Variance Risk Premium: Estimation, Term Structure and Equity Risk Premium Predictability

Ye Yue*
Tampere University of Technology

March 11, 2016

Abstract

In this paper, variance risk premium and its term structure are estimated. The risk-neutral variance is extracted from option prices using model-free methodology, and the physical variance is estimated using different GARCH models. The term structure of variance and variance risk premium are specified and the monotonicity of variance term structure curves are analysed. To model daily varying variance term structure curves, I propose a time-varying GARCH model to capture the stylized non-monotonicity by incorporating important macroeconomic news into classical GARCH model. Additionally, I investigate the predictability of equity risk premium with different terms via panel data regression on both VIX term structure data and variance risk premium term structure data. I confirm the prediction power of variance risk premium term structure on excess return term structure. A strong term structure leverage effect of VIX on excess return is also observed.

JEL classification: G12, G13.

Keywords: Variance risk premium; Equity risk premium; Variance Term Structure; GARCH

The research leading to these results has received funding from the European Union Seventh Framework Programme ([FP7/2007-2013] [FP7/2007-2011]) under grant agreement No. 289032.

*Department of Industrial Management, Tampere University of Technology & HPCFinance Project - FP7 Marie Curie ITN, Email: ye.yue@tut.fi & lshxxyueye@gmail.com, Phone: +358 503010130. I would like to thank Prof. Juho Kanniainen and Binhuan Lin who provide matlab codes for two-component GARCH model estimation. I also thank BSS of Aarhus University for allowing me the access to Bloomberg terminal during my secondment.

1. Introduction

Variance or volatility as a measure of risk is a crucial factor in finance, especially in asset pricing and risk management. Whether variance risk is placed a premium has been an interesting and disputable question. In the framework of risk-neutral pricing, there are strong empirical evidences, see Carr and Wu (2009), showing the inconsistency between the conditional variance under physical and risk neutral measure. The gap between them is defined as variance risk premium. Providing insights into variance risk premium is helpful for investors to understand market states and asset dynamics. Variance normally relates to a period, for instance, VIX index is constructed for one month variance. The relation between variance and its period introduces a term structure problem. Sufficiently investigating the term structure of variance risk premium is also beneficial to risk managers and traders relying equity return prediction. Ait-Sahalia, Karaman, and Mancini (2014) find investors would like to pay for larger variance risk premium after market drops. This phenomenon is stronger in short horizons, however significantly persistent in long horizons. Regarding return prediction, we show that the variance risk premium term structure is a robust factor playing important role in forecasting both short term and long term equity excess returns. Feunou, Fontaine, Taamouti, and Tédongap (2014) extract variance risk factors from term structure data and find the estimated variance risk premium not only predicts excess returns but also is countercyclical.

This article contributes to the literature by specifying the variance risk premium and its term structure in the frame of GARCH models, in which the complex shape of term structure curves can be modelled by a GARCH candidate with time varying persistent coefficients driven by market news indicator in a proper way. Another main contribution is that I validate the significant role of variance risk premium and its term structure in forecasting equity risk premium term structure by implementing panel data regression.

1.1. Variance Risk Premium

Apart from defining variance risk premium as a "gap" of variance under different probability measures, it can also be understood in the theory of pricing variance swap rate. There is a link of variance risk premium to variance derivatives, see Carr and Wu (2009). As discussed above, it is also an prominent economic predictor for excess asset return. In this section, I briefly review related definitions and results in the construction of variance risk premium and its application to return prediction in recent literature.

1.1.1. *Realized Variance and Quadratic Variation*

The rigorous definitions of variance risk premium and variance swap rely on realized variance and quadratic variation. I briefly introduce these two definitions in this section. All notations and terms in this section follow Andersen, Bollerslev, Diebold, and Labys (2009a).

Let the log-return of an asset be $r(t,h)$, $r(t,h) = \log(S(t)/S(t-h))$, where $S(t)$ is the asset price at time t. Let $QV(t,h)$ be the quadratic variation of $\log S$ in time interval $[t-h,t]$. If the process of $\log S(t)$ is a general semimartingale, then its Doob-Mayer decomposition is $\log S(t) = \mu(t) + M^C(t) + M^J(t)$, where $\mu(t)$ is a process with finite variation, $M^C(t)$ and $M^J(t)$ are the continuous part and jump part of the local martingale in the decomposition. The quadratic variation

$$QV(t,h) = [M^C, M^C]_t - [M^C, M^C]_{t-h} + \sum_{t-h<s<t} \Delta M^2(s)$$

The realized variance,

$$RV(t,h;n) = \sum_{i=1}^{n} r^2(t-h+i\Delta h, \Delta h)$$

where $\Delta h = h/n$, namely the interval $[t-h,t]$ is divided by $n-1$ equally spaced points. One important property links conditional expectation of quadratic variation to conditional variance of return given information up to time t is

$$\mathbb{E}_t QV(t+1,t) = \mathbb{VAR}_t r(t+1) \quad (1)$$

this will be applied to GARCH models in the following sections, see Andersen et al. (2009a).

1.1.2. *Payoff of Variance Swap*

The variance derivative market is developing at a high speed with sufficient liquidity recent years. Many research focus on pricing variance derivatives, such as variance swap and variance future, see Javaheri, Wilmott, and Haug (2004) and Zhu and Zhang (2007). I present the definitions of variance swap and variance risk premium as the same way in Carr and Wu (2009) and illustrate their relation.
The variance risk premium VRP at time t with maturity T is given by

$$VRP_t = \mathbb{E}_t^{\mathbb{P}} QV(T, T-t) - \mathbb{E}_t^{\mathbb{Q}} QV(T, T-t) \quad (2)$$

and the payoff of realized variance swap (SWP) for the long side is given by

$$SWP_t = RV(T, T-t; n) - \mathbb{E}_t^{\mathbb{Q}} RV(T, T-t; n) \tag{3}$$

where $\mathbb{P}$ is physical measure and $\mathbb{Q}$ is risk-neutral measure. It is observed that

$$VRP_t \approx \mathbb{E}_t^{\mathbb{P}} SWP_t$$

It can be proved RV consistently convergences to QV, which shows the relation between variance swap rate and variance risk premium.

1.1.3. Predictor of Asset Return

In financial literature, another role of variance risk premium is as a predictor for asset excess return. Bollerslev, Gibson, and Zhou (2011) investigate stock return predictability using volatility risk premium. They regress monthly S&P 500 excess returns on volatility risk premium and 29 macro-finance covariates. They claim that volatility risk premium has strong predictive power. Similar results are also presented in Carr and Wu (2009) and Feunou et al. (2014), where the latter considers variance and higher order moments and their term structure. A recent work given by Bollerslev, Xu, and Zhou (2015) shows volatility risk premium also provides effective prediction on dividend growth rates. However, to my best knowledge, no current research investigates the predictability of excess return term structure by applying panel regression technique on variance risk premium term structure data. In this paper, I run panel data regression with term and period fixed effects aiming to illustrate the excess return term structure predictability. Our methodology differs from conventional research on term structure where principal component analysis is often implemented. In principal component analysis, even though the independent components can be specified, the fixed effects still are latent.

1.2. Model-based and Model-free Implied Variance

Model-based implied variance is widely specified in both discrete time models and continuous time models.In discrete time models, regarding conditional variance, GARCH family is a typical representative. GARCH implied variance with different terms can be found in, e.g. Christoffersen, Jacobs, Ornthanalai, and Wang (2008). For continuous time models, stochastic volatility model is a popular class in literature, see e.g. Bollerslev et al. (2011). Ait-Sahalia et al. (2014) considers volatility dynamics as a general continuous time jump process. A more general specification of conditional variance with jumps is given in Carr and Wu

(2009).

The model-free implied variance can be extracted from option data. The development of this theory starts from Madan (1998) to Britten-Jones and Neuberger (2000), who consider continous stochastic volatility, then Jiang and Tian (2005) and Jiang and Tian (2007) extend the results in Britten-Jones and Neuberger (2000) to incorporate jumps in volatility process, see also Carr and Wu (2009).

In this article, I simply apply the methodology in Carr and Wu (2009) to estimate model-free risk-neutral variance. For variance under physical measure, I employ different GARCH models to extract GARCH implied variance. Most model-free implied variance theory is based on continuous time model, we may still apply the a discrete time setting for these models. Because, by equation 1, the consistency of conditional quadratic variation and conditional variance of daily return is guaranteed.

The article is organized as follows. Section 2 reviews the theory of model-free implied variance and analyze the error sources in model-free estimation using option prices. Three GARCH models are also presented in this section including single and multiple component GARCH models and a new time varing GARCH model which has a flexible variance term structure function. Section 3 demonstrates GARCH implied variance and Section 4 presents its term structure formula. Section 5 reports the empirical results for Component GARCH and NewsGARCH models using joint likelihood maximum estimation method. Fitting performance is discussed as well in this section. Section 6 investigates the equity risk premium predictability by sufficiently considering both VIX and variance risk premium and their term structure. Section 7 concludes.

2. Model Free Implied Variance using Market Option Prices

2.1. Theory of Model Free Implied Variance

The technique to extract financial information from option prices has been developing for several decades. Breeden and Litzenberger (1978) in their path-breaking work extract risk-neutral density using European option market prices with different strikes. Madan (1998) develop a formula to replicate any second order differentiable payoff, $f(S_T)$, of underlying asset using out-of-money European option prices

$$\mathbb{E}_t^{\mathbb{Q}} f(S_T) = f(\kappa)B_t + f'(\kappa)[C_t(\kappa) - P_t(\kappa)] + \int_0^{\kappa} f''(K)P_t(K)dK + \int_{\kappa}^{\infty} f''(K)C_t(K)dK \quad (4)$$

where B_t is the value of risk-free asset at time t, $C_t(K)$ and $P_t(K)$ are call and put option risk-neutral prices with strike k and maturity T. For the proof and further information see Madan (1998). The formula 4 is quite general and unrelated to the specification of underlying asset price process. Therefore, the conditional expectation $\mathbb{E}_t^{\mathbb{Q}} f(S_T)$ is called a "model-free" quantity in literature. In particular, the model-free variance, skewness and kurtosis draw more attention and are investigated widely, see Christoffersen, Jacobs, and Chang (2012). Especially in a quite general setting for the dynamics of underlying asset, the risk neutral quadratic variance can be extracted from out-of-money call and put option prices, which is a fundamental theory to this research.

The most interesting and widely adopted application of this model-free methodology is for risk-neutral conditional variance. Because variance on the one hand is a constant parameter or a latent stochastic process in most models; on the other hand, variance is of economic sense as a measure for market risk. Applying formula 4 to $\log(S_t)$ and Itô's formula to underlying asset price process, which is assumed a continuous semimartingale in Madan (1998), the conditional integrated variance is

$$\mathbb{E}_t^{\mathbb{Q}} \int_t^T \sigma^2(\tau) d\tau = 2 \int_0^\infty \frac{C_t(T,K) - C_t(t,K)}{K^2} dK \tag{5}$$

Britten-Jones and Neuberger (2000) consider path continuous stochastic volatility model and show that

$$\mathbb{E}_t^{\mathbb{Q}} \int_t^T (dS_\tau / S_\tau)^2 d\tau = 2 \int_0^\infty \frac{C_t(T,K) - C_t(t,K)}{K^2} dK \tag{6}$$

and Jiang and Tian (2005) extents the underlying asset process to incorporate jumps into log-return process, they obtain the same formula 6. Carr and Wu (2009) generalize the market model and allow jumps in both underlying asset price process and volatility process. In their work, the market includes a zero-coupon bond with 1 dollar at maturity T, an underlying asset with spot price S_t and forward price F_t, F_t is driven by

$$dF_t = F_{t-}\sigma_{t-} dW_t + \int_{\mathbb{R}^0} F_{t-}(e^x - 1)[\mu(dx, dt) - \mu_t(x) dx dt] \tag{7}$$

where W_t is a standard Brownian motion under risk-neutral measure, $\mathbb{R}^0$ denotes non-zero real line. The risk-neutral conditional variance is given by

$$\mathbb{E}_t^{\mathbb{Q}} \int_t^T (dS_\tau / S_\tau) d\tau = 2 \int_0^\infty \frac{Opt_t^{out}(T,K)}{B_t(T) K^2} dK + \epsilon \tag{8}$$

where the error term

$$\epsilon = -2\mathbb{E}_t^{\mathbb{Q}} \int_t^T \int_{\mathbb{R}^0} (e^x - 1 - x - \frac{x^2}{2}) \nu_t(x) dx dt \tag{9}$$

and $Opt_t^{out}(T, K)$ denotes the value of an out-of-money European option at time t. The proof of formula 8, 9 and further information can be found in appendix of Carr and Wu (2009).

2.2. Numerical Approximation and Errors

The theory of model free conditional variance extracted from market option prices presented in last section signals the success of using information from derivative markets. However, there are still implementation issues in practice we have to consider. Jiang and Tian (2005) analyzed several errors introduced in the numerical approximation of the integral in equation 5.

1. **Truncation error** : Strikes of option are bounded, $K \in [K^{min}, k^{max}]$, where K^{min} and K^{max} are minimum and maximum of strikes written in option contract from market, therefore the integral domain is limited to $[K^{min}, k^{max}]$ rather than $[0, \infty]$.
2. **Discretion error** : This error is generated by numerical integral in truncated domain $K \in [K^{min}, k^{max}]$. Since truncated domain is relative thin and strike sample is sparse, the difference between numerical integral and it theoretical value is significantly large.
3. **Taylor expansion error** : Carr and Wu (2009) and Jiang and Tian (2007) apply Taylor expansion for log function of terminal value of underlying asset until second order. Error is introduced due to dropping higher order residuals.
4. **Spot-Forward error** : Jiang and Tian (2005) consider the case underlying asset is forward price F_t. Instead of using market prices, F_t can be approximated by $S_t/B(t, T)$ adopting spot price S_t and a zero-coupon bond $B(t, T)$ which payoff is 1 dollar at maturity T.
5. **Maturity interpolation error** : This error is introduced in the approximation for certain maturity T, which is unavailable in market. Hence interpolation is needed. One typical treatment is collecting option prices with the most close maturities, T_1 and T_2, $T_1 < T < T_2$, to interpolate the desired conditional variance with maturity T.

Regarding the correction of truncation error, one way is implementing exploration in some format to recover the truncated tails. Andersen and Bondarenko (2007) investigated the truncated conditional variance, which is named after 'Corridor Implied Volatility'(CIV). They investigated the relation of CIV and model-free implied volatility(MFIV). The general way to reduce discretion error is adopting some sophisticated method in numerical integral. The Taylor expansion error and Spot-Forward error are relative small and determined intrinsically by model setting. However, the maturity interpolation error depends on the sparse maturities of option market data rather than interpolation methods. In the work of Carr and Wu (2009), different interpolation methods are applied. They find the approximated value is robust to different interpolation schemes. In this paper, I follow the interpolation-extrapolation method used in Carr and Wu (2009) to estimate the risk-neutral conditional variance with different maturities. The first step is to obtain sufficiently many implied volatility points in a relative large strike interval. Then I interpolate and extrapolate in all 2000 implied volatility points with a strike range is determined by average implied volatility and spot price. In the second step, option prices are calculated using Black-Scholes formula and the 2000 volatility points given other parameters. The last step is to interpolate the estimated conditional variances for certain maturity. This procedure is essential in the construction of VIX index.

2.3. VIX Index and Its Construction

VIX is an index measuring the forthcoming one month variance. It is first proposed in 1993 by the Chicago Broad Option Exchange(CBOE). The calculation is simply based on weighted average of At-the-Monety implied volatilities on S&P 100 options using Black-Scholes formula. The maturity of VIX is 30 calendar days. The new VIX was developed in 2003. The At-the-Monety S&P 100 option prices used in calculation are replaced by Out-of-the-Monety S&P 500 option prices. The VIX calculation formula is also updated based on model-free conditional variance methodology which I present in last section. It is given by

$$VIX = 100\sqrt{\frac{2}{T}\sum_i \frac{\Delta K_i}{K_i^2}e^{RT}Q(K_i) - \frac{1}{T}\left[\frac{F}{K_0}-1\right]^2} \tag{10}$$

where T is the time to maturity, F is forward price of underlying asset S&P 500, and R is risk-free interest rate. K_0 is the first strike below F, K denotes strikes and $Q(K)$ is middle option price with strike K, while $\Delta K_i = \frac{K_{i+1-K_{i-1}}}{2}$. The formula (10) is a discretized version of formula in Carr and Wu (2009), equation (8) with a modified error term. Further details can be found in CBOE (2009). A recent review written by Gonzalez-Perez (2015) summarizes the history of VIX and its many applications. Interested readers may refer to the literature thereby.

The two prominent stylized facts of variance are leverage effect and feedback effect, which reveal the relation between variance and past/future asset returns. There are vast evidences that variance is negatively correlated to asset returns, see Engle and Ng (1993) and

Campbell and Hentschel (1992). Bollerslev, Litvinova, and Tauchen (2006) confirm these two facts using high-frequency data. Monthly returns in daily rolling window(upper panel) and its absolute value(middle panel) of SPX, whose term structure are the same as VIX(lower panel), are plotted in Fig. 1. It can be observed that large negative returns and large VIX arrive almost together and the dynamics of absolute monthly return also shows similarity to the VIX time series. Our intuition is triggered by these observations which imply that conditional variance and its term structure may be good predictors for equity returns.

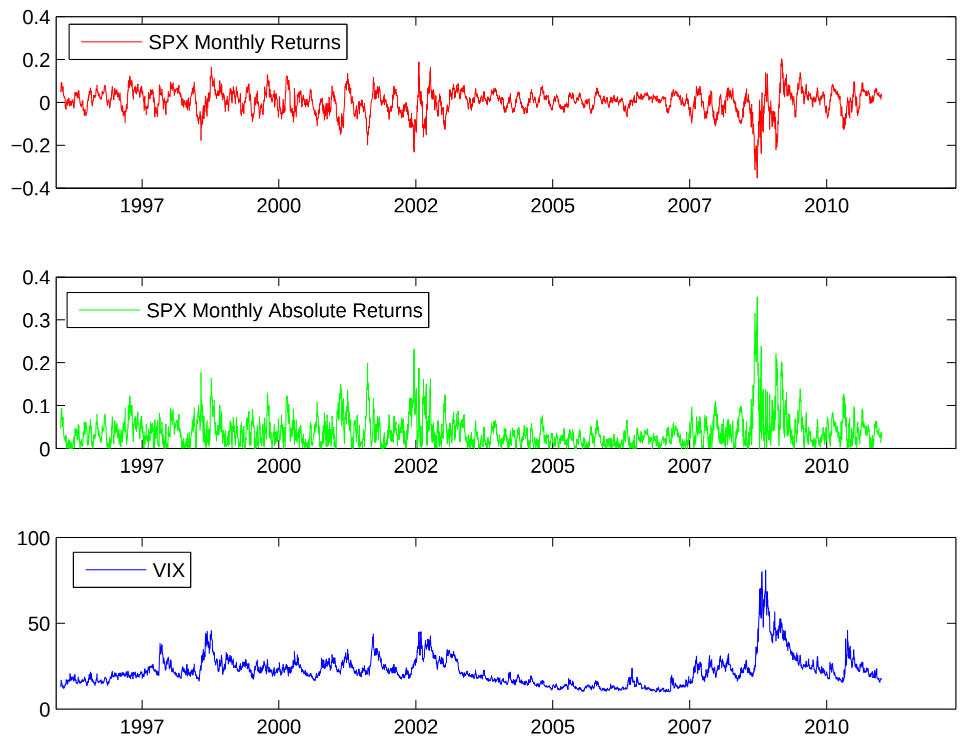

Fig. 1. Monthly Return and VIX

3. GARCH Model Implied Variance

In discrete time setting, the class of GARCH models play an important role in variance modeling. The success of this model family is due to its simple structure and the capability to capture several stylized facts observed in financial markets. For instance, volatility clustering and conditional heavy tail distributed shocks can be properly modeled by GARCH model, more general information on GARCH models can be found in Bollerslev, Chou, and Kroner (1992) and Andersen, Davis, Kreiss, and Mikosch (2009b).

3.1. *Single-Component GARCH Implied Variance*

In this section, I present the standard one-component GARCH model the same as in Christoffersen and Jacobs (2004) with different news impact curves. One component means there is only one variance variable in the variance equation. The model is given by

$$\ln(S_t/S_{t-1}) \equiv R_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \sqrt{h_t}\epsilon_t \tag{11}$$

$$h_{t+1} = \beta_0 + \beta_1 h_t + \beta_2 h_t f(\epsilon_t) \tag{12}$$

where ϵ_t is a standard normal innovation. r is risk-free rate, and λ is equity risk premium measuring the sensitivity of the change in return with respect to the change of variance. β_0, β_1, β_2 are all positive numbers, and additional stationarity condition requires $\beta_1 + \beta_2 < 1$. The market modeled by GARCH model is incomplete, therefore there are different specifications for risk-neutral probability measure. Duan (1995) introduces one specification called locally risk-neutral valuation relationship (LRNVR), which is a widely applied choice of pricing kernel using utility function. The pricing kernel is given by

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|F_t = \exp(-\sum_{i=1}^{t}(\epsilon_i\lambda + \frac{1}{2}h_i\lambda^2)) \tag{13}$$

Then the risk-neutral expected gross return is exactly the gross risk-free return

$$\mathbb{E}^{\mathbb{Q}}[\exp(R_t)|F_{t-1}] = \exp(r) \tag{14}$$

Since the innovation has conditional normal distribution, by Duan's LRNVR, there is no difference between the conditional variance under physical and risk-neutral probability measure.

$$\mathbb{VAR}^{\mathbb{P}}[R_t|F_{t-1}] = \mathbb{VAR}^{\mathbb{Q}}[R_t|F_{t-1}] = h_t \tag{15}$$

where F_t is the market information filtration. This equality of one step conditional variance is not realistic for it implies no risk premium for one step variance. However, there appears a great amount of literature on option pricing using non-normal GARCH models. For example, Christoffersen, Heston, and Jacobs (2006) model the innovation with inverse Gaussian distribution, Badescu, Kulperger, and Lazar (2008) considers normal mixture innovation, and Badescu, Elliott, Kulperger, Miettinen, and Siu (2011) and Chorro, Guégan, and Ielpo (2012) specify the innovation as general hyperbolic type distribution, which includes most common used non-Gaussian distribution. A quite general discussion on non-normal pricing kernel specification is summarized in Christoffersen, Elkamhi, Feunou, and Jacobs (2009).

For simplicity, in this paper I still follows Duan's LRNVR with normal innovation. Hence, the model (11) and (12) under risk-neutral measure are

$$\ln(S_t/S_{t-1}) \equiv R_t = r - \frac{1}{2}h_t + \sqrt{h_t}\epsilon_t^* \tag{16}$$

$$h_{t+1} = \beta_0 + \beta_1 h_t + \beta_2 h_t f(\epsilon_t^* - \lambda) \tag{17}$$

where $\epsilon_t^* = \epsilon_t + \lambda$, is a standard normal innovation under risk-neutral measure.

I consider two popular GARCH models. One is NGARCH model, which is introduced by Engle and Ng (1993) to allow the asymmetric news impact on variance. The dynamics of NGARCH is

$$R_{t+1} = r + \lambda\sqrt{h_{t+1}} - \frac{1}{2}h_{t+1} + \sqrt{h_{t+1}}\epsilon_{t+1} \tag{18}$$

$$h_{t+1} = \beta_0 + \beta_1 h_t + \beta_2 h_t(\epsilon_t - \beta_3)^2 \tag{19}$$

Under the LRNVR, the risk-neutral dynamics are

$$R_{t+1} = r - \frac{1}{2}h_{t+1} + \sqrt{h_{t+1}}\epsilon_{t+1}^* \tag{20}$$

$$h_{t+1} = \beta_0 + \beta_1 h_t + \beta_2 h_t(\epsilon_t^* - \beta_3 - \lambda)^2 \tag{21}$$

The other GARCH model I consider is Heston-Nandi model introduced in Heston and Nandi (2000), which is a discretized version of the well known Heston stochastic volatility model. The HN-GARCH is

$$R_{t+1} = r + \gamma h_{t+1} + \sqrt{h_{t+1}}\epsilon_{t+1} \tag{22}$$

$$h_{t+1} = \beta_0 + \beta_1 h_t + \beta_2(\epsilon_t - \beta_3\sqrt{h_t})^2 \tag{23}$$

Under the LRNVR, the risk-neutral dynamics are

$$R_{t+1} = r - \frac{1}{2}h_{t+1} + \sqrt{h_{t+1}}\epsilon_{t+1}^* \tag{24}$$

$$h_{t+1} = \beta_0 + \beta_1 h_t + \beta_2\left(\epsilon_t^* - \big(\beta_3 + \gamma + \frac{1}{2}\big)\sqrt{h_t}\right)^2 \tag{25}$$

where ϵ_t and ϵ_t^* are independent standard normal innovations under physical and risk-neutral measure respectively.

3.2. Multiple-Component GARCH Implied Variance

One-component GARCH model is widely used in modeling return dynamics under physical measure. However it still lacks of power to generate the "smirk" implied from option data. It means one-component conditional normal GARCH is of week competence in option pricing. Christoffersen et al. (2008) introduce a two-component GARCH model aiming to capture empirical characteristics of volatility term structure. In order to allow the term structure curve to be non-monotonic, a long run component is added in the second conditional moment equation. The mean equation in two-component GARCH model is the same as HN-GARCH, hence I name this component GARCH model by HN-CGARCH

$$\ln(S_t/S_{t-1}) \equiv R_t = r + \lambda\sqrt{h_t} + \sqrt{h_t}\epsilon_t \tag{26}$$

The variance equations are

$$h_{t+1} = q_{t+1} + \tilde{\beta}(h_t - q_t) + \alpha v_{1,t} \tag{27}$$

$$q_{t+1} = \sigma^2 + \rho(q_t - \sigma^2) + \phi v_{2,t} \tag{28}$$

where

$$v_{i,t} = (\epsilon_t^2 - 1) - 2\gamma_i\sqrt{h_t}\epsilon_t, \quad i = 1, 2 \tag{29}$$

ϵ_t is i.i.d normal innovation, q_t is called long run component. The HN-CGARCH model has the following risk-neutral form

$$\ln(S_t/S_{t-1}) \equiv R_t = r - \frac{1}{2}\sqrt{h_t} + \sqrt{h_t}\epsilon_t^* \tag{30}$$

The variance equations are

$$h_{t+1} = q_{t+1} + \tilde{\beta}^*(h_t - q_t) + \alpha v_{1,t} \tag{31}$$

$$q_{t+1} = \sigma^2 + \rho^*(q_t - \sigma^2) + \phi v_{2,t} \tag{32}$$

where

$$v_{i,t} = (\epsilon_t^* - \gamma_i^*\sqrt{h_t})^2 - (1 + \gamma_i^{*2}), \quad i = 1, 2 \tag{33}$$

$$\epsilon_t^* = \epsilon_t + (\lambda + \frac{1}{2})\sqrt{h_t} \tag{34}$$

$$\gamma_i^* = \gamma_i + \lambda + \frac{1}{2} \tag{35}$$

$$\tilde{\beta}^* = \tilde{\beta} + \alpha(\gamma_1^{*2} - \gamma_1^2) + \phi(\gamma_2^{*2} - \gamma_2^2) \tag{36}$$

$$\tilde{\rho}^* = \tilde{\rho} + \alpha(\gamma_1^{*2} - \gamma_1^2) + \phi(\gamma_2^{*2} - \gamma_2^2) \tag{37}$$

A recent work by Bormetti, Corsi, and Majewski (2015) extends two-component GARCH to incorporate multiple components into the conditional variance equation. In their general framework, the components can be any previsible variables including but not necessary conditional variance. However, reasonable economic senses for those additional components are needed to be identified in practice.

3.3. Time-Varying GARCH Implied Variance

Anther way to introduce a non-monotonic VIX term structure is to model the conditional variance as a standard GARCH multiplied by a deterministic function. It is so called time varying GARCH in the literature, e.g.Amado and Teräsvirta (2013).

The model is given by

$$\ln(S_{t+1}/S_t) \equiv R_{t+1} = r + \lambda\sigma_{t+1} - \frac{1}{2}\sigma_{t+1}^2 + \sigma_{t+1}\epsilon_{t+1} \tag{38}$$

$$\sigma_{t+1}^2 = g_{t+1}h_{t+1} \tag{39}$$

$$h_{t+1} = \beta_0 + \beta_1 h_t + \beta_2 h_t f(\epsilon_t) \tag{40}$$

where g_t is a deterministic function with respect to time t.

Under risk neutral measure

$$\ln(S_{t+1}/S_t) \equiv R_{t+1} = r - \frac{1}{2}\sigma_{t+1}^2 + \sigma_{t+1}\epsilon_{t+1}^* \tag{41}$$

$$\sigma_{t+1}^2 = g_{t+1}h_{t+1} \tag{42}$$

$$h_{t+1} = \beta_0 + \beta_1 h_t + \beta_2 h_t f(\epsilon_t^* - \lambda) \tag{43}$$

Where $\epsilon_t^* = \epsilon_t + \lambda$, is a standard normal innovation under risk-neutral measure.

3.4. Time-varying GARCH(1,1) driven by Market News

In order to obtain a VIX term structure curve with flexible shape, the key is to specify the multiplier g_t. One possible choice is to identify g_t as a function of certain classes of news. Intuitively, when important news flows arrive in the market, the change of multiplier function may result in a non-monotonic VIX term structure curve. More realistically, g_t can be relaxed to a function $g_t(News_t, \tau)$, where $\tau = 1, 2, ..., 252$, valued in the forecasting horizon for VIX. $News_t$ is a news set observed at t. On each day t, we have a news driving function $g_t(News_t, \tau)$

The model is

$$\ln(S_{t+1}/S_t) \equiv R_{t+1} = r + \lambda\sigma_{t+1} - \frac{1}{2}\sigma_{t+1}^2 + \sigma_{t+1}\epsilon_{t+1} \tag{44}$$

$$\sigma_{t+1}^2 = g_t(News_t, t+1)h_{t+1} \tag{45}$$

$$h_{t+1} = \beta_0 + \beta_1 h_t + \beta_2 h_t f(\epsilon_t) \tag{46}$$

where g_t is a deterministic function with respect to time t and news information $News_t$.

Under risk neutral measure

$$\ln(S_{t+1}/S_t) \equiv R_{t+1} = r - \frac{1}{2}\sigma_{t+1}^2 + \sigma_{t+1}\epsilon_{t+1}^* \tag{47}$$

$$\sigma_{t+1}^2 = g_t(News_t, t+1)h_{t+1} \tag{48}$$

$$h_{t+1} = \beta_0 + \beta_1 h_t + \beta_2 h_t f(\epsilon_t^* - \lambda) \tag{49}$$

Where $\epsilon_t^* = \epsilon_t + \lambda$, is a standard normal innovation under risk-neutral measure. Since $News_t$ needs to be specified clearly, we introduce a news indicator next section. it is available to calculate using news data.

3.4.1. News Indicator

My intuition to construct news indicator comes from the motivation to change variance term structure. Therefore, it consists two parts. The first part measures the intensity of news, the second part weights the impact of news on realized variance.

Suppose that there are n classes of important news, $\mathbb{N}^i$, $i = 1, 2, ..., n$ observed from market in time horizon $\mathbb{T} = \{1, 2, ..., T\}$. $\#\mathbb{N}_t^i$ is the number of pieces of news in class i on day t. Let $N_t \subseteq \{1, 2, ..., n\}$ be the news index set on day t. Let $\mathbb{T}^i \subseteq \mathbb{T}$ be a set of trading days when news $\mathbb{N}^i$ arrives on. We measure the impact of news i on variance by Kolmogorov-Smirnov Distance between empirical distributions of realized variance on days in $\mathbb{T}^i + 1$ and on days in $\mathbb{T} \setminus \mathbb{T}^i + 1$.

$$D_t^{KS,i} = KS(\hat{F}_{RV}^i, \hat{F}_{RV}^{-i})$$

where $\hat{F}_{RV}^i$ is the empirical distribution of realized variance using subsample $\mathbb{T}^i + 1$ containing all the following days after news i arrives. Whereas $\hat{F}_{RV}^{-i}$ denotes the empirical distribution of realized variance using subsample $\mathbb{T} \setminus \mathbb{T}^i + 1$.

$$\mathbb{I}_t^{News} = \omega \frac{|\sum_{i \in N_t} \#\mathbb{N}_t^i - \sum_{i \in N_{t-1}} \#\mathbb{N}_{t-1}^i|}{\sum_{i \in N_{t-1}} \#\mathbb{N}_{t-1}^i} + (1-\omega) \frac{\sum_{i \in N_t} D_i^{KS}}{\sum_{i=1}^{n} D_i^{KS}} \quad (50)$$

where $\omega \in [0, 1]$ is a given weight.

3.4.2. Time-varying News Impact Curves

We specify the time varying news impact curve as a nonnegative polynomial with time varying coefficients driven by news indicator $\mathbb{I}_t^{News}$. This specification provides us sufficient flexibility for modelling the dynamics of VIX term structure curves.

$$g_t(News_t, \tau) = \left([a_1 + a_5 \mathbb{I}_t^{News}]\tau^3 + [a_2 + a_5 \mathbb{I}_t^{News}]\tau^2 + [a_3 + a_5 \mathbb{I}_t^{News}]\tau + a_4 + a_5 \mathbb{I}_t^{News}\right)^2 \quad (51)$$

where $a_i,\ i = 1, 2, 3, 4$ are constant parameters determining the "fundamental shape" of VIX term structure, the shape of VIX term structure curves is time varying due to the impulse of daily changing news indicator $a_5 \mathbb{I}_t^{News}$.

4. Variance Term Structure

4.1. VIX Term Structure

The VIX term structure extracted from option data is plotted in Fig.2. The three dimensional plot demonstrates clearly that short term VIX is more volatile than the long term VIX. More detailed statistics is given in table 1.

It is worthy of noting that non-monotonicity is a stylized fact of daily VIX term structure curves. According to our investigation, the shape is more complex than a monotonic and even humped shape curves. Four typical VIX term structure curves are presented in Fig. 3. Hence the two-component GARCH is still not sufficiently model variance term structure. This is the motivation to introduce time-varying GARCH model.

4.2. GARCH Implied Term Structure

In this section, I deduce the VIX term structure formula of GARCH(1,1), CGARCH(1,1) and NewsGARCG(1,1) and I also analyze monotonicity of implied term structure curves from these three GARCH models.

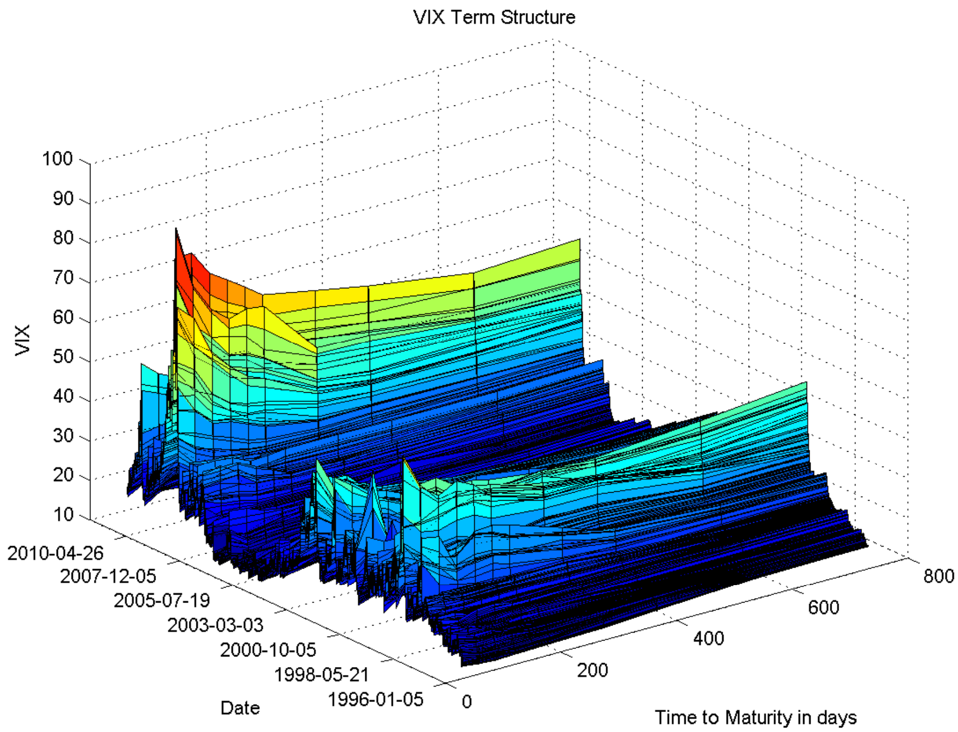

Fig. 2. VIX Term Structure

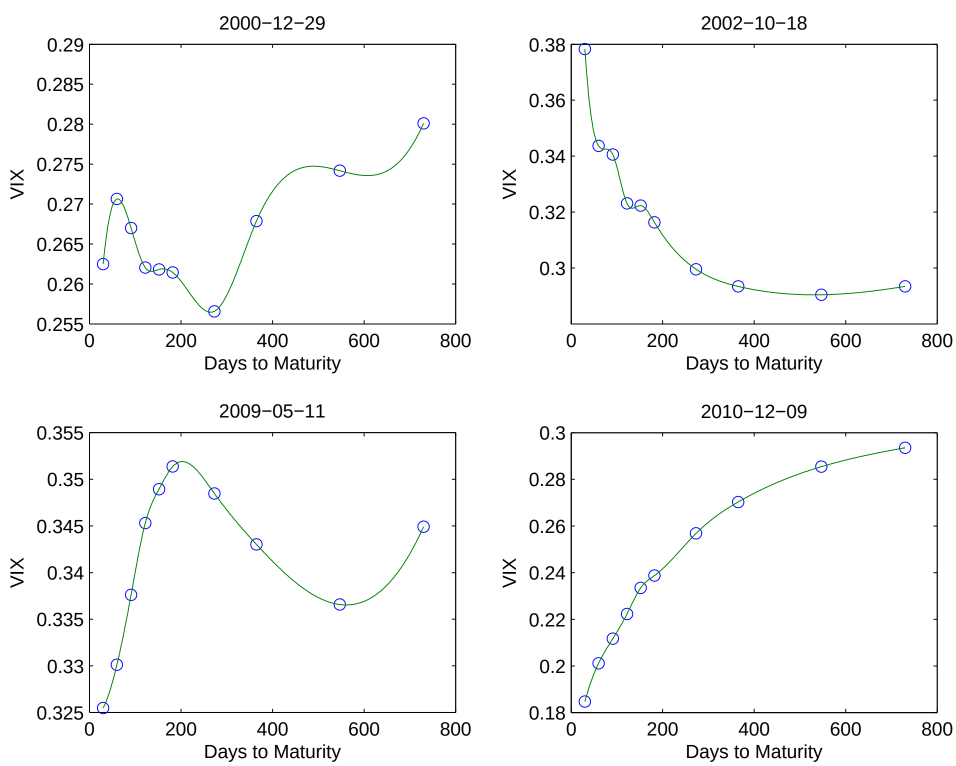

Fig. 3. Typical VIX Term Structure Curves

4.2.1. VIX formula deduced from GARCH(1,1)

Let $M = \beta_1 + \beta_2 \mathbb{E}^{\mathbb{Q}}[f(\epsilon^*)]$. Under stationary condition, the unconditional variance is $\sigma^2 = \frac{\beta_0}{1-M}$

$$\mathbb{E}_t^{\mathbb{Q}}[h_{t+k}] = \beta_0 \sum_{j=0}^{k-1} M^j + M^{k-1}(\beta_1 h_t + \beta_2 f(\epsilon^*)) \tag{52}$$

$$= \beta_0 \frac{1-M^{k-1}}{1-M} + M^{k-1}(\beta_1 h_t + \beta_2 f(\epsilon^*)) \tag{53}$$

$$= \frac{\beta_0}{1-M} + \left(\beta_1 h_t + \beta_2 f(\epsilon^*) - \frac{\beta_0}{1-M}\right) M^{k-1} \tag{54}$$

$$= \sigma^2 + (h_{t+1} - \sigma^2) M^{k-1} \tag{55}$$

The annualized VIX term structure formula is given by

$$VIX_t(\tau) = 100\sqrt{\frac{252}{\tau}\mathbb{E}_t^Q(\sum_{k=1}^{\tau} h_{t+k})} = 100\sqrt{\frac{252}{K}\sum_{k=1}^{\tau}\mathbb{E}_t^Q[h_{t+k}]} \tag{56}$$

$$= 100\sqrt{252\sigma^2 + \frac{252}{\tau}(h_{t+1} - \sigma^2)\frac{1-M^\tau}{1-M}}, \quad \tau \geq 1 \tag{57}$$

4.2.2. VIX formula deduced from Component GARCH(1,1)

$$\mathbb{E}_t^{\mathbb{Q}}[h_{t+k}] = \mathbb{E}_t^{\mathbb{Q}}[q_{t+k}] + \mathbb{E}_t^{\mathbb{Q}}[h_{t+k} - q_{t+k}] \tag{58}$$

$$= \sigma^2 + \rho^{k-1}(q_{t+1} - \sigma^2) + \tilde{\beta}^{k-1}(h_{t+1} - q_{t+1}) \tag{59}$$

$$\mathbb{E}_t^{\mathbb{Q}}[h_{t+k}] = \mathbb{E}_t^{\mathbb{Q}}[q_{t+k}] + \mathbb{E}_t^{\mathbb{Q}}[h_{t+k} - q_{t+k}] \tag{60}$$

$$= \sigma^2 + \rho^{*k-1}(q_{t+1} - \sigma^2) + \tilde{\beta}^{*k-1}(h_{t+1} - q_{t+1}) \tag{61}$$

$$VIX_t(\tau) = 100\sqrt{\frac{252}{\tau}\left[\tau\sigma^2 + (q_{t+1} - \sigma^2)\frac{1-\rho^{*\tau}}{1-\rho^*} + (h_{t+1} - q_{t+1})\frac{1-\tilde{\beta}^{*\tau}}{1-\tilde{\beta}^*}\right]} \tag{62}$$

4.2.3. VIX formula deduced from News GARCH(1,1)

Let $M = \beta_1 + \beta_2 \mathbb{E}^{\mathbb{Q}}[f(\epsilon^* - \lambda)]$. Under stationary condition, $0 < M < 1$, and the unconditional variance σ^2 is $\frac{\beta_0}{1-M}$

The annualized VIX term structure formula is given by

$$VIX_t(\tau) = 100\sqrt{\frac{252}{\tau}\mathbb{E}_t^{\mathbb{Q}}(\sum_{k=1}^{\tau} g_t(\mathbb{I}_t^{News}, k)h_{t+k})} = 100\sqrt{\frac{252}{\tau}\sum_{k=1}^{\tau}\mathbb{E}_t^{\mathbb{Q}}[g_t(\mathbb{I}_t^{News}, k)h_{t+k}]} \quad (63)$$

$$= 100\sqrt{\frac{252}{\tau}\sum_{k=1}^{\tau} g_t(\mathbb{I}_t^{News}, k)\big(\sigma^2 + (h_{t+1} - \sigma^2)M^{k-1}\big)}, \quad \tau \geq 1 \quad (64)$$

It is observed that the shape of VIX term structure curve $VIX_t(K)$ is determined by the multiplier $g_t(\mathbb{I}_t^{News}, k)$.

5. Data, Model Estimation and Calibration Criteria

5.1. Data

The empirical work in this paper relies on the following data. First, the asset prices are S&P 500 composite index in the period of 05 January 1996 to 03 January 2011. Second, to estimate model-free conditional variance, I use prices of European call and put options written on S&P 500 in the same sample period with different strikes and time to maturities. These two data are obtained from the database Option Metrics, which is a reliable high quality financial data source widely used in literature. Additionally, I also download VIX index from the Chicago Board of Options Exchange (CBOE) for a robust comparison for our 30-day model-free variance. Third, the risk-free interest rates I choose are US Treasury Constant Maturity Rates with maturities 1 month, 3 month, 6 month, 1 year, and 2 year. This data is provided by Board of Governors of the Federal Reserve System (US). However the sample period starts from 31 July 2001 to 03 January 2011.

For the construction of news indicator in News GARCH model, I collect 420 types of US macroeconomic announcements from Bloomberg World Economic Calendar. These 420 announcements include most important macroeconomic news, such as publication of GDP, PPI, CPI, national employment report and so on.

Table 1 presents four sample moments of VIX and realized variance term structure extracted from option prices and daily log-returns of S&P 500. It can be observed that the mean value increases as term turns to be longer, while the variation indicated by standard deviation decreases for both VIX and realized variance. This confirms that investors' predicted long term variance are larger than short term. Additionally, opinion of large long term variance holding by investors is consistent and relatively stable. Table 1 also illustrates that realized variances are systemically smaller than model-free variances. This is a typical fact

Table 1: This table reports the sample mean (***Mean***), standard deviation (***Std***), sample skewness (***Skew***) and sample kurtosis (***Kurt***) of VIX and Realized Variance across 10 different terms from 1 month to 24 month. The sample period is from 05 January 1996 to 03 January 2011. Realized Variance is calculated using sum squared daily log-returns of S&P 500. Note that all values are annualized with 252 days.

VIX and RV

Term	**VIX**				**RV**			
	Mean	***Std***	***Skew***	***Kurt***	***Mean***	***Std***	***Skew***	***Kurt***
1	23.2386	9.0770	1.9447	9.7532	21.6872	12.4441	2.7581	14.3750
2	23.6804	8.5618	1.7069	8.3885	22.0170	11.8498	2.6655	13.5355
3	23.9322	8.1687	1.5109	7.1163	22.2132	11.4784	2.4997	11.9554
4	24.1475	7.8402	1.4299	6.6987	22.3555	11.1965	2.3352	10.4944
5	24.2186	7.5797	1.3438	6.2521	22.4789	10.9469	2.2064	9.5116
6	24.2430	7.3522	1.2358	5.6083	22.5918	10.7225	2.0940	8.7689
9	24.3576	7.0506	1.1463	5.0432	22.9713	10.1417	1.7596	6.8654
12	24.6376	6.9709	1.0603	4.6421	23.3515	9.6399	1.4569	5.4637
18	25.0800	6.7819	0.8527	3.7985	23.8406	8.8180	0.9716	3.7990
24	25.5871	6.8347	0.8320	3.6104	24.1336	8.0954	0.6349	3.0286

leading to negative variance risk premium, which is studied and reported in Bollerslev et al. (2011) and Carr and Wu (2009). The sample skewness and kurtosis go lower for both VIX and RV with increasing time to maturity, which provides additional evidence that long term annualized variance is more asymmetrically volatile than short annualized term variance.

5.2. Joint Likelihood Estimation Using VIX Term Structure

VIX as the risk-neutral information sources has been used widely to estimate asset pricing models. For example, for continuous models, Duan and Yeh (2010) estimate a stochastic volatility model with jumps using VIX data. Regarding GARCH models calibration, Kanniainen, Lin, and Yang (2014) estimate three GARCH models with VIX and improve the model performance for option valuation. Hao and Zhang (2013) consider five GARCH models, and conclude that the GARCH implied VIX based on return data is systemically lower than VIX.[1] In this section, I concisely demonstrate a general joint likelihood function for GARCH model using both returns and VIX term structure data.

The joint likelihood function consists of two sub-likelihood functions, one is of the in-

[1]This might be due to the inconsistent ways used in counting days in $[t, T]$, namely from spot time to maturity. The authors use 252 days for one year but 30 days for one month in their paper.

novation in return equation, the other part is of the error between model-free VIX term structure and GARCH implied variance term structure.

$$\ln L = \ln L(\epsilon^R; \Theta) + \ln L(\epsilon^{VIX}; \Theta') \tag{65}$$

$$\epsilon^R_{t+1} = \frac{R_{t+1} - \mathbb{E}_t(R_{t+1})}{h_{t+1}} \tag{66}$$

where ϵ^R_t is $I.I.N(0,1)$ random variable, while ϵ^{VIX} is a zero mean normal vector whose dimension is as same as term structure vector. Θ and Θ' are parameters. In this paper, I consider two types of VIX term structure errors, $\epsilon^{Add-VIX}$ and $\epsilon^{Mul-VIX}$.

$$\epsilon^{Add-VIX}_t = \mathbf{VIX}_t - \mathbf{VIX}^{Model}_t \tag{67}$$

$$\epsilon^{Mul-VIX}_t = \ln(\mathbf{VIX}_t) - \ln(\mathbf{VIX}^{Model}_t) \tag{68}$$

where $\mathbf{VIX}_t$ are model-free VIX term structure market data extracted from option prices at t, $\mathbf{VIX}^{Model}_t$ is the corresponding model implied VIX term structure estimates. By our empirical analysis, the VIX term structure error series are strongly autocorrelated, therefore we allow the autocorrelation in $\epsilon^{Add-VIX}_t$ and $\epsilon^{Mul-VIX}_t$ and specify them as a vector AR(1) process

$$\epsilon^{Add-VIX}_t = \Psi^{Add}\epsilon^{Add-VIX}_{t-1} + e_t \tag{69}$$

$$\epsilon^{Mul-VIX}_t = \Psi^{Mul}\epsilon^{Mul-VIX}_{t-1} + e_t \tag{70}$$

where e_t is a standard normal vector. Ψ is a diagonal matrix controlling the autocorrelation for the VIX error series at each term. Under this setting, the joint likelihood functions is given by

$$\ln L = \ln L(\epsilon^R; \Theta) + \ln L(\epsilon^{VIX}; \Theta') \tag{71}$$

$$= -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{n}\left\{\ln h_t - \frac{R_t - \mathbb{E}_t(R_t)}{h_t}\right\} \tag{72}$$

$$-\frac{nT}{2}\ln(2\pi) - \frac{n-1}{2}\ln|\Sigma - \Psi\Sigma\Psi'| - \frac{1}{2}\sum_{t=2}^{n} e'_t(\Sigma - \Psi\Sigma\Psi')^{-1}e_t \tag{73}$$

where Σ is the contemporaneous variance of ϵ^{VIX}_t. For more details, see Beach and MacKinnon (1979).

I apply this joint estimation strategy to two multiple component GARCH model and

News GARCH. The estimates and standard errors are given in table 2. Strong similarity is displayed within estimates of correlated coefficents ψ in four cases of two models, which allow autocorrelated VIX term structure errors. Firstly, the long-run persistent coefficient ρ is nearly 1, which is higher than the short-run persistent coefficient $\tilde{\beta}$ for HNCGARCH. Secondly, the persistent coefficient estimates of additive VIX term structure error are larger than the multiplicative case. Thirdly, the standard errors of additive case are normally smaller than the corresponding multiplicative case. This implies that the VIX term structure error should not be specified as the log difference between model-free and model based VIX term structure data. Especially, for HNCGARCH model, the risk-premium parameter λ is negative, which conflicts most specification in literature. allowing autocorrelated VIX term structure errors.

5.3. A Remark on Model Calibration Criteria

Due to the complex shape of VIX term structure, more sophisticated GARCH model is demanded to fit risk-neutral variance term structure since market VIX term structure curves loss monotonicity and hump-shape in most trading days. Hence single and two-component GARCH models are not sufficient to model market VIX term structure.

In this section, I list several criteria for modeling term structure curve. Table 3 presents the monotonicity of model-free and CGARCH implied variance term structure curves. Obviously, the number of increasing curves for model-free VIX is much lower than CGARCH implied ones. All the GARCH models estimated in this paper are prone to imply an increasing term structure curve. This misspecification is fairly strong that approximately 800 non-monotonic term structure curves are identified as increasing, which counts half of our total sample size. In fact, most model-free term structure curves are of complex shape, normally two extremum, One maximum and one minimum. The purely decreasing term stucture curves are fewest among the three classes, however, CGARCH models still over specify this case. All in all, from table 3, both two-component GARCH model and News GARCH model are likely to mis-specify the term structure curves to be monotonic. From this point of view, more sophisticated model need to be developed to fit the market data better.

I also consider the performance of two-component HNGARCH and News GARCH estimated with additive and multiplicative error types. The root mean squared error of the difference between model-free VIX data and model based counterpart at each term are plotted in Fig. 4. Generally, RMSE of the model estimated with additive errors are smaller than model with multiplicative errors. This evidence is also illustrated in Fig. 5, where

Table 2: This table presents joint MLE using returns and VIX term structure data for HNCGARCH and NCGARCH model. The VIX term structure errors are specified as both $\epsilon_t^{Add-VIX}$ and $\epsilon_t^{Mul-VIX}$, namely the difference and log-difference of model-free and model based VIX term structure. For each parameter, estimate is in upper line, corresponding standard error is in lower line. $\psi(i)$ is the ith diagonal element in matrix Ψ. The sample period is from 05 January 1996 to 03 January 2011.

Joint MLE using Returns and VIX Term Structure Data

Model	HNCGARCH			NewsGARCH	
Parameters	***Additive Error***	***Multiplicative Error***	***Parameters***	***Additive Error***	***Multiplicative Error***
$\tilde{\beta}$	9.6539E-01	8.8924E-01	b_1	1.1263E+00	1.1716E+00
	7.0371E-04	5.0158E-03		6.6329E-02	1.4866E-01
α	1.5307E-06	2.6475E-06	b_2	-1.2715E+00	-1.6040E+00
	9.5082E-08	3.0699E-07		1.1351E-01	2.5094E-01
γ_1	5.6781E+02	4.2418E+02	b_3	4.6983E-01	6.2303E-01
	3.4242E+01	4.4562E+01		7.4489E-02	1.5179E-01
λ	8.0581E-02	-4.8748E-01	b_4	1.0009E+00	1.2231E+00
	6.1193E-02	8.9829E-02		3.2526E-02	4.0822E-02
ω	2.0209E-07	8.8738E-07	b_5	2.5865E-02	3.2791E-04
	1.0824E-08	4.1361E-08		1.1688E-03	1.3650E-03
ρ	9.9852E-01	9.9587E-01	β_0	1.0697E-06	1.0829E-06
	1.1222E-04	1.8091E-04		7.3381E-08	7.1641E-08
ϕ	2.0679E-06	1.2526E-06	β_1	9.6568E-01	9.7500E-01
	5.1752E-08	6.2903E-08		6.5873E-04	8.1880E-04
γ_2	1.0107E+02	1.1462E+02	β_2	1.1387E-02	2.2796E-02
	4.5474E+00	8.8548E+00		1.9045E-04	8.5126E-04
			λ	1.3507E+00	1.7975E-01
				1.4210E-02	2.2730E-02
$\psi(1)$	7.8075E-01	7.7846E-01	$\psi(1)$	8.3432E-01	8.3954E-01
	1.1179E-03	1.2770E-03		1.3292E-03	1.6470E-03
$\psi(2)$	8.0050E-01	8.2638E-01	$\psi(2)$	8.4252E-01	8.5028E-011
	4.0141E-04	4.3192E-04		7.2088E-04	8.7580E-04
$\psi(3)$	8.2838E-01	8.5337E-01	$\psi(3)$	8.4137E-01	8.5697E-01
	3.2789E-04	2.5946E-04		7.1155E-04	5.8736E-04
$\psi(4)$	8.7162E-01	8.8203E-01	$\psi(4)$	8.4251E-01	8.5913E-01
	2.2991E-04	2.3747E-04		7.7560E-04	8.3626E-04
$\psi(5)$	8.7460E-01	8.8979E-01	$\psi(5)$	8.5287E-01	8.7309E-01
	3.3784E-04	3.3440E-04		5.3493E-04	6.2266E-04
$\psi(6)$	8.7818E-01	8.9723E-01	$\psi(6)$	8.5263E-01	8.8015E-01
	1.9128E-04	2.3982E-04		3.8530E-04	4.2958E-04
$\psi(7)$	8.9648E-01	9.1434E-01	$\psi(7)$	8.6145E-01	8.9276E-01
	2.1760E-04	1.6904E-04		3.1001E-04	1.9414E-04
$\psi(8)$	9.0751E-01	9.2429E-01	$\psi(8)$	8.7373E-01	9.0318E-01
	1.7315E-04	1.3373E-04		3.2510E-04	1.8669E-04
$\psi(9)$	9.1476E-01	9.3641E-01	$\psi(9)$	8.9507E-01	9.2154E-01
	3.1654E-04	2.1760E-04		4.0338E-04	2.9522E-04
$\psi(10)$	9.1633E-01	9.4137E-01	$\psi(10)$	9.1025E-01	9.3084E-01
	4.6820E-04	2.7123E-04		7.4346E-04	5.7721E-04

Table 3: This table presents the number of classified variance term structure curves by their monotonicity. Three classes I consider are increasing, decreasing and non-monotonic for four model based and Model-free (Market) VIX term structure curves. The sample period is from 29 January 2001 to 03 January 2011.

Statistics of Monotonicity of Daily VIX Term Structure Curves

	Monotonicity of Variance Term Structure		
	Incerasing	Decreasing	Non-Monotonic
Market	787	33	1678
Add-HNGARCH	1548	52	898
Multi-HNGARCH	1467	349	682
Add-NewsGARCH	1652	156	690
Multi-NewsGARCH	1949	0	549

I plot the daily RMSE for VIX term structure fitting. The News GARCH model shows better performance of market data fitting than CGARCH, see Fig.6, for polynomial news impact function is of more flexibility than a linear combination of two exponential functions in CGARCH VIX formula.

6. The Predictability of Equity Risk Premium Term Structure

The variance risk premium is an effective factor in the prediction of equity risk premium. Carr and Wu (2009) estimated a realized version of variance risk premium (payoff of realized variance swap) and apply it as a regressor to predict the excess stock returns including both stock indices and individual stocks. They demonstrate the importance of variance risk premium in the prediction of equity risk premium. Bollerslev et al. (2011) also tested the predictability of excess stock returns by running a regression on monthly variance risk premium and other macroeconomic variables. Their work confirms the importance of variance risk premium. A recent work in Feunou et al. (2014) studies the term structure effect of several risk premiums including equity risk premium, variance, skewness and kurtosis risk premiums. They find empirical evidences that risk premium implied from higher order conditional moments have different excess return prediction power.

In this section, I reexamine the predictability of equity risk premium term structure using variance risk premium term structure data. Our aim is to reveal the prediction power of variance risk premium by considering term structure. It can be seen the short term risk

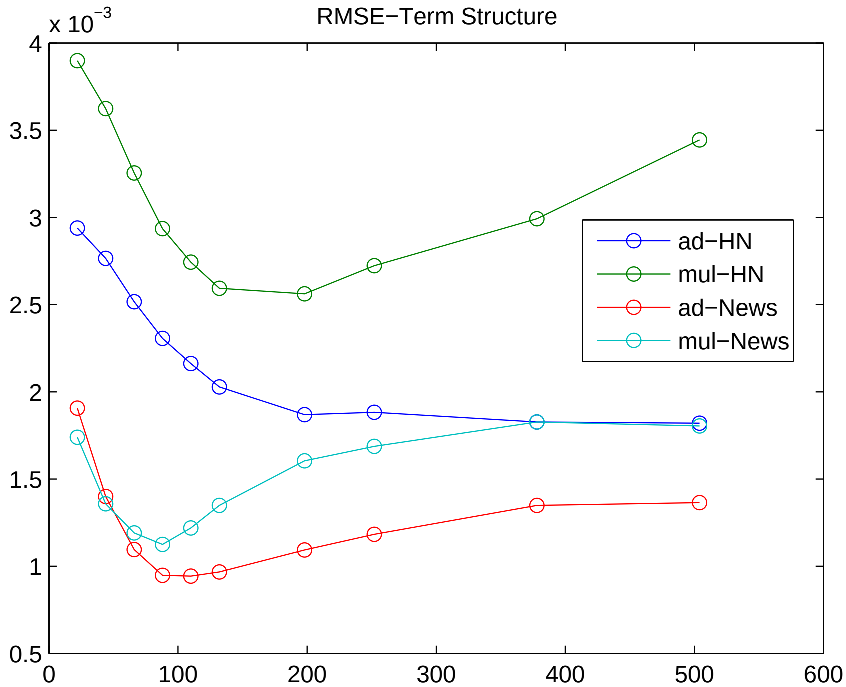

Fig. 4. Term Structure RMSE

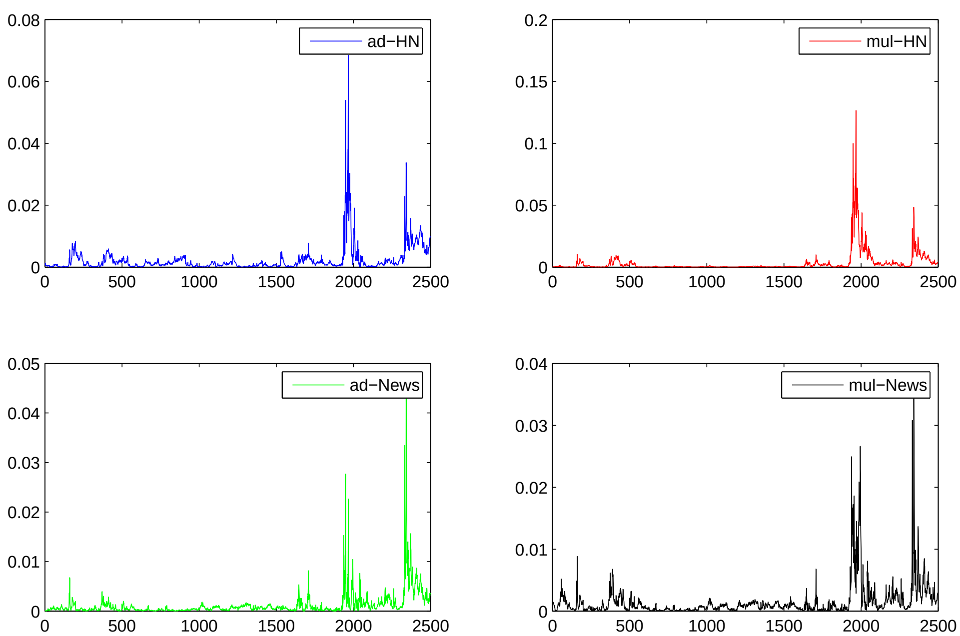

Fig. 5. Daily RMSE

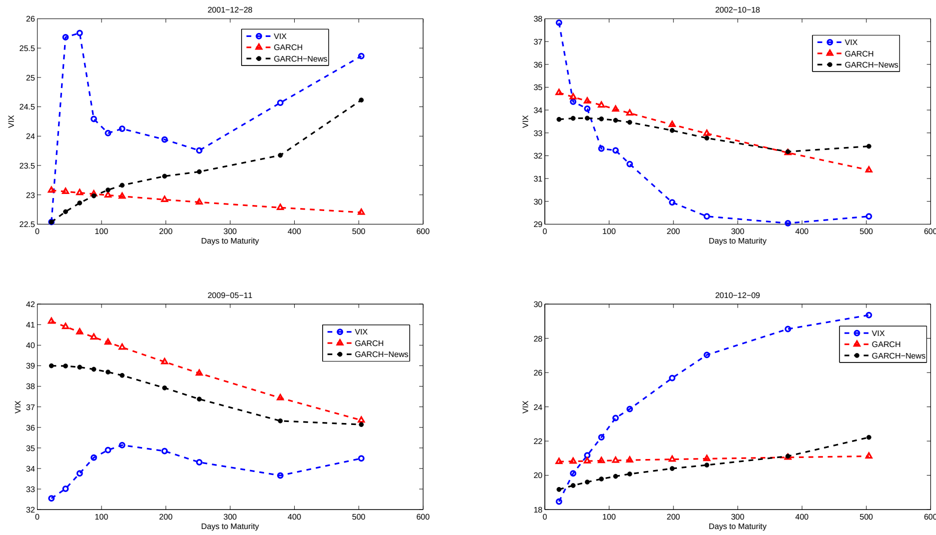

Fig. 6. Market and Model Implied VIX Term Structure

premium is much volatile and significant than long term.

6.1. *Equity Risk Premium, GARCH Implied Variance Risk Premium and Their Term Structure*

In this section, I present the estimated ERP and GARCH Implied VRP. Fig. 7 plots the ERP panel data with five terms, 1, 3, 6, 12, and 24 month. Fig. 8 plots the VRP panel data with ten terms, 1, 2, 3, 4, 5, 6, 9, 12, 18, and 24 month.

6.2. *Panel Regression*

6.2.1. *Term Structure Leverage Effect: Panel Regression of Excess Return on VIX*

In this section, I study the predictability of excess return term structure, which is also the equity risk premium term structure. It is defined as the difference between S&P 500 log-return panel and US treasury constant maturity rates panel. In order to reveal the term structure leverage effect, the only predictor variable I use is VIX term structure and the panel data in the same sample period as excess return term structure data. Other important predictor variables are not ignored. However, instead of introducing other independent variables in the regression equation, I assume these variables are summarized into two latent

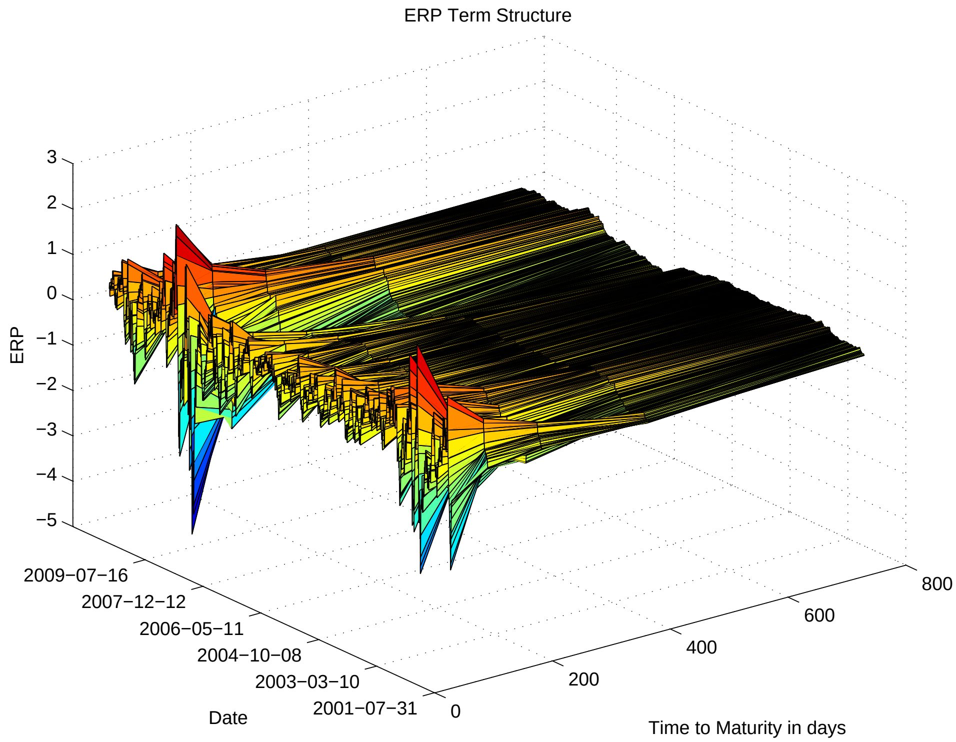

Fig. 7. ERP Term Structure

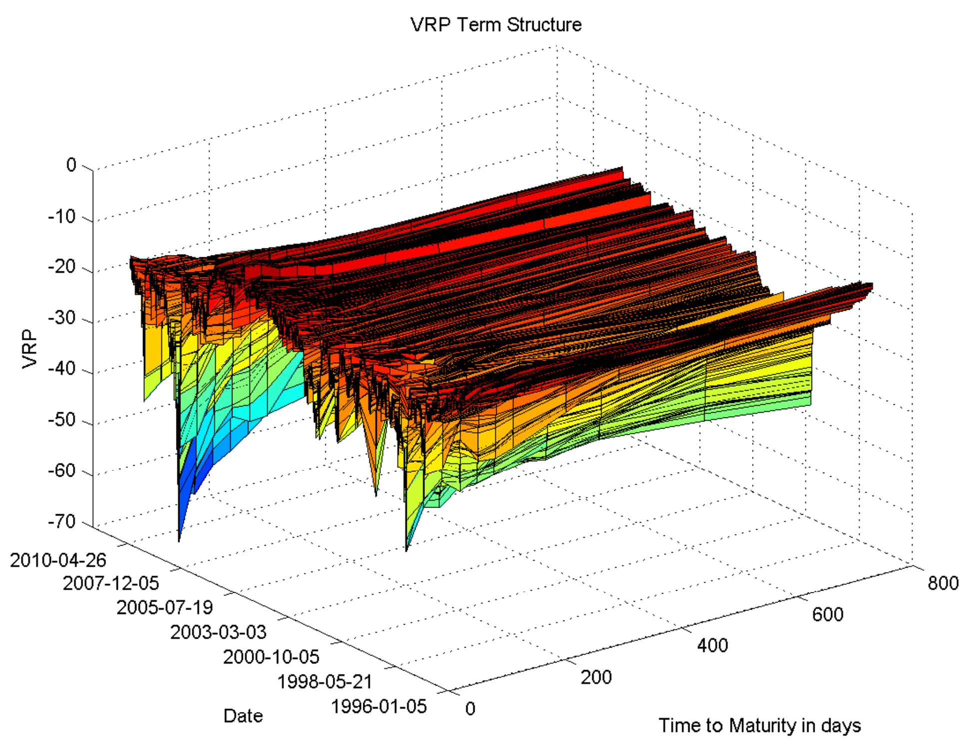

Fig. 8. HN2CGARCH(1,1) Implied VRP Term Structure

fixed effects in the panel regression, one is term fixed effect, the other is period fixed effect.

$$ERP_{t,i} = C + \beta VIX_{t,i} + \mu_i + \eta_t + \epsilon_{t,i} \quad (74)$$

where $ERP_{t,i} = r_{t,i} - r^f_{t,i}$, the difference between annualized log-return $r_{t,i} = \frac{T_i}{252}\log(S_{t+T_i}/S_t)$ and annualized risk free rate $r^f_{t,i}$, which is US treasury constant maturity rates with maturity T_i. $t = 1, 2, ..., 2371$ represents time, and $T_i = 22, 66, 132, 252, 528$ denotes term i, $i = 1, 2, ..., 5$. μ_i is the term structure fixed effect, η_t is the period fixed effect. Table 4 reports the panel regression results and related tests. By the significant negative value of estimated coefficient $\hat{\beta} = -1.08$, I conclude that there is a salient term structure leverage effect between excess return and model-free variance.

Table 4: This table presents the results of panel regression of equity risk premium term structure(ERP) on VIX term structure data. The sample period is from 31 July 2001 to 03 January 2011. ERP is calculated as the difference S&P 500 log-return with different terms and US treasury constant maturity rates. Maturities include 1 month, 3 month, 6 month, 1 year and 2 year.

Panel Regression of ERP on VIX

Dependent Variable: ERP
Method: Panel Least Squares
Sample: 1 11855
Periods included: 2371
Cross-sections included: 5
Total panel (balanced) observations: 11855

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.255218	0.032449	7.865077	0.0000
VIX	-1.083033	0.135543	-7.990333	0.0000

R-squared	0.490161	Mean dependent var	-0.003076
Adjusted R-squared	0.362419	S.D. dependent var	0.385967
S.E. of regression	0.308189	Akaike info criterion	0.660966
Sum squared resid	900.3202	Schwarz criterion	2.140180
Log likelihood	-1541.876	Hannan-Quinn criter.	1.157464
F-statistic	3.837121	Durbin-Watson stat	0.104562
Prob(F-statistic)	0.000000		

Table 5: This table presents the results of panel regression of equity risk premium term structure (ERP) on HNCGARCH implied VRP term structure. The sample period is from 31 July 2001 to 03 January 2011. ERP is calculated as the difference S&P 500 log-return with different terms and US treasury constant maturity rates. Maturities include 1 month, 3 month, 6 month, 1 year and 2 year.

Panel Regression of ERP on VRP

Dependent Variable: ERP
Method: Panel Least Squares
Sample: 1 11855
Periods included: 2371
Cross-sections included: 5
Total panel (balanced) observations: 11855

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.215339	0.029875	7.208096	0.0000
HN_VRP	1.122472	0.152839	7.344138	0.0000

R-squared	0.489632	Mean dependent var	-0.003076
Adjusted R-squared	0.361757	S.D. dependent var	0.385967
S.E. of regression	0.308349	Akaike info criterion	0.662005
Sum squared resid	901.2560	Schwarz criterion	2.141219
Log likelihood	-1548.034	Hannan-Quinn criter.	1.158503
F-statistic	3.828992	Durbin-Watson stat	0.104518
Prob(F-statistic)	0.000000		

6.2.2. Panel Regression: Excess Return on VRP

Similar to last section, I investigate the predictability of excess return term structure defined as the difference between S&P 500 log-return panel and US treasury constant maturity rates panel by using variance risk premium term structure data (HN_VRP) implied by model-free VIX term structure panel and HNGARCH implied variance term structure panel data. I do not neglect other important factors may impact on ERP. Nevertheless, I incorporate these factors into two latent fixed effects in the panel regression without introducing other independent variables.

$$ERP_{t,i} = C + \beta HN\_VRP_{t,i} + \mu_i + \eta_t + \epsilon_{t,i} \tag{75}$$

where $ERP_{t,i} = r_{t,i} - r^f_{t,i}$, the difference between annualized log-return $r_{t,i} = \frac{T_i}{252}\log(S_{t+T_i}/S_t)$ and the corresponding annualized risk free rates, which is US treasury constant maturity rates with maturity T_i. $t = 1, 2, ..., 2371$ represents time, and $T_i = 22, 66, 132, 252, 528$ denotes term i, $i = 1, 2, ..., 5$. μ_i is the term structure fixed effect, η_t is the period fixed effect. Table 5 reports the panel regression results and related tests. The significant positive value of estimated coefficient $\hat{\beta} = 1.12$, which shows the prediction power of VRP term structure on ERP term structure.

We are also concerned with the existence of term structure fixed effect and period effects in the two panel regression above. The fixed effect tests are reported in table 6. By F test and Chi-square test, the null hypothesis which claims no fixed effects is rejected. We are not surprised by this results since even the VIX term structure data and VRP term structure data are of important prediction power to ERP, they still can not completely determine ERP. In fact, there is latent term structure effect (term fixed effect μ_i) not varying with time are estimated in the lower panel of table 6. Interestingly, it is negative in short term and turns to be positive after 1 year. The period fixed effect (η_t) are plotted in Fig 9. It is unexpected that the period effects of two panel regression are nearly the same. One possible explanation is the latent period effect of ERP is completely unrelated to variance variables.

Table 6: This table presents the fixed effect test in panel regressions of ERP on VIX (left upper panel) and ERP on VRP (right upper panel). The lower panel illustrates estimated term structure fixed effects.

Fixed Effects Tests and Estimates

Fixed Effects Tests						
ERP V.S. VIX				**ERP V.S. VRP**		
Test cross-section and period fixed effects						
Effects Test	Statistic	d.f.	Prob.	Statistic	d.f.	Prob.
Cross-section F	3.575134	(4,9479)	6.4000E-03	3.781128	(4,9479)	4.5000E-03
Cross-section Chi-square	17.87162	4	1.3000E-03	18.90054	4	8.0000E-04
Period F	3.786437	(2370,9479)	0.0000E+00	3.807161	(2370,9479)	0.0000E+00
Period Chi-square	7897.094	2370	0.0000E+00	7928.606	2370	0.0000E+00
Cross-Section/Period F	3.780875	(2374,9479)	0.0000E+00	3.801492	(2374,9479)	0.0000E+00
Cross-Section/Period Chi-square	7898.34	2374	0.0000E+00	7929.74	2374	0.0000E+00
Fixed Effects						
1 Month	-0.01684			-0.01776		
3 Month	-0.00124			-0.00138		
6 Month	-0.00115			-0.00095		
1 Year	0.001332			0.001698		
2 Year	0.017902			0.01839		

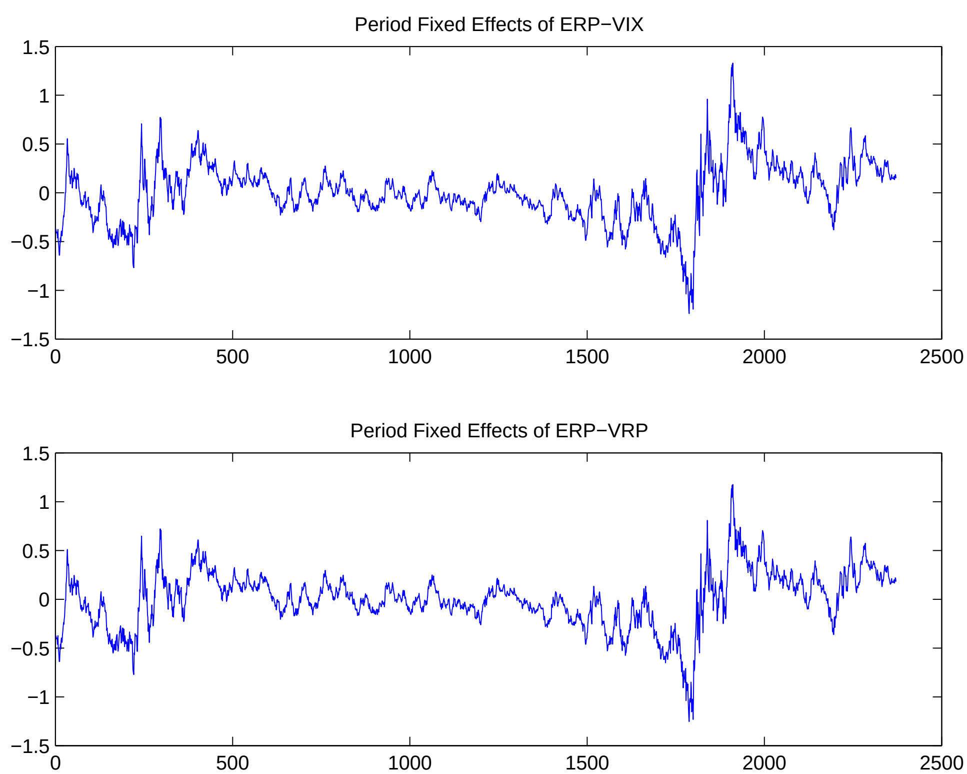

Fig. 9. Period Fixed Effects of ERP on VIX and HNGARCH VRP

7. Conclusion

Variance risk premium is an important economic factor not only in excess asset return prediction but also in its close link to variance derivatives. However, the term structure variance risk is not analyzed too much. This paper investigates one-component GARCH and two-component GARCH model and their capability to fit model-free VIX term structure data. I find that these models are prone to over-specify variance term structure curves to be monotonic, which deviates the model-free variance term structure extracted from option prices. In order to model the variance term structure properly, a new time-varying GARCH model is introduced whose implied VIX formula is controlled by a news impact function. The variance term structure curves implied by time-varying News GARCH model is of great flexibility to capture monotonic and non-monotonic variance term structure curves with multiple extrema due to the impact of important macroeconomic news. Additionally, the log joint likelihood function with additive error produces smaller VIX term structure market data fitting error than the log joint likelihood function multiplicative error by means of both daily RMSE and term RMSE. This suggests the VIX model error should be simple difference rather than log-difference between market VIX term structure data and model implied variance term structure values. Regarding equity risk premium predictablity, I empirically confirm that variance risk premium term structure and VIX term structure are important predictors for equity risk premium term structure, which is consistent to most literature considering the prediction power of single term variance risk premium and 30-day VIX index. The term structure panel data regression of ERP on VIX demonstrates strong term structure leverage effect. I also observe that the fixed effects (term fixed effect, period fixed effect) of panel regression of ERP on VIX are very close to the fixed effects extracted from panel regression of ERP on VRP. Term fixed effect is negative with term less than 1 year, while it turns positive in long term case.

Our analysis at least explore the following avenues to future research. First, different news indicators for News GARCH model are to be constructed. The news impact function is not unique and can be set in different functional forms. All these different forms show different way for certain market news to impact on variance term structure. Second, the homoscedastic one lag autocorrelated error vector in joint MLE method can be specified allowing heteroscedasticity to improve the performance of joint MLE. Finally, if jumps are introduced in our model, then jump risk premium (JRP) and its term structure need to be estimated and related empirical facts are to be investigated. It is also meaningful to answer the question: what is the relationship among ERP, VRP and JRP?

References

Ait-Sahalia, Y., Karaman, M., Mancini, L., 2014. The term structure of variance swaps and risk premia. Available at SSRN 2136820 .

Amado, C., Teräsvirta, T., 2013. Modelling volatility by variance decomposition. Journal of Econometrics 175, 142–153.

Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P., 2009a. Parametric and nonparametric volatility measurement. Handbook of financial econometrics 1, 67–138.

Andersen, T. G., Bondarenko, O., 2007. Construction and interpretation of model-free implied volatility. Tech. rep., National Bureau of Economic Research.

Andersen, T. G., Davis, R. A., Kreiss, J.-P., Mikosch, T. V., 2009b. Handbook of financial time series. Springer Science & Business Media.

Badescu, A., Elliott, R. J., Kulperger, R., Miettinen, J., Siu, T. K., 2011. A comparison of pricing kernels for garch option pricing with generalized hyperbolic distributions. International Journal of Theoretical and Applied Finance 14, 669–708.

Badescu, A., Kulperger, R., Lazar, E., 2008. Option valuation with normal mixture garch models. Studies in Nonlinear Dynamics & Econometrics 12.

Beach, C. M., MacKinnon, J. G., 1979. Maximum likelihood estimation of singular equation systems with autoregressive disturbances. International Economic Review pp. 459–464.

Bollerslev, T., Chou, R. Y., Kroner, K. F., 1992. Arch modeling in finance: A review of the theory and empirical evidence. Journal of econometrics 52, 5–59.

Bollerslev, T., Gibson, M., Zhou, H., 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., Litvinova, J., Tauchen, G., 2006. Leverage and volatility feedback effects in high-frequency data. Journal of Financial Econometrics 4, 353–384.

Bollerslev, T., Xu, L., Zhou, H., 2015. Stock return and cash flow predictability: The role of volatility risk. Journal of Econometrics .

Bormetti, G., Corsi, F., Majewski, A. A., 2015. Term structure of variance risk premium in multi-component garch models. Available at SSRN 2619278 .

Breeden, D. T., Litzenberger, R. H., 1978. Prices of state-contingent claims implicit in option prices. Journal of business pp. 621–651.

Britten-Jones, M., Neuberger, A., 2000. Option prices, implied price processes, and stochastic volatility. The Journal of Finance 55, 839–866.

Campbell, J. Y., Hentschel, L., 1992. No news is good news: An asymmetric model of changing volatility in stock returns. Journal of financial Economics 31, 281–318.

Carr, P., Wu, L., 2009. Variance risk premiums. Review of Financial Studies 22, 1311–1341.

CBOE, 2009. The cboe volatility index–vix. White Paper .

Chorro, C., Guégan, D., Ielpo, F., 2012. Option pricing for garch-type models with generalized hyperbolic innovations. Quantitative Finance 12, 1079–1094.

Christoffersen, P., Elkamhi, R., Feunou, B., Jacobs, K., 2009. Option valuation with conditional heteroskedasticity and nonnormality. Review of Financial studies p. hhp078.

Christoffersen, P., Heston, S., Jacobs, K., 2006. Option valuation with conditional skewness. Journal of Econometrics 131, 253–284.

Christoffersen, P., Jacobs, K., 2004. Which garch model for option valuation? Management science 50, 1204–1221.

Christoffersen, P., Jacobs, K., Chang, B. Y., 2012. Forecasting with option-implied information. Handbook of economic forecasting 2.

Christoffersen, P., Jacobs, K., Ornthanalai, C., Wang, Y., 2008. Option valuation with long-run and short-run volatility components. Journal of Financial Economics 90, 272–297.

Duan, J.-C., 1995. The garch option pricing model. Mathematical finance 5, 13–32.

Duan, J.-C., Yeh, C.-Y., 2010. Jump and volatility risk premiums implied by vix. Journal of Economic Dynamics and Control 34, 2232–2244.

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

Feunou, B., Fontaine, J.-S., Taamouti, A., Tédongap, R., 2014. Risk premium, variance premium, and the maturity structure of uncertainty. Review of Finance 18, 219–269.

Gonzalez-Perez, M. T., 2015. Model-free volatility indexes in the financial literature: A review. International Review of Economics & Finance .

Hao, J., Zhang, J. E., 2013. Garch option pricing models, the cboe vix, and variance risk premium. Journal of Financial Econometrics 11, 556–580.

Heston, S. L., Nandi, S., 2000. A closed-form garch option valuation model. Review of Financial Studies 13, 585–625.

Javaheri, A., Wilmott, P., Haug, E. G., 2004. Garch and volatility swaps. Quantitative Finance 4, 589–595.

Jiang, G. J., Tian, Y. S., 2005. The model-free implied volatility and its information content. Review of Financial Studies 18, 1305–1342.

Jiang, G. J., Tian, Y. S., 2007. Extracting model-free volatility from option prices: An examination of the vix index. Journal of Derivatives 14.

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

Madan, C., 1998. Towards a theory of volatility trading. Working Paper .

Zhu, Y., Zhang, J. E., 2007. Variance term structure and vix futures pricing. International Journal of Theoretical and Applied Finance 10, 111–127.