北京大学中国经济研究中心
China Center for Economic Research
讨论稿系列
Working Paper Series

E2022025 **2022-12-30**

Option Pricing with Overnight and Intraday Volatility

Fang Liang Lingshan Du Zhuo Huang

Abstract

Efficiently exploiting the volatility information contained in price variations is im- portant for pricing options and other derivatives. In this paper, we develop a new and flexible option-pricing model that explicitly specifies the joint dynamics of overnight and intraday returns. The application of multivariate Edgeworth-Sargan density en- ables us to derive analytical approximations for option valuation formulas. Empirically, the model improves significantly upon benchmark models using S&P 500 index options. In particular, its separate modeling of intraday and overnight return volatility leads to an out-of-sample gain of 7.24% in pricing accuracy compared with the modeling of the close-to-close return volatility as a whole. The improvements are more pronounced during highly volatile periods.

Keywords: Overnight Volatility, Multivariate Edgeworth-Sargan Density, Option Pricing

JEL Classification: G13, C58, C51

Option Pricing with Overnight and Intraday Volatility*

Fang Liang† Lingshan Du‡ Zhuo Huang§

Abstract

Efficiently exploiting the volatility information contained in price variations is important for pricing options and other derivatives. In this paper, we develop a new and flexible option-pricing model that explicitly specifies the joint dynamics of overnight and intraday returns. The application of multivariate Edgeworth-Sargan density enables us to derive analytical approximations for option valuation formulas. Empirically, the model improves significantly upon benchmark models using S&P 500 index options. In particular, its separate modeling of intraday and overnight return volatility leads to an out-of-sample gain of 7.24% in pricing accuracy compared with the modeling of the close-to-close return volatility as a whole. The improvements are more pronounced during highly volatile periods.

Keywords: Overnight Volatility, Multivariate Edgeworth-Sargan Density, Option Pricing
JEL Classification: G13, C58, C51

*Fang Liang and Zhuo Huang acknowledge funding support provided by the National Natural Science Foundation of China (72101278, 72271010).

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

‡Guanghua School of Management, Peking University, Beijing, China.

§Corresponding author: Zhuo Huang, China Center for Economic Research, National School of Development, Peking University, Beijing, China. Email: zhuohuang@nsd.pku.edu.cn

1 Introduction

The proper specification of underlying asset volatility dynamics is a key element of an option valuation framework. A growing body of literature advocates for the use of empirically grounded properties in option pricing models. These studies use observed (realized) quantities to update volatility, such that volatility is no longer latent. This modeling approach leads to more precise measurement and forecasting of asset volatility, and the theoretical and empirical justifications for constructing reliable realized variance measures based on intraday high-frequency observations are given by Andersen et al. (2001a), Andersen et al. (2001b), and Andersen et al. (2003), among others. Recent studies, such as Corsi et al. (2013), Christoffersen et al. (2014), and Christoffersen et al. (2015), jointly model the dynamics of returns and realized variances under circumstances of option pricing and verify that this type of option valuation model is superior to models optimized only on returns.

The aforementioned studies focus exclusively on total daily returns and do not distinguish information pertaining to trading periods from that pertaining to non-trading periods. However, a halt in trading may result in a price information process that differs from that resulting from continuous trading. As total daily (close-to-close) returns rely only on the last tick price on an exchange for each trading day, close-to-close returns are not capable of effectively reflecting news that arrives during market closure.

For instance, before opening, European investors submit orders based on information revealed in U.S. stock markets, and trading is performed at a single price that clears the market. This means that the opening price of the exchange reflects accumulated overnight information from oversea markets (see, for example, Taylor (2007), Tsiakas (2008), and Chan et al. (1996)). Another strand of literature finds that extended futures trading (usually 24h-trading, except on weekends) contains information that is useful for explaining subsequent overnight spot returns. Hasbrouck (2003) shows that S&P 500 E-mini futures, which are traded overnight when the underlying stock market is closed, account for approximately 90% of the movement of the S&P 500 index. Similarly, Craig et al. (1995) find close links between implied changes and actual overnight changes in the Nikkei index. It is clear from both of the above strands of literature that overnight information from overseas stock markets and futures markets is crucial for explaining observed patterns in opening price reactions and unobserved patterns

during non-trading periods.

Figure 1 plots the averaged 1-minute returns and variances of the S&P 500 index from market open to close. It is not surprising to find that the largest 1-minute variance for a trading day occurs when the market opens, as this variance reflects the opening price reaction to information accumulated overnight from futures markets and overseas stock markets. The figure shows that a standard U-shaped pattern exists (Hong and Wang (2000)). The closing variances, although tending to be higher than those earlier in the day, are not so dramatic as the opening variances.

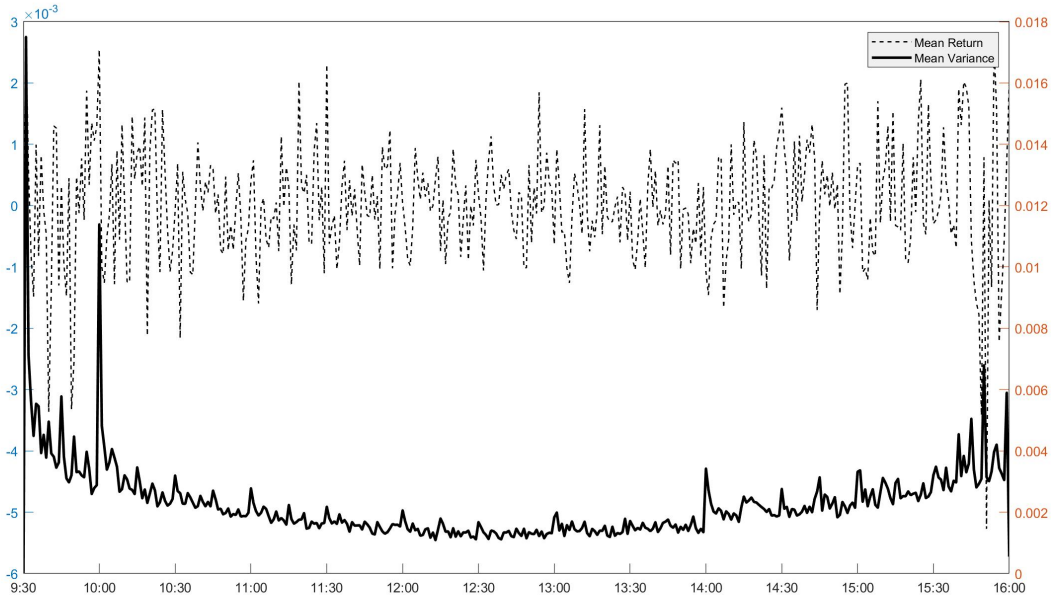

Figure 1 Opening Price Reaction to Accumulated Overnight Information

Notes: Figure 1 presents the averaged 1-minute returns and 1-minute variances (computed by 1-minute squared returns) of the S&P 500 index during trading hours. This sample covers the period from July 2, 2003 to December 18, 2019.

Moreover, macroeconomic and corporate announcements released during non-trading hours can be predictive of opening price, asset return and volatility (see, for example, Chan and Marsh (2022), Jiang et al. (2012), Akey et al. (2022), and Hu et al. (2021)). Moshirian et al. (2012) examine the impact of corporate news announcements released overnight on price discovery during the pre-opening period in the Australian Securities Exchange and find that prices respond immediately to overnight news upon the commencement of trading. Boudoukh

et al. (2019) find that fundamental information in firm-level announcements accounts for 49.6% of overnight idiosyncratic volatility (vs. 12.4% of trading-hour idiosyncratic volatility). Furthermore, Atilgan (2014) obtains evidence that compared with volatility spreads on non-earnings-announcement days, volatility spreads on earnings-announcement days[1] are more predictive of stock returns.

To show the immediate reactions of asset returns and variances to overnight news releases, we consider the most important U.S. macroeconomic announcement, the nonfarm payroll employment release. The nonfarm payroll employment is one of the most important announcements for all markets and it is often referred to as the "king" of announcements by market participants (for example, see Andersen and Bollerslev (1998) and Andersen et al. (2007)). Nonfarm payroll employment is released at 8:30 Eastern Standard Time (EST) when the futures markets are open but the equity markets are closed. Figure 2 plots the averaged 1-minute returns and variances of the S&P 500 E-mini futures in the period ranging from 1 hour before to 1 hour after the announcement. As can be seen, there is a sharp increase in both returns and variances immediately after the announcement, revealing an instantaneous market reaction to information incorporated into overnight announcements. The variances from the announcement time to the market opening time are higher than those before the announcement.

[1]Del Corral et al. (2003) find that nearly 93% of announcements are made during non-trading hours.

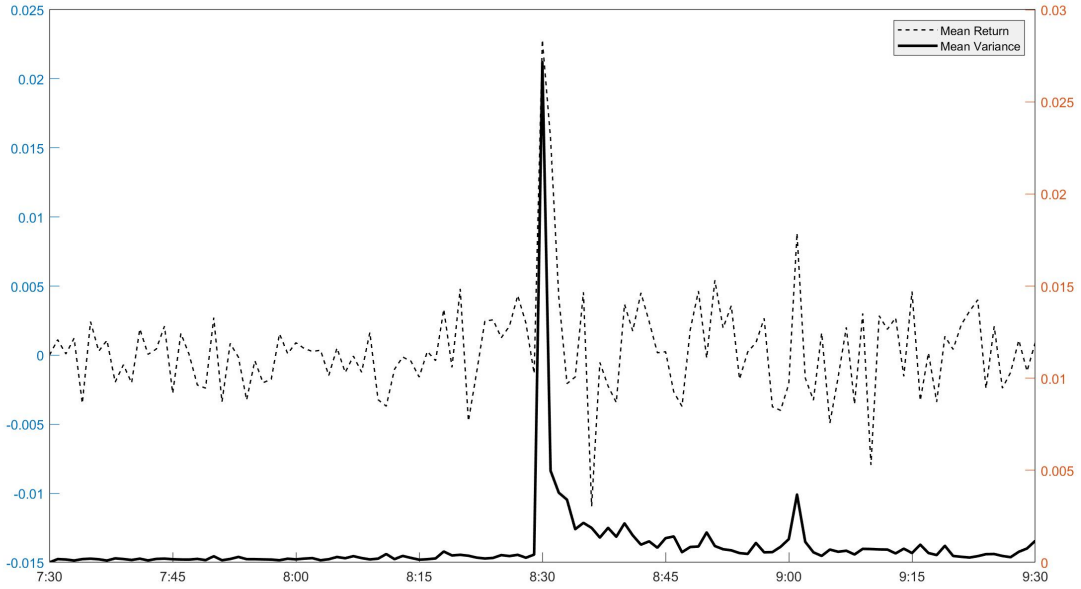

Figure 2 Instantaneous Market Reaction to Information in Overnight Macroeconomic Announcements

Notes: Figure 2 presents the averaged 1-minute returns and 1-minute variances (constructed by the square of 1-minute returns) of the S&P 500 E-mini futures on days that the U.S. Bureau of Labor Statistics announced the unemployment rate. The sample covers the period from July 2, 2003 to December 18, 2019. Unemployment announcement dates are obtained from the website of the U.S. Bureau of Labor Statistics (http://www.bls.gov).

The effect of market closure on stock (index) returns has been considered extensively in the literature. For example, weekend returns are lower than weekday returns (see, for example, French (1980), Gibbons and Hess (1981), and Keim and Stambaugh (1984)). Moreover, returns over trading periods are more volatile than returns over non-trading periods (see, for example, Fama (1965), French and Roll (1986), Oldfield and Rogalski (1980), and Amihud and Mendelson (1991)). The aforementioned studies examine the difference between the nature of overnight returns and that of intraday returns and have been extended to incorporate overnight information to enhance the accuracy of volatility forecasting (see, for example, Tsiakas (2008), Linton and Wu (2020), and Dhaene and Wu (2020)) and to manage risk within VaR models (Taylor (2007)).

However, few studies consider the influence of market closure on option pricing. Boes et al. (2007) model the overnight change in the stock prices by a single jump in addition to a standard affine model that allows for stochastic volatility and random jumps during

the day, and find that both overnight and intraday jumps are important for option pricing. Jones and Shemesh (2018) suggest that widespread and highly persistent option mispricing is driven by the incorrect treatment of stock return variance during periods of market closure. Muravyev and Ni (2020) find a remarkable day–night pattern: overnight average delta-hedged option returns are negative, whereas intraday average delta-hedged option returns are slightly positive. Wang et al. (2022) integrate overnight returns, intraday returns, and intraday realized volatility within an augmented autoregressive volatility model, in which overnight returns and intraday returns are assumed to be independent.

Building on these insights, this paper develops an option valuation model in which the underlying asset price features specific overnight and intraday variance dynamics. Our modeling framework explicitly prices options with distinct dynamics for overnight and intraday variations. To stress the difference in information, we use alternative model-free empirical measures driving overnight and intraday volatilities, respectively.

We find, for the SPX market over two separate periods, that both overnight and intraday volatilities are important for option pricing. Our new proposed model, the Bisected Realized GARCH, performs well in matching the historical and risk-neutral distributions of S&P 500 index returns. Specifically, when the Bisected Realized GARCH model is optimized on a dataset of S&P 500 index options, realized overnight and intraday variances, and returns, the model exhibits significantly better performance than various popular specifications. In particular, the overnight component accounts for an out-of-sample gain of 7.24% in pricing accuracy, relative to simply modeling total daily returns. This indicates that the overnight component should be taken into account when hedging risks during market closure.

One of the main contributions of this paper is a new and flexible option pricing model that can accommodate distinct overnight and intraday variance dynamics in the underlying asset price process. An important feature of this model is that the dynamics of overnight and intraday variances are governed by their nonparametric empirical proxies. Since these proxies are constructed in discrete time, our model contributes to the discrete-time family. The application of multivariate Edgeworth-Sargan density enables us to derive analytical approximations for option valuation formulas comprising several nested option-pricing specifications. This feature facilitates the estimation procedure, allows for a direct comparison between nested models, and avoids the need to resort to simulation techniques.

The remainder of this paper is organized as follows. Section 2 presents the theoretical background and empirical implementation of overnight and intraday volatility. Section 3 introduces our option pricing model. Section 4 describes the physical estimation. In Section 5, we specify the risk-neutralization procedure and derive analytical approximations for option valuation formulas. In Section 6, we evaluate empirical option pricing performance. Section 7 concludes.

2 Overnight and Intraday Volatility

This section outlines the theoretical background and empirical implementation of computing overnight and intraday components of the total daily variation. Hansen and Lunde (2005) construct overnight and intraday volatility with close-to-open return and intraday high-frequency data, which we describe first.

2.1 Realized Overnight and Intraday Volatility

For trading periods, high-frequency trades or quotes are ready to construct realized intraday variance,

$$RV_{OC,t} = \sum_{j=1}^{N_{OC,t}} r_{OC,t_j}^2, \tag{1}$$

where r_{OC,t_j}, $RV_{OC,t}$, and $N_{OC,t}$ are high-frequency returns, realized variances, and the number of observations during the open-to-close (OC in short) period. $r_{OC,t_j} = \log S_{OC,t_j} - \log S_{OC,t_{j-1}}, t < t_{OC,1} < t_{OC,2} < \cdots < t_{N_{OC,t}} < t+1$ are the time at which (trade or quote) prices are available.

To incorporate overnight information, Hansen and Lunde (2005) use close-to-open (hereafter CO) returns when constructing the total realized variance:

$$RV_t = r_{CO,t}^2 + RV_{OC,t}, \tag{2}$$

where $r_{CO,t}$ denotes the overnight return, computed as the log-difference between the open price of trading day t and the close price of the previous trading day:

$$r_{CO,t} = \log Open_t - \log Close_{t-1}. \tag{3}$$

Due to market closure, only two price observations are available to compute overnight return and variance.

In existing studies, it has been proved that most of the S&P 500 price discovery occurs in the S&P 500 E-mini futures market (Hasbrouck (2003)). Consequently, during non-trading hours of stock market, high-frequency price observations of its 24h-traded futures counterpart provide accurate measures of returns and variances. Heston and Nandi (2000) use S&P 500 futures prices to imply out S&P 500 index levels. Hu et al. (2021) calculate the realized volatility as the root of the sum of squared log returns of the S&P 500 E-mini futures sampled at 1-minute frequency, to further capture the uncertainty risk caused by macro-announcement during non-trading period. Craig et al. (1995) show that information during non-trading hours of stock market is efficiently reflected by its 24h-traded derivatives counterpart.

In light of this, we use overnight high-frequency prices of S&P 500 E-mini futures to imply out S&P 500 index levels during market closure, and construct implied overnight realized measures. The basic idea is that, the futures' price (F), is determined by its maturity (τ), risk-free rate (r), and the underlying spot price (S):

$$F = Se^{r\tau}, \tag{4}$$

thus, high-frequency index returns during non-trading hours can be implied by its 24h-traded futures counterpart (which matures at time T)

$$\begin{aligned} r^S_{t_j} &= \log S_{t_j} - \log S_{t_{j-1}} \\ &= \log\left(F_{t_j} e^{-r(T-t_j)}\right) - \log\left(F_{t_{j-1}} e^{-r(T-t_{j-1})}\right) \\ &= \log F_{t_j} - \log F_{t_{j-1}} + \underbrace{r(t_j - t_{j-1})}_{\approx 0} \\ &\approx r^F_{t_j}, \end{aligned} \tag{5}$$

which can be applied to construct the realized overnight variance:

$$RV_{CO,t} = \sum_{j=1}^{N_{CO,t}} \left(r^F_{CO,t_j}\right)^2, \tag{6}$$

where $r_{CO,t_j} = \log F_{CO,t_j} - \log F_{CO,t_{j-1}},\ t_{CO,1} < \cdots < t_{CO,N_{CO,t}}$ are the time at which high-frequency future prices are available during overnight period.

Since realized measures constructed using high-frequency data promote volatility forecasting accuracy (see Andersen et al. (2003) for example), the implied overnight realized measures

may contribute to better volatility forecasts of the underlying index, and thus to more accurate option prices.

2.2 Empirical Implementation

Following Hansen et al. (2022), the empirical measures of overnight and intraday variations on trading day t are given by

$$RV_{CO,t} = \sum_{j=1}^{N_{CO,t}} \left(r_{CO,t_j}\right)^2, \tag{7}$$

$$RV_{OC,t} = \sum_{j=1}^{N_{OC,t}} \left(r_{OC,t_j}\right)^2, \tag{8}$$

where $N_{CO,t}$ and $N_{OC,t}$ are the number of 5-minute overnight and intraday returns on trading day t, respectively. 5-minute returns are computed using the last tick price in each 5-minute interval. S&P 500 E-mini futures data come from TickData, and intraday realized variances are from the Realized Library of Oxford-Man.

2.3 Distinct Dynamics of Overnight and Intraday Volatilities

Figure 3 plots the series of total daily return and realized volatility (Graph A and B), overnight return and realized volatility (Graph C and D), and intraday counterparts (Graph E and F), from July 2, 2003 to December 18, 2019. We could see large swings in returns and high level of daily volatility during 2008 financial crisis. Comparing with Graph E and F, Graph C and D show rapidly growing importance of overnight variations after 2013, with averagely higher level of overnight volatility relative to the intraday counterpart. Besides, Graph D and F not only display synchronicity, but also show different dynamic patterns and evolution tracks of overnight and intraday volatility. For example, overnight periods are significantly more volatile than trading hours after 2018.

Table 1 reports summary statistics of returns and RV series, providing a straightforward access to distinct patterns and properties during trading and non-trading hours. Importantly, overnight variances have a higher persistence compared with intraday variances, indicated by a larger $AR(1)$ coefficient for RV_{CO}. In addition, overnight variances are averagely larger and more volatile than intraday variances. The above empirical regularities will help us to check the fitness of our model.

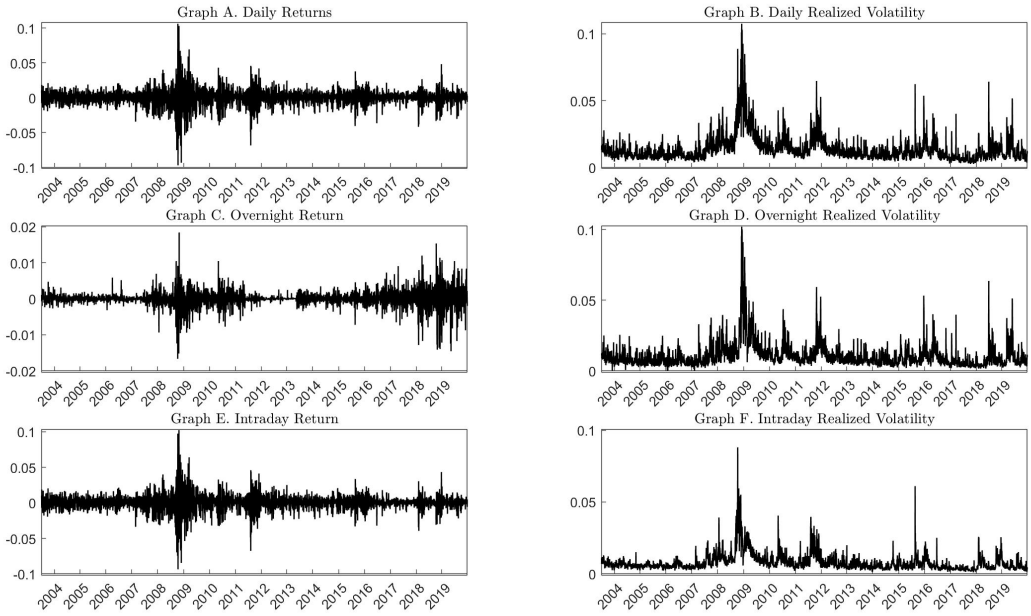

Figure 3 S&P 500 Returns and Realized Volatilities

Notes: Figure 3 presents the time trends of S&P 500 returns and volatilities. Specifically, we plot daily returns, r (Graph A); daily realized volatilities, $\sqrt{RV}$(Graph B); overnight returns, r_{CO} (Graph C); overnight realized volatilities computed using S&P 500 E-mini futures prices, $\sqrt{RV_{CO}}$ (Graph D); intraday returns, r_{OC} (Graph E); and intraday realized volatilities, $\sqrt{RV_{OC}}$ (Graph F). The sample period is from July 2, 2003 to December 18, 2019.

To sum up, different dynamic patterns of overnight and intraday variances provide supportive evidence for separate dynamic specifications. Thus, building a model that can accommodate distinct overnight and intraday variance dynamics is our next task.

Table 1 Summary Statistics of Historical Series

Summary Statistics of Historical Series

Series	Mean (%)	Median (%)	Std. Dev (%)	Skewness	Kurtosis	AR(1)
Return	7.2338	17.4072	17.7054	-0.3411	14.5768	-0.0867
Ovenight Return	2.2316	1.5509	3.3020	-0.4827	15.3239	-0.0445
Intraday Return	5.0022	15.1867	16.6365	-0.3176	15.3239	-0.0953
Realized Variance	22.0311	17.8777	15.2674	3.2678	19.8691	0.8414
Overnight Realized Variance	17.0182	13.3917	13.7093	3.6059	24.5662	**0.8219**
Intraday Realized Variance	12.3200	9.6300	9.4408	3.7296	26.3739	**0.7197**

Notes: Table 1 reports summary statistics of returns and realized variances. Sample mean, median, and standard deviation are annualized and presented in percentages. AR (1) notes the first-order autocorrelation coefficient. The sample covers from July 2, 2003 to December 18, 2019.

3 Modeling Overnight and Intraday Volatility

In this section, we develop an option pricing framework in which the underlying asset price features specific overnight and intraday variance dynamics. In our modeling framework, the dynamics of overnight and intraday variances are driven by their model-free realized measures constructed using high-frequency historical returns. The timely arrived overnight information and accurate description of overnight and intraday patterns lead to better volatility forecasts of the underlying index, and thus increase option pricing accuracy.

3.1 Key Objectives

In order to develop an option pricing model capturing distinct dynamics of overnight and intraday volatilities, following Hansen et al. (2022), we separate total daily return into overnight and intraday components:

$$r_t = r_{CO,t} + r_{OC,t}, \tag{9}$$

where $r_{CO,t}$ and $r_{OC,t}$ denote the overnight and intraday return on trading day t, respectively. Thus, the total daily variance, h_t, is expressed as

$$h_t = d_{CO,t}h_{CO,t} + d_{OC,t}h_{OC,t} + 2\rho_t\sqrt{d_{CO,t}h_{CO,t}d_{OC,t}h_{OC,t}}, \tag{10}$$

where $h_{CO,t}$ and $h_{OC,t}$ denote the overnight and intraday variance, respectively. Following Hansen et al. (2022), $d_{CO,t}$ and $d_{OC,t}$ are categorical variables capturing time-varying patterns of overnight and intraday variances. ρ_t is the correlation of $r_{CO,t}$ and $r_{OC,t}$. To obtain accurate volatility forecasts of the underlying index, dynamics of $h_{CO,t}$, $h_{OC,t}$, and ρ_t are key objectives of our discussion.

3.2 Return Equations

Under the framework of GARCH option pricing models (Duan (1995) and Christoffersen et al. (2013)), we consider the following specification:

$$r_{j,t} = r_j + \lambda_j\sqrt{d_{j,t}h_{j,t}} - \frac{1}{2}d_{j,t}h_{j,t} + \sqrt{d_{j,t}h_{j,t}}z_{j,t}, \tag{11}$$

where $j = \{CO, OC\}$, r_js are corresponding risk-free returns, and $d_{j,t}$s are categorical variables capturing time-varying patterns of overnight and intraday variances. Equation (11) enables us to interpret λ_{CO} and λ_{OC} as compensations for overnight and intraday volatility risk exposures, respectively. $z_{j,t}$s are stochastic terms following a bivariate joint normal distribution

$$\mathbf{z}_t = \begin{pmatrix} z_{CO,t} \\ z_{OC,t} \end{pmatrix} \overset{i.i.d}{\sim} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_t \\ \rho_t & 1 \end{pmatrix}\right), \tag{12}$$

where ρ_t is the correlation of $z_{CO,t}$ and $z_{OC,t}$.

3.2.1 Mapping the Correlations

Instead of modeling the conditional correlation, ρ_t, that is subject to $\rho_t \in [-1, 1]$, we will model ϱ_t that varies freely in $\mathbb{R}$. Following Hansen et al. (2014) and Archakov and Hansen (2020), we use the Fisher-Z transformation and relate to ϱ_t as:

$$\varrho_t = \operatorname{arctanh}\rho_t = \frac{1}{2}\log\left(\frac{1+\rho_t}{1-\rho_t}\right). \tag{13}$$

It is straightforward to find that the inverse transformation is also available in a closed form, because $\rho_t = (e^{2\varrho_t} - 1)/(e^{2\varrho_t} + 1)$.

The corresponding empirical quantities of correlations, Y_t, are subjected to the same transformation, and y_t is the empirical measure that we use to update the dynamics of ϱ_t:

$$y_t = \operatorname{arctanh} Y_t. \tag{14}$$

3.3 Incorporating Realized Measurements

Measurement equations describe the contemporaneous measurement relationship between latent variables and the corresponding observed empirical quantities:

$$\log RV_{j,t} = \xi_j + \phi_j(\log d_{j,t} + \log h_{j,t}) + \delta(z_{j,t}) + \sigma_{v_j} v_{j,t}, \tag{15}$$

$$y_t = \widetilde{\xi} + \widetilde{\phi}\varrho_t + \sigma_{\widetilde{v}}\widetilde{v}_t. \tag{16}$$

$\delta(z_{j,t})$ is a leverage function that captures dependencies between return and volatility innovations. This dependency is known to be empirically important, and is referred to as the leverage effect (Black (1976), Christie (1982), and Engle and Ng (1993)). Motivated by findings in Hansen et al. (2012) and Hansen and Huang (2016), a second-order Hermite polynomial is sufficient for capturing the asymmetry dependence between return and volatility innovations, we focus on a parsimonious leverage function written as a second-order Hermite polynomial in this paper, i.e.

$$\delta(z_{j,t}) = \delta_1 z_{j,t} + \delta_2(z_{j,t}^2 - 1). \tag{17}$$

This choice is convenient because it ensures that $\mathbb{E}(\delta(z_{j,t})) = 0$, for any distribution with $\mathbb{E}(z_{j,t}) = 0$ and $Var(z_{j,t}) = 1$. The second-order Hermite polynomial form is also convenient in the derivation of option pricing formula and quasi-likelihood analysis.

The measurement equations involve the error terms, $\mathbf{u}_t = (v_{CO,t}, v_{OC,t}, \widetilde{v}_t)'$. Following Archakov and Hansen (2020), we specify $\mathbf{u}_t$ to be

$$\mathbf{u}_t = \begin{pmatrix} v_{CO,t} \\ v_{OC,t} \\ \widetilde{v}_t \end{pmatrix} \overset{i.i.d}{\sim} N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{v_{CO},v_{OC}} & \rho_{v_{CO},\widetilde{v}} \\ \rho_{v_{CO},v_{OC}} & 1 & \rho_{v_{OC},\widetilde{v}} \\ \rho_{v_{CO},\widetilde{v}} & \rho_{v_{OC},\widetilde{v}} & 1 \end{pmatrix} \right), \tag{18}$$

and independent of $\mathbf{z}_t$,

$$corr(\mathbf{z}_t, \mathbf{u}_t) = 0. \tag{19}$$

$RV_{CO,t}$, $RV_{OC,t}$, and y_t are the observed realized measurements of the latent variables, $h_{CO,t}$, $h_{OC,t}$, and ϱ_t, respectively, and these will be used to inform the model about time variation in the latent variables.

3.4 Dynamic Equations

We are now ready to specify the dynamic equations that describe how latent variables depend on past observables. Following Huang et al. (2017), we use GARCH dynamics, where the information provided by realized measures at trading day t is used to update the forecast of latent variables on $t+1$. The dynamics for overnight and intraday variances and the transformed correlations are as follows:

$$\log h_{j,t+1} = \omega_j + \beta_j \log h_{j,t} + \tau(z_{j,t}) + \gamma_j(\log RV_{j,t} - \log d_{j,t}), \tag{20}$$

$$\varrho_t = \tilde{\omega} + \tilde{\beta}\varrho_{t-1} + \tilde{\gamma} y_{t-1}, \tag{21}$$

where $\tau(z_{j,t})$ is also a leverage function.

4 Estimation on the Underlying Asset

4.1 Quasi Maximum Likelihood Estimation

We estimate the models using Quasi Maximum Likelihood Estimation (QMLE). Under the Gaussian specification, the log-likelihood for returns at time $t+1$ conditional on observations at time t can be written as

$$\ln \mathcal{L}_t^{\mathbb{P}}\left(\mathrm{r}_{j,t+1}\right) = -\frac{1}{2}\ln\left(2\pi \operatorname{var}_t\left[\mathrm{r}_{t+1}\right]\right) - \frac{\left(\mathrm{r}_{j,t+1} - \mathbb{E}_t\left[\mathrm{r}_{j,t+1}\right]\right)^2}{2\operatorname{var}_t\left[\mathrm{r}_{j,t+1}\right]}. \tag{22}$$

Similarly, the log-likelihood for realized variances at time $t+1$ conditional on observations at time t is given by

$$\ln \mathcal{L}_t^{\mathbb{P}}\left(\mathrm{RV}_{j,t+1}\right) = -\frac{1}{2}\ln\left(2\pi \operatorname{var}_t\left[\mathrm{RV}_{j,t+1}\right]\right) - \frac{\left(\mathrm{RV}_{j,t+1} - \mathbb{E}_t\left[\mathrm{RV}_{j,t+1}\right]\right)^2}{2\operatorname{var}_t\left[\mathrm{RV}_{j,t+1}\right]}, \tag{23}$$

where $j \in \{CO, OC\}$. Thus, the log-likelihood function for returns and realized variances is given by summing the log likelihoods over all the observations, yielding

$$\ln \mathcal{L}_{\mathbb{P}} = \sum_{t=1}^{T} \sum_{j\in\{CO,OC\}} \ln \mathcal{L}_t^{\mathbb{P}}\left(\mathrm{r}_{j,t+1}\right) + \sum_{t=1}^{T} \sum_{j\in\{CO,OC\}} \ln \mathcal{L}_t^{\mathbb{P}}\left(\mathrm{RV}_{j,t+1}\right). \tag{24}$$

4.2 Nested Models

Before the empirical estimation, we discuss special cases of interest that are nested by the new proposed option pricing model. In the following empirical studies, we compare the performance of the Bisected Realized GARCH model to these nested specifications and other popular models.

4.2.1 The Realized GARCH Model

One of the main contributions of this paper is a new option pricing model that can accommodate distinct overnight and intraday variance dynamics in the underlying asset price process. To achieve that, we have separate dynamic specifications for overnight and intraday variances. If the patterns and properties of the overnight components are ignored, and simply model total daily variations, return equation will become

$$r_{t+1} = r + \lambda\sqrt{h_{t+1}} - \frac{1}{2}h_{t+1} + \sqrt{h_{t+1}}z_{t+1}, \tag{25}$$

where r denotes the risk-free return, and z_t are i.i.d. $\mathcal{N}(0,1)$. Similarly, dynamic equation and measurement equation will become

$$\log h_{t+1} = \omega + \beta \log h_t + \tau(z_t) + \gamma\sigma u_t, \tag{26}$$

$$\log RV_t = \xi + \phi \log h_t + \delta(z_t) + \sigma u_t. \tag{27}$$

Again, the model specifies $\text{corr}(z_t, u_t) = 0$. This specification is identical to the log-linear Realized GARCH (RG) model by Hansen et al. (2012) and Huang et al. (2017).

4.2.2 The EGARCH model

The realized measurements constructed by high-frequency data are used to inform the model about time variation in the latent variables. Furthermore, if the dynamics of all latent variables are only driven by daily observables, the dynamic equation become

$$\log h_{t+1} = \omega + \beta \log h_t + \tau(z_t),$$

which is precisely the EGARCH model (EG) by Duan et al. (1997) and Duan et al. (2006).

4.3 Full Sample Estimation

Following Hansen et al. (2022), we use $d_{CO,t}$ to explicitly capture the announcement effect of important macroeconomic releases on overnight volatilities. In our empirical study, we focus on the most important U.S. macroeconomic announcement, namely the nonfarm payroll employment release. The nonfarm payroll is among the most significant of the announcements for all of the markets, and it is often referred to as the "king" of announcements by market participants; see, e.g., Andersen and Bollerslev (1998). Since these news announcements are released at 8:30 Eastern Standard Time (EST) when the futures markets are open, but the equity markets closed, we set $d_{CO,t}$ to θ on announcement days, otherwise $d_{CO,t} = 1$, that is,

$$d_{CO,t} = \begin{cases} \theta, & \text{announcement day} \\ 1, & \text{otherwise} \end{cases} \tag{28}$$

and $d_{OC,t} \equiv 1$.

Table 2 reports the estimation results under $\mathbb{P}$ measure. We compare the new proposed model to nested specifications discussed in Section 4.2 as well as popular settings, the Generalized Affine Realized Volatility (GARV) model by Christoffersen et al. (2014) and the Heston-Nandi GARCH (HNG) model by Heston and Nandi (2000). For each estimation, we report the coefficients, and the corresponding standard errors are presented in parentheses. To allow for a direct comparison, we also include log likelihoods, $\mathcal{L}$, and volatility persistence, $\pi^{\mathbb{P}}$. The sample period is from July 2, 2003 to December 18, 2019.

For all models, the estimates of λ are positive and statistically significant, indicating positive volatility risk compensations for both trading and non-trading hours. Moreover, for the BRG and the RG model, the estimates of τ and δ describe the leverage effect, which is in line with previous studies (see Hansen et al. (2012) and Hansen et al. (2014) for example).

Table 2 provides supportive evidence for distinct dynamics in overnight and intraday volatilities. The volatility persistence, $\pi^{\mathbb{P}}$, is higher for overnight period, which is consistent with the higher $AR(1)$ coefficients of overnight variance reported in Table 1. The volatility persistence estimates are all above 0.96, which is in line with the established empirical results. Besides, the coefficients of the "OC" components of the BRG model exhibit high similarity with the parameters of the RG model using close-to-open returns to update overnight variations (i.e. $RV_t = r^2_{CO,t} + RV_{OC,t}$). This similarity further confirms that r^2_{CO} contains only little overnight information, and thus we need realized measures constructed using high-frequency

Table 2 Full Sample Estimation Under $\mathbb{P}$ Measure

Estimation on Historical Returns and Realized Measurements

Parameters	BRG		RG	GARV	HNG	EG
	CO	OC				
ω	0.1584	-0.0134	-0.0287		9.19E-07	-0.1256
	(0.0132)	(0.0083)	(0.0163)		(1.14E-05)	(0.0106)
λ	0.2620	0.3479	0.3467	0.4935	0.5436	0.4673
	(0.0158)	(0.0161)	(0.0050)	(1.61E-04)	(0.0183)	(0.0161)
β	**0.6433**	**0.7077**	0.9704	0.9692	0.6061	0.9860
	(0.0056)	(0.0137)	(0.0044)	(0.0002)	(0.0031)	(0.0034)
τ_1	-0.0706	-0.0467	-0.1088	9.42E-11	1.54E-06	0.1522
	(0.0056)	(0.0059)	(0.0071)	(8.76E-13)	(5.21E-09)	(0.0122)
τ_2	0.0002	0.0199	0.0516	473.9283	505.0916	-0.8489
	(0.0005)	(0.0031)	(0.0038)	2.3032	(2.7838)	(0.0813)
γ	**0.3345**	**0.2360**	0.2606			
	(0.0141)	(0.0121)	(0.0129)			
ξ	0.3716	-0.0576	-0.0171			
	(0.0367)	(0.0291)	(0.0346)			
ϕ	1.0038	1.1178	1.1743	1.50E-06		
	(0.0368)	(0.0242)	(0.0263)	(1.10E-08)		
δ_1	-0.0268	-0.1194	-0.1193	5.97E-07		
	(0.0080)	(0.0093)	(0.0092)	(3.46E-11)		
δ_2	0.0710	0.0836	0.0962	538.5526		
	(0.0033)	(0.0047)	(0.0051)	(1.47E-02)		
σ	0.4831	0.5557	0.5485	0.5194		
	(0.0053)	(0.0063)	(0.0060)	(2.13E-03)		
κ				0.4423		
				(2.15E-07)		
θ	1.0506			0.3535		
	(0.0304)			(5.37E-03)		
ρ	0.3636					
	(0.0154)					
$\log h_1$	-2.4867	-1.9583	0.0655	0.0130	0.6581	-0.0361
	(0.4299)	(0.4078)	(0.3212)	(3.93E-04)	0.0635	0.4587
$\pi^{\mathbb{P}}$	**0.9791**	**0.9715**	0.9704	0.9692	0.9970	0.9860
$\mathcal{L}$	-12299		-7650	-8007	-6171	-1454

Notes: Table 2 reports estimation results under $\mathbb{P}$ measure. The Bisected Realized GARCH (BRG) model is estimated on close-to-open returns, open-to-close returns, close-to-open realized variances, and open-to-close realized variances. The Realized GARCH (RG) model and the Generalized Affine Realized Volatility (GARV) model are estimated on close-to-close returns and daily realized variances. The Heston-Nandi GARCH (HNG) model and the EGARCH (EG) model are estimated on close-to-close returns. We report the estimated coefficients, and the corresponding standard errors are presented in parentheses. Note that for GARV and HNG model, the parameters ω can be inferred by $\mathbb{E}[\bar{h}]$.

data from contemporaneous futures trading to timely and accurately incorporate overnight market information. In addition, the macroeconomic announcements released during non-trading hours lead to a higher overnight volatility, indicated by $\theta > 1$.

5 Risk Neutralization and Option Valuation

5.1 Risk Neutralization

To derive the option valuation formula, we need to obtain the dynamics under the risk neutral measure. Following Christoffersen et al. (2010), we apply the exponentially affine stochastic discount factor (SDF) as below to derive the dynamics under $\mathbb{Q}$ measure.

$$Z_{t+1} = \frac{\exp(\nu_{CO,t} z_{CO,t+1} + \nu_{OC,t} z_{OC,t+1} + \theta_{CO,t} v_{CO,t+1} + \theta_{OC,t} v_{OC,t+1} + \widetilde{\theta}_t \widetilde{v}_{t+1})}{\mathbb{E}_t[\exp(\nu_{CO,t} z_{CO,t+1} + \nu_{OC,t} z_{OC,t+1} + \theta_{CO,t} v_{CO,t+1} + \theta_{OC,t} v_{OC,t+1} + \widetilde{\theta}_t \widetilde{v}_{t+1})]},$$

The non-arbitrage conditions, namely that the expectations of returns under $\mathbb{Q}$ measure equal to the risk-free returns r_j in the corresponding period, yield

$$\mathbb{E}_t^Q[\exp(r_{j,t+1})] = \exp(r_j), \quad j \in \{CO, OC\}. \tag{29}$$

Note that Z_{t+1} builds the relationship between $\mathbb{P}$ measure and $\mathbb{Q}$ measure,

$$\mathbb{E}_t[Z_{t+1} \exp(r_{j,t+1})] = \mathbb{E}_t^Q[\exp(r_{j,t+1})], \quad j \in \{CO, OC\}, \tag{30}$$

we have

$$\mathbb{E}_t[Z_{t+1} \exp(r_{j,t+1})] = \exp(r_j), \quad j \in \{CO, OC\}. \tag{31}$$

Then we consider the risk-neutral moment generating function (MGF) and the sufficient condition to derive the model under $\mathbb{Q}$ measure. The moment generating function of the five random shocks can be written as

$$\begin{aligned} &\mathbb{E}_t^Q[\exp(s_1 z_{CO,t+1} + s_2 z_{OC,t+1} + s_3 v_{CO,t+1} + s_4 v_{OC,t+1} + s_5 \widetilde{v}_{t+1})] \\ &= \mathbb{E}_t[Z_{t+1} \exp(s_1 z_{CO,t+1} + s_2 z_{OC,t+1} + s_3 v_{CO,t+1} + s_4 v_{OC,t+1} + s_5 \widetilde{v}_{t+1})] \end{aligned} \tag{32}$$

We maintain the nature of the model by mapping innovations from physical measures to risk-neutral measures, as follows:

$$z^*_{CO,t+1} = z_{CO,t+1} + \lambda_{CO}, \tag{33}$$

$$z^*_{OC,t+1} = z_{OC,t+1} + \lambda_{OC}, \tag{34}$$

$$v^*_{CO,t+1} = v_{CO,t+1} - \chi_1, \tag{35}$$

$$v^*_{OC,t+1} = v_{OC,t+1} - \chi_2, \tag{36}$$

$$\widetilde{v}^*_{t+1} = \widetilde{v}_{t+1} - \chi_3, \tag{37}$$

where χ_1, χ_2, χ_3 are risk compensations.

Now we are ready to obtain the dynamic model under $\mathbb{Q}$ measure. Similar to the Bisected Realized GARCH model under $\mathbb{P}$ measure, the model under $\mathbb{Q}$ measure also consists of return equations (1* to 3*), dynamic equations (4* to 6*), and measurement equations (7* to 9*). $\mathbf{z}^*_\mathbf{t}$ and $\mathbf{u}^*_\mathbf{t}$ are independently normally distributed under $\mathbb{Q}$ measure.

$$r_t = r_{CO,t} + r_{OC,t}, \tag{1*}$$

$$r_{CO,t} = r_{CO} - \frac{1}{2} d_{CO,t} h_{CO,t} + \sqrt{d_{CO,t} h_{CO,t}}\, z^*_{CO,t}, \tag{2*}$$

$$r_{OC,t} = r_{OC} - \frac{1}{2} d_{OC,t} h_{OC,t} + \sqrt{d_{OC,t} h_{OC,t}}\, z^*_{OC,t}, \tag{3*}$$

$$\log h_{CO,t} = \omega_{CO} + \beta_{CO} \log h_{CO,t-1} + \tau(z^*_{CO,t}) + \gamma_{CO}(\log RV_{CO,t-1} - \log d_{CO,t-1}), \tag{4*}$$

$$\log h_{OC,t} = \omega_{OC} + \beta_{OC} \log h_{OC,t-1} + \tau(z^*_{OC,t}) + \gamma_{OC}(\log RV_{OC,t-1} - \log d_{OC,t-1}), \tag{5*}$$

$$\varrho_t = \widetilde{\omega} + \widetilde{\beta} \varrho_{t-1} + \widetilde{\gamma} \widetilde{y}_{t-1}, \tag{6*}$$

$$\log RV_{CO,t} = \xi^*_{CO} + \phi_{CO}(\log d_{CO,t} + \log h_{CO,t}) + \delta(z^*_{CO,t}) + \sigma_{v_{CO}} v^*_{CO,t}, \tag{7*}$$

$$\log RV_{OC,t} = \xi^*_{OC} + \phi_{OC}(\log d_{OC,t} + \log h_{OC,t}) + \delta(z^*_{OC,t}) + \sigma_{v_{OC}} v^*_{OC,t}, \tag{8*}$$

$$y_t = \widetilde{\xi}^* + \widetilde{\phi} q_t + \sigma_{\widetilde{v}} \widetilde{v}^*_t. \tag{9*}$$

The above risk neutralization process is detailed in Appendix A.

5.2 Multivariate Approximation Density

Due to the log-linear specification, models under the Realized GARCH framework do not have an analytical moment generating function (MGF). Following Huang et al. (2017), we apply the analytical approximation method to derive an option pricing formula for the Bisected Realized GARCH model. Furthermore, the Bisected Realized GARCH model accommodate distinct dynamics in overnight and intraday components, therefore we resort to a multivariate approximation approach. Following Jarrow and Rudd (1982), Duan et al. (1997), and Duan et al. (2006), we consider the multivariate version of the Edgeworth expansion, the

multivariate Edgeworth-Sargan density, to analytically approximate the joint density of the cumulative close-to-open and open-to-close returns. According to Perote (2004), the multivariate Edgeworth-Sargan density shows better performance than many other densities (for example, the multivariate Student-t density) in fitting financial data.

Specifically, the joint density of $\mathbf{z} = (z_{CO}, z_{OC})'$ can be approximated by

$$p(z_{CO}, z_{OC}) = \phi(z_{CO}, z_{OC}) + g(z_{CO})g(z_{OC})(q(z_{CO}) + q(z_{OC})), \quad -\infty < z_{CO}, z_{OC} < +\infty \tag{38}$$

where $\phi(z_{CO}, z_{OC})$ is a zero-mean bivariate normal density with a covariance matrix

$$\mathbf{\Sigma}_{\phi} = \begin{pmatrix} \sigma^2_{z_{CO}} & \sigma_{z_{CO},z_{OC}} \\ \sigma_{z_{CO},z_{OC}} & \sigma^2_{z_{OC}} \end{pmatrix},$$

and its marginal densities are $g(z_{CO})$ and $g(z_{OC})$. In addition,

$$q(z_{CO}) = \sum_{s=2}^{S_{CO}} w_{CO,s} H_s(v_{CO}),\ q(z_{OC}) = \sum_{s=2}^{S_{OC}} w_{OC,s} H_s(v_{OC})$$

are linear combinations of the first S_j Hermite polynomials of $g(z_{CO})$ and $g(z_{OC})$, where $v_{CO} = \frac{z_{CO}}{\sigma_{z_{CO}}}$ and $v_{CO} = \frac{z_{OC}}{\sigma_{z_{OC}}}$. We follow Huang et al. (2017) and keep only cumulants below fourth-order in our expansion.

5.3 Option Valuation Formula

Under the risk neutral measure, the European call option price is given by

$$e^{-rT}\mathbb{E}_0^{\mathbb{Q}}\left(\max(S_T - K), 0\right) = e^{-rT} \iint_D (S_T - K)\widetilde{p}(\mathbf{z}) d\mathbf{z} \tag{39}$$

where T denotes the maturity of the option and K is the strike price. $\widetilde{p}(\mathbf{z})$ describes the joint PDF of normalized cumulative overnight and intraday return under $\mathbb{Q}$ measure. S_T denotes the price of underlying asset at time T, whose randomness comes from $\mathbf{z}$ completely.

$$\begin{aligned} S_T &= S_0\ \exp(R_{CO,T} + R_{OC,T}) \\ &= S_0\ \exp(\overline{\mu}_{CO} - \sigma_{CO} z_{CO} + \overline{\mu}_{OC} - \sigma_{OC} z_{OC}) \end{aligned} \tag{40}$$

where $R_{j,T}$ is the cumulative return under $\mathbb{Q}$ measure, $\overline{\mu}_j$ and σ_j^2 are the corresponding mean and variance, respectively.

$$R_{j,T} = \sum_{t=1}^{T} \left(r_j - \frac{1}{2} d_{j,t} h_{j,t} + \sqrt{d_{j,t} h_{j,t}} z^*_{j,t} \right), \tag{41}$$

$$\overline{\mu}_j = \mathbb{E}_0^Q[R_{j,T}], \quad \sigma_j^2 = \mathbb{E}_0^Q[R_{j,T}^2] - (\mathbb{E}_0^Q[R_{j,T}])^2. \tag{42}$$

The variable of integration, $\mathbf{z}$, is the normalized cumulative return under $\mathbb{Q}$ measure, whose joint density can be approximated by the multivariate Edgeworth-Sargan density discussed in Section 5.2.

Proposition 1. *The price of a European call option matures at time T is given by*

$$C_{approx} = C + \kappa_{CO,3}A_{CO,3} + \kappa_{OC,3}A_{OC,3} + (\kappa_{CO,4}-3)A_{CO,4} + (\kappa_{OC,4}-3)A_{OC,4} + \kappa_{CO,3}^2 A_{CO,6} + \kappa_{OC,3}^2 A_{OC,6},$$

where

$$\begin{aligned}
C &= S_0 e^{\Delta\sigma}\Phi(D';\Sigma) - Ke^{-rT}\Phi(D;\Sigma),\\
A_{CO,3} &= \frac{1}{6}S_0 e^{\delta\sigma}\sigma_{CO}\left(\sigma_{CO}^2\Phi(D'';\Sigma_0) + \frac{\sigma_{CO}^2}{\sigma^2}(2\sigma - d)\phi(d)\right),\\
A_{OC,3} &= \frac{1}{6}S_0 e^{\delta\sigma}\sigma_{OC}\left(\sigma_{OC}^2\Phi(D'';\Sigma_0) + \frac{\sigma_{OC}^2}{\sigma^2}(2\sigma - d)\phi(d)\right),\\
A_{CO,4} &= \frac{1}{24}S_0 e^{\delta\sigma}\sigma_{CO}\left(\sigma_{CO}^3\Phi(D'';\Sigma_0) + \frac{\sigma_{CO}^3}{\sigma^3}(d^2 - 1 - 3\sigma(d-\sigma))\phi(d)\right),\\
A_{OC,4} &= \frac{1}{24}S_0 e^{\delta\sigma}\sigma_{OC}\left(\sigma_{OC}^3\Phi(D'';\Sigma_0) + \frac{\sigma_{OC}^3}{\sigma^3}(d^2 - 1 - 3\sigma(d-\sigma))\phi(d)\right),\\
A_{CO,6} &= \frac{1}{72}S_0 e^{\delta\sigma}\sigma_{CO}\left(\sigma_{CO}^5\Phi(D'';\Sigma_0) + \frac{\sigma_{CO}^5}{\sigma^5}(3 - 6d^2 + d^4 + 5\sigma(d - (d-\sigma)(\sigma d - 2) - (d-\sigma)^3))\phi(d)\right),\\
A_{OC,6} &= \frac{1}{72}S_0 e^{\delta\sigma}\sigma_{OC}\left(\sigma_{OC}^5\Phi(D'';\Sigma_0) + \frac{\sigma_{OC}^5}{\sigma^5}(3 - 6d^2 + d^4 + 5\sigma(d - (d-\sigma)(\sigma d - 2) - (d-\sigma)^3))\phi(d)\right).
\end{aligned} \tag{43}$$

The expressions and notations are given in Appendix B.

Proof. See Appendix C. □

6 Option Pricing Performance

To examine whether overnight variance dynamics and implied overnight realized measures help to improve option pricing accuracy, in this section, we compare the option pricing performance of the Bisected Realized GARCH model to a set of benchmarks. We first describe the option dataset, and then estimate models to obtain coefficients which will be used to conduct out-of-sample evaluations. Finally, we compare the pricing performance across models in different out-of-sample windows.

6.1 Option Dataset

We use daily close prices of European options on the S&P 500 index. Option data are from OptionMetrics and VIX data are from CBOE. The sample period covers from July 2, 2003 to December 18, 2019. Following Christoffersen et al. (2014), we use out-of-the-money (OTM) put and call options with positive trading volume and with maturity between 15 and 180 days, and out-of-money put option prices are converted to in-the-money call option prices using the put-call parity. For each maturity quoted on each Wednesday and Thursday, we use the six most liquid strike prices, which finally yields a sample of 74,848 option contracts.

Table 3 presents summary statistics, and the option data are sorted by moneyness (Panel A), maturity (Panel B), and the VIX level (Panel C). In each panel, we report the number of contracts, the average option price, the average Black-Scholes implied volatility, and the average bid-ask spread in dollars.

In Panel A, the measure of moneyness is the Black-Scholes delta defined as

$$\Delta = \Phi\left(\frac{\ln(S_0/K) + rT + 1/2(IV^{MKT})^2 T/365}{IV^{MKT}\sqrt{T/365}}\right), \tag{44}$$

where $\Phi(\cdot)$ stands for the normal cumulative distribution function (CDF), r is the annualized risk-free rate, T is maturity, and IV^{MKT} is the annualized implied Black-Scholes volatility computed using the market price of the option. A few empirical regularities emerge at this point. We observe that the deep out-of-the-money (OTM) puts, i.e., contracts with deltas exceeding 0.7, are relatively expensive. The implied volatility for those options, 20.42%, is higher compared with 15.07% for options with delta between 0.4 and 0.5, and 13.08% for options with delta below 0.3. The data provide supportive evidence for the smirk pattern across moneyness.

In Panel B, the option data are sorted by calendar days to maturity. We observe that options with longer maturities are generally more expensive, which is in line with previous findings, see Christoffersen et al. (2014), among others.

Panel C reports results sorted by the VIX level. As expected, option prices, implied volatilities, and dollar spreads increase with VIX. We conclude from Panel C that most of our sample days come from mild periods, with VIX levels less than 25%.

The full dataset includes three subsamples. Data on Wednesdays from July 2, 2003 to December 30, 2015 (in-sample) are used to estimate models, and obtain coefficients which will

Table 3 Summary Statistics of the S&P 500 Index Option Data

S&P 500 Index Option data (2003-2019)

	Delta<0.3	0.3⩽Delta<0.4	0.4⩽Delta<0.5	0.5⩽Delta<0.6	0.6⩽Delta<0.7	Delta⩾0.7	All
Panel A. By Moneyness							
Number of contracts	10,218	5,639	8,844	13,372	8,813	27,962	74,848
Average price	11.2995	22.0336	36.1877	52.9316	76.7431	174.0350	90.9878
Average implied volatility	13.0793	13.9505	15.0700	16.4939	17.6432	20.4244	17.2715
Average bid-ask spread	1.0713	1.3280	1.5645	1.7116	1.5371	1.0729	1.3188
	DTM<30	**30⩽DTM<60**	**60⩽DTM<90**	**90⩽DT<120**	**120⩽DTM<150**	**DTM⩾150**	**All**
Panel B. By Maturity							
Number of contracts	15,581	23,155	13,355	9,988	6,954	5,815	74,848
Average price	42.8841	74.2270	97.1204	121.4386	148.0314	152.0155	90.9878
Average implied volatility	14.6657	16.6491	18.3513	19.3903	18.9056	18.6584	17.2715
Average bid-ask spread	0.7992	1.1160	1.5509	1.7245	1.6927	1.8414	1.3188
	VIX<15	**15⩽VIX<20**	**20⩽VIX<25**	**25⩽VIX<30**	**30⩽VIX<35**	**VIX⩾35**	**All**
Panel C. By VIX Level							
Number of contracts	36,003	22,588	8,712	3,728	1,503	2,314	74,848
Average price	86.8972	94.8711	97.6983	94.7407	97.4913	81.1917	90.9878
Average implied volatility	13.1344	17.4413	21.9962	25.2239	29.7880	41.2515	17.2715
Average bid-ask spread	0.9354	1.3614	1.7218	2.0618	2.5252	3.3691	1.3188

Notes: This table reports summary statistics of the S&P 500 index option data sorted by moneyness (Panel A), maturity (Panel B), and the VIX level (Panel C). In Panel A, moneyness is measured by delta computed from the Black-Scholes model. In Panel B, DTM denotes calendar days to maturity. In Panel C, VIX level stands for CBOE Volatility Index. Option data are from OptionMetrics and VIX data are from CBOE. The sample period is from July 2, 2003 to December 18, 2019. We keep data on Wednesdays and Thursdays, and option prices were preprocessed following Christoffersen et al. (2014).

be used to compare the out-of-sample option pricing performance across models. We conduct the out-of-sample evaluations in two different windows. The first is Thursdays from July 3, 2003 to December 31, 2015, and the second is Wednesdays from January 6, 2016 to December 18, 2019.

Table 4 describes the subsamples. For each subsample, we sort data by moneyness and maturity, and report the number of options as well as the average of implied volatilities in each category. Similar to the full sample, one-third to a half of option contracts are deep out-of-money puts (i.e. $\Delta > 0.7$), and options with maturity less than 60 days account for over a half for each subsample. Among different subsamples, the market volatility is lower after 2016. To be more specific, the average of implied volatilities on Wednesdays after 2016 is 0.1466, which is 21.18% lower than that from 2003 to 2015.

6.2 The Joint Likelihood Function

Following a large body of the derivatives literature (see Trolle and Schwartz (2009) and Kanniainen et al. (2014), among others), we minimize the vega-weighted RMSE (VWRMSE) in the option-based Gaussian likelihood component:

$$VWRMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\frac{P_n^{MKT} - P_n^{MOD}}{BSV_n^{MKT}}\right)^2}, \tag{45}$$

where P_n^{MOD} represents the model-based price for option n, P_n^{MKT} is the market price for option n, and BSV_n^{MKT} denotes the Black-Scholes vega (Black and Scholes (1973)), the derivative with respect to volatility, computed using the market implied volatility.

Now we are ready to estimate the model with the joint likelihood function of the observed variables and pricing errors. Specifically, we solve the following joint optimization problem:

$$\max \ln \mathcal{L}_P + \ln \mathcal{L}_o, \tag{46}$$

where

$$\begin{aligned} \ln \mathcal{L}_o &= -\frac{N}{2}\log(2\pi) - \frac{N}{2}\sum_{n=1}^{N}\log(\sigma_e^2) - \sum_{n=1}^{N}\frac{((P_n^{MOD} - P_n^{MKT})/BSV_n^{MKT})^2}{2\sigma_e^2} \\ &\propto -\frac{1}{2}\sum_{n=1}^{N}\left(\log(\sigma_e^2) + \frac{VWRMSE^2}{\sigma_e^2}\right). \end{aligned} \tag{47}$$

Table 4 Option Dataset Summary for Subsamples

Panel A. WED, 2003-2015

		Maturity						
		<30	30-60	60-90	90-120	120-150	>150	Total
Delta	<0.3	414	946	671	534	361	348	3,274
		(0.1459)	(0.1403)	(0.1458)	(0.1514)	(0.1484)	(0.1462)	(0.1455)
	0.3-0.4	438	619	318	252	150	131	1,908
		(0.1396)	(0.1467)	(0.1615)	(0.1780)	(0.1691)	(0.1694)	(0.1550)
	0.4-0.5	640	986	607	442	239	238	3,152
		(0.1449)	(0.1622)	(0.1697)	(0.1829)	(0.1748)	(0.1754)	(0.1650)
	0.5-0.6	789	1449	1141	738	384	382	4,885
		(0.1566)	(0.1663)	(0.1833)	(0.2018)	(0.1928)	(0.1947)	(0.1784)
	0.6-0.7	615	977	609	462	285	305	3,254
		(0.1688)	(0.1821)	(0.1911)	(0.2103)	(0.2039)	(0.2012)	(0.1890)
	>0.7	1,064	2,869	2,063	1,355	878	791	9,020
		(0.1973)	(0.2090)	(0.2230)	(0.2347)	(0.2263)	(0.2241)	(0.2177)
	Total	3,960	7,846	5,409	3,783	2,297	2,198	25,493
		(0.1645)	(0.1787)	(0.1918)	(0.2037)	(0.1966)	(0.1949)	(0.1860)

Panel B. THU, 2003-2015

		Maturity						
		<30	30-60	60-90	90-120	120-150	>150	Total
Delta	<0.3	537	855	737	584	393	291	3,397
		(0.1352)	(0.1393)	(0.1431)	(0.1513)	(0.1497)	(0.1454)	(0.1433)
	0.3-0.4	535	508	346	246	183	134	1,952
		(0.1349)	(0.1487)	(0.1584)	(0.1758)	(0.1673)	(0.1649)	(0.1529)
	0.4-0.5	720	928	590	422	275	186	3,121
		(0.1395)	(0.1621)	(0.1679)	(0.1799)	(0.1761)	(0.1788)	(0.1626)
	0.5-0.6	915	1,376	1,179	798	409	333	5,010
		(0.1539)	(0.1697)	(0.1788)	(0.1976)	(0.1967)	(0.1855)	(0.1767)
	0.6-0.7	664	894	612	424	328	256	3,178
		(0.1618)	(0.1829)	(0.1938)	(0.2138)	(0.2033)	(0.1998)	(0.1882)
	>0.7	1,302	2,636	1,890	1,363	926	680	8,797
		(0.1979)	(0.2087)	(0.2199)	(0.2341)	(0.2263)	(0.2251)	(0.2166)
	Total	4,673	7,197	5,354	3,837	2,514	1,880	25,455
		(0.1607)	(0.1796)	(0.1876)	(0.2020)	(0.1967)	(0.1934)	(0.1839)

Panel C. WED, 2016-2019

		Maturity						
		<30	30-60	60-90	90-120	120-150	>150	Total
Delta	<0.3	1,124	1,223	340	318	270	271	3,547
		(0.1022)	(0.1048)	(0.1069)	(0.1079)	(0.1094)	(0.1114)	(0.1053)
	0.3-0.4	780	530	118	115	115	121	1,779
		(0.0995)	(0.1089)	(0.1158)	(0.1219)	(0.1244)	(0.1254)	(0.1082)
	0.4-0.5	962	829	275	206	167	132	2,571
		(0.1068)	(0.1189)	(0.1316)	(0.1317)	(0.1347)	(0.1371)	(0.1187)
	0.5-0.6	889	1,150	532	395	261	250	3,477
		(0.1145)	(0.1246)	(0.1380)	(0.1384)	(0.1476)	(0.1493)	(0.1292)
	0.6-0.7	749	717	275	224	232	184	2,381
		(0.1283)	(0.1413)	(0.1540)	(0.1572)	(0.1585)	(0.1631)	(0.1435)
	>0.7	2,444	3,663	1,052	1,110	1,097	779	10,145
		(0.1592)	(0.1724)	(0.1966)	(0.2033)	(0.2068)	(0.2082)	(0.1816)
	Total	6,948	8,112	2,592	2,368	2,143	1,737	23,900
		(0.1270)	(0.1431)	(0.1577)	(0.1651)	(0.1720)	(0.1687)	(0.1466)

Notes: For each category, the number of options is provided, and the average of implied volatilities is in parentheses.

We optimize on the joint likelihood function, which consists of two parts, the likelihood of returns and realized variances, $\ln \mathcal{L}_P$ in equation (24), and an option-based likelihood component, $\ln \mathcal{L}_o$.

6.3 In-sample Estimation Results

Table 5 reports the in-sample estimation results on returns, realized measures, and option data. Similar to Table 2, $\log h_1$ represents the initial value of variances and are estimated as free parameters. We find that $\log h_{CO,1}$ is lower than $\log h_{OC,1}$, indicating that the level of overnight variance is averagely lower than the intraday counterpart, which is in accordance with French and Roll (1986).

The estimated coefficients of λ_{CO} and λ_{OC} in the BRG model, measuring the price of overnight and intraday volatility risk, are both significantly positive. Importantly, λ_{CO} is larger than λ_{OC}, which is consistent with the intuition that investors ask for a higher compensation for exposure to overnight (non-trading) volatility risk. The same pattern has been illustrated in, for example, Bogousslavsky (2021), who emphasizes that the margin requirements are higher during overnight period, and lending fees are typically charged only on positions held overnight. In addition, λ_{OC} is again close to λ reported in the RG model, maintaining the findings in Table 2.

Moreover, Table 5 shows similar overnight and intraday dynamic patterns with those in Table 2. Furthermore, the significantly positive γ_{CO} verifies the importance of the high-frequency overnight information in modeling overnight variance dynamics. Our findings highlight the importance of distinct variation patterns in overnight and intraday volatility, timely arrived overnight information, and volatility risk compensation for non-trading hours.

6.4 Evaluating Pricing Performance

6.4.1 In-sample Pricing Performance

In terms of option pricing accuracy, we compare the performance of different models using the implied volatility root-mean-squared error (IVRMSE). We refer to Renault (1997) for a discussion on the benefits of using the IVRMSE for evaluating option pricing performance. To compute IVRMSE, we first obtain the model-based implied volatility, IV_n^{MOD}, by using the

Table 5 In-sample Joint Estimation: 2003-2015

Estimations on Historical Returns, Realized Measures, and Option Data

Parameters	BRG		RG	GARV	HNG	EG
	CO	OC				
λ	0.1131	0.0257	0.0322	2.0000	1.6220	-0.0013
	(0.0264)	(0.0022)	(0.0389)	(0.0117)	(2.19E-05)	(9.12E-05)
β	**0.6608**	**0.7095**	0.9427	0.9002	0.0783	0.9856
	(0.0417)	(0.0463)	(0.0163)	(0.1435)	(3.85E-05)	(0.0007)
τ_1	-0.0001	0.0009	-0.0883	2.02E-06	5.22E-06	0.1381
	(-0.0009)	(0.0233)	(0.0239)	(4.17E-07)	(1.94E-10)	(0.0023)
τ_2	0.0284	0.0372	0.0487	400.7900	398.5555	-0.5130
	(0.0089)	(0.0111)	(0.0134)	(0.7565)	(0.0099)	(0.0010)
γ	**0.3741**	**0.2803**	0.2805			
	(0.0438)	(0.0368)	(0.0406)			
ξ	0.3829	-2.0231	-2.0089			
	(0.0431)	0.2452	(0.6239)			
ϕ	0.7869	0.8411	0.8475	2.06E-06		
	(0.0587)	(0.0724)	(0.0669)	(3.02E-08)		
δ_1	-0.0185	-0.0252	-0.0007	4.26E-06		
	(0.0265)	(0.0349)	(0.0037)	(6.81E-07)		
δ_2	0.0225	0.0044	8.43E-05	680.7891		
	(0.0109)	(0.0063)	(0.0147)	(0.2586)		
σ	0.6089	0.6987	0.6979	2.77E-06		
	(0.0176)	(0.0202)	(0.0203)	(3.58E-07)		
χ	-0.0108	1.0052	-0.2387	-0.0391		
	(1.89E-04)	(1.64E-05)	(0.0022)	(0.0093)		
κ				0.2003		
				(0.1574)		
$\log h_1$	-13.2891	-9.3743	-9.2997	-9.0172	-9.6195	-9.0366
	(0.7554)	(0.5328)	(0.5487)	(0.0312)	(0.0189)	(0.0001)
$\pi^{\mathbb{P}}$	0.9552	0.9453	0.9427	0.9652	0.9075	0.9856
$\pi^{\mathbb{Q}}$	0.9552	0.9453	0.9427	0.9658	0.9142	0.9856
$\mathcal{L}$	96374		30715	44892	39489	31153

Notes: Table 5 reports estimation results on returns, realized measures, and option data. The sample period is from July 2, 2003 to December 30, 2015. We report the estimated coefficients, and their corresponding standard errors are presented in parentheses. Note that for the GARV and the HNG model, the parameters ω can be inferred by $\mathbb{E}[\bar{h}]$. $\pi^{\mathbb{P}}$ and $\pi^{\mathbb{Q}}$ stand for the volatility persistence under $\mathbb{P}$ and $\mathbb{Q}$ measure, respectively.

model-based option prices and inverting the BS formula

$$IV_n^{MOD} = BS^{-1}(P_n^{MOD}).$$

We compare the model-based implied volatility to the market implied volatility, IV_n^{MKT}, obtained from the option dataset, and compute the implied volatility error

$$e_n = IV_n^{MKT} - IV_n^{MOD},$$

with N denoting the total number of options in the sample. The IVRMSE is now computed as

$$IVRMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N} e_n^2}.$$

Table 6 reports the in-sample evaluation results. Overall, the IVRMSE of the BRG model shows a 6.65% improvement over the RG model, a nested model of the BRG which focuses on total daily variations. Compared to the other nested specification, the EG model, the Bisected Realized GARCH model achieves an impressive 19.02% reduction of the IVRMSE.

In addition to evaluating the overall option pricing accuracy, we also compare performance in pricing options with different moneyness, maturities, and VIX levels. Specifically, Panel A of Table 6 reports the IVRMSE of the five models sorted by moneyness. The BRG model performs the best in five of the six moneyness categories considered, thus, the overall improvement by the BRG model is not driven by any particular range of the moneyness. Notice that the BRG model performs the best for options with the highest Delta (Delta $\geqslant 0.7$), and these options also have the highest implied volatility (see Table 3).

Panel B reports the IVRMSE sorted by calendar days to maturity. Again, the BRG model performs the best in five of the six maturity groups. Therefore, the overall improvement the BRG model is not driven by any particular subset of the maturities.

Panel C reports the IVRMSE sorted by the VIX levels. We observe that the BRG model, the RG model, and the EG model show better performance in highly volatile times (when VIX $\geqslant 35$). This is attributed to the log-linear specification of the volatility dynamics which is flexible to capture the sharp change in volatility during highly volatile times. The above finding is in line with previous empirical results, see Huang et al. (2017) for example. Moreover, the BRG model shows a 6.83% improvement in IVRMSE, compared to the RG model when

Table 6 In-sample Evaluations: WED, 2003-2015

IVRMSE by Moneyness, Maturity, and VIX

Model	Total	Delta<0.3	0.3⩽Delta<0.4	0.4⩽Delta<0.5	0.5⩽Delta<0.6	0.6⩽Delta<0.7	Delta⩾0.7
Panel A. Partitioned by Moneyness							
BRG	**3.8150**	**3.7822**	**3.7692**	**3.8565**	**4.0277**	4.3438	**3.5887**
RG	4.0866	3.9438	3.9625	4.0835	4.3813	4.4897	3.8408
GARV	4.3023	4.6850	4.3533	4.1867	4.4256	4.3580	4.1595
HNG	4.3102	4.3811	4.1556	3.9003	4.1705	**4.2405**	4.5100
EG	4.7110	4.3865	4.2429	4.2969	4.5669	4.7704	4.9923
Model	Total	T<30	30⩽T<60	60⩽T<90	90⩽T<120	120⩽T<150	T⩾150
Panel B. Partitioned by Maturity							
BRG	**3.8150**	**3.7473**	**3.7854**	4.0211	**4.0917**	**3.9286**	**3.8658**
RG	4.0866	3.8833	4.0028	4.1230	4.3585	4.1877	3.9355
GARV	4.3023	4.8397	4.2055	**3.9312**	4.3710	4.3850	4.2228
HNG	4.3102	4.4602	4.0724	4.1378	4.7057	4.4395	4.2206
EG	4.7110	4.4521	4.6634	4.6071	5.0619	4.8678	4.6705
Model	Total	VIX<15	15⩽VIX<20	20⩽VIX<25	25⩽VIX<30	30⩽VIX<35	VIX⩾35
Panel C. Partitioned by VIX Level							
BRG	**3.8150**	3.4555	2.1056	3.9027	**6.4479**	8.9339	**6.5491**
RG	4.0866	3.7785	**2.0674**	4.1692	6.8892	9.6294	7.0289
GARV	4.3023	2.7932	3.1015	**3.1015**	6.5812	8.8928	8.9272
HNG	4.3102	**2.1935**	3.1214	5.3943	6.9742	**8.7806**	8.7714
EG	4.7110	4.1346	3.0203	5.4055	7.6490	9.6607	7.1057

Notes: Table 6 reports the IVRMSEs to compare the in-sample option pricing performance across models. The IVRMSEs are sorted by moneyness, maturity, and the CBOE VIX level, as in Table 3. Panel A reports the IVRMSEs for contracts sorted by the Black–Scholes delta (see equation (44)). Panel B reports the IVRMSEs for contracts sorted by days to maturity. Panel C reports the IVRMSEs for contracts sorted by the VIX level on the quote day. For each category, the smallest IVRMSE is marked in bold. The sample covers options on Wednesdays from July 2, 2003 to December 30, 2015. The IVRMSEs are expressed in percentages.

VIX $\geqslant$ 35. Thus, the Bisected Realized GARCH model achieves an even better performance during highly volatile times.

6.4.2 Out-of-sample Pricing Performance

Compared with the in-sample evaluation, studying the out-of-sample option pricing performance is a more convincing way for measuring the additional improvement by the Bisected Realized GARCH model. We conduct the out-of-sample evaluations in two different windows. The first is Thursdays from July 3, 2003 to December 31, 2015, and the second is Wednesdays from January 6, 2016 to December 18, 2019. The corresponding results are reported in Table 7 and Table 8, respectively.

Results reported in Table 7 are similar to those in Table 6, but the BRG model achieves a larger improvement of 8.24% in out-of-sample test, compared to the RG model. Moreover, the BRG model has a better performance in all of the three Panels in Table 7, relative to Table 6. To be more specific, the BRG model performs the best in every maturity group, five of the six moneyness categories, and also five of the six VIX levels considered. Moreover, the BRG model again outperforms the other models in highly volatile times.

Table 8 report results of the second out-of-sample examination, using options during a pure out-of-sample period. Overall, the BRG model gains a 7.24% reduction of IVRMSE, relative to the RG model. Across different categories, the BRG model performs the best in all of the categories sorted by moneyness. For options on Wednesdays from 2016 to 2019, the BRG model has a good performance for options with shorter maturities and higher VIX levels. As expected, all of the five models have the most difficulty in pricing options in highly volatile times, and the BRG model achieves at least 27.90% improvement in IVRMSE compared with the other models. We conclude here that the Bisected Realized GARCH model has an even better performance in highly volatile times.

Table 7 Out-of-sample Evaluations: THU, 2003-2015

IVRMSE by Moneyness, Maturity and VIX

Model	Total	Delta<0.3	0.3⩽Delta<0.4	0.4⩽Delta<0.5	0.5⩽Delta<0.6	0.6⩽Delta<0.7	Delta⩾0.7
Panel A. Partitioned by Moneyness							
BRG	**3.7832**	**3.6915**	3.5901	**3.8029**	**3.8423**	**4.1987**	**3.3186**
RG	4.1228	3.9812	**3.7964**	4.2626	4.3209	4.4217	3.9566
GARV	4.5781	4.9208	4.6485	4.4977	4.5544	4.5880	4.5254
HNG	4.7769	4.7003	4.6694	4.4382	4.6271	4.7993	4.9512
EG	4.6555	4.3391	4.0125	4.3850	4.4549	4.6776	4.9443
Model	Total	T<30	30⩽T<60	60⩽T<90	90⩽T<120	120⩽T<150	T⩾150
Panel B. Partitioned by Maturity							
BRG	**3.7832**	**3.5493**	**3.7331**	**4.1275**	**4.3343**	**3.8919**	**3.8532**
RG	4.1228	3.8859	4.0186	4.1941	4.4535	4.1076	4.0194
GARV	4.5781	5.4091	4.5484	4.2033	4.6311	4.4414	4.2955
HNG	4.7769	5.4095	4.8114	4.5598	4.7773	4.6273	4.3708
EG	4.6555	4.4066	4.6314	4.6481	4.9412	4.6553	4.5819
Model	Total	VIX<15	15⩽VIX<20	20⩽VIX<25	25⩽VIX<30	30⩽VIX<35	VIX⩾35
Panel C. Partitioned by VIX Level							
BRG	**3.7832**	3.4051	**1.9567**	**4.0369**	**6.5613**	**8.6470**	**6.2541**
RG	4.1228	3.8429	2.0341	4.2220	7.0676	9.3297	6.7042
GARV	4.5781	3.2113	3.1096	5.4269	7.4232	9.3140	9.0120
HNG	4.7769	**2.8204**	3.7474	5.9045	7.5679	9.2148	8.2828
EG	4.6555	4.1651	3.0692	5.3325	7.7883	8.6598	7.1319

Notes: Table 7 reports the IVRMSEs to compare the out-of-sample option pricing performance across models. The IVRMSEs are sorted by moneyness, maturity, and the CBOE VIX level, as in Table 3. Panel A reports the IVRMSEs for contracts sorted by the Black–Scholes delta (see equation (44)). Panel B reports the IVRMSEs for contracts sorted by days to maturity. Panel C reports the IVRMSEs for contracts sorted by the VIX level on the quote day. For each category, the smallest IVRMSE is marked in bold. The sample covers options on Thursdays from July 3, 2003 to December 31, 2015. The IVRMSEs are expressed in percentages.

Table 8 Out-of-sample Evaluations: WED, 2016-2019

IVRMSEs for Moneyness, Maturity and VIX

Model	Total	Delta<0.3	0.3≤Delta<0.4	0.4≤Delta<0.5	0.5≤Delta<0.6	0.6≤Delta<0.7	Delta≥0.7
Panel A. Partitioned by Moneyness							
BRG	**3.3933**	**3.1476**	**2.9811**	**2.9286**	**3.1265**	**3.2607**	**3.6821**
RG	3.6580	3.3760	3.1021	3.4360	3.4490	3.6123	3.7824
GARV	3.6635	3.8512	3.3770	3.1350	3.4677	3.7926	3.7829
HNG	3.8422	4.2091	3.9010	3.8903	3.6052	3.8801	4.5128
EG	4.2821	3.2502	3.5068	3.8192	3.8921	3.9134	4.5775

Model	Total	T<30	30≤T<60	60≤T<90	90≤T<120	120≤T<150	T≥150
Panel B. Partitioned by Maturity							
BRG	**3.3933**	**3.2046**	**3.4073**	3.5079	3.4705	3.3694	2.8546
RG	3.6580	3.5399	3.6606	3.8498	3.7680	3.6522	3.4858
GARV	3.6635	5.0386	3.9901	**3.2692**	3.0933	**2.7867**	3.3750
HNG	3.8422	4.4786	3.7830	3.4280	**2.9443**	3.7472	**2.6039**
EG	4.2821	4.3920	4.4146	4.2561	4.3195	4.0414	3.8281

Model	Total	VIX<15	15≤VIX<20	20≤VIX<25	25≤VIX<30	30≤VIX<35
Panel C. Partitioned by VIX Level						
BRG	**3.3933**	3.8790	2.2660	**2.4584**	**2.9792**	**6.0426**
RG	3.6580	4.1504	**2.2505**	2.7183	4.9659	8.3814
GARV	3.6635	3.8586	2.3584	4.1098	7.6277	7.6262
HNG	3.8422	**2.2244**	3.6329	6.7507	9.5375	9.9501
EG	4.2821	4.2790	3.4495	4.4266	6.0297	8.9491

Notes: Table 8 reports the IVRMSEs to compare the out-of-sample option pricing performance across models. The IVRMSEs are sorted by moneyness, maturity, and the CBOE VIX level, as in Table 3. Panel A reports the IVRMSEs for contracts sorted by the Black–Scholes delta (see equation (44)). Panel B reports the IVRMSEs for contracts sorted by days to maturity. Panel C reports the IVRMSEs for contracts sorted by the VIX level on the quote day. For each category, the smallest IVRMSE is marked in bold. The sample covers options on Wednesdays from January 6, 2016 to December 18, 2019. The IVRMSEs are expressed in percentages.

7 Conclusion

This study develops a new and flexible option-pricing model that can accommodate distinct overnight variance dynamics and intraday variance dynamics into its underlying asset-pricing process. An important feature of the model is that the dynamics of overnight and intraday 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 multivariate Edgeworth-Sargan density enable us to derive analytical approximations for option pricing 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 overnight and intraday variances, and returns. In particular, the overnight component accounts for an out-of-sample gain of 7.24% in pricing accuracy, relative to simply modeling total daily returns.

References

P. Akey, V. Grégoire, and C. Martineau. Price revelation from insider trading: Evidence from hacked earnings news. *Journal of Financial Economics*, 143(3):1162–1184, 2022.

Y. Amihud and H. Mendelson. Liquidity, maturity, and the yields on us treasury securities. *The Journal of Finance*, 46(4):1411–1425, 1991.

T. G. Andersen and T. Bollerslev. Deutsche mark–dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies. *the Journal of Finance*, 53 (1):219–265, 1998.

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.

T. G. Andersen, T. Bollerslev, F. X. Diebold, and P. Labys. Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625, 2003.

T. G. Andersen, T. Bollerslev, F. X. Diebold, and C. Vega. Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of international Economics*, 73 (2):251–277, 2007.

I. Archakov and P. R. Hansen. A canonical representation of block matrices with applications to covariance and correlation matrices. *arXiv preprint arXiv:2012.02698*, 2020.

Y. Atilgan. Volatility spreads and earnings announcement returns. *Journal of Banking & Finance*, 38:205–215, 2014.

F. Black. Studies of stock price volatility changes. *proceedings of the 1976 meetings of the business& economics statistics section, american statistical association*, 81:177–181, 1976.

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

M.-J. Boes, F. C. Drost, and B. J. Werker. The impact of overnight periods on option pricing. *Journal of Financial and Quantitative Analysis*, 42(2):517–533, 2007.

V. Bogousslavsky. The cross-section of intraday and overnight returns. *Journal of Financial Economics*, 141(1):172–194, 2021.

J. Boudoukh, R. Feldman, S. Kogan, and M. Richardson. Information, trading, and volatility: Evidence from firm-specific news. *The Review of Financial Studies*, 32(3):992–1033, 2019.

K. C. Chan, W.-M. Fong, B.-C. Kho, and R. Stulz. Information, trading and stock returns: Lessons from dually-listed securities. *Journal of Banking & Finance*, 20(7):1161–1187, 1996.

K. F. Chan and T. Marsh. Asset pricing on earnings announcement days. *Journal of Financial Economics*, 144(3):1022–1042, 2022.

A. A. Christie. The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of financial Economics*, 10(4):407–432, 1982.

P. Christoffersen, R. Elkamhi, B. Feunou, and K. Jacobs. Option valuation with conditional heteroskedasticity and nonnormality. *The Review of Financial Studies*, 23(5):2139–2183, 2010.

P. Christoffersen, K. Jacobs, and C. Ornthanalai. Garch option valuation: theory and evidence. *The Journal of Derivatives*, 21(2):8–41, 2013.

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.

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

A. Craig, A. Dravid, and M. Richardson. Market efficiency around the clock some supporting evidence using foreign-based derivatives. *Journal of Financial Economics*, 39(2-3):161–180, 1995.

W. Del Corral, D. Colwell, D. Michayluk, and L. Woo. News releases when the markets are closed. *unpublished, University of Technology, Sidney*, 2003.

G. Dhaene and J. Wu. Incorporating overnight and intraday returns into multivariate garch volatility models. *Journal of Econometrics*, 217(2):471–495, 2020.

J. Duan, G. Gauthier, J. Simonato, and C. Sasseville. Approximating the gjr-garch and egarch option pricing models analytically. *Journal of Computational Finance*, 9(3):41, 2006.

J.-C. Duan. The garch option pricing model. *Mathematical finance*, 5(1):13–32, 1995.

J.-C. Duan, G. Gauthier, and J.-G. Simonato. An analytical approximation for the garch option pricing model. 1997.

R. F. Engle and V. K. Ng. Measuring and testing the impact of news on volatility. *The journal of finance*, 48(5):1749–1778, 1993.

E. F. Fama. The behavior of stock-market prices. *The journal of Business*, 38(1):34–105, 1965.

K. R. French. Stock returns and the weekend effect. *Journal of financial economics*, 8(1): 55–69, 1980.

K. R. French and R. Roll. Stock return variances: The arrival of information and the reaction of traders. *Journal of financial economics*, 17(1):5–26, 1986.

M. R. Gibbons and P. Hess. Day of the week effects and asset returns. *Journal of business*, pages 579–596, 1981.

P. R. Hansen and Z. Huang. Exponential garch modeling with realized measures of volatility. *Journal of Business & Economic Statistics*, 34(2):269–287, 2016.

P. R. Hansen and A. Lunde. A realized variance for the whole day based on intermittent high-frequency data. *Journal of Financial Econometrics*, 3(4):525–554, 2005.

P. R. Hansen, Z. Huang, and H. H. Shek. Realized garch: a joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, 27(6):877–906, 2012.

P. R. Hansen, A. Lunde, and V. Voev. Realized beta garch: A multivariate garch model with realized measures of volatility. *Journal of Applied Econometrics*, 29(5):774–799, 2014. doi: https://doi.org/10.1002/jae.2389. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.2389`.

P. R. Hansen, Z. Huang, and F. Liang. The weekly rhythm of volatility. *working paper*, 2022.

J. Hasbrouck. Intraday price formation in us equity index markets. *The Journal of Finance*, 58(6):2375–2400, 2003.

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

H. Hong and J. Wang. Trading and returns under periodic market closures. *The Journal of Finance*, 55(1):297–354, 2000.

G. X. Hu, J. Pan, J. Wang, and H. Zhu. Premium for heightened uncertainty: Explaining pre-announcement market returns. *Journal of Financial Economics*, 2021.

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.

R. Jarrow and A. Rudd. Approximate option valuation for arbitrary stochastic processes. *Journal of financial Economics*, 10(3):347–369, 1982.

C. X. Jiang, T. Likitapiwat, and T. H. McInish. Information content of earnings announcements: Evidence from after-hours trading. *Journal of Financial and Quantitative Analysis*, 47(6):1303–1330, 2012.

C. S. Jones and J. Shemesh. Option mispricing around nontrading periods. *The Journal of Finance*, 73(2):861–900, 2018.

J. Kanniainen, B. Lin, and H. Yang. Estimating and using garch models with vix data for option valuation. *Journal of Banking & Finance*, 43:200–211, 2014.

D. B. Keim and R. F. Stambaugh. A further investigation of the weekend effect in stock returns. *The journal of finance*, 39(3):819–835, 1984.

O. Linton and J. Wu. A coupled component dcs-egarch model for intraday and overnight volatility. *Journal of Econometrics*, 217(1):176–201, 2020.

F. Moshirian, H. G. L. Nguyen, and P. K. Pham. Overnight public information, order placement, and price discovery during the pre-opening period. *Journal of Banking & Finance*, 36(10):2837–2851, 2012.

D. Muravyev and X. C. Ni. Why do option returns change sign from day to night? *Journal of Financial Economics*, 136(1):219–238, 2020.

G. S. Oldfield and R. J. Rogalski. A theory of common stock returns over trading and non-trading periods. *The Journal of Finance*, 35(3):729–751, 1980.

J. Perote. The multivariate edgeworth-sargan density. *Spanish Economic Review*, 6(1):77–96, 2004.

E. Renault. Econometric models of option pricing errors. *Econometric Society Monographs*, 28:223–278, 1997.

N. Taylor. A note on the importance of overnight information in risk management models. *Journal of Banking & Finance*, 31(1):161–180, 2007.

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.

I. Tsiakas. Overnight information and stochastic volatility: A study of european and us stock exchanges. *Journal of Banking & Finance*, 32(2):251–268, 2008.

T. Wang, S. Cheng, F. Yin, and M. Yu. Overnight volatility, realized volatility, and option pricing. *Journal of Futures Markets*, 42(7):1264–1283, 2022.

8 Appendix

8.1 Appendix A. Risk-Neutralization

Given the independence of $\mathbf{z_t}$ and $\mathbf{u_t}$, and the correlation of components in $\mathbf{z_t}$ and $\mathbf{u_t}$ follows

$$\mathbf{\Sigma}_z = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad \mathbf{\Sigma}_u = \begin{pmatrix} 1 & \rho_{v_{CO},v_{OC}} & \rho_{v_{CO},\widetilde{v}} \\ \rho_{v_{CO},v_{OC}} & 1 & \rho_{v_{OC},\widetilde{v}} \\ \rho_{v_{CO},\widetilde{v}} & \rho_{v_{OC},\widetilde{v}} & 1 \end{pmatrix},$$

the pricing kernel corresponding to five random shocks

$$Z_{t+1} = \frac{\exp(\nu_{CO,t} z_{CO,t+1} + \nu_{OC,t} z_{OC,t+1} + \theta_{CO,t} v_{CO,t+1} + \theta_{OC,t} v_{OC,t+1} + \widetilde{\theta}_t \widetilde{v}_{t+1})}{\mathbb{E}_t[\exp(\nu_{CO,t} z_{CO,t+1} + \nu_{OC,t} z_{OC,t+1} + \theta_{CO,t} v_{CO,t+1} + \theta_{OC,t} v_{OC,t+1} + \widetilde{\theta}_t \widetilde{v}_{t+1})]}$$

can be reduced into

$$Z_{t+1} = \exp\begin{pmatrix} \nu_{CO,t} z_{CO,t+1} + \nu_{OC,t} z_{OC,t+1} - \frac{\nu_{CO,t}^2}{2} - \frac{\nu_{OC,t}^2}{2} - \nu_{CO,t}\nu_{OC,t}\rho_{t+1} \\ \\ +\theta_{CO,t} v_{CO,t+1} + \theta_{OC,t} v_{OC,t+1} + \widetilde{\theta}_t \widetilde{v}_{t+1} \\ \\ -\frac{\theta_{CO,t}^2}{2} - \frac{\theta_{OC,t}^2}{2} - \frac{\widetilde{\theta}_t^2}{2} \\ \\ -\theta_{CO,t}\theta_{OC,t}\rho_{v_{CO},v_{OC}} - \theta_{CO,t}\widetilde{\theta}_t \rho_{v_{CO},\widetilde{v}} - \theta_{OC,t}\widetilde{\theta}_t \rho_{v_{OC},\widetilde{v}} \end{pmatrix}. \tag{A-1}$$

We impose the no-arbitrage condition, namely that the return under $\mathbb{Q}$-measure equals to risk-free return r_j in that specific period,

$$\mathbb{E}_t^Q[\exp(r_{j,t+1})] = \exp(r_j), \quad j \in \{CO, OC\}. \tag{A-2}$$

Note that Z_{t+1} builds the relationship between $\mathbb{P}$ measure and $\mathbb{Q}$ measure,

$$\mathbb{E}_t[Z_{t+1}\exp(r_{j,t+1})] = \mathbb{E}_t^Q[\exp(r_{j,t+1})], \quad j \in \{CO, OC\}. \tag{A-3}$$

we have

$$\mathbb{E}_t[Z_{t+1}\exp(r_{j,t+1})] = \exp(r_j), \quad j \in \{CO, OC\}. \tag{A-4}$$

One can derive by plugging in the explicit form of pricing kernel Z_t and the return dynamic under $\mathbb{P}$ measure into the first term of the equation above that:

$$\begin{aligned}
&\mathbb{E}_t\left[\exp\left(\begin{array}{c}
\nu_{CO,t}z_{CO,t+1}+\nu_{OC,t}z_{OC,t+1}-\frac{\nu_{CO,t}^2}{2}-\frac{\nu_{OC,t}^2}{2}-\nu_{CO,t}\nu_{OC,t}\rho_{t+1}\\
+\theta_{CO,t}v_{CO,t+1}+\theta_{OC,t}v_{OC,t+1}+\widetilde{\theta}_t\widetilde{v}_{t+1}\\
-\frac{\theta_{CO,t}^2}{2}-\frac{\theta_{OC,t}^2}{2}-\frac{\widetilde{\theta}_t^2}{2}\\
-\theta_{CO,t}\theta_{OC,t}\rho_{v_{CO},v_{OC}}-\theta_{CO,t}\widetilde{\theta}_t\rho_{v_{CO},\widetilde{v}}-\theta_{OC,t}\widetilde{\theta}_t\rho_{v_{OC},\widetilde{v}}\\
+r_j+\lambda_j\sqrt{d_{j,t}h_{j,t}}-\frac{1}{2}d_{j,t}h_{j,t}+\sqrt{d_{j,t}h_{j,t}}z_{j,t+1}
\end{array}\right)\right]\\
&=\exp\left(r_j+\lambda_j\sqrt{d_{j,t}h_{j,t}}+(\nu_{j,t}+\rho_{t+1}\nu_{-j,t})\sqrt{d_{j,t}h_{j,t}}\right)\\
&=\exp(r_j),
\end{aligned}\tag{A-5}$$

which implies

$$\nu_{j,t}+\rho_{t+1}\nu_{-j,t}=-\lambda_j,\tag{A-6}$$

where $j\in\{CO,OC\}$ and $-j=\{CO,OC\}\backslash j$.

Then we consider the risk-neutral moment generating function (MGF) and the sufficient condition to derive the model under $\mathbb{Q}$ measure. The moment generating function of the five random shocks can be written as $\mathbb{E}_t^{\mathbb{Q}}[\exp(\mathbf{s}'(\mathbf{z'_{t+1}},\mathbf{u'_{t+1}})')]$:

$$\begin{aligned}
&\mathbb{E}_t^Q[\exp(s_1z_{CO,t+1}+s_2z_{OC,t+1}+s_3v_{CO,t+1}+s_4v_{OC,t+1}+s_5\widetilde{v}_{t+1})]\\
&=\mathbb{E}_t[Z_{t+1}\exp(s_1z_{CO,t+1}+s_2z_{OC,t+1}+s_3v_{CO,t+1}+s_4v_{OC,t+1}+s_5\widetilde{v}_{t+1})]\\
&=\exp\left(\begin{array}{c}
s_1(\nu_{CO,t}+\rho_{t+1}\nu_{OC,t})+s_2(\nu_{OC,t}+\rho_{t+1}\nu_{CO,t})+\frac{s_1^2}{2}+\frac{s_2^2}{2}+s_1s_2\rho_{t+1}\\
+s_3(\theta_{CO,t}+\rho_{v_{CO},v_{OC}}\theta_{OC,t}+\rho_{v_{OC},\widetilde{v}}\widetilde{\theta}_t)\\
+s_4(\rho_{v_{CO},v_{OC}}\theta_{CO,t}+\theta_{OC,t}+\rho_{v_{OC},\widetilde{v}}\widetilde{\theta}_t)\\
+s_5(\rho_{v_{CO},\widetilde{v}}\theta_{CO,t}+\rho_{v_{OC},\widetilde{v}}\theta_{OC,t}+\widetilde{\theta}_t)\\
+\frac{s_3^2}{2}+\frac{s_4^2}{2}+\frac{s_5^2}{2}+s_3s_4\rho_{v_{CO},v_{OC}}+s_3s_5\rho_{v_{CO},\widetilde{v}}+s_4s_5\rho_{v_{OC},\widetilde{v}}
\end{array}\right)
\end{aligned}\tag{A-7}$$

We set

$$\begin{aligned}
\theta_{CO,t} + \rho_{v_{CO},v_{OC}}\theta_{OC,t} + \rho_{v_{CO},\widetilde{v}}\widetilde{\theta}_t &= \chi_1, \\
\rho_{v_{CO},v_{OC}}\theta_{CO,t} + \theta_{OC,t} + \rho_{v_{OC},\widetilde{v}}\widetilde{\theta}_t &= \chi_2, \\
\rho_{v_{CO},\widetilde{v}}\theta_{CO,t} + \rho_{v_{OC},\widetilde{v}}\theta_{OC,t} + \widetilde{\theta}_t &= \chi_3,
\end{aligned} \tag{A-8}$$

thus implying that we have the following mapping relationship:

$$\begin{aligned}
z^*_{CO,t+1} &= z_{CO,t+1} + \lambda_{CO}, \\
z^*_{OC,t+1} &= z_{OC,t+1} + \lambda_{OC}, \\
v^*_{CO,t+1} &= v_{CO,t+1} - \chi_1, \\
v^*_{OC,t+1} &= v_{OC,t+1} - \chi_2, \\
\widetilde{v}^*_{t+1} &= \widetilde{v}_{t+1} - \chi_3.
\end{aligned} \tag{A-9}$$

and we can obtain the dynamic model under $\mathbb{Q}$-measure.

8.2 Appendix B. Expressions and Notations in the Option Pricing Formula

Here are the notations applied in the option pricing formula. To begin with, the domain of integration involving

$$\begin{aligned}
D' &= \{(z_{CO}, z_{OC}) | \sigma_{CO} z_{CO} + \sigma_{OC} z_{OC} \leqslant \log(S_0/K) + \overline{\mu}_{CO} + \overline{\mu}_{OC} + (\sigma^2_{CO} + \sigma^2_{OC}) + 2\rho\sigma_{CO}\sigma_{OC}\}, \\
D'' &= \{(z_{CO}, z_{OC}) | \sigma_{CO} z_{CO} + \sigma_{OC} z_{OC} \leqslant \log(S_0/K) + \overline{\mu}_{CO} + \overline{\mu}_{OC} + (\sigma^2_{CO} + \sigma^2_{OC})\},
\end{aligned} \tag{B-1}$$

To ease notation, we denote

$$\begin{aligned}
\sigma &= \sqrt{\sigma^2_{CO} + \sigma^2_{OC}}, \\
\delta &= \frac{1}{\sigma}(-rT + \overline{\mu}_{CO} + \overline{\mu}_{OC} + \frac{\sigma^2_{CO} + \sigma^2_{OC}}{2}), \\
\Delta &= \frac{1}{\sigma}(-rT + \overline{\mu}_{CO} + \overline{\mu}_{OC} + \frac{\sigma^2_{CO} + \sigma^2_{OC}}{2} + \rho\sigma_{CO}\sigma_{OC}), \\
k &= \frac{1}{\sigma}(\log(S_0/K) + \overline{\mu}_{CO} + \overline{\mu}_{OC}), \\
d &= k + \sigma.
\end{aligned} \tag{B-2}$$

where

$$z_{CO} = \frac{R_{CO,T} - \overline{\mu}_{CO}}{\sigma_{CO}}, \quad z_{OC} = \frac{R_{OC,T} - \overline{\mu}_{OC}}{\sigma_{OC}}$$

are both standardized cumulate returns with cumulants κ_{CO} and κ_{OC}, and

$$\rho = cov_0^{\mathbb{Q}}(z_{CO}, z_{OC}),\ \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},\ \Sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Φ and ϕ denotes cumulative density function (CDF) of bivariate normal distribution, and denotes probability density function (PDF) of normal distribution, respectively. And $\kappa_{j,i} = \mathbb{E}_0^{\mathbb{Q}}[R_j^i]$, which can be expressed as

$$\kappa_{j,3} = \frac{1}{\sigma_j^3}\left[\mathbb{E}_0^{\mathbb{Q}}\left(R_j^3\right) - \overline{\mu}_j^3\right] - 3\frac{\overline{\mu}_j}{\sigma_j}, \quad \kappa_{j,4} = \frac{1}{\sigma_j^4}\left[\mathbb{E}_0^{\mathbb{Q}}\left(R_j^4\right) - \overline{\mu}_j^4\right] - 2\frac{\overline{\mu}_j}{\sigma_j}\left(2\kappa_{j,3} + 3\frac{\overline{\mu}_j}{\sigma_j}\right).$$

8.3 Appendix C. Proof of Proposition 1

Lemma 1. *Let z_1, z_2 are orthogonal standard distributed random variables. That is, $corr(z_1, z_2) = 0$. Let*

$$G_1(k, \sigma_1, n) = \iint\limits_{\sigma_1 z_1 + \sigma_2 z_2 \leqslant \sigma k} (z_1 - \sigma_1)^n \phi(z_1, z_2) dz_1 dz_2,$$
$$G_2(k, \sigma_2, n) = \iint\limits_{\sigma_1 z_1 + \sigma_2 z_2 \leqslant \sigma k} (z_2 - \sigma_2)^n \phi(z_1, z_2) dz_1 dz_2.$$

Then we have the following recurrence relation for any $\sigma_1, \sigma_2 > 0$ and $n \in \mathbb{R}^+$:

$$G_1(k, \sigma_1, n) = (n+1)G_1(k, \sigma_1, n) - \sigma_1 G_1(k, \sigma_1, n+1) - \frac{\sigma_1 \sigma_2^{n+1}}{\sigma^{n+2}}\mathbb{E}\left(z + \frac{\sigma_1(k-\sigma)}{\sigma_2}\right)^{n+1}\phi(k),$$
$$G_2(k, \sigma_2, n) = (n+1)G_2(k, \sigma_2, n) - \sigma_2 G_2(k, \sigma_2, n+1) - \frac{\sigma_2 \sigma_1^{n+1}}{\sigma^{n+2}}\mathbb{E}\left(z + \frac{\sigma_2(k-\sigma)}{\sigma_2}\right)^{n+1}\phi(k),$$

where

$$G_1(k, \sigma_1, 0) = \Phi(k),\ G_1(k, \sigma_1, 1) = -\frac{\sigma_1}{\sigma}\phi(k) - \sigma_1\Phi(k),$$
$$G_2(k, \sigma_2, 0) = \Phi(k),\ G_2(k, \sigma_2, 1) = -\frac{\sigma_2}{\sigma}\phi(k) - \sigma_2\Phi(k).$$

Proof. Integrating z_1 first, we have

$$G_1(k, \sigma_1, n) = \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{\frac{\sigma k - \sigma_2 z_2}{\sigma_1}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z_1^2}{2}} d\frac{1}{n+1}(z_1 - \sigma_1)^{n+1}\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z_2^2}dz_2,$$

Integrating by parts yields

$$G_1(k, \sigma_1, n) = \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(\frac{\sigma k - \sigma_2 z_2}{\sigma_1})^2} \cdot \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z_2^2}\frac{1}{n+1}\left(\frac{\sigma k - \sigma_2 z_2}{\sigma_1} - \sigma_1\right)^{n+1} dz_2$$
$$+ \frac{1}{n+1}G_1(k, \sigma_1, n+2) + \frac{\sigma_1}{n+1}G_1(k, \sigma_1, n+1),$$

where

$$
\begin{aligned}
&\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{\sigma k-\sigma_2 z_2}{\sigma_1})^2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_2^2} \frac{1}{n+1} \left(\frac{\sigma k-\sigma_2 z_2}{\sigma_1} - \sigma_1 \right)^{n+1} dz_2 \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}[(\frac{\sigma}{\sigma_2}t - \frac{\sigma_1 k}{\sigma_2})^2 + k^2]} \frac{1}{\sqrt{2\pi}} (t-\sigma_1)^{n+1} \frac{\sigma_1}{\sigma_2} dt \ (\text{Let } \frac{\sigma k-\sigma_2 z_2}{\sigma_1} = t) \\
&= \phi(k) \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \left(z + \frac{\sigma_1(k-\sigma)}{\sigma_2} \right)^{n+1} \left(\frac{\sigma_2}{\sigma} \right)^{n+1} \cdot \frac{\sigma_1}{\sigma} e^{-\frac{1}{2}z^2} dz \ (\text{Let } \frac{\sigma}{\sigma_2}t - \frac{\sigma_1 k}{\sigma_2} = z) \\
&= \frac{\sigma_1 \sigma_2^{n+1}}{\sigma^{n+2}} \mathbb{E} \left(z + \frac{\sigma_1(k-\sigma)}{\sigma_2} \right)^{n+1} \phi(k).
\end{aligned}
\tag{C-1}
$$

Rearranging the terms can we derive Lemma 1. □

Proof Outline of Proposition 1. The price of European option matures at time T can be given by

$$
e^{-rT}\mathbb{E}_0^Q[\max(S_T - K, 0)] = e^{-rT} \iint_D [S_0 \exp(\mu - \boldsymbol{\sigma}'\mathbf{z}) - K]\, \widetilde{g}(\mathbf{z}) d\mathbf{z},
$$

where $\widetilde{g}(z_{CO}, z_{OC}) = g(-z_{CO}, -z_{OC})$. and

$$
\begin{aligned}
& e^{-rT} \iint_D [S_0 \exp(\mu - \boldsymbol{\sigma}'\mathbf{z}) - K]\, \widetilde{g}(\mathbf{z}) \mathrm{d}\mathbf{z} \\
&= e^{-rT} \iint_D [S_0 \exp(\mu - \boldsymbol{\sigma}'\mathbf{z}) - K] \cdot \frac{1}{2\pi\sqrt{1-\rho^2}} \cdot \exp(-\frac{1}{2}\mathbf{z}'\boldsymbol{\Sigma}^{-1}\mathbf{z}) \mathrm{d}\mathbf{z} && \text{(A)} \\
&\quad - e^{-rT} \iint_D \frac{1}{2\pi} [S_0 \exp(\mu - \boldsymbol{\sigma}'\mathbf{z}) - K]\, e^{-\frac{1}{2}(z_{CO}^2 + z_{OC}^2)} \cdot \frac{\kappa_{CO,3}}{6} (z_{CO}^3 - 3z_{CO}) \mathrm{d}\mathbf{z} && \text{(B)} \\
&\quad + e^{-rT} \iint_D \frac{1}{2\pi} [S_0 \exp(\mu - \boldsymbol{\sigma}'\mathbf{z}) - K]\, e^{-\frac{1}{2}(z_{CO}^2 + z_{OC}^2)} \cdot \frac{\kappa_{CO,4} - 3}{24} (z_{CO}^4 - 6z_{CO}^2 + 3) \mathrm{d}\mathbf{z} && \text{(C)} \\
&\quad + e^{-rT} \iint_D \frac{1}{2\pi} [S_0 \exp(\mu - \boldsymbol{\sigma}'\mathbf{z}) - K]\, e^{-\frac{1}{2}(z_{CO}^2 + z_{OC}^2)} \cdot \frac{\kappa_{CO,3}^2}{72} (z_{CO}^6 - 15z_{CO}^4 + 45z_{CO}^2 - 15) \mathrm{d}\mathbf{z} && \text{(D)} \\
&\quad - e^{-rT} \iint_D \frac{1}{2\pi} [S_0 \exp(\mu - \boldsymbol{\sigma}'\mathbf{z}) - K]\, e^{-\frac{1}{2}(z_{CO}^2 + z_{OC}^2)} \cdot \frac{\kappa_{OC,3}}{6} (z_{OC}^3 - 3z_{OC}) \mathrm{d}\mathbf{z} && \text{(E)} \\
&\quad + e^{-rT} \iint_D \frac{1}{2\pi} [S_0 \exp(\mu - \boldsymbol{\sigma}'\mathbf{z}) - K]\, e^{-\frac{1}{2}(z_{CO}^2 + z_{OC}^2)} \cdot \frac{\kappa_{OC,4} - 3}{24} (z_{OC}^4 - 6z_{OC}^2 + 3) \mathrm{d}\mathbf{z} && \text{(F)} \\
&\quad + e^{-rT} \iint_D \frac{1}{2\pi} [S_0 \exp(\mu - \boldsymbol{\sigma}'\mathbf{z}) - K]\, e^{-\frac{1}{2}(z_{CO}^2 + z_{OC}^2)} \cdot \frac{\kappa_{OC,3}^2}{72} (z_{OC}^6 - 15z_{OC}^4 + 45z_{OC}^2 - 15) \mathrm{d}\mathbf{z} && \text{(G)}
\end{aligned}
$$

Here term (A) involves a bivariate normal distributed probability density function, while term (B) to (G) involve higher order terms in Edgeworth-Sargan density, which are utilized to slightly modify the asymptotic distribution.

For the first term (A), we apply linear transformation to match the normal distribution probability density function:

$$
\begin{aligned}
& e^{-rT} \iint\limits_{D} \left[S_0 \exp\left(\mu - \boldsymbol{\sigma}'\mathbf{z}\right) - K\right] \widetilde{g}(\mathbf{z}) \cdot \frac{1}{2\pi\sqrt{1-\rho^2}} \cdot \exp(-\frac{1}{2}\mathbf{z}'\boldsymbol{\Sigma}^{-1}\mathbf{z}) \mathrm{d}z_{CO} \mathrm{d}z_{OC} \\
= & e^{-rT} \iint\limits_{D} S_0 \exp\left(\overline{\mu}_{CO} + \overline{\mu}_{OC} + \frac{1}{2}\boldsymbol{\upsilon}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\upsilon} - \frac{1}{2}\mathbf{z}^{\dagger\prime}\boldsymbol{\Sigma}^{-1}\mathbf{z}^{\dagger}\right) \frac{1}{2\pi\sqrt{1-\rho^2}} \cdot \mathrm{d}z_{CO} \mathrm{d}z_{OC} \\
& - \iint\limits_{D} K e^{-rT} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(\frac{1}{2}\mathbf{z}'\boldsymbol{\Sigma}^{-1}\mathbf{z}\right) \mathrm{d}z_{CO} \mathrm{d}z_{OC} \\
= & S_0 e^{\Delta\sigma} \Phi(D'; \Sigma) - K e^{-rT} \Phi(D; \Sigma),
\end{aligned}
\tag{C-2}
$$

where

$$
\boldsymbol{\upsilon} = \begin{pmatrix} \sigma_{CO} + \rho\sigma_{OC} \\ \rho\sigma_{CO} + \sigma_{OC} \end{pmatrix}, \quad \mathbf{z}^{\dagger} = \mathbf{z} + \boldsymbol{\upsilon}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},
$$

$$
\Delta = \frac{1}{\sigma}(-rT + \overline{\mu}_{CO} + \overline{\mu}_{OC} + \frac{\sigma_{CO}^2 + \sigma_{OC}^2}{2} + \rho\sigma_{CO}\sigma_{OC}),
$$

$$
D = \{(z_{CO}, z_{OC}) | \sigma_{CO} z_{CO} + \sigma_{OC} z_{OC} \leqslant \log(S_0/K) + \overline{\mu}_{CO} + \overline{\mu}_{OC}\},
$$

$$
D' = \{(z_{CO}, z_{OC}) | \sigma_{CO} z_{CO} + \sigma_{OC} z_{OC} \leqslant \log(S_0/K) + \overline{\mu}_{CO} + \overline{\mu}_{OC} + (\sigma_{CO}^2 + \sigma_{OC}^2) + 2\rho\sigma_{CO}\sigma_{OC}\}.
$$

For term (B) and (E), one can derive by lemma 1 that

$$
\begin{aligned}
& -\frac{1}{6}\left\{ e^{-rT} \iint\limits_{D} \frac{1}{2\pi} \left[S_0 \exp\left(\mu - \boldsymbol{\sigma}'\mathbf{z}\right) - K\right] e^{-\frac{1}{2}(z_{CO}^2 + z_{OC}^2)} \cdot (z_j^3 - 3z_j) \mathrm{d}\mathbf{z} \right\} \\
= & -\frac{1}{6}\left\{ S_0 e^{\delta\sigma} \left(G_j(d, \sigma_j, 3) - 3G_j(d, \sigma_j, 1)\right) - K e^{-rT} \left(G_j(k, 0, 3) - 3G_j(k, 0, 1)\right) \right\} \\
= & -\frac{1}{6}\left\{ -S_0 e^{\delta\sigma} \sigma_j (\sigma_j^2 \Phi(D''; \Sigma_0) + \frac{\sigma_j^2(2\sigma - d)}{\sigma^2} \phi(d)) \right\} \\
=: & A_{j,3}.
\end{aligned}
$$

where d, σ_j are given in Proposition 1, and

$$
D'' = \{(z_{CO}, z_{OC}) | \sigma_{CO} z_{CO} + \sigma_{OC} z_{OC} \leqslant \log(S_0/K) + \overline{\mu}_{CO} + \overline{\mu}_{OC} + (\sigma_{CO}^2 + \sigma_{OC}^2)\}.
$$

Apply the same technique to term (C) to (G) can we obtain

$$
\begin{aligned}
A_{j,4} \coloneqq & \frac{1}{24} \begin{Bmatrix} S_0 e^{\delta\sigma} (G_j(d, \sigma_j, 4) - 6G_j(d, \sigma_j, 2) + 3G_j(d, \sigma_j, 0)) \\ \\ -K e^{-rT} (G_j(k, 0, 4) - 6G_j(k, 0, 2) + 3G_j(k, 0, 0)) \end{Bmatrix} \\
= & \frac{1}{24} S_0 e^{\delta\sigma} \sigma_j (\sigma_j^3 \Phi(D''; \Sigma_0) + \frac{\sigma_j^3}{\sigma^3}(d^2 - 1 - 3\sigma(d - \sigma))\phi(d)),
\end{aligned}
\tag{C-3}
$$

$$A_{j,6} \coloneqq \frac{1}{72} S_0 e^{\delta\sigma} \sigma_j (\sigma_j^5 \phi(D''; \Sigma_0) + \frac{\sigma_j^5}{\sigma^5}(3 - 6d^2 + d^4 + 5\sigma(d - (d-\sigma)(\sigma d - 2) - (d-\sigma)^3))\phi(d)). \tag{C-4}$$

Additionally, based on the Bisected Realized GARCH model, the covariance of standardized cumulative return $R_{CO,T}$ and $R_{OC,T}$ under $\mathbb{Q}$-measure can be written as

$$\begin{aligned}
\rho = & \frac{1}{\sigma_{CO}\sigma_{OC}} \cdot (Tr_{CO} - \overline{\mu}_{CO})(Tr_{OC} - \overline{\mu}_{OC}) - \frac{Tr_{CO} - \overline{\mu}_{CO}}{2\sigma_{CO}\sigma_{OC}} \sum_{i=1}^{T} \mathbb{E}_0^Q \left[d_{OC,i} h_{OC,i} \right] - \frac{Tr_{OC} - \overline{\mu}_{OC}}{2\sigma_{CO}\sigma_{OC}} \sum_{i=1}^{T} \mathbb{E}_0^Q \left[d_{CO,i} h_{CO,i} \right] \\
& + \frac{1}{4\sigma_{CO}\sigma_{OC}} \cdot \sum_{i=1}^{T} \sum_{j=1}^{T-i} \mathbb{E}_0^Q [d_{CO,i} h_{CO,i} d_{OC,i+j} h_{OC,i+j}] + \frac{1}{4\sigma_{CO}\sigma_{OC}} \cdot \sum_{i=1}^{T} \sum_{j=1}^{T-i} \mathbb{E}_0^Q [d_{OC,i} h_{OC,i} d_{CO,i+j} h_{CO,i+j}] \\
& - \frac{1}{2\sigma_{CO}\sigma_{OC}} \cdot \sum_{i=1}^{T} \sum_{j=1}^{T-i} \mathbb{E}_0^Q \left[\sqrt{d_{OC,i} h_{OC,i}} z_{OC,i} d_{CO,i+j} h_{CO,i+j} \right] \\
& - \frac{1}{2\sigma_{CO}\sigma_{OC}} \cdot \sum_{i=1}^{T} \sum_{j=1}^{T-i} \mathbb{E}_0^Q \left[\sqrt{d_{CO,i} h_{CO,i}} z_{CO,i} d_{OC,i+j} h_{OC,i+j} \right] \\
& + \frac{5}{4\sigma_{CO}\sigma_{OC}} \sum_{i=1}^{T} \mathbb{E}_0^Q \left[d_{CO,i} h_{CO,i} d_{OC,i} h_{OC,i} \right].
\end{aligned} \tag{C-5}$$

the detailed expressions are discussed in Appendix D.

8.4 Appendix D. Analytical Terms

Following Duan et al. (1997), the moment of cumulative return based on the Bisected Realized GARCH under $\mathbb{Q}$-measure can be written as follows:

$$\mathbb{E}_0^Q (R_{j,T}^s) = \mathbb{E}_0^Q \left[\left(\sum_{i=1}^{T} (r_j - \frac{d_{j,i} h_{j,i}}{2} + \sqrt{d_{j,i} h_{j,i}} z_{j,i}^*) \right)^s \right], \tag{D-1}$$

so the first four moments are

$$\mathbb{E}_0^Q(R_{j,T}) = \mathbb{E}_0^Q\left[\sum_{i=1}^{T}(r_j - \frac{d_{j,i}h_{j,i}}{2} + \sqrt{d_{j,i}h_{j,i}}z_{j,i}^*)\right] = Tr_j - \frac{1}{2}\sum_{i=1}^{T}\mathbb{E}_0^Q[d_{j,i}h_{j,i}],$$

$$\mathbb{E}_0^Q(R_{j,T}^2) = T^2r_j^2 - Tr_j\sum_{i=1}^{T}\mathbb{E}_0^Q\left[d_{j,i}h_{j,i}\right] + \frac{1}{4}S_{j,D_1} + S_{j,D_2} - S_{j,D_3},$$

$$\begin{aligned}\mathbb{E}_0^Q(R_{j,T}^3) = {} & T^3r_j^3 - \frac{3}{2}T^2\cdot r_j^2\cdot\sum_{i=1}^{T}\mathbb{E}_0^Q\left[d_{j,i}h_{j,i}\right] + 3Tr_j\left(\frac{1}{4}S_{j,D_1} + S_{j,D_2} - S_{j,D_3}\right)\\ & + \left(-\frac{1}{8}S_{j,T_1} + S_{j,T_2} + \frac{3}{4}S_{j,T_3} - \frac{3}{2}S_{j,T_4}\right),\end{aligned}$$

$$\begin{aligned}\mathbb{E}_0^Q(R_{j,T}^4) = {} & T^4r_j^4 - 2T^3\cdot r_j^3\sum_{i=1}^{T}\mathbb{E}_0^Q[d_{j,i}h_{j,i}] + 6T^2r_j^2\left(\frac{1}{4}S_{j,D_1} + S_{j,D_2} - S_{j,D_3}\right)\\ & + Tr_j\left(-\frac{1}{2}S_{j,T_1} + 4S_{j,T_2} + 3S_{j,T_3} - 6S_{j,T_4}\right)\\ & + \left(\frac{1}{16}S_{j,Q_1} + S_{j,Q_2} - \frac{1}{2}S_{j,Q_3} + \frac{3}{2}S_{j,Q_4} - 2S_{j,Q_5}\right).\end{aligned}$$

We simplify the notation of the risk-neutral dynamic and have

$$\log h_{j,t+1} = \overline{\omega}_j + \gamma_j(\phi_j - 1)\log d_{j,t} + \overline{\beta}_j\log h_{j,t} + \epsilon_{j,t}, \tag{D-2}$$

where[2]

$$\epsilon_{j,t} = \overline{\tau}_{j,1}(z_{j,t}^* - \lambda_j) + \overline{\tau}_{j,2}((z_{j,t}^* - \lambda_j)^2 - 1) + \gamma_j\sigma_{v_j}v_{j,t}^*,$$

$$\overline{\omega}_j = \omega_j + \gamma_j\xi_j^*, \overline{\beta}_j = \beta_j + \gamma_j\phi_j, \overline{\tau}_{j,1} = \tau_{j,1} + \gamma_j\delta_{j,1}, \overline{\tau}_{j,2} = \tau_{j,2} + \gamma_j\delta_{j,2}.$$

The terms needed for $S_{D_i}s, S_{T_i}s$, and $S_{Q_i}s^3$ can be divided into three classes, that is, terms concerning only CO part, terms concerning only OC part, and terms concerning both of them. We will give the detailed derivation of the third kind of terms, and the former follow the structure of those for daily volatility (h_t) in Huang et al. (2017) and Duan et al. (1997).

Now In order to compute ρ by equation (C-5), we need to calculate

$$\mathbb{E}_0^Q[d_{j,i}h_{j,i}d_{-j,i+k}h_{-j,i+k}],\ \mathbb{E}_0^Q\left[\sqrt{d_{j,i}h_{j,i}}z_{j,i}d_{-j,i+k}h_{-j,i+k}\right],$$

together with

$$\mathbb{E}_0^Q[d_{CO,i}h_{CO,i}d_{OC,i}h_{OC,i}],\ \mathbb{E}_0^Q\left[\sqrt{d_{j,i}h_{j,i}}z_{j,i}d_{-j,i+k}h_{-j,i+k}\right],$$

[2]We also denote $Y_{j,t} = \overline{\omega}_j + \epsilon_{j,t}$.

[3]See Duan et al. (2006).

where $j \in \{CO, OC\}, -j = \{CO, OC\} \backslash j$, and $i + k \leqslant T$.

Referring to the recursive relationship in equation (D-2), one can derive that

$$
\begin{aligned}
&\mathbb{E}_0^Q[d_{j,i}h_{j,i}d_{-j,i+k}h_{-j,i+k}] \\
&= \mathbb{E}_0^Q[d_{j,i}h_{j,i}d_{-j,i}^{\overline{\beta}_{-j}^k}h_{-j,i}^{\overline{\beta}_{-j}^k}] \cdot \mathbb{E}_0^{\mathbb{Q}}[d_{-j,k+1}h_{-j,k+1}] \cdot (d_{-j,1}h_{-j,1})^{-\overline{\beta}_j^k} \cdot \frac{d_{-j,i+k}d_{-j,1}^{\overline{\beta}_{-j}^k}}{d_{-j,k+1} \cdot d_{-j,i}^{\overline{\beta}_{-j}^k}} \cdot \frac{\prod_{l=i}^{i+k-1} d_{-j,l}^{\gamma_{-j}(\phi_{-j}-1)\overline{\beta}_{-j}^{i+k-1-l}}}{\prod_{l=i}^{i+k-1} d_{-j,i+k-l}^{\gamma_{-j}(\phi_{-j}-1)\overline{\beta}_{-j}^{l-i}}} \\
&= \mathbb{E}_0^Q[d_{j,i}h_{j,i}d_{-j,i}^{\overline{\beta}_{-j}^k}h_{-j,i}^{\overline{\beta}_{-j}^k}] \cdot \mathbb{E}_0^{\mathbb{Q}}[d_{-j,k+1}h_{-j,k+1}] \cdot (d_{-j,1}h_{-j,1})^{-\overline{\beta}_j^k} \cdot \frac{d_{-j,i+k}d_{-j,1}^{\overline{\beta}_{-j}^k}}{d_{-j,k+1} \cdot d_{-j,i}^{\overline{\beta}_{-j}^k}} \cdot \frac{\prod_{l=0}^{k-1} d_{-j,i+l}^{\gamma_{-j}(\phi_{-j}-1)\overline{\beta}_{-j}^{k-1-l}}}{\prod_{l=0}^{k-1} d_{-j,k-l}^{\gamma_{-j}(\phi_{-j}-1)\overline{\beta}_{-j}^{l}}}.
\end{aligned}
\tag{D-3}
$$

Similarly,

$$
\begin{aligned}
&\mathbb{E}_0^Q\left[\sqrt{d_{j,i}h_{j,i}}z_{j,i}d_{-j,i+k}h_{-j,i+k}\right] \\
&= \mathbb{E}_0^{\mathbb{Q}}[\sqrt{d_{j,i}h_{j,i}}d_{-j,i}^{\overline{\beta}_{-j}^k}h_{-j,i}^{\overline{\beta}_{-j}^k}] \cdot \mathbb{E}_0^{\mathbb{Q}}[z_{j,i}\exp\left(\overline{\beta}_{-j}^{k-1}Y_{-j,i}\right)] \cdot \mathbb{E}_0^{\mathbb{Q}}[d_{-j,k}h_{-j,k}] \cdot (d_{-j,1}h_{-j,1})^{-\overline{\beta}_j^{k-1}} \cdot \frac{d_{-j,i+k}d_{-j,1}^{\overline{\beta}_{-j}^{k-1}}}{d_{-j,k}d_{-j,i}^{\overline{\beta}_{-j}^{k-1}}} \cdot \frac{\prod_{l=i}^{i+k-1} d_{-j,l}^{\gamma_{-j}(\phi_{-j}-1)\overline{\beta}_{-j}^{i+k-1-l}}}{\prod_{l=0}^{j-2} d_{-j,k-1-l}^{\gamma_{-j}(\phi_{-j}-1)\overline{\beta}_{-j}^{l}}} \\
&= \mathbb{E}_0^{\mathbb{Q}}[\sqrt{d_{j,i}h_{j,i}}d_{-j,i}^{\overline{\beta}_{-j}^j}h_{-j,i}^{\overline{\beta}_{-j}^j}] \cdot \mathbb{E}_0^{\mathbb{Q}}[z_{j,i}\exp\left(\overline{\beta}_{-j}^{k-1}Y_{-j,i}\right)] \cdot \mathbb{E}_0^{\mathbb{Q}}[d_{-j,k}h_{-j,k}] \cdot (d_{-j,1}h_{-j,1})^{-\overline{\beta}_j^{k-1}} \cdot \frac{d_{-j,i+k}d_{-j,1}^{\overline{\beta}_{-j}^{k-1}}}{d_{-j,k}d_{-j,i}^{\overline{\beta}_{-j}^{k-1}}} \cdot \frac{\prod_{l=0}^{k-1} d_{-j,i+l}^{\gamma_{-j}(\phi_{-j}-1)\overline{\beta}_{-j}^{k-1-l}}}{\prod_{l=0}^{k-2} d_{-j,k-1-l}^{\gamma_{-j}(\phi_{-j}-1)\overline{\beta}_{-j}^{l}}}.
\end{aligned}
\tag{D-4}
$$

Then, to find $\mathbb{E}_0^{\mathbb{Q}}[h_{CO,i}^m h_{OC,i}^n]$, we may refer to the following useful result: for bivariate normal distributed random vector

$$
\begin{pmatrix} x \\ y \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right),
$$

it can be shown that

$$
\begin{aligned}
&\mathbb{E}\left[\exp(a_1x^2 + b_1x + a_2y^2 + b_2y)\right] \\
=&\frac{1}{\sqrt{1 - 2(a_1 + a_2) + 4a_1a_2(1 - \rho^2)}} \exp\left(\frac{b_1^2(1 - 2a_2(1 - \rho^2)) + 2b_1b_2\rho + b_2^2(1 - 2a_1(1 - \rho^2))}{2 \cdot (1 - 2(a_1 + a_2) + 4a_1a_2(1 - \rho^2))}\right),
\end{aligned}
$$

so

$$\mathbb{E}_0^Q[h_{CO,i}^m h_{OC,i}^n] = h_{CO,1}^{m\overline{\beta}_{CO}^{i-1}} h_{OC,1}^{n\overline{\beta}_{OC}^{i-1}} \prod_{k=0}^{i-2} d_{CO,i-1-k}^{m\overline{\beta}_{CO}^k \gamma_{CO}(\phi_{CO}-1)} d_{OC,i-1-k}^{n\overline{\beta}_{OC}^k \gamma_{OC}(\phi_{OC}-1)} \prod_{k=0}^{i-2} e^{m\overline{\beta}_{CO}^k \overline{\omega}_{CO} + n\overline{\beta}_{OC}^k \overline{\omega}_{OC}}$$

$$\cdot \prod_{k=0}^{i-2} \exp \begin{pmatrix} -\frac{1}{2}\log\left(1-2(u_1+u_2)+4u_1u_2(1-(\mathbb{E}_0^Q[\rho_{1+k}])^2)\right) \\ +\frac{v_1^2(1-2u_2(1-(\mathbb{E}_0^Q[\rho_{1+k}])^2))+2v_1v_2\mathbb{E}_0^Q[\rho_{1+k}]+v_2^2(1-2u_1(1-(\mathbb{E}_0^Q[\rho_{1+k}])^2))}{2\cdot(1-2(u_1+u_2)+4v_1v_2(1-(\mathbb{E}_0^Q[\rho_{1+k}])^2))} \\ -m\overline{\beta}_{CO}^k\overline{\tau}_{CO,1}\lambda_{CO} + m\overline{\beta}_{CO}^k\overline{\tau}_{CO,2}(\lambda_{CO}^2-1) - n\overline{\beta}_{OC}^k\overline{\tau}_{OC,1}\lambda_{OC} + n\overline{\beta}_{OC}^k\overline{\tau}_{OC,2}(\lambda_{OC}^2-1) \\ +\frac{m^2\overline{\beta}_{CO}^{2k}\gamma_{CO}^2\sigma_{v_{CO}}^2}{2} + mn\overline{\beta}_{CO}^k\overline{\beta}_{OC}^k\gamma_{CO}\gamma_{OC}\sigma_{v_{CO}}\sigma_{v_{OC}}\rho_{v_{CO},v_{OC}} + \frac{n^2\overline{\beta}_{OC}^{2k}\gamma_{OC}^2\sigma_{v_{OC}}^2}{2} \end{pmatrix}, \tag{D-5}$$

where

$$u_1 = m\overline{\beta}_{CO}^k\overline{\tau}_{CO,2}, u_2 = m\overline{\beta}_{CO}^k(2\overline{\tau}_{CO,2}\lambda_{CO} + \overline{\tau}_{CO,1}),$$
$$v_1 = n\overline{\beta}_{OC}^k\overline{\tau}_{OC,2}, v_2 = n\overline{\beta}_{OC}^k(2\overline{\tau}_{OC,2}\lambda_{OC} + \overline{\tau}_{OC,1}).$$

Therefore, when $m = n = 1$,

$$\mathbb{E}_0^Q[h_{CO,i}h_{OC,i}] = h_{CO,1}^{\overline{\beta}_{CO}^{i-1}} h_{OC,1}^{\overline{\beta}_{OC}^{i-1}} \prod_{k=0}^{i-2} d_{CO,i-1-k}^{\overline{\beta}_{CO}^k \gamma_{CO}(\phi_{CO}-1)} d_{OC,i-1-k}^{\overline{\beta}_{OC}^k \gamma_{OC}(\phi_{OC}-1)} \prod_{k=0}^{i-2} e^{\overline{\beta}_{CO}^k \overline{\omega}_{CO} + \overline{\beta}_{OC}^k \overline{\omega}_{OC}}$$

$$\cdot \prod_{k=0}^{i-2} \exp \begin{pmatrix} -\frac{1}{2}\log\left(1-2(a_1+a_2)+4a_1a_2(1-(\mathbb{E}_0^Q[\rho_{1+k}])^2)\right) \\ +\frac{b_1^2(1-2a_2(1-(\mathbb{E}_0^Q[\rho_{1+k}])^2))+2b_1b_2\mathbb{E}_0^Q[\rho_{1+k}]+b_2^2(1-2a_1(1-(\mathbb{E}_0^Q[\rho_{1+k}])^2))}{2\cdot(1-2(a_1+a_2)+4a_1a_2(1-(\mathbb{E}_0^Q[\rho_{1+k}])^2))} \\ -\overline{\beta}_{CO}^k\lambda_{CO} + \overline{\beta}_{CO}^k\overline{\tau}_{CO,2}(\lambda_{CO}^2-1) - \overline{\beta}_{OC}^k\lambda_{OC} + \overline{\beta}_{OC}^k\overline{\tau}_{OC,2}(\lambda_{OC}^2-1) \\ +\frac{\overline{\beta}_{CO}^{2k}\gamma_{CO}^2\sigma_{v_{CO}}^2}{2} + \overline{\beta}_{CO}^k\overline{\beta}_{OC}^k\gamma_{CO}\gamma_{OC}\sigma_{v_{CO}}\sigma_{v_{OC}}\rho_{v_{CO},v_{OC}} + \frac{\overline{\beta}_{OC}^{2k}\gamma_{OC}^2\sigma_{v_{OC}}^2}{2} \end{pmatrix}. \tag{D-6}$$

Similarly,

$$\mathbb{E}_0^Q[h_{CO,i}^{\frac{1}{2}}h_{OC,i}^{\frac{1}{2}}] = \left(h_{CO,1}^{\overline{\beta}_{CO}^{i-1}}h_{OC,1}^{\overline{\beta}_{OC}^{i-1}}\prod_{k=0}^{t-2}d_{CO,i-1-k}^{\overline{\beta}_{CO}^{k}\gamma_{CO}(\phi_{CO}-1)}d_{OC,i-1-k}^{\overline{\beta}_{OC}^{k}\gamma_{OC}(\phi_{OC}-1)}\prod_{k=0}^{i-2}e^{\overline{\beta}_{CO}^{k}\overline{\omega}_{CO}+\overline{\beta}_{OC}^{k}\overline{\omega}_{OC}}\right)^{\frac{1}{2}}$$
$$\cdot\prod_{k=0}^{i-2}\exp\left(\begin{array}{c} -\frac{1}{2}\left(1-2(a_3+a_4)+4a_3a_4(1-(\mathbb{E}_0^Q[\rho_{1+k}])^2)\right) \\ +\frac{b_3^2(1-2a_4(1-(\mathbb{E}_0^Q[\rho_{1+k}])^2))+2b_3b_4\mathbb{E}_0^Q[\rho_{1+k}]+b_4^2(1-2a_3(1-(\mathbb{E}_0^Q[\rho_{1+k}])^2))}{2\cdot(1-2(a_3+a_4)+4a_3a_4(1-(\mathbb{E}_0^Q[\rho_{1+k}])^2))} \\ +\frac{1}{2}(-\overline{\beta}_{CO}^{k}\lambda_{CO}+\overline{\beta}_{CO}^{k}\overline{\tau}_{CO,2}(\lambda_{CO}^2-1)-\overline{\beta}_{OC}^{k}\lambda_{OC}+\overline{\beta}_{OC}^{k}\overline{\tau}_{OC,2}(\lambda_{OC}^2-1)) \\ +\frac{1}{4}(\frac{\overline{\beta}_{CO}^{2k}\gamma_{CO}^2\sigma_{v_{CO}}^2}{2}+\overline{\beta}_{CO}^{k}\overline{\beta}_{OC}^{k}\gamma_{CO}\gamma_{OC}\sigma_{v_{CO}}\sigma_{v_{OC}}\rho_{v_{CO},v_{OC}}+\frac{\overline{\beta}_{OC}^{2k}\gamma_{OC}^2\sigma_{v_{OC}}^2}{2}) \end{array}\right), \tag{D-7}$$

where

$$\begin{aligned} &a_1=\overline{\beta}_{CO}^{k}\overline{\tau}_{CO,2}, b_1=\overline{\beta}_{CO}^{k}(-2\overline{\tau}_{CO,2}\lambda_{CO}+\overline{\tau}_{CO,1}),\\ &a_2=\overline{\beta}_{OC}^{k}\overline{\tau}_{OC,2}, b_2=\overline{\beta}_{OC}^{k}(-2\overline{\tau}_{OC,2}\lambda_{OC}+\overline{\tau}_{OC,1}),\\ &a_3=\frac{1}{2}a_1, b_3=\frac{1}{2}b_1, a_4=\frac{1}{2}a_2, b_4=\frac{1}{2}b_2. \end{aligned}$$

Additionally,

$$\begin{aligned} &\mathbb{E}_0^{\mathbb{Q}}[z_{j,i}\exp\left(kY_{-j,i}\right)]\\ &=\mathbb{E}_0^{\mathbb{Q}}\left[z_{j,i}\exp\left(k\overline{\omega}_{-j}+k\overline{\tau}_{-j,1}(z_{-j,i}-\lambda_{-j})+k\overline{\tau}_{-j,2}((z_{-j,i}-\lambda_{-j})^2-1)+k\tilde{\sigma}_{v_{-j}}v_{-j,i}\right)\right]\\ &=\mathbb{E}_0^{\mathbb{Q}}\left[z_{j,i}\exp\left(k\overline{\tau}_{-j,2}z_{-j,i}^2+k(\overline{\tau}_{-j,1}-2\overline{\tau}_{-j,2}\lambda_{-j})z_{-j,i}\right)\right]\exp\left(k\overline{\omega}_{-j}-k\overline{\tau}_{-j,1}\lambda_{-j}+k\overline{\tau}_{-j,2}(\lambda_{-j}^2-1)+\frac{k^2\tilde{\sigma}_{v_{-j}}^2}{2}\right)\\ &=\exp\left(k\overline{\omega}_{-j}-k\overline{\tau}_{-j,1}\lambda_{-j}+k\overline{\tau}_{-j,2}(\lambda_{-j}^2-1)+\frac{b_{-j}^2}{4a_{-j}}+\frac{k^2\tilde{\sigma}_{v_{-j}}^2}{2}\right)\mathbb{E}_0^{\mathbb{Q}}[\rho_i]\frac{b_{-j}}{(2a_{-j})^{\frac{3}{2}}}, \end{aligned}$$

where $a_{-j}=\frac{1}{2}-k\overline{\tau}_{-j,2}, b_{-j}=k\left(\overline{\tau}_{-j,1}-2\overline{\tau}_{-j,2}\lambda_{-j}\right).$

Since $\varrho=\frac{1}{2}\log\left(\frac{1+\rho_i}{1-\rho_i}\right)$, we have $\mathbb{E}_0^Q[\rho_i]=\mathbb{E}_0^Q[\frac{e^{2\varrho_i}-1}{e^{2\varrho_i}+1}]$ and

$$\varrho_i=(\widetilde{\omega}+\widetilde{\gamma}\widetilde{\xi})(1+(\widetilde{\beta}+\widetilde{\gamma}\widetilde{\phi})+\cdots(\widetilde{\beta}+\widetilde{\gamma}\widetilde{\phi})^{i-2})+(\widetilde{\beta}+\widetilde{\gamma}\widetilde{\phi})^{t-1}\varrho_1+\widetilde{\gamma}\sigma_{\widetilde{v}}(\widetilde{v}_{i-1}+\cdots+(\widetilde{\beta}+\widetilde{\gamma}\widetilde{\phi})^{i-2}\widetilde{v}_1).$$

Taking expectation, and we have

$$\mathbb{E}_0^Q[\rho_i]=\int_{-\infty}^{+\infty}\frac{e^{2((\widetilde{\omega}+\widetilde{\gamma}\widetilde{\xi})(1+(\widetilde{\beta}+\widetilde{\gamma}\widetilde{\phi})+\cdots(\widetilde{\beta}+\widetilde{\gamma}\widetilde{\phi})^{i-2})+(\widetilde{\beta}+\widetilde{\gamma}\widetilde{\phi})^{t-1}q_1+\widetilde{\gamma}\sigma_{\widetilde{v}}\widetilde{V}_{i-1})}-1}{e^{2((\widetilde{\omega}+\widetilde{\gamma}\widetilde{\xi})(1+(\widetilde{\beta}+\widetilde{\gamma}\widetilde{\phi})+\cdots(\widetilde{\beta}+\widetilde{\gamma}\widetilde{\phi})^{i-2})+(\widetilde{\beta}+\widetilde{\gamma}\widetilde{\phi})^{t-1}q_1+\widetilde{\gamma}\sigma_{\widetilde{v}}\widetilde{V}_{i-1})}+1}\cdot\frac{1}{\sqrt{2\pi}\sigma_i}\exp\left(-\frac{\widetilde{V}_{i-1}^2}{2\sigma_i^2}\right)d\widetilde{V}_{i-1}, \tag{D-8}$$

where

$$\widetilde{V}_{i-1} = \widetilde{v}_{i-1} + \cdots + (\widetilde{\beta} + \widetilde{\gamma}\widetilde{\phi})^{i-2}\widetilde{v}_1 \sim N\left(0, \sigma_i^2\right), \quad \sigma_i = \frac{1 - (\widetilde{\beta} + \widetilde{\gamma}\widetilde{\phi})^{2i-2}}{1 - (\widetilde{\beta} + \widetilde{\gamma}\widetilde{\phi})^2}.$$

Finally, following Duan et al. (2006), we approximate fractional h based on the following Taylor expansion in the empirical part.

$$\begin{aligned}
\mathbb{E}_0^Q[h_{j,t}^a] \approx (1 &+ \frac{32}{12}a - \frac{23}{8}a^2 + \frac{13}{12}a^3 - \frac{1}{8}a^4)\mathbb{E}_0^Q[h_{j,t}]^a \\
&+ (-3a + \frac{19}{4}a^2 - 2a^3 + \frac{1}{4}a^4)\mathbb{E}_0^Q[h_{j,t}]^{a-2}\mathbb{E}_0^Q[h_{j,t}^2] \\
&+ (\frac{4}{3}a - \frac{7}{3}a^2 + \frac{7}{6}a^3 - \frac{1}{6}a^4)\mathbb{E}_0^Q[h_{j,t}]^{a-3}\mathbb{E}_0^Q[h_{j,t}^3] \\
&+ (-\frac{1}{4}a + \frac{11}{24}a^2 - \frac{1}{4}a^3 + \frac{1}{24}a^4)\mathbb{E}_0^Q[h_{j,t}]^{a-4}\mathbb{E}_0^Q[h_{j,t}^4].
\end{aligned} \tag{D-9}$$