IMES DISCUSSION PAPER SERIES

Stock Return Predictability and Variance Risk Premia around the ZLB

Toshiaki Ogawa, Masato Ubukata, and Toshiaki Watanabe

Discussion Paper No. 2020-E-9

IMES

INSTITUTE FOR MONETARY AND ECONOMIC STUDIES

BANK OF JAPAN

2-1-1 NIHONBASHI-HONGOKUCHO
CHUO-KU, TOKYO 103-8660
JAPAN

IMES Discussion Paper Series 2020-E-9
July 2020

Stock Return Predictability and Variance Risk Premia around the ZLB

Toshiaki Ogawa*, Masato Ubukata, and Toshiaki Watanabe*****

Abstract

We make an empirical analysis of whether and how variance risk premia (VRP) contribute to predicting excess stock returns in the US and Japan. Our new findings to be added to the literature are that (i) the correlation between VRP and future excess returns in the US is insignificant when the risk-free rate is close to zero, and (ii) the correlation in Japan is significantly negative. To explain these findings, we also conduct a preliminary theoretical analysis with a structural model of asset pricing based on two assumptions: the zero lower bound (ZLB) for the risk-free rate, and a negative correlation between the consumption growth rate and the volatility-of-volatility. These allow excess returns to follow a hump-shaped pattern. This affects the sign and significance of the correlation of the returns with the VRP.

Keywords: Excess returns; Heterogeneous autoregressive model; Nikkei 225; Realized volatility; S&P500; Variance risk premium; Zero lower bound
JEL classification: C52, C53, G17

* Deputy Director and Economist, Institute for Monetary and Economic Studies, Bank of Japan
(E-mail: toshiaki.ogawa@boj.or.jp)
** Professor, Faculty of Economics, Meiji Gakuin University
(E-mail: ubukata@eco.meijigakuin.ac.jp)
*** Professor, Institute of Economic Research, Hitotsubashi University
(E-mail: watanabe@ier.hit-u.ac.jp)

We thank the participants in the Workshop "Macro- and Financial Econometrics" in the 5th Hitotsubashi Summer Institute in 2019 and the 2019 Finance Workshop held by the Institute for Monetary and Economic Studies, the Bank of Japan. Ubukata acknowledges financial support from the Ministry of Education, Culture, Sports, Science and Technology of the Japanese Government through Grant-in-Aid for Scientific Research No.18K01690. Watanabe acknowledges financial support from the Ministry of Education, Culture, Sports, Science and Technology of the Japanese Government through Grant-in-Aid for Scientific Research No. 17H00985, 19H00588 and 20H00073, and from Hitotsubashi Institute for Advanced Study. The views expressed in this paper are those of the authors and do not necessarily reflect the official views of the Bank of Japan.

1 Introduction

Variance risk premium (VRP) is defined as the difference between the expectation of the quadratic variation (QV) of financial returns under the risk-neutral measure Q and that under the physical measure P, which we call the conditional variance (CV). Using an asset pricing model with time-varying volatility-of-volatility, Bollerslev et al. (2009) show that the VRP can help in predicting excess stock returns. Bollerslev et al. (2009) and Bekaert and Hoerova (2014) provide empirical evidence for this implication, using the VRP of S&P500 in the US. Ubukata and Watanabe (2014a) and Andersen et al. (2019), however, find that the VRP lacks the excess return predictability for the Nikkei 225 in Japan.

Bollerslev et al. (2014, Table 3) examine whether the VRP contributes to predicting excess stock returns in eight countries (France, Germany, Japan, Switzerland, the Netherlands, Belgium, the UK and the US). Specifically, they regress the monthly-averaged excess return (ER) of a stock index in each of the eight countries from month t to month $t+h$ on the VRP of that index for month t; namely, $\mathrm{ER}_{t:t+h} = \alpha + \beta \mathrm{VRP}_t + \epsilon_t$, where ϵ_t is residuals. They consider the cases of h =1, 2,...,6, 9, and 12 for the countries, and report the adjusted R^2 of these regressions. They find that the country-specific VRPs predict the returns in all of the countries except for Japan. The relevance of the VRP to future excess stock returns differs between countries. To investigate these differences in relation to countries' risk-free rates, we plot the adjusted R^2 for selected hs and the risk-free rates in the eight countries in Figure 1.[1] In Bollerslev et al.'s (2014) regressions for the selected hs, the VRPs are estimated to have statistically significant coefficients for six countries at one percent levels of statistical significance as indicated with $\bullet$, and for one country at a five percent level of statistical significance as indicated with $\circ$. Japan's VRP is estimated to have a statistically insignificant coefficient as indicated with $\times$. Figure 1 shows that a country with a higher (lower) risk-free rate has a higher (lower) adjusted R^2, and that Japan's adjusted R^2 is quite small.[2]

This article investigates the relevance of VRPs to predicting excess returns for Japan and the US. The period of our sample for Japan ranges from 2000.2 to 2018.9. In this period, Japan's risk-free rate is always close to zero, except for a few points in time. The sample

[1] For each country, we select a specific h which has the largest value of the adjusted R^2. The risk-free rates are "short-term interest rates" in *Main Economic Indicators* published by the OECD, and the averages of these interest rates for the 12 years from 2000−2011, the period analyzed in Bollerslev et al. (2014). Bollerslev et al. (2014, Table 3) also report the Newey-West t−statistics for the coefficients of the VRPs.

[2] Apart from the above-mentioned selection rule, we straightforwardly select the cases where hs are 3, 4, and 5. Even in these cases, we find the same positive relation between the adjusted R^2 and the risk-free rates across the eight countries.

period for the US ranges from 1992.1 to 2017.7. The US risk-free rate is close to zero between 2008.11 and 2016.10. We examine this subsample period separately.

[Figure 1 about here.]

The expectation of QV under Q can be obtained as the squared implied volatility from options prices. However, it is not appropriate to use the Black-Scholes model, because it assumes that volatility is constant over time. We use a model-free implied volatility, which is not based on the Black-Scholes model and is allowed to change over time. Specifically, we use the Volatility Index Japan (VXJ) published by Osaka University for the Nikkei 225 stock index in Japan, and VIX index from the Chicago Board Options Exchange (CBOE) for the S&P500 in the US. There are a number of methods for calculating CV. Traditional methods include the GARCH and the stochastic volatility models with daily returns. A recent method is to fit a model to realized variance (RV), which is the estimate of QV calculated using intraday returns. Bollerslev et al. (2009, 2014) assume that RV follows a simple random-walk process. Beyond this simple model, Corsi (2009) proposes a heterogeneous autoregressive (HAR) model, which has been extended and is now widely used as surveyed by Watanabe (2020). In this model, RV is a linear function of RVs with different frequency, such as daily, weekly and monthly RVs. Ubukata and Watanabe (2014a) calculate CV using this model to examine the excess return predictability of RV in Japan. If asset price is subject to discontinuous jumps, RV will have both continuous and jump components, and the persistence in these two components may be different. Andersen et al. (2007) extend this model to the HAR model with continuous and jump components (HAR-CJ) by separating RV into these two components and introducing them separately as explanatory variables. Bekaert and Hoerova (2014) extend the HAR-CJ model by introducing the squared VIX as an explanatory variable and taking account of the negative correlation between today's return and tomorrow's volatility, which is a well-known phenomenon in stock markets. To calculate CV, we use this model, which we call the asymmetric HAR-CJ model with the squared VIX.

The main empirical findings to be added to the literature are as follows. First, the correlation between the VRP and future excess returns in the US is not statistically significant in the subsample (2008.11–2016.10) when the risk-free rate is close to zero, while it is statistically significant in the full sample (1992.1-2017.7), which is consistent with the previous literature, such as Bollerslev et al. (2009, 2014) and Bekaert and Hoerova (2014). Second, the correlation between the VRP and future excess returns in Japan is negative and, unlike Bollerslev et al.

(2014), Ubukata and Watanabe (2014a), Uchiyama and Yamanaka (2015) and Andersen et al. (2019), statistically significant.

Finally, to mark a first step toward a theoretical explanation of these findings, we provide a simple asset pricing model by incorporating two new ingredients to Bollerslev et al. (2009): the zero lower bound (ZLB) for the risk-free rate, and a negative correlation between the consumption growth rate and the volatility-of-volatility. Since the risk-free rate can not be lower than zero, excess stock returns follow a hump-shaped pattern and have a kink when the risk-free rate hits the ZLB. On the left side of this kink (before the risk-free rate hits the ZLB), the correlation between the VRP and the excess returns is significantly positive, which may correspond to the case of the US when the risk-free rate is not close to zero. On the right hand side of the kink, the correlation is significantly negative, which may correspond to the case of Japan. Around the kink, the correlation becomes insignificant, which may correspond to the case of the US in the ZLB period.

The article proceeds as follows. Sections 2 reviews the VRP and how we estimate it. Section 3 explains the data used and the estimates of the VRP. Section 4 examines the excess return predictability afforded by the VRP in Japan and the US. Section 5 makes a model to understand our findings and discusses its associated limitations and potential improvements for future theoretical research on the VRP. Section 6 concludes.

2 Variance Risk Premium

Suppose that the log-price p of an asset follows a typical jump diffusion process below:

$$dp(t) = \mu(t)dt + \sigma(t)dW(t) + \kappa(t)dN(t), \tag{1}$$

where $\mu(t)$ is the drift, $\sigma(t)$ is the volatility, $W(t)$ is a standard Brownian motion, $\kappa(t)$ is the jump size, and $N(t)$ is a counting measure for the jumps.

Suppose that we observe n intraday returns $\{r_{t+1/n}, r_{t+2/n}, \ldots, r_{t+1}\}$ during day t. Then, daily quadratic variation on day t, QV_t, is defined as:

$$\mathrm{QV}_t = \operatorname*{plim}_{n\to\infty} \sum_{i=1}^{n} r_{t+i/n}^2. \tag{2}$$

Quadratic variation between day $t+1$ and $t+\tau$ is defined as:

$$\text{QV}_{t+1:t+\tau} = \sum_{i=1}^{\tau} \text{QV}_{t+i}. \tag{3}$$

Variance risk premium (VRP) is defined as the difference between the conditional expectation of QV under Q and that under P as:

$$\text{VRP}_t = E_t^Q[\text{QV}_{t+1:t+\tau}] - E_t^P[\text{QV}_{t+1:t+\tau}]. \tag{4}$$

As usual, we set $\tau = 22$ (one month).

$E_t^Q[\text{QV}_{t+1:t+\tau}]$ is calculated from option prices. However, it is not appropriate to use the Black-Scholes model, which assumes that volatility is constant over time. If the asset price follows the jump-diffusion process (1), the following equation holds approximately, even if the volatility changes over time (see Britten-Jones and Neuberger [2000] and Jiang and Tian [2005] for the proof):

$$E_t^Q[\text{QV}_{t+1:t+\tau}] \approx 2\exp(r\tau)\left[\int_0^F \frac{P(\tau,K)}{K^2}dK + \int_F^\infty \frac{C(\tau,K)}{K^2}dK\right], \tag{5}$$

where r is risk-free rate, τ is time to maturity, K is strike price, $P(\tau,K)$ is price of put option, $C(\tau,K)$ is price of call option and $F = S\exp(r\tau)$. Equation (5) is subject to a minor approximation error when the price jumps. However, this is immaterial in the context of predictive regressions (see Andersen et al. [2019] for details). Since the integrals in the right-hand-side cannot be solved analytically, they are approximated by the summation. In the US, CBOE calculates VIX from the prices of S&P500 options using this method. In Japan, Osaka University and Nikkei Publishing Inc. calculate VXJ and Nikkei VI respectively from the prices of Nikkei 225 options. In VIX and Nikkei VI, the summation with respect to K is calculated using the options traded in the market, which may cause a non-negligible bias because the number of options with different K traded in the market is small (Jiang and Tian [2005, 2007]). For Japan, we use VXJ because Osaka University calculates it by interpolating options with the strike prices which are not traded in the market (see Fukasawa et al. [2011] for the detail of VXJ). As such data is not available for the US, we use VIX. VIX and VXJ are the square root of equation (5), annualized. We take the squares of VIX and VXJ and divide them by 12, with the resulting figures being used for $E_t^Q[\text{QV}_{t+1:t+22}]$.

There are several methods for the computation of CV, which researchers tend to calculate by fitting a model to the realized variance (RV) using intraday returns. The simple daily RV is the sum of squared intraday returns over a day, and is written as:

$$\mathrm{RV}_t = \sum_{i=1}^{n} r_{t+i/n}^2. \tag{6}$$

RV between days $t+1$ and $t+\tau$ is

$$\mathrm{RV}_{t+1:t+\tau} = \sum_{i=1}^{\tau} \mathrm{RV}_{t+i}. \tag{7}$$

This simple RV runs the risk of being biased by market microstructure noises such as bid-ask bounce and nonsynchronous trading.[3] As discussed by Ubukata and Watanabe (2014b), some researchers have developed new methods to deal with any such bias, for example the realized kernel estimator of Barndorff-Nielsen et al. (2008, 2009), and the two- or multi-scale estimator of Zhang et al. (2005) and Zhang (2006).[4] Nevertheless, Liu et al. (2015) show that these methods have no advantage over the simple RV calculated using 5-minutes returns, in terms of ability to predict volatility. Therefore, we use such returns for the simple RV, rather than using these more complex methods.

Bollerslev et al. (2014) assume simply that $E_t^P[\mathrm{QV}_{t+1:t+22}] = \mathrm{QV}_{t-21:t}$ and estimate $\mathrm{QV}_{t-21:t}$ as $\mathrm{RV}_{t-21:t}$. However, this assumption is true only if monthly QV follows a random walk. Andersen et al. (2001a,b, 2003) have documented that daily RV may follow a long-memory process instead of a random walk, so they use an autoregressive fractionally integrated moving average (ARFIMA) model for daily RV (see Beran [1994] for long-memory and ARFIMA models). A more widely used model is the heterogeneous autoregressive (HAR) model proposed by Corsi (2009), written as:

$$\log \mathrm{RV}_{t+1} = \alpha_0 + \beta_d \log \mathrm{RV}_t + \beta_w \log \mathrm{RV}_{t-4:t} + \beta_m \log \mathrm{RV}_{t-21:t} + v_t,\ \ v_t \sim N(0, \sigma_v^2), \tag{8}$$

where RV_t, $\mathrm{RV}_{t-4:t}$, and $\mathrm{RV}_{t-21:t}$ are daily, weekly and monthly RVs. In this article, we express

[3] See Campbell et al. (1997) for details of market microstructure noise and Hansen and Lunde (2006) for the bias in RV caused by market microstructure noise.

[4] Takahashi et al. (2009, 2016) extend stochastic volatility models so as to take account of the bias in RV by modeling the daily returns and RV jointly. Similarly, Hansen et al. (2012) and Hansen and Huang (2016) extend GARCH and EGARCH models to realized GARCH and EGARCH models.

all variables in monthly units, so that we define RV_t, $\mathrm{RV}_{t-4:t}$ and $\mathrm{RV}_{t-21:t}$ as:

$$\mathrm{RV}_t = 22\sum_{i=1}^{n} r_{t-1+i/n}^2, \quad \mathrm{RV}_{t-4:t} = \frac{22}{5}\sum_{i=1}^{5} \mathrm{RV}_{t+1-i}, \quad \mathrm{RV}_{t-21:t} = \sum_{i=1}^{22} \mathrm{RV}_{t+1-i}. \tag{9}$$

The HAR model (8) is not a long-memory model. However, it is known to accurately approximate a long-memory process. The advantage of using this model is that it can be estimated by OLS and we have only to replace the left-hand-side of equation (8) by $\log \mathrm{RV}_{t+1:t+22}$ when we wish to forecast monthly QV.

If the asset price follows equation (1), QV is divided into two components as:

$$\mathrm{QV}_{t+1:t+\tau} = \int_t^{t+\tau} \sigma^2(s)ds + \sum_{t<s\le t+\tau} \kappa^2(s)ds. \tag{10}$$

The first and second terms on the right hand side are called the continuous and jump components, respectively. The continuous component is usually more persistent than the jump component. Andersen et al. (2007) extend the HAR model to the HAR-CJ model by separating RV into continuous and jump components and introducing them separately as explanatory variables. Let C_t and J_t denote the continuous and jump components on day t in monthly units and define weekly and monthly continuous components as:

$$C_{t-4:t} = \frac{22}{5}\sum_{i=1}^{5} C_{t-i+1}, \quad C_{t-21:t} = \sum_{i=1}^{22} C_{t-i+1}, \tag{11}$$

and weekly and monthly jump components as:

$$J_{t-4:t} = \frac{22}{5}\sum_{i=1}^{5} J_{t-i+1}, \quad J_{t-21:t} = \sum_{i=1}^{22} J_{t-i+1}. \tag{12}$$

Then, the HAR-CJ model has the following form:

$$\begin{aligned} &\log \mathrm{RV}_{t+1:t+22} \\ &= \alpha_0 + \beta_d \log C_t + \beta_w \log C_{t-4:t} + \beta_m \log C_{t-21:t} \\ &\quad + \gamma_d \log(1+J_t) + \gamma_w \log(1+J_{t-4:t}) + \gamma_m \log(1+J_{t-21:t}) + v_t,\ v_t \sim N(0,\sigma_v^2). \end{aligned} \tag{13}$$

To separate RV into continuous and jump components, Andersen et al. (2007) use the bipower variation (BV) method proposed by Barndorff-Nielsen and Shephard (2006), which we also use in this article. Appendix A details this method.

The negative correlation between today's return and tomorrow's volatility is a common phenomenon in stock markets. The HAR and HAR-CJ models above do not take into account this asymmetry in volatility. Although VIX and VXJ are under Q, they may include additional information on future QV under P because they are the expectation of QV. To consider this additional information and that asymmetry in volatility, Bekaert and Hoerova (2014) insert $\log(\mathrm{VIX}_t^2/12)$ and R_t^- into the HAR-CJ model (13). They gain the following:

$$\begin{aligned}
&\log \mathrm{RV}_{t+1:t+22} \\
&= \alpha_0 + \alpha_1 \log(\mathrm{VIX}_t^2/12) + \beta_d \log C_t + \beta_w \log C_{t-4:t} + \beta^m \log C_{t-21:t} \\
&\quad + \gamma_d \log(1+J_t) + \gamma_w \log(1+J_{t-4:t}) + \gamma_m \log(1+J_{t-21:t}) \\
&\quad + \delta R_t^- + v_t,\ v_t \sim N(0,\sigma_v^2),
\end{aligned} \tag{14}$$

where R_t is daily return and $R_t^- = \min[R_t, 0]$, which captures the asymmetry in volatility. If $\delta < 0$, this model is consistent with the above-mentioned negative correlation between today's return and tomorrow's volatility. We use this model, which we call the asymmetric HAR-CJ (AHAR-CJ) model with the squared VIX (VXJ).[5]

3 Data

The New York stock exchange is open between 9:00 and 15:00 and does not have a lunch break, whereas the Tokyo Stock Exchange is open between 9:00 and 11:30 and between 12:30 and 15:00, with a one-hour lunch break. We obtain the prices of S&P500 at 9:00, 9:05, ..., 15:00 from the Cash Indices of Tick Data, LLC, and those of Nikkei 225 at 9:00, 9:05, ..., 11:30 and 12:30, 12:35, ..., 15:00 from the Stock Indices of Nikkei NEEDS Tick Data. We calculate 5-minute returns of S&P500 and Nikkei 225 as the log difference of the consecutive prices. We omit the overnight returns in both the indices and the lunch time returns between 11:30 and 12:30 in Nikkei 225 because their time intervals are long.

As mentioned above, we first calculate RVs of Nikkei 225 and S&P500 as the sum of the squares of 5-minute returns without overnight and lunch-time returns. Since a whole day's (24

[5]Bekaert and Hoerova (2014) also add

$$R_{t-4:t}^- = \min\left[R_t, (22/5)\sum_{t=1}^{5} R_{t+1-i}\right], \quad R_{t-21:t}^- = \min\left[R_t, \sum_{t=1}^{22} R_{t+1-i}\right]$$

as explanatory variables. However, we remove them because they are not statically significant, either in Japan or the US.

hours) volatility is underestimated when the RV is calculated with overnight and lunch-time returns excluded, we use the method proposed by Hansen and Lunde (2005) to convert the RVs to the estimate of a whole day's volatility:

$$\mathrm{RV}_t = c^* \mathrm{RV}_t^{(o)}, \quad c^* = \frac{\sum_{t=1}^{T}(R_t - \bar{R})^2}{\sum_{t=1}^{T} \mathrm{RV}_t^{(o)}}, \tag{15}$$

where R_t is the daily return and $\bar{R}$ is the sample mean of daily returns $\{R_1, \ldots, R_T\}$. This case verifies the following:

$$\frac{1}{T}\sum_{t=1}^{T} \mathrm{RV}_t = \frac{1}{T}\sum_{t=1}^{T}(R_t - \bar{R})^2. \tag{16}$$

This adjustment makes the sample mean of RV equal to the sample variance of daily returns.

As in equation (14), we separate a whole day's RV into continuous and jump components using the bipower variation method proposed by Barndorff-Nielsen and Shephard (2006). We set J_t equal to $\mathrm{RV}_t - \mathrm{BV}_t$ if it is statistically significant at the 1% significance level, and equal to zero otherwise.[6] The daily continuous component C_t is set equal to $\mathrm{RV}_t - J_t$.

Table 1 shows the OLS estimation results of the AHAR-CJ model with the squared VIX (VXJ) using the full samples. Following Andersen et al. (2007), we calculate the Newey-West (1987) $t-$statistics with a Bartlett kernel and the lag length of 44 taking account of the overlapping observations problem. The adjusted R^2 figures for Nikkei 225 and S&P500 are 0.598 and 0.715, respectively, indicating the high explanatory power of this model. All parameters except γ_m in Nikkei 225 and γ_d and γ_w in S&P500 are statistically significant at the 5% significance level. Bekaert and Hoerova (2014) also find that γ_d and γ_w in S&P500 are not statistically significant. We also conduct all analyses below by assuming $\gamma_m = 0$ in Nikkei 225 and $\gamma_d = \gamma_w = 0$ in S&P500, but the result does not change qualitatively. δ is significantly negative both for Nikkei 225 and S&P500, which is consistent with the common phenomenon in stock markets of a negative correlation between today's return and tomorrow's volatility.

[Table 1 about here.]

We estimate the AHAR-CJ model with the squared VIX or VXJ at the end of the last trading day of each month using 500 samples up to that day. Using the estimates of parameters

[6] We also calculate C_t and J_t using the 0.1% and 5% significance levels, but the results are unaltered. We report the 1%-significant-level cases in the text.

and error variance ($\hat{\alpha}_0$, $\hat{\alpha}_1$, $\hat{\beta}_d$, $\hat{\beta}_w$, $\hat{\beta}_m$, $\hat{\gamma}_d$, $\hat{\gamma}_w$, $\hat{\gamma}_m$, $\hat{\delta}$, $\hat{\sigma}_v^2$), we calculate $E_t^P[\text{RV}_{t+1:t+22}]$ as:

$$\begin{aligned} & E_t^P[\text{RV}_{t+1:t+22}] \\ &= \exp\left[E_t^P[\log \text{RV}_{t+1:t+22}] + \frac{1}{2}Var_t^P[\log \text{RV}_{t+1:t+22}]\right] \\ &= \exp\Big[\hat{\alpha}_0 + \hat{\alpha}_1 \log(\text{VIX}_t^2/12) + \hat{\beta}_d \log C_t + \hat{\beta}_w \log C_{t-4:t} + \hat{\beta}_m \log C_{t-21:t} \\ &\quad + \hat{\gamma}_d \log(1+J_t) + \hat{\gamma}_w \log(1+J_{t-4:t}) + \hat{\gamma}_m \log(1+J_{t-21:t}) + \hat{\delta} R_t^- + \frac{1}{2}\hat{\sigma}_v^2\Big]. \end{aligned} \tag{17}$$

We can then calculate the VRP as the difference between either the squared VIX or VXJ, and $E_t^P[\text{RV}_{t+1:t+22}]$. Daily data on VIX and VXJ are downloaded from the CBOE and Osaka University websites. We obtain the VRP of Nikkei 225 at the end of each month between 2000.2 and 2018.9, and of S&P500 between 1992.1 and 2017.7. Figure 2 plots the squared VIX or VXJ, CV and the VRP of Nikkei 225 and S&P500. If traders are risk averse, the VRP would have to be positive. However, we sometimes obtain negative values for the VRP.

[Figure 2 about here.]

4 The Excess Return Predictability and the VRP

We regress the h-month-ahead excess stock return, $\text{ER}_{t:t+h} = \frac{1}{h}\sum_{i=1}^{h}\text{ER}_{t+i}$, on VRP_t to examine the excess return predictability afforded by the VRP as:

$$\text{ER}_{t:t+h} = \alpha + \beta \text{VRP}_t + \epsilon_t, \tag{18}$$

where the excess returns of Nikkei 225 and S&P500 are calculated as the differences between the returns of Nikkei 225 and S&P500, and the risk-free rates. Following the previous literature, such as Bollerslev et al. (2009, 2014) and Bekaert and Hoerova (2014), we use the 3-month T-bill rate as the risk-free rate for the US. We use the 1-month call rate for Japan because the 3-month call market is so illiquid that there are some missing data due to no trading. The 3-month T-bill and 1-month call money rates are downloaded from the websites of the Federal Reserve Bank of St. Louis and the Bank of Japan (BOJ), respectively.

This regression is subject to the overlapping observations problem. Bollerslev et al.'s (2014) simulation result shows that the Newey-West (1987) $t-$statistic with a Bartlett kernel and a lag length determined by $h + 4((T-h)/100)^{2/9}$ is marginally more powerful than the Hodrick

(1992) type $t-$statistic.[7] Table 2 shows the estimates of α and β with the Newey-West-based $t-$statistics and adjusted R^2 for Nikkei 225. While the previous literature, such as Bollerslev et al. (2014), Ubukata and Watanabe (2014a), Uchiyama and Yamanaka (2015) and Andersen et al. (2019), does not find any significant correlations between the VRP and the future excess returns in Japan for any h, Table 2(a) shows a significantly negative correlation for $h = 3, 4$ at the 5% significance level (one-sided test). This does not mean that the VRP contributes to predicting the excess returns for Nikkei 225, because the adjusted R^2 is very small. Notice that the adjusted R^2 is denoted in percentage in Tables 2 and 3, while it is not in Table 1. The negative correlation between the VRP and the excess returns contradicts the theoretical results in Bollerslev et al. (2009). Table2(b) shows significantly positive correlation between the VRP and the excess returns for $h = 1, 3, 4$ at the 5% significance level (one-sided test), which is consistent with the theoretical results in Bollerslev et al. (2009) and the empirical results in Bollerslev et al. (2009, 2014) and Bekaert and Hoerova (2014). The adjusted R^2 is still small but larger than that for Nikkei 225.

[Table 2 about here.]

[Figure 3 about here.]

Figure 3 plots the 1-month call money rate in Japan and 3-month T-bill rate in the US, which we use as risk-free rates. The 1-month call rate in Figure 3(a) is always close to zero, except for some short periods. The 3-month T-bill rate in the US is close to zero only between 2008.11 and 2016.10. No one has examined the excess return predictability of the VRP during this period in the US. If the lack of excess return predictability of the VRP in Japan is due to the ZLB for the risk-free rate, we would expect the VRP in the US to not predict the excess returns during this period. Table 3 shows the estimation results using the subsample between 2008.11 and 2016.10. As is expected, β is not statistically significant for any h at any standard significance level.

[Table 3 about here.]

Figure 2 shows that VIX and VRP in Japan increase around the end of 2008 after the Lehman shock in 2008.9, while CV increases more than VIX in the US, so that VRP has

[7] The lag-length selection for the Newey-West (1987) $t-$statistics is still under debate (see Lazarus et al. [2018]). We set the lag-length for the Newey-West $t-$statistics in Table 1 following Andersen et al. (2007), and that in Table 2 and 3 following Bollerslev et al. (2014).

negative values. To examine the effect of the dramatic increase in VRP around the end of 2008 in Japan from our results, we excluded the data between 2008.10 and 2008.12 from all analyses for Japan in this section. The results are, however, qualitatively unaltered.

5 Model

In this section, we conduct a theoretical analysis to explain the results obtained in the previous section by using a simple asset pricing model. Bollerslev et al. (2009) provide an asset pricing model with time-varying volatility-of-volatility to produce a positive correlation between the VRP and excess stock returns. In the previous sections, we obtain the following results.

1. In the US, the correlation between the VRP and excess stock returns is significantly positive when we use the full sample (1992.1-2017.7), while it is not statistically significant when we restrict the sample to the ZLB period (2008.11-2016.10).

2. In Japan, the correlation between the VRP and excess stock returns is significantly negative.

These results can not be explained by Bollerslev et al.'s (2009) model, which always produces a positive correlation between the VRP and excess stock returns. To explain the above results, we extend the model of Bollerslev et al. (2009) by incorporating two new ingredients: (a) the ZLB for the risk-free rate and (b) a negative correlation between the consumption growth rate and the volatility-of-volatility.

Pricing a higher-order risk premium, or VRP, requires the use of a recursive utility function (e.g., Epstein and Zin [1989], Bansal and Yaron [2004]). To obtain an analytical solution to a model using this utility function, we make some unconventional assumptions. In reality, the ZLB is imposed on nominal interest rates, but we impose it on real interest rates by assuming zero inflation that equalizes nominal with real interest rates. We exclude the monetary authority and assume that real interest rates are equal to zero if the values determined by the Euler equation are less than zero. Thanks to these assumptions, our model can explain why the correlation between the VRP and future excess returns can be zero or negative. We close this section by discussing the theoretical potentials and technical difficulties of incorporating our model to New-Keynesian frameworks, general equilibrium models with nominal rigidities, and ways to render those unconventional assumptions unnecessary.

5.1 Environment

Suppose that geometric growth rate of consumption in the economy, $g_{t+1} = \log\left(\frac{C_{t+1}}{C_t}\right)$, follows a stochastic process such as:

$$g_{t+1} = \mu_g(q_t) + \sigma_{g,t} z_{g,t+1}, \tag{19}$$

$$\sigma^2_{g,t+1} = a_\sigma + \rho_\sigma \sigma^2_{g,t} + \sqrt{q_t} z_{\sigma,t+1}, \tag{20}$$

$$q_{t+1} = a_q + \rho_q q_t + \varphi_q \sqrt{q_t} z_{q,t+1}, \tag{21}$$

where $\sigma_{g,t}$ is the conditional variance of growth rate, q_t is a parameter associated with the volatility-of-volatility, $\{z_{g,t}, z_{\sigma,t}, z_{q,t}\}$ are jointly independent i.i.d. N(0,1) innovation processes, and μ_g is the mean growth rate. Parameters satisfy $a_\sigma > 0$, $a_q > 0$, $|\rho_\sigma| < 1$, $|\rho_q| < 1$, $\varphi_q > 0$. Notice that there are two state variables in the economy, namely, $\sigma_{g,t}$ and q_t.

We assume a negative correlation between the consumption growth rate and the volatility-of-volatility as:

$$\mu_g(q_t) = -\alpha q_t + \beta, \tag{22}$$

where α and β are associated parameters. This assumption is based on what Oya (2011) and Ubukata and Watanabe (2014a) find for Japan; that is, changes in the VRP are counter-cyclical. These changes are directly proportional to changes in the volatility-of-volatility in our model, as will be demonstrated later. That assumption diverges from Bollerslev et al. (2009) in which $\mu_g(q_t)$ is constant. The assumption contributes to creating a negative correlation between the VRP and excess stock returns.

We assume that the representative agent in the economy has the Epstein-Zin-Weil utility function (Epstein and Zin [1989]) such that:

$$V_t = \left[(1-\delta) C_t^{\frac{1-\gamma}{\theta}} + \delta \left(E_t[V_{t+1}^{1-\gamma}]\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}}, \tag{23}$$

where δ is the subjective discount factor, γ is the coefficient of risk aversion, θ is defined as $\theta \equiv \frac{1-\gamma}{1-\frac{1}{\psi}}$ in which ψ is the intertemporal elasticity of substitution.

Then, the logarithm of the stochastic discount factor, $m_{t+1} \equiv \log(M_{t+1})$, can be expressed as:

$$m_{t+1} = \theta \ln \delta - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{t+1}, \tag{24}$$

where $r_{t+1} \equiv \log(R_{t+1})$ is the return at $t+1$ on the consumption claim (equity).

5.2 Equilibrium Solution

5.2.1 Stock Returns

By using the stochastic discount factor m_{t+1} derived in the previous subsection, the equilibrium condition for equity, $E_t[M_{t+1} R_{t+1}] = 1$, leads to

$$E_t \left[\exp \left(\theta \ln \delta - \frac{\theta}{\psi} g_{t+1} + \theta r_{t+1} \right) \right] = 1. \tag{25}$$

We conjecture that the log (equity) price-consumption ratio, w_t, follows:

$$w_t = A_0 + A_\sigma \sigma_{g,t}^2 + A_q q_t. \tag{26}$$

Furthermore, following Campbell and Shiller (1988), the stock return (r_{t+1}) is approximated by using this price-consumption ratio as:

$$r_{t+1} = \kappa_0 + \kappa_1 w_{t+1} - w_t + g_{t+1}. \tag{27}$$

Parameters (A_0, A_σ, A_q) are pinned down by solving the above conditions.

r_{t+1} can be expressed explicitly by using state variables $(\sigma_{g,t}, q_t)$ and innovations as:

$$\begin{aligned} r_{t+1} = & -\ln \delta - \alpha q_t + \frac{\beta}{\psi} - \frac{(1-\gamma)^2}{2\theta} \sigma_{g,t}^2 + (\kappa_1 \rho_q - 1) A_q q_t + \sigma_{g,t} z_{g,t+1} \\ & + \kappa_1 \sqrt{q_t} \left(A_\sigma z_{\sigma,t+1} + A_q \varphi_q z_{q,t+1} \right). \end{aligned} \tag{28}$$

See Appendix B for details of the derivation.

5.2.2 Risk-Free Rate

We derive the risk-free rate (r^f) in a similar way to that used by Bansal and Yaron (2004). Let $r_t^f \equiv \log(R_t^f)$, where R_t^f is the return on the risk-free asset at $t+1$. The equilibrium condition for this asset, $E_t[M_{t+1}R_t^f] = 1$, leads to

$$E_t\left[\exp\left(\theta\ln\delta - \frac{\theta}{\psi}g_{t+1} + (\theta-1)r_{t+1} + r_t^f\right)\right] = 1. \tag{29}$$

Solving this equation for r_t^f yields:

$$\begin{aligned} r_t^f =& -\ln\delta - \alpha q_t\left(\frac{\theta}{\psi} + 1 - \theta\right) + \frac{\beta}{\psi} \\ &+ \sigma_{g,t}^2\left[\frac{(1-\gamma)^2}{2\theta}(\theta-1) - \frac{1}{2}\left(\frac{\theta}{\psi}\right)^2 - \frac{1}{2}(\theta-1)^2 - \frac{\theta(1-\theta)}{\psi}\right] \\ &+ q_t\left[(1-\theta)(\kappa_1\rho_q - 1)A_q - \frac{1}{2}(\theta-1)^2\kappa_1^2\left(A_\sigma^2 + (A_q\varphi_q)^2\right)\right]. \end{aligned} \tag{30}$$

See Appendix C for details of the derivation.

5.2.3 Excess Stock Returns without the ZLB

We compute excess stock returns by using equations (28) and (30) as:

$$E_t\left[r_{t+1} - r_t^f\right] = \underbrace{\sigma_{g,t}^2\left(\gamma - \frac{1}{2}\right)}_{\text{Risk premium from volatility}} + \underbrace{q_t\left(\frac{1}{2} - \theta\right)\kappa_1^2\left(A_\sigma^2 + (A_q\varphi_q)^2\right)}_{\text{Risk premium from volatility-of-volatility}}. \tag{31}$$

The excess returns consist of two kinds of risk premia. The first comes from volatility, and the second comes from volatility-of-volatility.

See Appendix D for details of the derivation.

5.2.4 Zero Lower Bound for the Risk-Free Rate

We assume that the risk-free rate r_t^f described by equation (30) is not less than zero, as shown below:

$$\hat{r}_t^f = \max\{r_t^f, 0\}. \tag{32}$$

Now, the excess stock return with the ZLB can be described as $E_t\left[r_{t+1} - \hat{r}_t^f\right]$.

We interpret here a situation where the ZLB works for the risk-free asset as meaning that the representative consumer has infinite demand for risk-free assets when the risk-free rate is zero. In this case, the market for risk-free assets does not clear, since the risk-free rate cannot decrease any further: excess demand for risk-free assets arises. In our model, risk-free assets correspond to securities and bank deposits. When the risk-free rate is zero, the representative consumer wants to invest more in those risk-free assets, however, supply cannot keep pace with demand.

5.2.5 VRP (Variance Risk Premium)

In line with Bollerslev et al. (2009), we describe the VRP as:

$$\begin{aligned} \text{VRP}_t &\equiv E_t^Q\left(\sigma_{t+1}^2\right) - E_t^P\left(\sigma_{t+1}^2\right) \\ &= (\theta-1)\kappa_1\left[A_\sigma + A_q\kappa_1^2\left(A_\sigma^2 + A_q^2\varphi_q^2\right)\varphi_q^2\right]q_t, \end{aligned} \tag{33}$$

where the VRP is proportional to the volatility-of-volatility.

5.3 Calibration

To calibrate our model, we set the parameter values with reference to the ones used in Bollerslev et al. (2009). Our κ_1 is 0.8 and different from that used by Bollerslev et al. (2009). This is because the solution of A_q in the quadratic equation (B4) becomes a complex number that prevents an equilibrium from emerging when we use the value 0.9 used by Bollerslev et al. (2009). Two parameters (α, β) are specific to our model. These are set so that the average consumption growth rate becomes 0.15%. Table 4 shows the value of all parameters.

[Table 4 about here.]

5.4 Model Results

Figure 4 shows the solutions of the calibrated model. We fix one of the state variables $\sigma_{g,t}$ at the mean value and change the other state variable q_t, then observe any change in excess stock returns and VRP.

Figure 4(a) shows how the stock return (r_t), the risk-free rate (r_t^f), and the excess stock return ($r_t - r_t^f$) change over the volatility-of-volatility (q_t). When q_t increases, both the stock return and the risk-free rate decrease initially because we assume that the consumption growth rate (μ_g) has a negative correlation with the volatility-of-volatility (q_t). One part of the excess stock return, or a risk premium coming from q_t, increases because investors require more premium for higher-order uncertainty as shown in equation (31). However, after the risk-free rate hits the ZLB, the excess stock return begins to decline. As a result, the excess stock return follows a hump-shaped pattern over q_t, as shown in Figure 4(a). Figure 4(b) shows the counterparts in Bollerslev et al. (2009). Their excess stock return monotonically increases as q_t increases.

Figure 4(c) shows how the VRP varies over q_t. As shown in equation (33), it monotonically increases when q_t increases. On the left hand side of the kink (before the risk-free rate hits the ZLB), the correlation between the VRP and excess stock returns is significantly positive. This may correspond to the US case. On the right hand side of this kink (after the risk-free rate hits the ZLB), the correlation is significantly negative. This may correspond to the case of Japan. Finally, around this kink, the correlation becomes insignificant. This corresponds to the US case in the ZLB period.

[Figure 4 about here.]

5.5 Discussion

To explain the two important observations on the VRP mentioned above, we extend the model of Bollerslev et al. (2009), while keeping their underlying model specific, a real endowment economy with exogenous production and consumption. To obtain an analytical solution to our extended model, we need to make the following three assumptions.

First, the ZLB is imposed not on the nominal but on the real interest rate. In our model, the real interest rate is equal to the nominal interest rate simply because there is no distinction between the nominal and real terms. This equalization can be verified in an extreme situation where the final goods price is perfectly rigid.

Second, we abstract from the monetary authority. In line with Bollerslev et al. (2009), our model does not provide the monetary authority with any role, because the model describes an endowment economy without any nominal frictions. The absence of such frictions removes the need for a policy-making body that affects the real allocation and social welfare.

Last, equation (32) means that real interest rates are equal to zero if the value of r_t^f determined by the Euler equation (30) is less than zero. This assumption causes excess demand for risk-free assets when real interest rates are zero, because the Euler equation for the assets is not satisfied. When the risk-free rate hits zero, it will not decrease any further, thereby failing to clear the market for the assets. Because production is exogenous in our endowment economy, there is no way for the economy to change production to restore equilibrium in the risk-free asset market. Securities and bank deposits are real-world examples of our model's risk-free assets. When the risk-free rate is zero, the representative consumer wants to invest in the risk-free assets as much as possible to store their value. Not all of this demand, however, can be accommodated, because the supply of these assets is limited.

One may regard these assumptions as unconventional with reference to the New-Keynesian framework. Within this framework, a conventional monetary economy would be one in which production is determined endogenously, where the monetary authority works in response to inflation based on a specific rule such as the Taylor rule, and where the nominal interest rate is equal to zero if the value determined by that rule is less than zero. New-Keynesian models with the ZLB have the potential to relax the three assumptions made in our model. Some New-Keynesian models do not need our first assumption because they can make a distinction between nominal and real terms by embedding money in the economy. This distinction helps impose the ZLB on nominal interest rate. Some New-Keynesian models work without our model's second assumption because they generate nominal frictions, calling for a monetary authority that can mitigate the inefficiencies caused by those frictions. Regarding the last unconventional assumption, when the risk-free rate hits zero, the New-Keynesian framework allows the new equilibrium to arise endogenously in the risk-free asset market in response to the reduction in demand for risk-free assets due to the decline in production, income, and then consumption (see, e.g., Eggertsson and Woodford [2003], Caballero and Farhi [2017], Basu and Bundick [2017]). In our model, this mechanism will not work, simply because production is exogenous, as in the model of Bollerslev et al. (2009).

Inspired by the potential of the New-Keynesian framework, Basu and Bundick (2017) and Strobel et al. (2019) employ the framework to analyze the second-order risk premium, and propose numerical methods for solving the equilibrium with the ZLB imposed on the nominal rate. Unfortunately, their methods are not applicable in this study because the VRP is a fourth-order risk premium, which requires the use of a recursive utility function for pricing.

6 Conclusion

We investigate the relevance of VRPs to predicting excess stock returns for Japan and the US. The main findings are as follows. First, the correlation between the VRP and the future excess returns in the US is statistically significant in the full sample (1990.12-2017.7), which is consistent with the previous literature, while it is not statistically significant in the subsample (2008.11–2016.10) during the period of zero interest policy. Second, the correlation between the VRP and the future excess returns in Japan is negative and, unlike the previous literature, statistically significant. To provide an account of these findings, we conduct a preliminary theoretical analysis with a structural model of asset pricing based on two assumptions: the ZLB for the risk-free rate and a negative correlation between the consumption growth rate and the volatility-of-volatility.

Our study merits several extensions. First, as we examine only Japan and the US, our study invites application to other countries too. Second, it may be interesting to examine the effect of the BOJ's ETF purchases on the VRP and excess stock returns, and on the relation between these two. Third, although we do not cover production and financial intermediaries, these would both merit consideration in constructing an equilibrium model for the VRP. Oya (2011) and Ubukata and Watanabe (2014a) show that the VRP contributes to predicting business cycles for Japan. Ubukata and Watanabe (2014a) also find that the VRP has predictive power for credit spreads in Japan. In the US, by contrast, rather than the VRP, it is conditional variance (CV) that predicts business cycles and financial instability (Bekaert and Hoerova [2014]). Fourth, to explain fully the mechanics of the correlation between the VRP and excess stock returns, our unconventional assumptions made to straightforwardly extend the model of Bollerslev et al. (2009) should be relaxed. The New-Keynesian framework has the potential to help this relaxation. However, to make it feasible for fourth-order risk premia, further technical efforts are needed to secure the approximation accuracy and computability for such risk premia. Last, there are some other extensions of the HAR model, such as those proposed by Bollerslev et al. (2016) and Chen et al. (2018). Stroud and Johannes (2014) show that the QV's ability to predict excess stock returns increases by modelling the dynamics of intraday volatility. It would be interesting to calculate the VRP using such new methods and compare it with that calculated based on the conventional HAR model.

References

Andersen, T. G., Bollerslev, T. and Diebold, F. X. (2007), "Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility," *Review of Economics and Statistics*, **89**, 701–720.

Andersen, T. G., Bollerslev T., Diebold F. X. and Ebens, H. (2001a), "The distribution of stock returns volatilities," *Journal of Financial Economics*, **61**, 43-76.

Andersen, T. G., Bollerslev T., Diebold F. X. and Labys, P. (2001b), "The distribution of realized exchange rate volatility," *Journal of the American Statistical Association*, **96**, 42-55.

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

Andersen, T. G., Todorov, V. and Ubukata, M. (2019), "Tail risk and return predictability for the Japanese equity market," *Journal of Econometrics*, Forthcoming.

Bansal, R., and Yaron, A. (2004), "Risks for the long run: A potential resolution of asset pricing puzzles," *The Journal of Finance*, **59**, 1481-1509.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2008) "Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise," *Econometrica*, **76**, 1481–1536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2009), "Realized kernels in practice: Trades and quotes," *Econometrics Journal*, **12**, C1–C32.

Barndorff-Nielsen, O. E. and Shephard, N. (2004), "Power and bipower variation with stochastic volatility and jumps (with discussion)," *Journal of Financial Econometrics*, **2**, 1–37.

Barndorff-Nielsen O. E. and Shephard, N. (2006), "Econometrics of testing for jumps in financial economics using bipower variation," *Journal of Financial Econometrics*, **4**, 1–30.

Basu, S., and Bundick, B. (2017), "Uncertainty Shocks in a Model of Effective Demand," *Econometrica*, **85**, 937-958.

Bekaert, G. and Hoerova, M. (2014), "The VIX, the variance premium and stock market volatility," *Journal of Econometrics*, **183**, 181–192.

Beran, J. (1994), *Statistics for long-memory processes*, Chapman & Hall.

Bollerslev, T., Marrone, J., Xu, L. and Zhou, H. (2014), "Stock return predictability and variance risk premia: Statistical inference and international evidence," *Journal of Financial and Quantitative Analysis*, **49**, 633–611.

Bollerslev, T., Patton, J. A. and Quaedvlieg, R. (2016), "Exploiting the errors: A simple approach for improved volatility forecasting," *Journal of Econometrics*, **192**, 1–18.

Bollerslev, T., Tauchen, G. and Zhou, H. (2009), "Expected stock returns and variance risk premia," *The Review of Financial Studies*, **22**, 4463–4492.

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

Caballero, R. J. and Farhi, E. (2017), "The safety trap," *The Review of Economic Studies*, **85**, 223-274.

Campbell, J. Y., Lo, A. W. and MacKinlay, A. C. (1997), *The Econometrics of Financial Markets*, Princeton: Princeton University Press.

Campbell, J. Y., and Shiller, R. J. (1988), "The dividend-price ratio and expectations of future dividends and discount factors," *The Review of Financial Studies*, **1**, 195–228.

Chen, X. B., Gao, J., Li, D. and Silvaplle, P. (2018), "Nonparametric estimation and forecasting for time-varying coefficient realized volatility models," *Journal of Business & Economic Statistics*, **36**, 88–100.

Corsi, F. (2009), "A simple approximate long-memory model of realized volatility," *Journal of Financial Econometrics*, **7**, 174–196.

Eggertsson, Gauti B., and Woodford, M., (2003), "The Zero Bound on Interest Rates and Optimal Monetary Policy," *Brookings Papers on Economic Activity*, 139–211.

Epstein, Larry G and Zin, Stanley E. (1989), "Risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework," *Econometrica*, **57**, 937-969.

Fukasawa, M, Ishida, I., Maghrebi, N., Oya, K., Ubukata, M. and Yamazaki, K. (2011), "Model-free implied volatility: From surface to index," *International Journal of Theoretical and Applied Finance*, **14**, 433–463.

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

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

Hansen, P. R. and Lunde, A. (2005), “A forecast comparison of volatility models: Does anything beat a GARCH(1,1)?” *Journal of Applied Econometrics*, **20**, 873–889.

Hansen, P. R. and Lunde, A. (2006), “Realized variance and market microstructure noise,” *Journal of Business & Economic Statistics*, **24**, 1271-61.

Hodrick, R. J. (1992), “Dividend yields and expected stock returns: Alternative procedures for inference and measurement,” *The Review of Financial Studies*, **5**, 357-386.

Huang, X. and G. Tauchen (2005) “The relative contribution of jumps to total price variance,” *Journal of Financial Econometrics*, **3**, 456-499.

Jiang, G. J. and Tian, Y. S. (2005), “Model-free implied volatility and its information content,” *The Review of Financial Studies*, **18**, 1305–1342.

Jiang, G. J. and Tian, Y. S. (2007), “Extracting model-free volatility from option prices: An examination of the VIX index,” *Journal of Derivatives*, **14**, 35–60.

Lazarus E, Lewis D.J., Stock J. H. and Watson, M. W. (2018), “HAR inference: Recommendations for practice,” *Journal of Business & Economic Statistics*, **36**, 541–574.

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

Newey, W. K. and West K. D. (1987), “A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix,” *Econometrica*, **55**, 703-708.

Oya, K. (2011), “Business cycle predictability with variance risk premium,” *Great Recession in the Global Economy and Business Cycle Analyses (Sekai douji fukyo to keiki junkan bunseki)*, University of Tokyo Press, 141–157 (in Japanese).

Strobel, J., Lee, G., Dorofeenko, V. and Salyer, K. (2019), “Time-varying risk shocks and the zero lower bound,” Annual Conference 2019 (Leipzig): 30 Years after the Fall of the Berlin Wall - Democracy and Market Economy 203491, Verein für Socialpolitik / German Economic Association.

Stroud, J. R. and Johannes, M. S. (2014), “Bayesian modeling and forecasting of 24-hour high-frequency volatility,” *Journal of the American Statistical Association*, **109**, 1368–1384.

Takahashi, M., Omori, Y. and Watanabe, T. (2009), "Estimating stochastic volatility models using daily returns and realized volatility simultaneously," *Computational Statistics and Data Analysis*, **53**, 2404–2426.

Takahashi, M., Watanabe, T. and Omori, Y. (2016), "Volatility and quantile forecasts by realized stochastic volatility models with generalized hyperbolic distribution," *International Journal of Forecasting*, **32**, 437–457.

Ubukata, M. and Watanabe, T. (2014a), "Market variance risk premiums in Japan for asset predictability," *Empirical Economics*, **47**, 169–198.

Ubukata, M. and Watanabe, T. (2014b), "Option pricing using realized volatility," *Japanese Economic Review*, **65**, 431–467

Uchiyama, T. and Yamanaka, S. (2015), "Volatility risk premia: Size and predictability of factor returns," *Securities Analysts Journal*, **53**, 69–79 (in Japanese).

Watanabe, T. (2020), "Heterogeneous autoregressive models: Survey with the application to the realized volatility of Nikkei 225 stock index," *HUE Journal of Economics and Business*, **42**, 5–18 (in Japanese).

Zhang, L. (2006), "Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach," *Bernoulli*, **12**, 1019–1043.

Zhang, L., Mykland, P. A. and Aït-Sahalia, Y., (2005), "A tale of two time scales: Determining integrated volatility with noisy high-frequency data," *Journal of the American Statistical Association*, **100**, 1394–1411.

Appendix

A Method for Dividing RV into the Continuous and Jump Components

To divide RV_t into the continuous component C_t and the jump component J_t, we use the method proposed by Barndorff-Nielsen and Shephard (2004) and Barndorff-Nielsen and Shephard (2006). The following Bipower Variation (BV_t) converges to the continuous component C_t as $n \to \infty$.

$$\mathrm{BV}_t = \mu_1^{-2} \frac{n}{n-1} \sum_{i=2}^{n} |r_{t-1+i/n}||r_{t-1+(i-1)/n}|, \tag{A1}$$

where $\mu_1 = \sqrt{2/\pi}$.

We first test for the null hypothesis of no jump. Following Barndorff-Nielsen and Shephard (2006) and Huang and Tauchen (2005), we use the following test statistic.

$$Z_t^* = \frac{\sqrt{n}(\mathrm{RV}_t - \mathrm{BV}_t)\mathrm{RV}_t^{-1}}{\sqrt{(\mu_1^{-4} + 2\mu_1^{-2} - 5)\mathrm{Max}(1, \mathrm{TQ}_t \mathrm{BV}_t^{-2})}}, \tag{A2}$$

where TQ_t is the tri-power quarticity defined by

$$\mathrm{TQ}_t = n\mu_{4/3}^{-3} \frac{n}{n-4} \sum_{i=3}^{n} |r_{t-1+i/n}|^{4/3} |r_{t-1+(i-1)/n}|^{4/3} |r_{t-1+(i-2)/n}|^{4/3},$$
$$\mu_{4/3} = 2^{2/3}\Gamma(7/6)\Gamma(1/2)^{-1}.$$

The asymptotic distribution of Z_t^* is the standard normal under the null hypothesis of no jump. We conduct this test at the 1% significance level.

Then, we set $\mathrm{C}_t = \mathrm{RV}_t - \mathrm{J}_t$ if the null hypothesis is rejected and $\mathrm{C}_t = \mathrm{RV}_t$ otherwise.

B Derivation for Stock Returns

By substituting equations (19), (20), (21), and (26) into (27), we can express the stock return r_{t+1} by using the state variables $(\sigma_{g,t}, q_t)$ and innovations as:

$$
\begin{aligned}
r_{t+1} &= \kappa_0 + \kappa_1(A_0 + A_\sigma \sigma_{g,t+1}^2 + A_q q_{t+1}) - (A_0 + A_\sigma \sigma_{g,t}^2 + A_q q_t) + g_{t+1} \\
&= \beta + \kappa_0 + (\kappa_1 - 1)A_0 + \kappa_1 A_q a_q + \kappa_1 A_\sigma a_\sigma + (\kappa_1 \rho_\sigma - 1)A_\sigma \sigma_{g,t}^2 \\
&+ [(\kappa_1 \rho_q - 1)A_q - \alpha] q_t + \kappa_1 A_\sigma \sqrt{q_t} z_{\sigma,t+1} + \kappa_1 A_q \varphi_q \sqrt{q_t} z_{q,t+1} + \sigma_{g,t} z_{g,t+1}. \qquad \text{(B1)}
\end{aligned}
$$

Using the above equation, we can rewrite the term in equation (25) as:

$$
\begin{aligned}
\theta \ln \delta - \frac{\theta}{\psi} g_{t+1} + \theta r_{t+1} = \theta \Bigg[& \ln \delta + \kappa_0 + \beta(1 - \frac{1}{\psi}) + (\kappa_1 - 1)A_0 + \kappa_1(A_\sigma a_\sigma + A_q a_q) \\
& + (\kappa_1 \rho_\sigma - 1)A_\sigma \sigma_{g,t}^2 + \left\{ (\kappa_1 \rho_q - 1)A_q - \alpha(1 - \frac{1}{\psi}) \right\} q_t \Bigg] \\
& + \theta \kappa_1 A_\sigma \sqrt{q_t} z_{\sigma,t+1} + \theta \kappa_1 A_q \varphi_q \sqrt{q_t} z_{q,t+1} + \theta \sigma_{g,t} \left(1 - \frac{1}{\psi} \right) z_{g,t+1}.
\end{aligned}
$$

Using the formula, $E[\exp(X)] = \exp[E(X) + \frac{1}{2}Var(X)]$, LHS of equation (25) becomes

$$
\begin{aligned}
\exp \Bigg[\theta \Bigg(& \ln \delta + \kappa_0 + \beta(1 - \frac{1}{\psi}) + (\kappa_1 - 1)A_0 + \kappa_1(A_\sigma a_\sigma + A_q a_q) \\
& + (\kappa_1 \rho_\sigma - 1)A_\sigma \sigma_{g,t}^2 + \left\{ (\kappa_1 \rho_q - 1)A_q - \alpha(1 - \frac{1}{\psi}) \right\} q_t \Bigg) \\
& + \frac{1}{2} q_t (\theta \kappa_1)^2 \left(A_\sigma^2 + (A_q \varphi_q)^2 \right) + \frac{1}{2} (\theta \sigma_{g,t})^2 \left(1 - \frac{1}{\psi} \right)^2 \Bigg].
\end{aligned}
$$

Since equilibrium condition (25) holds for any state variables $(\sigma_{g,t}, q_t)$, we have following three conditions:

$$
\beta \left(1 - \frac{1}{\psi} \right) + \ln \delta + \kappa_0 + (\kappa_1 - 1)A_0 + \kappa_1(A_\sigma a_\sigma + A_q a_q) = 0, \qquad \text{(B2)}
$$

$$
(\kappa_1 \rho_\sigma - 1)A_\sigma + \frac{1}{2}\theta \left(1 - \frac{1}{\psi} \right)^2 = 0, \qquad \text{(B3)}
$$

$$
(\kappa_1 \rho_q - 1)A_q - \alpha \left(1 - \frac{1}{\psi} \right) + \frac{1}{2}\theta \kappa_1^2 \left(A_\sigma^2 + (A_q \varphi_q)^2 \right) = 0. \qquad \text{(B4)}
$$

Equations (B2), (B3), and (B4) jointly yield the solution for (A_0, A_σ, A_q). In particular, we have the explicit solution for A_σ:

$$A_\sigma = \frac{(1-\gamma)^2}{2\theta(1-\kappa_1\rho_\sigma)}. \tag{B5}$$

By substituting (B2) and (B5) into (B1), we have the solution for the stock return r_{t+1} as equation (28).

C Derivation for the Risk-Free Rate

The derivation for the risk-free rate r_t^f is as follows:

Using the formula, $E[\exp(X)] = \exp[E(X) + \frac{1}{2}Var(X)]$, equation (29) becomes

$$\exp E_t\left[\left(\theta \ln \delta - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{t+1} + r_t^f\right) + \frac{1}{2} Var_t\left(\frac{\theta}{\psi} g_{t+1} + (1-\theta) r_{t+1}\right)\right] = 1.$$

Solving the above equation with respect to r_t^f yields:

$$\begin{aligned}
r_t^f &= -\theta \ln \delta + \frac{\theta}{\psi} E_t\left[g_{t+1}\right] + (1-\theta) E_t\left[r_{t+1}\right] - \frac{1}{2} Var_t\left(\frac{\theta}{\psi} g_{t+1} + (1-\theta) r_{t+1}\right) \\
&= -\theta \ln \delta + \frac{\theta}{\psi}\left(-\alpha q_t + \beta\right) + (1-\theta)\left[-\ln \delta - \alpha q_t + \frac{\beta}{\psi} - \frac{(1-\gamma)^2}{2\theta} \sigma_{g,t}^2 + (\kappa_1 \rho_q - 1) A_q q_t\right] \\
&- \frac{1}{2}\left[\left(\frac{\theta \sigma_{g,t}}{\psi}\right)^2 + (\theta-1)^2\left(\sigma_{g,t}^2 + \kappa_1^2 q_t\left(A_\sigma^2 + (A_q \varphi_q)^2\right)\right) + \frac{2\theta(1-\theta)}{\psi} \sigma_{g,t}^2\right] \\
&= -\ln \delta - \alpha q_t\left(\frac{\theta}{\psi} + 1 - \theta\right) + \frac{\beta}{\psi} + (1-\theta)\left[-\frac{(1-\gamma)^2}{2\theta} \sigma_{g,t}^2 + (\kappa_1 \rho_q - 1) A_q q_t\right] \\
&- \frac{1}{2}\left[\left(\frac{\theta \sigma_{g,t}}{\psi}\right)^2 + (\theta-1)^2\left(\sigma_{g,t}^2 + \kappa_1^2 q_t\left(A_\sigma^2 + (A_q \varphi_q)^2\right)\right) + \frac{2\theta(1-\theta)}{\psi} \sigma_{g,t}^2\right] \\
&= -\ln \delta - \alpha q_t\left(\frac{\theta}{\psi} + 1 - \theta\right) + \frac{\beta}{\psi} + \sigma_{g,t}^2\left[\frac{(1-\gamma)^2}{2\theta}(\theta - 1) - \frac{1}{2}\left(\frac{\theta}{\psi}\right)^2 - \frac{1}{2}(\theta-1)^2 - \frac{\theta(1-\theta)}{\psi}\right] \\
&+ q_t\left[(1-\theta)(\kappa_1 \rho_q - 1) A_q - \frac{1}{2}(\theta-1)^2 \kappa_1^2\left(A_\sigma^2 + (A_q \varphi_q)^2\right)\right].
\end{aligned}$$

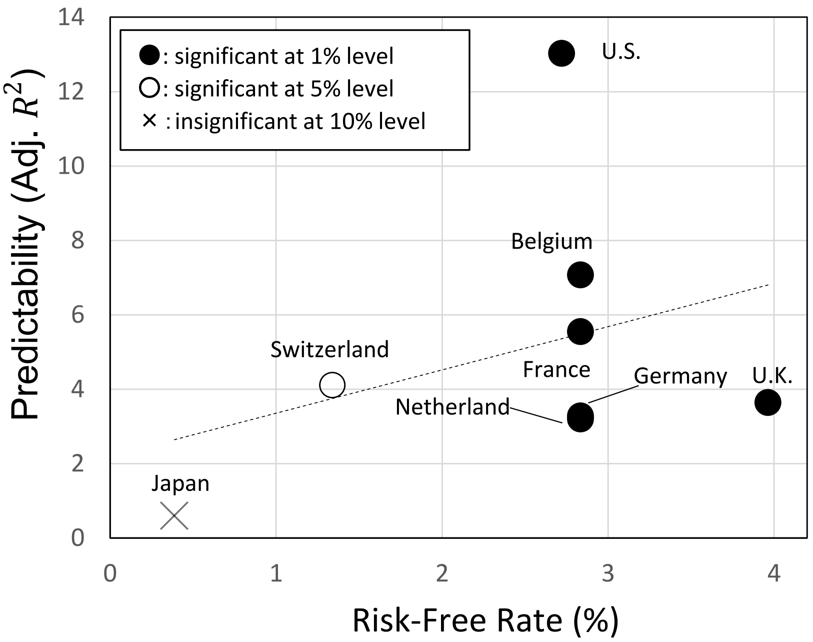
D Derivation for the Excess Stock Return without the ZLB

Excess stock return can be computed by using equations (28) and (30) as:

$$
\begin{aligned}
E_t\left[r_{t+1}-r_t^f\right] =& -\ln\delta-\alpha q_t+\frac{\beta}{\psi}-\frac{(1-\gamma)^2}{2\theta}\sigma_{g,t}^2+(\kappa_1\rho_q-1)A_q q_t \\
&+\ln\delta+\alpha q_t\left(\frac{\theta}{\psi}+1-\theta\right)-\frac{\beta}{\psi} \\
&-\sigma_{g,t}^2\left[\frac{(1-\gamma)^2}{2\theta}(\theta-1)-\frac{1}{2}\left(\frac{\theta}{\psi}\right)^2-\frac{1}{2}(\theta-1)^2-\frac{\theta(1-\theta)}{\psi}\right] \\
&-q_t\left[(1-\theta)(\kappa_1\rho_q-1)A_q-\frac{1}{2}(\theta-1)^2\kappa_1^2\left(A_\sigma^2+(A_q\varphi_q)^2\right)\right] \\
=&\,\theta\alpha q_t\left(\frac{1}{\psi}-1\right)+\sigma_{g,t}^2\left[\frac{1}{2}\left(\frac{\theta}{\psi}\right)^2+\frac{1}{2}(\theta-1)^2+\frac{\theta(1-\theta)}{\psi}-\frac{(1-\gamma)^2}{2}\right] \\
&+q_t\left[\theta(\kappa_1\rho_q-1)A_q+\frac{1}{2}(\theta-1)^2\kappa_1^2\left(A_\sigma^2+(A_q\varphi_q)^2\right)\right] \\
=&\,\theta\alpha q_t\left(\frac{1}{\psi}-1\right)+\sigma_{g,t}^2\left(\gamma-\frac{1}{2}\right)+q_t\Big[\theta(\kappa_1\rho_q-1)A_q+\frac{1}{2}\theta^2\kappa_1^2\left(A_\sigma^2+(A_q\varphi_q)^2\right) \\
&-(\theta-\frac{1}{2})\kappa_1^2\left(A_\sigma^2+(A_q\varphi_q)^2\right)\Big] \\
=&\,\theta\alpha q_t\left(\frac{1}{\psi}-1\right)+\sigma_{g,t}^2\left(\gamma-\frac{1}{2}\right)-q_t\left[(\theta-\frac{1}{2})\kappa_1^2\left(A_\sigma^2+(A_q\varphi_q)^2\right)\right] \\
&+\theta\alpha q_t\left(1-\frac{1}{\psi}\right) \\
=&\,\sigma_{g,t}^2\left(\gamma-\frac{1}{2}\right)+q_t\left(\frac{1}{2}-\theta\right)\kappa_1^2\left(A_\sigma^2+(A_q\varphi_q)^2\right),
\end{aligned}
$$

where we use the condition (B4).

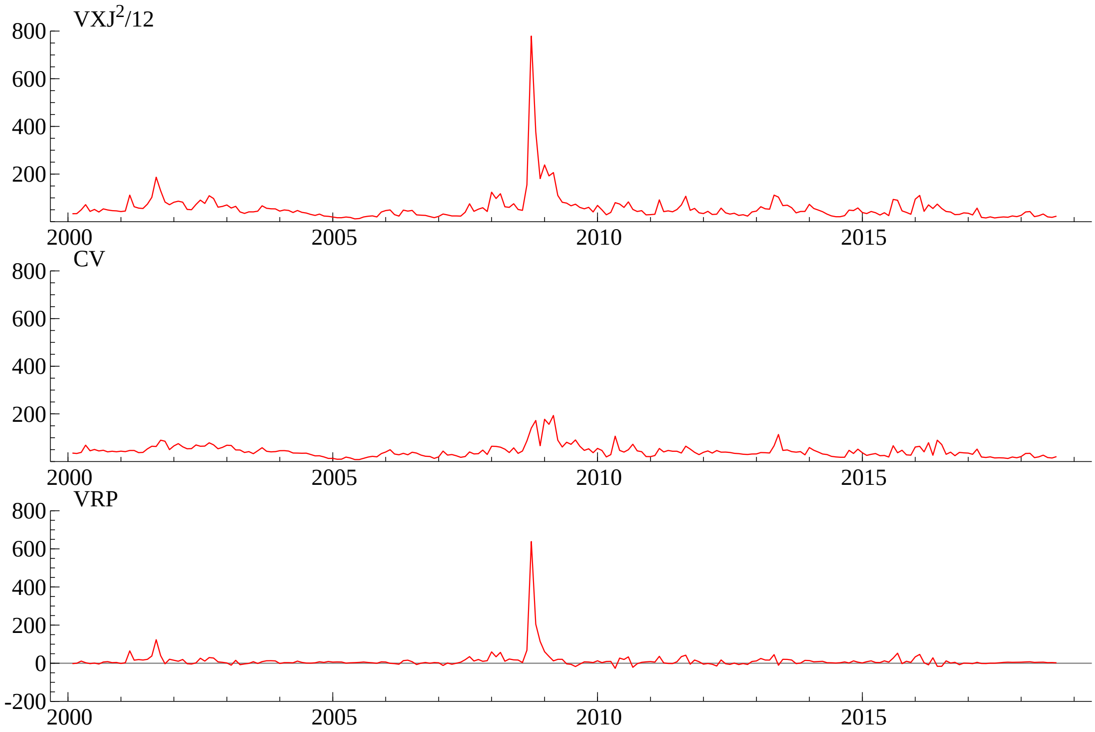
Figure 1: Relation between the Predictability and the Risk-Free Rate by Country



Note: Bollerslev et al. (2014, Table 3) regress the monthly-averaged excess return of a stock index in each of the eight countries from month t to month $t+h$ on the VRP of that index for month t and report the adjusted R^2 for h =1, 2,...,6, 9, and 12. This figure plots the adjusted R^2 for selected hs and the risk-free rates in the eight countries. For each country, we select a specific h which has the largest value of adjusted R^2. The risk-free rates are "short-term interest rates" in $Main\ Economic\ Indicators$ published by OECD, and the averages of these interest rates for the 12 years from 2000−2011, the period analyzed in Bollerslev et al. (2014). Bollerslev et al. (2014, Table 3) also report the Newey-West $t-$statistics for the coefficients of the VRPs. In Bollerslev et al.'s (2014) regressions for the selected hs, the VRPs are estimated to have statistically significant coefficients for six countries at one percent levels of statistical significance as indicated with $\bullet$, and for one country at a five percent level of statistical significance as indicated with $\circ$. Japan's VRP is estimated to have a statistically insignificant coefficient as indicated with $\times$.

Figure 2: Squared VIX (VXJ), CVs and VRPs

(a) Japan (Nikkei 225)



(b) US (S&P500)

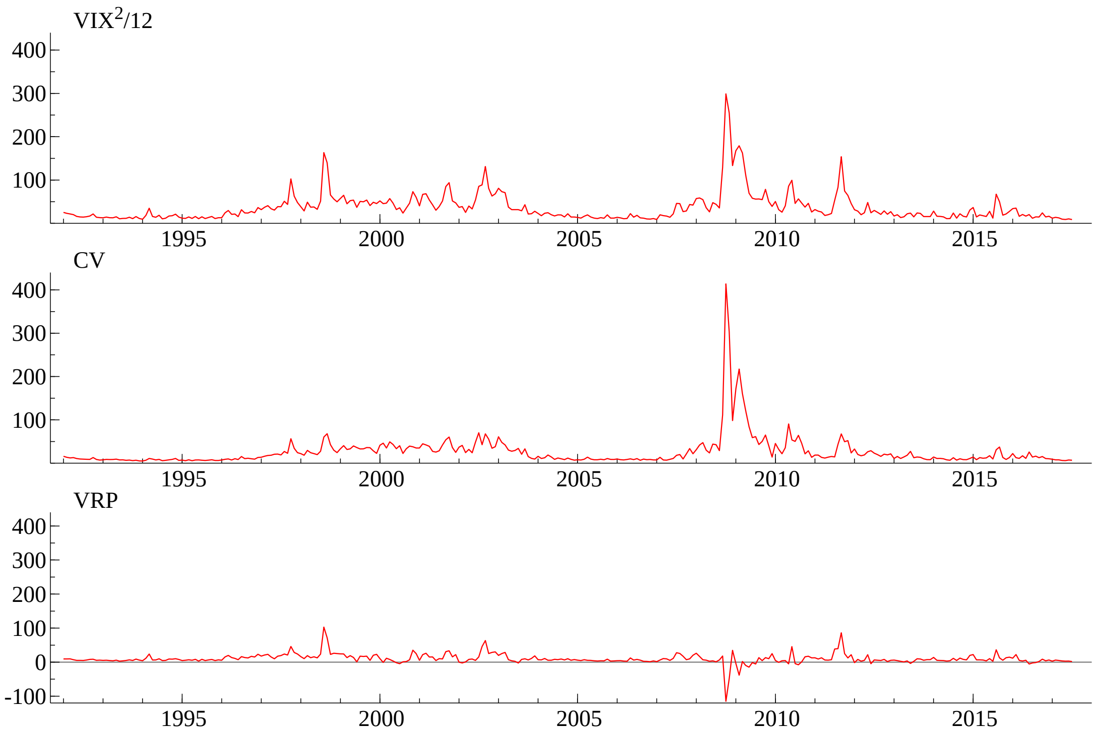
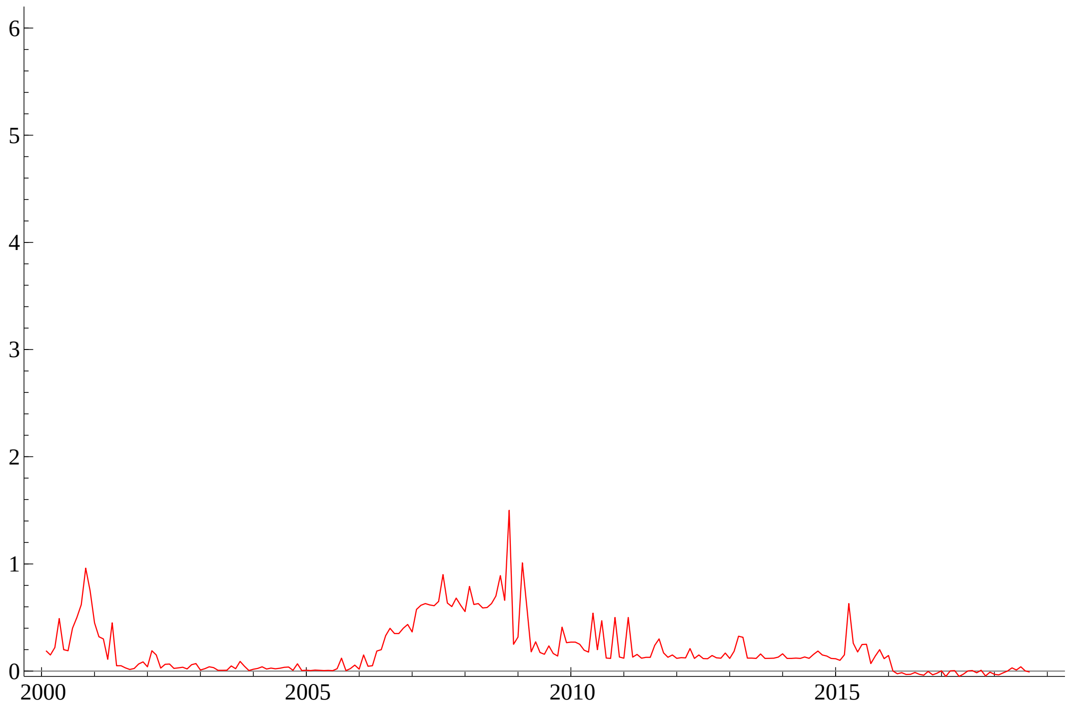


Figure 3: Risk-Free Rates in Japan and US

(a) 1-month call money rates for Japan (%)



(b) 3-month T-bill rates for US (%)

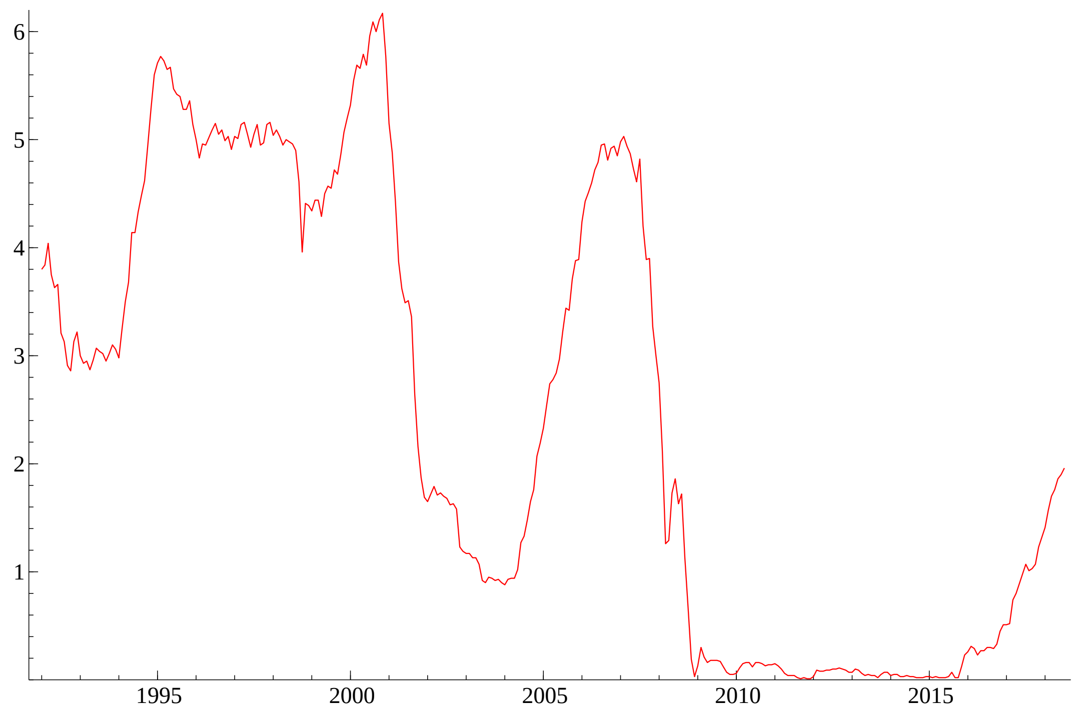
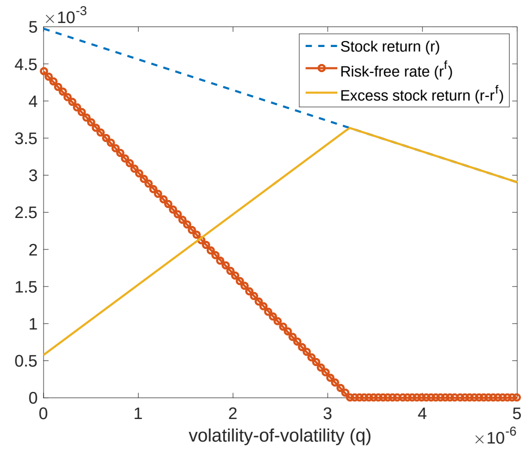
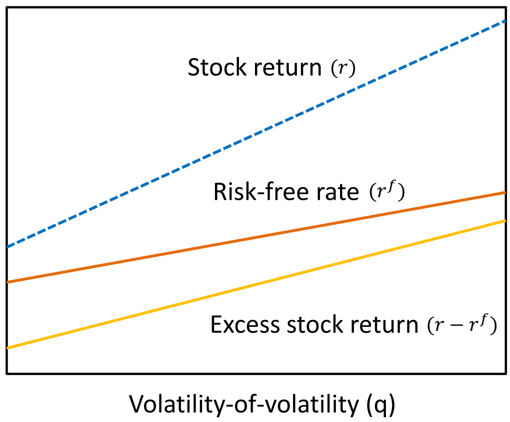


Figure 4: Solutions of Calibrated Model

(a) Excess stock return over q_t (our model)



(b) Excess stock return over q_t (Bollerslev)



(c) VRP over q_t

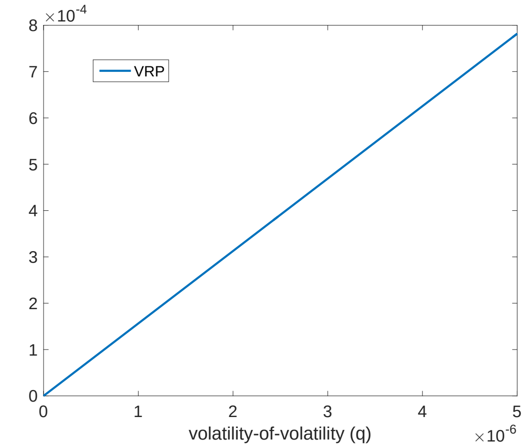


Table 1: Estimation Results of the AHR-CJ Model with the Squared VIX (VXJ)

(a) Nikkei 225

	α_0	α_1	β_d	β_w	β_m	γ_d	γ_w	γ_m	δ	$\overline{\mathrm{R}}^2$
Estimates	0.654	0.276	0.125	0.228	0.157	0.016	0.033	0.060	−0.001	0.598
$t-$stat.	3.637	3.317	7.121	4.562	2.256	3.132	2.230	1.480	−2.687	

The sample period is between 1998.2.5 and 2018.9.28 and the sample size is 5,073. The Newey-West (1987) $t-$statistics with a Bartlett kernel and the lag length of 44 are calculated taking account of the overlapping observations problem. $\overline{R}^2$ is the adjusted R^2.

(b) S&P500

	α_0	α_1	β_d	β_w	β_m	γ_d	γ_w	γ_m	δ	$\overline{\mathrm{R}}^2$
Estimates	0.219	0.316	0.106	0.209	0.214	0.007	0.007	0.109	−0.003	0.715
$t-$stat.	1.678	4.437	7.787	5.778	3.800	1.618	0.309	2.358	−4.706	

The sample period is between 1990.2.1 and 2017.7.31 and the sample size is 6,924. The Newey-West (1987) $t-$statistics with a Bartlett kernel and the lag length of 44 are calculated taking account of the overlapping observations problem. $\overline{R}^2$ is the adjusted R^2.

Table 2: Results of the Regressions (Full Sample)

(a) Nikkei 225 (2000.2–2018.9)

h	1	2	3	4	5	6
α	0.073	0.036	0.097	0.127	0.096	0.069
	(0.163)	(0.080)	(0.226)	(0.300)	(0.225)	(0.161)
β	−0.004	−0.001	−0.005	−0.006	−0.003	−0.001
	(−1.119)	(−0.305)	(−2.612)	(−3.317)	(−1.390)	(−0.234)
$\overline{R}^2$(%)	−0.362	−0.443	0.017	0.401	−0.137	−0.439
h	7	8	9	10	11	12
α	0.055	0.055	0.068	0.080	0.093	0.105
	(0.127)	(0.127)	(0.160)	(0.188)	(0.220)	(0.251)
β	0.001	0.001	0.001	0.001	0.000	−0.000
	(0.376)	(0.480)	(0.341)	(0.244)	(0.103)	(−0.104)
$\overline{R}^2$(%)	−0.416	−0.380	−0.405	-0.427	−0.447	−0.446

The numbers in parentheses are the Newey-West (1987) $t-$statistics with a Bartlett kernel and a lag length determined by $h + 4((T - h)/100)^{2/9}$. $\overline{R}^2$ is the adjusted R^2.

(b) S&P500 (1992.1–2017.7)

h	1	2	3	4	5	6
α	0.051	0.098	0.021	0.039	0.171	0.248
	(0.164)	(0.311)	(0.066)	(0.124)	(0.574)	(0.867)
β	0.032	0.028	0.036	0.034	0.022	0.015
	(2.183)	(1.805)	(2.957)	(3.024)	(1.975)	(1.348)
$\overline{R}^2$(%)	1.047	1.579	4.473	5.307	2.339	1.065
h	7	8	9	10	11	12
α	0.267	0.295	0.342	0.371	0.367	0.367
	(0.951)	(1.075)	(1.308)	(1.454)	(1.468)	(1.469)
β	0.013	0.010	0.006	0.003	0.003	0.003
	(1.223)	(1.046)	(0.571)	(0.286)	(0.342)	(0.379)
$\overline{R}^2$(%)	0.906	0.533	−0.045	−0.251	−0.225	−0.207

The numbers in parentheses are the Newey-West (1987) $t-$statistics with a Bartlett kernel and a lag length determined by $h + 4((T - h)/100)^{2/9}$. $\overline{R}^2$ is the adjusted R^2.

Table 3: Results of the Regressions for S&P5 00

S&P500 (2008.11–2016.10: Zero Interest Rate Period)

h	1	2	3	4	5	6
α	0.930	1.035	0.886	0.970	1.088	1.099
	(2.062)	(2.447)	(2.071)	(2.563)	(3.180)	(3.370)
β	−0.001	−0.013	0.011	0.008	−0.005	−0.007
	(−0.014)	(−0.472)	(0.482)	(0.424)	(−0.277)	(−0.356)
$\overline{R}^2$(%)	−1.063	−0.595	−0.525	−0.631	−0.803	−0.505
h	7	8	9	10	11	12
α	1.092	1.136	1.121	1.114	1.067	1.053
	(3.686)	(4.543)	(4.718)	(4.905)	(4.974)	(5.033)
β	−0.007	−0.013	−0.012	−0.013	−0.008	−0.007
	(−0.424)	(−1.012)	(−1.037)	(−1.126)	(−0.908)	(−0.791)
$\overline{R}^2$(%)	−0.354	2.293	2.614	3.837	1.438	1.170

The number in parenthesis is the Newey-West (1987) $t-$statistics with a Bartlett kernel and a lag length determined by $h + 4((T - h)/100)^{2/9}$. $\overline{R}^2$ is the adjusted R^2.

Table 4: Model Parameters

Parameter	Value	Source
γ	10	Bollerslev et al. (2009)
δ	0.997	Bollerslev et al. (2009)
ψ	1.5	Bollerslev et al. (2009)
κ_1	0.8	In Bollerslev et al. (2009) $\kappa_1 = 0.9$
ρ_σ	0.978	Bollerslev et al. (2009)
ρ_q	0.8	Bollerslev et al. (2009)
$a_q(1-\rho_q)^{-1}$	10^{-6}	Bollerslev et al. (2009)
φ_q	10^{-3}	Bollerslev et al. (2009)
$E(\sigma_g)$	0.0078	Bollerslev et al. (2009)
$E(\mu_g)$	0.0015	Bollerslev et al. (2009)
α	$1.3 * 10^3$	Model-specific parameter
β	0.0028	$E(\mu_g) = \alpha E(q) + \beta$