GARCH Option Pricing Models, the CBOE VIX, and Variance Risk Premium

JINJI HAO
Department of Economics, Washington University in St. Louis

JIN E. ZHANG
Department of Accountancy and Finance, Otago Business School, University of Otago and School of Economics and Finance, The University of Hong Kong

ABSTRACT
In this article, we derive the corresponding implied VIX formulas under the locally risk-neutral valuation relationship (LRNVR) proposed by Duan (1995) when a class of square-root stochastic autoregressive volatility (SR-SARV) models are proposed for S&P 500 index. The empirical study shows that the GARCH implied VIX is consistently and significantly lower than the CBOE VIX for all kinds of GARCH model investigated when they are estimated with returns only. When jointly estimated with both returns and VIX, the parameters are distorted unreasonably, and the GARCH implied VIX still cannot fit the CBOE VIX from various statistical aspects. The source of this discrepancy is then theoretically analyzed. We conclude that the GARCH option pricing under the LRNVR fails to incorporate the price of volatility or variance risk premium. (JEL: G13, C52)

KEYWORDS: GARCH option pricing models, GARCH implied VIX, the CBOE VIX, Variance risk premium

We are especially grateful to Eric Renault (editor) and two anonymous referees whose helpful comments substantially improved the paper. We also acknowledge helpful comments from Andrew Carverhill (our AsianFA discussant), Werner Ploberger, Jinghong Shu, Tianyi Wang, Dongming Zhu and seminar participants at Peking University, Asian Finance Association (AsianFA) 2010 International Conference in Hong Kong. Jin E. Zhang has been supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. HKU 7549/09H). Address correspondence to Jin E. Zhang, School of Economics and Finance, The University of Hong Kong, Pokfulam Road, Hong Kong, P. R. China, or e-mail: jinzhang@hku.hk

© The Author, 2013. Published by Oxford University Press. All rights reserved.
For Permissions, please email: journals.permissions@oup.com
1 INTRODUCTION

Finance literature has put much effort on studying the premia that investors require for compensating various risks in financial market, especially the equity risk premium for price risk (volatility). However, instead of a constant volatility assumed in the Black–Scholes framework, a lot of research has confirmed that the volatility itself is time-varying, which is termed volatility risk. Many stochastic volatility models and various GARCH models go along this line.

Then a natural question is whether this volatility risk is priced and compensated in financial market. One of the possible rationales for the existence of volatility risk premium is the negative correlation between the volatility and the index, which has been verified in many literatures. In the context of asset pricing theory, the source of risk is the correlation with the market portfolio, aggregate consumption or pricing kernel. Theoretically, the negative correlation between volatility and index suggests a negative risk premium. If so, the premia required by investors should be reflected in the prices of volatility-dependent assets such as options and volatility products. The pursuit of empirical evidence generally proceeds in two directions. One is to study the phenomenon that the implied volatility of options exceeds the realized volatility. Various delta-neutral portfolios of options are constructed to test whether significant gains or losses could be produced. Another one is to investigate the difference between the variance swap rate and the realized variance, which is coined variance premium. Variance swap rate, the risk-neutral expectation of the future variance, can be replicated with European options (See Demeterfi et al. 1999; Carr and Wu 2009). Methodologies for calculating a model-free realized variance have also been developed (see Andersen et al. 2003).

Since 1980s, option pricing models with stochastic volatility have introduced the market price of volatility risk when changing from the physical probability measure to the risk-neutral measure. These papers include Wiggins (1987), Johnson and Shanno (1987), Hull and White (1987), Scott (1987) and Heston (1993). However, they set the market price of volatility risk to either zero or a constant and discussed little about the size, sign or dynamics of this parameter.

Since the beginning of this century, the evidence of the existence of volatility risk premium has been well documented. Coval and Shumway (2001) studied the expected option return under the framework of classic asset pricing theory. They showed that the zero-beta, at the money straddles that are in long positions of volatility suffer from average weekly losses of about three percent. Bakshi and Kapadia (2003) constructed a correspondence between the sign and magnitude of volatility risk premium and the mean delta-hedged portfolio returns. Their empirical results indicated a negative market volatility risk premium. Carr and Wu (2009) calculated the variance premia for several stock market indexes through replications with options, and average negative variance premium was shown.

The dynamics and driving forces of variance premium are studied in recent literature. Vilkov (2008) used the synthetic variance swap returns to approximate the
variance risk premium and studied the dynamics and cross-sectional properties of variance premia embedded in index options and individual stock options. Todorov (2010) studied variance risk in terms of stochastic volatility and jumps. Model-free realized variance and realized jumps are constructed using high-frequency data. He found that price jumps play an important role in explaining the variance risk premium. Specifically, the estimated variance risk premium increases after a big market jump and slowly reverts to its long-run mean thereafter. Eraker (2008) captured the volatility premium and the large negative correlation between shocks to volatility and stock price with a general equilibrium based on long-run risk.

In this article, we investigate whether the GARCH option pricing model can capture the variance premium. Since the seminal autoregressive conditional heteroscedasticity (ARCH) model of Engle (1982) and the generalized autoregressive conditional heteroscedasticity (GARCH) model of Bollerslev (1986), the GARCH models have attracted huge attention from the academics and practitioners and have been intensively used to model the financial times series. This is mostly because they can capture the volatility clustering and fat tails that are typical properties of the financial time series. The family of GARCH models has also been enriched to capture the stylized fact that negative returns have higher impact on the volatility than positive ones, which is called leverage effect. This class of GARCH models includes the exponential GARCH (EGARCH) of Nelson (1991), the threshold GARCH (TGARCH) of Glosten, Jagannathan, and Runkle (1993) and the non-linear asymmetric GARCH (AGARCH) of Engle and Ng (1993) and the like. Engle and Lee (1993) introduced the component GARCH (CGARCH) that separates the conditional variance into a transitory component and a permanent component.

Duan (1995) pioneered in employing the GARCH model in the option pricing theory. He put forward an equilibrium argument that the options can be priced under a locally risk-neutral valuation relationship (LRNVR) with some assumptions on the utility function when the price of the underlying asset follows a GARCH process. Kallsen and Taqqu (1998) considered a broad class of ARCH-type models embedded into a continuous-time framework and derived the same result by a no-arbitrage argument. The GARCH option pricing model has some linkage with those bivariate diffusion option pricing models. Duan (1996, 1997), showed that most variants of GARCH model mentioned above converge to the bivariate diffusion processes commonly used for modeling the stochastic volatility. Ritchken and Trevor (1999) developed a lattice algorithm that is applicable for option pricing under both GARCH models and bivariate diffusions. We will further discuss this limiting property in this article.

The concept of LRNVR in Duan (1995) has been mainly followed by the subsequent GARCH option pricing literature, for instance, Heston and Nandi (2001), Christoffersen and Jacobs (2004) and Christoffersen et al. (2008) among the others. Garcia, Ghysels, and Renault (2010) provided an interesting and
HAO & ZHANG | GARCH Option Pricing Models

A comprehensive review on the econometrics of option pricing models. We will also focus on Duan’s option pricing model in this article. There are several papers discussing the weakness of traditional GARCH option pricing models. Garcia and Renault [1998] pointed out that the hedging formula given in Duan [1995] is not consistent with the fact that the GARCH option pricing is not homogeneous of degree one with respect to the spot price and the strike price. A new GARCH option pricing model with filtered historical simulation developed by Barone-Adesi, Engle, and Mancini [2008] avoided a change of probability measure by directly calibrating a new risk neutral GARCH model on option prices. The most relevant research to our work is Christoffersen, Heston, and Jacobs [2011] in which they propose a variance-dependent pricing kernel for GARCH model accounting for both the equity risk premium and the variance risk premium.

To study the variance premium captured by GARCH option pricing models, we will derive the GARCH implied VIX under the LRNVR proposed by Duan [1995]. To do this, a more general class of square-root stochastic autoregressive volatility (SR-SARV) models Meddahi and Renault [2004] which subsumes many specific GARCH models is proposed for the S&P 500 index. We calculate the (squared) VIX as a risk-neutral expectation of the average variance over the next 21 trading days under the LRNVR. In this article, GARCH(1,1) and four other variants of GARCH(1,1) model are numerically estimated with different data sets.

We then compare the GARCH implied VIX with the CBOE VIX. We find that the GARCH implied VIX is significantly and consistently lower than the CBOE VIX when only returns are used for estimation. The difference is around 3.6 that is consistent with the empirical evidence of the size of volatility premium. When VIX is used for estimation, the parameters, especially the equity risk premium, from physical to risk-neutral ones is still under development. For example, Duan [1999] uses Box-Cox transformation to introduce conditional fat-tailed innovations; Siu, Tong, and Yang [2004] rely on the conditional Esscher transformation to incorporate infinitely divisible distributed innovation; Christoffersen, Heston, and Jacobs [2004] develop a model with Inverse Gaussian innovation allowing for conditional skewness; Christoffersen et al. [2011] characterize the Radon-Nikodym derivative for neutralizing a class of GARCH models with more general innovations. It is an interesting research topic to examine the performance of these non-Gaussian GARCH option pricing models in capturing the variance risk premium. However, under normality, Duan [1995] and Siu, Tong, and Yang [2004] give the same result as Duan [1995] with LRNVR, and Christoffersen et al. [2011] subsumes Duan [1995] and Heston and Nandi [2000] as special cases. It would be reasonable to suspect that these models might have similar problem observed in this article. Certain amount of research needs to be done in order to make any conclusion. We decide to leave this topic for the future research.

We share the same view that ‘The filtering problem in these models (GARCH dynamics) is straightforward because the distribution of one-period return has a known conditional variance, ... it has implication for option pricing. Because the models do not contain an independent adjustment for variance risk, they do not offer much flexibility in the modeling of variance risk premia’ Christoffersen, Heston, and Jacobs [2011].

VIX is the Chicago Board of Exchange (CBOE) listed volatility index, which is updated in 2004 and reflects expectations of the volatility of the S&P 500 index over the next 30 calendar days. Demeterci et al. [1999] showed that the squared VIX is actually the variance swap rate and can be replicated with a portfolio of options written on the S&P 500 index.
are distorted to unreasonable levels, and the statistics of the CBOE VIX remain unmatched.

The source of this discrepancy is then theoretically analyzed. We investigate the LRNVR and the diffusion limits of the GARCH model under both the physical measure and the LRNVR. It is shown that the innovation of volatility is invariant with respect to the change of probability measure with the LRNVR, and no premium for the volatility risk is compensated.

The article is constructed as follows. In Section 2 we first review the LRNVR of Duan (1995) for changing probability measure for the GARCH(\(p,q\)) model. We then derive VIX formulas under the LRNVR for a broad class of GARCH models and its extension, which subsume GARCH, TGARCH, AGARCH, CGARCH, and EGARCH models. In Section 3 we estimate these models using time series of the S&P 500 index and the CBOE VIX. The GARCH implied VIX is computed and compared with the CBOE VIX in Section 4. In Section 5 the failure of the GARCH option pricing under the LRNVR to incorporate the price of volatility risk is analyzed. We conclude in Section 6.

2 THEORETICAL RESULTS ON GARCH IMPLIED VIX

2.1 GARCH Option Pricing Models

Duan (1995) utilized a linear GARCH process for modeling the underlying asset and pricing the options written on it. In that article, the return of the asset in each period is modeled to follow a conditional lognormal distribution under the physical measure \(P\),

\[
\ln \frac{X_t}{X_{t-1}} = r + \lambda \sqrt{h_t} - \frac{1}{2} h_t + \epsilon_t, \tag{1}
\]

where \(X_t\) is the price of the asset, \(r\) is the constant interest rate, and \(\lambda\) is the risk premium; \(\epsilon_t\) follows a GARCH(\(p,q\)) process

\[
\epsilon_t | \phi_{t-1} \sim N(0,h_t) \quad \text{under measure } P,
\]

\[
h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i}, \tag{2}
\]

where \(\phi_t\) is the information set of all information up to and including time \(t\); \(p \geq 0, q \geq 0; \alpha_0 > 0, \alpha_i \geq 0, i = 1, \ldots, q; \beta_i \geq 0, i = 1, \ldots, p\).

With assumptions made on the utility function and the aggregated consumption growth, Duan (1995) proposed a new locally risk neutral valuation relationship, \(Q\), under which

\[
\ln \frac{X_t}{X_{t-1}} = r - \frac{1}{2} h_t + \xi_t, \tag{3}
\]
and

\[ \xi_t | \phi_{t-1} \sim N(0, h_t), \]

\[ h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \left( \xi_{t-i} - \lambda \sqrt{h_{t-i}} \right)^2 + \sum_{i=1}^{p} \beta_i h_{t-i}. \] (4)

### 2.2 GARCH Implied VIX

VIX reflects investors’ expectation of the volatility of the S&P 500 index in the following 30 calendar days, that is

\[ \left( \frac{\text{VIX}_t}{100} \right)^2 = \mathbb{E}_t^Q \left[ \frac{1}{\tau_0} \int_{t+	au_0}^{t+\tau_0} \tilde{h}_s ds \right], \] (5)

where \( \tau_0 = 30 \) calendar days or 21 trading days, and \( \tilde{h}_s \) is the instantaneous annualized variance of the rate of return of S&P 500. In this article, we calculate VIX as an expected arithmetic average of the variance in the \( n \) subperiods of the following 30 calendar days, that is

\[ \left( \frac{\text{VIX}_t}{100} \right)^2 = \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}_t^Q \left[ \tilde{h}_{t+k} \right]. \] (6)

Especially, we will use data with daily frequency, that is \( \tau_0 = n \), then

\[ \text{Vix}_t = \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}_t^Q \left[ h_{t+k} \right], \] (7)

where \( \text{Vix}_t = \frac{1}{252} \left( \frac{\text{VIX}_t}{100} \right)^2 \) is defined as a proxy for \( \text{VIX}_t \) in terms of daily variance.

The conditional mean of future variance can be calculated under a broad class of GARCH models. We will consider the square-root stochastic autoregressive volatility (SR-SARV) models (Meddahi and Renault, 2004).

**Definition 1:** (Discrete time SR-SARV(p) model, Meddahi and Renault, 2004). A stationary square-integrable process \( \{ \varepsilon_t, t \in \mathbb{Z} \} \) is called a SR-SARV(p) process with respect to a filtration \( J_t, t \in \mathbb{Z} \), if:

1. \( \varepsilon_t \) is a martingale difference sequence (m.d.s.) w.r.t. \( J_{t-1} \), that is \( E[\varepsilon_t | J_{t-1}] = 0 \),
2. the conditional variance process \( f_t \) of \( \varepsilon_{t+1} \) given \( J_t \) is a marginalization of a stationary \( J_t \)-adapted VAR(1) of dimension \( p \):

\[ f_t = \text{Var}[\varepsilon_{t+1} | J_t] = \varepsilon_t F_t, \] (8)

\[ F_t = \Omega + \Gamma F_{t-1} + V_t, \quad \text{with} \ E[V_t | J_{t-1}] = 0, \] (9)

\footnote{Indeed, \( f_t = h_{t+1} \). But we adopt the notations in Meddahi and Renault (2004) here.}
where \( e \in \mathbb{R}^p, \Omega \in \mathbb{R}^p \) and the eigenvalues of \( \Gamma \) have modulus less than one.

If the S&P 500 follows a SR-SARV\((p)\) process under the risk neutral measure, an analytical formula for the GARCH implied VIX can be obtained.

**Proposition 1:** If the S&P 500 follows a SR-SARV\((p)\) process under the locally risk neutral valuation relationship \( Q \) proposed by Duan [1995], then the implied VIX at time \( t \) is affine in \( F_t \), i.e.,

\[
V_{ixt} = \zeta + \Psi F_t, \quad \Psi \in \mathbb{R}^p.
\]  \( (10) \)

In particular, if \( p = 1 \) (then \( e = 1 \)), the implied VIX at time \( t \) is a linear function of the conditional variance of the next period,

\[
V_{ix1} = \zeta + \psi f_t, \quad \psi \in \mathbb{R},
\]  \( (11) \)

where

\[
\zeta = \frac{\Omega}{1 - \Gamma} (1 - \psi),
\]

\[
\psi = \frac{1 - \Gamma^n}{n(1 - \Gamma)}.
\]

**Proof.** See Appendix.  \( \blacksquare \)

In particular, the threshold GARCH\((1,1)\) (TGARCH) of Glosten, Jagannathan, and Runkle [1993], the non-linear asymmetric GARCH\((1,1)\) (AGARCH) of Engle and Ng [1993] and the component GARCH\((1,1)\) (CGARCH) of Engle and Lee [1993], which are widely used, are special cases of SR-SARV\((p)\) models. In specific, they take the forms of the following:

TGARCH\((1,1)\):

**Physical measure:**

\[
h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \theta \epsilon_{t-1}^2 1(\epsilon_{t-1} < 0) + \beta_1 h_{t-1},
\]  \( (12) \)

**LRNVR:**

\[
h_t = \alpha_0 + \left( \xi_{t-1} - \lambda \sqrt{h_{t-1}} \right) \left[ \alpha_1 + \theta 1(\xi_{t-1} - \lambda \sqrt{h_{t-1}} < 0) \right] + \beta_1 h_{t-1}.
\]  \( (13) \)

AGARCH\((1,1)\):

**Physical measure:**

\[
h_t = \alpha_0 + \alpha_1 \left( \epsilon_{t-1} - \theta \sqrt{h_{t-1}} \right)^2 + \beta_1 h_{t-1},
\]  \( (14) \)

**LRNVR:**

\[
h_t = \alpha_0 + \alpha_1 \left( \xi_{t-1} - \lambda \sqrt{h_{t-1}} - \theta \sqrt{h_{t-1}} \right)^2 + \beta_1 h_{t-1}.
\]  \( (15) \)

CGARCH\((1,1)\):

**Physical measure:**

\[
h_t - q_t = \alpha_1 \left( \epsilon_{t-1}^2 - q_{t-1} \right) + \beta_1 \left( h_{t-1} - q_{t-1} \right),
\]  \( (16) \)
Proposition 2: Let \( \{\xi_t, t \in \mathbb{Z}\} \) be a m.d.s. with the conditional variance \( h_t \equiv \text{Var}[\xi_t|\xi_{\tau}, \tau \leq t-1] \) under the LRNVR. If \( h_t \) is given by (4) or (15), then \( \xi_t \) is a SR-SARV(1) process. If \( h_t \) is given by (17), then \( \xi_t \) is a SR-SARV(2) process. Furthermore, if \( u_t = \xi_t / \sqrt{h_t} \) is i.i.d., the TGARCH model (13) is also a SR-SARV(1) process.\(^5\)

Proof. See Appendix. \( \blacksquare \)

In this article, we assume \( \epsilon_t / \sqrt{h_t} \) and \( \xi_t / \sqrt{h_t} \) are i.i.d. under the physical measure and the LRNVR, respectively, which is sufficient for this proposition to be true.

Corollary 1: Under the locally risk neutral valuation relationship \( Q \) proposed by Duan (1995), if the S&P 500 follows a GARCH(1,1), TGARCH(1,1) or AGARCH(1,1) process, the implied VIX is a linear function of the conditional variance of the next period, \( h_{t+1} \). If the S&P 500 follows a CGARCH(1,1) process, the implied VIX is a linear function of the transitory and permanent components of the conditional variance of the next period, \( h_{t+1} \) and \( q_{t+1} \).

The specific VIX formulas are given in Appendix.

Now we extend the SR-SARV model to a square-root stochastic exponential autoregressive volatility (SR-SEARV) model:

Definition 2: (Discrete time SR-SEARV(1) model). A stationary square-integrable process \( \{\epsilon_t, t \in \mathbb{Z}\} \) is called a SR-SEARV(1) process with respect to a filtration \( J_t, t \in \mathbb{Z} \), if:

(i) \( \epsilon_1 \) is a martingale difference sequence w.r.t. \( J_{t-1}, \) that is \( E[\epsilon_1|J_{t-1}] = 0, \)

(ii) the logarithm of the conditional variance process \( f_t \) of \( \epsilon_{t+1} \) given \( J_t \) is a stationary \( J_t \)-adapted AR(1):

\[
\ln f_t = \omega + \gamma \ln f_{t-1} + v_t, \text{ with } v_t \text{ i.i.d.} \tag{18}
\]

where \( |\gamma| < 1. \)

\(^5\)Under the LRNVR, the VGARCH model discussed in Meddahi and Renault (2008) is also a SR-SARV(1), while the Asymmetric GARCH model, a different version used in Meddahi and Renault (2008), and the Heston and Nandi model are not because \( \sqrt{h_t} \) terms would appear in the expressions of conditional variance after the risk neutralization.
Proposition 3: If the S&P 500 follows a SR-SEARV(1) process under the locally risk neutral valuation relationship Q proposed by Duan (1995), then the implied VIX is a polynomial function of the conditional variance of the next period, $h_{t+1}$.

We will show that the EGARCH(1,1) model is a SR-SEARV(1) below.

**EGARCH(1,1):**

Physical measure:

\[
\ln h_t = \alpha_0 + \beta_1 \ln h_{t-1} + g(z_{t-1}), \quad z_t = \epsilon_t / \sqrt{h_t}, \tag{19}
\]

\[
g(z_{t-1}) = \alpha_1 z_{t-1} + \kappa \left( |z_{t-1}| - \sqrt{2/\pi} \right).
\]

LRNVR:

\[
\ln h_t = \alpha_0 + \beta_1 \ln h_{t-1} + g(u_{t-1} - \lambda), \quad u_t = \xi_t / \sqrt{h_t}, \tag{20}
\]

\[
g(u_{t-1} - \lambda) = \alpha_1 (u_{t-1} - \lambda) + \kappa \left( |u_{t-1} - \lambda| - \sqrt{2/\pi} \right).
\]

Proposition 4: If \( \{\xi_t, t \in \mathbb{Z}\} \) is a m.d.s. under the LRNVR with the conditional variance $h_t \equiv \text{Var}[\xi_t]$, $t \leq t - 1$ given by (20) and $u_t = \xi_t / \sqrt{h_t}$ i.i.d., then $\xi_t$ is a SR-SEARV(1) process.

Let $v_t = g(u_t - \lambda)$. Since $u_t$ is i.i.d., $v_t$ is also i.i.d., and the EGARCH(1,1) model is a SR-SEARV(1).

Corollary 2: If the S&P 500 follows an EGARCH(1,1) process under the locally risk neutral valuation relationship Q proposed by Duan (1995), the implied VIX is a polynomial function of the conditional variance of the next period, $h_{t+1}$.

The specific VIX formula is also given in Appendix.

3 DATA AND ESTIMATION

Based on the theoretical results obtained in last section, a natural question is whether the GARCH implied VIX fits the market VIX well. In this section, we will investigate it by estimating GARCH models and calculating corresponding VIX times series.

In this article, we will use two time series data from the closing price of S&P 500 index and the CBOE VIX ranging from January 2, 1990 to August 10, 2009. We also use the daily 3-month U.S. Treasury bills (secondary market) rate as the risk-free rate and we get this time series data from the Federal Reserve website.
We may have different approaches in estimating the models. The first straightforward one is to run a maximum likelihood estimation of the GARCH models under the physical measure using returns only. This is feasible since there is no separate parameter for the risk-neutral process. The log-likelihood function for all the five versions of GARCH (1,1) model is,

\[ \ln L_R = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^{T} \left\{ \ln(h_t) + \left[ \ln\left(\frac{X_t}{X_{t-1}}\right) - r - \lambda \sqrt{h_t} + \frac{1}{2} h_t \right]^2 / h_t \right\}, \]

(21)

with \( h_t \) updated by respective conditional variance processes.

The second approach is to run a joint maximum likelihood estimation using both returns and VIX since the market VIX may contain additional information about the underlying return process. Indeed, there is an existing literature on combining data from both the underlying and option markets for model estimation. However, the problem is that the daily return innovation \( z_t \) simultaneously determines the current price level \( X_t \) by (1) and the conditional variance of the next period \( h_{t+1} \) by (2) and its variants. The later then determines the current VIX level \( Vix_t \) by VIX formulas we derived in last section. To accommodate both time series, we will allow for a difference between the CBOE VIX and the implied VIX by specifying the following model on a daily basis.

\[ VIX_{Mkt} = VIX_{Imp} + \mu, \quad \mu \sim i.i.d. N\left(0, s^2\right). \]

(22)

where \( s^2 \) is estimated with sample variance \( \hat{s}^2 = \text{var}(VIX_{Mkt} - VIX_{Imp}) \). We then have the log-likelihood function corresponding to the CBOE VIX

\[ \ln L_V = -\frac{T}{2} \ln\left(2\pi \hat{s}^2\right) - \frac{1}{2\hat{s}^2} \sum_{t=1}^{T} \left( VIX_{Mkt}^t - VIX_{Imp}^t \right)^2, \]

(23)

which is also the log-likelihood function when we use VIX time series only. The joint estimation of the parameters can be obtained by maximizing the joint log-likelihood function

\[ \ln L_T = \ln L_R + \ln L_V. \]

(24)

An estimation using only the market VIX is also reported for comparison. We then evaluate the performance of models in terms of the goodness of fit of VIX from various aspects.

In particular, we set the conditional variance for the first period as the variance of the rate of return of S&P 500 index over the whole sample period. It is noted

---

6For example, Pan (2002), Chernov and Ghysels (2000), and Jones (2003).

7TGARCH by (12), AGARCH by (14), CGARCH by (16), and EGARCH by (19).

8This may be rationalized as a measurement error.

9Note that \( \sqrt{Vix_t} \) expresses the VIX index in terms of daily variance, so we use \( \sqrt{Vix_t} \) here.
that the stationary conditions for the GARCH processes under physical and risk-neutral measures are different, with the later having more strict constraints on the parameters. Thus, when estimating the GARCH parameters under the physical measure, we maximize the log-likelihood functions subject to the stationary conditions under the risk-neutral measure. Specifically, for the linear GARCH(1,1), the stationary condition constraint is

\[ \alpha_1 \left(1 + \lambda^2\right) + \beta_1 < 1. \]

For the TGARCH(1,1) process, the stationary condition constraint is

\[ \alpha_1 \left(1 + \lambda^2\right) + \beta_1 + \theta \left[ \frac{\lambda}{\sqrt{2\pi}} e^{-\frac{\lambda^2}{2}} + \left(1 + \lambda^2\right) N(\lambda) \right] < 1. \]

For the AGARCH(1,1) process, the stationary condition constraint is

\[ \alpha_1 \left[1 + (\lambda + \theta)^2\right] + \beta_1 < 1. \]

For the CGARCH(1,1) process, the stationary condition constraint is that the eigenvalues of the coefficient matrix

\[
\left(\begin{array}{cc}
\alpha + \beta + (\phi + \alpha)\lambda^2 & \rho - \alpha - \beta \\
\phi\lambda^2 & \rho
\end{array}\right)
\]

have modulus less than 1. For the EGARCH(1,1) process, the stationary condition constraint is

\[ |\beta_1| < 1. \]

4 NUMERICAL RESULTS

In this section, we examine the performance of the models estimated in fitting VIX time series. Table 1 displays the maximum likelihood estimates and standard errors of GARCH models. The last three columns report the log-likelihood values. Although the contributions from both returns and VIX as well as the total are reported, we maximize \( \ln L_R \) when only returns are used, \( \ln L_V \) when only VIX levels are used and the total when both time series are used.

The most notable finding in Table 1 is that the equity risk premium increases significantly in the GARCH, TGARCH, and CGARCH models when the VIX data is considered, especially when it is used alone. It increases from 0.0523 (returns used) to 0.2068 (returns and VIX used) and 0.7914 (VIX used) in the GARCH model, from 0.0231 (returns used) to 0.0751 (returns and VIX used) and 0.4441 (VIX used) in the TGARCH model, and from 0.0529 (returns used) to 0.2651 (returns and VIX used) and 0.8764 (VIX used) in the CGARCH model. In the AGARCH model, it slightly
<table>
<thead>
<tr>
<th>Model &amp; data</th>
<th>$\alpha_0$</th>
<th>$\alpha_1$</th>
<th>$\beta_1$</th>
<th>$\kappa$</th>
<th>$\theta$</th>
<th>$\rho$</th>
<th>$\phi$</th>
<th>$\lambda$</th>
<th>Log-likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GARCH</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>Returns</td>
<td>7.2449e-7</td>
<td>0.0632</td>
<td>0.9312</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.0523</td>
<td>36649</td>
</tr>
<tr>
<td></td>
<td>(1.6103e-7)</td>
<td>(0.0064)</td>
<td>(0.0069)</td>
<td>(0.0010)</td>
<td>(0.0058)</td>
<td>(0.0139)</td>
<td></td>
<td></td>
<td>16097</td>
</tr>
<tr>
<td>VIX</td>
<td>1.7109e-6</td>
<td>0.0866</td>
<td>0.9387</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.7914</td>
<td>38345</td>
</tr>
<tr>
<td></td>
<td>(0.0406e-6)</td>
<td>(0.0009)</td>
<td>(0.0018)</td>
<td>(0.0005)</td>
<td>(0.0008)</td>
<td>(0.0018)</td>
<td></td>
<td></td>
<td>14472</td>
</tr>
<tr>
<td>Both</td>
<td>1.6766e-6</td>
<td>0.0473</td>
<td>0.9488</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.2068</td>
<td>39507</td>
</tr>
<tr>
<td></td>
<td>(0.0959e-6)</td>
<td>(0.0011)</td>
<td>(0.0012)</td>
<td>(0.0002)</td>
<td>(0.0002)</td>
<td>(0.0121)</td>
<td></td>
<td></td>
<td>15872</td>
</tr>
<tr>
<td><strong>TGARCH</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>Returns</td>
<td>1.0955e-6</td>
<td>0.0012</td>
<td>0.9333</td>
<td>–</td>
<td>0.1088</td>
<td>–</td>
<td>–</td>
<td>0.0231</td>
<td>36424</td>
</tr>
<tr>
<td></td>
<td>(0.1686e-6)</td>
<td>(0.0049)</td>
<td>(0.0066)</td>
<td>(0.0010)</td>
<td>(0.0010)</td>
<td>(0.0141)</td>
<td></td>
<td></td>
<td>16165</td>
</tr>
<tr>
<td>VIX</td>
<td>1.6109e-6</td>
<td>0.0047</td>
<td>0.9516</td>
<td>–</td>
<td>0.0426</td>
<td>–</td>
<td>–</td>
<td>0.4441</td>
<td>39416</td>
</tr>
<tr>
<td></td>
<td>(0.0405e-6)</td>
<td>(0.0026)</td>
<td>(0.0019)</td>
<td>(0.0003)</td>
<td>(0.0003)</td>
<td>(0.0041)</td>
<td></td>
<td></td>
<td>15500</td>
</tr>
<tr>
<td>Both</td>
<td>1.5091e-6</td>
<td>0.0036</td>
<td>0.9600</td>
<td>–</td>
<td>0.0619</td>
<td>–</td>
<td>–</td>
<td>0.0751</td>
<td>39815</td>
</tr>
<tr>
<td></td>
<td>(0.0989e-6)</td>
<td>(0.0015)</td>
<td>(0.0010)</td>
<td>(0.0022)</td>
<td>(0.0022)</td>
<td>(0.0033)</td>
<td></td>
<td></td>
<td>15958</td>
</tr>
<tr>
<td><strong>AGARCH</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>Returns</td>
<td>1.1283e-6</td>
<td>0.0552</td>
<td>0.8802</td>
<td>–</td>
<td>1.0143</td>
<td>–</td>
<td>–</td>
<td>0.0150</td>
<td>36960</td>
</tr>
<tr>
<td></td>
<td>(0.1574e-6)</td>
<td>(0.0056)</td>
<td>(0.0050)</td>
<td>(0.0097)</td>
<td>(0.0097)</td>
<td>(0.0142)</td>
<td></td>
<td></td>
<td>16179</td>
</tr>
<tr>
<td>Both</td>
<td>1.8976e-6</td>
<td>0.0393</td>
<td>0.9349</td>
<td>–</td>
<td>0.7718</td>
<td>–</td>
<td>–</td>
<td>0.0156</td>
<td>39849</td>
</tr>
<tr>
<td></td>
<td>(0.0410e-6)</td>
<td>(0.0009)</td>
<td>(0.0017)</td>
<td>(0.0028)</td>
<td>(0.0028)</td>
<td>(0.0042)</td>
<td></td>
<td></td>
<td>15988</td>
</tr>
<tr>
<td><strong>CGARCH</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>Returns</td>
<td>2.0515e-7</td>
<td>0.0470</td>
<td>0.9180</td>
<td>–</td>
<td>–</td>
<td>0.9983</td>
<td>0.0227</td>
<td>0.0529</td>
<td>36476</td>
</tr>
<tr>
<td></td>
<td>(0.3271e-7)</td>
<td>(0.0085)</td>
<td>(0.0040)</td>
<td>(0.0010)</td>
<td>(0.0010)</td>
<td>(0.0039)</td>
<td>(0.0027)</td>
<td></td>
<td>16104</td>
</tr>
<tr>
<td>VIX</td>
<td>2.7668e-7</td>
<td>0.0504</td>
<td>0.8902</td>
<td>–</td>
<td>–</td>
<td>0.9939</td>
<td>0.0029</td>
<td>0.8564</td>
<td>38467</td>
</tr>
<tr>
<td></td>
<td>(0.1304e-7)</td>
<td>(0.0003)</td>
<td>(0.0029)</td>
<td>(0.0004)</td>
<td>(0.0004)</td>
<td>(0.0017)</td>
<td>(0.0018)</td>
<td></td>
<td>14195</td>
</tr>
<tr>
<td>Both</td>
<td>9.6898e-7</td>
<td>0.0602</td>
<td>0.8819</td>
<td>–</td>
<td>–</td>
<td>0.9975</td>
<td>0.0331</td>
<td>0.2651</td>
<td>39828</td>
</tr>
<tr>
<td></td>
<td>(0.3156e-7)</td>
<td>(0.0039)</td>
<td>(0.0086)</td>
<td>(0.0004)</td>
<td>(0.0004)</td>
<td>(0.0038)</td>
<td>(0.0018)</td>
<td></td>
<td>15889</td>
</tr>
<tr>
<td><strong>EGARCH</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>Returns</td>
<td>-0.1329</td>
<td>-0.0921</td>
<td>0.9854</td>
<td>0.1105</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.0167</td>
<td>36661</td>
</tr>
<tr>
<td></td>
<td>(0.0178)</td>
<td>(0.0079)</td>
<td>(0.0019)</td>
<td>(0.0000)</td>
<td>(0.0010)</td>
<td>(0.0040)</td>
<td>(0.0014)</td>
<td></td>
<td>16169</td>
</tr>
<tr>
<td>VIX</td>
<td>-0.0761</td>
<td>-0.0641</td>
<td>0.8895</td>
<td>0.0906</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>-0.0690</td>
<td>40480</td>
</tr>
<tr>
<td></td>
<td>(0.0077)</td>
<td>(0.0029)</td>
<td>(0.0006)</td>
<td>(0.0034)</td>
<td>(0.0034)</td>
<td>(0.0051)</td>
<td></td>
<td></td>
<td>16002</td>
</tr>
<tr>
<td>Both</td>
<td>-0.0838</td>
<td>-0.0614</td>
<td>0.8981</td>
<td>0.0955</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.0096</td>
<td>40506</td>
</tr>
<tr>
<td></td>
<td>(0.0020)</td>
<td>(0.0017)</td>
<td>(0.0022)</td>
<td>(0.0021)</td>
<td>(0.0021)</td>
<td>(0.0019)</td>
<td></td>
<td></td>
<td>16038</td>
</tr>
</tbody>
</table>

The bold values indicate the log-likelihood that is being maximized. In parentheses are standard errors.
increases from 0.0150 to 0.0156 (returns and VIX used). However, the equity risk premium is not significantly different from zero no matter which data set is used in the EGARCH model. It can also be seen that in all models except the CGARCH model the persistence of conditional variance, $\beta_1$, increases. This helps to raise the long-run variance of the risk-neutral SR-SARV(1) process in together with the increase in $\lambda$. The log-likelihood values indicate that more weights are attached to the VIX data when both returns and VIX are used, which will also be confirmed by Table 2.

With the estimates of parameters, we can figure out the conditional variance and compute the corresponding GARCH implied VIX. Table 2 shows how the implied VIX fits the CBOE VIX in various respects. The first five columns report the statistics of the errors. The sixth column is a $t$-test of whether their means are significantly different. The seventh column shows the probability that the daily implied VIX lies out of one standard deviation of the CBOE VIX, and the eighth column is their linear correlation coefficient. It is not surprising that the models estimated using VIX data only fit the market VIX much better than those estimated using returns only. The models estimated using both returns and VIX are in between but much closer to the former ones, which is consistent with the log-likelihood values in Table 1.

When only returns are considered, although the correlation coefficients between the implied VIX and the CBOE VIX are very high, ranging from 0.92 to 0.94 for different GARCH models, the implied VIX is significantly lower than the market VIX with very low $P$-values in Table 2. The mean errors (CBOE VIX minus implied VIX) are very close among the five GARCH model, with the minimum 3.47 of the AGARCH model and the maximum 3.78 of the TGARCH model. The EGARCH model attains the minimum mean absolute error of 3.76, and the TGARCH model has a maximum of 4.08. In terms of the root mean squared error, the AGARCH model gets the lowest value of 4.73 and the TGARCH model performs worst with a value of 4.98. It is important to note that the magnitude of mean error between the CBOE VIX and the implied VIX, ranging from 3.47 to 3.78, is almost consistent with that of variance premium in standard deviation unit, which is about 3.3. Thus, the GARCH implied VIX undervalues the CBOE VIX by an amount close to the variance premium. Figure 1 shows the trends of the CBOE VIX and the implied VIX of the five models estimated using returns only.

When the VIX data is considered, the difference between the market VIX and the implied VIX drops substantially. The EGAECCH model performs best, achieving a mean error near zero, a standard error 2.73, a mean absolute error 2.10, and a root mean squared error 2.73. Figure 2 shows the trends of the CBOE VIX and the implied VIX of the five models estimated using both returns and VIX.

10 In the AGARCH model, when only VIX time series is used, $\lambda$ and $\theta$ play the same role, and only $\lambda + \theta$ matters. So the model is not identifiable and degenerates to the regular GARCH model.

11 Note that even a negative value is reported when only VIX data is used.
Table 2 Model fit: VIX levels and statistical properties

<table>
<thead>
<tr>
<th>Model &amp; Data</th>
<th>ME</th>
<th>Std.Err.</th>
<th>MAE</th>
<th>MSE</th>
<th>RMSE</th>
<th>P-value</th>
<th>Violation of one-sigma band</th>
<th>Corr.Coeff.</th>
<th>AR1</th>
<th>AR10</th>
<th>AR30</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Returns</td>
<td>3.63</td>
<td>3.31</td>
<td>4.02</td>
<td>24.13</td>
<td>4.91</td>
<td>0.0000</td>
<td>7.86%</td>
<td>0.9943</td>
<td>0.9355</td>
<td>0.7761</td>
<td>70.64</td>
<td>3.10</td>
<td>17.06</td>
<td></td>
</tr>
<tr>
<td>VIX</td>
<td>0.12</td>
<td>3.08</td>
<td>2.36</td>
<td>9.51</td>
<td>3.08</td>
<td>0.4813</td>
<td>1.61%</td>
<td>0.9961</td>
<td>0.9551</td>
<td>0.8221</td>
<td>65.23</td>
<td>3.27</td>
<td>17.66</td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>0.26</td>
<td>3.23</td>
<td>2.39</td>
<td>10.47</td>
<td>3.24</td>
<td>0.1256</td>
<td>2.30%</td>
<td>0.9967</td>
<td>0.9554</td>
<td>0.8162</td>
<td>66.80</td>
<td>3.27</td>
<td>17.95</td>
<td></td>
</tr>
<tr>
<td>TGARCH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Returns</td>
<td>3.78</td>
<td>3.24</td>
<td>4.08</td>
<td>24.78</td>
<td>4.98</td>
<td>0.0000</td>
<td>8.27%</td>
<td>0.9901</td>
<td>0.9096</td>
<td>0.7358</td>
<td>69.13</td>
<td>3.22</td>
<td>17.89</td>
<td></td>
</tr>
<tr>
<td>VIX</td>
<td>0.10</td>
<td>3.06</td>
<td>2.33</td>
<td>9.35</td>
<td>3.06</td>
<td>0.5324</td>
<td>1.54%</td>
<td>0.9961</td>
<td>0.9564</td>
<td>0.8303</td>
<td>65.49</td>
<td>3.22</td>
<td>17.14</td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>0.27</td>
<td>3.08</td>
<td>2.31</td>
<td>9.57</td>
<td>3.09</td>
<td>0.1055</td>
<td>1.60%</td>
<td>0.9960</td>
<td>0.9548</td>
<td>0.8267</td>
<td>67.54</td>
<td>3.23</td>
<td>17.23</td>
<td></td>
</tr>
<tr>
<td>AGARCH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Returns</td>
<td>3.47</td>
<td>3.22</td>
<td>3.79</td>
<td>22.39</td>
<td>4.73</td>
<td>0.0000</td>
<td>7.44%</td>
<td>0.9988</td>
<td>0.9106</td>
<td>0.7296</td>
<td>72.82</td>
<td>3.29</td>
<td>18.73</td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>0.26</td>
<td>3.08</td>
<td>2.34</td>
<td>9.56</td>
<td>3.09</td>
<td>0.1180</td>
<td>1.76%</td>
<td>0.9957</td>
<td>0.9518</td>
<td>0.8134</td>
<td>67.42</td>
<td>3.28</td>
<td>17.88</td>
<td></td>
</tr>
<tr>
<td>CGARCH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Returns</td>
<td>3.66</td>
<td>3.06</td>
<td>3.91</td>
<td>22.77</td>
<td>4.77</td>
<td>0.0000</td>
<td>6.71%</td>
<td>0.9941</td>
<td>0.9428</td>
<td>0.8236</td>
<td>64.45</td>
<td>2.67</td>
<td>13.38</td>
<td></td>
</tr>
<tr>
<td>VIX</td>
<td>0.23</td>
<td>2.84</td>
<td>2.18</td>
<td>8.09</td>
<td>2.84</td>
<td>0.1647</td>
<td>1.17%</td>
<td>0.9927</td>
<td>0.9361</td>
<td>0.8028</td>
<td>62.01</td>
<td>2.93</td>
<td>15.49</td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>0.25</td>
<td>3.03</td>
<td>2.28</td>
<td>9.26</td>
<td>3.04</td>
<td>0.1274</td>
<td>1.78%</td>
<td>0.9923</td>
<td>0.9378</td>
<td>0.8220</td>
<td>64.21</td>
<td>2.98</td>
<td>15.35</td>
<td></td>
</tr>
<tr>
<td>EGARCH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Returns</td>
<td>3.62</td>
<td>3.12</td>
<td>3.76</td>
<td>22.81</td>
<td>4.78</td>
<td>0.0000</td>
<td>8.08%</td>
<td>0.9889</td>
<td>0.9068</td>
<td>0.7457</td>
<td>47.16</td>
<td>2.19</td>
<td>10.92</td>
<td></td>
</tr>
<tr>
<td>VIX</td>
<td>0.00</td>
<td>2.73</td>
<td>2.10</td>
<td>7.45</td>
<td>2.73</td>
<td>0.9964</td>
<td>0.75%</td>
<td>0.9953</td>
<td>0.9510</td>
<td>0.8295</td>
<td>63.13</td>
<td>2.17</td>
<td>10.53</td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>0.09</td>
<td>2.73</td>
<td>2.10</td>
<td>7.48</td>
<td>2.73</td>
<td>0.5748</td>
<td>0.71%</td>
<td>0.9949</td>
<td>0.9475</td>
<td>0.8203</td>
<td>64.04</td>
<td>2.18</td>
<td>10.65</td>
<td></td>
</tr>
<tr>
<td>CBOE VIX</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.9844</td>
<td>0.9162</td>
<td>0.7846</td>
<td>70.65</td>
<td>2.06</td>
<td>10.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This table shows how the implied VIX fits the CBOE VIX in levels as well as other statistical properties for the five GARCH models investigated during the time period from January 2, 1990 to August 8, 2009. The error is calculated as the CBOE VIX minus the implied VIX. The mean error (ME) calculates the daily average error between the implied VIX and the CBOE VIX. The standard error (Std.Err.) calculates the standard deviation of the error. The mean absolute error (MAE) calculates the daily average absolute error between the implied VIX and the CBOE VIX. The mean squared error (MSE) calculates the daily average squared error between the implied VIX and the CBOE VIX. The root mean squared error (RMSE) calculates the square root of the mean squared error. The P-value is for the null hypothesis that the means of the implied VIX and the CBOE VIX are equal. Violation of one-sigma band represents the probability that the implied VIX lies out of the one-standard-deviation band of the CBOE VIX. The correlation coefficient (Corr. Coef.) calculates the linear correlation between the implied VIX and the CBOE VIX. Autocorrelation coefficients with lag of 1, 10, and 30 days and higher moments are reported.
Figure 1 The comparison between the implied VIX and the CBOE VIX (estimated with returns).

The last six columns of Table 2 are autocorrelations and higher moments of the implied VIX with the statistics of the CBOE VIX displayed at the bottom. With VIX data considered, the autocorrelations of the fitted VIX in all the models except the CGARCH model increase. In general, the autocorrelations of the fitted VIX are much higher than the market data. In terms of moments, only the EGARCH model can fit the skewness and kurtosis well but its fitted variance is relatively lower.

In sum, if the VIX data is not considered, the implied VIX is significantly lower than the CBOE VIX for all the models. If the VIX data is considered, the parameters are “distorted” to fit the levels of VIX such that the GARCH, TGARCH (VIX used), and CGARCH models report too high equity risk premium, and the EGARCH
model does not report a significant equity risk premium. Moreover, they still cannot fit the autocorrelations and higher moments well.

5 ANALYSIS OF THE MODEL SPECIFICATION

The empirical results show that the GARCH option pricing model under the LRNVR cannot capture the variance premium. We will illustrate this point with the case of AGARCH(1,1) model. Under the physical probability measure $P$, the model is specified as

$$
\ln \frac{X_t}{X_{t-1}} = r + \lambda \sqrt{h_t} - \frac{1}{2} h_t + \sqrt{h_t} \eta_t,
$$

(25)
\[ h_t = \alpha_0 + \alpha_1 h_{t-1} (\nu_{t-1} - \theta)^2 + \beta_1 h_{t-1}, \quad (26) \]

where \( \nu_t \) is standard normal, conditional on the information at time \( t-1; \alpha_0 > 0, \alpha_1 \geq 0, \beta_1 \geq 0 \) and \( (1 + \theta^2) \alpha_1 + \beta_1 < 1 \) for a covariance stationary process. A positive \( \theta \) can capture the negative correlation between the return and the conditional variance as

\[
\text{Cov}^P(\nu_t, h_{t+1}) = -2\theta \alpha_0 \alpha_1 \left[ 1 - (1 + \theta^2) \alpha_1 - \beta_1 \right]^{-1}. \quad (27)
\]

If \( \theta = 0 \), this AGARCH model degenerates to a linear GARCH(1,1) model discussed in Duan (1995), where the return and the conditional variance are uncorrelated. Under the LRNVR \( Q \), the prices evolve in a risk neutral world

\[
\ln \frac{X_t}{X_{t-1}} = r - \frac{1}{2} h_t + \sqrt{h_t} \epsilon_t, \quad (28)
\]

\[
h_t = \alpha_0 + \alpha_1 h_{t-1} (\epsilon_{t-1} - \theta^*)^2 + \beta_1 h_{t-1}, \quad (29)
\]

where \( \epsilon_t \) is standard normal under the LRNVR \( Q \), conditional on the information at time \( t-1 \), and \( \theta^* = \theta + \lambda \).

Duan (1996, 1997) studied the diffusion limit of the GARCH model. Divide each time period ("day") into \( n \) subperiods of width \( s = 1/n \). For \( k = 1, 2, \ldots, n \), an approximating process is constructed as

\[
\ln X^{(n)}_{k,s} = \ln X^{(n)}_{(k-1)s} + \left( r + \lambda \sqrt{h^{(n)}_{k,s}} \left( \frac{1}{2} h^{(n)}_{k,s} \right) + \sqrt{h^{(n)}_{k,s}} \sqrt{s} \right) s + \sqrt{h^{(n)}_{k,s}} \sqrt{s} \nu_k, \quad (30)
\]

\[
h^{(n)}_{(k+1)s} - h^{(n)}_{ks} = \alpha_0 s + h^{(n)}_{ks} [\alpha_1 q + \beta_1 - 1] s + h^{(n)}_{ks} \alpha_1 \sqrt{s} \left[ (\nu_k - \theta)^2 - q \right], \quad (31)
\]

where \( \nu_k, k = 1, 2, \ldots, \) is a sequence of i.i.d. standard normal random variables; \( q = 1 + \theta^2 \). And the corresponding process under the LRNVR \( Q \) is

\[
\ln X^{(n)}_{ks} = \ln X^{(n)}_{(k-1)s} + \left( r - \frac{1}{2} h^{(n)}_{ks} \right) s + \sqrt{h^{(n)}_{ks}} \sqrt{s} \epsilon_k, \quad (32)
\]

\[
h^{(n)}_{(k+1)s} - h^{(n)}_{ks} = \alpha_0 s + h^{(n)}_{ks} [\alpha_1 q + \beta_1 - 1] s + h^{(n)}_{ks} \alpha_1 \sqrt{s} \left[ (\epsilon_k - \theta - \lambda \sqrt{s})^2 - q \right], \quad (33)
\]

where \( \epsilon_k = \nu_k + \lambda \sqrt{s}, k = 1, 2, \ldots, \) is a sequence of i.i.d. standard normal random variables under the LRNVR \( Q \).

Duan shows that the limiting diffusion process of the approximating process \( 30 \) and \( 31 \) of AGARCH(1,1) under the physical measure \( P \) is

\[
d\ln X_t = \left( r + \lambda \sqrt{h_t} - \frac{1}{2} h_t \right) dt + \sqrt{h_t} dW_t, \quad (34)
\]
HAO & ZHANG | GARCH Option Pricing Models

\[
dh_t = \left[ \alpha_0 + \left( \alpha_1 q + \beta_1 - 1 \right) h_t \right] dt - 2 \theta \alpha_1 h_t dW_{1t} + \sqrt{2} \alpha_1 h_t dW_{2t},
\]

where \(dW_{1t}\) and \(dW_{2t}\) are independent Wiener processes. And the limiting diffusion process of the approximating process (32) and (33) of AGARCH(1,1) under the LRNVR \(Q\) is

\[
d\ln X_t = \left( r - \frac{1}{2} h_t \right) dt + \sqrt{h_t} dZ_{1t},
\]

\[
dh_t = \left[ \alpha_0 + \left( \alpha_1 q + \beta_1 - 1 + 2 \lambda \alpha_1 \theta \right) h_t \right] dt - 2 \theta \alpha_1 h_t dZ_{1t} + \sqrt{2} \alpha_1 h_t dZ_{2t},
\]

where \(dZ_{1t}\) and \(dZ_{2t}\) are independent Wiener processes. In particular, \(dZ_{1t} = dW_{1t} + \lambda dt, dZ_{2t} = dW_{2t}\).

We have two comments on the LRNVR proposed by Duan (1995) and the diffusion limit properties of the GARCH model.

Firstly, for the true bivariate diffusion model, a price of volatility risk is usually introduced to the volatility process when we move from the physical measure to the risk neutral measure. This is because volatility has its own risk that has to be compensated. However, the diffusion limit of the GARCH model under the LRNVR \(Q\) does not reflect this kind of compensation for volatility risk. Actually, under the probability measure change from the physical measure to the LRNVR \(Q\), the innovation in volatility process is invariant, \(dZ_{2t} = dW_{2t}\). The failure of the diffusion limit to incorporate a price of volatility risk results from the inability of the GARCH option pricing model to account for the volatility risk.

Secondly, the presence of the equity premium \(\lambda\) under the LRNVR \(Q\) in the volatility process of both the GARCH model (29) and its diffusion limit (37) does not represent incorporation of a variance premium. Consider the linear GARCH(1,1) with \(\theta = 0\). As shown in the diffusion limit of the volatility process (37) under the LRNVR \(Q\), the term containing the equity premium \(\lambda\) will disappear. Some may argue that the variance premium derives from the negative correlation between the price and the volatility, and the absence of a variance premium in the linear GARCH is because they are uncorrelated as shown in (29). However, even the price process and the volatility process are uncorrelated in bivariate diffusions, a price of volatility risk will still show up under the risk neutral measure. Moreover, literatures show that very little of the volatility risk premium can be explained by the market risk or the correlation of volatility with prices. Instead, it may be driven by some other risk factors including jump risk. Thus, the equity risk premium does not simultaneously represent the volatility risk premium.

By now we have theoretically demonstrated that the GARCH option pricing under the LRNVR is not capable of capturing the variance premium. This suggests the inappropriateness of the model now widely used in the GARCH option pricing literature.

In this section, we are trying to investigate the essence of changing probability measures for GARCH models with LRNVR. We accomplish this task by comparing
the continuous-time limits of the GARCH processes under physical and the risk-neutral measures. We demonstrate our logic by using the diffusion limit of GARCH models in Duan (1996, 1997). Similar analysis can be done with Heston and Nandi (2000). By comparing the continuous-time limits instead of GARCH processes themselves, we can easily find out what is happening during the process of changing probability measures. In continuous-time model, we have an adjustment of equity risk premium for the innovation in the return process and another adjustment of variance risk premium for the innovation in the variance process, as demonstrated in Wiggins (1987), Johnson and Shanno (1987), Hull and White (1982), Scott (1987), and Heston (1993). However, when we look at the continuous-time limits of the GARCH processes, we find that there is only one risk adjustment for the innovation in returns, and the innovations in the variance process are the same under both physical and risk-neutral measures. We then argue that there is no risk adjustment for the variance risk when changing the GARCH process from physical to the risk-neutral measures.

6 CONCLUSION

In this article, we follow the GARCH option pricing model of Duan (1995) and calculate the VIX squared as the expected arithmetic average of the conditional variance over the next 21 trading days under the LRNVR. GARCH implied VIX formulas are derived for a class of square-root stochastic autoregressive volatility (SR-SARV) models.

Numerical results for five specific GARCH models are obtained. We use the time series of the closing price of S&P 500 index and the CBOE VIX to run the maximum likelihood estimation of the GARCH models. The corresponding VIX time series are then calculated. The comparison of the GARCH implied VIX with the CBOE VIX shows that the GARCH implied VIX is significantly and consistently lower than the CBOE VIX when only returns are used for estimation. Moreover, the magnitude of the difference is coincident with the empirical variance premium. When the CBOE VIX is used for estimation, the parameters are distorted to match the levels of VIX, but the implied VIX is still unable to fit the statistical properties of the CBOE VIX. This indicates that the GARCH option pricing under the LRNVR cannot price volatility properly.

With the case of AGARCH(1,1), we illustrate the reason that the GARCH option pricing model under the LRNVR fails to fit the CBOE VIX. Comparing the diffusion limits of the GARCH process under the physical measure and the LRNVR, we find that the innovation of volatility is invariant with respect to the change of probability measure. Moreover, we point out that the equity risk premium cannot serve to capture the variance premium, which is usually misunderstood in the literature. Therefore, the GARCH option pricing model under the LRNVR does not incorporate a premium for the volatility risk.
The empirical results and the theoretical arguments both indicate that the GARCH option pricing model under the LRNVR is not capable of capturing the variance premium. This suggests that the LRNVR is not completely specified, and kind of fully risk neutral measure for the GARCH option pricing is called for. It is an interesting topic to establish an equilibrium GARCH option pricing model that is able to incorporate Christoffersen, Heston, and Jacobs’s (2011) variance-dependent pricing kernel. Zhang, Zhao, and Chang’s (2012) production-based general equilibrium is a possible setup to develop the model.

APPENDIX

A. PROOFS

Proof of Proposition 1. For \( k \geq 1 \),

\[
\begin{align*}
E_t^Q (f_{t+k}) &= e^t E_t^Q (F_{t+k}) = e^t E_t^Q \left( \Omega + \Gamma F_{t+k-1} + V_t + k \right) \\
&= e^t E_t^Q \left[ \Omega + E_t^Q \left( \Gamma F_{t+k-1} + V_t + k \right) \right] \\
&= e^t E_t^Q \left( \Omega + \Gamma F_{t+k-1} \right) \\
&= e^t \left[ \Omega + \Gamma E_t^Q (F_{t+k-1}) \right],
\end{align*}
\]

(A1)

and continuing this iterating process, we have

\[
E_t^Q (f_{t+k}) = e^t \left( \sum_{i=0}^{k-1} \Gamma^i \Omega + \Gamma^k F_t \right).
\]

(A2)

With \( Vix_t = \frac{1}{n} \sum_{k=1}^{n} E_t^Q (h_{t+k}) = \frac{1}{n} \sum_{k=0}^{n-1} E_t^Q (f_{t+k}) \), we have

\[
Vix_t = \zeta + \Psi F_t,
\]

(A3)

with

\[
\begin{align*}
\zeta &= e^t \sum_{k=1}^{n-1} \sum_{i=0}^{k-1} \Gamma^i \Omega, \\
\Psi &= e^t \sum_{k=1}^{n-1} \Gamma^k,
\end{align*}
\]

which is affine in \( F_t \).
For $p = 1$, we can get $\text{VIX}_t$ as a linear function of the conditional variance of the next period, $f_t$,

$$\text{VIX}_t = \zeta + \psi f_t,$$

where

\begin{equation}
\zeta = \frac{\Omega}{1-\Gamma}(1-\psi),
\end{equation}

\begin{equation}
\psi = \frac{1-\Gamma^n}{n(1-\Gamma)}.
\end{equation}

Proof of Proposition 2. Let $u_t = \xi_t / \sqrt{h_t}$. The results for GARCH(1,1) and AGARCH(1,1) have been proved in Proposition 3.2 of [Meddahi and Renault (2004)]. Following the same idea, we rewrite the first three models as $h_t = \omega + \gamma h_{t-1} + \epsilon_{t-1}$ with:

GARCH(1,1): $\omega = \alpha_0$, $\gamma = \alpha_1 (1 + \lambda^2) + \beta_1$, $\epsilon_{t-1} = \alpha_1 h_{t-1} (u_{t-1}^2 - 1 - 2\lambda u_{t-1})$. TARCH(1,1): $\omega = \alpha_0$, $\gamma = \alpha_1 (1 + \lambda^2) + \beta_1 + \theta S$, $\epsilon_{t-1} = \alpha_1 h_{t-1} (u_{t-1}^2 - 1 - 2\lambda u_{t-1}) + \theta h_{t-1} [u_{t-1} - \lambda]^2 (u_{t-1} < \lambda) - S$, where $S = E^{\gamma}_{t-1} (u_{t-1} - \lambda)^2 (u_{t-1} < \lambda)$. AGARCH(1,1): $\omega = \alpha_0$, $\gamma = \alpha_1 [1 + \lambda^2] + \beta_1$, $\epsilon_{t-1} = \alpha_1 h_{t-1} (u_{t-1}^2 - 1 - 2\lambda + \theta u_{t-1})$.

We can rewrite the CGARCH(1,1) model as $h_t = \epsilon_t F_t$ and $F_t = \Omega + \Gamma F_{t-1} + V_{t-1}$ with: $\epsilon_t = (0, 1)$, $F_t = \left(\frac{h_t}{\psi_{t}}\right)$, $\Omega = \alpha_0 \left(\frac{1}{1}\right)$, $\Gamma = \left(\frac{1}{1}\right)$, $F_t = h_{t-1} (u_{t-1}^2 - 2\lambda u_{t-1} - 1)$ $\left(\frac{\phi + \alpha_1}{\phi}\right)$. Since $E^{\gamma}_{t-1} (u_{t-1}) = 0$ and $E^{\gamma}_{t-1} (u_{t-1}^2) = 1$, we have $E^{\gamma}_{t-1} (\epsilon_{t-1}) = 0$ and $E^{\gamma}_{t-1} (V_{t-1}) = 0$. Thus, $\xi_t$ is a SR-SARV(1) for the first three models and a SR-SARV(2) for the CGARCH model.

Proof of Proposition 3. Denote $\epsilon^i = E_i^Q (\epsilon^i_{t+i+1}) = \xi_i$. Under the LRNVR $Q$, the expectation of the conditional variance $k \geq 1$ periods ahead can be expressed as

\begin{equation}
E_t^Q (\xi_{t+k}) = E_t^Q \left[ E_{t+k-1}^Q (\epsilon_{t+k-1}^{i+1} \epsilon_{t+k}^{i+1}) \right]
= E_t^Q (\epsilon_{t+k-1}^{i+1} E_{t+k-1}^Q (\epsilon_{t+k-1}^{i+1}))
= E_t^Q (\epsilon_{t+k-1}^{i+1} E_{t+k-1}^Q (\epsilon_{t+k}^{i+1}))
= E_t^Q (\epsilon_{t+k-1}^{i+1}).
\end{equation}

For $0 \leq i \leq k - 1$, we have

\begin{equation}
\gamma^{i} \ln \xi_{t+k-i} = \gamma^{i} \omega + \gamma^{i+1} \ln \xi_{t+k-i-1} + \gamma^{i} \epsilon_{t+k-i}.
\end{equation}
Thus,

\[ E_t^Q (f_{t+k-i}) = e^{\gamma_i \omega} E_t^Q (f_{t+k-i+1}) E_{t+k-i-1}^Q (e^{\gamma_i \eta_{i+k-i}}) \]

\[ = \iota_i E_t^Q (f_{t+k-i-1}). \]  

(A7)

Then starting from formula (A5) and iterating with formula (A7), we have

\[ E_t^Q (f_{t+k}) = f_t^{\gamma_k} \prod_{i=0}^{k-1} \iota_i. \]  

(A8)

And the implied VIX formula is

\[ \text{VIX}_t = \frac{1}{n} \left[ f_t + \sum_{k=1}^{n-1} \left( \prod_{i=0}^{k-1} f_t^{\eta_i} \right) \right]. \]  

(A9)

B. IMPLIED VIX FORMULAS

Substituting their parameters into the general formula (11), we get the following VIX formulas:

**GARCH(1,1):**

\[ \text{VIX}_t = A + B h_{t+1}, \]  

(A10)

where

\[ A = \frac{\alpha_0}{1 - \eta} (1 - B), \]
\[ B = \frac{1 - \eta^n}{n(1 - \eta)}, \]
\[ \eta = \alpha_1 (1 + \lambda^2) + \beta_1. \]

**TGARCH(1,1):**

\[ \text{VIX}_t = C + Dh_{t+1}, \]  

(A11)

where

\[ C = \frac{\alpha_0}{1 - \eta} (1 - D), \]
\[ D = \frac{1 - \eta^n}{n(1 - \eta)}. \]
\[ \eta = \alpha_1 \left( 1 + \lambda^2 \right) + \beta_1 + \theta S. \]

If \( u_t = \xi_t / \sqrt{h_t} \) follows i.i.d. \( N(0,1) \), \( S = \left[ \sqrt{2 \pi} e^{-\lambda^2 / 2} + (1 + \lambda^2) N(\lambda) \right]. \)

**AGARCH(1,1):**

\[ \text{Vix}_t = E + F h_{t+1}, \]  \hspace{1cm} \text{(A12)}

where

\[ E = \frac{\alpha_0}{1 - \eta} (1 - F), \]
\[ F = \frac{1 - \eta^2}{\eta(1 - \eta)}, \]
\[ \eta = \alpha_1 \left[ 1 + (\lambda + \theta)^2 \right] + \beta_1. \]

**CGARCH(1,1):**

For CGARCH(1,1), we cannot give an explicit formula, but we can refer to equations (7) and (A2) to get the result numerically.

**EGARCH(1,1):**

For EGARCH(1,1), the \( \eta \) in (A2) is given by

\[ \eta = e^{\beta_1 \left( \alpha_0 - \kappa \right) \sqrt{2 / \pi}} \left\{ e^{-\beta_1 \left( \alpha_1 - \kappa \right) \lambda \sqrt{2 / \pi}} N \left[ \lambda - \beta_1 \left( \alpha_1 - \kappa \right) \right] + e^{-\beta_1 \left( \alpha_1 + \kappa \right) \lambda \sqrt{2 / \pi}} N \left[ \beta_1 \left( \alpha_1 + \kappa \right) - \lambda \right] \right\}. \]  \hspace{1cm} \text{(A13)}

**REFERENCES**


