WORKING PAPER SERIES

Kenichiro Shiraya, Tomohisa Yamakami
and Akira Yamazaki

Forecasting Realized Volatility from Option Prices

July 9, 2025

No. 273

The Research Institute for Innovation Management, HOSEI UNIVERSITY

Forecasting Realized Volatility from Option Prices

Kenichiro Shiraya* Tomohisa Yamakami† Akira Yamazaki‡

This Version: July 9, 2025

Abstract

This paper provides a comprehensive evaluation of volatility forecasters, each reflecting a distinct risk preference, for predicting realized volatility of the S&P 500 index from option prices. These forecasters are benchmarked against a risk-neutral counterpart, corresponding to the VIX index. Our empirical analysis shows that the forecaster drawn from Chabi-Yo and Loudis (2020) consistently delivers the strongest performance: it forms rational expectations for realized volatility, achieves superior out-of-sample predictive accuracy, and substantially improves trading outcomes in variance swap strategies. The remaining forecasters also provide effective predictors, whereas the risk-neutral benchmark exhibits relatively weak predictive performance. Incorporating skewness and kurtosis fails to enhance out-of-sample performance, suggesting that the original predictors are sufficient.

Keywords: realized volatility, predictive power, S&P 500 index, risk preference, option prices

Classification codes: G12, G13, G17

1 Introduction

We evaluate a set of volatility forecasting agents, referred to as *forecasters*, each embodying a distinct risk preference. These forecasters extract their predictions for realized volatility of the S&P 500 index from option prices.

The question of whether financial market dynamics are predictable has long drawn the interest of both academics and practitioners, given its implications for asset pricing and investment decisions. First and foremost, the literature has addressed the predictability of aggregate market returns, which has been the subject of extensive debate. Early evidence suggesting return predictability posed a challenge to the efficient market hypothesis, prompting researchers to investigate predictive variables—most notably the dividend-price ratio—that might contain information about future returns. Panel A of Figure 1 illustrates results from

*Graduate School of Economics, The University of Tokyo. Kenichiro Shiraya is supported by Center for Advanced Research in Finance (CARF).

†Graduate School of Economics, The University of Tokyo.

‡Graduate School of Business Administration, Hosei University. Center for Advanced Research in Finance (CARF), The University of Tokyo. Akira Yamazaki is supported by the Research Institute for Innovation Management at Hosei University.

a simple regression of S&P 500 excess returns over the subsequent 180 days on the dividend-price ratio. The estimated R^2 of 0.109 indicates only modest explanatory power. Moreover, the predictive content of the dividend-price ratio, as well as that of many other proposed predictors, is limited by well-documented issues such as small-sample bias (Stambaugh, 1999), temporal instability, and poor out-of-sample performance (Goyal and Welch, 2008).

In contrast, the forecasting of realized volatility has garnered increasing attention. A large body of research documents that volatility is time-varying (Engle, 1982; Bollerslev, 1986) and responds asymmetrically to unanticipated shocks (Black, 1976; Nelson, 1991). These characteristics complicate the modeling of volatility, prompting the development of nonlinear models capable of capturing its complex dynamics (McAleer and Medeiros, 2008). However, such models often involve difficult estimation challenges: their log-likelihood functions are typically non-convex and may contain multiple local optima, making inference unreliable. Furthermore, Pavlidis et al. (2012) show that these models rely heavily on historical time-series data and tend to perform poorly in out-of-sample forecasts.

As a forward-looking alternative, the VIX index is widely used as a predictor of S&P 500 realized volatility. Unlike time-series models, the VIX index reflects market participants' expectations, rather than being based on past returns. Panel B of Figure 1 presents regression results of realized S&P 500 volatility over the subsequent 30 days on the VIX index. The VIX index demonstrates substantial predictive power for future volatility—arguably stronger than the dividend-price ratio's ability to forecast returns. This empirical contrast motivates our focus on realized volatility forecasting, which appears to offer a more robust and empirically tractable alternative to return predictability.

Nevertheless, relying on the VIX index as a volatility forecast raises both theoretical and empirical concerns. By construction, the VIX index represents the square root of the risk-neutral expectation of future variance and thus reflects the forecast of a risk-neutral investor rather than that under the physical measure. This introduces a systematic upward bias when the VIX index is used to predict realized volatility. Indeed, the slope coefficient of 0.83 in Panel B suggests that the VIX index consistently overstates subsequent volatility. Empirical evidence supports this observation: Carr and Wu (2009) find that variance swaps earn negative risk premia, implying that risk-neutral forecasts tend to exceed realized outcomes. Similarly, Cumby et al. (1993), Jorion (1995), and Adhikari and Hilliard (2014) report that the VIX index performs no better than simple historical volatility as a forecasting tool. These limitations underscore the need for alternative volatility forecasts that incorporate forecaster risk preferences and improve predictive accuracy under the physical measure.

Our approach builds on recent work that derives ex ante market expectations from option prices, beginning with Martin (2017). Martin, as well as Chabi-Yo and Loudis (2020), derives lower bounds on market risk premia from option prices in a manner consistent with asset pricing theory. Martin and Wagner (2019) and Kadan and Tang (2020) extend this framework to estimate bounds on the risk premia of individual stocks. These studies argue that such bounds represent a significant advance in forecasting stock returns. In this paper, we adopt a similar methodology to forecast realized volatility. We consider several forecasters drawn from existing literature, each characterized by a distinct risk preference and constructing a volatility predictor in real time from observed option prices.

We carry out three sets of comparisons using the data on the S&P 500 index and its option prices. First, we benchmark predictive performance against historical volatility, which is sometimes replaced by an expanding-window sample average of realized volatility. Second, we assess performance relative to the risk-neutral forecaster, whose volatility predictor corresponds to the VIX index. A central objective of this paper is to identify a forecaster that outperforms this risk-neutral benchmark. Third, we compare predictive performance across two targets—realized volatility versus market excess returns—using the same set of forecasters, some of whose predictors for market returns correspond to the market premium bounds described above. As shown in Figure 1, predictor variables for realized volatility generally exhibit stronger predictive power than those for market excess returns. Our aim is to quantify the magnitude of this difference.

We assess whether each forecaster forms rational expectations of future realized volatility by regressing realized volatility on the forecaster's volatility predictor. Specifically, we test the following two hypotheses: (i) the volatility predictor is unbiased and (ii) it is a statistically significant predictor of future realized volatility. Our results show that, for most combinations of forecaster and forecasting horizon, both hypotheses cannot be rejected, indicating that these forecasters produce unbiased and informative forecasts of realized volatility. In contrast, when the same tests are applied to predictions of market excess returns, most forecasters fail to generate a significant forecasting signal.

Next, we evaluate each forecaster's predictive power using out-of-sample R^2 alongside a mean squared predictive error (MSPE)-based test. The results reveal that certain forecasters deliver statistically significant predictive power for realized volatility across all forecasting horizons, consistently outperforming the historical average of realized volatility. In contrast, the risk-neutral benchmark becomes insignificant at horizons beyond three months, suggesting that the VIX-like index is not necessarily a better predictor than the historical benchmark. Moreover, the risk-neutral forecaster exhibits a systematic bias, tending to overstate future volatility. By comparison, forecasting performance for market excess returns is uniformly weak across nearly all forecaster-horizon combinations.

To enhance predictive accuracy, we augment our realized volatility regressions by including the skewness and kurtosis of future market returns—derived from option prices and interpreted as proxies for price jump risk—as additional regressors. Although these higher-moment variables are statistically significant in in-sample regressions, the multivariate models deliver predictive performance comparable to the original univariate specifications. Furthermore, their out-of-sample performance deteriorates markedly due to overfitting. These results suggest that the forecasters' unmodified volatility predictors offer superior forecasting performance.

Finally, to assess the economic value of realized volatility forecasts, we conduct a variance swap trading exercise. In this setting, a mean-variance investor relies on a selected forecaster to determine her position in the variance swap. Trading performance is evaluated using both Sharpe ratios and gains in certainty equivalent return. Our results show that strategies based on volatility predictors achieve substantially higher Sharpe ratios than those using a historical volatility benchmark and yield meaningful improvements in investor utility. For comparison, we also perform a parallel exercise in which market return predictors are used to guide investments in the S&P 500 index. In this case, investment performance largely mirrors that of the historical benchmark. Taken together, these findings suggest that volatility predictors generate robust returns in variance swap trading strategies, whereas market return predictors offer limited benefits.

The related literature not cited above is briefly summarized as follows. A substantial body of research has examined the role of implied volatility in forecasting realized volatility.

For instance, Day and Lewis (1992), who study S&P 100 index options, and Lamoureux and Lastrapes (1993), who investigate options on ten individual stocks, find that implied volatility is biased and contains less predictive information about future realized volatility than historical volatility. In contrast, Christensen and Prabhala (1998) and Christensen and Hansen (2002), both of whom examine implied volatility on the S&P 100 index, conclude that implied volatility is an unbiased and efficient predictor for realized volatility.

In parallel, a growing number of studies have extended traditional time-series models to enhance volatility forecasting. Ghysels et al. (2006) propose mixed data sampling (MIDAS) regressions that integrate information from variables observed at different frequencies. Corsi (2009) introduces the Heterogeneous Autoregressive (HAR) model, which captures volatility components realized over different horizons. Kambouroudis et al. (2016, 2021) incorporate implied volatility into time-series models and report improvements in predictive accuracy. More recently, Bucci (2020) explore the use of neural networks for volatility forecasting, demonstrating their competitive performance. In addition, Andersen et al. (2001, 2003) employ high-frequency intraday data to estimate realized volatility at daily and lower frequencies, offering insights into its statistical properties. Caporin (2023) investigates the role of jumps in underlying asset prices and finds that incorporating jump components enhances the in-sample performance of standard volatility models.

Studies on option-implied financial econometrics are directly relevant to this paper. Bakshi et al. (2003) calculate risk-neutral skewness from option prices and document its variation over time and across individual stocks. Back et al. (2022) evaluate the validity, tightness, and predictive power of option-implied risk premium bounds proposed by Martin (2017), Chabi-Yo and Loudis (2020), Martin and Wagner (2019), and Kadan and Tang (2020). They find that while these bounds are valid, they are not tight, and their predictive power remains inconclusive. Bakshi et al. (2023) estimate the probability of market disasters using S&P 500 index options, adopting an investor who dislikes higher volatility states—an agent whose behavior we model as a forecaster in this paper. Yamazaki (2022) proposes a method for recovering stock return distributions from option prices, and Yamazaki (2025) extends this approach to recover subjective probabilities for nonlinear payoffs.

The remainder of this paper is organized as follows. Section 2 outlines the methodology used to forecast realized volatility and introduces the forecasters, who are distinguished by their differing risk preferences. Section 3 describes the data employed in the empirical analysis. Section 4 presents the empirical findings, and Section 5 concludes. Proofs and technical details are provided in the Appendix, while additional empirical results are reported in the Internet Appendix.

2 Methodology

This section presents our methodology for forecasting realized volatility. Our forecasting target is the realized volatility of a stock over the horizon $[t, T]$, defined as

$$\mathrm{RV}_{t,T} := \sqrt{\frac{1}{T-t} \sum_{i \in I} \left(\frac{S_i}{S_{i-\Delta t}} - 1 \right)^2}, \tag{2.1}$$

where S_t denotes the stock price at time t. Here, $I := \{t + \Delta t, t + 2\Delta t, \ldots, T\}$ is the set of monitoring dates, and Δt represents the monitoring interval, which corresponds to one trading day.

2.1 Realized Volatility Predictor

We assume that a forecaster considers the stock price process as follows:

$$\frac{dS_u}{S_u} = \mu_t du + \sigma_u dW_u \qquad \text{for all} \quad u \in [t, T], \tag{2.2}$$

where W_u is a standard Brownian motion under a physical probability measure $\mathbb{P}$, and σ_u is an unknown stochastic volatility process. The forecaster assumes the drift term to be

$$\mu_t := \frac{1}{T-t} \log \mathbb{E}_t^{\mathbb{P}} \left[\frac{S_T}{S_t} \right],$$

where $\mathbb{E}_t^{\mathbb{P}}[\,\cdot\,]$ denotes the time-t conditional expectation operator under $\mathbb{P}$.

We define a realized volatility predictor (hereafter RVP) as

$$\mathrm{RVP}_t := \sqrt{\mathbb{E}_t^{\mathbb{P}} \left[\frac{1}{T-t} \int_t^T \sigma_u^2 du \right]}. \tag{2.3}$$

The aim of this paper is to evaluate the forecasting performance of the RVP with respect to realized volatilities. Forecasters with differing characteristics hold heterogeneous beliefs about the physical probability measure $\mathbb{P}$, and consequently adopt distinct RVPs when predicting future volatility.

Our basic approach to computing the RVP is as follows. By applying Itô's lemma to (2.2), the RVP can be rewritten as

$$\mathrm{RVP}_t = \sqrt{\frac{2}{T-t} \left(\log \mathbb{E}_t^{\mathbb{P}} \left[\frac{S_T}{S_t} \right] - \mathbb{E}_t^{\mathbb{P}} \left[\log \frac{S_T}{S_t} \right] \right)}. \tag{2.4}$$

We assume an arbitrage-free economy, which ensures the existence of a strictly positive stochastic discount factor (hereafter SDF). Let $m_{t,T}$ denote the SDF from time t to T. The projected SDF conditional on S_T, defined by $m_t(S_T) := \mathbb{E}_t^{\mathbb{P}}[m_{t,T}|S_T]$, is assumed to be twice continuously differentiable. For any twice differentiable time-t measurable function $f_t(x)$, the following representation holds:

$$\mathbb{E}_t^{\mathbb{P}}[f_t(S_T)] = \frac{1}{R_{t,T}^f} \left(\frac{f_t(F)}{m_t(F)} \right) + \int_0^F \left(\frac{f_t(K)}{m_t(K)} \right)'' P_t(K) dK + \int_F^\infty \left(\frac{f_t(K)}{m_t(K)} \right)'' C_t(K) dK, \tag{2.5}$$

where $R_{t,T}^f$ denotes the gross return on the risk-free asset from time t to T, and F is the time-t forward price of the stock maturing at time T. $P_t(K)$ and $C_t(K)$ represent the time-t prices of European put and call options written on the stock with strike K and maturity T, respectively.

The formula in (2.5), whose proof is provided in Appendix A, is applied to the right-hand side of (2.4) to compute the RVP. Given a specification of the SDF projection, the RVP can thus be derived from observed option prices. Crucially, the characteristics of forecasters are embedded within the specification of the SDF projection.

It is important to note that, although realized volatility is inherently path-dependent, the RVP is constructed based on information embedded in the terminal distribution of the stock price, as implied by option prices. One of the key research questions of this paper is whether realized volatility can be predicted from the option-implied terminal distribution of the underlying asset price.

In general, the SDF projection can be decomposed[1] as $m_t(S_T) = c_t n_t(S_T)$, where c_t is a time-varying parameter and $n_t(x)$ is a time-t measurable function of x. The parameter c_t can be expressed as

$$c_t = \frac{1}{R^f_{t,T}} \left(\frac{1}{n_t(F)} \right) + \int_0^F \left(\frac{1}{n_t(K)} \right)'' P_t(K) dK + \int_F^\infty \left(\frac{1}{n_t(K)} \right)'' C_t(K) dK. \tag{2.6}$$

The derivation of (2.6) can be found in Appendix A. This expression allows us to avoid unnecessary parameter estimation, as c_t can be directly inferred from observed option prices at time t.

Some potential criticisms may be raised against the RVP as a predictor of realized volatility. One concern is that the assumed stock price process omits jumps. However, it is well known that incorporating jumps into a model-free framework is challenging when evaluating variance swaps and their subspecies. Instead of directly incorporating jumps, we introduce a proxy variable to capture jump risk in Section 4.3.

Another concern is that the RVP is defined as the square root of the expected integral of the variance process, rather than the expected value of the square root of the integral, which might more closely match the definition of realized volatility. Although the latter could, in principle, yield a more accurate predictor, the RVP aligns with the theoretical foundation of the VIX index. To address the bias that arises from the non-commutativity between the square root and the expectation operator, we also assess the predictive performance of the squared RVP as a predictor of realized variance. The results show that using squared RVP leads to the same qualitative conclusions as using the RVP itself for forecasting realized volatility. Detailed results on the forecasting performance for realized variances are provided in the Internet Appendix.

These considerations highlight the theoretical and practical challenges associated with the use of the RVP. Nonetheless, the appeal of the RVP lies in its direct link to option-implied information, offering for a market-based, forward-looking estimate of future volatility. In contrast to purely historical measures, the RVP can incorporate both the beliefs implied by current option prices and the forecaster's risk preferences, thereby reflecting prevailing risk perceptions, market expectations, and the heterogeneity of forecaster views.

For comparison, we also evaluate the forecasting performance of the future annualized excess return on the stock, defined as

$$\mathrm{ER}_{t,T} := \frac{R_{t,T} - R^f_{t,T}}{T - t},$$

where $R_{t,T}$ denotes the gross return on the stock from time t to T. As a predictor of excess

[1]Ghosh et al. (2017) consider a similar decomposition within the framework of consumption-based asset pricing models, referring to c_t as a potentially unobservable component. They demonstrate that most of the time variation in the SDF is attributable to fluctuations in this component.

return, we use the equity risk premium,

$$\mathrm{ERP}_t := \frac{\mathbb{E}_t^{\mathbb{P}}\left[R_{t,T}\right] - R_{t,T}^f}{T-t}, \tag{2.7}$$

which we refer to as the excess return predictor (hereafter ERP) throughout this paper. The expected stock return in (2.7) is computed according to the formula given in (2.5).

2.2 Types of Forecasters

A forecaster with a subjective view of market expectations and distinct risk preferences yields a different RVP, as defined in (2.3). We consider a set of forecasters, each of whom has a different RVP, to predict realized volatility. In the empirical analysis, we assess which forecasters demonstrate superior predictive performance.

2.2.1 Risk-Neutral Forecaster

A risk-neutral forecaster relies on a risk-neutral distribution for the future dynamics of the stock price. The corresponding SDF for such a forecaster is given by $m_{t,T} = 1/R_{t,T}^f$. The resulting risk-neutral RVP coincides with the square root of the standard pricing formula for a variance swap, and is theoretically equivalent to the VIX index. We use the risk-neutral forecaster as a benchmark in our analysis.

2.2.2 CRRA Forecaster

A constant relative risk aversion (CRRA) forecaster is characterized by the isoelastic utility function. The corresponding SDF projection for such a forecaster is given by

$$m_t(S_T) = c_t \left(\frac{S_T}{S_t}\right)^{-\eta},$$

where η is a positive constant. The time-varying parameter c_t is interpreted, in standard asset pricing theory, as a subjective discount factor reflecting the forecaster's rate of time preference. In our framework, c_t can be obtained from (2.6) by setting $n_t(x) = (x/S_t)^{-\eta}$. This specification implies that the forecaster invests his entire wealth in the stock, and his relative risk aversion is equal to η.

When $\eta = 1$, so that the forecaster has log-utility, the resulting ERP coincides with the lower bound of the equity premium proposed by Martin (2017) (see Example 2 of Section III in that paper). Accordingly, the RVP generated by the log-utility forecaster can be viewed as a counterpart to Martin's bound. In addition to the log-utility case, we also employ CRRA forecasters[2] with $\eta = 2$ and $\eta = 3$, which we denote as CRRA2 and CRRA3, respectively, in our empirical analysis.

[2]Bliss and Panigirtzoglou (2004) estimate the risk aversion coefficient of a CRRA investor by maximizing the predictive performance of risk-adjusted probability density functions for the S&P 500 index. At a 6-week forecasting horizon, they report an estimated coefficient of 3.37, which is statistically significant at the 1% level. The estimated values gradually decline as the forecasting horizon increases.

2.2.3 CYL Forecaster

Chabi-Yo and Loudis (2020) adopt a representative investor framework and derive the physical moments of stock excess returns by applying a Taylor series expansion to the reciprocal of the investor's marginal utility. They impose sign restrictions on the risk-neutral moments of excess returns—specifically, requiring that the odd-order moments are weakly negative. By further constraining the investor's tolerance for risk, skewness, and kurtosis, they obtain an approximate formula for the physical moments that avoids the need to estimate preference parameters. The first-order moment in this formula yields a lower bound on the equity risk premium. Empirically, they demonstrate that, at longer horizons, this bound exhibits better predictive power than the one proposed by Martin (2017).

We adopt the representative investor considered by Chabi-Yo and Loudis (2020) as a forecaster of realized volatility. Hereafter, we refer to this forecaster as the CYL forecaster, named after Chabi-Yo and Loudis (2020). In the case of the CYL forecaster, the formula in (2.5) cannot be directly applied to compute the RVP. Instead, following the results of Chabi-Yo and Loudis (2020), we employ the following alternative approximate expressions:

$$\mathbb{E}_t^{\mathbb{P}}\left[\frac{S_T}{S_t}\right] \approx R_{t,T}^f + BR_1,$$

and

$$\mathbb{E}_t^{\mathbb{P}}\left[\log\frac{S_T}{S_t}\right] \approx \log R_{t,T}^f - \sum_{k=1}^{4}\frac{(-R_{t,T}^f)^{-k}}{k}\left(M_k^{\mathbb{Q}} + BR_k\right),$$

where $M_n^{\mathbb{Q}}$ denotes the n-th order risk-neutral moment of the stock's excess price return, and BR_k is the restricted bound defined in (B.2) of Appendix B.3. Further details are provided in Appendix B.

2.2.4 BGX Forecaster

Bakshi et al. (2023) propose the following SDF projection:

$$m_t(S_T) = c_t \exp\left(\eta_0 - 1 + \eta_1 Z_{t,T}\right), \tag{2.8}$$

where η_0 and η_1 are constants, and $Z_{t,T}$ represents the excess return on a volatility contract, defined as

$$Z_{t,T} := \frac{\left\{\log\frac{S_T}{S_t}\right\}^2}{v_t} - R_{t,T}^f.$$

Here, v_t is the time-t price of a volatility contract that pays $\{\log(S_T/S_t)\}^2$ at maturity T. While their original specification assumes $c_t = 1$ for all time t, so that the SDF is simply $n_t(S_T) = \exp\left(\eta_0 - 1 + \eta_1 Z_{t,T}\right)$, we instead apply (2.6) to allow for a time-varying parameter c_t in the SDF projection. The model-free pricing formula for the volatility contract and the parameter estimation procedure are detailed in Appendix C. We refer to the forecaster associated with the SDF in (2.8) as the BGX forecaster, named after Bakshi et al. (2023).

The BGX SDF incorporates a state variable representing the excess return on the volatility contract. When $\eta_1 > 0$, the BGX forecaster exhibits aversion to high-volatility states.

It is worth noting that the original specification proposed by Bakshi et al. (2023) also incorporates the excess return on the underlying stock as a state variable. However, based on empirical analysis using S&P 500 index data, Bakshi et al. (2023) find this state variable to be statistically insignificant. In light of this result, we exclude it from our implementation. Variance-dependent SDFs of this type have also been proposed by Christoffersen et al. (2013) and Song and Xiu (2016).

Utilizing the BGX SDF, Bakshi et al. (2023) estimate the probability of equity market disasters, defined as substantial declines in the S&P 500 index. Their findings indicate that the disaster probabilities, inferred from S&P 500 index options, not only exhibit predictive power for realized disaster events but are also systematically linked to the likelihood of extreme positive returns.

2.3 Risk Preferences

The forecasters introduced in the previous subsection can be distinguished by their risk preferences. We classify the characteristics of these forecasters using the concepts of relative risk aversion $\mathcal{A}$, relative prudence $\mathcal{P}$, and relative temperance $\mathcal{T}$, defined as

$$\mathcal{A} := -\frac{x m_t'(x)}{m_t(x)}, \qquad \mathcal{P} := -\frac{x m_t''(x)}{m_t'(x)}, \quad \text{and} \quad \mathcal{T} := -\frac{x m_t'''(x)}{m_t''(x)}. \tag{2.9}$$

These definitions are not based on a conventional utility function, but rather follow Rosenberg and Engle (2002), who define risk preference measures in terms of the SDF.

Both the CRRA and CYL forecasters are characterized by constant relative risk preferences. The CRRA forecaster exhibits relative risk aversion $\mathcal{A} = \eta$, relative prudence $\mathcal{P} = \eta+1$, and relative temperance $\mathcal{T} = \eta + 2$. Similarly, the CYL forecaster has constant values of $\mathcal{A} = 1$, $\mathcal{P} = 4$, and $\mathcal{T} = 6$, with the derivations provided in Appendix B. In contrast, the BGX forecaster exhibits state-dependent relative risk preferences.

Figure 2 illustrates the relative risk preferences of five forecasters—log-utility (Log-U), CRRA2, CRRA3, CYL, and BGX—excluding the risk-neutral forecaster. The CYL forecaster exhibits the same level of relative risk aversion as the log-utility forecaster, implying that its degree of risk aversion aligns with that of the forecaster consistent with Martin's (2017) lower bound. However, the CYL forecaster displays higher relative prudence than both the log-utility and CRRA2 forecasters, and a level comparable to that of the CRRA3 forecaster.

The relative risk preferences of the BGX forecaster shown in Figure 2 are based on the parameter value $\eta_1 = 0.17$, which corresponds to the estimate of the BGX SDF at the 3-month horizon, and the volatility contract price $v_t = 0.0121$, reflecting the historical average for contracts with 3 months to maturity. Further details on these values are provided in Appendix C. Panel A shows that the relative risk aversion of the BGX forecaster declines as the market return increases and becomes negative when the net market return turns positive. This indicates that the BGX forecaster becomes risk-seeking in states where the net market return is positive. Consequently, as shown in Panel B, the relative prudence of the BGX forecaster diverges when the market is nearly flat—that is, when the gross market return is close to one.

3 Data

3.1 Market Data

In our empirical analysis, we use the S&P 500 Index as the underlying asset for predicting realized volatility. To compute the RVP, we construct a dataset using daily historical option price data on the S&P 500 Index, obtained from the Cboe DataShop (`https://datashop.cboe.com`). The sample period covers January 2007 to August 2023. We collect bid and ask quotes of out-of-the-money option prices at 3:45 p.m. U.S. Eastern Time, as listed on the Chicago Board Options Exchange.

For interest rate data, we use U.S. Treasury bill secondary market rates obtained from the Federal Reserve Economic Data (FRED) (`https://fred.stlouisfed.org`). Dividend yields on the S&P 500 Index are calculated by the put-call parity, using prices of near-the-money call and put options. This approach to extracting implied dividend yields has been employed in previous studies, such as Aït-Sahalia and Lo (1998) and Polkovnichenko and Zhao (2013).

Our analysis considers forecasting horizons of 1, 2, 3, 4, and 6 months. To align option data with these horizons, we first compute the Black-Scholes implied volatilities using the midpoints of bid and ask quotes. We then apply linear interpolation to the squared implied volatilities of options with maturities close to the target horizon and identical strike prices. Interest rates and dividend yields corresponding to each horizon are also obtained through linear interpolation. These procedures enable us to construct constant-maturity market data suitable for our analysis.

To eliminate data errors and ensure reliable empirical results, we apply several screening criteria. First, we discard any options with bid quotes less than \$0.025 or with negative bid-ask spreads. We also exclude options with zero open interest or trading volume. For constructing 1-month market data, we eliminate options with fewer than 8 trading days remaining until expiration. To ensure no-arbitrage conditions, we further discard any options that violate the monotonicity of option prices across strikes or fall below the lower bounds of option prices.

3.2 RVP and Realized Volatility

The procedure for calculating the RVP, based on (2.4) and (2.5), proceeds as follows. First, implied volatility curves are constructed for each observation date by interpolating implied volatilities from the constant-maturity market data. To achieve this, we apply a cubic smoothing spline technique, which provides a flexible and robust method for smoothing noisy data while avoiding overfitting. The spline is fitted to the observed implied volatilities as a function of log-moneyness, defined as $\log K/S_t$.

For strike prices beyond the available market range, we extrapolate using the spline-fitted values at the boundaries. Specifically, when log-moneyness falls below the lowest quoted level, we adopt the implied volatility estimated at the lowest strike; conversely, when it exceeds the highest quoted level, we use the value at the highest strike. This extrapolation approach follows the methodology proposed by Carr and Wu (2009).

The resulting implied volatility curves—combining interpolation within the observed range and extrapolation beyond it—are then used to compute synthetic option prices via the Black-Scholes formula. These prices serve as inputs for the formulas in (2.5) and (2.6). For the numerical integration involved in computing these quantities, we truncate the range of strike prices to within ±10 standard deviations of the current index level, where the standard deviation is approximated using the historical average of near-the-money implied volatilities. The ERP is computed using the same procedure.

We consider 30 forecaster-horizon combinations, comprising six types of forecasters—the risk-neutral (RN), log-utility (Log-U), CRRA with $\eta = 2$ (CRRA2) and $\eta = 3$ (CRRA3), CYL, and BGX—and five forecasting horizons (1-, 2-, 3-, 4-, and 6-month). Table 1 reports summary statistics for their respective RVPs, along with realized volatility, denoted as "Realized" in the table. The sample comprises daily observations on the RVPs from January 2007 to August 2023, along with the corresponding realized volatility. The means of the risk-neutral, log-utility, and CYL RVPs exceed that of realized volatility by approximately 0.2% to 4.6% per year, depending on the type of forecaster and the forecasting horizon. In contrast, the mean of the BGX RVP is approximately 1.9% to 2.2% per year lower than that of realized volatility, depending on the horizon. The standard deviations of all RVPs are smaller than that of realized volatility by about 2.1% to 3.8%. Moreover, while the RVPs exhibit positive skewness and substantial excess kurtosis, both are less pronounced than those observed in realized volatility. However, the standard deviations, skewness, and kurtosis of RVPs may not be directly comparable to those of realized volatility due to differences in their statistical construction.

Table 2 presents summary statistics for realized excess returns, denoted as "Realized" in the table, along with the ERPs corresponding to the RVP data reported in Table 1. The risk-neutral ERP is omitted, as it is identically zero. A comparison of the two tables, focusing on the CRRA forecasters, reveals a clear pattern: higher relative risk aversion is associated with lower RVPs and higher ERPs.

Figure 3 presents the time series of the RVPs (left panels) and ERPs (right panels). The RVPs exhibit strong co-movement. For example, the correlation between the risk-neutral RVP and the BGX RVP at the 3-month horizon is as high as 0.97. Nevertheless, an ordering in the levels of the RVPs is observed. Across all horizons, the risk-neutral RVP consistently displays higher values than the other RVPs throughout most of the sample period, whereas the BGX RVP tend to remain the lowest. These differences in the levels of the RVPs become more pronounced as the forecasting horizon increases. As expected, the RVPs exhibit pronounced peaks during episodes of market distress, such as the 2008 financial crisis and the COVID-19 shock in early 2020. The RVPs also appear to be strongly correlated with the ERPs. For instance, the correlation between the RVP and the ERP for the log-utility forecaster is 0.95. This high correlation is theoretically consistent, as the ERP for the log-utility forecaster—equivalent to Martin's (2017) bound—corresponds to the discounted value of *the risk-neutral variance* of the gross market return, as shown by Martin (2017).

4 Empirical Results

4.1 Expectation Hypothesis Test

To examine the validity of a given RVP as a forecast of realized volatility, we run the following expectation hypothesis regression over the full sample period:

$$\mathrm{RV}_{t,T} = \alpha + \beta\,\mathrm{RVP}_t + \varepsilon_T.$$

Under the null hypothesis that the RVP represents the expected value of realized volatility, we expect $\alpha = 0$ and $\beta = 1$. In addition, we test whether the RVP is a statically significant predictor by evaluating the null hypothesis of $\beta = 0$. For comparison, we also include historical volatility, which is measured at time t over a backward-looking window of the same length as the forecasting horizon, as an alternative explanatory variable in the regression. To evaluate the expectation hypothesis, we identify which RVPs satisfy both of the following conditions: failure to reject the null hypothesis that $\alpha = 0$ and $\beta = 1$, and rejection of the null hypothesis that $\beta = 0$.

Table 3 presents the estimated intercept and slope coefficients from the regression using daily data with overlapping forecasting horizons. Regression results based on monthly data are reported in the Internet Appendix. The values in parentheses show the t-statistics for testing the null hypotheses that $\alpha = 0$ and $\beta = 1$, while the values in square brackets correspond to the t-statistics for testing the null hypothesis that $\beta = 0$. All t-statistics are adjusted for serial dependence using the Newey and West (1987) method, with the number of lags set to 1.5 times the number of days in the forecasting horizon, following the lag-length choice adopted by Back et al. (2022). The R^2 values reported in the table are expressed as percentages.

Table 3 shows that, for the 1-month horizon, the null hypothesis of $\alpha = 0$ is rejected at the 10% significance level for the log-utility, CRRA2, CRRA3, and CYL forecasters, with t-statistics of -1.65, -1.80, -1.91, and -1.81, respectively. For the 6-month horizon, the null hypothesis of $\beta = 1$ is rejected at the 5% and 10% levels for the risk-neutral and BGX forecasters, with t-statistics of -2.40 and -1.68, respectively. These results indicate that the corresponding RVPs are biased as forecasts of realized volatility at these horizons. On the other hand, the null hypothesis of $\beta = 0$ is rejected at the 1% level for all forecasters across all horizons, indicating that each RVP is a statistically significant predictor. Taken together, 24 out of the 30 forecaster-horizon combinations satisfy the expectation hypothesis, with the remaining six cases failing to do so.

For comparison, Table 4 reports the results of the expectation hypothesis regression of realized excess returns for the S&P 500 index on the ERP, replacing the RVP as the independent variable. The table shows that the null hypothesis of $\alpha = 0$ and $\beta = 1$ is not rejected for most forecasters, consistent with previous studies such as Martin (2017) and Chabi-Yo and Loudis (2020). However, with the exception of the 6-month horizon and a few other cases, the null hypothesis of $\beta = 0$ is also not rejected, indicating that most ERPs lack statistical significance as predictors of realized excess returns. Taken together, only four of the 30 forecaster-horizon combinations satisfy the expectation hypothesis for the S&P 500 index excess returns: the CRRA2, CRRA3, and BGX forecasters at the 6-month horizon, and the BGX forecaster at the 2-month horizon. The remaining 26 combinations fail the test. It is also noteworthy that the R^2 values from the excess return regressions are substantially lower than those from the realized volatility regressions, ranging from 0.1% to 12%, compared with 20% to 53% for realized volatility.

Figure 4 plots the point estimates of the slope coefficients along with their 95% confidence intervals from regressions based on monthly data. The confidence intervals are calculated using the Newey and West (1987) estimator, with the lag length set to the forecasting horizon (in months) plus 12 months. The left panels display estimates from regressions of realized volatility on the RVPs, while the right panels show those from regressions of excess returns on the ERPs. The figure demonstrates that slope estimates for realized volatility are substantially more precise than those for excess returns. Notably, the 95% confidence intervals for the slope coefficients on the risk-neutral RVPs at the 1-, 3-, and 6-month horizons do not include one, suggesting that these RVPs—corresponding to the VIX—are biased as forecasts of realized volatility.

4.2 Forecasting Performance

In this subsection, we evaluate the forecasting performance of the RVPs for realized volatility. A key requirement in forecasting realized volatility at time t is that only information available up to time t may be used; no future information is permitted. To assess forecasting performance, we employ the R^2_{OS} statistic proposed by Campbell and Thompson (2008), which is defined as:

$$R^2_{OS} := 1 - \frac{\sum_t (\mathrm{RV}_{t,T} - \hat{\mathrm{RV}}_{t,T})^2}{\sum_t (\mathrm{RV}_{t,T} - \bar{\mathrm{RV}}_{t,T})^2}, \tag{4.1}$$

where $\hat{\mathrm{RV}}_{t,T}$ represents the forecasted realized volatility, and $\bar{\mathrm{RV}}_{t,T}$ denotes the historical average of realized volatility based on an expanding window using information available up to time t. The R^2_{OS} statistics compares the mean squared predictive error (MSPE) of the forecasted realized volatility to that of the historical average. It ranges from $-\infty$ to 1, with $R^2_{OS} > 0$ indicating that the forecasted realized volatility $\hat{\mathrm{RV}}_{t,T}$ outperforms the historical average benchmark $\bar{\mathrm{RV}}_{t,T}$ in terms of MSPE.

In the full-sample forecasting test, we use the RVPs as the forecasted realized volatility:

$$\hat{\mathrm{RV}}_{t,T} = \mathrm{RVP}_t, \tag{4.2}$$

excluding the BGX RVP. This exclusion is due to the fact that estimating the parameters of the BGX SDF requires market data from portions of the sample period that are not yet available at time t, thereby violating the information constraint inherent in the forecasting test. In contrast, the other RVPs do not involve any parameter estimation and thus satisfy the requirement. The expanding window used to calculate the historical average benchmark begins in January 1997, which is 10 years prior to the start of the full-sample period in January 2007.

In the out-of-sample forecasting test, we divide the sample period into two subperiods: the in-sample period is set from January 2007 to December 2016, and the out-of-sample period spans from January 2017 to August 2023. For this test, we estimate the predictive regression model recursively as follows:

$$\hat{\mathrm{RV}}_{t,T} = \alpha_t + \beta_t \mathrm{RVP}_t, \tag{4.3}$$

where α_t and β_t denote the intercept and slope coefficient, respectively, estimated using all data available up to time t. The forecasted realized volatility in (4.3) can be interpreted as a bias-adjusted version of the RVP. For comparison, we also use the forecast based on the raw RVP, as in (4.2), in the out-of-sample forecasting test. For the BGX RVP, the parameters of the BGX SDF are estimated using an expanding window over the sample period. The resulting parameter estimates are reported in Table 12 of Appendix C. The expanding window for computing the historical average benchmark in the out-of-sample test begins in January 2007.

We evaluate forecast accuracy using an MSPE-based test. The null hypothesis is that the MSPE of the historical average benchmark is less than or equal to that of the forecasted realized volatility, against the one-sided alternative that the benchmark MSPE is greater than the forecasted volatility MSPE. Equivalently, this corresponds to testing $R^2_{OS} \leq 0$ versus $R^2_{OS} > 0$. The p-values are calculated using standard errors corrected for autocorrelation following Newey and West (1987). When the forecasted realized volatility is the bias-adjusted RVP from (4.3), we employ the MSPE-adjusted statistic proposed by Clark and West (2007) to account for the fact that the forecast model nests the historical average benchmark.

Table 5 reports the forecasting performance for realized volatility based on daily data. Results using monthly data, reported in the Internet Appendix, are very similar. All R^2_{OS} values are expressed as percentages.

According to the full-sample results, all forecasters produce positive R^2_{OS} values except for the risk-neutral forecaster at the 6-month horizon. For this forecaster, R^2_{OS} is not statistically significantly positive at horizons of 3 months or longer. Hence, beyond a 3-month horizon, the VIX-like index does not necessarily outperform the historical average benchmark in forecasting realized volatility. By contrast, the CRRA2, CRRA3, and CYL forecasters consistently yield higher R^2_{OS} values than the risk-neutral forecaster, and their performance is statistically significant across all horizons. These results indicate that, in the full-sample test, these three forecasters demonstrate superior predictive power relative to both the risk-neutral forecaster and historical average benchmark.

Turning to the out-of-sample results using the bias-adjusted RVP, R^2_{OS} is significantly positive for all forecasters at every forecasting horizon. When using the raw RVP instead, only the CRRA2 and CYL forecasters maintain significantly positive R^2_{OS} across all horizons. Notably, for the risk-neutral forecaster, R^2_{OS} based on the raw RVP is 6% to 30% lower than its bias-adjusted counterpart, indicating a systematic bias in the VIX-like index. In contrast, the CRRA2, CRRA3, and CYL forecasters show equal or even higher R^2_{OS} values when using the raw RVP compared to the bias-adjusted RVP, suggesting that the bias adjustment may lead to overfitting in these cases.

Taken together, the full-sample and out-of-sample results suggest that the CRRA2 and CYL forecasters provide the most reliable and unbiased forecasts of realized volatility.

For comparison, we conduct the same forecasting performance tests on excess returns, using the ERP in place of the RVP as the predictor. The results are reported in Table 6. Across both the full-sample and out-of-sample tests, the R^2_{OS} values for forecasted excess returns are substantially lower than those for forecasted realized volatility. In many cases, the values are close to zero or negative and fail to reach statistical significance. The lone exception occurs the 6-month horizon in the full-sample test, and in the out-of-sample test for the CRRA2, CRRA3, and BGX forecasters at the 1-month horizon, as well as for the CRRA2 forecaster at the 2-month horizon. These findings indicate that, relative to realized volatility, none of the forecasters consistently produces reliable predictions of excess returns.

4.3 Skewness and Kurtosis as Predictors of Realized Volatility

A limitation of the RVP is that the stock price process in (2.2) does not account for price jumps. Incorporating jumps in a model-free framework, however, remains a challenging task. As an alternative, this subsection explores the use of skewness and kurtosis of the net price return distribution perceived by a forecaster as predictive variables to enhance the RVP.

It is well established that negative price jump risk induces negative skewness and elevated excess kurtosis, typically accompanied by heightened volatility (Merton, 1976; Kou, 2002). Therefore, skewness and kurtosis may be regarded as practical proxies for jump risk.

4.3.1 SKW and KRT

For a forecaster operating under the physical measure $\mathbb{P}$, skewness (hereafter SKW) and excess kurtosis (hereafter KRT) of the net price return at time t are defined respectively as

$$\mathrm{SKW}_t := \frac{m_3^{\mathbb{P}} - 3m_1^{\mathbb{P}}\,\mathrm{SD}_t^2 - (m_1^{\mathbb{P}})^3}{\mathrm{SD}_t^3},$$

and

$$\mathrm{KRT}_t := \frac{m_4^{\mathbb{P}} - 4m_1^{\mathbb{P}}m_3^{\mathbb{P}} + 6(m_1^{\mathbb{P}})^2\,m_2^{\mathbb{P}} - 3(m_1^{\mathbb{P}})^4}{\mathrm{SD}_t^4} - 3,$$

where

$$\mathrm{SD}_t := \sqrt{m_2^{\mathbb{P}} - (m_1^{\mathbb{P}})^2},$$

and

$$m_n^{\mathbb{P}} := \mathbb{E}_t^{\mathbb{P}}\left[\left(\frac{S_T}{S_t} - 1\right)^n\right].$$

Here, SD_t denotes the standard deviation of the net price return, and $m_n^{\mathbb{P}}$ is the n-th non-central moment of the net price return under the physical measure. For the risk-neutral, CRRA, and BGX forecasters, these moments are computed using the formula in (2.5) with $f_t(x) = (x/S_t - 1)^n$, whereas the corresponding procedure for the CYL forecaster is detailed in Appendix B.

Summary statistics for SKWs and KRTs across forecasters and forecasting horizons are provided in the Internet Appendix. The means of SKWs are negative, ranging from -1.37 to -0.56, whereas the means of KRTs range from 1.56 to 4.17. SKWs exhibit smaller standard deviations than KRTs. The distributions of SKWs are positively skewed for all forecasters except the risk-neutral and log-utility forecasters at the 1-month horizon, while the distributions of KRTs exhibit positive skewness across all forecasters and horizons. Both SKW and KRT distributions display pronounced excess kurtosis. Correlations among RVPs, SKWs, and KRTs are non-negligible. For example, for the CYL forecaster at the 1-month horizon, the correlation between the RVP and SKW is 0.70, while that between the RVP and KRT is -0.58. Overall, RVPs tend to be positively correlated with SKWs and negatively correlated with KRTs.

4.3.2 In-Sample Predictive Regression

To assess whether skewness and kurtosis enhance predictive power, we run a multivariate predictive regression using the full-sample data, including SKW and KRT alongside RVP as predictors. To address multicollinearity among the predictors and remove irrelevant variation, we employ the partial least squares (PLS) regression method. PLS is a linear dimension-reduction technique that projects both the dependent variable and the set of predictors onto a new space that maximizes their covariance, thereby identifying components most relevant for prediction. Originally proposed by Wold (1966, 1975), PLS regression has been successfully applied in return forecasting by Kelly and Pruitt (2013) and Huang et al. (2015).

Table 7 reports the estimated predictor loadings of the PLS components for each forecaster at the 1- and 2-month horizons. Results for other forecasting horizons, presented in the Internet Appendix, exhibit similar patterns. All forecasters and horizons display a common structure across the first two PLS components. For the first component, both RVP and SKW have positive loadings, while KRT has a negative loading. By contrast, the second component assigns positive loadings to RVP and KRT and a negative loading to SKW. This second component thus captures a dynamic in which an increase in RVP is associated with more negative SKW and higher KRT, and vice versa—behavior consistent with negative jump risk in the S&P 500 index. Consequently, we interpret the second component[3] as a proxy for negative jump risk.

In Table 7, PctVar1 denotes the percentage of variance in the predictor set (RVP, SKW, and KRT) explained by each PLS component, while PctVar2 represents the percentage of variance in the dependent variable (realized volatility) explained by each component. For all forecasters and horizons, PctVar1 for the third PLS component is substantially lower than for the first two components. Moreover, PctVar2 for the third component remains below 1%, indicating it contributes no meaningful predictive power for realized volatility. Therefore, we discard the third component and retain only the first two components in the subsequent predictive regressions.

We estimate the following bivariate regression model:

$$\mathrm{RV}_{t,T} = \alpha + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \varepsilon_T,$$

where $\{X_{1,t}\}$ and $\{X_{2,t}\}$ denote the scores on the first and second PLS components, respectively.

Table 8 reports the estimated coefficients β_1 and β_2, along with adjusted R^2 values (expressed as percentages), based on daily data with overlapping forecasting horizons. Values in parentheses are t-statistics computed using Newey and West (1987) standard errors, with the number of lags set to 1.5 times the number of days in the forecasting horizon. Results for the monthly predictive regressions are presented in the Internet Appendix.

As shown in Table 8, with the only exception of the β_2 estimate for the BGX forecaster at the 6-month horizon, all coefficients are positive and statistically significant. Since RVPs have positive loadings on both the first and second PLS components, these bivariate regression results are consistent with the univariate regression findings in Section 4.1—namely, RVPs possess genuine predictive power for realized volatility. At first glance, the significantly positive β_2 estimates suggest that the second component—serving as a proxy for negative jump risk—may offer additional predictive power. However, the adjusted R^2 values from the bivariate regressions remain comparable to the ordinary R^2 values from the univariate regression shown in Table 3, indicating no substantial improvement in overall fit.

[3] As reported in the Internet Appendix, the CYL forecaster at the 4-month horizon and the CRRA2, CRRA3, CYL, and BGX forecasters at the 6-month horizon deviate from the second component loading pattern observed in Table 7. Consequently, for these forecaster-horizon combinations, the second PLS component cannot be interpreted as a proxy for negative jump risk.

4.3.3 Out-Of-Sample Forecasting Performance

We next evaluate the out-of-sample performance of the PLS regression model using the same framework as in Section 4.2. The out-of-sample period spans January 2017 to August 2023. Forecasting performance is assessed using the R^2_{OS} statistics defined in (4.1) and the MSPE-based test. In the out-of-sample test, the forecasted realized volatility is given by

$$\hat{\mathrm{RV}}_{t,T} = \alpha_t + \beta_{1,t}\hat{X}_{1,t} + \beta_{2,t}\hat{X}_{2,t}, \tag{4.4}$$

where $\{\hat{X}_{1,u}\}_{u\leq t}$ and $\{\hat{X}_{2,u}\}_{u\leq t}$ represent the first and second PLS component scores, respectively, extracted from data available up to time t, and α_t, $\beta_{1,t}$, and $\beta_{2,t}$ are the corresponding coefficient estimates.

Table 9 presents the out-of-sample forecasting performance of the PLS regression models for realized volatility, based on daily data. A large proportion of the R^2_{OS} values are negative, and even when the values are positive, they do not exceed 2.5% and fail to achieve statistical significance under the MSPE-based test. Compared to the in-sample results in Table 8, the out-of-sample forecasting power deteriorates markedly. These findings suggest that incorporating SKW and KRT into the PLS regression leads to pronounced overfitting. In conclusion, the raw RVP proves to be a more reliable predictor of realized volatility than the more complex multivariate models. Results for the monthly-based out-of-sample forecasting performance are presented in the Internet Appendix.

4.4 Implications for Variance Swap Trading

In this subsection, we examine the implications of the RVP for volatility trading. Although realized volatility itself is not directly tradable, realized variance—defined as the square of realized volatility in (2.1)—can be traded via variance swaps. We evaluate the performance of a mean-variance investor who allocates capital between a variance swap and the risk-free asset, using the squared RVP as a predictor for future realized variance. Following Kandel and Stambaugh (1996), Campbell and Thompson (2008), Ferreira and Santa-Clara (2011), and Huang et al. (2015), we assess the strategy's performance based on its Sharpe ratio and the gain in certainty equivalent return (hereafter CER).

We consider variance swap trading over the horizon from time t to T. At time t, a mean-variance investor allocates a fraction of wealth

$$w_t = \frac{1}{\gamma}\frac{\hat{R}^v_{t,T}}{\hat{\sigma}^2_{t,T}},$$

to a variance swap maturing at time T, where γ denotes the investor's risk aversion coefficient, $\hat{R}^v_{t,T}$ is the forecasted excess return on the variance swap, and $\hat{\sigma}_{t,T}$ is the forecasted standard deviation of the variance swap return. The forecasted excess return is given by

$$\hat{R}^v_{t,T} := \frac{\mathrm{RVP}_t^2}{\frac{1}{R^f_{t,T}}(\mathrm{RVP}_t^{\mathbb{Q}})^2} - R^f_{t,T} = R^f_{t,T}\left(\frac{\mathrm{RVP}_t^2 - (\mathrm{RVP}_t^{\mathbb{Q}})^2}{(\mathrm{RVP}_t^{\mathbb{Q}})^2}\right),$$

where $\mathrm{RVP}_t^{\mathbb{Q}}$ denotes the risk-neutral RVP, whose square serves as the variance swap rate. Thus, we assume that the investor is not necessarily identical to the forecaster but instead takes a trading position in the variance swap based on the RVP provided by the selected forecaster. The remaining weight, $1 - w_t$, is invested in the risk-free asset. The resulting realized portfolio return at time T is given by

$$R^p_{t,T} = w_t R^v_{t,T} + R^f_{t,T},$$

where $R^v_{t,T}$ denotes the realized excess return on the variance swap, defined as

$$R^v_{t,T} := R^f_{t,T} \left(\frac{\mathrm{RV}^2_{t,T} - (\mathrm{RVP}^{\mathbb{Q}}_t)^2}{(\mathrm{RVP}^{\mathbb{Q}}_t)^2} \right).$$

This framework enables us to evaluate how effectively RVP-based forecasts translate into trading gains.

We estimate the forecasted standard deviation of variance swap returns using a rolling five-year window of past monthly returns, following Campbell and Thompson (2008). To avoid extreme positions, the portfolio weight w_t is constrained to lie between -1.5 to 1.5. Unlike previous studies (e.g., Campbell and Thompson, 2008; Huang et al., 2015), we allow both long and short positions in variance swaps. As Carr and Wu (2009) demonstrate, variance swaps tend to carry negative risk premia on average, rendering short positions essential for achieving consistently profitable trading outcomes. To assess the sensitivity of performance to risk aversion, we consider γ of 1, 3, and 5, in line with Huang et al. (2015).

This trading performance exercise also serves as a comparison between the predictive performance of each RVP and the risk-neutral RVP. When a forecaster's RVP exceeds the risk-neutral RVP, the investor takes a long position in the variance swap ($w_t > 0$); if the forecast is accurate ($R^v_{t,T} > 0$), the investor earns a positive excess portfolio return ($w_t R^v_{t,T} > 0$). Conversely, when a forecaster's RVP is below the risk-neutral RVP, the investor takes a short position in the variance swap ($w_t < 0$), and a correct forecast ($R^v_{t,T} < 0$) again yields a positive excess return. Hence, a forecaster whose RVP consistently outperforms the risk-neutral benchmark should generate a positive Sharpe ratio for the corresponding trading strategy. We calculate the Sharpe ratio as the sample mean of excess portfolio returns divided by their sample standard deviation, using observations from January 2013 through August 2023. The preceding window—from January 2007 to December 2012—is reserved for estimating the forecasted standard deviation of variance swap returns.

The CER of the trading strategy is defined as

$$\mathrm{CER} := \mu_p - \frac{1}{2}\gamma\sigma^2_p,$$

where μ_p and σ^2_p denote the sample mean and variance of portfolio returns over the evaluation period, respectively. We measure the CER gain as the difference between the CER obtained using an RVP-based forecast and the CER obtained using the historical average of realized variance as the forecast. The historical benchmark is calculated using a rolling five-year window of past monthly realized variance observations.

Table 10 summarizes the variance swap trading results, reporting both Sharpe ratios and CER gains on an annualized basis. All Sharpe ratios are positive, including those obtained using the historical average of realized variance as the forecast. Across all levels of risk aversion, strategies based on RVPs consistently yield higher Sharpe ratios than the historical benchmark. For example, with $\gamma = 3$ and the CYL RVP, the Sharpe ratio exceeds the benchmark by 0.15 to 0.36 points across the five trading horizons. Moreover, with exception of the BGX RVP, Sharpe ratios tend to increase monotonically with the degree of risk aversion, suggesting enhanced performance for more risk-averse investors.

Table 10 also shows that while many RVP-based strategies generate positive CER gains, some result in CER losses. Across all trading horizons and levels of risk aversion, strategies based on the log-utility and CYL RVPs consistently deliver positive CER gains and perform as well as or better than those using other RVPs. This result indicates that the RVPs of these two forecasters offer superior predictive power for guiding variance swap trading. With the exception of the BGX forecaster, CER gains generally increase with the investor's risk aversion coefficient, implying that more risk-averse investors benefit more from accurate volatility forecasts. For instance, under $\gamma = 1$, the CYL-based CER gains range from 0.3 to 6.1, whereas under $\gamma = 5$, they increase substantially, ranging from 1.5 to 94.4. Notably, high levels of risk aversion combined with short trading horizons tends to yield particularly large CER gains.

For comparison, we conduct a parallel trading performance exercise, in which a mean-variance investor allocates capital between the S&P 500 index and the risk-free asset, using the ERP as a predictor for future realized returns on the S&P 500. The results are reported in Table 11. All Sharpe ratios presented in the table are positive, and their magnitudes generally exceed those observed in Table 10 for variance swap trading.

However, the Sharpe ratios achieved using ERP-based forecasts differ only marginally from those obtained using the historical average return, suggesting that the ERP adds limited predictive value for market returns. In line with this observation, the CER gains in Table 11 are uniformly small and never exceed 0.03. These values are more than an order of magnitude lower than those documented in Table 10, where RVP-based variance swap strategies frequently yield CER gains of several points or more. Taken together, these results suggest that ERP-based forecasts fail to meaningfully improve investor utility relative to a simple historical average benchmark. In contrast, RVP-based forecasts applied to variance swap trading produce substantial enhancements in both Sharpe ratios and CERs. The economic implication is that while the predictability of equity returns via ERPs appears limited, volatility predictability—when appropriately captured through RVPs—provides a more robust foundation for active trading strategies.

5 Conclusion

This paper investigates which type of forecaster, differentiated by risk preferences, exhibits superior predictive power for the realized volatility of the S&P 500 index. We evaluate five forecasters: the log-utility forecaster following Martin (2017), CRRA forecasters with relative risk aversion coefficients of 2 and 3, the CYL forecaster of Chabi-Yo and Loudis (2020), and the BGX forecaster of Bakshi et al. (2023). The risk-neutral forecast, corresponding to the VIX index, serves as a benchmark. Our empirical analysis spans five forecasting horizons: 1, 2, 3, 4, and 6 months.

The results show that the CYL forecaster consistently performs at least as well as, and often better than, the other forecasters. This forecaster exhibits a relative risk aversion of 1, identical to that of the log-utility forecaster, and a relative prudence of 4, matching that of the CRRA3 forecaster. Expectation hypothesis regressions indicate that the CYL forecaster forms rational expectations of future realized volatility at all horizons except for the 1-month case. Forecasting performance test reveals that the CYL forecaster, along with the CRRA2 forecaster, delivers statistically significant predictive performance across all horizons. In the evaluation of variance swap trading strategies, forecasts generated by the CYL and log-utility forecasters consistently enhance trading performance, as measured by both the Sharpe ratio and gains in certainty equivalent return, across all trading horizons.

These findings underscore the value of incorporating risk preferences into volatility forecasting. In particular, the CYL forecaster offers a robust and practically effective alternative to traditional forecasting approaches. Moreover, other forecasters also demonstrate predictive advantages over the risk-neutral benchmark in most cases, further validating the relevance of risk-preference-based perspectives. Importantly, forecasts that reflect investor preferences produce unbiased predictions of realized volatility—meaning that their RVPs do not require ex-post bias correction. By contrast, the risk-neutral forecast—namely, the VIX-like index—exhibits a systematic bias.

We also explore the role of price jump risk in volatility forecasting by augmenting the PLS regression models to include skewness and kurtosis of S&P 500 returns as proxies for jump risk. While the in-sample regression results suggest that these higher-order moments may provide incremental predictive value, the out-of-sample performance of the augmented models deteriorates markedly, likely due to overfitting. This outcome reinforces the practical usefulness of raw RVPs as parsimonious and robust predictors of realized volatility.

Future research could pursue at least two important extensions. First, extending the analysis beyond the U.S. equity market to include major international stock indices, such as the Nikkei 225, FTSE 100, or EURO STOXX 50, would allow for testing the robustness of risk-preference-based forecasts across different market structures and volatility regimes. Such an investigation would help determine whether the relative advantages of risk-preference-based forecasts persist in the presence of variations in trading volume, institutional participation, and country specific market conditions.

Second, comparing risk-preference-based forecasts with those generated by established time-series models, such as GARCH, HAR, or stochastic volatility models, would clarify the relative strengths and limitations of preference-sensitive versus purely statistical approaches. By pursuing these directions, future work can deepen our understanding of the interplay between investor heterogeneity and realized volatility dynamics and potentially guide the development of more accurate and practically useful forecasting methods.

A Derivations of (2.5) and (2.6)

Consider the time-t price of a contingent claim that pays $f_t(S_T)/m_t(S_T)$ at maturity T. Under the physical measure $\mathbb{P}$, the price can be expressed as

$$\mathbb{E}_t^{\mathbb{P}}\left[m_{t,T}\frac{f_t(S_T)}{m_t(S_T)}\right] = \mathbb{E}_t^{\mathbb{P}}\left[\mathbb{E}_t^{\mathbb{P}}\left[m_{t,T}\frac{f_t(S_T)}{m_t(S_T)} \mid S_T\right]\right] = \mathbb{E}_t^{\mathbb{P}}\left[f_t(S_T)\right]. \tag{A.1}$$

This follows from the law of iterated expectations. On the other hand, under a risk-neutral measure, denoted as $\mathbb{Q}$, the price is given by

$$
\begin{aligned}
\frac{1}{R^f_{t,T}}\mathbb{E}^{\mathbb{Q}}_t\left[\frac{f_t(S_T)}{m_t(S_T)}\right] = {} & \frac{1}{R^f_{t,T}}\left(\frac{f_t(F)}{m_t(F)}\right) + \int_0^F \left(\frac{f_t(K)}{m_t(K)}\right)'' P_t(K)dK \\
& + \int_F^\infty \left(\frac{f_t(K)}{m_t(K)}\right)'' C_t(K)dK,
\end{aligned} \tag{A.2}
$$

where $\mathbb{E}^{\mathbb{Q}}_t[\,\cdot\,]$ denotes the time-t conditional expectation operator under $\mathbb{Q}$. This expression utilizes the static replication formula from Carr and Madan (1998). Since (A.1) and (A.2) represent the same price, it follows that the expression in (2.5) holds. By setting $f_t(S_T) = 1$ and $m_t = c_t n_t(S_T)$, we obtain (2.6). □

B Supplementary Notes on CYL Forecaster

B.1 Risk Aversion, Prudence, and Temperance

Chabi-Yo and Loudis (2020) define the following preference parameters:

$$
\tau := \frac{1}{\mathcal{A}}, \qquad \rho := \frac{1}{2}\frac{\mathcal{P}}{\mathcal{A}}, \qquad \kappa := \frac{1}{2}\frac{\mathcal{P}}{\mathcal{A}}\frac{1}{3}\frac{\mathcal{T}}{\mathcal{A}}.
$$

The parameter τ represents the risk tolerance, while ρ and κ are referred to as skewness and kurtosis tolerance, respectively. To derive a parameter-free lower bound for the equity premium, Chabi-Yo and Loudis (2020) impose the following restrictions on the preference parameters:

$$
\frac{1}{\tau} = 1, \qquad \frac{(1-\rho)}{\tau^2} = -1, \quad \text{and} \quad \frac{(1-2\rho+\kappa)}{\tau^3} = 1.
$$

These restrictions imply that the CYL forecaster exhibits relative risk aversion of $\mathcal{A} = 1$, relative prudence of $\mathcal{P} = 4$, and relative temperance of $\mathcal{T} = 6$.

B.2 Risk-Neutral Moment

The n-th oder risk-neutral moment of the excess price return on the stock is defined as

$$
M^{\mathbb{Q}}_n := \mathbb{E}^{\mathbb{Q}}_t\left[\left(\frac{S_T}{S_t} - R^f_{t,T}\right)^n\right],
$$

and can be expressed as

$$
\begin{aligned}
M^{\mathbb{Q}}_n = {} & \frac{n(n-1)R^f_{t,T}}{S_t^2}\left[\int_0^{R^f_{t,T}S_t}\left(\frac{K}{S_t} - R^f_{t,T}\right)^{n-2} P_t(K)dK\right. \\
& \left. + \int_{R^f_{t,T}S_t}^\infty \left(\frac{K}{S_t} - R^f_{t,T}\right)^{n-2} C_t(K)dK\right].
\end{aligned} \tag{B.1}
$$

A detailed derivation of (B.1) is provided in Appendix B.2 of Chabi-Yo and Loudis (2020).

B.3 Physical Moment

According to Chabi-Yo and Loudis (2020), the n-th order physical moment of the excess price return on the stock, defined as

$$M_n^{\mathbb{P}} := \mathbb{E}_t^{\mathbb{P}}\left[\left(\frac{S_T}{S_t} - R_{t,T}^f\right)^n\right],$$

can be approximated by

$$M_n^{\mathbb{P}} \approx M_n^{\mathbb{Q}} + BR_n,$$

where BR_n is the restricted bound as derived in Chabi-Yo and Loudis (2020), and is given by

$$BR_n := \frac{\sum_{k=1}^{3}(-R_{t,T}^f)^{-k}\left(M_n^{\mathbb{Q}} M_k^{\mathbb{Q}} - M_{n+k}^{\mathbb{Q}}\right)}{1 - \sum_{k=1}^{3}(-R_{t,T}^f)^{-k} M_k^{\mathbb{Q}}}. \tag{B.2}$$

To obtain the skewness and kurtosis of the net price return on the stock for the PLS regression analysis in Section 4.3, physical moments of the net price return up to the fourth order are required. Using the binomial theorem, the n-th order physical moment is given by

$$m_n^{\mathbb{P}} = \mathbb{E}_t^{\mathbb{P}}\left[\left(\frac{S_T}{S_t} - 1\right)^n\right] = \sum_{k=0}^{n}\binom{n}{k}\left(R_{t,T}^f - 1\right)^k M_{n-k}^{\mathbb{P}}.$$

C Supplementary Notes on BGX Forecaster

C.1 Pricing Formula for Volatility Contract

Bakshi et al. (2023) provide the model-free pricing formula for the volatility contract with payoff $\{\log(S_T/S_t)\}^2$ as follows:

$$\begin{aligned} v_t &:= \frac{1}{R_{t,T}^f}\mathbb{E}_t^{\mathbb{Q}}\left[\left\{\log\frac{S_T}{S_t}\right\}^2\right] \\ &= 2\int_0^{S_t}\frac{1-\log\frac{K}{S_t}}{K^2}P_t(K)dK + 2\int_{S_t}^{\infty}\frac{1-\log\frac{K}{S_t}}{K^2}C_t(K)dK. \end{aligned}$$

C.2 Parameter Estimation

The original SDF projection in Bakshi et al. (2023) is not $m_t(S_T)$ as defined in (2.8), but rather $n_t(S_T)$. Following their methodology, we estimate the parameters η_0 and η_1 associated with the original SDF $n_t(S_T)$ by solving the following relative entropy minimization problem:

$$\inf_{(\eta_0,\eta_1)} \mathbb{E}^{\mathbb{P}}\left[n_t(S_T)\log n_t(S_T)\right], \tag{C.1}$$

subject to the constraints

$$\mathbb{E}^{\mathbb{P}}\left[n_t(S_T)\right] = \mathbb{E}^{\mathbb{P}}\left[1/R_{t,T}^f\right] \quad \text{and} \quad \mathbb{E}^{\mathbb{P}}\left[n_t(S_T)Z_{t,T}\right] = 0.$$

We implement the Lagrangian dual problem of (C.1), given by

$$\inf_{(\eta_0,\eta_1)} -\mathbb{E}^{\mathbb{P}}\left[1/R^f_{t,T}\right]\eta_0 + \mathbb{E}^{\mathbb{P}}\left[\exp\left(\eta_0 - 1 + \eta_1 Z_{t,T}\right)\right]. \tag{C.2}$$

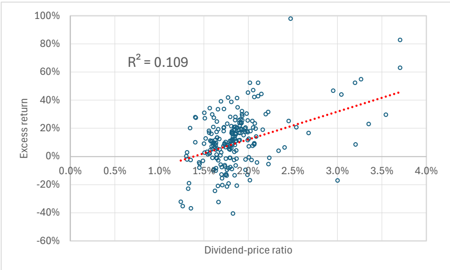
To obtain (η_0, η_1), we approximate the unconditional expectations in (C.2) using historical averages.

We use monthly market data from January 2007 through August 2023 to estimate the parameters. Table 12 presents the estimates of (η_0, η_1), obtained from expanding-window samples. To assess the statistical significance of these estimates, we employ a stationary bootstrap procedure, resampling from the historical data of $(R^f_{t,T}, Z_{t,T})$. We generate 1,000 bootstrap replications and construct 90% confidence intervals for (η_0, η_1) based on the empirical percentiles of the bootstrap distribution. The table shows that most of the parameter estimates for η_1 are statistically significantly positive at the 10% level, suggesting that the BGX forecaster exhibits aversion to volatility risk. An exception is observed for the 1-month horizon over the sample periods from 2007 to 2018, where the estimates are not statistically significant. For the full-sample analysis, we use the estimates obtained from the 2007–2023 period, whereas the estimates based on expanding-window samples are employed for the out-of-sample analysis.

Figure 1: Market Forecasting

The figure displays naïve predictive regressions for forecasting future S&P 500 index dynamics, based on monthly data from October 2003 to January 2023 and presented in annualized terms. Panel A shows the result from regressing excess returns over the subsequent 180 days on the dividend-price ratio. Panel B shows the result from regressing realized volatility over the subsequent 30 days on the VIX index.

Panel A: Excess Return



Panel B: Realize Volatility

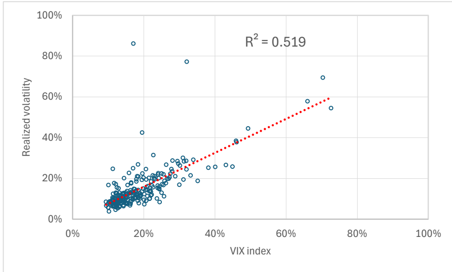
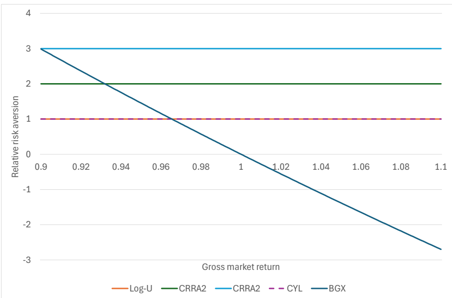


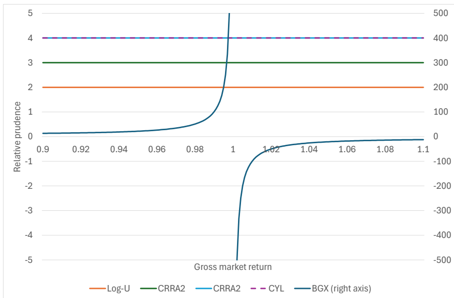
Figure 2: Risk Preferences of Forecasters

The figure displays the levels of risk preference for five forecasters (Log-U, CRRA2, CRRA3, CYL, and BGX) with respect to gross market returns. Panels A, B, and C present relative risk aversion, relative prudence, and relative temperance, respectively, as defined in (2.9).

Panel A: Relative Risk Aversion



Panel B: Relative Prudence



Panel C: Relative Temperance

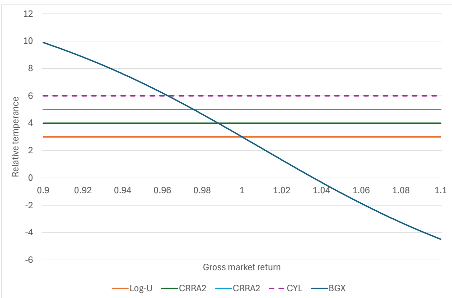


Figure 3: Time Series of RVP and ERP

The left panels of the figure display the time series of RPVs for six forecasters (RN, Log-U, CRRA2, CRRA3, CYL, and BGX). The right panels present the time series of ERPs for the five forecasters, excluding RN. The data consist of daily observations from January 2007 to August 2023.

Left Panels: RVP

Right Panels: ERP

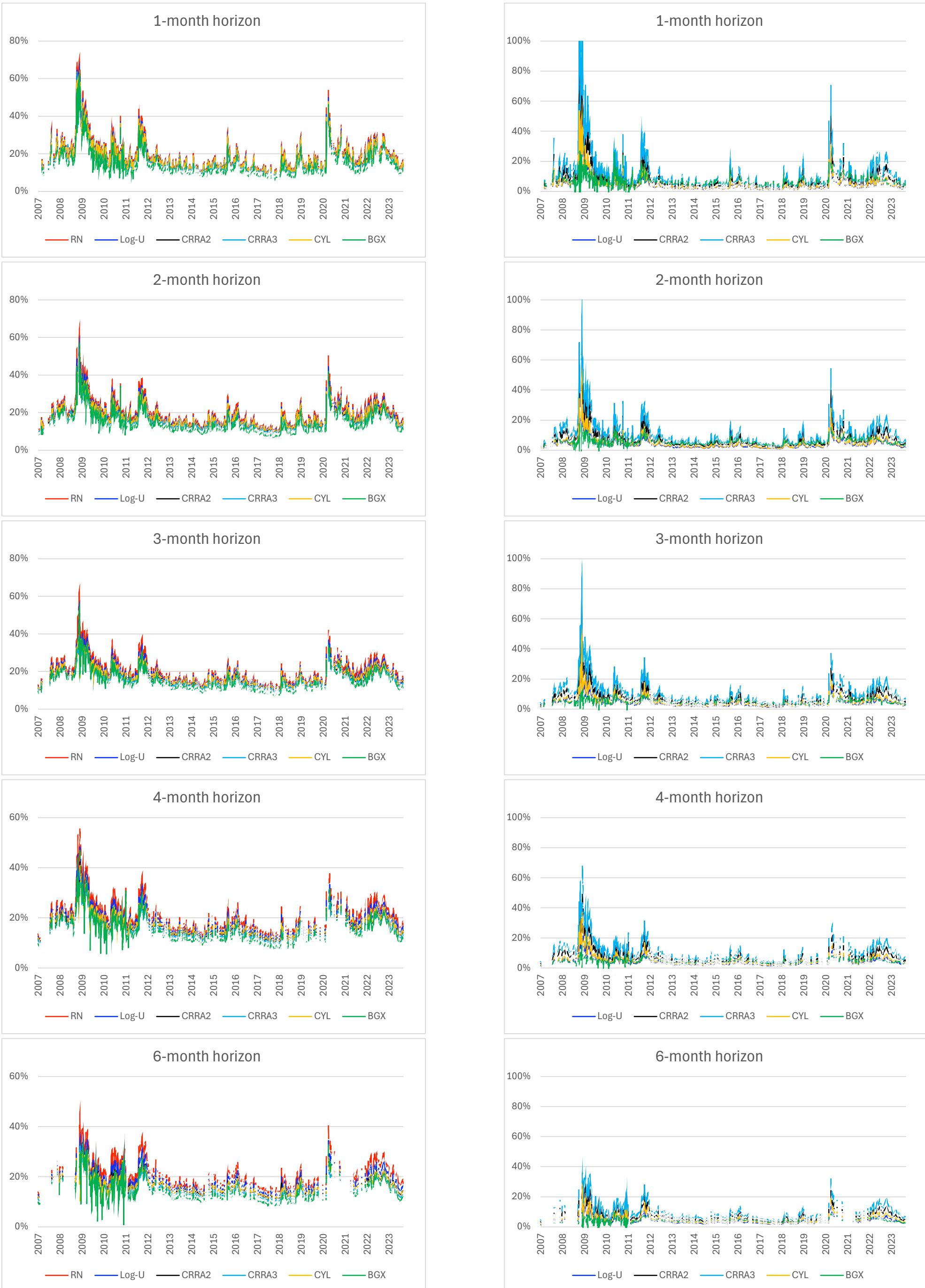


Figure 4: Slope Estimates of Expectation Hypothesis Regression

The figure displays the slope estimates from the expectation hypothesis regressions based on monthly data. The left panels show the slope estimates from regressions of realized volatility on the RPVs of six forecasters (RN, Log-U, CRRA2, CRRA3, CYL, and BGX), as well as on historical volatility (Hist). The right panels show the slope estimates from regressions of excess returns on the ERPs of the five forecasters, excluding RN, as well as on historical excess returns (Hist). Dots represent point estimates, and error bars indicate 95% confidence intervals.

Left Panels: Realized Volatility

Right Panels: Excess Return

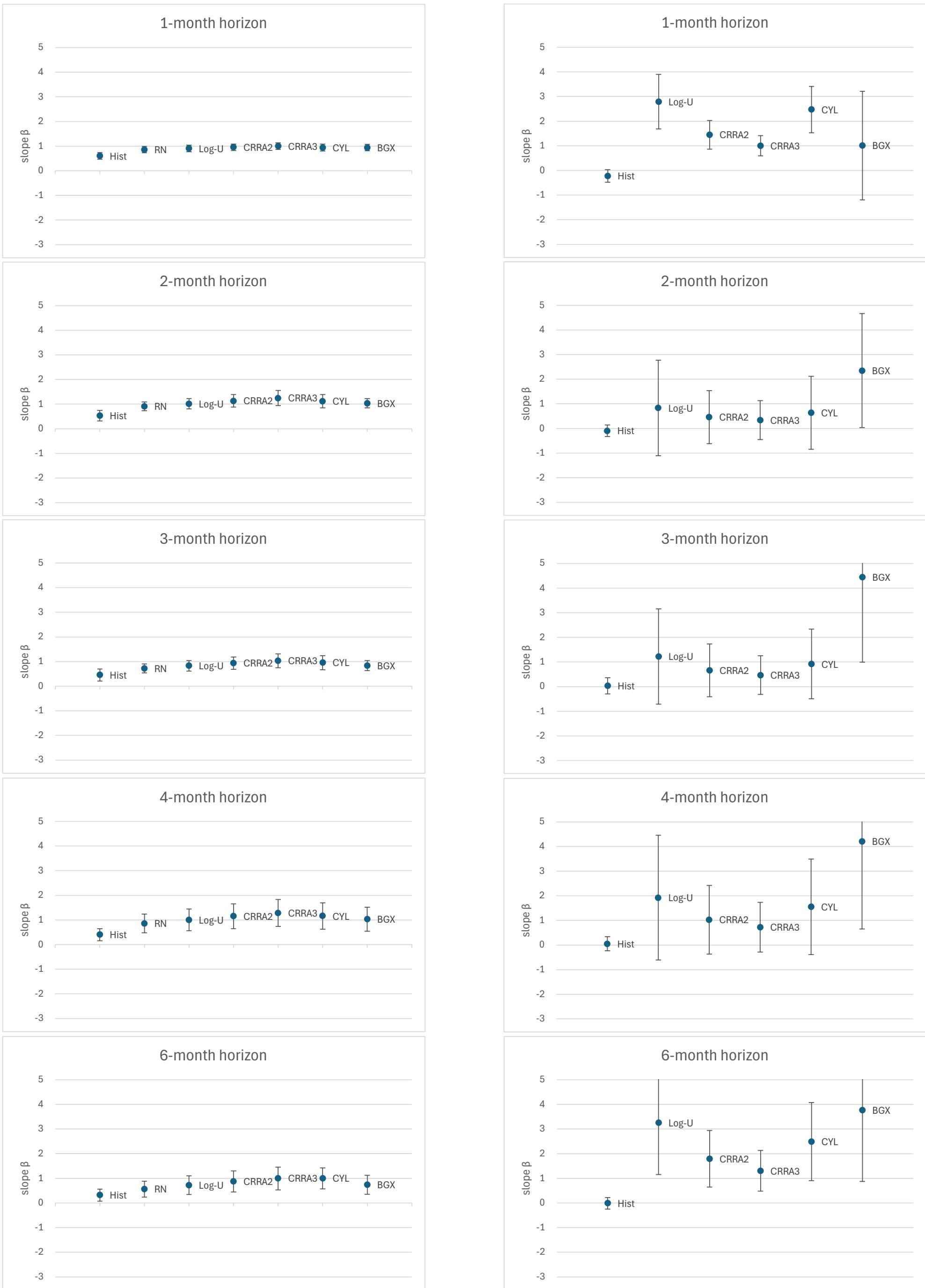


Table 1: Summary Statistics for Realized Volatility Predictors (RVPs)

The table presents summary statistics for the RPVs of six forecasters (RN, Log-U, CRRA2, CRRA3, CYL, and BGX) and the realized volatility (Realized) of the S&P 500 index, based on daily observations from January 2007 to August 2023. Means, standard deviations, and percentiles are annualized and expressed as percentages.

					Percentile				
	Mean	Stdv	Skew	Kurt	P10	P25	P50	P75	P90
Panel A: 1-month horizon									
Realized	17.33	3.38	2.96	15.23	7.80	9.95	14.31	20.65	28.75
RN	19.97	2.44	2.10	9.31	12.36	14.12	17.74	23.03	29.82
Log-U	19.12	2.29	2.05	9.08	11.93	13.59	17.05	22.13	28.38
CRRA2	18.36	2.15	2.03	9.01	11.53	13.14	16.37	21.30	27.08
CRRA3	17.69	2.05	2.03	9.20	11.17	12.72	15.80	20.55	25.89
CYL	18.81	2.19	1.96	8.47	11.82	13.46	16.84	21.82	27.81
BGX	15.16	2.14	2.19	9.85	8.63	10.10	13.00	17.94	23.79
Panel B: 2-month horizon									
Realized	16.56	4.17	3.04	15.89	8.44	10.48	13.58	19.84	26.15
RN	19.53	2.68	1.65	7.31	13.13	14.70	17.87	22.77	27.70
Log-U	18.30	2.44	1.56	6.82	12.40	13.88	16.75	21.36	25.79
CRRA2	17.24	2.24	1.52	6.63	11.75	13.16	15.81	20.09	24.15
CRRA3	16.33	2.09	1.53	6.84	11.22	12.53	15.02	19.02	22.72
CYL	17.75	2.27	1.41	5.94	12.12	13.59	16.33	20.72	24.86
BGX	14.52	2.34	1.70	7.61	9.01	10.38	12.92	17.24	21.69
Panel C: 3-month horizon									
Realized	17.56	5.09	2.65	12.17	9.05	11.19	14.32	20.57	26.89
RN	20.49	3.22	1.52	6.97	13.97	15.61	19.07	23.81	28.37
Log-U	18.88	2.86	1.42	6.37	13.01	14.49	17.59	21.97	25.87
CRRA2	17.52	2.58	1.36	5.98	12.18	13.56	16.37	20.39	23.71
CRRA3	16.40	2.37	1.34	5.93	11.51	12.77	15.34	19.11	22.06
CYL	17.96	2.55	1.21	5.18	12.58	13.97	16.86	20.95	24.19
BGX	15.23	2.83	1.61	7.39	9.64	11.00	13.83	18.27	21.90
Panel D: 4-month horizon									
Realized	17.57	5.84	2.29	9.39	8.96	11.43	14.49	20.81	26.80
RN	21.01	3.74	1.47	6.11	14.44	16.15	19.65	24.27	28.93
Log-U	19.09	3.26	1.37	5.64	13.29	14.79	17.88	22.09	25.97
CRRA2	17.51	2.90	1.30	5.31	12.30	13.66	16.47	20.35	23.52
CRRA3	16.27	2.64	1.26	5.14	11.46	12.75	15.36	18.95	21.66
CYL	17.79	2.80	1.15	4.73	12.63	14.02	16.86	20.73	23.77
BGX	15.41	3.30	1.52	6.35	9.76	11.15	14.05	18.53	22.13
Panel E: 6-month horizon									
Realized	16.26	5.11	1.86	8.49	9.83	11.67	13.87	19.82	24.96
RN	20.82	3.82	1.12	4.20	15.26	16.64	19.58	23.79	28.80
Log-U	18.54	3.26	1.09	4.00	13.77	15.04	17.42	21.19	25.16
CRRA2	16.75	2.88	1.13	4.22	12.52	13.72	15.72	19.16	22.33
CRRA3	15.40	2.65	1.23	4.76	11.56	12.67	14.41	17.66	20.22
CYL	16.63	2.69	1.11	4.28	12.65	13.82	15.71	18.93	21.72
BGX	14.38	3.29	1.23	4.71	9.76	11.08	13.10	16.90	20.85

Table 2: Summary Statistics for Excess Return Predictors (ERPs)

The table reports summary statistics for the ERPs of five forecasters (Log-U, CRRA2, CRRA3, CYL, and BGX) and the realized excess returns (Realized) of the S&P 500 index, based on daily observations from January 2007 to August 2023. Means, standard deviations, and percentiles are annualized and expressed as percentages.

					Percentile				
	Mean	Stdv	Skew	Kurt	P10	P25	P50	P75	P90
Panel A: 1-month horizon									
Realized	6.99	17.57	-1.26	8.99	-67.76	-21.42	15.52	42.75	66.51
Log-U	4.45	1.37	4.08	25.88	1.46	1.90	2.99	5.05	8.34
CRRA2	8.54	2.61	4.10	26.20	2.82	3.66	5.76	9.73	15.89
CRRA3	12.35	3.77	4.14	26.87	4.10	5.32	8.33	14.13	23.05
CYL	4.92	1.58	4.27	28.06	1.58	2.06	3.26	5.50	9.23
BGX	7.02	0.93	1.44	7.76	4.05	5.00	6.39	8.21	11.16
Panel B: 2-month horizon									
Realized	7.47	15.92	-1.11	7.65	-42.28	-8.22	12.69	31.25	45.30
Log-U	3.91	1.24	3.34	22.36	1.60	2.01	2.94	4.79	7.04
CRRA2	7.38	2.33	3.31	21.83	3.04	3.82	5.56	9.06	13.24
CRRA3	10.50	3.29	3.32	21.94	4.34	5.46	7.92	12.89	18.75
CYL	4.50	1.51	3.74	28.36	1.81	2.27	3.37	5.47	8.16
BGX	5.55	0.83	0.98	6.45	3.46	4.18	5.20	6.51	8.08
Panel C: 3-month horizon									
Realized	7.53	15.89	-1.06	6.69	-33.50	-6.20	12.97	27.00	39.13
Log-U	4.17	1.49	3.18	21.41	1.78	2.22	3.28	5.14	7.19
CRRA2	7.77	2.74	3.13	20.77	3.35	4.15	6.13	9.61	13.34
CRRA3	10.93	3.83	3.11	20.51	4.73	5.88	8.65	13.52	18.70
CYL	4.97	1.89	3.68	28.40	2.07	2.59	3.88	6.07	8.69
BGX	4.91	0.79	0.87	5.17	3.27	3.83	4.63	5.73	6.95
Panel D: 4-month horizon									
Realized	7.81	15.76	-1.04	5.79	-26.47	-4.83	11.97	23.53	37.65
Log-U	4.30	1.72	2.67	13.45	1.88	2.34	3.44	5.27	7.31
CRRA2	7.94	3.14	2.63	13.15	3.50	4.34	6.36	9.77	13.51
CRRA3	11.09	4.35	2.61	12.92	4.89	6.08	8.92	13.67	18.63
CYL	5.28	2.25	2.98	16.20	2.23	2.80	4.16	6.40	9.18
BGX	4.50	0.83	0.94	5.50	3.07	3.52	4.23	5.22	6.50
Panel E: 6-month horizon									
Realized	11.33	13.38	-0.12	5.59	-11.89	1.63	12.35	21.30	32.40
Log-U	4.03	1.58	1.78	6.85	2.05	2.45	3.33	4.96	7.11
CRRA2	7.35	2.87	1.80	6.95	3.76	4.49	6.04	9.00	12.78
CRRA3	10.16	3.98	1.85	7.26	5.21	6.24	8.36	12.50	17.60
CYL	5.14	2.13	1.96	8.42	2.54	3.03	4.20	6.29	9.36
BGX	4.11	1.01	0.04	6.13	2.85	3.31	3.91	4.77	6.03

Table 3: Expectation Hypothesis Regression of Realized Volatility

The table reports the estimated intercepts (α) and slopes (β) from the expectation hypothesis regressions, where the realized volatility of the S&P 500 index is regressed on the RPVs. The estimations are based on daily data with overlapping forecasting horizons. Parentheses report t-statistics for testing the null hypotheses that $\alpha = 0$ and $\beta = 1$, while square brackets report t-statistics for testing the null hypothesis that $\beta = 0$. All t-statistics are computed using the method of Newey and West (1987), with the number of lags set to 1.5 times the number of days in the forecasting horizon. Values of the ordinary R^2 are expressed as percentages.

	1-month horizon			2-month horizon			3-month horizon			4-month horizon			6-month horizon		
	α	β	R^2	α	β	R^2	α	β	R^2	α	β	R^2	α	β	R^2
Historical	0.06	0.65	42.4	0.08	0.53	28.6	0.10	0.45	20.4	0.11	0.40	16.0	0.12	0.32	10.0
	(4.43)	(-3.91)		(4.66)	(-4.70)		(4.49)	(-4.74)		(4.58)	(-4.88)		(4.94)	(-5.52)	
		[7.28]			[5.39]			[3.91]			[3.26]			[2.55]	
RN	-0.03	1.01	53.2	-0.02	0.94	36.7	-0.01	0.87	30.0	-0.01	0.88	31.7	0.03	0.63	22.1
	(-1.48)	(0.09)		(-0.87)	(-0.44)		(-0.22)	(-0.97)		(-0.32)	(-0.74)		(0.92)	(-2.40)	
		[8.90]			[7.39]			[6.25]			[5.43]			[4.07]	
Log-U	-0.03	1.08	53.2	-0.03	1.04	37.1	-0.01	0.99	30.8	-0.02	1.02	32.4	0.02	0.75	23.2
	(-1.65)	(0.64)		(-1.10)	(0.29)		(-0.54)	(-0.09)		(-0.60)	(0.10)		(0.64)	(-1.36)	
		[8.81]			[7.20]			[6.00]			[5.21]			[4.19]	
CRRA2	-0.04	1.14	53.0	-0.03	1.14	37.4	-0.02	1.11	31.4	-0.03	1.16	33.1	0.02	0.87	24.0
	(-1.80)	(1.10)		(-1.28)	(0.86)		(-0.77)	(0.56)		(-0.81)	(0.70)		(0.47)	(-0.65)	
		[8.79]			[7.08]			[5.81]			[5.05]			[4.34]	
CRRA3	-0.04	1.20	52.8	-0.03	1.22	37.5	-0.03	1.21	31.8	-0.03	1.28	33.7	0.02	0.96	24.5
	(-1.91)	(1.48)		(-1.41)	(1.29)		(-0.90)	(1.00)		(-0.93)	(1.10)		(0.43)	(-0.21)	
		[8.87]			[7.06]			[5.71]			[4.97]			[4.49]	
CYL	-0.04	1.12	52.9	-0.03	1.12	37.0	-0.03	1.10	30.6	-0.03	1.17	31.5	0.01	0.90	22.5
	(-1.81)	(0.94)		(-1.31)	(0.72)		(-0.86)	(0.52)		(-0.88)	(0.69)		(0.35)	(-0.49)	
		[8.58]			[6.80]			[5.50]			[4.70]			[4.35]	
BGX	0.00	1.14	51.8	0.01	1.08	36.9	0.02	1.01	31.2	0.02	1.00	31.9	0.06	0.70	20.1
	(0.06)	(1.03)		(0.46)	(0.54)		(0.94)	(0.03)		(0.85)	(0.00)		(2.13)	(-1.68)	
		[8.56]			[7.08]			[6.21]			[5.06]			[3.86]	

Table 4: Expectation Hypothesis Regression of Excess Return

The table reports the estimated intercepts (α) and slopes (β) from the expectation hypothesis regressions, in which the excess returns of the S&P 500 index are regressed on the ERPs. The estimations are based on daily data with overlapping forecasting horizons. Parentheses report t-statistics for testing the null hypotheses that $\alpha = 0$ and $\beta = 1$, while square brackets report t-statistics for testing the null hypothesis that $\beta = 0$. All t-statistics are computed using the method of Newey and West (1987), with the number of lags set to 1.5 times the number of days in the forecasting horizon. Values of the ordinary R^2 are expressed as percentages.

	1-month horizon			2-month horizon			3-month horizon			4-month horizon			6-month horizon		
	α	β	R^2	α	β	R^2	α	β	R^2	α	β	R^2	α	β	R^2
Historical	0.08	-0.07	0.5	0.08	-0.07	0.6	0.07	0.07	0.5	0.06	0.10	1.0	0.07	0.02	0.0
	(1.97)	(-13.59)		(1.88)	(-11.00)		(1.57)	(-6.41)		(1.51)	(-6.58)		(1.62)	(-6.88)	
		[-0.88]			[-0.76]			[0.52]			[0.72]			[0.11]	
Log-U	0.05	0.52	0.2	0.04	0.96	0.6	0.04	0.91	0.7	0.05	0.76	0.7	0.00	2.91	11.8
	(1.02)	(-0.38)		(0.72)	(-0.02)		(0.79)	(-0.06)		(1.02)	(-0.16)		(-0.10)	(2.08)	
		[0.41]			[0.59]			[0.61]			[0.51]			[3.17]	
CRRA2	0.05	0.28	0.2	0.04	0.50	0.5	0.04	0.47	0.7	0.05	0.42	0.7	-0.01	1.62	12.0
	(1.00)	(-1.08)		(0.72)	(-0.57)		(0.80)	(-0.68)		(0.99)	(-0.72)		(-0.14)	(1.22)	
		[0.41]			[0.57]			[0.58]			[0.51]			[3.21]	
CRRA3	0.05	0.19	0.2	0.04	0.35	0.5	0.04	0.33	0.6	0.05	0.30	0.7	-0.01	1.17	12.0
	(0.99)	(-1.72)		(0.72)	(-1.05)		(0.81)	(-1.15)		(0.98)	(-1.18)		(-0.13)	(0.46)	
		[0.41]			[0.56]			[0.55]			[0.50]			[3.23]	
CYL	0.05	0.42	0.1	0.04	0.82	0.6	0.04	0.73	0.8	0.05	0.55	0.6	0.00	2.12	11.3
	(1.13)	(-0.54)		(0.77)	(-0.13)		(0.88)	(-0.24)		(1.18)	(-0.41)		(0.12)	(1.66)	
		[0.40]			[0.62]			[0.64]			[0.50]			[3.14]	
BGX	0.00	0.93	0.2	-0.06	2.51	1.7	-0.12	3.98	3.9	-0.04	2.72	2.0	0.00	2.74	4.2
	(0.07)	(-0.06)		(-0.83)	(1.05)		(-1.44)	(1.78)		(-0.54)	(0.85)		(0.02)	(1.48)	
		[0.86]			[1.74]			[2.38]			[1.34]			[2.33]	

Table 5: Forecasting Performance for Realized Volatility

The table reports the results of the forecasting performance test for realized volatility of the S&P 500 index. The full-sample period spans from January 2007 to August 2023, while the out-of-sample period covers January 2017 to August 2023. R^2_{OS} denotes the out-of-sample R^2 statistic of Campbell and Thompson (2008), expressed as percentages. Statistical significance is assessed using the MSPE-based test, with asterisks (*, **, ***) indicating the 10%, 5%, and 1% levels, respectively.

	Full-sample	Out-of-sample	
Forecasted volatility	$\hat{\mathrm{RV}}_{t,T} = \mathrm{RVP}_t$	$\hat{\mathrm{RV}}_{t,T} = \alpha_t + \beta_t \mathrm{RVP}_t$	$\hat{\mathrm{RV}}_{t,T} = \mathrm{RVP}_t$
	R^2_{OS}	R^2_{OS}	R^2_{OS}
Panel A: 1-month horizon			
RN	48.4**	32.9***	26.8***
Log-U	50.8**	32.8***	30.5***
CRRA2	51.7***	32.7***	32.7***
CRRA3	51.4***	32.6***	33.6***
CYL	50.9**	32.6***	33.6***
BGX		33.0***	31.4***
Panel B: 2-month horizon			
RN	29.7*	25.2***	13.7
Log-U	35.5**	25.3***	22.5**
CRRA2	37.7***	25.3***	26.8***
CRRA3	37.5***	25.3***	28.0***
CYL	36.6**	25.1***	25.1***
BGX		25.7***	24.5**
Panel C: 3-month horizon			
RN	20.8	23.6***	6.1
Log-U	29.7*	23.2***	19.5
CRRA2	32.5**	22.6***	25.2**
CRRA3	31.7**	22.1***	26.0**
CYL	31.3**	22.1***	23.7***
BGX		23.3***	22.4
Panel D: 4-month horizon			
RN	21.8	32.7***	11.8
Log-U	32.0*	32.0***	28.5*
CRRA2	34.2**	30.9**	34.0*
CRRA3	32.2**	29.7**	33.0*
CYL	32.6**	30.5**	33.2*
BGX		31.1**	30.1
Panel E: 6-month horizon			
RN	−14.9	21.6***	−8.3
Log-U	18.0	21.0***	18.4
CRRA2	29.1**	19.8***	23.7**
CRRA3	29.2*	18.1**	18.4
CYL	28.2**	19.9***	24.2**
BGX		18.6***	13.7

Table 6: Forecasting Performance for Excess Return

The table reports the results of the forecasting performance test for excess returns on the S&P 500 index. The full-sample period spans from January 2007 to August 2023, while the out-of-sample period covers January 2017 to August 2023. R^2_{OS} denotes the out-of-sample R^2 statistic of Campbell and Thompson (2008), expressed as percentages. Statistical significance is assessed using the MSPE-based test, with asterisks (*, **, ***) indicating the 10%, 5%, and 1% levels, respectively.

	Full-sample	Out-of-sample	
Forecasted volatility	$\hat{\text{ER}}_{t,T} = \text{ERP}_t$	$\hat{\text{ER}}_{t,T} = \alpha_t + \beta_t \text{ERP}_t$	$\hat{\text{ER}}_{t,T} = \text{ERP}_t$
	R^2_{OS}	R^2_{OS}	R^2_{OS}
Panel A: 1-month horizon			
Log-U	0.6	0.3^{**}	1.6
CRRA2	−0.3	0.3^{**}	3.3^{**}
CRRA3	−2.9	0.3^{**}	4.1^{**}
CYL	0.5	0.3^{**}	1.8
BGX		−0.1	1.3^{**}
Panel B: 2-month horizon			
Log-U	1.2	0.9	1.3
CRRA2	1.5	0.8	5.0^{*}
CRRA3	−0.4	0.8	6.4
CYL	1.5	1.0	2.2
BGX		1.2	1.5
Panel C: 3-month horizon			
Log-U	2.4	1.2	1.0
CRRA2	2.6	1.1	5.5
CRRA3	−0.3	1.1	6.6
CYL	2.8	1.1	2.2
BGX		−3.2	0.5
Panel D: 4-month horizon			
Log-U	3.9	−0.3	−1.4
CRRA2	4.2	−0.2	2.0
CRRA3	0.5	−0.1	1.0
CYL	4.2	−0.5	−0.3
BGX		−16.2	−2.1
Panel E: 6-month horizon			
Log-U	11.5^{*}	−3.0	5.5
CRRA2	22.8^{**}	−1.9	10.2
CRRA3	27.4^{**}	−0.9	9.0
CYL	16.1^{**}	−3.6	7.7
BGX		−15.9	3.9

Table 7: Predictor Loadings of PLS Components

The table reports the estimated predictor loadings of the PLS components for RVP, SKW, and KRT. PctVar1 denotes the percentage of variance in the predictor set (RVP, SKW, and KRT) explained by each PLS component, while PctVar2 refers to the percentage of variance in the dependent variable (realized volatility) explained by each component. The table presents results for each forecaster at the 1- and 2-month horizons. Results for other forecasting horizons are available in the Internet Appendix.

	Component	RVP	SKW	KRT	PctVar1	PctVar2
Panel A: 1-month horizon						
RN	1st	49.4	50.1	-46.7	0.79	0.41
	2nd	23.9	-16.3	25.6	0.16	0.13
	3rd	3.1	-15.9	-13.7	0.05	0.00
Log-U	1st	49.8	49.3	-45.3	0.77	0.41
	2nd	23.0	-15.7	26.5	0.16	0.13
	3rd	3.6	-18.7	-16.5	0.07	0.00
CRRA2	1st	50.3	48.2	-43.6	0.74	0.41
	2nd	22.0	-14.9	27.2	0.16	0.13
	3rd	4.0	-22.0	-19.7	0.10	0.00
CRRA3	1st	50.8	46.8	-41.8	0.72	0.41
	2nd	20.8	-14.3	27.2	0.15	0.13
	3rd	4.1	-25.1	-23.2	0.13	0.00
CYL	1st	50.2	49.8	-45.2	0.78	0.41
	2nd	22.0	-13.0	26.7	0.15	0.12
	3rd	4.6	-19.5	-16.3	0.07	0.00
BGX	1st	54.7	11.8	-6.6	0.35	0.51
	2nd	4.9	-44.4	48.3	0.48	0.01
	3rd	-3.4	30.4	25.5	0.17	0.00
Panel B: 2-month horizon						
RN	1st	56.8	59.1	-57.8	0.85	0.31
	2nd	26.4	-14.3	21.3	0.11	0.06
	3rd	-4.3	15.6	11.7	0.03	0.00
Log-U	1st	57.0	58.5	-57.2	0.84	0.31
	2nd	26.0	-14.5	21.7	0.11	0.06
	3rd	-4.0	17.5	13.9	0.04	0.00
CRRA2	1st	57.3	57.8	-56.4	0.83	0.31
	2nd	25.3	-14.3	22.0	0.11	0.07
	3rd	-3.7	19.8	16.6	0.06	0.00
CRRA3	1st	57.7	56.9	-55.3	0.82	0.31
	2nd	24.5	-14.0	22.0	0.11	0.07
	3rd	2.9	-22.3	-19.9	0.08	0.00
CYL	1st	57.6	59.0	-57.3	0.85	0.32
	2nd	24.2	-11.0	21.9	0.10	0.06
	3rd	-5.7	-18.4	13.2	0.05	0.00
BGX	1st	58.1	36.5	-33.8	0.50	0.33
	2nd	22.5	-35.0	45.2	0.32	0.05
	3rd	-7.4	37.1	27.4	0.18	0.00

Table 8: In-Sample PLS Regression of Realized Volatility

The table reports the estimated coefficients β_1 and β_2 from the PLS regressions, in which the realized volatility of the S&P 500 index is regressed on the scores of the first two PLS components derived from RVP, SKW, and Kurt. The estimations are based on daily data with overlapping forecasting horizons. Parentheses report t-statistics computed using the method of Newey and West (1987), with the number of lags set to 1.5 times the number of days in the forecasting horizon. Adjusted R^2 values are expressed as percentages.

	1-month horizon			2-month horizon			3-month horizon			4-month horizon			6-month horizon		
	β_1	β_2	R^2	β_1	β_2	R^2	β_1	β_2	R^2	β_1	β_2	R^2	β_1	β_2	R^2
RN	4.15	2.30	53.4	3.62	1.57	37.3	3.13	1.17	31.6	2.91	1.12	33.2	1.66	0.83	24.4
	(8.92)	(7.49)		(6.41)	(5.88)		(5.08)	(3.87)		(4.39)	(3.94)		(4.07)	(3.35)	
Log-U	4.13	2.32	53.4	3.61	1.62	37.5	3.14	1.19	31.8	2.92	1.13	33.5	1.68	0.84	25.1
	(8.84)	(7.50)		(6.26)	(6.17)		(4.91)	(3.95)		(4.31)	(4.16)		(4.23)	(3.31)	
CRRA2	4.13	2.31	53.3	3.60	1.66	37.5	3.14	1.17	31.8	2.95	1.02	33.4	1.72	0.75	25.1
	(8.85)	(7.48)		(6.17)	(6.45)		(4.82)	(3.96)		(4.24)	(3.64)		(4.29)	(2.77)	
CRRA3	4.14	2.30	53.2	3.58	1.69	37.6	3.14	1.15	31.7	2.98	0.98	33.5	1.75	0.65	24.8
	(8.95)	(7.51)		(6.18)	(6.75)		(4.81)	(3.96)		(4.28)	(3.55)		(4.29)	(2.23)	
CYL	4.14	2.29	53.2	3.63	1.57	37.6	3.19	1.10	32.3	3.01	0.92	33.8	1.78	0.70	26.2
	(8.71)	(7.11)		(6.14)	(5.76)		(4.88)	(4.16)		(4.22)	(3.57)		(4.65)	(3.32)	
BGX	4.65	0.62	52.2	3.69	1.47	37.8	3.21	1.16	32.9	3.01	0.73	32.7	1.71	0.11	21.0
	(8.50)	(0.62)		(6.08)	(7.68)		(4.95)	(4.16)		(4.84)	(3.90)		(3.98)	(1.09)	

Table 9: Out-of-sample Forecasting Performance of PLS Regression

The table reports the results of the out-of-sample forecasting test for the PLS regression model in (4.4), where the realized volatility of the S&P 500 index is regressed on the scores of the first two components derived from RVP, SKW, and KRT. The out-of-sample period spans from January 2017 to August 2023. R^2_{OS} denotes the out-of-sample R^2 statistic of Campbell and Thompson (2008), expressed as percentages. Although we conduct the MSPE-based test to assess the statistical significance of R^2_{OS}, none of the values are statistically significantly greater than zero.

Horizon	1-month	2-month	3-month	4-month	6-month
	R^2_{OS}	R^2_{OS}	R^2_{OS}	R^2_{OS}	R^2_{OS}
RN	-0.2	-7.4	-14.5	-8.5	-14.1
Log-U	0.5	-7.2	-16.1	-12.9	-15.7
CRRA2	1.4	-6.7	-16.7	-16.2	-16.7
CRRA3	2.5	-6.1	-16.1	-15.4	-16.5
CYL	0.5	-6.8	-15.9	-16.2	-17.2
BGX	-0.1	-5.8	-13.5	-7.0	-12.5

Table 10: Trading Performance of Variance Swap

The table reports performance measures of variance swap trading for a mean-variance investor with risk aversion coefficients $\gamma = 1$, 3, and 5. The investor allocates capital between a variance swap and the risk-free asset, using the squared RVP as a predictor of future realized variance. The portfolio weight on the variance swap is constrained to lie between -1.5 and 1.5. In this table, SR denotes the Sharpe ratio, computed as the sample mean of excess portfolio returns divided by their sample standard deviation. ΔCER denotes the CER gain, defined as the difference between the certainty equivalent return (CER) based on the RVP-based forecast and that based on the historical average (Historical) of realized variance. The sample period spans from January 2013 to August 2023. Both measures are reported in annualized terms.

	$\gamma = 1$		$\gamma = 3$		$\gamma = 5$	
	SR	ΔCER	SR	ΔCER	SR	ΔCER
Panel A: 1-month horizon						
Historical	0.02		0.08		0.15	
Log-U	0.17	15.3	0.23	58.3	0.24	46.3
CRRA2	0.19	-1.9	0.21	39.2	0.23	34.6
CRRA3	0.22	-3.5	0.18	14.0	0.22	19.2
CYL	0.16	6.1	0.23	53.7	0.24	43.5
BGX	0.23	-4.4	0.21	-54.3	0.17	-99.8
Panel B: 2-month horizon						
Historical	0.11		0.09		0.06	
Log-U	0.21	1.2	0.25	53.5	0.25	100.6
CRRA2	0.23	0.9	0.21	22.9	0.25	81.6
CRRA3	0.23	0.9	0.21	4.6	0.23	57.5
CYL	0.23	0.9	0.25	43.3	0.27	94.4
BGX	0.23	0.9	0.23	0.6	0.20	2.3
Panel C: 3-month horizon						
Historical	0.16		0.16		0.18	
Log-U	0.36	3.6	0.56	22.6	0.60	31.8
CRRA2	0.29	0.9	0.45	15.4	0.54	25.2
CRRA3	0.29	0.7	0.39	10.1	0.48	18.9
CYL	0.32	1.9	0.52	18.9	0.61	29.0
BGX	0.29	0.7	0.30	1.8	0.38	8.1
Panel D: 4-month horizon						
Historical	0.32		0.32		0.32	
Log-U	0.50	1.1	0.62	6.9	0.66	11.2
CRRA2	0.49	0.6	0.56	4.2	0.62	7.9
CRRA3	0.49	0.6	0.52	2.3	0.59	5.8
CYL	0.50	0.8	0.58	5.0	0.64	8.8
BGX	0.49	0.6	0.49	0.8	0.52	2.4
Panel E: 6-month horizon						
Historical	0.43		0.46		0.53	
Log-U	0.62	0.5	0.77	2.1	0.83	2.3
CRRA2	0.61	0.3	0.68	1.1	0.76	1.3
CRRA3	0.61	0.3	0.66	0.6	0.71	0.5
CYL	0.60	0.3	0.69	1.2	0.78	1.5
BGX	0.61	0.4	0.62	0.0	0.66	-0.3

Table 11: Trading Performance of Stock Index

The table reports performance measures of stock market index trading for a mean-variance investor with risk aversion coefficients $\gamma = 1$, 3, and 5. The investor allocates capital between the S&P 500 index and the risk-free asset, using the ERP as a predictor of future realized returns on the index. The portfolio weight on the stock index is constrained to lie between -1.5 and 1.5. In this table, SR denotes the Sharpe ratio, computed as the sample mean of excess portfolio returns divided by their sample standard deviation. ΔCER denotes the CER gain, defined as the difference between the certainty equivalent return (CER) based on the ERP-based forecast and that based on the historical average (Historical) of realized returns. The sample period spans from January 2013 to August 2023. Both measures are reported in annualized terms.

	$\gamma = 1$		$\gamma = 3$		$\gamma = 5$	
	SR	ΔCER	SR	ΔCER	SR	ΔCER
Panel A: 1-month horizon						
Historical	0.56		0.41		0.35	
Log-U	0.58	0.00	0.61	0.02	0.57	0.02
CRRA2	0.61	0.01	0.63	0.04	0.62	0.03
CRRA3	0.63	0.02	0.59	0.03	0.63	0.03
CYL	0.58	0.00	0.62	0.02	0.58	0.02
BGX	0.63	0.01	0.56	0.03	0.54	0.02
Panel B: 2-month horizon						
Historical	0.72		0.61		0.54	
Log-U	0.68	-0.01	0.66	0.00	0.62	0.00
CRRA2	0.75	0.01	0.68	0.01	0.66	0.01
CRRA3	0.78	0.01	0.67	0.01	0.68	0.02
CYL	0.69	-0.01	0.66	0.00	0.63	0.01
BGX	0.76	0.01	0.68	0.01	0.65	0.01
Panel C: 3-month horizon						
Historical	0.73		0.62		0.58	
Log-U	0.70	-0.01	0.67	-0.01	0.64	0.00
CRRA2	0.77	0.01	0.68	0.01	0.68	0.01
CRRA3	0.80	0.01	0.69	0.01	0.68	0.01
CYL	0.72	0.00	0.67	0.00	0.64	0.00
BGX	0.76	0.01	0.67	0.00	0.65	0.01
Panel D: 4-month horizon						
Historical	0.78		0.68		0.66	
Log-U	0.76	-0.01	0.65	-0.01	0.60	-0.01
CRRA2	0.83	0.01	0.70	0.00	0.66	0.00
CRRA3	0.87	0.02	0.74	0.01	0.68	0.00
CYL	0.78	0.00	0.64	-0.01	0.61	-0.01
BGX	0.81	0.00	0.70	0.00	0.67	0.00
Panel E: 6-month horizon						
Historical	0.80		0.66		0.62	
Log-U	0.77	-0.01	0.61	-0.01	0.57	-0.01
CRRA2	0.86	0.01	0.69	0.00	0.62	0.00
CRRA3	0.91	0.02	0.74	0.01	0.67	0.01
CYL	0.80	0.00	0.63	-0.01	0.58	0.00
BGX	0.81	0.00	0.65	0.00	0.61	0.00

Table 12: Parameter Estimates for BGX

The table reports the estimated parameters η_0 and η_1 of the BGX SDF in (2.8), obtained by solving the optimization problem in (C.2) based on monthly data of expanding window samples. It also presents 90% confidence intervals for the estimates, constructed from the empirical percentiles of the bootstrap distribution based on 1,000 bootstrap replications.

Expanding samples		η_0	Bootstrap CIs		η_1	Bootstrap CIs	
Start	End	Estimate	Lower	Upper	Estimate	Lower	Upper
Panel A: 1-month horizon							
2007	2016	1.02	1.00	1.14	0.10	-0.05	0.49
2007	2017	1.02	1.00	1.14	0.11	-0.03	0.53
2007	2018	1.02	1.00	1.11	0.12	-0.03	0.46
2007	2019	1.02	1.00	1.15	0.13	0.00	0.55
2007	2020	1.03	1.00	1.15	0.13	0.00	0.53
2007	2021	1.03	1.00	1.17	0.14	0.00	0.56
2007	2022	1.03	1.00	1.15	0.15	0.01	0.53
2007	2023	1.03	1.00	1.17	0.15	0.02	0.56
Panel B: 2-month horizon							
2007	2016	1.03	1.00	1.14	0.12	0.02	0.40
2007	2017	1.03	1.00	1.16	0.13	0.02	0.47
2007	2018	1.03	1.00	1.13	0.13	0.02	0.43
2007	2019	1.03	1.00	1.10	0.14	0.03	0.25
2007	2020	1.04	1.00	1.15	0.15	0.05	0.49
2007	2021	1.04	1.00	1.16	0.16	0.04	0.49
2007	2022	1.04	1.01	1.13	0.16	0.05	0.42
2007	2023	1.04	1.00	1.11	0.17	0.05	0.27
Panel C: 3-month horizon							
2007	2016	1.04	1.00	1.15	0.14	0.01	0.40
2007	2017	1.04	1.00	1.27	0.15	0.01	0.44
2007	2018	1.03	1.00	1.13	0.14	0.02	0.38
2007	2019	1.04	1.00	1.14	0.15	0.03	0.41
2007	2020	1.04	1.00	1.11	0.15	0.03	0.34
2007	2021	1.04	1.00	1.13	0.16	0.04	0.37
2007	2022	1.04	1.00	1.13	0.17	0.05	0.39
2007	2023	1.04	1.00	1.13	0.17	0.04	0.37
Panel D: 4-month horizon							
2007	2016	1.03	1.00	1.11	0.13	0.01	0.30
2007	2017	1.03	1.00	1.12	0.14	0.01	0.33
2007	2018	1.04	1.00	1.12	0.15	0.03	0.33
2007	2019	1.04	1.00	1.11	0.16	0.02	0.33
2007	2020	1.04	1.00	1.12	0.16	0.04	0.34
2007	2021	1.04	1.00	1.12	0.17	0.04	0.34
2007	2022	1.04	1.00	1.12	0.18	0.05	0.35
2007	2023	1.05	1.00	1.14	0.18	0.06	0.37
Panel E: 6-month horizon							
2007	2016	1.06	1.00	1.81	0.19	0.03	2.30
2007	2017	1.06	1.00	1.72	0.19	0.02	2.20
2007	2018	1.06	1.00	1.80	0.21	0.05	2.28
2007	2019	1.07	1.00	1.78	0.22	0.04	2.26
2007	2020	1.07	1.00	1.79	0.22	0.06	2.33
2007	2021	1.07	1.01	1.79	0.22	0.06	2.30
2007	2022	1.08	1.01	1.87	0.23	0.07	2.48
2007	2023	1.08	1.01	1.85	0.24	0.08	2.45

References

[1] Adhikari, Binay and Hilliard, Jimmy. The VIX, VXO and realised volatility: A test of lagged and contemporaneous relationships. *International Journal of Financial Markets and Derivatives*, 3(3):222–240, 2014.

[2] Aït-Sahalia, Yacine and Lo, Andrew. Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance*, 53(2):499–547, 1998.

[3] Andersen, Torben., Bollerslev, Tim., Diebold, Francis., and Ebens, Heiko. The distribution of realized stock return volatility. *Journal of Financial Economics*, 61(1):43–76, 2001.

[4] Andersen, Torben., Bollerslev, Tim., Diebold, Francis., and Labys, Paul. Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625, 2003.

[5] Back, Kerry., Crotty, Kevin., and Kazempour, Seyed Mohammad. Validity, tightness, and forecasting power of risk premium bounds. *Journal of Financial Economics*, 144(3):732–760, 2022.

[6] Bakshi, Gurdip., Gao, Xiaohui., and Xue, Jinming. Recovery with applications to forecasting equity disaster probability and testing the spanning hypothesis in the treasury market. *Journal of Financial and Quantitative Analysis*, 58(4):1808–1842, 2023.

[7] Bakshi, Gurdip., Kapadia, Nikunj., and Madan, Dilip. Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies*, 16(1):101–143, 2003.

[8] Black, Fischer. Studies of stock price volatility changes. In *Proceedings from the American Statistical Association, Business and Economic Statistics Section*, page 177, 1976.

[9] Bliss, Robert R and Panigirtzoglou, Nikolaos. Option-implied risk aversion estimates. *Journal of Finance*, 59(1):407–446, 2004.

[10] Bollerslev, Tim. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

[11] Bucci, Andrea. Realized volatility forecasting with neural networks. *Journal of Financial Econometrics*, 18(3):502–531, 2020.

[12] Campbell, John and Thompson, Samuel. Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies*, 21(4):1509–1531, 2008.

[13] Caporin, Massimiliano. The role of jumps in realized volatility modeling and forecasting. *Journal of Financial Econometrics*, 21(4):1143–1168, 2023.

[14] Carr, Peter and Madan, Dilip. Towards a theory of volatility trading. In Robert. Jarrow, editor, *Volatility*, pages 417–427. Risk Publications, 1998.

[15] Carr, Peter and Wu, Liuren. Variance risk premiums. *Review of Financial Studies*, 22(3):1311–1341, 2009.

[16] Chabi-Yo, Fousseni and Loudis, Johnathan. The conditional expected market return. *Journal of Financial Economics*, 137(3):752–786, 2020.

[17] Christensen, Bent and Hansen, Charlotte. New evidence on the implied-realized volatility relation. *European Journal of Finance*, 8(2):187–205, 2002.

[18] Christensen, Bent and Prabhala, Nagpurnanand. The relation between implied and realized volatility. *Journal of Financial Economics*, 50(2):125–150, 1998.

[19] Christoffersen, Peter., Heston, Steven., and Jacobs, Kris. Capturing option anomalies with a variance-dependent pricing kernel. *Review of Financial Studies*, 26(8):1963–2006, 2013.

[20] Clark, Todd and West, Kenneth. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291–311, 2007.

[21] Corsi, Fulvio. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2):174–196, 2009.

[22] Cumby, Robert., Figlewski, Stephen., and Hasbrouck, Joel. Forecasting volatilities and correlations with EGARCH models. *Journal of Derivatives*, 1(2):51–63, 1993.

[23] Day, Theodore and Lewis, Craig. Stock market volatility and the information content of stock index options. *Journal of Econometrics*, 52(1-2):267–287, 1992.

[24] Engle, Robert. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987–1007, 1982.

[25] Ferreira, Miguel and Santa-Clara, Pedro. Forecasting stock market returns: The sum of the parts is more than the whole. *Journal of Financial Economics*, 100(3):514–537, 2011.

[26] Ghosh, Anisha., Julliard, Christian., and Taylor, Alex. What is the consumption-CAPM missing? An information-theoretic framework for the analysis of asset pricing models. *Review of Financial Studies*, 30(2):442–504, 2017.

[27] Ghysels, Eric., Santa-Clara, Pedro., and Valkanov, Rossen. Predicting volatility: Getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 131(1-2):59–95, 2006.

[28] Huang, Dashan., Jiang, Fuwei., Tu, Jun., and Zhou, Guofu. Investor sentiment aligned: A powerful predictor of stock returns. *Review of Financial Studies*, 28(3):791–837, 2015.

[29] Jorion, Philippe. Predicting volatility in the foreign exchange market. *Journal of Finance*, 50(2):507–528, 1995.

[30] Kadan, Ohad and Tang, Xiao. A bound on expected stock returns. *Review of Financial Studies*, 33(4):1565–1617, 2020.

[31] Kambouroudis, Dimos., McMillan, David., and Tsakou, Katerina. Forecasting stock return volatility: A comparison of GARCH, implied volatility, and realized volatility models. *Journal of Futures Markets*, 36(12):1127–1163, 2016.

[32] Kambouroudis, Dimos., McMillan, David., and Tsakou, Katerina. Forecasting realized volatility: The role of implied volatility, leverage effect, overnight returns, and volatility of realized volatility. *Journal of Futures Markets*, 41(10):1618–1639, 2021.

[33] Kandel, Shmuel and Stambaugh, Robert. On the predictability of stock returns: An asset-allocation perspective. *Journal of Finance*, 51(2):385–424, 1996.

[34] Kelly, Bryan and Pruitt, Seth. Market expectations in the cross-section of present values. *Journal of Finance*, 68(5):1721–1756, 2013.

[35] Kou, Steven. A jump-diffusion model for option pricing. *Management Science*, 48(8):1086–1101, 2002.

[36] Lamoureux, Christopher and Lastrapes, William. Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities. *Review of Financial Studies*, 6(2):293–326, 1993.

[37] Martin, Ian. What is the expected return on the market? *Quarterly Journal of Economics*, 132(1):367–433, 2017.

[38] Martin, Ian and Wagner, Christian. What is the expected return on a stock? *Journal of Finance*, 74(4):1887–1929, 2019.

[39] McAleer, Michael and Medeiros, Marcelo. A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries. *Journal of Econometrics*, 147(1):104–119, 2008.

[40] Merton, Robert. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1-2):125–144, 1976.

[41] Nelson, Daniel. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2):347–370, 1991.

[42] Newey, Whitney and West, Kenneth. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708, 1987.

[43] Pavlidis, Efthymios., Paya, Ivan., and Peel, David. Forecast evaluation of nonlinear models: The case of long-span real exchange rates. *Journal of Forecasting*, 31(7):580–595, 2012.

[44] Polkovnichenko, Valery and Zhao, Feng. Probability weighting functions implied in options prices. *Journal of Financial Economics*, 107(3):580–609, 2013.

[45] Rosenberg, Joshua and Engle, Robert. Empirical pricing kernels. *Journal of Financial Economics*, 64(3):341–372, 2002.

[46] Song, Zhaogang and Xiu, Dacheng. A tale of two option markets: Pricing kernels and volatility risk. *Journal of Econometrics*, 190(1):176–196, 2016.

[47] Stambaugh, Robert. Predictive regressions. *Journal of Financial Economics*, 54(3):375–421, 1999.

[48] Welch, Ivo and Goyal, Amit. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4):1455–1508, 2008.

[49] Wold, Herman. Estimation of principal components and related models by iterative least squares. *Multivariate Analysis*, pages 391–420, 1966.

[50] Wold, Herman. Path models with latent variables: The NIPALS approach. In *Quantitative Sociology*, pages 307–357. Elsevier, 1975.

[51] Yamazaki, Akira. Recovering subjective probability distributions. *Journal of Futures Markets*, 42(7):1234–1263, 2022.

[52] Yamazaki, Akira. Subjective probability distributions of nonlinear payoffs: Recovering option payoff, agent's utility, and pricing kernel distributions. *North American Journal of Economics and Finance*, 76:102362, 2025.

本ワーキングペーパーの掲載内容については、著編者が責任を負うものとします。

法政大学イノベーション・マネジメント研究センター
The Research Institute for Innovation Management, HOSEI UNIVERSITY

〒102-8160 東京都千代田区富士見 2-17-1
TEL: 03(3264)9420　FAX: 03(3264)4690
URL: https://riim.ws.hosei.ac.jp
E-mail: cbir@adm.hosei.ac.jp