The New Variation Method of Volatility Risk Premium

Yoshihiko Sugihara *†

First draft: May 2, 2008
This draft: October 17, 2008

Abstract

Integrated implied volatility is an expected quadratic variation that can be derived from European option prices under the risk neutral measure, whereas integrated realized volatility is a historical quadratic variation that can be derived from high-frequency ex-post stock returns under the physical measure. Earlier studies used the difference as the volatility risk premium and showed the premium was negative for buyers of options in most periods. Their method, however, has a filtration inconsistency problem when we evaluate an expected risk premium. This problem results from the difference of the filtration required for the computation of the two volatility estimates. We proposes a new method of evaluating the expected risk premium with consistent filtration. In our method, the premium is identified as the gap between the current integrated implied volatility and the integrated volatility estimated by a time series model of realized volatilities. We calculate the premium under the condition that the one-day realized volatility obeys either the Heston model or the ARFIMAX model. We found the premium based on our method correlates with market risk indicators, such as Citi Macro Risk Index or iTraxx Japan, more strongly than those based on the earlier works. This indicates our proposed measure of expected volatility risk premium reflects the market sentiment well and it is a valid indicator for the market risk aversion. As a byproduct, we can quantify the probability where the realized volatility goes below the implied volatility. This probability might also be an indicator for market's volatility risk aversion.

Keywords: volatility risk premium, model-free implied volatility, realized volatility, VIX, Heston model, ARFIMAX model
JEL Classification: G12, G13, G14

*I would like to thank Peter Carr, Sacha Stoikov, Toshiaki Watanabe and seminar participants at Diawa International Workshop on Financial Engineering 2008 for useful comments and suggestions. Views expressed in this paper are those of the author and do not necessarily reflect the official views of the Bank of Japan.

†Institute for Monetary and Economic Studies, Bank of Japan (E-mail: yoshihiko.sugihara@boj.or.jp).

1 Introduction

Volatilities in asset prices often vary roughly. This generates a premium for a volatility implied in a derivative. The existence of the premium is widely known and that is called a volatility risk premium, hereafter VRP. Various earlier works studied the VRP in a parametric way under a stochastic volatility model, or SV model.

Scott [1987], Wiggins [1987] and Hull and White [1987] introduced the SV model and Wiggins [1990] firstly proved the SV model fitted using the option prices with various strikes and term to maturities. In their works, a VRP was derived by the partial differential equation of a derivative price under the SV model (see Appendix A.1 and A.2). Various studies on a VRP also applied parametric models for volatilities. One of the most recent studies, Bakshi and Kapadia [2003], introduced a parametric method to capture the VRP from stock option buyers' delta-hedged gains. They showed that the VRP was negative for option buyers and the absolute value was likely to stay large in periods of higher volatilities. The negative premium for option buyers implies that the option sellers gain a positive premium, an excess return beyond the risk neutral volatility value. That is frequently observed in the US stock market. Uchida and Miyazaki [2008] applied this method to the Japanese stock market and found similar results.

Recently, some works has examined non-parametric estimation for a VRP defined by the gap between the model-free integrated implied volatility and the integrated realized volatility. The model-free integrated implied volatility is an aggregated implied volatility derived in non-parametric way from put and call European option prices with different strikes. On the other hand, the integrated realized volatility is a non-parametric consistent estimator of a realized variance computed from high-frequency intraday price data. Carr and Wu [2005] examined this type of VRPs on some stock indices and individual stocks in the US market. They found that these VRPs were negative for option buyers in most periods, and the VRPs in individual stock options can primarily be explained by the VRP in the stock index options. Bollerslev *et al.* [2007] estimated the VRP using the realized volatility of the S&P500 and the VIX index, the model-free integrated implied volatility index on the S&P500. They assumed the VRP can be explained by macro economic variables and lagged VRP, and showed that the VRP was negative for option buyers and it well predicted the future returns of the S&P500.

These non-parametric estimation of the VRP owes to the theory and computation method of the model-free integrated implied volatility and the realized volatility. Neuberger [1990], Dupire [1992], Demeterfi *et al.* [1999] and Britten-Jones and Neuberger [2000] generated and developed the theory of the model-free implied volatility, and Jiang and Tian [2007] developed its computation methodology. The VIX index, traded on the Chicago Board Options Exchange, is based on the theory. Sugihara [2008 *to appear*] summarised the theory and computation method, and showed the empirical characteristics of the model-free integrated implied volatility on the Nikkei 225 Japanese Stock Index, hereafter Nikkei. The availability of high-frequency intraday prices data led to the development of the estimation methodology of the realized volatility. Andersen and Bollerslev [1997, 1998] proposed the concept of realized volatility and showed the realized volatility was a theoretically reliable estimator of volatility. McAleer and Mederios [2008] or Shibata [2008] surveyed recent developments of the realized volatility.

The non-parametric estimation for the VRP is much simpler and easier to implement than the estimation based on earlier works under the SV model. However, a filtration inconsistency problem arises in the non-parametric estimation. This problem results from the difference of the filtration required for the computation of the two volatility measures. At a given day, the realized volatility computed from past market prices is an ex post volatility, whereas the model-free integrated implied volatility on the day is an expected future volatility. If we try to match their integration periods under the same filtration to compute an expected VRP instead of a realized VRP, we have to work out future realized volatilities which cannot be observed at the day.

In order to determine the expected VRP, we propose a new and simple method which addresses this filtration inconsistency problem. In our method, the expected VRP is identified as the gap between the model-free integrated implied volatility estimated at a day t and the integrated volatility estimated by a time series model of realized volatilities using data up to the day t. We calculate the premium under the condition that the realized volatility obeys either the Heston model or the ARFIMAX model, and obtain different results from earlier works. The estimated risk premium by our method shows stickier developments than those by earlier works. It also shows co-movement with risk indices, proxies of investors' risk aversion. As a byproduct, we can compute the probability where the integrated implied volatility stays higher than the integrated realized volatility. This probability might indicate the risk aversion of the option market.

Note that there are some viewpoints of the VRP in earlier works. According to Lee [2001], those differences mainly came from two matters. One is whether the VRP is based on the correlated stochastic factor between stock prices and volatilities or uncorrelated one, and the other is whether the VRP measures risk premiums for derivative sellers or buyers. In this paper, we assume the VRP reflects the correlation, and is captured from option buyers' standpoint. Hence, the sign of our VRP is opposite of VRPs based on option sellers' standpoint in some earlier works.

The rest of this paper is organised as follows. Section 2 defines the integrated implied and realized volatilities. Section 3 describes the four measures of VRP and discusses the filtration inconsistency problem arisen from the difference in those VRPs. Section 4 formulates our proposing measure of an expected VRP with its calculation procedure. Section 5 shows the empirical results of the expected VRP and discusses the difference of conventional VRPs and the expected VRP. Lastly, Section 6 concludes the paper.

2 Integrated Implied and Realized Volatilities

Suppose an asset price at time t, S_t, moves stochastically with i.i.d. instantaneous returns. The integrated quadratic variation at time t until time T $(T > t)$ is defined as:

$$\langle S \rangle_{t,T} = \frac{1}{T-t} \int_t^T \left(\frac{dS_u}{S_{u^-}} \right)^2, \tag{1}$$

where $T - t$ is an annualizing conversion coefficient and S_{t^-} is the stock price just before time t. The expected value of the future quadratic valuation can be estimated by options prices, and the realized value of the past quadratic variation can be estimated by the underlying asset processes. The former estimator is called the integrated implied variance, hereafter IIV, which is an expected variance under the risk neutral measure $\mathcal{Q}$,

$$\sigma^2_{IIV}(t,T) = E^{\mathcal{Q}}[\langle S \rangle_{t,T} | \mathcal{F}_t], \tag{2}$$

where $\mathcal{F}_t$ is the filtration generated by the asset prices process until time t. The latter is called the integrated realized variance, hereafter IRV, which is a consistent estimator of the realized variance under the physical measure $\mathcal{P}$,

$$\sigma^2_{IRV}(t,T) = E^{\mathcal{P}}[\langle S \rangle_{t,T} | \mathcal{F}_T] = \langle S \rangle_{t,T}. \tag{3}$$

Note that the IRV is a time-T value, not a time-t value.

The IIV is estimated from out-of-the-money European put and call options prices across all strikes. Demeterfi *et al.* [1999] and Britten-Jones and Neuberger [2000] derived the IIV by option prices as:

$$\sigma^2_{IIV}(t,T) = \frac{2}{(T-t)B(t,T)} \left(\int_0^{F(t,T)} \frac{P(t,T,K)}{K^2} dK + \int_{F(t,T)}^{\infty} \frac{C(t,T,K)}{K^2} dK \right), \tag{4}$$

where $B(t,T)$, $F(t,T)$, $P(t,T,K)$ and $C(t,T,K)$ denote a risk free discounted bond price, a forward rate, a put and a call option price at time t with a maturity T and a strike price K, respectively. The IIV represents the aggregated expectation of the future volatility in the options market. The volatility index on the S&P500, the VIX index, traded on the Chicago Board Options Exchange, is based on a discrete formula of eq.(4).

The IRV is estimated from high-frequency ex post intraday stock returns by discretizing the integral in eq.(1). Assume we have M days from time t to time T and let $\{t_i\}_{i=1}^M$ denotes set of time of days between time t to time T where $t_0 = t$ and $t_M = T$, the IRV can be estimated as:

$$\sigma^2_{IRV}(t,T) = \frac{1}{T-t} \sum_{i=1}^{M} RV_{t_i}, \tag{5}$$

where RV_{t_i} is the one-day realized variance (RV) on day i. Further assuming there are N i.i.d. intraday returns on each day i, $\{r_{j,i}\}_{j=1}^{N}$, we can estimate RV_{t_i} as:

$$RV_{t_i} = \sum_{j=1}^{N} r_{j,i}^2. \tag{6}$$

Theoretically, when N is large enough, the right hand side of eq.(5) converges in probability to eq.(1), i.e.:

$$\operatorname*{plim}_{N\to\infty} \sigma_{IRV}^2(t,T) = \langle S \rangle_{t,T}. \tag{7}$$

However, many earlier studies showed that, when the data frequency, or N, is too large, eq.(5) tends to overestimate the IRV because of an increase in market microstructure noises such as the price reciprocal development between the bid and offer prices within a short period, so-called "bid ask bounce."

3 Volatility Risk Premium

3.1 Definitions and filtration inconsistency problem

Using the IIV and IRV defined in section 2, the VRP is considered to be the gap between them, i.e.,

$$\mathrm{VRP} = \mathrm{IRV} - \mathrm{IIV}. \tag{8}$$

The VRP is equivalent with the covariance between pricing kernel of volatility and the quadratic variation. Let ξ_t denotes the Radon-Nikodým derivative process as

$$\xi_t = E\left[\left.\frac{d\mathcal{Q}}{d\mathcal{P}}\right| \mathcal{F}_t\right], \tag{9}$$

and let $M_{t,T}$ the pricing kernel as

$$M_{t,T} = \xi_T/\xi_t. \tag{10}$$

Note that $E[M_{t,T}|\mathcal{F}_t] = 1$. Then the IIV and IRV have the following relationship.

$$\begin{aligned}
\sigma_{IIV}^2(t,T) &= E^{\mathcal{Q}}[\langle S \rangle_{t,T}|\mathcal{F}_t] \\
&= E^{\mathcal{P}}[M_{t,T}\langle S \rangle_{t,T}|\mathcal{F}_t] \\
&= E^{\mathcal{P}}[\langle S \rangle_{t,T}|\mathcal{F}_t] + \mathrm{Cov}^{\mathcal{P}}(M_{t,T}, \langle S \rangle_{t,T}|\mathcal{F}_t) \\
&= \tilde{\sigma}_{IRV}^2(t,T) - \tilde{\lambda}(t,T),
\end{aligned} \tag{11}$$

where $\tilde{\sigma}_{IRV}^2(t,T) = E^{\mathcal{P}}[\langle S \rangle_{t,T}|\mathcal{F}_t]$ and $\tilde{\lambda}(t,T) = -\mathrm{Cov}^{\mathcal{P}}(M_{t,T}, \langle S \rangle_{t,T}|\mathcal{F}_t)$ denotes the VRP at time t integrated from time t to time T. Earlier works showed the VRP is related to the representative investor's risk aversion. See Appendix A.5 for the details. In eq.(11), we add a tilde ˜ on the IRV to emphasise filtration difference from our previous defined IRV; $\sigma^2_{IRV}(t,T)$ defined in eq.(3) is $\mathcal{F}_T$-measurable, whereas $\tilde{\sigma}^2_{IRV}(t,T)$ in eq.(11) is $\mathcal{F}_t$-measurable. $\tilde{\sigma}^2_{IRV}(t,T)$, and hence $\tilde{\lambda}(t,T)$, cannot be computed from historical market data.

Alternatively, we define a realized VRP, $\overline{\lambda}(t,T)$, hereafter rVRP, using $\mathcal{F}_T$-measurable variances as:

$$\begin{aligned}\overline{\lambda}(t,T) &= E^{\mathcal{P}}[\langle S\rangle_{t,T}|\mathcal{F}_T] - E^{\mathcal{Q}}[\langle S\rangle_{t,T}|\mathcal{F}_T] \\ &= \sigma^2_{IRV}(t,T) - \sigma^2_{IIV}(t,T), \end{aligned} \tag{12}$$

which can be computed from historical market data any time after T. As we show in Appendix A.4, the rVRP is related to the delta-hedged gain of options. Though we can conduct ex post calculation of eq.(12), the rVRP is still not a market expected VRP at time t.

In order to estimate the market expected VRP, we have to resolve the filtration inconsistency problem which arises from the difference in the measurement periods of IIV and IRV. Because this problem stems from the prediction of the future physical measure, we introduce a model to predict the future IRV based on the past time series of the RV. This approach gives a new VRP, namely an expected VRP, or eVRP.

In the estimation for the eVRP, we model RV process rather than IRV process to avoid the overlapping problem proposed by Christensen *et al.* [2002], that is the adjacent IRVs share the same RVs for integration. We simply assume that the model that best fit to the past RV's time series is the best estimate for the future physical measure. Let $\mathcal{G}_{t,T}$ denotes the model-based estimation of the future RV under the physical measure from time t to time T. The eVRP, $\lambda(t,T)$, is defined under the filtration $\mathcal{F}_t \vee \mathcal{G}_{t,T}$ as:

$$\lambda(t,T) = E^{\mathcal{P}}[\langle S\rangle_{t,T}|\mathcal{F}_t \vee \mathcal{G}_{t,T}] - \sigma^2_{IIV}(t,T), \tag{13}$$

or

$$\lambda(t,T) = -\mathrm{Cov}^{\mathcal{P}}(M_{t,T}, \langle S\rangle_{t,T}|\mathcal{F}_t \vee \mathcal{G}_{t,T}). \tag{14}$$

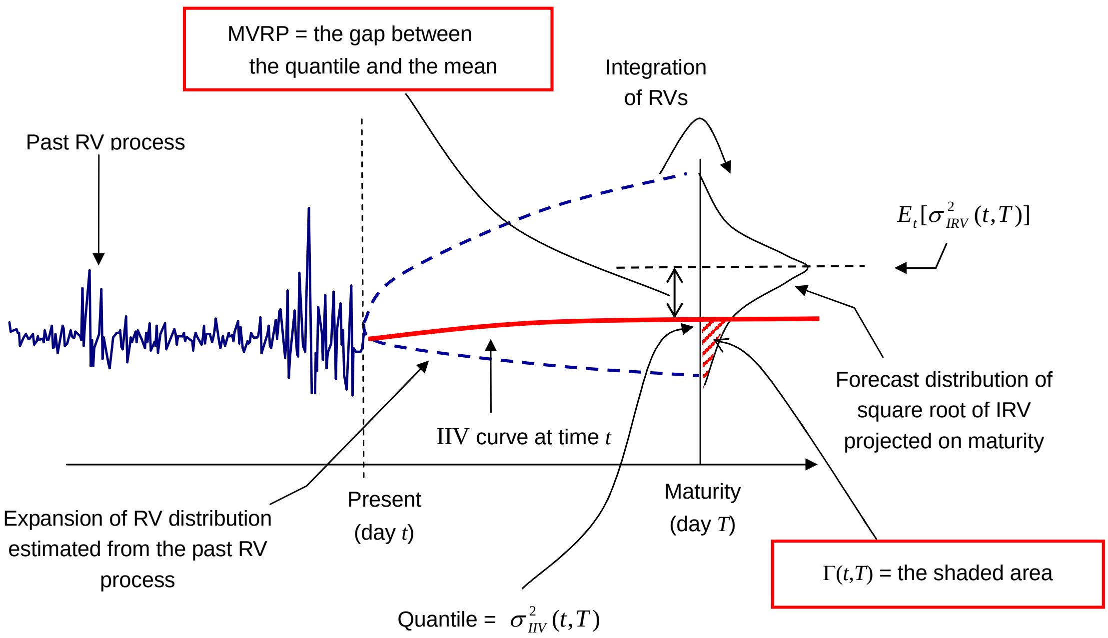
The concept of the eVRP is displayed in Figure 1.

As a byproduct, the estimation procedure of the eVRP gives the probability where the forecasted $\sigma^2_{IRV}(t,T)$ is lower than the $\sigma^2_{IIV}(t,T)$. We define the degree of volatility risk aversion for option sellers: Γ as

$$\Gamma(t,T) = \mathrm{Prob}(\sigma^2_{IRV}(t,T) \le \sigma^2_{IIV}(t,T)) = F_{IRV}(\sigma^2_{IIV}(t,T)), \tag{15}$$

where F_{IRV} is the cumulative density function of the model expected IRV. The VRA is given by the shaded area in the Figure 1. It is close to 100% when option sellers are very risk averse and require a large premium. Note that, since both option sellers and buyers are risk averse investors, Γ shows an imbalance of risk aversion between option sellers and buyers.

Figure 1: The concept of eVRP and VRA



The simplest estimation of $\mathcal{G}_{t,T}$ is the linear estimation; simply set $\mathcal{G}_{t,T} = \sigma(\{S_u\}^t_{u=t-(T-t)})$ using lagged samples of the daily RV. We call the IRV the trailed IRV denoted as $\underline{\sigma}^2_{IRV}(t,T)$ and defined as:

$$\underline{\sigma}^2_{IRV}(t,T) = \langle S \rangle_{t-(T-t),t}, \tag{16}$$

and a trailed VRP, or tVRP, is defined as:

$$\underline{\lambda}(t,T) = \underline{\sigma}^2_{IRV}(t,T) - \sigma^2_{IIV}(t,T). \tag{17}$$

The tVRP is an expected VRP under the assumption that the IRV in the previous month is the best prediction of the IRV in the following month.

So far, we have introduced four types of VRPs. The definitions and the differences of VRPs are summarised in Table 1.

Note that the filtration inconsistency problem has not been recognised in earlier studies that uses parametric estimation, because most of those discussed on the instantaneous VRP.

Table 1: The definition, notation and difference of VRPs

term	notation	explanation
true VRP	$\tilde{\lambda}(t,T)$	True expected VRP, unobservable at any time.
realized VRP (rVRP)	$\overline{\lambda}(t,T)$	Ex-post VRP, related to delta-hedged gain of options, observable after time T.
trailed VRP (tVRP)	$\underline{\lambda}(t,T)$	Expected VRP under static expectation assumption on future IRV, observable at time t.
expected VRP (eVRP)	$\lambda(t,T)$	Our proposed expected VRP using model-base forcast of IRV, observable at time t.

Since $\lambda(t,T) = \frac{1}{T-t}\int_t^T \lambda(t,s)ds$, the instantaneous VRP, $\lambda(t,t)$, is defined as $\lambda(t,t) = \lim_{T\to t}\lambda(t,T) = \frac{\partial}{\partial T}\lambda(t,T)\big|_{T=t}$ for any VRPs, and those converges to the same value whichever VRPs we use, i.e.:

$$\tilde{\lambda}(t,t) = \overline{\lambda}(t,t) = \underline{\lambda}(t,t) = \lambda(t,t). \tag{18}$$

See Appendix A.3 for the instantaneous VRP defined under the SV model.

3.2 Volatility risk premiums in the options market

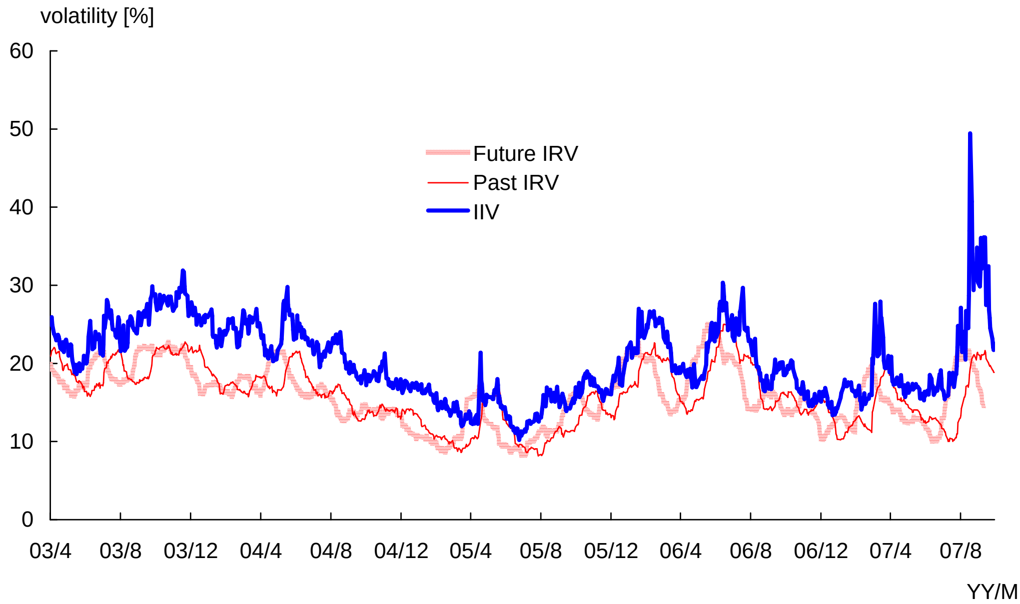
Figure 2 shows the movements of the one-month IIV computed using option prices with one-month maturity and the one-month IRV on Nikkei from April 2003 to September 2007. They are plotted in the volatility dimension, i.e., the square root of the integrated variance in eq.(4) and eq.(5). Two types of IRV are shown: future one-month IRV and past one-month IRV. The former covers the same period as the IIV, whereas the latter uses lagged samples of the daily RV. In other words, the future IRV is $\sigma_{IRV}(t, t + 1\text{month})$, and the past IRV is $\sigma_{IRV}(t - 1\text{month}, t)$.

The IRVs and IIV normally move parallel to each other, and the level of the IIV is greater than the IRVs in most of the periods. The gap is corresponds to the VRP [1]. We consider the gap between the future IRV and the IIV to be the rVRP, and that between the past IRV and the IIV to be the tVRP. Since IIV moves over IRV in most periods, the signs on the VRPs are negative.

The negative sign on the VRPs indicates that option sellers require a premium for uncertainty of the future volatility, or the volatility risk. We find that the past IRV seems to move more consistently with the IIV rather than the future IRV, and the future IRV sometimes exceeds the IIV. These observations lead to the following hypotheses: (a) option sellers require a premium by observing past IRV to compensate for the future volatility

[1] These values are computed in the volatility dimension and may be little smaller than the square root of the VRP. We neglect those difference in the discussion here.

Figure 2: Time series of IRV and IIV on Nikkei from April 2003 to September 2007



Notes: The values are one-month term to maturity in the volatility dimension. The IIV is calculated using the method proposed by Jiang and Tian [2007].

risk, and option buyers are willing to pay the premium, and (b) the premium covers the volatility risk for option sellers in most periods, but sometimes the premium is undervalued.

These results show the risk aversion of the financial markets, and reveal how the market participants form their expectation. The hypothesis (a) supports our assumption on the calculation of the eVRP; the expectation on the future RV seems to be based on the past development of RV. We guess the model-base eVRP is a better estimate of expected VRP than the tVRP that based on a static expectation on the future RV.

Under SV model, the risk premium on a derivative is composed of those on the stochastic factors in the process of underlying asset and the volatility, as we discussed in eq.(A-17) in Appendix A.2. Decomposition of the risk premium gives

$$\text{(Risk premium on a derivative)} = w_1 \cdot \text{(Risk premium on the underlying)} + w_2 \lambda(t,t), \quad (19)$$

where w_1, w_2 are weights which are proportional to the delta and the vega of the derivative respectively. Since the vega is positive for European options, the negative VRP reduces the risk premium on an option itself.

4 Variation Method of Expected Volatility Risk Premium

In this section, We show two models to forecast future RV process, which is used in the estimation of the eVRP, and describe the calculation procedures of the eVRP and Γ defined in Section 3.1.

4.1 Two Models for RV forecast

The first model is the Heston type model defined as:

$$dS_t = \mu S_t dt + \sqrt{RV_t} S_t dW_t^1,$$

$$d(RV_t) = \kappa(\theta - RV_t)dt + \sigma_V \sqrt{RV_t}\left(\rho dW_t^1 + \sqrt{1-\rho^2} dW_t^2\right), \tag{20}$$

where dW_t^1, dW_t^2 are independent Weiner processes and $\kappa, \theta, \sigma_V, \rho$ ($\sigma_V > 0, |\rho| \leq 1$) are parameters. The original model of eq.(20) was proposed by Heston [1993], which is the stochastic volatility model with mean-reverting square root volatility process. This model substitutes the volatility process in the original Heston model for RV. So, we call this model the Heston type model for RV hereafter. The Heston model is commonly used among practitioners, and this Heston type model for RV may well represent forcasting models which major market participants use.

The second model for RV forecast is the ARFIMAX model proposed by Giot and Laurent [2004] as:

$$R_t = \sqrt{RV_t} z_t,$$

$$(1-L)^d\{\ln(RV_t) - \mu_0 - \mu_1|R_{t-1}| - \mu_2 \mathbf{1}_{\{R_{t-1}<0\}}|R_{t-1}|\} = (1+\delta L)u_t, \tag{21}$$

$$z_t \backsim N(0, \sigma_z^2), \quad u_t \backsim N(0, \sigma_u^2),$$

where L is a lag operator[2], $R_t = \ln(S_t/S_{t-1})$ is the one-day return on day t, and $\mu_0, \mu_1, \mu_2, \delta, \sigma_z$ and σ_u ($\sigma_z, \sigma_u > 0$) are parameters. According to Watanabe and Sasaki [2007] and Shibata [2008], the ARFIMAX model provides a good fit of the past movement of the RV for Nikkei.

Note that both models have mean reverting property of RV. The Heston type model has fixed and deterministic mean of RV, whereas the ARFIMAX model has stochastic mean which changes based on the previous day's return. The ARFIMAX model expressed the volatility movement's asymmetric feature, that is the volatility is normally higher when the previous day's return was negative than when positive. This feature is expressed by the μ_3 parameter and $\mathbf{1}_{\{R_{t-1}<0\}}$ in eq.(21) which takes non-zero value only when the previous day's return was negative. The "X" in ARFIMAX is added to address this asymmetry. We assume the mean reverting property of RV in those models better expresses the market's

[2] $Lx_t = x_{t-1}$.

expectation for the future RV than the simple static expectation assumption in tVRP.

4.2 The calculation procedure

The eVRP and Γ can be calculated by the following procedure.

First, estimate the IIV in eq.(4) from the option prices data, following the methodology proposed by Jiang and Tian [2007]. The one- to six-month IIVs are estimated by interpolating market-traded terms of IIV using a cubic spline function. For the risk free discount bond price $B(t, T)$, the interpolated rate of LIBOR and swap rates is used.

Second, compute the RV from intraday stock price data by eq.(6). The appropriate choice of N, or the choice of a time interval to measure returns, is an ongoing issue for RV estimation. Because five-minute returns are widely used in earlier studies including those on Nikkei, we follow this choice of N. Note that the return during lunch break and the return from closing to the next morning's opening are included in the daily RV. The first sample, $j = 1$ in eq.(6), corresponds to the nighttime return from the last day's closing to the next morning's opening.

Third, estimate the parameters of the models for the RV. The parameters are estimated at each day t_i by the quasi-maximum likelihood method using the previous one-year of ex post RV data including the day t_0 itself [3]. Since estimation of IIV and IRV in Section 3.2 suggests that market volatility expectation relies on the recent development of RV, the parameters are updated daily for the better tracking of the development.

Forth, simulate RV process from the day t_0 to day t_M using the estimated parameters, and set the estimated future physical measure $\mathcal{G}_{t,T}$ using the simulated RV process. Then compute the model expected value of future IRV under the filtration $\mathcal{F}_t \vee \mathcal{G}_{t,T}$ by ten thousand times of Monte Carlo simulation each day.

Lastly, compute the eVRP, $\lambda(t, T)$, and the $\Gamma(t, T)$ from the IIV on day t_0 and the simulated IRV. The Monte Carlo simulation is used for the calculation of $\Gamma(t, T)$. We repeat the above process for each day to get the time series of eVRP and Γ.

5 Empirical Analysis

5.1 Volatility risk premium

5.1.1 Differences in three VRPs: eVRP, tVRP and rVRP

The upper and middle panels in Figure 3 display the eVRPs on Nikkei under the Heston and ARFIMAX models respectively. The VRPs from one, three and six months term to maturity are shown in square percent point scale. The lower panel shows one month rVRP

[3] The reason for the inclusion of the day t_0 is that the options market for the Nikkei closes at 15:10 whereas the stock market closes at 15:00. By comparing the closing IIV with the IRV, which are computed with past data including the current day, an unbiased VRP can be computed.

and tVRP for comparison. Comparing the eVRPs with tVRP and rVRP, we observe the following differences.

First, we focus the difference of RV forecast models. The eVRPs based on the Heston model and that on ARFIMAX model move parallel to each other. The directions of those changes are almost the same, but the levels differ occasionally.

Second, the eVRPs are less volatile and stickier than the tVRP and rVRP. The eVRPs suggest that the risk premium may primarily fluctuate with inertia. The eVRPs exhibits the cyclical development of the risk premium, compared with the tVRP and rVRP, and the eVRP also seems to detect the turning points of the long-term risk premium trends which are hard to be detected by the tVRP or rVRP.

Third, the eVRPs stay stably around zero while the realized volatility stays low. That may indicate the bargaining power balances between options buyers and sellers while the volatility stays low. In periods from October 2004 to July 2005 and from December 2006 to June 2007 when the IRV stayed below 15% as seen in Figure 2, the eVRPs stayed stably around zero. In the same periods, however, the tVRP moved above zero and the rVRP fluctuates. The tVRP and rVRP did not detect the bargaining power balance.

Forth, comparing eVRPs with different terms, we see the shorter term eVRPs are more volatile than the longer term eVRPs. This indicates that the shorter term VRP is more sensitive to changes in market sentiments. This is consistent with the characteristic of the implied volatility term structure showed in Sugihara [2008 *to appear*].

5.1.2 Consistency with other market risk indicators

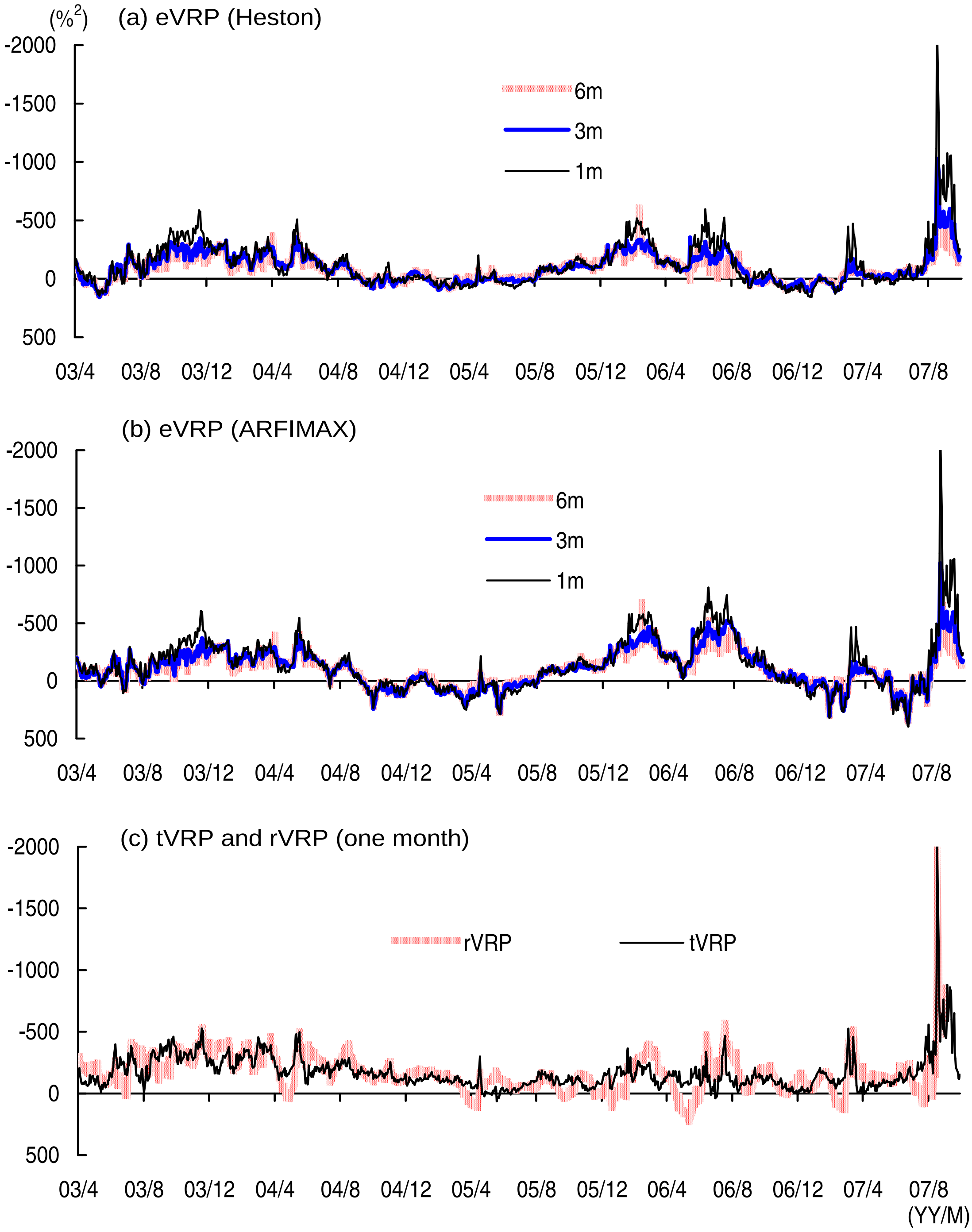
In order to determine whether the eVRP is better estimate of the risk premium than the tVRP and rVRP, we compare those VRPs with market risk indicators to be regarded as a proxy of risk preference. The VRPs are expected to move consistent with the risk indicators derived from the other markets, because investors in the options market normally invest in other asset classes such as government bonds or credit default swaps, and their risk aversion in the other market affects the options market.

The Citi Macro Risk Index, hereafter CMRI [4], and iTraxx Japan [5] are chosen for the risk indicators. The former measures global investors' overall risk aversion. Because more than half of investment into the Japanese stock market comes from overseas, the VRP in Nikkei is expected to be consistent with the global investors' risk preference such as CMRI. The iTraxx Japan measures credit risk in the referenced Japanese enterprises. It reflects

[4] The Citi Macro Risk Index is a equally weighted index of emerging market sovereign spreads, US credit spreads, US swap spreads and implied FX, equity and swap rate volatilities. The index is expressed in a rolling historical percentile and ranges between 0 (low risk aversion) to 1 (high risk aversion). Data source is Citibank Ltd.

[5] The iTraxx Japan refers 50 most liquid investment grade credit default swap premiums on Japanese entities. The indication is arithmetic average of index CDS premiums collected from broker dealers in Tokyo. It starts to be computed from July 2004. Data source is Markit Group Ltd.

Figure 3: eVRP, tVRP and rVRP



Notes: (a) The eVRP on Nikkei where RV obeys the Heston model, (b) the eVRP on Nikkei where RV obeys the ARFIMAX model, (c) the rVRP and tVRP. One-, three-, six-months-terms of eVRPs, and one-month-term rVRP and tVRP are shown. The ordinates are inverted to show the negative risk premium.

the downside risk aversion of the Japanese credit markets. Since the risk preference in the credit market is generally shared with the stock market, the VRP is expected to have a positive correlation with iTraxx Japan.

Table 2 displays the correlation between the VRPs and the risk indicators. Every VRP has a negative correlation with risk indicators and the eVRPs show stronger correlation than the rVRP and tVRP. The negative correlation indicates when the option sellers' risk aversion agrees with overall investors' risk aversion. And those stronger correlations results from coherent cycle of eVRPs and the risk indicator. Figure 4 displays the time series of eVRPs, tVRP and risk indicators. The upper panel displays the CMRI and VRPs, and the lower panel displays the iTraxx Japan and VRPs. During the periods from August to December 2004 and from July 2006 to January 2007, the eVRPs co-move well with the decrease in the risk indicators. The eVRP goes down in the periods while the tVRP almost unchanges. The eVRP also detects the rise of the risk aversion better than the tVRP in the period from May 2005 to February 2006. The same feature also holds when compared iTraxx Japan. In the periods from October to November 2004, from January to July 2006, from July to December 2006 and from April to June 2007, the eVRP tracks uprise and downturn of the risk aversion while the tVRP almost unchanges. These results suggest the eVRP is the better indicator of the risk aversion in the stock market rather than the tVRP or rVRP.

Table 2: Correlation between the VRPs and risk indicators

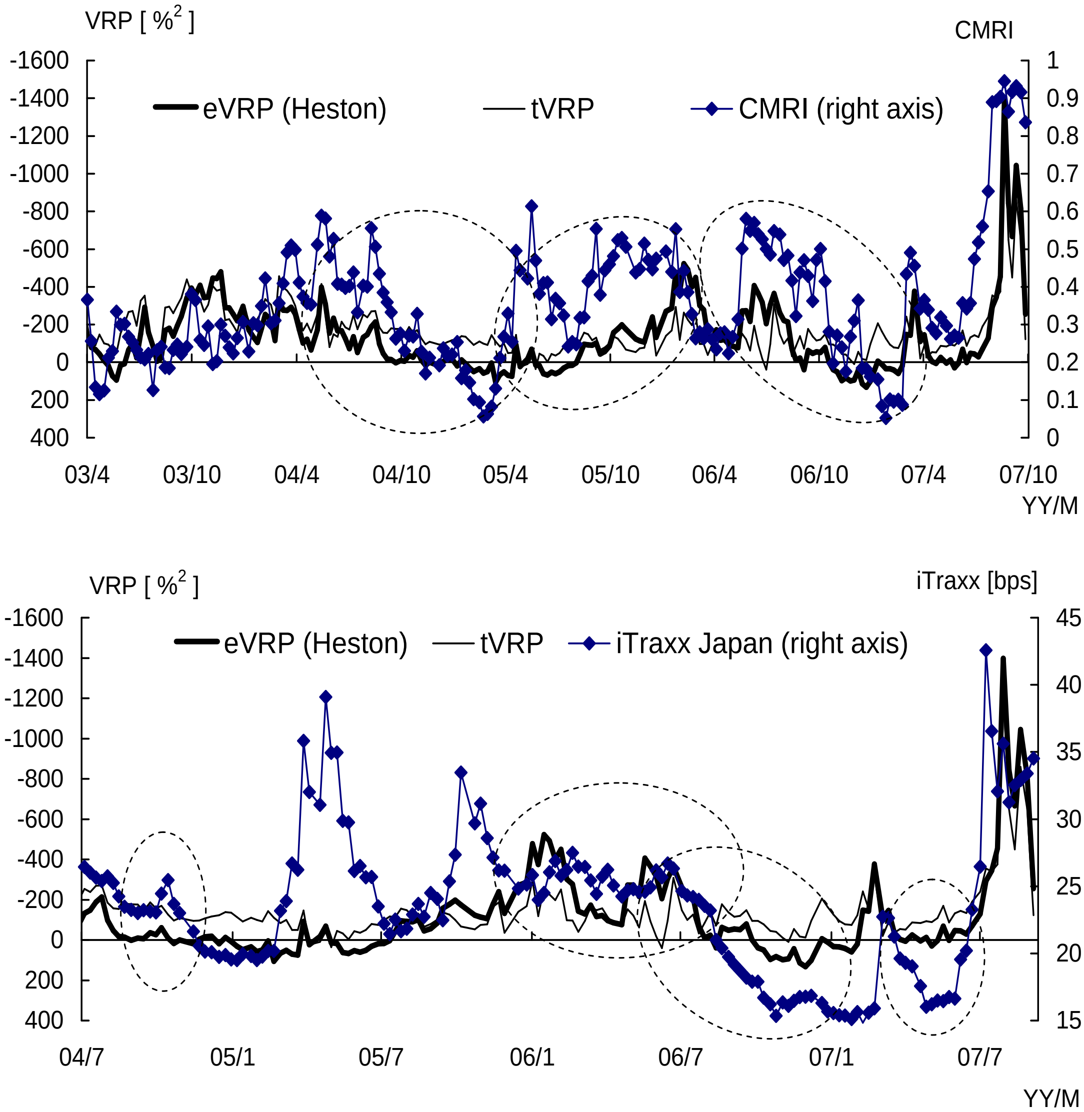
	CMRI	iTraxx Japan
eVRP (Heston)	-0.5935	-0.5045
eVRP (ARFIMAX)	-0.5680	-0.4424
tVRP	-0.4008	-0.3078
rVRP	-0.2283	-0.2057

Notes: The correlations with the CMRI are computed using data from April 2003 to September 2007, whereas those with iTraxx use data from July 2004 to September 2007. VRPs with one month term to maturity are used.

5.2 Prediction for the future

We further investigate the prediction accuracy of the eVRP and tVRP for the realized VRP, the future IRV and the future increase in IRV. We will look into the root mean square error, or RMSE, for the prediction accuracy of VRP for the rVRP, the future IRV

Figure 4: Time series of eVRPs, tVRP and the risk indicators



Notes: The upper and lower panel displays the eVRP, tVRP with the CMRI and the iTraxx Japan respectively. Those are plotted weekly to show the direction of the move. For the eVRP, the Heston model case is shown and for every VRP, one-month-term to maturity VRP are shown. Note that the periods in horizontal axes differ between upper and lower panels due to the data availability. The dotted circles indicate the periods when the movements of eVRPs move more consistently with the risk indicators.

Table 3: Prediction error of the eVRPs and tVRP for the rVRP measured by RMSE defined in eqs.(22) to (24)

	realized VRP	future IRV	future increase in IRV
eVRP(Heston)	0.4394	0.7548	0.8138
eVRP(ARFIMAX)	0.5590	0.7977	0.9575
tVRP	0.3464	0.6015	0.7411

and the increase in IRV, defined respectively as:

$$\text{RMSE(realized VRP)} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(\lambda(t_i, T_i) - \overline{\lambda}(t_i, T_i)\right)^2}, \tag{22}$$

$$\text{RMSE(future IRV)} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(-\lambda(t_i, T_i) - \sigma_{IRV}^2(t_i, T_i)\right)^2}, \tag{23}$$

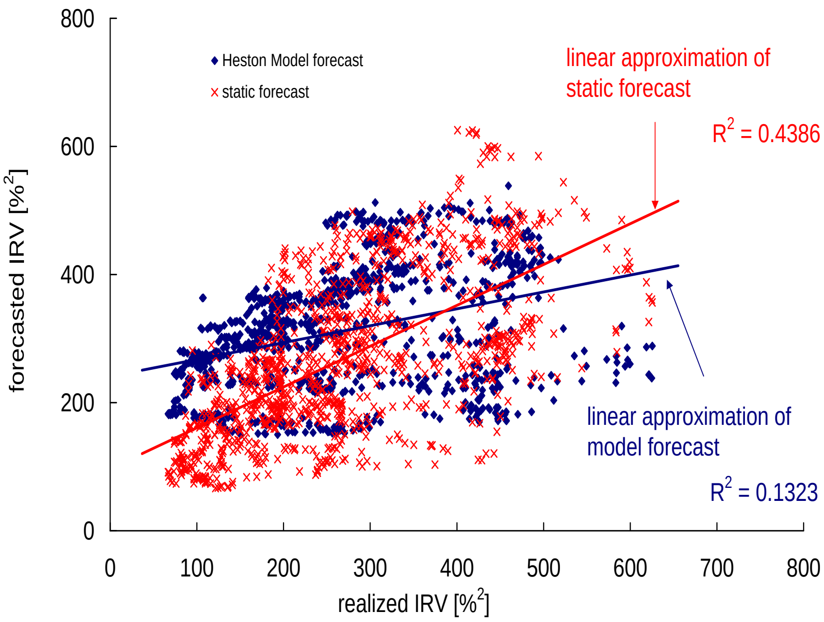
$$\text{RMSE(future }\Delta\text{IRV)} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(-\lambda(t_i, T_i) - \sigma_{IRV}^2(t_i, T_i) + \sigma_{IRV}^2(2t_{i-1} - T_i, t_i)\right)^2}, \tag{24}$$

where m is a number of observations. In eq.(23) and eq.(24), the sign of λ is minus as we assume negative VRP predicts the future level or future increase in IRV. Note that the prediction accuracy does not matter for the evaluation of expected VRP, since unpredictable shocks in future volatility may diverse realized RV from forecasted RV.

Table 3 displays the RMSE of the eVRPs with one to six months term to maturity. As seen, the eVRP with Heston model has a higher prediction accuracy than that with ARFIMAX for all cases. Table 3 also displays the tVRP's prediction accuracy. The tVRP has a smaller error and more accurately predicts the rVRP than the eVRPs for each case. Though those difference in error is relatively small, especially between eVRP in Heston's and tVRP, this result indicates that the market expectation of future volatility risk does not always correctly predict the future, and the simple linear estimation gives the better prediction. Since the tVRP simply extends the present level of RV for the future and neglect the mean reverting property of RV process, the tVRP fluctuates largely enough to cover the wide range of realized RV (see Figure 5). On the other hand, the forecasted RV stays around middle range of realized one in Figure 5, which may cause the worse performance in RMSE.

Now, we face a question that why the eVRPs show stronger correlation with risk indicators in spite of their low prediction power. Possible hypotheses are the following: (a) the formation of the market expectation cannot be measured only by ex post performance of prediction. Because there are some future unpredictable shocks, the prediction error of

Figure 5: Realized IRV and forecasted IRV: comparison between the Heston model forecasts and static forecasts in the tVRP



the expected VRP become high. (b) The mean reverting property of RV models in the eVRP might explain the formation of the market expectation better than the static expectation assumption on the tVRP, and (c) even though the models has better prediction accuracy for one day ahead, the prediction accuracy up to maturity become worsen due to the accumulation of uncertainty in the model on each day up to maturity.

5.3 Volatility risk aversion

In Section 4, we define another volatility risk aversion indicator: $\Gamma(t, T)$. Figure 6 displays the estimation results of Γ. The upper and lower panels display the Γ computed under the Heston model and the ARFIMAX model respectively.

First, low Γ period is much shorter than high Γ period. This indicates that a risk seeking market vanishes easier than a risk averse market and once the market goes into risk averse the state tends to continue.

Second, the shorter-term-to-maturity Γ precedes the longer-term Γ when the Γ is low, whereas the longer-term Γ precede the shorter-term Γ when the Γ is high. This indicates that the longer the term of the Γ the more risk sensitive the market.

Third, the ARFIMAX-Γ moves more drastically than Heston-Γ. That indicates the the distribution of model projected IRV is narrower in ARFIMAX than that in Heston.

6 Conclusion

This paper addresses the filtration inconsistency problem regarding the estimation of expected volatility risk premium, and proposes a new evaluation method which resolves the problem. We find that the volatility risk premium computed by our method has stronger correlation with risk indicators for general market risk preferences than those computed by the conventional methods. This suggests that our measure is a better indicator of markets' aversion of volatility risk. We also propose a new measure of markets' volatility risk aversion.

Our method depends on RV models to describe how market participants form their expectation on the future volatility. We do not directly examine the relevance of the RV models due to difficulty in obtaining the expectation. This issue is remained for our future challenge.

Although our volatility risk measures are assessed to be the better risk aversion indicator, we also find that the tVRP, the VRP computed under the simple linear assumption of RV process, has more accurately predicts the future than the eVRP. Those results indicates that even the model-base expectation for future risk has difficulty in predicting the future correctly, and much simpler linear estimation provides a better prediction accuracy.

References

Andersen, T.G. and T. Bollerslev, "Intraday Periodicity and Volatility Persistence in Financial Markets," *Journal of Empirical Finance*, 4, pp. 115–158, 1997.

——— "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts," *International Economic Review*, 39, pp. 885–905, 1998.

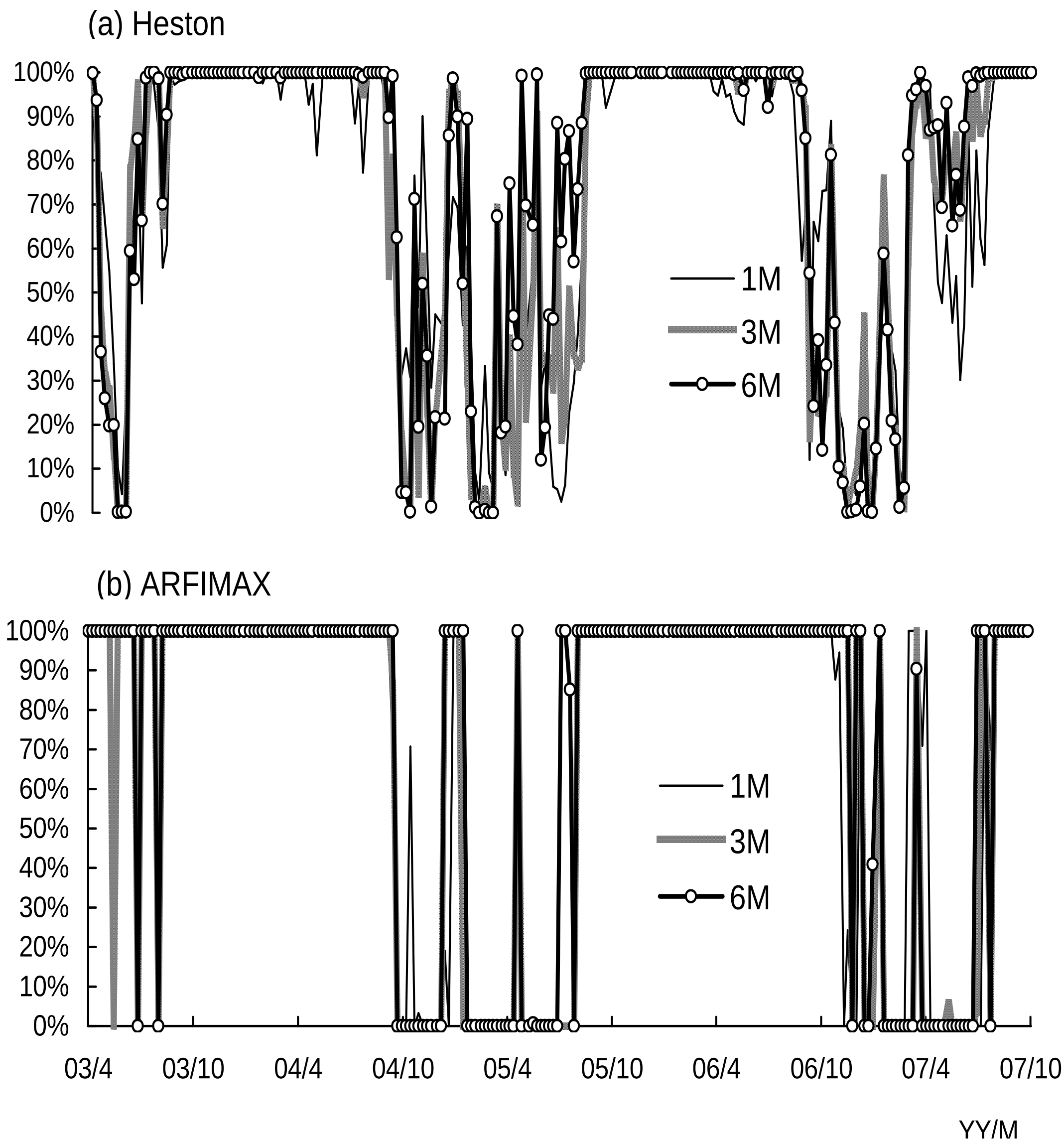
Bakshi, G. and N. Kapadia, "Delta-Hedged Gains and the Negative Market Volatility Risk Premium," *The Review of Financial Studies*, 16(2), pp. 527–566, 2003.

Bollerslev, T., M. Gibson, and H. Zhou, "Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities," Finance and Economics Discussion Series, 2007.

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

Carr, P. and L. Wu, "Variance Risk Premia," Working Paper, 2005.

Figure 6: Time series of Γ



Notes: (a) displays the Γ with Heston model and (b) displays Γ with the ARFIMAX model. One, three, six months term of Γ are shown.

Christensen, B.J., C.S. Hansen, and N.R. Prabhala, "The Telescoping Overlapping Problem in Options Data," Working Paper, University of Maryland, 2002.

Demeterfi, K., E. Derman, M. Kamal, and J. Zou, "More Than You Ever Wanted To Know About Volatility Swaps," *Quantitative Strategies Research Notes, Goldman Sachs*, March, 1999.

Derman, E., *Volatility Smile Lecture Notes*, Master of Financial Engineering Program, Columbia University, Lecture 10, pp.10–14, 2007.

Dupire, B., "Arbitrage Pricing with Stochastic Volatility," *Societe Generale*, 1992.

Giot, P. and S. Laurent, "Modelling Daily Value-at-Risk Using Realized Volatility and ARCH Type Models," *Journal of Empirical Finance*, 11(3), pp. 379–398, 2004.

Heston, S., "A Closed Form Solutiob for Options with Stochastic Volatility with Applications to Bond and Currency Options," *The Review of Financial Studies*, 6, pp. 327–343, 1993.

Hull, J. and A. White, "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance*, 42(2), pp. 281–300, 1987.

Jiang, G.J. and Y.S. Tian, "Extracting Model-Free Volatility from Option Prices: An Examination of the VIX Index," *Journal of Derivatives*, Spring, pp. 35–60, 2007.

Lee, R.W., "Implied and Local Volatility under Stochastic Volatility," *International Journal of Theoretical and Applied Finance*, 4(1), pp. 45–89, 2001.

McAleer, M. and M.C. Mederios, "Realized Volatility: A Review," *Econometric Review*, 27(1-3), pp. 10–45, 2008.

Neuberger, A., "Volatility Trading," *Working Paper, London Business School*, 1990.

Scott, L.O., "Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application," *Journal of Financial and Quantitative Analysis*, 22(4), pp. 419–438, 1987.

Shibata, M., "Estimation for Volatility using Intraday Market Data: Survey for Realized Volatility and Application for Japanese Stock Index and Stock Index Futures," *Kin'yu kenkyu (Monetary and Economic Studies)*, 27(1), pp. 1–54, 2008. in Japanese.

Sugihara, Y., "Model Free Implied Volatility in Nikkei Index Options Market," IMES Discussion Paper Series, Institute for Monetary Economic Studies, Bank of Japan, 2008 *to appear*.

Uchida, Y. and K. Miyazaki, "Volatility Risk Premium in Nikkei 225 Index Options Market," *Gendai Finance*, 23, pp. 35–59, 2008. in Japanese.

Watanabe, T. and K. Sasaki, "The Volatility Estimation by an ARCH howpublished Model and the Realized Volatility, and the Value at Risk," *Kin'yu kenkyu (Monetary and Economic Studies)*, pp. 39–74, 2007. in Japanese.

Wiggins, J.B., "Option Values under Stochastic Volatilities," *Journal of Financial Economics*, 19, pp. 351–372, 1987.

——— "The Pricing of Foreign Currency Options with Stochastic Volatility," *Journal of Econometrics*, 45, pp. 239–265, 1990.

A Volatility Risk Premium under Stochastic Volatility Model

In this section, we consider the volatility risk premium using SV model. Many of earlier works captured VRP under the SV model and we show how their works. We also show our VRP is equivalent with the volatility risk premium under SV model.

A.1 Derivation of PDE of derivatives on SV model

First, we derive the partial differential equation of a derivative on an asset under SV model, following Scott [1987]. Suppose an asset S_t follows a general SV model,

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma_t S_t dW_t^{(S)}, \\ d\sigma_t^2 &= \alpha(S_t, \sigma_t^2, t)dt + \beta(S_t, \sigma_t^2, t)dW_t^{(\sigma)}, \\ dW_t^{(S)} dW_t^{(\sigma)} &= \rho_t dt, \end{aligned} \tag{A-1}$$

where $dW_t^{(S)}$ and $dW_t^{(\sigma)}$ denote correlated Weiner processes, and μ, α, β and ρ_t denote parameters. We assume the asset is not a dividend paying asset.

Considering two derivatives on the asset, which has the value $f_1(S_t, \sigma_t^2, t)$ and $f_2(S_t, \sigma_t^2, t)$. Both depend on the asset price S_t and volatility σ_t, but has different parameters.

Suppose an investor holding one of the derivatives, who hedge off the risk of the volatility and the volatility's volatility by the asset and the other derivative. The investor longs one contract of the derivative f_1 and shorts δ_1 of the underlying asset and δ_2 contracts of the derivative f_2. The value of the portfolio Π_t is

$$\Pi_t = f_1(S_t, \sigma_t^2, t) - \delta_1 S_t - \delta_2 f_2(S_t, \sigma_t^2, t). \tag{A-2}$$

By Ito's lemma, the growth of the portfolio value become

$$\begin{aligned} d\Pi_t &= \left(\frac{\partial f_1}{\partial t} dt + \frac{\partial f_1}{\partial S_t} dS_t + \frac{\partial f_1}{\partial \sigma_t^2} d\sigma_t^2 + \frac{1}{2}\sigma_t^2 S_t^2 \frac{\partial^2 f_1}{\partial S_t^2} dt + \frac{1}{2}\beta^2 \frac{\partial^2 f_1}{(\partial \sigma_t^2)^2} dt + \sigma_t S_t \beta \rho_t \frac{\partial^2 f_1}{\partial \sigma_t^2 \partial S_t} dt \right) \\ &\quad -\delta_1 dS_t \\ &\quad -\delta_2 \left(\frac{\partial f_2}{\partial t} dt + \frac{\partial f_2}{\partial S_t} dS_t + \frac{\partial f_2}{\partial \sigma_t^2} d\sigma_t^2 + \frac{1}{2}\sigma_t^2 S_t^2 \frac{\partial^2 f_2}{\partial S_t^2} dt + \frac{1}{2}\beta^2 \frac{\partial^2 f_2}{(\partial \sigma_t^2)^2} dt + \sigma_t S_t \beta \rho_t \frac{\partial^2 f_2}{\partial \sigma_t^2 \partial S_t} dt \right) \\ &= \left[\left(\frac{\partial f_1}{\partial t} + \frac{1}{2}\sigma_t^2 S_t^2 \frac{\partial^2 f_1}{\partial S_t^2} + \frac{1}{2}\beta^2 \frac{\partial^2 f_1}{(\partial \sigma_t^2)^2} + \sigma_t S_t \beta \rho_t \frac{\partial^2 f_1}{\partial \sigma_t^2 \partial S_t} \right) \right. \\ &\qquad \left. - \delta_2 \left(\frac{\partial f_2}{\partial t} + \frac{1}{2}\sigma_t^2 S_t^2 \frac{\partial^2 f_2}{\partial S_t^2} + \frac{1}{2}\beta^2 \frac{\partial^2 f_2}{(\partial \sigma_t^2)^2} + \sigma_t S_t \beta \rho_t \frac{\partial^2 f_2}{\partial \sigma_t^2 \partial S_t} \right) \right] dt \\ &\qquad\qquad + \left(\frac{\partial f_1}{\partial S_t} - \delta_2 \frac{\partial f_2}{\partial S_t} - \delta_1 \right) dS_t + \left(\frac{\partial f_1}{\partial \sigma_t^2} - \delta_2 \frac{\partial f_2}{\partial \sigma_t^2} \right) d\sigma_t^2. \end{aligned} \tag{A-3}$$

In order to hedge the risk in the dS_t and $d\sigma_t$, the investor should chooses δ_1 and δ_2 so as to the last two terms in eq.(A-3) become zero, i.e.,

$$\frac{\partial f_1}{\partial S_t} - \delta_2 \frac{\partial f_2}{\partial S_t} - \delta_1 = 0, \quad \frac{\partial f_1}{\partial \sigma_t^2} - \delta_2 \frac{\partial f_2}{\partial \sigma_t^2} = 0. \tag{A-4}$$

then, the δ_1 and δ_2 are computed as

$$\delta_1 = \frac{\partial f_1}{\partial S_t} - \frac{\partial f_2}{\partial S_t} \frac{\partial f_1}{\partial \sigma_t^2} / \frac{\partial f_2}{\partial \sigma_t^2}, \quad \delta_2 = \frac{\partial f_1}{\partial \sigma_t^2} / \frac{\partial f_2}{\partial \sigma_t^2}. \tag{A-5}$$

This makes the portfolio value Π_t riskless. Since the return of the risk free portfolio should be equal to the risk free rate,

$$d\Pi_t = r_t \Pi_t dt, \tag{A-6}$$

where r_t is the risk free rate. Applying eq.(A-2), eq.(A-3) and eq.(A-5) to eq.(A-6) yields

$$\begin{aligned} & r_t f_1 - \left(\frac{\partial f_1}{\partial S_t} - \frac{\partial f_2}{\partial S_t} \frac{\partial f_1}{\partial \sigma_t^2} / \frac{\partial f_2}{\partial \sigma_t^2} \right) r_t S_t - \frac{\partial f_1}{\partial \sigma_t^2} / \frac{\partial f_2}{\partial \sigma_t^2} r_t f_2 \\ & = \left(\frac{\partial f_1}{\partial t} + \frac{1}{2} \sigma_t^2 S_t^2 \frac{\partial^2 f_1}{\partial S_t^2} + \frac{1}{2} \beta^2 \frac{\partial^2 f_1}{(\partial \sigma_t^2)^2} + \sigma_t S_t \beta \rho_t \frac{\partial^2 f_1}{\partial \sigma_t^2 \partial S_t} \right) \\ & - \frac{\partial f_1}{\partial \sigma_t^2} / \frac{\partial f_2}{\partial \sigma_t^2} \cdot \left(\frac{\partial f_2}{\partial t} + \frac{1}{2} \sigma_t^2 S_t^2 \frac{\partial^2 f_2}{\partial S_t^2} + \frac{1}{2} \beta^2 \frac{\partial^2 f_2}{(\partial \sigma_t^2)^2} + \sigma_t S_t \beta \rho_t \frac{\partial^2 f_2}{\partial \sigma_t^2 \partial S_t} \right). \end{aligned} \tag{A-7}$$

Defining an operator $\mathcal{V}$ as

$$\mathcal{V} = \left(\frac{\partial}{\partial \sigma_t^2} \right)^{-1} \left(\frac{\partial}{\partial t} + r_t S_t \frac{\partial}{\partial S_t} + \frac{1}{2} \sigma_t^2 S_t^2 \frac{\partial^2}{\partial S_t^2} + \frac{1}{2} \beta^2 \left(\frac{\partial}{\partial \sigma_t^2} \right)^2 + \sigma_t S_t \beta \rho_t \frac{\partial^2}{\partial \sigma_t^2 \partial S_t} - r_t \right), \tag{A-8}$$

eq.(A-7) can be rewritten as

$$\mathcal{V} f_1 = \mathcal{V} f_2 \tag{A-9}$$

Since the left and right hand side of eq.(A-9) are function of derivative f_1 and f_2 respectively, the value should be completely independent from the variables of the derivative, such as strike price or maturity. We can write the value as $\mathcal{V} f = -\varphi(S_t, \sigma_t^2, t)$.

Then any derivative f on the asset S satisfy the following PDE.

$$\frac{\partial f}{\partial t} + r_t S_t \frac{\partial f}{\partial S_t} + \varphi \frac{\partial f}{\partial \sigma_t^2} + \frac{1}{2} \sigma_t^2 S_t^2 \frac{\partial^2 f}{\partial S_t^2} + \frac{1}{2} \beta^2 \frac{\partial^2 f}{(\partial \sigma_t^2)^2} + \sigma_t S_t \beta \rho_t \frac{\partial^2 f}{\partial \sigma_t^2 \partial S_t} = r_t f. \tag{A-10}$$

That is the PDE for an derivative value with stochastic volatility, hereafter SV-PDE, which considered to be the SV version of the Black-Scholes-Merton's PDE.

A.2 Definition of VRP

The VRP is defined under SV model with same analogy with the market price of risk. We follow the Derman [2007] in the following discussion. The change of the value of an derivative f can be written by Ito's lemma as:

$$\begin{aligned} df &= \left[\frac{\partial f}{\partial t} + \mu S_t \frac{\partial f}{\partial S_t} + \alpha \frac{\partial f}{\partial \sigma_t^2} + \frac{1}{2}\sigma_t^2 S_t^2 \frac{\partial^2 f}{\partial S_t^2} + \frac{1}{2}\beta^2 \frac{\partial^2 f}{(\partial \sigma_t^2)^2} + S_t \sigma_t \beta \rho_t \frac{\partial^2 f}{\partial S_t \partial \sigma_t^2}\right] dt \\ &\quad + \sigma_t S_t \frac{\partial f}{\partial S_t} dW_t^{(S)} + \beta \frac{\partial f}{\partial \sigma_t^2} dW_t^{(\sigma)}. \end{aligned} \quad \text{(A-11)}$$

Letting μ_f, $\sigma_f^{(S)}$, $\sigma_f^{(\sigma)}$ denote f's drift, volatility for $dW_t^{(S)}$ and for $dW_t^{(\sigma)}$ respectively, we view the eq.(A-11) as the two dimensional geometric Brownian motion as

$$\frac{df}{f} = \mu_f dt + \sigma_f^{(S)} dW_t^{(S)} + \sigma_f^{(\sigma)} dW_t^{(\sigma)}, \quad \text{(A-12)}$$

where

$$\mu_f = \frac{1}{f}\left[\frac{\partial f}{\partial t} + \mu S_t \frac{\partial f}{\partial S_t} + \alpha \frac{\partial f}{\partial \sigma_t^2} + \frac{1}{2}\sigma_t^2 S_t^2 \frac{\partial^2 f}{\partial S_t^2} + \frac{1}{2}\beta^2 \frac{\partial^2 f}{(\partial \sigma_t^2)^2} + S_t \sigma_t \beta \rho \frac{\partial^2 f}{\partial S_t \partial \sigma_t^2}\right], \quad \text{(A-13)}$$

$$\sigma_f^{(S)} = \frac{S_t \sigma_t}{f} \frac{\partial f}{\partial S_t}, \quad \text{(A-14)}$$

$$\sigma_f^{(\sigma)} = \frac{\beta}{f} \frac{\partial f}{\partial \sigma_t^2}. \quad \text{(A-15)}$$

Applying eq.(A-10) to eq.(A-13), we get

$$\mu_f = \frac{1}{f}\left[(\mu - r_t) S_t \frac{\partial f}{\partial S_t} + (\alpha - \varphi)\frac{\partial f}{\partial \sigma_t^2} + r_t f\right]. \quad \text{(A-16)}$$

Eq.(A-16) can be written using eq.(A-14) and eq.(A-15) as:

$$\mu_f - r_t = \frac{\sigma_f^{(S)}}{\sigma_t}(\mu - r_t) + \frac{\sigma_f^{(\sigma)}}{\beta}(\alpha - \varphi). \quad \text{(A-17)}$$

We consider $\mu - r_t$ to be a risk premium of $dW_t^{(S)}$ or that on the asset price itself. By the same logic, $\alpha - \varphi$ is considered to be a risk premium of $dW_t^{(\sigma)}$ or the volatility risk premium. In consequence, φ is considered to be a kind of "risk free rate" of volatility. The eq.(A-17) shows the risk premium on the derivative is the weighted sum of the risk premium of the asset and the volatility. Eq.(A-17) is called "the market price of risk equation".

We guess our VRP, λ, equals to the volatility risk premium instantaneously. That means $\lambda(t,t)$ has the form

$$\lambda(t,t) = \alpha - \varphi. \quad \text{(A-18)}$$

We have showed λ is negative in most periods. The overall risk premium in a derivative with positive vega is smaller than the risk premium of the underlying asset itself, since the overall risk premium of a derivative is composed of the risk premium of underlying asset and its volatility as in eq.(A-17) and $\alpha - \varphi$ in eq.(A-17) is negative.

Further setting

$$\sigma_f = \sqrt{\left(\sigma_f^{(S)}\right)^2 + \left(\sigma_f^{(\sigma)}\right)^2 + 2\rho_t \sigma_f^{(S)} \sigma_f^{(\sigma)}}, \tag{A-19}$$

we get the Sharpe ratio version of eq.(A-17) as

$$\frac{\mu_f - r_t}{\sigma_f} = \frac{\sigma_f^{(S)}}{\sigma_f} \frac{\mu - r_t}{\sigma_t} + \frac{\sigma_f^{(\sigma)}}{\sigma_f} \frac{\alpha - \varphi}{\beta}. \tag{A-20}$$

The Sharpe ratio, or the market value of risk, of the derivative is weighted sum of those of the underlying asset and its volatility. Since $\sigma_f \leq \sigma_f^{(S)} + \sigma_f^{(\sigma)}$, the Sharpe ratio of the derivative is smaller than the weighted average of the Sharpe ratios of the asset and the volatility unless the asset and the volatility is perfectly correlated.

A.3 Relation between $\alpha - \varphi$ and λ

The risk neutral representation of the stochastic volatility model, eq.(A-2), changes to the following:

$$\begin{aligned} dS_t &= r_t S_t dt + \sigma_t S_t dW_t^{(\tilde{S})}, \\ d\sigma_t^2 &= \varphi(S_t, \sigma_t^2, t)dt + \beta(S_t, \sigma_t^2, t)dW_t^{(\tilde{\sigma})}, \\ dW_t^{(\tilde{S})} dW_t^{(\tilde{\sigma})} &= \rho_t dt, \end{aligned} \tag{A-21}$$

where

$$\begin{aligned} dW_t^{(\tilde{S})} &= dW_t^{(S)} + \frac{\mu - r_t}{\sigma_t} dt, \\ dW_t^{(\tilde{\sigma})} &= dW_t^{(\sigma)} + \frac{\alpha - \varphi}{\beta} dt. \end{aligned} \tag{A-22}$$

For simplicity, we assume α and β in eq.(A-2) and φ in eq.(A-22) are function of t and S_t, and do not depend on σ_t^2 from now. The integrated volatility under the risk neutral measure $\mathcal{Q}$ can be computed as

$$\begin{aligned} E^{\mathcal{Q}}[\langle S \rangle_{t,T} | \mathcal{F}_t] &= \frac{1}{T-t} E^{\mathcal{Q}} \int_t^T \sigma_u^2 du \\ &= \frac{1}{T-t} \int_t^T \left[\sigma_t^2 + \int_t^u \varphi dv \right] du, \end{aligned} \tag{A-23}$$

whereas the integrated volatility under the physical measure $\mathcal{P}$ is

$$
\begin{aligned}
E^{\mathcal{P}}[\langle S\rangle_{t,T}|\mathcal{F}_t] &= \frac{1}{T-t}E^{\mathcal{P}}\int_t^T \sigma_u^2 du \\
&= \frac{1}{T-t}\int_t^T \left[\sigma_t^2 + \int_t^u \alpha dv\right] du. \qquad \text{(A-24)}
\end{aligned}
$$

These difference (eq.(A-23)−eq.(A-24)) is considered to be the true VRP, $\tilde{\lambda}(t,T)$, as defined in eq.(11), i.e.:

$$
\begin{aligned}
\tilde{\lambda}(t,T) &= E^{\mathcal{P}}[\langle S\rangle_{t,T}|\mathcal{F}_t] - E^{\mathcal{Q}}[\langle S\rangle_{t,T}|\mathcal{F}_t] \\
&= \frac{1}{T-t}\int_t^T\int_t^u (\alpha_v - \varphi_v) dv du. \qquad \text{(A-25)}
\end{aligned}
$$

Instantaneously,

$$
\begin{aligned}
\lambda(t,t) &= \frac{\partial}{\partial T}\tilde{\lambda}(t,T)\bigg|_{T\to t} \\
&= \frac{\partial}{\partial T}\left[\lim_{T\to t}\frac{1}{T-t}\int_t^T\int_t^u (\alpha_v - \varphi_v) dv du\right] \\
&= \frac{\partial}{\partial u}\int_t^u (\alpha_v - \varphi_v) dv\bigg|_{u\to t} \\
&= \alpha_t - \varphi_t, \qquad \text{(A-26)}
\end{aligned}
$$

as we guess in eq.(A-18).

A.4 Relation between dynamic delta hedge and VRP

From eq.(A-17), we can compute the value of the delta hedged portfolio. In Section A.1, we assume a portfolio that consists of two derivatives and one underlying asset to hedge off the risk of both the asset and the volatility. Here, we assume a portfolio that hedges the risk of asset only. We call this a delta-hedged portfolio. In the stochastic volatility setting, the delta of a derivative is also stochastic, therefore we assume the portfolio manager dynamically rebalances the position every time the delta changes. Let Π_t denotes the value of the delta-hedged portfolio at time t. The change of the portfolio value is computed as:

$$
\begin{aligned}
d\Pi_t &= f_t + df_t - \frac{\partial f_t}{\partial S_t}(S_t + dS_t) + (1 + r_t dt)\left(\frac{\partial f_t}{\partial S_t}S_t - f_t\right) \\
&= df_t - \frac{\partial f_t}{\partial S_t}dS_t - r_t\left(f_t - S_t\frac{\partial f_t}{\partial S_t}\right)dt. \qquad \text{(A-27)}
\end{aligned}
$$

By eq.(A-12), eq.(A-13), eq.(A-14) and eq.(A-15),

$$\begin{aligned} df_t &= \mu_f f_t dt + \sigma_f^{(S)} f_t dW_t^{(S)} + \sigma_f^{(\sigma)} f_t dW_t^{(\sigma)} \\ &= \left[(\mu - r_t) S_t \frac{\partial f}{\partial S_t} + (\alpha - \varphi) \frac{\partial f}{\partial \sigma_t^2} + r_t f \right] dt \\ &\qquad + S_t \sigma_t \frac{\partial f}{\partial S_t} dW_t^{(S)} + \beta \frac{\partial f}{\partial \sigma_t^2} dW_t^{(\sigma)}. \end{aligned} \tag{A-28}$$

Eq.(A-27) can be rewritten using eq.(A-28) and eq.(A-2) as:

$$d\Pi_t = (\alpha_t - \varphi_t) \frac{\partial f}{\partial \sigma_t^2} dt + \beta \frac{\partial f}{\partial \sigma_t^2} dW_t^{(\sigma)} \tag{A-29}$$

Hence, the value of the portfolio hedged from time t to time T is given by integrating both sides of eq.(A-29) as:

$$\Pi_T = \int_t^T (\alpha_u - \varphi_u) \frac{\partial f}{\partial \sigma_u^2} du + \int_t^T \beta \frac{\partial f}{\partial \sigma_u^2} dW_u^{(\sigma)}. \tag{A-30}$$

The expected value of eq.(A-30) become

$$E_t \Pi_T = E_t \left[\int_t^T (\alpha_u - \varphi_u) \frac{\partial f}{\partial \sigma_u^2} du \right] = E_t \left[\int_t^T \lambda(t,u) \frac{\partial f}{\partial \sigma_u^2} du \right]. \tag{A-31}$$

Eq.(A-31) indicates that the growth of the value of the delta-hedged portfolio is proportional to the volatility risk premium instantaneously, when we assume the portfolio manager can perfectly and dynamically hedge off the risk of the asset. When considering an equity option for the derivative f, the delta hedge gain may be negative, because the option has a positive vega and we know $\alpha - \varphi = \lambda$ is negative in most periods.

A.5 Relation between investors' risk aversion and VRP

Many earlier works, such as Heston [1993], Bakshi and Kapadia [2003] and Bollerslev *et al.* [2007], discussed the VRP related to the investors' utility. In their works, the pricing kernel in eq.(10) is supposed to be equal to the marginal utility of the representative investor whose utility function is assumed to be a power function of the investor's wealth w_t, which is defined as:

$$U(w_t) = e^{-rt} \frac{w_t^{1-\gamma}}{1-\gamma} \tag{A-32}$$

where the parameter γ indicates the investor's risk aversion; the larger the γ the more risk averse the investor. In wealth based asset pricing model, the pricing kernel defined in eq.(10) is equal to the inter-temporal marginal rate of substitutions, i.e.:

$$M_{t,T} = \frac{U'(w_T)}{U'(w_t)} = e^{-r(T-t)} \left(\frac{w_T}{w_t} \right)^{-\gamma} \tag{A-33}$$

Suppose the representative investor invests in one share of a stock whose price obeys standard SV model in eq.(A-2). Then, the pricing kernel become $M_{t,T} = e^{-r(T-t)}(S_T/S_t)^{-\gamma}$. The instantaneous risk premium for the investor can be computed from eq.(1), eq.(11) and Ito's lemma as:

$$\begin{aligned} \lambda(t,t) &= -\frac{\partial}{\partial T} \text{Cov}\left(M_{t,T}, \langle S \rangle_{t,T}\right)\Big|_{T=t} \\ &= -\text{Cov}\left(\frac{\partial}{\partial T} M_{t,T}\Big|_{T=t}, \frac{\partial}{\partial T} \langle S \rangle_{t,T}\Big|_{T=t} \right) \\ &= \gamma \text{Cov}\left(\frac{dS_t}{S_t}, \left(\frac{dS_t}{S_t}\right)^2 \right) \\ &= \gamma \beta_t \rho_t dt. \end{aligned} \tag{A-34}$$

Integrating both sides yields

$$\overline{\lambda}(t,T) = \gamma \int_t^T \beta_s \rho_s ds. \tag{A-35}$$

That shows the VRP is proportional to the investors' risk aversion γ.