{2A}}}\right)^2 \\ &- \log(2A) + \log\left(1-\frac{C-RV_{t+1}}{\sqrt{\Delta}}\right),\end{aligned} \tag{D-2}$$

where

$$\begin{aligned}
&\Delta = (C - RV_{t+1})^2 - 4AB, \\
&A = cd, \quad B = ad\left(h_t^{RV}\right)^2, \quad C = h_t^{RV} - d\left(\left(\frac{c}{\eta^2} + a\eta^2\right)h_t^{RV} + a\eta^4\right).
\end{aligned} \tag{D-3}$$

References

S. Aboura and D. Maillard. Option pricing under skewness and kurtosis using a cornish–fisher expansion. *Journal of Futures Markets*, 36(12):1194–1209, 2016.

T. G. Andersen, T. Bollerslev, F. X. Diebold, and H. Ebens. The distribution of realized stock return volatility. *Journal of financial economics*, 61(1):43–76, 2001a.

T. G. Andersen, T. Bollerslev, F. X. Diebold, and P. Labys. The distribution of realized exchange rate volatility. *Journal of the American statistical association*, 96(453):42–55, 2001b.

G. Bakshi, C. Cao, and Z. Chen. Empirical performance of alternative option pricing models. *The Journal of finance*, 52(5):2003–2049, 1997.

O. E. Barndorff-Nielsen. Normal inverse gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of statistics*, 24(1):1–13, 1997a.

O. E. Barndorff-Nielsen. Processes of normal inverse gaussian type. *Finance and stochastics*, 2: 41–68, 1997b.

O. E. Barndorff-Nielsen and N. Shephard. Non-gaussian ornstein–uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):167–241, 2001.

O. E. Barndorff-Nielsen and N. Shephard. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 64(2):253–280, 2002.

G. Barone-Adesi, R. F. Engle, and L. Mancini. A garch option pricing model with filtered historical simulation. *The review of financial studies*, 21(3):1223–1258, 2008.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637–654, 1973.

M. Chernov and E. Ghysels. A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation. *Journal of financial economics*, 56(3):407–458, 2000.

P. Christoffersen and K. Jacobs. Which garch model for option valuation? *Management science*, 50(9):1204–1221, 2004.

P. Christoffersen, S. Heston, and K. Jacobs. Option valuation with conditional skewness. *Journal of Econometrics*, 131(1-2):253–284, 2006.

P. Christoffersen, C. Dorion, K. Jacobs, and Y. Wang. Volatility components, affine restrictions, and nonnormal innovations. *Journal of Business & Economic Statistics*, 28(4):483–502, 2010.

P. Christoffersen, B. Feunou, K. Jacobs, and N. Meddahi. The economic value of realized volatility: Using high-frequency returns for option valuation. *Journal of Financial and Quantitative Analysis*, 49(3):663–697, 2014.

P. Christoffersen, B. Feunou, and Y. Jeon. Option valuation with observable volatility and jump dynamics. *Journal of Banking & Finance*, 61:S101–S120, 2015.

C. Corrado. The hidden martingale restriction in gram-charlier option prices. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 27(6):517–534, 2007.

F. Corsi, N. Fusari, and D. La Vecchia. Realizing smiles: Options pricing with realized volatility. *Journal of Financial Economics*, 107(2):284–304, 2013.

B. Feunou and C. Okou. Good volatility, bad volatility, and option pricing. *Journal of financial and quantitative analysis*, 54(2):695–727, 2019.

B. Feunou and R. Tédongap. A stochastic volatility model with conditional skewness. *Journal of Business & economic statistics*, 30(4):576–591, 2012.

L. Forsberg and T. Bollerslev. Bridging the gap between the distribution of realized (ecu) volatility and arch modelling (of the euro): the garch-nig model. *Journal of Applied Econometrics*, 17(5): 535–548, 2002.

C. R. Harvey and A. Siddique. Conditional skewness in asset pricing tests. *The Journal of finance*, 55(3):1263–1295, 2000.

S. L. Heston and S. Nandi. A closed-form garch option valuation model. *The review of financial studies*, 13(3):585–625, 2000.

Z. Huang, T. Wang, and P. R. Hansen. Option pricing with the realized garch model: An analytical approximation approach. *Journal of Futures Markets*, 37(4):328–358, 2017.

J. Hull and A. White. The pricing of options on assets with stochastic volatilities. *The journal of finance*, 42(2):281–300, 1987.

N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous univariate distributions, volume 2*, volume 289. John wiley & sons, 1995.

E. Jondeau and M. Rockinger. Gram–charlier densities. *Journal of Economic Dynamics and Control*, 25(10):1457–1483, 2001.

S. Kocherlakota. The bivariate inverse gaussian distribution: An introduction. *Communications in Statistics-Theory and Methods*, 15(4):1081–1112, 1986.

A. A. Majewski, G. Bormetti, and F. Corsi. Smile from the past: A general option pricing framework with multiple volatility and leverage components. *Journal of Econometrics*, 187(2):521–531, 2015.

E. Renault. Econometric models of option pricing errors", advances in economics and econometrics: theory and applications, eds d. kreps and k. wallis, vol. iii, 221-278, 1997.

E. Schlögl. Option pricing where the underlying assets follow a gram/charlier density of arbitrary order. *Journal of Economic Dynamics and Control*, 37(3):611–632, 2013.

L. Stentoft. American option pricing using garch models and the normal inverse gaussian distribution. *Journal of Financial Econometrics*, 6(4):540–582, 2008.

C. Tong and Z. Huang. Pricing vix options with realized volatility. *Journal of Futures Markets*, 41 (8):1180–1200, 2021.

A. B. Trolle and E. S. Schwartz. Unspanned stochastic volatility and the pricing of commodity derivatives. *The Review of Financial Studies*, 22(11):4423–4461, 2009.

M. C. Tweedie. Statistical properties of inverse gaussian distributions. i. *The Annals of Mathematical Statistics*, 28(2):362–377, 1957.

A. Wald. *Sequential Analysis*. John Wiley & Sons, New York, 1947.

T. Wang, F. Liang, Z. Huang, and H. Yan. Do realized higher moments have information content?-var forecasting based on the realized garch-rsrk model. *Economic Modelling*, 109:105781, 2022.

L. Zhang, P. A. Mykland, and Y. Aït-Sahalia. Edgeworth expansions for realized volatility and related estimators. *Journal of Econometrics*, 160(1):190–203, 2011.