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

E2022015 2022-08-16

Realized GARCH, CBOE VIX, and the Volatility Risk Premium*†

Peter Reinhard Hansen[a] & **Zhuo Huang**[b] & **Chen Tong**[c] & **Tianyi Wang**[d]

[a] *University of North Carolina & Copenhagen Business School*

[b] *National School of Development, Peking University*

[c] *Department of Finance, School of Economics & Wang Yanan Institute for Studies in Economics (WISE), Xiamen University*

[d] *School of Banking and Finance, University of International Business and Economics*

May 17, 2022

Abstract

We show that the Realized GARCH model yields closed-form expression for both the Volatility Index (VIX) and the volatility risk premium (VRP). The Realized GARCH model is driven by two shocks, a return shock and a volatility shock, and these are natural state variables in the stochastic discount factor (SDF). The volatility shock endows the exponentially affine SDF with a compensation for volatility risk. This leads to dissimilar dynamic properties under the physical and risk-neutral measures that can explain time-variation in the VRP. In an empirical application with the S&P 500 returns, the VIX, and the VRP, we find that the Realized GARCH model significantly outperforms conventional GARCH models.

Keywords: Realized GARCH; High Frequency Data; Volatility Risk Premium; Realized Variance; VIX.

JEL Classification: C10; C22; C80

Realized GARCH, CBOE VIX, and the Volatility Risk Premium*†

Peter Reinhard Hansen[a] **& Zhuo Huang**[b] **& Chen Tong**[c] **& Tianyi Wang**[d]

[a] *University of North Carolina & Copenhagen Business School*

[b] *National School of Development, Peking University*

[c] *Department of Finance, School of Economics & Wang Yanan Institute for Studies in Economics (WISE), Xiamen University*

[d] *School of Banking and Finance, University of International Business and Economics*

May 17, 2022

Abstract

We show that the Realized GARCH model yields closed-form expression for both the Volatility Index (VIX) and the volatility risk premium (VRP). The Realized GARCH model is driven by two shocks, a return shock and a volatility shock, and these are natural state variables in the stochastic discount factor (SDF). The volatility shock endows the exponentially affine SDF with a compensation for volatility risk. This leads to dissimilar dynamic properties under the physical and risk-neutral measures that can explain time-variation in the VRP. In an empirical application with the S&P 500 returns, the VIX, and the VRP, we find that the Realized GARCH model significantly outperforms conventional GARCH models.

Keywords: Realized GARCH; High Frequency Data; Volatility Risk Premium; Realized Variance; VIX.
JEL Classification: C10; C22; C80

*Correspondence author: Chen Tong (tongchen@xmu.edu.cn). Department of Finance, School of Economics & Wang Yanan Institute for Studies in Economics (WISE), Xiamen University, Fujian, 361005, China.

†Zhuo Huang acknowledges financial support by the National Natural Science Foundation of China (71671004) and Tianyi Wang acknowledges financial support by the National Natural Science Foundation of China (71871060). Chen Tong's research was supported by the Fundamental Research Funds for the Central Universities (20720221025).

1 Introduction

The variance risk premium or volatility risk premium (VRP) has been the focus of much research since the seminal papers by Coval and Shumway (2001) and Bakshi and Kapadia (2003). The VRP is the difference between the expected return variance under the risk-neutral measure and the expected return variance under the physical measure,[1] where the former can be inferred from option prices. The leading example is the VIX, which is a model-free measure of the expected variance over the next 30 days under the risk-neutral measure. Expectations under the physical measure can be based on a suitable volatility model.[2] The VRP is a measure of volatility risk compensation and it is typically positive. This is to be expected because large increases in volatility tend to coincide with large negative returns. This relationship is observed for a broad range of financial assets, see e.g. Coval and Shumway (2001), Bakshi and Kapadia (2003) and Carr and Wu (2009). The VRP is also recognized as a distinct risk factor that predicts both aggregate stock returns and the cross-section of stock returns, see e.g. Bollerslev et al. (2009), Bekaert and Hoerova (2014), and Cremers et al. (2015). Moreover, the VRP plays an important role in option pricing, see e.g. Byun et al. (2015) and Song and Xiu (2016). Given its importance, it is interesting to develop an econometric model that can explain the time variation in the VRP while being coherent with other features of the data.

Conventional GARCH models specify expectations under the physical measure, $\mathbb{P}$, and additional structure is needed before expectations can be computed under the risk-neutral measure, $\mathbb{Q}$. Duan (1995) pioneered the use of GARCH models for option pricing by introducing a locally risk neutral valuation relationship (LRNVR). The LRNVR defines a link between the expected volatility under the $\mathbb{P}$ and $\mathbb{Q}$. With this additional structure in place, GARCH models can be used to price options and the corresponding VRP can be inferred. Unfortunately, standard GARCH models combined with LRNVR cannot adequately explain the VRP, as shown by Hao and Zhang (2013). They found that the VIX implied by GARCH models is substantially below the observed VIX. Hao and Zhang (2013) explored if this shortcoming could be amended by modifying the objective function to also target the VIX. Unfortunately, this leads to parameter values (in the GARCH model) that contradict the empirical properties under $\mathbb{P}$. In our empirical application, we also reach the conclusion that GARCH models in conjunction with LRNVR cannot explain the dynamic properties under both probability measures. A strong argument for looking beyond standard GARCH models is provided by their diffusion limits.

[1] This definition of the VRP follows that in Bollerslev et al. (2009). Other definitions are used in the literature.

[2] For instance a GARCH model or a reduced-form model for the realized volatility. For alternative methods for computing the expected variance under both the physical and risk-neutral measures, see Bollerslev et al. (2009), Bollerslev et al. (2011), Wu (2011) and Conrad and Loch (2015).

These reveal that GARCH models are unable to fully compensate for volatility risk, because GARCH models lack a separate volatility shock variable. Stochastic volatility (SV) models, such as those by Taylor (1986) and Kim et al. (1998), are better suited for this situation because they have a dual-shock structure with distinct shocks to returns and volatilities. This property facilitates a distinct compensation for volatility risk, see e.g. Bollerslev et al. (2011). The main drawback of SV models is that they are more involved to estimate than observation-driven models, such as GARCH models.[3]

It is evident that the GARCH framework must be generalized in order to become a coherent model of $\mathbb{P}$ and $\mathbb{Q}$. This requires either a more flexible volatility model or a more sophisticated risk neutralization method. In this paper, we pursue both extensions by combining the Realized GARCH model with an exponentially affine stochastic discount factor. This framework includes compensation for both equity risk and volatility risk and it yields closed-form expressions for both the VIX and the VRP. The Realized GARCH model is an observation-driven model that conveniently has a dual-shock structure that is similar to that of SV models. This model is simple to estimate and easy to combine with an exponentially affine stochastic discount factor. The parameter estimation can be adapted to include VIX pricing errors in the objective function, similar to the estimation method used in Bardgett et al. (2019), see also Andersen et al. (2019) who links the realized volatility to volatility in an SV model. The estimated model has several interesting properties. It delivers a higher level of volatility, a higher volatility-of-volatility, and a stronger (more negative) leverage correlation under $\mathbb{Q}$ than under $\mathbb{P}$. The estimated model also generates higher levels of skewness and kurtosis in cumulative returns under $\mathbb{Q}$ than under $\mathbb{P}$. The difference between log-volatility under $\mathbb{P}$ and $\mathbb{Q}$ can conveniently be decomposed into two terms, where the first term is compensation for equity risk through the leverage effect and the second term is compensation for volatility specific risk.

In an empirical analysis of daily S&P 500 returns, realized volatilities, and the VIX over 15 years, we compare the proposed model with a range of alternative specifications. These include the EGARCH model by Nelson (1991), the GARCH model by Bollerslev (1986), Heston-Nandi GARCH by Heston and Nandi (2000). These models are combined with either the LRNVR by Duan (1995) or the variance dependent SDF by Christoffersen et al. (2013).[4] We find that the new model has the best in-sample and out-of-sample VIX pricing performance, and the proposed model does particularly well during the turmoil period with the global financial crisis. We find that the Realized GARCH model provides the best empirical fit and, importantly, provides superior out-of-sample forecast of all variables of interest.

[3]The same complication arises with Jump-GARCH models, because they typically rely on particle filters for estimation, see e.g. Ornthanalai (2014).

[4]Christoffersen et al. (2013), introduced a variance-dependent SDF to improve the option pricing performance of the Heston-Nandi GARCH model. The idea was also used in Byun et al. (2015) in the context of Jump-GARCH models.

The improved empirical results are driven by the inclusion of realized volatility in the modeling. This helps in two ways. First, the inclusion improves volatility forecasts and this greatly improve the log-likelihood of returns under $\mathbb{P}$. Second, the inclusion of a realized volatility enables us to define a volatility shock that serves as a second state variables in the SDF. This state variable characterizes the compensation for volatility risk, which is important for explaining key differences between $\mathbb{P}$ and $\mathbb{Q}$ and time-variation therein.[5] The VIX pricing formula is deduced from restrictive distributional assumptions and we explore its sensitivity to distributional misspecification. The situation with non-Gaussian innovation is assessed by amending the VIX pricing formula using a Gram-Charlier expansion. This leads to a VIX pricing formula that depends on higher-order moments. However, in our empirical analysis we find the difference between the original and amended expressions to be negligible, even though skewness and excess kurtosis are significantly different from zero.

The remainder of this paper is organized as follows. We present the Realized GARCH model, risk-neutralization, and the model implied VIX/VRP formula in Section 2 and discuss the distinct model dynamics under $\mathbb{P}$ and $\mathbb{Q}$ in Section 3. We introduce the set of competing models in Section 4 and present our empirical analysis in Section 5. We explore distributional misspecification in Section 6, and conclude in Section 7. We present all the proofs in Appendix A and include a range of additional robustness checks in Appendix B. These including results based on a different definition of VIX pricing errors and alternative choices for the realized measure of volatility.

2 The Realized GARCH Model and VIX Pricing

The Realized GARCH framework is a joint model of returns and realized volatility measures. Returns are modeled with a GARCH model, which is augmented to include a realized measure of volatility, and the Realized GARCH framework is characterized by a measurement equation that ties the realized measure to the conditional variance. Realized measures of volatility are computed from high frequency data where the realized variance (RV) and the realized kernel (RK) by Barndorff-Nielsen et al. (2008) are prime examples. Realized GARCH models are generally found to outperform conventional GARCH models in terms of modeling returns as well as forecasting volatility. The reason is simply that the realized measures provide more accurate measurements of volatility than daily returns, and conventional GARCH models rely on the latter for "updating" the time-varying volatility. The Realized GARCH framework was introduced by Hansen et al. (2012) and later refined in Hansen and Huang (2016) to

[5] The need for a model to simultaneously explain the variation under $\mathbb{P}$ and $\mathbb{Q}$ was pointed out in Bates (1996), and has since received much attention in the option pricing literature, see, e.g., Pan (2002), Eraker (2004), and Santa-Clara and Yan (2010).

have a more flexible leverage function and to allow for the inclusion of multiple realized measures.

2.1 Model under the Physical Measure

We adapt the model in Hansen and Huang (2016) to the present context, by adding an appropriate compensation for equity risk. Under the physical measure the model is characterized by:

$$r_t = r + \lambda\sqrt{h_t} - \tfrac{1}{2}h_t + \sqrt{h_t}z_t, \quad (1)$$

$$\log h_{t+1} = \omega + \beta \log h_t + \tau(z_t) + \gamma\sigma u_t, \quad (2)$$

$$\log x_t = \kappa + \phi \log h_t + \delta(z_t) + \sigma u_t, \quad (3)$$

where r_t is the logarithmic return, λ is the price of equity risk, $h_t = \mathrm{var}_{t-1}(r_t)$ is the conditional variance, r is the risk-free rate, $z_t = (r_t - \mathbb{E}_{t-1}r_t)/\sqrt{h_t}$ is the standardized return, and x_t is the realized measure of volatility. Our addition to this model framework, is a compensation for equity risk that adds the term, $\lambda\sqrt{h_t} - \frac{1}{2}h_t$, to the return equation (1). The two random innovations, z_t and u_t, that are assumed to be i.i.d standard Gaussian, $N(0, 1)$. The quadratic functions $\tau(z) = \tau_1 z + \tau_2(z^2 - 1)$ and $\delta(z) = \delta_1 z + \delta_2(z^2 - 1)$ are *leverage functions* that capture dependence between return shocks and volatility shocks. The parameter σ can be interpreted as the volatility-of-volatility shock. The model simplifies to a variant of the classical EGARCH model of Nelson (1991) when $\gamma = 0$.[6]

A key property of the model is that two shocks, z_t and u_t, are included in the GARCH equation, (2). This is contrast to conventional GARCH models, where the conditional volatility is solely driven by lagged daily returns.[7] The dual-shock structure is important for describing the dynamic properties under both $\mathbb{P}$ and $\mathbb{Q}$ simultaneously, such that the dynamic properties of returns, the VIX, and the VRP, can be explained within a unified coherent framework.

2.2 Risk Neutralization and Properties under the Risk-Neutral Measure

Before we can price the VIX, we need to state how the physical measure, $\mathbb{P}$, relates to the risk-neutral counterpart, $\mathbb{Q}$. In the literature on option pricing with GARCH models, the most commonly used risk neutralization methods are the locally risk-neutral valuation relationship (LRNVR) by Duan (1995)

[6]We use z_t^2 in place of $|z_t|$ that was used the original EGARCH model, which has some advantages, see Hansen et al. (2012) and Hansen and Huang (2016). For completeness, we have also estimated a Realized GARCH model with $\tau_1 z_t + \tau_2(|z_t| - \sqrt{2/\pi})$, which led to very similar qualitative and quantitative results.

[7]For derivative pricing, this "single shock" structure is a serious limitation because the equity risk premium parameter, λ, must be increased to unreasonable high levels in order to explain the variance risk premium, see Hao and Zhang (2013) and our empirical results in Table 2.

and the variance-dependent SDF by Christoffersen et al. (2013). These methods are applicable to a single-shock GARCH models and do not apply to the dual-shock structure in our framework. We will instead adopt an exponentially affine stochastic discount factor, which has previously been used for risk neutralization with multiple shocks in Corsi et al. (2013).

The stochastic discount factor, M_{t+1}, must satisfy $\mathbb{E}_t^{\mathbb{Q}}[X_{t+1}] = \mathbb{E}_t^{\mathbb{P}}[M_{t+1}X_{t+1}]$ for all asset prices, X_{t+1}. In the Realized GARCH framework it is natural to use z_{t+1} and u_{t+1} as state variables, and we will adopt the SDF defined by:

$$M_{t+1} = \frac{\exp(-\lambda z_{t+1} - \xi u_{t+1})}{\mathbb{E}_t^{\mathbb{P}} \exp(-\lambda z_{t+1} - \xi u_{t+1})} = \exp\left\{-\lambda z_{t+1} - \xi u_{t+1} - \tfrac{1}{2}(\lambda^2 + \xi^2)\right\}. \tag{4}$$

Empirically, one would expect λ to be positive and ξ to be negative, which correspond to a positive equity risk premium and a negative variance risk premium, respectively. It should be noted that the parameter, λ, that appears in (4) is identical to the λ in the return equation, (1). This is not by assumption but an implication of a no-arbitrage condition. If we, as a starting point, permitted the λ in (4), to be a free and, possibly, time-varying parameter, then it can be shown that this coefficient must be constant and equal to λ in (1). This is a consequence of a no-arbitrage condition, see Lemma A.1 in the Appendix.

2.2.1 Dynamic Properties under the Risk-Neutral Measure

Under the risk-neutral measure, $\mathbb{Q}$, it follows that the moment generating function (MGF) is given by

$$\begin{aligned}\Psi(s_1, s_2) = \mathbb{E}_t^{\mathbb{Q}}[\exp(s_1 z_{t+1} + s_2 u_{t+1})] &= \mathbb{E}_t^{\mathbb{P}}[M_{t+1}\exp(s_1 z_{t+1} + s_2 u_{t+1})] \\ &= \exp[-s_1\lambda - s_2\xi + \tfrac{1}{2}(s_1^2 + s_2^2)].\end{aligned}$$

This MGF is identical to $\mathbb{E}_t^{\mathbb{P}}[\exp(s_1 z_{t+1}^* + s_2 u_{t+1}^*)]$, where $z_{t+1}^* = z_{t+1} + \lambda$ and $u_{t+1}^* = u_{t+1} + \xi$, and it implies the following dynamic model under the risk-neutral measure:

$$r_{t+1} = r - \tfrac{1}{2}h_{t+1} + \sqrt{h_{t+1}}z_{t+1}^*, \tag{5}$$

$$\log h_{t+1} = \tilde{\omega} + \beta \log h_t + \tilde{\tau}_1 z_t^* + \tau_2(z_t^{*2} - 1) + \gamma\sigma u_t^*, \tag{6}$$

$$\log x_t = \tilde{\kappa} + \phi \log h_t + \tilde{\delta}_1 z_t^* + \delta_2(z_t^{*2} - 1) + \sigma u_t^*, \tag{7}$$

where (z_t^*, u_t^*) has a bivariate Gaussian distribution, $N(0, I)$, under $\mathbb{Q}$. The relationships between parameters (under $\mathbb{P}$ and $\mathbb{Q}$) are: $\tilde{\omega} = \omega - \tau_1\lambda + \tau_2\lambda^2 - \gamma\sigma\xi$, $\tilde{\tau}_1 = \tau_1 - 2\tau_2\lambda$, $\tilde{\kappa} = \kappa - \delta_1\lambda + \delta_2\lambda^2 - \sigma\xi$, and $\tilde{\delta}_1 = \delta_1 - 2\delta_2\lambda$, see Lemma A.2 for details.

The mapping $(z_t, u_t) \mapsto (z_t^*, u_t^*)$ can be viewed as a generalization of LRNVR to the bivariate case. In the present context, this bivariate structure rely on the inclusion of the realized measure in the modeling, which facilitates a more complex dynamic structure than can be achieved with conventional GARCH models. The conventional GARCH model emerges as a special case when $\gamma = 0$ and $\xi = 0$. In this situation, (5) and (6) simplify to an EGARCH model with the change of measure, $z_{t+1}^* = z_{t+1} + \lambda$. This reveals a close relation between the GARCH models with exponentially affine SDF and the simple change of measure that was proposed by Duan (1995). This connection appears to have been overlooked in the exiting literature.

2.3 The VIX Pricing Formula

The Chicago Board Options Exchange's (CBOE) VIX index can be computed as the square-root of the annualized expected variance over the next 30 calendar days, where the expectation is calculated under the risk-neutral measure. We will model returns and realized measure using daily data and we will use 22 trading days and 252 trading days to represent a month and a year, respectively. The model-based VIX formula is therefore given by

$$\mathrm{VIX}_t^{\mathrm{model}} = 100 \times \sqrt{\tfrac{252}{22} \sum_{k=1}^{22} \mathbb{E}_t^{\mathbb{Q}}(h_{t+k})},$$

where the expression for $\mathbb{E}_t^{\mathbb{Q}}(h_{t+k})$ is model specific. The combination of the Realized GARCH model and the exponentially affine SDF leads to the following expression:

Theorem 1. *For the Realized GARCH model (1)-(3) and the SDF (4), the model-implied VIX is given by:*

$$\mathrm{VIX}_t^{\mathrm{RG}} = 100 \times \sqrt{\frac{252}{22} \left[h_{t+1} + \sum_{k=2}^{22} \left(\prod_{i=0}^{k-2} F_i \right) h_{t+1}^{\beta^{k-1}} \right]}, \tag{8}$$

where $F_i = (1 - 2\beta^i \tau_2)^{-1/2} \exp\left\{ \beta^i(\tilde{\omega} - \tau_2) + \frac{1}{2}\beta^{2i}[\frac{\tilde{\tau}_1^2}{1-2\beta^i\tau_2} + \gamma^2\sigma^2] \right\}$*, with* $\tilde{\omega} = \omega - \tau_1\lambda + \tau_2\lambda^2 - \delta\sigma\xi$ *and* $\tilde{\tau}_1 = \tau_1 - 2\lambda\tau_2$*.*

The expression (8) facilitates an easy comparison of the model implied VIX with the actual VIX index, and it is analogous to the expressions obtained for a range of conventional GARCH models in Hao and Zhang (2013). The proof of Theorem 1 is given in Appendix A.

2.4 Volatility Risk Premium

The literature has proposed several definitions of the VRP, see Bollerslev et al. (2009).[8] We adopt the following definition,

$$\mathrm{VRP}_t^{\texttt{modelfree}} = \mathrm{VIX}_t - \sqrt{\tfrac{252}{22}\mathbb{E}_t^{\mathbb{P}}\left(\sum_{i=1}^{22}\mathrm{RVcc}_{t+i}\right)} \times 100,$$

where $\mathrm{RVcc}_t = \mathrm{RV}_t + r_{co,t}^2$, RV_t is the realized variance estimator for the hours with active trading on day t, and $r_{co,t}^2$ is the squared overnight return, which is computed from the closing price on day $t-1$ and the opening price of day t. This difference between the observed VIX and the expected realized measure of volatility (for the corresponding 22 trading days) is a model-free measure of the VRP. This is a theoretical quantity, because the expectation operator depends on $\mathbb{P}$ that is unknown in practice. The following empirical VRP

$$\mathrm{VRP}_t^{\texttt{market}} = \mathrm{VIX}_t - \sqrt{\tfrac{252}{22}\sum_{i=1}^{22}\mathrm{RVcc}_{t-i+1}} \times 100,$$

was proposed in Bollerslev et al. (2009). This quantity relies on the assumption that realized volatility follows a martingale process, such that the expected monthly realized volatility is given by the observed realized volatility over the most recent month. As a robustness check, we also consider a second, alternative, empirical measure, which is based on the heterogeneous autoregressive (HAR) model by Corsi (2009), see Appendix B.1.

Our model-implied VRP is simply

$$\mathrm{VRP}_t^{\texttt{model}} = \left(\sqrt{\tfrac{252}{22}\sum_{k=1}^{22}\mathbb{E}_t^{\mathbb{Q}}(h_{t+k})} - \sqrt{\tfrac{252}{22}\sum_{k=1}^{22}\mathbb{E}_t^{\mathbb{P}}(h_{t+k})}\right) \times 100,$$

which is the annualized, one-month ahead, expected volatility using the risk-neutral measure less the corresponding quantity under the physical probability measure.

3 Key Model Properties under $\mathbb{P}$ and $\mathbb{Q}$

In this section, we analyze the Realized GARCH model with the exponentially affine SDF, under both $\mathbb{P}$ and $\mathbb{Q}$, and we derive key properties of volatility, leverage, and returns under both measures. These

[8]See Bollerslev et al. (2009) for detailed discussion of VRP measures, including ex-post and ex-ante measures, and measures in units of variances and in units of volatilities.

results provide theoretical insight about importance of various model parameters and their interpretations. We derive the properties under the assumption that the parameters in the Realized GARCH model and the SDF satisfy:

$$|\beta| < 1, \quad \lambda, \gamma, \sigma, \tau_2, \delta_2 > 0, \quad \xi, \tau_1, \delta_1 < 0.$$

These inequalities guarantee the following properties: (a) the volatility process is stationary, (b) the equity premium is positive, (c) the volatility premium is negative, and (d) the model has a leverage effect. The stated parameter restrictions are consistent with our empirical results in Section 5.

3.1 Average Volatility

For the average log-volatility (unconditional mean of $\log h_t$) we have that

$$\mathbb{E}^{\mathbb{P}}(\log h) = \frac{\omega}{1-\beta} \qquad \text{and} \qquad \mathbb{E}^{\mathbb{Q}}(\log h) = \frac{\omega - \tau_1\lambda + \tau_2\lambda^2 - \gamma\sigma\xi}{1-\beta}.$$

Given our assumptions above, it follows that the average log-volatility is higher under the $\mathbb{Q}$-measure than under the $\mathbb{P}$-measure. Thus, the logarithmic variant of the VRP,

$$\mathbb{E}^{\mathbb{Q}}(\log h)\text{-}\mathbb{E}^{\mathbb{P}}(\log h) = \frac{-\tau_1\lambda + \tau_2\lambda^2 - \gamma\sigma\xi}{1-\beta} = \frac{-\tau_1\lambda + \tau_2\lambda^2}{1-\beta} + \frac{-\gamma\sigma\xi}{1-\beta}, \tag{9}$$

is positive. The logarithmic variant of the VRP was used in Carr and Wu (2009) and Ammann and Buesser (2013).

From (9) we see that the logarithmic VRP can be decomposed into two terms. One that is driven by the equity risk premium from leverage effect and a second term driven by the compensation for volatility risk and volatility-of-volatility due to volatility shocks. Their relative contributions to the log VRP are given by $\frac{-\tau_1\lambda+\tau_2\lambda^2}{-\tau_1\lambda+\tau_2\lambda^2-\gamma\sigma\xi}$ and $\frac{-\gamma\sigma\xi}{-\tau_1\lambda+\tau_2\lambda^2-\gamma\sigma\xi}$, respectively, which, conveniently, do not depend on β.

It is worth noting that a positive VRP does not require the equity risk premium to be positive if ξ is sufficiently negative. The literature typically finds $\lambda > 0$, e.g. French et al. (1987), but negative values have also been reported, see e.g. Jensen and Lunde (2001).

3.2 Volatility-of-Volatility

The volatility-of-volatility (in log-h) can be derived similarly and is, under $\mathbb{P}$ and $\mathbb{Q}$, given by

$$\text{var}^{\mathbb{P}}(\log h) = \frac{\tau_1^2 + 2\tau_2^2 + \gamma^2\sigma^2}{1-\beta^2} \qquad \text{and} \qquad \text{var}^{\mathbb{Q}}(\log h) = \frac{(\tau_1 - 2\tau_2\lambda)^2 + 2\tau_2^2 + \gamma^2\sigma^2}{1-\beta^2},$$

respectively. Since $\tau_1 < 0$ it follows that their difference, $\text{var}^{\mathbb{Q}}(\log h) - \text{var}^{\mathbb{P}}(\log h) = 4\lambda(\lambda\tau_2^2 - \tau_1\tau_2)/(1-\beta^2)$ is positive, such that the unconditional variance of volatility is higher under $\mathbb{Q}$ than under $\mathbb{P}$.

3.3 Dependence between Returns and Volatility (Leverage)

The dependence between returns and volatility is another important aspect of asset pricing. Here we follow Christoffersen et al. (2014) and quantify this dependence using the conditional correlation between $\log h_{t+1}$ and r_t, (leverage correlations) under $\mathbb{P}$ and $\mathbb{Q}$. Under $\mathbb{P}$ we have

$$\text{cov}_t^{\mathbb{P}}(\log h_{t+1}, r_t) = \mathbb{E}_t^{\mathbb{P}}[(\tau_1 z_t + \tau_2 z_t^2 + \gamma\sigma u_t) z_t \sqrt{h_t}] = \tau_1\sqrt{h_t},$$

such that the conditional correlation is

$$\rho_{\mathbb{P}} = \text{corr}_t^{\mathbb{P}}(\log h_{t+1}, r_t) = \frac{\tau_1}{\sqrt{\tau_1^2 + 2\tau_2^2 + \gamma^2\sigma^2}}.$$

Under the risk-neutral measure, $\mathbb{Q}$, we find that $\text{cov}_t^{\mathbb{Q}}(\log h_{t+1}, r_t) = (\tau_1 - 2\tau_2\lambda)\sqrt{h_t}$ such that

$$\rho_{\mathbb{Q}} = \text{corr}_t^{\mathbb{Q}}(\log h_{t+1}, r_t) = \frac{\tau_1 - 2\lambda\tau_2}{\sqrt{(\tau_1 - 2\tau_2\lambda)^2 + 2\tau_2^2 + \gamma^2\sigma^2}}.$$

These correlations are, as expected, both negative, and it can be shown that $\rho_{\mathbb{Q}}^2 - \rho_{\mathbb{P}}^2 > 0$, such that the leverage effect is more pronounce under $\mathbb{Q}$ than under $\mathbb{P}$.

3.4 Skewness and Kurtosis of Multi-period Returns

While VIX pricing only requires the expectations of future volatility, many other problems, such as option pricing, require an accurate description of the distribution of cumulative returns.[9] Figure 1 presents the unconditional skewness and kurtosis of cumulative returns for the Realized GARCH model, for cumulative returns spanning a period from 1 to 250 days (approximately one year). For comparison, we also include the corresponding results based on the EGARCH model. The simulation designs for the

[9] Duan et al. (1999) provided a method to price options with the skewness and kurtosis of cumulative returns.

two models are the estimates we obtained in our empirical analysis, see Table 2. Because closed-form expressions for skewness and kurtosis of cumulative returns are not readily available, these results are based on simulation methods with 1,000,000 replications. The first 750 days were discarded in each simulation in order to minimize the influence of initial values.

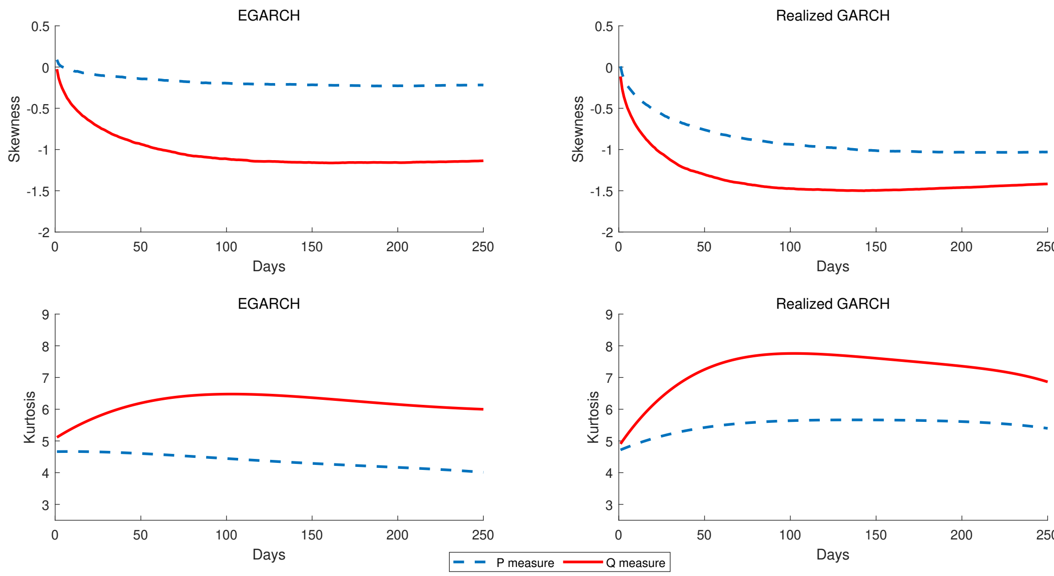

Figure 1: Skewness and kurtosis of cumulative returns under $\mathbb{P}$ and $\mathbb{Q}$ for the EGARCH and the Realized GARCH model.

The results in Figure 1 show that cumulative returns are more left-skewed (have a more negative skewness) under $\mathbb{Q}$ than under $\mathbb{P}$, and the tails are also thicker (larger kurtosis) under $\mathbb{Q}$ than under $\mathbb{P}$. This is true for both the Realized GARCH model and the EGARCH model. However, the magnitude of skewness and kurtosis is much larger for the Realized GARCH model especially under the risk-neutral measures.

In Figure 2, we present the simulated densities for standardized cumulative returns over one month (left panels) and six months (right panels). The densities under $\mathbb{P}$ are in the upper panels and those under $\mathbb{Q}$ are presented in the lower panels, where the solid red lines are for the Realized GARCH model and the dashed blue line are for the EGARCH model based on the parameter estimates we obtained in our empirical analysis. A left skew can be seen for both models and it is more pronounced at longer horizons (six months), especially for the Realized GARCH model. The skewness is also more pronounced under the risk-neutral measure, $\mathbb{Q}$, which is consistent with the results in Figure 1.

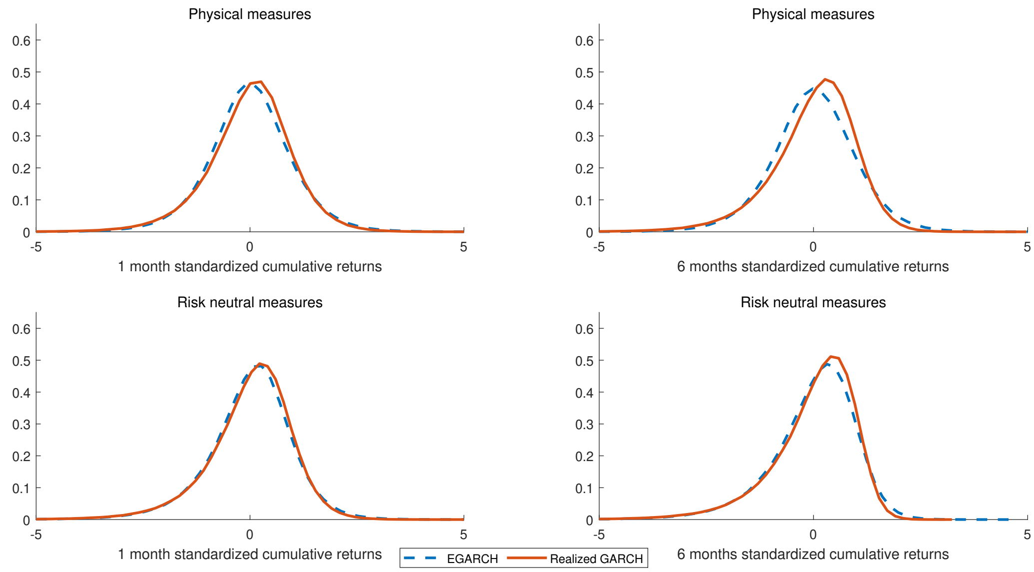

Figure 2: Density of Cumulative Returns via Simulation

4 Some Competing Models and their Properties

In this section, we introduce four alternative models, which we will use to benchmark the Realized GARCH model against. We compare the models ability to explain and predict the three variables: return volatility, the market VIX, and the volatility risk premium. The four alternative models are: the GARCH model by Bollerslev (1986), the EGARCH model by Nelson (1991), the Heston-Nandi GARCH model by Heston and Nandi (2000), which are combined with Duan's LRNVR, and the Heston-Nandi GARCH combined with the variance dependent SDF, as proposed by Christoffersen et al. (2013).[10]

4.1 GARCH and EGARCH Model

GARCH and EGARCH model are commonly used as benchmarks in comparisons of volatility models. The original GARCH model tends to perform well with exchange rate data, but it is typically outperformed by models that can accommodate a leverage effect when applied to equity returns, see Hansen and Lunde (2005a). The volatility dynamics for the GARCH(1,1) is given by

$$h_{t+1} = \omega + \beta h_t + \alpha h_t z_t^2,$$

[10] Hao and Zhang (2013) examined GARCH, EGARCH, TGARCH, AGARCH and CGARCH models. To conserve space, we focus on the GARCH and EGARCH models because the EGARCH had the best performance in the study by Hao and Zhang (2013), and the original GARCH model is a natural benchmark.

and that of the EGARCH(1,1) is given by

$$\log h_{t+1} = \omega + \beta \log h_t + \tau_1 z_t + \tau_2(|z_t| - \sqrt{2/\pi}).$$

From Hao and Zhang (2013, propositions 1-4), these models can be risk-neutralized using the LRNVR and yield the following model-based pricing formulae for VIX:

$$\text{VIX}_t^{\text{G}} = 100 \times \sqrt{252\sigma_h^2 + \tfrac{252}{22}\frac{1-\tilde{\beta}^{22}}{1-\tilde{\beta}}(h_{t+1} - \sigma_h^2)}, \qquad \text{with} \quad \sigma_h^2 = \omega/(1-\tilde{\beta}), \quad \tilde{\beta} = \beta + \alpha(1+\lambda^2)$$

and

$$\text{VIX}_t^{\text{EG}} = 100 \times \sqrt{\tfrac{252}{22}\left[h_{t+1} + \sum_{k=2}^{22}\left(\prod_{i=0}^{k-2} F_i\right) h_{t+1}^{\beta^{k-1}}\right]},$$

respectively, where

$$\begin{aligned} F_i &= \exp\left[\beta\left(\omega - \tau_2\sqrt{\frac{2}{\pi}}\right)\right]\left\{\exp\left[-\beta^i(\tau_1-\tau_2)\lambda + \frac{\beta^{2i}(\tau_2-\tau_1)^2}{2}\right]\Phi[\lambda - \beta^i(\tau_1-\tau_2)]\right. \\ &\quad \left. + \exp\left[-\beta^i(\tau_1+\tau_2)\lambda + \frac{\beta^{2i}(\tau_2+\tau_1)^2}{2}\right]\Phi[\beta^i(\tau_1+\tau_2) - \lambda]\right\}. \end{aligned}$$

4.2 Heston-Nandi GARCH Model under LRNVR

The Heston-Nandi GARCH model is a popular discrete-time model for option pricing. The equity premium is assumed to be proportional to the conditional variance and a specific leverage term is adopted in the GARCH equation,

$$\begin{aligned} r_{t+1} &= r + \lambda h_{t+1} - \tfrac{1}{2}h_{t+1} + \sqrt{h_{t+1}}z_{t+1}, \\ h_{t+1} &= \omega + \beta h_t + \alpha(z_t - \delta\sqrt{h_t})^2. \end{aligned}$$

This structure conveniently yields a closed-form option pricing formula. Under LRNVR risk neutralization the corresponding dynamics under $\mathbb{Q}$ is:

$$\begin{aligned} r_{t+1} &= r - \tfrac{1}{2}h_{t+1} + \sqrt{h_{t+1}}z_{t+1}^*, \\ h_{t+1} &= \omega + \beta h_t + \alpha(z_t^* - (\delta+\lambda)\sqrt{h_t})^2, \\ &= \underbrace{\omega + \alpha}_{=\tilde{\omega}} + \underbrace{\left[\beta + \alpha(\delta+\lambda)^2\right]}_{=\tilde{\beta}} h_t - 2\alpha(\delta+\lambda)\sqrt{h_t}z_t^* + \alpha(z_t^{*2} - 1). \end{aligned}$$

where $z_t^* = z_t + \lambda\sqrt{h_t}$. If $|\tilde{\beta}| < 1$, then the unconditional mean of h_t under $\mathbb{Q}$ is $\sigma_h^2 = \tilde{\omega}/(1-\tilde{\beta})$, and the k-step ahead expected conditional variance is

$$\mathbb{E}_t^{\mathbb{Q}}(h_{t+k}) = \sigma_h^2 + \tilde{\beta}^{k-1}(h_{t+1} - \sigma_h^2).$$

The model implied VIX pricing formula is therefore given by:

$$\text{VIX}_t^{\text{HN}} = \sqrt{\tfrac{252}{22}\sum_{k=1}^{22}\mathbb{E}_t^{\mathbb{Q}}(h_{t+k})} \times 100 = \sqrt{252\sigma_h^2 + \tfrac{252}{22}\tfrac{1-\tilde{\beta}^{22}}{1-\tilde{\beta}}(h_{t+1} - \sigma_h^2)} \times 100.$$

4.3 Heston-Nandi GARCH under Variance Dependent SDF

An alternative to LRNVR is the variance dependent SDF by Christoffersen et al. (2013). As suggested by its name, this SDF depends on h_t, and this dependence has been shown to improve the option pricing performance of the Heston-Nandi GARCH model.

Christoffersen et al. (2013), show that the dynamic properties under $\mathbb{Q}$ are given by

$$\begin{aligned} r_{t+1} &= r - \tfrac{1}{2}h_{t+1}^* + \sqrt{h_{t+1}^*}z_{t+1}^*, \\ h_{t+1}^* &= \omega^* + \beta h_t^* + \alpha^*(z_t^* - \delta^*\sqrt{h_t^*})^2, \end{aligned}$$

where z_t^* has a standard normal distribution and

$$\begin{aligned} h_t^* &= h_t/(1+2\alpha\xi), \qquad & \omega^* = \omega/(1+2\alpha\xi), \\ \alpha^* &= \alpha/(1+2\alpha\xi)^2, \qquad & \delta^* = (\lambda + \delta - \tfrac{1}{2})(1+2\alpha\xi) + \tfrac{1}{2}. \end{aligned}$$

Here ξ is the variance risk aversion parameter, see Christoffersen et al. (2013). The resulting model-implied VIX pricing formula is given by

$$\text{VIX}_t^{\text{HN}_{\text{vd}}} = \sqrt{252\sigma_h^{*2} + \tfrac{252}{22}\tfrac{1-\beta^{*22}}{1-\beta^*}(h_{t+1}^* - \sigma_h^{*2})},$$

where $\sigma_h^{*2} = (\omega^* + \alpha^*)/(1-\beta^*)$ with $\beta^* = \beta + \alpha^*\delta^{*2}$. Unlike LRNVR, the variance dependent SDF will induce a transformation of one-step-ahead conditional variance after the change of measure.[11]

[11] Christoffersen et al. (2013) suggested to estimate $\tilde{\xi} = 1/(1+2\alpha\xi)$ in place of ξ, when estimating the model.

5 Empirical Analysis

For the estimation it is convenient to use a different expression of the GARCH equation,

$$\log h_{t+1} = (\omega - \gamma\kappa) + (\beta - \gamma\phi)\log h_t + (\tau(z_t) - \gamma\delta(z_t)) + \gamma \log x_t, \tag{2'}$$

which is obtained by substituting (3) into (2). This formulation highlights the observation-driven structure of the model, as it shows how the conditional volatility depends on the observable realized measure and (a function of) the lagged standardized return. This makes evaluation and maximization of the log-likelihood straight forward.[12]

5.1 Model Estimation

We estimate the unknown parameters (of the model and the SDF) by maximizing a joint log-likelihood function that is composed of the log-likelihood function of the (Realized) GARCH model and the log-likelihood for VIX pricing errors.

The log-likelihood function for the Realized GARCH model specifies the dynamics for (r_t, x_t) while the GARCH models (GARCH, EGARCH, Heston-Nandi) specifies the dynamics for r_t. These likelihood terms are combined with a log-likelihood for the VIX pricing errors, where the latter is influenced by both the choice of volatility model and the choice of SDF. Following Hao and Zhang (2013) we adopt at Gaussian specification for the pricing error, where $\mathrm{VIX}_t^{\mathrm{Model}} - \mathrm{VIX}_t \sim i.i.d\ N(0, \sigma^2_{\mathtt{vix}})$, and as a robustness check we also estimate the parameters using a second (multiplicative) specification: $\mathrm{VIX}_t = \mathrm{VIX}_t^{\mathrm{Model}}\eta_t$, where it is assumed that $\log \eta_t \sim i.i.d\ N(-\sigma^2_{\mathtt{vix}}/2, \sigma^2_{\mathtt{vix}})$, such that $\mathbb{E}(\eta_t) = 1$. The two specifications produce very similar estimates and similar pricing errors, see Appendix B.2.

For the Realized GARCH model, the total (quasi) log-likelihood is given by

$$\ell_r + \ell_x + \ell_{\mathtt{vix}},$$

[12] For the related stochastic volatility models, direct maximization of the likelihood for is typically impractical, and other estimation methods, such as GMM and simulation based methods, are often employed for this type of models.

where

$$\ell_r = -\tfrac{1}{2}\sum_{t=1}^{T}\{\log 2\pi + \log h_t^{\texttt{RG}} + (r_t - \mu_t^{\texttt{RG}})^2/h_t^{\texttt{RG}}\},$$

$$\ell_x = -\tfrac{1}{2}\sum_{t=1}^{T}\{\log 2\pi + \log \sigma^2 + [\log x_t - \omega - \beta \log h_t^{\texttt{RG}} - \delta(z_t)]^2/\sigma^2\},$$

$$\ell_{\texttt{vix}} = -\tfrac{1}{2}\sum_{t=1}^{T}\{\log 2\pi + \log \sigma_{\texttt{vix}}^2 + (\mathrm{VIX}_t^{\texttt{RG}} - \mathrm{VIX}_t)^2/\sigma_{\texttt{vix}}^2\},$$

with $h_t^{\texttt{RG}}$ given from the GARCH equation (2) and $\mu_t^{\texttt{RG}} = r + \lambda\sqrt{h_t^{\texttt{RG}}} - \frac{1}{2}h_t^{\texttt{RG}}$.

The likelihood of the other models are define similarly with model-specific definitions of μ_t, h_t, and the model-implied VIX. The conventional GARCH models do not have the second term of the log-likelihood, because they do include the realized measure, x_t, in the modeling.

The idea of combining the likelihood of a time-series model with a second likelihood for option pricing errors is now standard in this literature. Some papers including pricing errors for the a range of options, see e.g. Christoffersen et al. (2013) and Christoffersen et al. (2014), or pricing errors for volatility derivatives, see e.g. Wang et al. (2017), Bardgett et al. (2019), or VIX pricing errors as in Hao and Zhang (2013). This is in contrast to an earlier literature that implicitly assumed pricing errors to be zero and adopted the VIX as the volatility variables, see, e.g., Duan and Yeh (2010) who estimated a stochastic volatility model with jumps by exploiting the theoretical link between the VIX and the latent volatility.

Table 1: Summary of Statistics

	Mean	Median	Min	Q1	Q3	Max
VIX	18.410	15.690	9.140	12.940	20.948	80.860
$\sqrt{AnnRV}$	13.573	10.726	4.211	8.517	15.157	73.553
$Ret(\%)$	0.017	0.058	-9.127	-0.378	0.485	10.246
VRP	4.837	4.737	-25.284	3.147	6.510	28.316
	Std	Skew	Kurt	AR1	AR10	AR22
VIX	8.812	2.657	12.662	0.980	0.898	0.810
$\sqrt{AnnRV}$	8.757	3.082	16.113	0.998	0.924	0.771
$Ret(\%)$	1.094	-0.410	15.224	-0.091	0.030	0.039
VRP	3.300	-0.214	10.088	0.860	0.286	0.030

Note: Variables are measure in percent of annualized volatility. For instance, $\sqrt{AnnRV} = 100\sqrt{\frac{252}{22}\sum_{i=1}^{22}\mathrm{RVcc}_{t-i+1}}$.

5.2 Data

Our empirical analysis is based on daily data for S&P 500 stock index and CBOE VIX. We obtain the daily VIX index and the daily returns from Yahoo Finance while the realized measures are downloaded from the Realized Library at Oxford-Man Institute. The primary realized measure is the realized variance from the hours with active trading with the squared overnight return added, see Hansen and Lunde (2005b). As another robustness check of our main results, we have also used different choices of realized measures, see Appendix B.3.

Our full sample spans 15 years, from January 2004 to December 2018. We will present empirical results based on the full sample period as well as out-of-sample results where the model is estimated recursively using a rolling window sample with 750 days. The out-of-sample performance is evaluated over the years 2007 to 2018. We also present separate out-of-sample results for two subsamples: the years 2007-2012, which include the global financial crisis period, and the years 2013-2018, which span the post-crisis period.

5.2.1 CBOE VIX calculation

The VIX index is a model-free measure of volatility. Prior to being annualized, it is computed as

$$\mathrm{VIX}_t = \sqrt{\frac{2}{T}\sum_i \frac{\Delta K_i}{K_i^2}\exp\left(rT\right)Q\left(K_i\right) - \frac{1}{T}\left(\frac{F}{K_0}-1\right)^2},$$

where T is the time to maturity, F is the forward index level, K_0 is the first strike below F, K_i is the strike price of the i-th out-of-the-money option, ΔK_i is the interval between strike prices, r is the risk-free rate associated with time to maturity, $Q(K_i)$ is the midpoint of the bid-ask spread for options with strike K_i. See Jiang and Tian (2005) for a detail discussion on the VIX formula, and Gonzalez-Perez (2015) for a review of model-free measures.[13]

We present summary statistics for the full sample period in Table 1. The data consists of daily returns, the daily realized variances (measured in units of annualized standard deviation), the CBOE VIX, and the VRP. A number of interesting observation can be made from Table 1. First, the distribution of VIX is skewed to the right with one (or more) extremely large values, and the same is seen for the realized volatility. Second, both time series of volatility are highly persistent with large and slowly decaying autocorrelations. Third, the VRP also has a large first-order autocorrelation but its higher-

[13]The VIX formula is described in the CBOE white paper, http://www.cboe.com/micro/vix/vixwhite.pdf, and is based on earlier results in, Carr and Madan (1998), Demeterfi et al. (1999), and Britten-Jones and Neuberger (2000), who applied similar methods to approximate the expected volatility under $\mathbb{Q}$.

order autocorrelations decay much faster than is the case for the VIX and the realized variance. This suggest that the two variables that the VRP is composed of, have a common stochastic trend that cancels out in the difference between the two variables (a type of "cointegration"). On average, the VIX is larger than the realized volatility, with the average VRP being around 4.8%.

5.3 Model Estimates (Full Sample)

In the following, we use the following abbreviations for the models: RG for the Realized GARCH model, EG for EGARCH, G for GARCH, HN for Heston-Nandi GARCH, and $\text{HN}_{\texttt{vd}}$ for Heston-Nandi GARCH with the variance dependent SDF.

We present the parameter estimates for the full sample period for each of the five models in Table 2 along with the QMLE robust standard errors in parentheses and some additional statistics.

An interesting observation can be made about the market price of equity risk, λ. This parameter is similar for the first three models, however the estimated of λ in the EGARCH and GARCH models are 10-20 times larger than the estimate for the Realized GARCH model. The estimates of λ in the EGARCH and GARCH models are in line with those reported in Hao and Zhang (2013). If the model are estimated solely from return data, then the estimates of λ are much smaller, see Hao and Zhang (2013). The reason is that the EG and G models lack a separate volatility risk parameters, and the models inflate the value of λ in order to compensate for the volatility risk that is embedded in the VIX. The λ for the Heston-Nandi model is not directly comparable to those of the other models, because this coefficient is associated with h in the return equation, rather than $\sqrt{h}$ (for the other models).

The persistence parameters under the $\mathbb{P}$-measure is denoted $\pi^{\mathbb{P}}$ and is defined by β for RG and EG, by $\alpha + \beta$ for G, and by $\beta + \alpha\delta^2$ for HN and $\text{HN}_{\texttt{vd}}$. The persistence is quite similar across models and close to unity in all cases. The estimates of τ_1 and δ_1 are negative for both RG model and the EG model, which reflect a negative correlation between return and volatility shocks. This is the so-called leverage effect and these findings are consistent with the existing literature.

The estimate of the volatility risk parameter in the RG model, ξ, is negative and significant. From the decomposition of the (log) VRP we can compute the relative contributions of the two terms in (9) using the estimated RG model. The first term is compensation for the equity risk premium and its contribution ($\propto -\tau_1\lambda + \tau_2\lambda^2$) is estimated to be 2.2%. The second term is the separate compensation for volatility risk and its contribution ($\propto -\gamma\sigma\xi$) is estimated to be 97.8%. This suggests that the majority of VRP is due to compensation for the volatility shock, u_t, and only a small of fraction of the VRP can

Table 2: Parameter estimates (full sample)

Model	RG	EG	G	HN	$HN_{\mathtt{vd}}$
λ	0.015	0.153	0.305	4.518	9.128
	(0.010)	(0.004)	(0.018)	(0.259)	(1.635)
ω	-0.088	-0.086	1.60E-06	-1.44E-06	-1.39E-06
	(0.014)	(0.002)	(1.09E-07)	(9.27E-08)	(1.13E-07)
β	0.991	0.990	0.940	0.870	0.895
	(0.001)	2.22E-04	(0.004)	(0.010)	(0.009)
α			0.054	3.10E-06	2.24E-06
			(0.004)	(1.50E-07)	(2.16E-07)
δ				197.183	202.167
				(12.852)	(18.782)
τ_1	-0.073	-0.062			
	(0.005)	(0.002)			
τ_2	0.012	0.096			
	(0.002)	(0.001)			
γ	0.080				
	(0.009)				
κ	0.427				
	(0.278)				
ϕ	1.078				
	(0.029)				
δ_1	-0.083				
	(0.010)				
δ_2	0.129				
	(0.010)				
σ^2	0.325				
	(0.010)				
ξ	-1.07				
	(0.130)				
η					1.143
					(0.061)
$\pi^{\mathbb{P}}$	0.991	0.990	0.993	0.990	0.986
ℓ_r	12864	12662	12482	12610	12693
		[12.607]	[18.226]	[12.466]	[6.311]
$\ell_{\mathtt{vix}}$	-8812	-9363	-9509	-10145	-10079
		[11.022]	[11.513]	[12.094]	[11.720]
$\ell_{r,\mathtt{vix}}$	4052	3299	2973	2466	2614
		[14.057]	[16.390]	[13.865]	[12.548]
$\ell_{r,x,\mathtt{vix}}$	823				

Note: Parameter estimates are reported with QMLE robust standard errors in parenthesis. The persistence parameter $\pi^{\mathbb{P}} = \beta$ is in the RG and EG models, $\pi^{\mathbb{P}} = \alpha + \beta$ in the model, and $\pi^{\mathbb{P}} = \beta + \alpha\delta^2$ in the two Heston-Nandi models. For the $HN_{\mathtt{vd}}$ model we report , $\eta = (1 + 2\alpha\xi)^{-1}$ (which is the variance risk ratio h_t^*/h_t) instead of ξ. (The implied value for ξ is here -55,768.95). The values reported in square brackets under a log-likelihood term is the corresponding Voung statistic relative to RG.

be attributed to the leverage effect and the equity premium.[14] This empirical finding supports the view in Hao and Zhang (2013) who argued that equity risk cannot justify the observed market VRP.

The value of the maximized log-likelihood function is a measure of the model's ability to fit the empirical distribution of the observed data. The Realized GARCH model with the affine exponential SDF clearly has the best fit for all terms of the log-likelihood that are directly comparable. Both ℓ_r and ℓ_{vix} and their sum $\ell_{r,\text{vix}}$, are much larger for the Realized GARCH model than any of the other models. This is confirmed by the Voung statistics for non-nested model comparisons. This is despite the fact that the other models seek to maximize $\ell_{r,\text{vix}}$ while the objective of the Realized GARCH model entails a tradeoff between this term and the log-likelihood for the realized measures, ℓ_x. Following the Realized GARCH model, the $\text{HN}_{\texttt{vd}}$ has the second best performance in terms of describing returns, ℓ_r, whereas the EGARCH takes the second spot in terms of explaining the variation in the VIX, ℓ_{vix}. Below, we evaluate the model's ability to describe the VIX in greater details.

5.4 Model Performance's for VIX, VRP, and Volatility

In this section, we focus on the models' ability to explain the variation in the VIX, VRP, and the volatility of cumulative returns. First, we report summary statistics for the full sample, then we report results for various subsamples – in-sample results as well as out-of-sample results. Most of the existing literature has focused on a single variable. For instance, the focus in Hao and Zhang (2013), Christoffersen et al. (2014), and Majewski et al. (2015) was VIX and derivative pricing, whereas Wang et al. (2017) focused on volatility under the physical measure.

5.4.1 Comparison of market and model-based VIX

In this section, we evaluate the model's ability to describe the VIX in greater details beyond the log-likelihood term, ℓ_{vix}, listed above. Table 3 reports a range of summary statistics based on the full sample, where we compare the model-based measures of VIX with the observed VIX.

In this comparison, the Realized GARCH model is also consistently the best model. It has the smallest bias, the smallest mean squared error, and the smallest mean absolute error. The models: EGARCH, GARCH, and Heston-Nandi with LRNVR tend to underestimate the VIX, whereas the Heston-Nandi GARCH with the variance dependent SDF tends to overestimate the VIX. The Realized GARCH model also has the highest correlation between the model-implied VIX and the market-based

[14] This finding is specific to the RG model structure, that only includes a short-term leverage effect. So it is possible that that models with a more sophisticated leverage effect and/or long memory feature, would result in different weights on the two terms.

Table 3: VIX pricing performance (full sample)

Model	RG	EG	G	HN	HN_{vd}	VIX
Bias	-0.048	-0.203	-0.318	-0.203	0.164	
MAE	1.866	2.222	2.199	2.439	2.419	
RMSE	2.504	2.898	3.012	3.565	3.504	
Corr	0.959	0.945	0.941	0.916	0.919	
AR1	0.994	0.994	0.996	0.993	0.992	0.980
AR10	0.901	0.935	0.943	0.933	0.926	0.898
AR22	0.775	0.836	0.841	0.854	0.836	0.810
Mean	18.362	18.207	18.093	18.207	18.574	18.410
Var	73.173	71.775	71.594	59.951	58.190	77.652
Skew	2.560	2.560	3.556	1.714	1.779	2.657
Kurt	12.301	12.282	18.909	6.580	6.862	12.662

Note: Summary statistics for the VIX errors, $e_t = \text{VIX}_t^{\text{model}} - \text{VIX}_t^{\text{market}}$. We report the sample average of e_t (Bias), the mean absolute errors (MAE), the root of mean squared errors (RMSE), the sample correlation between $\text{VIX}_t^{\text{model}}$ and $\text{VIX}_t^{\text{market}}$ (Corr), and the sample autocorrelations of e_t for lags 1, 10, and 22, that are denoted AR1, AR10, and AR22, respectively. For $\text{VIX}_t^{\text{model}}$ we report its sample average (Mean), its sample variance (Var), its sample skewness, (Skew), and its sample excess kurtosis (Kurt).

VIX. With the Realized GARCH model, the resulting statistical properties of the model-based VIX are closer to those of the market-based VIX, that those of other models.

5.4.2 In-Sample Comparison of the VRP and and Its Components

Table 4 provides the in-sample pricing performance for the variance risk premium and its two components: the volatility index (VIX) and the annualized model-based volatility, and we report the bias for each of the models.[15]

In terms of the volatility risk premium, the Realized GARCH model provides the smallest bias, the smallest root mean square error (RMSE), and the smallest mean absolute error (MAE). The reduction in pricing errors relative to other models ranges from 15.0% to 30.6% in terms of RMSE and 27.1% to 44.7% in terms of MAE. Among the competing models, the Heston-Nandi GARCH model with variance dependent SDF appears to be the best alternative. In contrast, the Heston-Nandi model with LRNVR, which is arguably a very popular option pricing model, does not fair well in terms of explaining the volatility risk premium. The EGARCH model performs significantly better than other GARCH models using LRNVR, especially in terms of the RMSE.

We observe very similar patterns across models in terms of their ability to price the VIX. The

[15]The annualized volatility is calculated based on the martingale process assumption made by Bollerslev et al. (2009) and our results are robust when annualized volatility is calculated based on the forecast value using HAR model (method used in Bekaert and Hoerova (2014) etc.). See section B.1 for details.

Table 4: In-sample Statistics of Model Performance

Model	RG	EG	G	HN	$HN_{\texttt{vd}}$
Volatility Risk Premium					
Bias	- 0.497 (10.3%)	- 3.827 (79.1%)	- 4.397 (90.9%)	- 4.350 (89.9%)	- 3.053 (63.1%)
RMSE	3.825	5.078	5.509	5.456	4.499
△%		24.7%	30.6%	29.9%	15.0%
MAE	2.635	4.316	4.766	4.720	3.612
△%		39.0%	44.7%	44.2%	27.1%
Volatility Index (VIX)					
Bias	- 0.048 (0.26%)	- 0.203 (1.10%)	- 0.318 (1.73%)	- 0.203 (1.10%)	- 0.164 (0.89%)
RMSE	2.504	2.898	3.012	3.565	3.504
△%		13.6%	16.9%	29.8%	28.5%
MAE	1.866	2.222	2.199	2.439	2.419
△%		16.0%	15.1%	23.5%	22.9%
Annualized Volatility					
Bias	0.449 (3.30%)	3.624 (26.7%)	4.080 (30.1%)	4.146 (30.5%)	3.217 (23.7%)
RMSE	2.943	4.380	4.658	5.666	4.920
△%		32.8%	36.8%	48.1%	40.2%
MAE	2.002	3.898	4.285	4.839	3.979
△%		48.6%	53.3%	58.6%	49.7%

Note: Bias denotes the difference between the model-implied quantity and the market-based quantity. The absolute bias in percent of average VRP, VIX and annualized volatility is given in parentheses. The rows indicated with "△%" present the increase in RMSE and MAE for all models relative to the RG. The market-based VRP in this table is defined by $\text{VRP}_t^{\text{market}} = \text{VIX}_t - \sqrt{\frac{252}{22}\sum_{i=1}^{22} \text{RVcc}_{t-i+1}} \times 100$.

Realized GARCH delivers the best performance while the Heston-Nandi GARCH takes last place. In fact, the non-affine models (RG, EG and G) perform substantially better that the two affine models (HN and $HN_{\texttt{vd}}$), which is consistent with the existing literature on option pricing with GARCH models. The main advantage of the affine models is their analytical expressions for the moment generating function. Fortunately, these are not needed for VIX pricing, so the non-affine model are clearly preferred for this problem.[16] The performance gain for the Realized GARCH model ranges from 13.6% to 29.8% in terms of RMSE and between 15.1% and 23.5% in terms of MAE.

All models tend to over-estimate the expected volatility under the physical measure. However, the bias is much smaller for the Realized GARCH model. This indicates that the other models, in order to

[16] It is very difficult, if not impossible, to obtain an analytical moment generation function for cumulative returns or the k-step ahead conditional volatility for these non-affine models. Both are needed for quasi-analytical pricing formula for derivatives using a Fourier inverse transformation. For this reason, computationally intensive simulation methods and analytical expansions are commonly used for pricing derivatives with non-affine models, see Huang et al. (2017) and Tong and Huang (2021) for the use of an Edgeworth expansion to price SPX and VIX options with the Realized GARCH model.

price the VIX, inadvertently increase the level of volatility to compensate for their shortcomings in risk neutralization. The RG and $HN_{\mathtt{vd}}$ both have additional parameter to compensate for volatility risk, which likely explain their smaller bias. In terms of explaining the annualized volatility, the Realized GARCH model reduces the RMSE by 32.8% to 48.1% and the MAE is reduced by 48.6% to 58.6%.

It is worth emphasizing that the parameter estimation does not target the volatility risk premium directly, the superior performance of Realized GARCH highlights the model's ability of reconcile the physical and risk-neutral dynamics within a single model framework. This is some accomplishment by the Realized GARCH framework, because this was considered to be a very difficult empirical problem, see Bates (1996).

5.4.3 Out-of-sample pricing performance

The proposed pricing model, which is based on the Realized GARCH model and the affine exponential SDF, requires a larger number of parameters to be estimated than the methods based on the conventional GARCH models. The larger number of parameters could entail some overfitting of the model, and this might explain some of the observed empirical improvements. It is therefore important to document that the model also provides improvements out-of-sample. In this section, we compare the models in terms of their out-of-sample pricing errors using a rolling estimation window, based on the past 750 daily observations. The first forecast is made for the first month (22 trading days) of 2007, and this forecasts is based on parameters that were estimated with the previous 750 daily observations (January 6th 2004 to December 29th, 2006). We report out-of-sample pricing errors for 2007-2018 and two sub-sample periods, 2007-2012 and 2013-2018. Splitting the out-of-sample period into two subsamples is interesting because some results could potentially be specific to the global financial crisis, which had high volatility and high volatility-of-volatility. The global financial crisis is contained in the first subsample.

We report a range of performance statistics for each of the models and each of the sample periods in Table 5. The significance of relative performance is evaluated with Diebold-Mariano (DM) statistics where we compare each of the alternative models to the Realized GARCH model. For this purpose, we first compute the tracking errors for the i-th model,

$$e_{it} = X_{i,t}^{\mathtt{model}} - X_t^{\mathtt{market}},$$

where X represents the volatility risk premium, the volatility index, or the annualized volatility. These errors are translated into losses using either the mean square error, $g(e_{it}) = e_{it}^2$ or the mean absolute error $g(e_{it}) = |e_{it}|$. The loss of model i, relative to the Realized GARCH model ($i = 0$) is defined by

$$d_{i,t} = g(e_{it}) - g(e_{0t}),$$

and we proceed to tests the hypothesis, $H_0 : \mathbb{E}(d_{i,t}) = 0$ using the Diebold-Mariano (DM) statistic, $\mathrm{DM}_i = \sqrt{T}\bar{d}_{i,\cdot}/\hat{\sigma}_{d_i}$, where $\bar{d}_{i,\cdot} = \frac{1}{T}\sum_t d_{i,t}$ and $\hat{\sigma}_{d_i}^2$ is an estimate of the long-run variance of $\{d_{i,t}\}$. Our estimates of $\sigma_{d_i}^2$, are based on the Parzen kernel with bandwidth $H = 42$. Under suitable regularity conditions, it can be shown that $\mathrm{DM}_i \stackrel{d}{\to} N(0, 1)$ under the null hypothesis $\mathbb{E}(d_{i,t}) = 0$, see Diebold and Mariano (1995), and the 10% and 5% critical values are therefore given by 1.64 and 1.96, respectively.

Once again the Realized GARCH provides the best out-of-sample pricing performance for the VRP, the VIX, and the annualized volatility, and this is found in all three sample periods. The model also provides the smallest bias in most cases and the improvement in RMSE/MAE ranges from 10% to 40% in most cases, which is significant in most cases. This shows that superior in-sample performance of the Realized GARCH model cannot be attributed to overfitting. The RMSE and MAE are, as expected, larger in the subsample with the financial crisis. The Realized GARCH model really stands out in terms of explaining the volatility under the physical measure (Annualized Volatility), where the reduction in out-of-sample loss is always larger than 20% and as larger as 52.2%.

The HN is always the worst model for tracking the VIX out-of-sample, as was the case in-sample. The picture is similar for the annualized volatility, albeit $\mathrm{HN}_{\mathtt{vd}}$ is the "best of the rest" in terms of tracking of annualized volatility in the post crisis period. This might be explained by the variance dependent SDF being less misspecified when volatility is low, while it cannot generate enough discrepancy between $\mathbb{P}$ and $\mathbb{Q}$ when volatility is high. Interestingly, even though the $\mathrm{HN}_{\mathtt{vd}}$ is clearly inferior to both the GARCH and EGARCH models in terms of forecasting volatility under $\mathbb{P}$ and $\mathbb{Q}$, it is actually better than these two models in terms of forecasting the VRP.

6 Robustness Check: The Case with Non-Gaussian Innovations.

A key assumption behind the simple and tractable expression for the SDF in (4) is the assumed normality of z_t and u_t. Normality is commonly assumed in this literature, see e.g. Hao and Zhang (2013) who derived VIX pricing for a range of GARCH models. This distributional assumptions is clearly rejected

Table 5: Out-of-sample Statistics of Model Performance

	Volatility Risk Premium					Volatility Index (VIX)					Annualized Volatility				
Full: 2007.01.03 - 2018.12.30															
Model	RG	EG	G	HN	HN_{vd}	RG	EG	G	HN	HN_{vd}	RG	EG	G	HN	HN_{vd}
Bias	- 0.688 (13.8%)	- 3.824 (76.8%)	- 3.762 (75.5%)	- 3.814 (76.6%)	- 2.753 (55.3%)	0.155 (0.79%)	- 0.212 (1.08%)	0.063 (0.32%)	- 0.533 (2.72%)	- 0.132 (0.67%)	0.844 (5.78%)	3.612 (24.7%)	3.825 (26.2%)	3.282 (22.5%)	2.885 (19.8%)
RMSE	3.975	5.420	5.378	5.387	4.668	2.870	3.259	3.278	4.298	3.903	3.271	5.154	4.961	5.856	5.360
DM stat.		8.266	8.064	8.386	4.853		3.976	3.579	2.659	1.887		6.609	5.400	5.314	4.066
△%		26.7%	26.1%	26.2%	14.8%		11.9%	12.4%	33.2%	26.5%		36.5%	34.1%	44.2%	39.0%
MAE	2.921	4.565	4.512	4.543	3.676	2.113	2.402	2.387	2.842	2.615	2.466	4.420	4.422	4.750	4.203
DM stat.		10.179	10.170	10.621	6.250		3.995	3.333	4.222	3.137		9.034	8.894	11.298	8.624
△%		36.0%	35.3%	35.7%	20.5%		12.0%	11.5%	25.6%	19.2%		44.2%	44.2%	48.1%	41.3%
Crisis Period: 2007.01.03 - 2012.12.30															
Model	RG	EG	G	HN	HN_{vd}	RG	EG	G	HN	HN_{vd}	RG	EG	G	HN	HN_{vd}
Bias	- 1.011 (17.8%)	- 4.311 (76.0%)	- 4.439 (78.2%)	- 4.315 (76.0%)	- 3.376 (59.5%)	- 0.294 (1.20%)	- 0.824 (3.38%)	- 0.484 (1.98%)	- 1.427 (5.85%)	- 0.321 (1.32%)	0.717 (3.83%)	3.487 (18.6%)	3.956 (21.1%)	2.888 (15.4%)	3.055 (16.3%)
RMSE	4.570	6.060	6.187	6.014	5.339	3.383	3.972	3.944	5.497	4.866	3.493	5.520	5.158	6.844	6.471
DM stat.		6.012	6.489	5.940	3.756		4.028	3.277	2.515	1.676		4.307	3.018	4.141	3.540
△%		24.6%	26.1%	24.0%	14.4%		14.8%	14.2%	38.5%	30.5%		36.7%	32.3%	49.0%	46.0%
MAE	3.420	5.164	5.261	5.142	4.301	2.557	3.069	2.946	3.798	3.327	2.514	4.505	4.476	5.261	4.992
DM stat.		7.760	8.200	7.997	5.092		4.254	2.820	3.906	2.541		5.559	5.287	8.462	7.740
△%		33.8%	35.0%	33.5%	20.5%		16.7%	13.2%	32.7%	23.1%		44.2%	43.8%	52.2%	49.6%
Post-crisis Period: 2013.01.03-2018.12.30															
Model	RG	EG	G	HN	HN_{vd}	RG	EG	G	HN	HN_{vd}	RG	EG	G	HN	HN_{vd}
Bias	- 0.365 (8.52%)	- 3.336 (77.9%)	- 3.084 (72.0%)	- 3.313 (77.3%)	- 2.129 (49.7%)	0.605 (4.09%)	0.401 (2.72%)	0.611 (4.14%)	0.362 (2.45%)	0.585 (3.96%)	0.970 (9.25%)	3.737 (35.6%)	3.695 (35.2%)	3.675 (35.1%)	2.714 (25.9%)
RMSE	3.271	4.694	4.423	4.676	3.881	2.242	2.339	2.435	2.591	2.603	3.032	4.758	4.755	4.663	3.946
DM stat.		5.903	5.400	6.254	3.188		1.256	1.704	2.879	2.846		6.162	6.981	6.516	4.014
△%		30.3%	26.0%	30.0%	15.7%		4.2%	7.9%	13.5%	13.9%		36.3%	36.2%	35.0%	23.2%
MAE	2.421	3.966	3.763	3.943	3.050	1.670	1.734	1.828	1.884	1.902	2.418	4.335	4.367	4.238	3.414
DM stat.		6.644	6.276	7.005	3.767		0.984	1.821	2.679	2.749		7.965	8.344	8.247	5.231
△%		39.0%	35.7%	38.6%	20.6%		3.7%	8.7%	11.4%	12.2%		44.2%	44.6%	42.9%	29.2%

Note: Bias denotes the average "model error", which is the difference between the model-implied quantity and the market-based quantity. The absolute bias in percent of the average VRP, VIX and annualized volatility is given in parentheses. Rows indicated with "△%" state who much larger the RMSE or MAE is for each alternative models, measured relative to RG. Diebold and Mariano statistics (DM stat.) are computed for the relative MSE or MAE losses, where the standard errors are calculated with the Parzen kernel with $H = 42$ as bandwidth. The market VRP in this table is defined as $\mathrm{VRP}_t^{\mathrm{market}} = \mathrm{VIX}_t - \sqrt{\frac{252}{22}\sum_{i=1}^{22} \mathrm{RVcc}_{t-i+1}} \times 100$.

by the normality test of Bontemps and Meddahi (2005),[17] and it is therefore relevant to investigate if this misspecification severely distorts our expressions for VRP and VIX.

The normality assumption is a fundamental pillar of most expressions derived under $\mathbb{Q}$. A rigorous analysis of the case with non-normality under the RG framework, which leads to tractable results, would be a stupendous feat. So, we will merely probe this problem by amending Theorem 1 to account for non-normality in the computation of the F_i-terms. We preserve other parts of model and directly assume the relationships between $\mathbb{P}$ and $\mathbb{Q}$ through λ (the equity risk premium) and ξ (the variance risk premium), thus bypassing the need for an SDF. Specifically, we assume that $z_t^* = z_t + \lambda$ and $u_t^* = u_t + \xi$ have the same distribution under $\mathbb{Q}$, as z_t and u_t have under $\mathbb{P}$.

The F_i-terms are key for the VIX pricing formula (8), and we explore the effect of non-normality on these using second-order Gram-Charlier expansions[18] to approximate the densities of z_t and u_t under $\mathbb{P}$. For a random variables with zero mean, unit variance, skewness equal to ς, and excess kurtosis given by κ, the approximating density is given by

$$f(\varsigma, \kappa; x) = \left[1 + \frac{\varsigma}{6} H_3(x) + \frac{\kappa}{24} H_4(x)\right] \phi(x),$$

where $H_3(x) = x^3 - 3x$, $H_4(x) = x^4 - 6x^2 + 3$, and $\phi(x)$ is the standard normal density. Thus, for z_t and u_t under $\mathbb{P}$ we have $f(\varsigma_z, \kappa_z; z_t)$ and $f(\varsigma_u, \kappa_u; u_t)$, which are also the approximating densities for z_t^* and u_t^* under $\mathbb{Q}$, i.e. $f(\varsigma_z, \kappa_z; z_t^*)$ and $f(\varsigma_u, \kappa_u; u_t^*)$.

We seek an expression for $F_i = \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(\beta^i (v_{t+k-1-i}^* + \tilde{\omega})\right)\right]$, where $v_t^* = \tilde{\tau}_1 z_t^* + \tau_2 (z_t^{*2} - 1) + \gamma\sigma u_t^*$. To this end, we introduce

$$\Psi(a, b, \varsigma, \kappa) = \int_{-\infty}^{+\infty} \exp(ax + bx^2) f(\varsigma, \kappa; x) \mathrm{d}x,$$

for which we have the following mathematical result.

Lemma 1. *For $b < \frac{1}{2}$ we have*

$$\Psi(a, b, \varsigma, \kappa) = \frac{e^{A_2}}{\sqrt{1-2b}} \left[1 + \varsigma C_3 + \kappa C_4\right],$$

where $C_3 = A_3 + A_1 B_1$ and $C_4 = A_4 + A_2 B_1 + B_2$, with $A_q = \frac{1}{q!}\left(\frac{a}{1-2b}\right)^q$ and $B_q = \frac{1}{q!}\left(\frac{b}{1-2b}\right)^q$.

[17] The Bontemps-Meddahi test statistics for z_t (full sample) are: 40.25 , 36.74, 26.74, 28.11, and 26.64 for RG, EG, G, HN, $\text{HN}_{\texttt{vd}}$, respectively, and for u_t in RG this test statistic is 24.17. These are all highly significant (1% critical value is 9.21).

[18] The Gram-Charlier expansion is commonly used in the literature on derivative pricing, albeit often mislabelled an Edgeworth expansions, see Huang et al. (2017) for details.

By applying Lemma to the expectation of

$$e^{\beta^i(\tilde{\tau}_1 z^* + \tau_2(z^{*2}-1) + \gamma\sigma u^* + \tilde{\omega})} = e^{\beta^i \tilde{\tau}_1 z^* + \beta^i \tau_2 z^{*2}} e^{\beta^i \gamma\sigma u^*} e^{\beta^i(\tilde{\omega}-\tau_2)},$$

with $(a,b) = (\beta^i \tilde{\tau}_1, \beta^i \tau_2)$ for z^* and $(a,b) = (\beta^i \gamma\sigma, 0)$ for u^*, we arrive at the amended expression,

$$F_i = e^{\beta^i(\tilde{\omega}-\tau_2)} \Psi(\beta^i \tilde{\tau}_1, \beta^i \tau_2, \varsigma_z, \kappa_z) \Psi(\beta^i \gamma\sigma, 0, \varsigma_u, \kappa_u),$$

which are used in (8) to compute VIX_t^{RG}.

Table 6: RG-Normal and RG-GramCharlier

Model	RG-Normal	RG-GramCharlier
	Parameters Estimation	
ς_z		-0.348
		(0.038)
κ_z		0.993
		(0.114)
ς_u		0.189
		(0.048)
κ_u		0.748
		(0.113)
ℓ_r	12864	12951
ℓ_x	-3229	-3193
$\ell_{\texttt{vix}}$	-8812	-8805
$\ell_{r,x,\texttt{vix}}$	823	953
	Volatility Risk Premium	
Bias	-0.497	-0.561
RMSE	3.825	3.823
MAE	2.635	2.648
	Volatility Index (VIX)	
Bias	-0.048	-0.049
RMSE	2.504	2.499
MAE	1.866	1.863
	Annualized Volatility	
Bias	0.449	0.511
RMSE	2.943	2.925
MAE	2.002	2.013

Note: Parameter estimates are reported with QMLE robust standard errors in parenthesis.

6.1 Empirical Analysis based on Amended VIX Expression

As in the previous section, we estimate the amended model by maximizing the full quasi log-likelihood function and report the estimates of the new parameters ς_z, κ_z, ς_u, and κ_u in Table 6 along with robust standard errors. Each of the four new parameters are significantly different from zero, while estimates of all parameters are all very similar to those of the original specification, see Table B.6, in Appendix B. The standardized returns, z_t^*, are negatively skewed, $\hat{\varsigma}_z = -0.348$, whereas the volatility shocks, u_t^*, are positively skewed, $\hat{\varsigma}_u = 0.189$. The excess kurtosis is positive for both distributions with is associated with heavy tail distributions. The gain in the log-likelihood from adding the four parameters is substantially at about 130 units, which corresponds to a likelihood ratio statistic of about 260. However most of the improvement stems from gains in the log-likelihood function for returns and the realized measure, ℓ_r and ℓ_x, whereas the improvement in the log-likelihood term for VIX pricing, $\ell_{\texttt{vix}}$, is just 7. We also report the Bias, RMSE and MAE for the VRP, the VIX, and annualized volatility for the full sample in Table 6. The gains are quite small in all cases. For the VIX the RMSE is reduced by just 0.2%, which is a minuscule improvement compared to the gains we reported for the original specification in Section 5. Therefore, amending the VIX pricing formula to the non-Gaussian case may not be worthwhile for the purpose of VIX pricing. This conclusion is solely based on amending the VIX pricing formula to permit non-normality with a 2nd order Gram-Charlier expansion. The VIX formula only requires a good approximation of the conditional expectation of future volatility under $\mathbb{Q}$, and it appears that the model based on conditional expectation of h_{t+h}^*, is relatively insensitive to deviations from Gaussianity. However, it is important to note that even if VIX pricing is fairly insensitive to distributional misspecification, this may not be true for other applications, such as option pricing. Option pricing relies heavily on the distribution of cumulative returns under $\mathbb{Q}$, and distributional misspecification could therefore be very important for option pricing, which we have not explored in this paper.

7 Conclusion

We have developed a Realized GARCH model for the simultaneous modeling of returns, the VIX, and the VRP, using an exponentially affine SDF that takes advantage of the dual-shock structure in the Realized GARCH model. This framework has several attractive features. First, its dual-shock structure lead to a distinct compensation for volatility risk, which is empirically important. Second, it takes advantage of the information contained in realized measures of volatility. Third, it has a flexible leverage function that captures the empirically important return-volatility dependence in a parsimonious manner. Fourth, the model combined with the exponentially affine SDF, conveniently, yields analytical formulae for the VIX and the volatility risk premium. Fifth, the model is an observation-driven model, which makes estimation straight forward. Sixth, its dynamic properties under the physical and risk-neutral measures offer intuitive and theory-consistent explanations for the excellent empirical performance offered by the Realized GARCH model.

References

Ammann, M. and R. Buesser (2013): "Variance risk premiums in foreign exchange markets," *Journal of Empirical Finance*, 23, 16 – 32.

Andersen, T., D. Dobrev, and E. Schaumburg (2012): "Jump-robust volatility estimation using nearest neighbor truncation," *Journal of Econometrics*, 169, 75–93.

Andersen, T. G. and T. Bollerslev (1998): "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts," *International Economic Review*, 39, 885–905.

Andersen, T. G., N. Fusari, and V. Todorov (2019): "The Pricing of Tail Risk and the Equity Premium: Evidence From International Option Markets," *Journal of Business & Economic Statistics*.

Bakshi, G. and N. Kapadia (2003): "Delta-Hedged Gains and the Negative Market Volatility Risk Premium," *Review of Financial Studies*, 16, 527–566.

Bardgett, C., E. Gourier, and M. Leippold (2019): "Inferring volatility dynamics and risk premia from the S&P 500 and VIX markets," *Journal of Financial Economics*, 131, 593 – 618.

Barndorff-Nielsen, O. E. (2004): "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2, 1–37.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008): "Designing Realized Kernels to Measure the ex post Variation of Equity Prices in the Presence of Noise," *Econometrica*, 76, 1481–1536.

Bates, D. S. (1996): "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options," *Review of Financial Studies*, 9, 69–107.

Bekaert, G. and M. Hoerova (2014): "The VIX, the Variance Premium and Stock Market Volatility," *Journal of Econometrics*, 183, 181–192.

Bollerslev, T. (1986): "Generalized autoregressive heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

Bollerslev, T., M. Gibson, and H. Zhou (2011): "Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities," *Journal of Econometrics*, 160, 235–245.

Bollerslev, T., G. Tauchen, and H. Zhou (2009): "Expected Stock Returns and Variance Risk Premia," *The Review of Financial Studies*, 22, 4463–4492.

Bontemps, C. and N. Meddahi (2005): "Testing normality: a GMM approach," *Journal of Econometrics*, 124, 149–186.

Britten-Jones, M. and A. Neuberger (2000): "Option Prices, Implied Price Processes, and Stochastic Volatility," *The Journal of Finance*, 55, 839–866.

Byun, S. J., B. H. Jeon, B. Min, and S.-J. Yoon (2015): "The Role of the Variance Premium in Jump-GARCH Option Pricing Models," *Journal of Banking & Finance*, 59, 38–56.

Carr, P. and D. Madan (1998): *Towards A Theory of Volatility Trading*, Risk Books, chap. 29 (in Volatility: New Estimation Techniques for Pricing Derivatives), 417–427.

Carr, P. and L. Wu (2009): "Variance Risk Premiums," *Review of Financial Studies*, 22, 1311–1341.

Christoffersen, P., B. Feunou, K. Jacobs, and N. Meddahi (2014): "The Economic Value of Realized Volatility: Using High-Frequency Returns for Option Valuation," *Journal of Financial and Quantitative Analysis*, 49, 663–697.

Christoffersen, P., S. L. Heston, and K. Jacobs (2013): "Capturing Option Anomalies with a Variance-Dependent Pricing Kernel," *Review of Financial Studies*, 26(8), 1963–2006.

Conrad, C. and K. Loch (2015): "The Variance Risk Premium and Fundamental Uncertainty," *Economics Letters*, 132, 56–60.

Corsi, F. (2009): "A Simple Approximate Long-Memory Model of Realized Volatility," *Journal of Financial Econometrics*, 7, 174–196.

CORSI, F., N. FUSARI, AND D. L. VECCHIA (2013): "Realizing Smiles: Pricing Options with Realized Volatility," *Journal of Financial Economics*, 107, 284–304.

COVAL, J. AND T. SHUMWAY (2001): "Expected Option Returns," *Journal of Finance*, 56, 983–1009.

CREMERS, M., M. HALLING, AND D. WEINBAUM (2015): "Aggregate Jump and Volatility Risk in the Cross-Section of Stock Returns," *The Journal of Finance*, 70, 577–614.

DEMETERFI, K., E. DERMAN, M. KAMAL, AND J. ZOU (1999): "A Guide to Volatility and Variance Swaps," *The Journal of Derivatives*, 6, 9–32.

DIEBOLD, F. X. AND R. S. MARIANO (1995): "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13, 253–263.

DUAN, J., G. GAUTHIER, AND J. SIMONATO (1999): "An Analytical Approximation for the GARCH Option Pricing Model," *Journal of Computational Finance*, 2, 75–116.

DUAN, J.-C. (1995): "The GARCH option pricing model," *Mathematical finance*, 5, 13–32.

DUAN, J.-C. AND C.-Y. YEH (2010): "Jump and Volatility Risk Premiums Implied by VIX," *Journal of Economic Dynamics and Control*, 34, 2232–2244.

ERAKER, B. (2004): "Do Stock Prices and Voltility Jump? Reconciling Evidence from Spot and Option Prices," *The Journal of Finance*, 59, 1367–1403.

FRENCH, K. R., G. SCHWERT, AND R. F. STAMBAUGH (1987): "Expected stock returns and volatility," *Journal of Financial Economics*, 19, 3 – 29.

GONZALEZ-PEREZ, M. T. (2015): "Model-free volatility indexes in the financial literature: A review," *International Review of Economics & Finance*, 40, 141–159.

HANSEN, P. R. AND Z. HUANG (2016): "Exponential GARCH Modeling with Realized Measures of Volatility," *Journal of Business & Economic Statistics*, 34, 269–287.

HANSEN, P. R., Z. HUANG, AND H. SHEK (2012): "Realized GARCH: A Joint Model of Returns and Realized Measures of Volatility," *Journal of Applied Econometrics*, 27, 877–906.

HANSEN, P. R. AND A. LUNDE (2005a): "A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?" *Journal of Applied Econometrics*, 20, 873–889.

——— (2005b): "A Realized Variance for the Whole Day Based on Intermittent High-Frequency Data," *Journal of Financial Econometrics*, 3, 525–554.

Hao, J. and J. E. Zhang (2013): "GARCH option pricing models, the CBOE VIX, and variance risk premium," *Journal of Financial Econometrics*, 11, 556–580.

Heston, S. L. and S. Nandi (2000): "A Closed-Form GARCH Option Valuation Model," *Review of Financial Studies*, 13, 585–625.

Huang, Z., T. Wang, and P. R. Hansen (2017): "Option Pricing with the Realized GARCH Model: An Analytical Approximation Approach," *Journal of Futures Markets*, 37, 313–428.

Jensen, M. B. and A. Lunde (2001): "The NIG-S&ARCH model: a fat-tailed, stochastic, and autoregressive conditional heteroskedastic volatility model," *The Econometrics Journal*, 4, 319–342.

Jiang, G. and Y. Tian (2005): "The Model-Free Implied Volatility and Its Information Content," *The Review of Financial Studies*, 18, 1305–1342.

Kim, S., N. Shephard, and S. Chib (1998): "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economic Studies*, 65, 361–93.

Liu, L. Y., A. J. Patton, and K. Sheppard (2015): "Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes," *Journal of Econometrics*, 187, 293–311.

Majewski, A. A., G. Bormetti, and F. Corsi (2015): "Smile from the Past: A General Option Pricing Framework with Multiple Volatility and Leverage Components," *Journal of Econometrics*, 187, 521–531.

Nelson, D. B. (1991): "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59, 347–370.

Ornthanalai, C. (2014): "Levy jump risk: Evidence from options and returns," *Journal of Financial Economics*, 112, 69–90.

Pan, J. (2002): "The jump-risk premia implicit in options: evidence from an integrated time-series study," *Journal of Financial Economics*, 63, 3 – 50.

Santa-Clara, P. and S. Yan (2010): "Crashes, Volatility, and the Equity Premium: Lessons from S&P 500 Options," *The Review of Economics and Statistics*, 92, 435–451.

Song, Z. and D. Xiu (2016): "A Tale of Two Option Markets: Pricing Kernels and Volatility Risk," *Journal of Econometrics*, 190, 176–196.

Taylor, S. (1986): *Modelling Financial Time Series.*, Chichester: John Wiley.

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

Wang, T., Y. Shen, Y. Jiang, and Z. Huang (2017): "Pricing the CBOE VIX Futures with the Heston-Nandi GARCH Model," *Journal of Futures Markets*, 37, 641–659.

Wu, L. (2011): "Variance dynamics: Joint evidence from options and high-frequency returns," *Journal of Econometrics*, 160, 280–287.

A Appendix of Proofs

Lemma A.1. *Suppose*

$$M_{t+1} = \frac{\exp(-\lambda_t z_{t+1} - \xi_t u_{t+1})}{\mathbb{E}\exp(-\lambda_t z_{t+1} - \xi_t u_{t+1})} = \exp\left\{-\lambda_t z_{t+1} - \xi_t u_{t+1} - \tfrac{1}{2}(\lambda_t^2 + \xi_t^2)\right\},$$

then by non-arbitrage we have $\lambda_t = \lambda$.

Proof. The non-arbitrage condition is $\mathbb{E}_t^{\mathbb{Q}}(\exp(r_{t+1})) = \exp(r)$ and the result follows by

$$\begin{aligned} \mathbb{E}_t^{\mathbb{Q}}(\exp(r_{t+1})) &= \mathbb{E}_t(M_{t+1}\exp(r_{t+1})) \\ &= \mathbb{E}_t \exp\left\{(\sqrt{h_{t+1}} - \lambda_t)z_{t+1} - \xi_t u_{t+1} - \tfrac{1}{2}(\lambda_t^2 + \xi_t^2) + r + \lambda\sqrt{h_{t+1}} - \tfrac{1}{2}h_{t+1}\right\} \\ &= \mathbb{E}_t \exp\left\{r + (\lambda - \lambda_t)\sqrt{h_{t+1}}\right\}. \end{aligned}$$

In order to have the above equation hold for all h_t, we need $\lambda_t = \lambda$. □

Lemma A.2. *The Realized GARCH model defined by (1)-(3) and the affine exponential SDF, implied the model (5)-(7) under the risk-neutral measure, $\mathbb{Q}$.*

Proof. Substituting $(z_t^* - \lambda, u_t^* - \xi)$ for (z_t, u_t) immediately yields

$$\begin{aligned} \log h_{t+1} &= \omega + \beta \log h_t + \tau_1(z_t^* - \lambda) + \tau_2[(z_t^* - \lambda)^2 - 1] + \gamma\sigma(u_t^* - \xi) \\ &= (\omega - \tau_1\lambda + \tau_2\lambda^2 - \gamma\sigma\xi) + \beta\log h_t + (\tau_1 - 2\tau_2\lambda)z_t^* + \tau_2(z_t^{*2} - 1) + \gamma\sigma u_t^*, \\ \log x_t &= \kappa + \phi \log h_t + \delta_1(z_t^* - \lambda) + \delta_2[(z_t^* - \lambda)^2 - 1] + \sigma(u_t^* - \xi) \\ &= (\kappa - \delta_1\lambda + \delta_2\lambda^2 - \sigma\xi) + \phi\log h_t + (\delta_1 - 2\delta_2\lambda)z_t^* + \delta_2(z_t^{*2} - 1) + \sigma u_t^*. \end{aligned}$$

□

Lemma A.3. *Suppose that $X \sim N(0,1)$ then for $b < \frac{1}{2}$ $\mathbb{E}\exp\{aX + bX^2\} = \frac{1}{\sqrt{1-2b}}\exp\{\frac{a^2/2}{1-2b}\}$*

Proof. We have

$$e^{ax+bx^2}e^{-\frac{x^2}{2}} = e^{ax - \frac{x^2}{2}(1-2b)} = e^{\frac{a}{\sqrt{1-2b}}u - \frac{u^2}{2}},$$

where $u = \sqrt{1-2b}x$. So integration by substitution yields

$$\begin{aligned} \mathbb{E}e^{aX+bX^2} &= \int \tfrac{1}{\sqrt{2\pi}} e^{ax+bx^2} e^{-\frac{x^2}{2}} \mathrm{d}x = \int \tfrac{1}{\sqrt{2\pi}} e^{\frac{a}{\sqrt{1-2b}}u - \frac{u^2}{2}} \tfrac{1}{\sqrt{1-2b}} \mathrm{d}u \\ &= \tfrac{1}{\sqrt{1-2b}} \mathrm{E} e^{\frac{a}{\sqrt{1-2b}}U} = \tfrac{1}{\sqrt{1-2b}} e^{\frac{1}{2}\frac{a^2}{1-2b}}. \end{aligned}$$

□

Proof of Theorem 1. From the risk-neutral dynamics, we have $\log h_{t+1} = \tilde{\omega} + \beta \log h_t + v_t^*$ where

$$v_t^* \quad = \quad \tilde{\tau}_1 z_t^* + \tau_2 (z_t^{*2} - 1) + \gamma \sigma u_t^*,$$

so that $\log h_{t+k} = \beta^{k-1} \log h_{t+1} + \sum_{i=0}^{k-2} \beta^i (v_{t+k-1-i}^* + \tilde{\omega})$. It follows that

$$\mathbb{E}_t^Q[h_{t+k}] = \mathbb{E}_t^Q[\exp\{\beta^{k-1} \log h_{t+1} + \sum_{i=0}^{k-2} \beta^i (v_{t+k-1-i}^* + \tilde{\omega})\}] = h_{t+1}^{\beta^{k-1}} \prod_{i=0}^{k-2} F_i,$$

where $F_i = \mathbb{E}_t^{\mathbb{Q}} \left[e^{\beta^i (v_{t+k-1-i}^* + \tilde{\omega})} \right]$. Using the expression for v_t^*, and applying Lemma A.3, we have,

$$\begin{aligned}
F_i &= \mathbb{E}_t^{\mathbb{Q}} \left[\exp\left\{ \beta^i (\tilde{\omega} - \tau_2) + \beta^i \tilde{\tau}_1 z + \beta^i \tau_2 z^2 + \beta^i \gamma \sigma u \right\} \right] \\
&= \exp\left\{ \beta^i (\tilde{\omega} - \tau_2) \right\} \mathbb{E}_t^{\mathbb{Q}} \left[\exp\left\{ \beta^i \tilde{\tau}_1 z + \beta^i \tau_2 z^2 \right\} \right] \mathbb{E}_t^{\mathbb{Q}} \left[\exp\left\{ \beta^i \gamma \sigma u \right\} \right] \\
&= \exp\left\{ \beta^i (\tilde{\omega} - \tau_2) \right\} (1 - 2\beta^i \tau_2)^{-1/2} \exp\{ \tfrac{1}{2} \tfrac{\beta^{2i} \tilde{\tau}_1^2}{1 - 2\beta^i \tau_2} \} \exp\left\{ \tfrac{1}{2} \beta^{2i} \gamma^2 \sigma^2 \right\} \\
&= (1 - 2\beta^i \tau_2)^{-1/2} \exp\left\{ \beta^i (\tilde{\omega} - \tau_2) + \tfrac{1}{2} \beta^{2i} [\tfrac{\tilde{\tau}_1^2}{1 - 2\beta^i \tau_2} + \gamma^2 \sigma^2] \right\},
\end{aligned}$$

where we suppressed the superscripts and subscripts on z and u to simplify the exposition. □

Proof of Lemma 1. Define $A_q = \frac{1}{q!} \left(\frac{a}{1-2b} \right)^q$ and $B_q = \frac{1}{q!} \left(\frac{b}{1-2b} \right)^q$, then, from

$$\begin{aligned}
ax + bx^2 - \frac{x^2}{2} &= ax - \tfrac{1}{2}(1-2b)x^2 = -\tfrac{1}{2}(1-2b)(x - \frac{a}{1-2b})^2 + \tfrac{1}{2} \left(\frac{a}{1-2b} \right)^2 \\
&= -\tfrac{1}{2} \frac{\left(x - \frac{a}{1-2b} \right)^2}{1/(1-2b)} + A_2,
\end{aligned}$$

we have

$$\begin{aligned}
\int e^{ax+bx^2} f(\varsigma, \kappa; x) \mathrm{d}x &= \int e^{ax+bx^2} \left(1 + \tfrac{\varsigma}{6} H_3(x) + \tfrac{\kappa}{24} H_4(x) \right) \tfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \mathrm{d}x \\
&= e^{A_2} \int \left(1 + \tfrac{\varsigma}{6} H_3(x) + \tfrac{\kappa}{24} H_4(x) \right) \tfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x - \frac{a}{1-2b}}{1/\sqrt{1-2b}} \right)^2} \mathrm{d}x \\
&= \frac{e^{A_2}}{\sqrt{1-2b}} \int \left(1 + \tfrac{\varsigma}{6} H_3(x) + \tfrac{\kappa}{24} H_4(x) \right) \frac{1}{\sqrt{2\pi \frac{1}{1-2b}}} e^{-\frac{1}{2} \left(\frac{x - \frac{a}{1-2b}}{1/\sqrt{1-2b}} \right)^2} \mathrm{d}x.
\end{aligned}$$

Evaluating this integral simple amounts to computing higher-order moments of a normally distributed random variable, $X \sim N(\mu, \sigma^2)$, with mean $\mu = \frac{a}{1-2b}$ and variance $\sigma^2 = \frac{1}{1-2b}$. For $H_3(x) = x^3 - 3x$, we have

$$
\begin{aligned}
\mathbb{E}[H_3(X)] &= \mu^3 + 3\mu\sigma^2 - 3\mu = \mu^3 + 3\mu(\sigma^2 - 1) \\
&= \left(\tfrac{a}{1-2b}\right)^3 + 3\tfrac{a}{1-2b}\tfrac{2b}{1-2b} = 6(A_3 + A_1B_1) = 6C_3,
\end{aligned}
$$

where we used that $\sigma^2 - 1 = 2\frac{b}{1-2b}$. For $H_4(x) = x^4 - 6x^2 + 3$ we have

$$
\begin{aligned}
\mathbb{E}[H_4(X)] &= \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4 - 6(\mu^2 + \sigma^2) + 3 = \mu^4 + 6\mu^2(\sigma^2 - 1) + 3(\sigma^2 - 1)^2 \\
&= \left(\tfrac{a}{1-2b}\right)^4 + 6\left(\tfrac{a}{1-2b}\right)^2\tfrac{2b}{1-2b} + 3\left(\tfrac{2b}{1-2b}\right)^2 = 4!A_4 + 12\cdot 2!A_2B_1 + 12\cdot 2!B_2 = 24C_4
\end{aligned}
$$

Combining these terms yields $\Psi(a, b, \varsigma, \kappa) = \frac{e^{A_2}}{\sqrt{1-2b}}\left[1 + \varsigma C_3 + \kappa C_4\right]$, which completes the proof.$\square$

B Additional Empirical Results: Robustness Checks

This appendix provides additional empirical results. These results primarily serve to demonstrate that the main empirical results are robust to a variety of alternative approaches to the empirical analysis, such as alternative measures of the market VRP, alternative objective functions, and different choice for the realized measures.

B.1 Robustness Check: HAR type VRP results

The volatility risk premium used in the main body of the paper used the realized volatility over the past months as the predictor of volatility for the following month, which implicitly use a martingale assumption for volatility. Here we consider an alternative regression-based approach, which is known as the HAR model, see Corsi (2009).[19] With this approach the monthly volatility is regressed on lagged daily, weekly, and monthly volatility,

$$
\mathrm{RVcc}_{t+1:t+22} = \beta_0 + \beta_d \mathrm{RVcc}_t + \beta_w \mathrm{RVcc}_{t-4:t} + \beta_m \mathrm{RVcc}_{t-21:t} + \epsilon_{t+1:t+22},
$$

where $\mathrm{RVcc}_{a:b} = \frac{1}{b-a+1}\sum_{t=a}^{b}\mathrm{RVcc}_t$. We also estimate this model using a rolling window with 750 observations (and lagged an additional 22 days to avoid look-ahead bias in our estimates).[20] Table B.1 and table B.2 provide in-sample and out-of-sample results, and all main results are identical to

[19] For other ways other ways to construct the expected volatility, see Bekaert and Hoerova (2014)

[20] For time t the is estimated with data located in $[t-750, t-22]$, otherwise the dependent variable will contain future realizations of RVcc.

those obtained with the martingale definition of the VRP. The Realized GARCH model continues to outperform all other models.

Table B.1: In-Sample Statistics of Model Performance with Alternative VRP Specification

Model	RG	EG	G	HN	$\text{HN}_{\texttt{vd}}$
Volatility Risk Premium					
Bias	0.774	-2.556	-3.126	-3.078	-1.782
RMSE	3.682	4.404	4.812	4.746	4.133
$\triangle\%$		16.4%	23.5%	22.4%	10.9%
MAE	2.184	3.070	3.514	3.458	2.750
$\triangle\%$		28.9%	37.8%	36.8%	20.6%
Volatility Index (VIX)					
Bias	-0.048	-0.203	-0.318	-0.203	0.164
RMSE	2.504	2.898	3.012	3.565	3.504
$\triangle\%$		13.6%	16.9%	29.8%	28.5%
MAE	1.866	2.222	2.199	2.439	2.419
$\triangle\%$		16.0%	15.1%	23.5%	22.9%
Annualized Volatility					
Bias	-0.823	2.353	2.808	2.875	1.946
RMSE	3.783	4.753	4.547	5.893	5.516
$\triangle\%$		20.4%	16.8%	35.8%	31.4%
MAE	2.393	3.653	3.596	4.285	3.881
$\triangle\%$		34.5%	33.4%	44.1%	38.3%

Note: Bias is defined as model-based quantity minus the corresponding market-based quantity; The rows starting with $\triangle\%$ present the increase in percentage of RMSE and MAE for the competing models relative to RG. In this table, the market VRP is defined by $\text{VRP}_t^{\texttt{market}} = \text{VIX}_t - \sqrt{\widehat{\text{RVcc}}_{t+1:t+22}} \times 100$, where the annualized volatility, $\sqrt{\widehat{\text{RVcc}}_{t+1:t+22}} \times 100$, is the predicted value from the HAR model.

B.2 Robustness Check: Estimation with logarithmic VIX

The VIX index is occasionally very volatile and it reached very high values during the financial crises. We inspect if some of our results are driven by outliers. Rather than measuring pricing errors using the level of VIX, we define VIX pricing errors for the logarithmically transformed VIX. This transformation is a well known method for dimming the influence of outliers. Table B.3 provides parameters estimated with logVIX error specifications while in-sample pricing performances are summarized in B.4. Again, we do not find significant changes in those results.

Table B.2: Out-of-sample Statistics of Model Performance with Alternative VRP Specification

	Volatility Risk Premium					**Volatility Index (VIX)**					**Annualized Volatility**				
Full Sample Period: 20070103 - 20181230															
Model	RG	EG	G	HN	HN_{vd}	RG	EG	G	HN	HN_{vd}	RG	EG	G	HN	HN_{vd}
Bias	0.450	-2.685	-2.624	-2.676	-1.614	0.155	-0.212	0.063	-0.533	0.132	-0.295	2.473	2.687	2.143	1.746
RMSE	4.189	5.017	4.783	4.901	4.841	2.870	3.259	3.278	4.298	3.903	3.681	5.068	4.284	5.764	5.663
DM stat.		3.168	2.621	2.931	3.112		3.976	3.579	2.659	1.887		4.285	1.377	3.180	2.898
△%		16.5%	12.4%	14.5%	13.5%		11.9%	12.4%	33.2%	26.5%		27.4%	14.1%	36.1%	35.0%
MAE	2.883	3.588	3.378	3.542	3.519	2.113	2.402	2.387	2.842	2.615	2.283	3.777	3.288	3.928	3.743
DM stat.		3.127	2.370	3.095	3.951		3.995	3.333	4.222	3.137		6.273	4.484	7.322	6.394
△%		19.7%	14.6%	18.6%	18.1%		12.0%	11.5%	25.6%	19.2%		39.6%	30.6%	41.9%	39.0%
Financial Crisis Period: 2007.01.03 - 2012.12.30															
Model	RG	EG	G	HN	HN_{vd}	RG	EG	G	HN	HN_{vd}	RG	EG	G	HN	HN_{vd}
Bias	0.257	-3.043	-3.172	-3.048	-2.109	-0.294	-0.824	-0.484	-1.427	-0.321	-0.551	2.219	2.688	1.621	1.788
RMSE	4.981	5.867	5.744	5.649	5.779	3.383	3.972	3.944	5.497	4.866	4.668	6.087	4.985	7.258	7.318
DM stat.		2.229	2.212	1.824	2.476		4.028	3.277	2.515	1.676		2.841	0.448	2.562	2.557
△%		15.1%	13.3%	11.8%	13.8%		14.8%	14.2%	38.5%	30.5%		23.3%	6.4%	35.7%	36.2%
MAE	3.340	4.341	4.220	4.210	4.383	2.557	3.069	2.946	3.798	3.327	2.849	4.552	3.868	4.908	4.954
DM stat.		2.978	2.761	2.819	4.246		4.254	2.820	3.906	2.541		4.309	2.683	5.542	5.344
△%		23.1%	20.8%	20.6%	23.8%		16.7%	13.2%	32.7%	23.1%		37.4%	26.3%	41.9%	42.5%
Post-Crisis Period: 2013.01.03-2018.12.30															
Model	RG	EG	G	HN	HN_{vd}	RG	EG	G	HN	HN_{vd}	RG	EG	G	HN	HN_{vd}
Bias	0.645	-2.326	-2.075	-2.304	-1.120	0.606	0.403	0.611	0.363	0.585	-0.039	2.729	2.686	2.667	1.705
RMSE	3.206	3.989	3.571	4.014	3.670	2.242	2.339	2.434	2.591	2.602	2.302	3.782	3.442	3.709	3.250
DM stat.		2.696	1.567	2.819	2.281		1.260	1.704	2.882	2.848		5.427	4.927	5.640	4.597
△%		19.6%	10.2%	20.1%	12.6%		4.2%	7.9%	13.5%	13.9%		39.1%	33.1%	37.9%	29.2%
MAE	2.426	2.835	2.536	2.874	2.653	1.669	1.735	1.828	1.885	1.902	1.715	3.003	2.709	2.948	2.532
DM stat.		1.389	0.426	1.543	1.206		0.990	1.819	2.683	2.751		4.903	4.168	5.177	4.337
△%		14.4%	4.3%	15.6%	8.6%		3.8%	8.7%	11.4%	12.2%		42.9%	36.7%	41.8%	32.3%

Note: Let the "model error" be the difference between the model-implied quantity and the market-based quantity. It sample average is denoted Bias and the rows indicated with "△%" state who much larger the RMSE or MAE is for each alternative models, measured relative to RG. Diebold and Mariano statistics (DM stat.) are computed for the relative MSE or MAE losses, where the standard errors are calculated with the Parzen kernel with $H = 42$ as bandwidth. The market VRP in this table is defined by $\text{VRP}_t^{\texttt{market}} = \text{VIX}_t - \sqrt{\widehat{\text{RVcc}}_{t+1:t+22}} \times 100$, where the annualized volatility, $\sqrt{\widehat{\text{RVcc}}_{t+1:t+22}} \times 100$, is the predicted value from the HAR model.

Table B.3: Full Sample Parameter Estimation with Alternative Pricing Errors for log-VIX

Model	RG	EG	G	HN	$HN_{\texttt{vd}}$
λ	0.014	0.137	0.303	4.673	4.969
	(0.036)	(0.010)	(0.021)	(0.339)	(0.796)
ω	-0.148	-0.120	1.60E-06	-6.08E-07	-9.63E-07
	(0.014)	(0.009)	(8.88E-08)	(1.54E-08)	(2.73E-08)
β	0.985	0.986	0.930	0.955	0.950
	(0.001)	9.09E-04	(0.005)	(0.001)	(0.000)
α			0.064	1.70E-06	1.33E-06
			(0.004)	(1.84E-08)	(2.87E-09)
δ				150.120	182.711
				(2.206)	(0.988)
τ_1	-0.082	-0.052			
	(0.005)	(0.004)			
τ_2	0.018	0.103			
	(0.003)	(0.003)			
γ	0.117				
	(0.010)				
κ	0.303				
	(0.230)				
ϕ	1.063				
	(0.025)				
δ_1	-0.088				
	(0.012)				
δ_2	0.127				
	(0.010)				
σ^2	0.308				
	(0.009)				
$-\xi$	0.790				1.191
	(0.086)				(0.022)
$\pi^{\mathbb{P}}$	0.985	0.986	0.994	0.994	0.995
ℓ_r	12895	12660	12523	12518	12652
		[13.963]	[18.044]	[15.533]	[8.733]
$\ell_{\texttt{vix}}$	2361	1703	1884	1435	1574
		[15.843]	[10.203]	[12.186]	[11.770]
$\ell_{r,\texttt{vix}}$	15256	14363	14406	13953	14226
		[19.688]	[16.322]	[16.207]	[13.429]
$\ell_{r,x,\texttt{vix}}$	12126				

Note: QMLE robust standard errors are in parenthesis. The persistence parameter $\pi^{\mathbb{P}}$ is measured by $\beta + \phi\gamma$ for $RG_{\texttt{r}}$ and β for all other models. For the model $HN_{\texttt{vd}}$ we report the value of $(1 + 2\alpha\xi)^{-1}$ in place of $-\xi$ (the implied value for ξ is here -121098.46). The values reported in square brackets under the likelihood are the Voung statistics for all models relative to the RG.

Table B.4: In-sample Statistics of Model Performance with Alternative Error Specification

Model	RG	EG	G	HN	HN_{vd}
Volatility Risk Premium					
Bias	-0.561	-4.092	-4.326	-4.627	-3.697
RMSE	3.664	5.268	5.456	5.679	4.971
△%		30.5%	32.9%	35.5%	26.3%
MAE	2.547	4.509	4.712	4.944	4.105
△%		43.5%	45.9%	48.5%	37.9%
Volatility Index (VIX)					
Bias	-0.202	-0.434	-0.398	-0.856	0.052
RMSE	2.729	3.134	3.077	4.334	3.754
△%		12.9%	11.3%	37.0%	27.3%
MAE	1.958	2.287	2.191	2.528	2.386
△%		14.4%	10.6%	22.5%	17.9%
Annualized Volatility					
Bias	0.359	3.658	3.928	3.771	3.748
RMSE	3.531	4.608	4.524	5.984	5.665
△%		23.4%	22.0%	41.0%	37.7%
MAE	2.340	4.139	4.101	5.009	4.601
△%		43.5%	42.9%	53.3%	49.1%

Note: Bias is defined as model generated value minus their market counterpart; The rows starting with △% present the increase in percentage of RMSE and MAE for the competing models relative to RG.

B.3 Robustness Check: Alternative Realized Measures

A wide range of realized measures of volatility have been proposed since the use of the realized variance was popularized by Andersen and Bollerslev (1998). As alternative measures we use the realized variance based on both 5-minute and 10-minute returns, denoted RV_{5m} and RV_{10m}, respectively. We also consider the bi-power variation, BV_{5m}, by Barndorff-Nielsen (2004), which is computed from 5-minute returns, and three different realized kernels, denoted RK_P, RK_{TH}, and RK_B, that are based on different kernel function. RK_P is based on the Parzen kerne, RK_{TH}, on the modified Tukey-Hanning, which is denoted TH_2 in Barndorff-Nielsen et al. (2008), and RK_B is the realized kernel estimator based on the Bartlett kernel function. Finally, the median-based realized variance., by Andersen et al. (2012) is denoted MedRV. The realized measures were obtained from the Realized Library at Oxford-Man institute.

Table B.5 provides in-sample pricing performances of the Realized GARCH model based on different realized measures. The results are quite similar across with other realized measures. For the VRP we observed some minor improvements over RV_{5m}, but the relative differences are quite small. Interestingly, the two jump robust estimators, BV_{5m} and MedRV, are somewhat worse at fitting the VIX and Annualize Volatility, which indicates that including the jump component is useful in this context. However, for the VRP both jump-robust measures perform similarly to other realized measures. The RV_{5m} seems to be adequate for the present modeling problem, as previously argued in Liu et al. (2015) for a range of problems.

Table B.5: In-sample Statistics of Model Performance with Alternative Realized Measures

Model	RV_{5m}	RV_{10m}	BV_{5m}	RK_P	RK_{TH}	RK_B	MedRV
Volatility Risk Premium							
Bias	-0.498	-0.497	-0.442	-0.514	-0.486	-0.482	-0.520
RMSE	3.825	3.838	3.818	3.843	3.764	3.764	3.844
△%		0.34%	-0.18%	0.48%	-1.62%	-1.60%	0.49%
MAE	2.635	2.646	2.613	2.649	2.586	2.585	2.657
△%		0.42%	-0.81%	0.55%	-1.90%	-1.91%	0.84%
Volatility Index (VIX)							
Bias	-0.048	-0.047	-0.036	-0.032	-0.047	-0.046	-0.037
Bias%	-0.09%	-0.05%	0.29%	0.23%	0.01%	0.02%	0.38%
RMSE	2.504	2.520	2.668	2.575	2.548	2.547	2.808
△%		0.63%	6.15%	2.77%	1.75%	1.72%	10.84%
MAE	1.866	1.896	1.985	1.920	1.917	1.912	2.078
△%		1.59%	6.00%	2.80%	2.64%	2.38%	10.20%
Annualized Volatility							
Bias	0.449	0.450	0.406	0.481	0.439	0.436	0.483
Bias%	8.35%	8.39%	8.33%	8.82%	8.31%	8.30%	9.00%
RMSE	2.942	2.911	3.066	2.909	3.171	3.165	3.048
△%		-1.08%	4.02%	-1.16%	7.22%	7.02%	3.47%
MAE	2.002	1.977	2.130	2.013	2.190	2.185	2.134
△%		-1.28%	6.02%	0.53%	8.57%	8.40%	6.17%

Note: Empirical results based on alternative realized volatility measures. RV_{5m} and RV_{10m} are the realized variances based on 5 minute returns and 10 minute returns, respectively, BV_{5m} is the bipower variation based on five-minute returns,RK_P, RK_{TH}, and RK_B are realized kernels estimator based the Parzen, a modified Tukey-Hanning, and the Bartlett kernel functions. Finally, MedRV is the median based realized variance. The rows starting with △% present the increase in percentage of RMSE and MAE for the competing models relative to the one using 5 minute RV.

B.4 Robustness Check: Non-Gaussian VIX pricing

Table B.6 reports all parameter estimates of the model in which the VIX pricing formula has been amended to permit non-Gaussian distributions for z_t^* and u_t^*. We have included the estimates for the original specification to simplify a direct comparison. The parameters, which the two specifications have in common, are nearly identical in all cases, whereas the four new parameters for skewness and excess kurtosis are all significantly different from zero.

Table B.6: RG-Normal and RG-GramCharlier

Model	RG-Normal	RG-GramCharlier
λ	0.015	0.018
	(0.010)	(0.010)
ω	-0.088	-0.087
	(0.014)	(0.013)
β	0.991	0.991
	(0.001)	(0.001)
τ_1	-0.073	-0.073
	(0.005)	(0.005)
τ_2	0.012	0.012
	(0.002)	(0.002)
γ	0.080	0.079
	(0.009)	(0.009)
κ	0.427	0.420
	(0.278)	(0.245)
ϕ	1.078	1.077
	(0.029)	(0.026)
δ_1	-0.083	-0.082
	(0.010)	(0.009)
δ_2	0.129	0.129
	(0.010)	(0.008)
σ^2	0.325	0.326
	(0.010)	(0.010)
ξ	-1.07	-1.06
	(0.130)	(0.129)
ς_z		-0.348
		(0.038)
κ_z		0.993
		(0.114)
ς_u		0.189
		(0.048)
κ_u		0.748
		(0.113)
ℓ_r	12864	12951
ℓ_x	-3229	-3193
$\ell_{\texttt{vix}}$	-8812	-8805
$\ell_{r,x,\texttt{vix}}$	823	953

Note: Parameter estimates are reported with QMLE robust standard errors in parenthesis.