# Return Predictability of Variance Differences: a fractionally co-integrated approach

**Abstract**

This paper examines the fractional co-integration between downside (upside) components of realized and implied variances. A positive association is found between the strength of their co-fractional relation and the return predictability of their differences. That association is established via the common long-memory component of the variances that are fractionally co-integrated, which represents the volatility-of-volatility factor that determines the variance premium. Our results indicate that market fears play a critical role not only in driving the long-run equilibrium relationship between implied-realized variances but also in understanding the return predictability. A simulation study further verifies these claims.
**Keywords**: Fractional co-integration; Variance risk premium; Return predictability.
**JEL Classification**: C51; C32; C14.

## 1 Introduction

Numerous empirical studies suggest that, unlike the variance, the variance risk premium ($VRP$) carries non-trivial predictive power for aggregate stock market returns over quarterly to annual horizons, where the degree of return predictability is greater than that afforded by more conventional predictors, see, Bollerslev, Tauchen, and Zhou (2009), Drechsler and Yaron (2010), Bollerslev et al. (2014), and Bali and Zhou (2016), among others. Those studies also provide strong evidence that much of the predictability implicit in the $VRP$ may be attributed to its close relation with macroeconomic uncertainty and aggregate risk aversion.

The $VRP$ is formally defined as the wedge between option-implied and realized variances. It is measured as the difference between (the square of) the CBOE VIX index and the statistical expectation of the future return variation. In Bollerslev, Tauchen, and Zhou (2009), Drechsler and Yaron (2010) and Bollerslev, Sizova, and Tauchen (2012), the return predictability of the $VRP$ is investigated using a self-contained general equilibrium model. This is in the spirit of the long-run risks (LRR) model pioneered by Bansal and Yaron (2004). Specifically, Bollerslev, Tauchen, and Zhou (2009) and Bollerslev, Sizova, and Tauchen (2012) extend the LRR model by allowing the time-varying volatility-of-volatility (vol-of-vol) within the economy to be generated by its own stochastic process. They further show that the difference between the risk-neutral and the objective expectations of return variation isolates the vol-of-vol factor, which then serves as the sole source of the true $VRP$. The $VRP$ therefore displays good predictive power for future returns in cases where the vol-of-vol plays a more dominant role in determining variation in the equity premium. A direct link between the $VRP$ and the vol-of-vol is also demonstrated in a purely probabilistic model introduced by Barndorff-Nielsen and Veraart (2013).

By incorporating jumps in the LRR model, Drechsler and Yaron (2010) disentangle the difference in the drifts of conditional variance from the $VRP$. That drift difference is related to the vol-of-vol associated with the level of uncertainty, leading to the covariation between the $VRP$ and the conditional equity premium that is partially dependent on the vol-of-vol factor. As indicated by Bollerslev, Tauchen, and Zhou (2009), the vol-of-vol is inherently latent and may be variously defined across different volatility models. Recent literature has achieved limited progress in seeking an appropriate proxy of the vol-of-vol factor which is necessary to explain the return predictability suggested by the $VRP$. A notable exception is the work of Conrad and Loch (2015) where the vol-of-vol is represented by the long-term volatility component of the GARCH-MIDAS model.

Consistent with the generalized LRR model discussed above, results suggested by Bollerslev et al. (2013) provide new evidence for return predictability of the $VRP$ based on the idea of fractional co-integration. Given the mounting empirical studies that favor fractional integration in

implied and realized variances (see Bandi and Perron (2006) and Nielsen (2007) for instance), fast mean reversion in the $VRP$, i.e. the difference between the two variances, points to the existence of fractional co-integration. Compared to the use of the univariate variances as a proxy for risk, Bollerslev et al. (2013) show that the less persistent $VRP$ rebalances the risk-return regression in terms of the order of fractional integration, and that this appears to be more informative for future returns. In addition, the $VRP$ is estimated as the co-integrating relation between the two variances, which is highly linked to the vol-of-vol and aggregate economic uncertainty.

Another strand of the literature on the properties of the $VRP$ focuses on the decomposition of the $VRP$ into its downside and upside components, so differentiating between investors' attitudes toward uncertainty risks associated with the left and right tail of the return distribution. Feunou, Jahan-Parvar, and Okou (2017) find that the downside $VRP$ often dominates total $VRP$ in predicting future excess returns. They rationalize this result by showing (i) that the two components exhibit opposite relationships to the equity premium and (ii) that the link between the upside $VRP$ and the equity premium is insignificant. In similar vein, Kilic and Shaliastovich (2018) adopt an alternative decomposition of the $VRP$: they argue that both good and bad components contain important information about future excess returns, which results in higher return predictability than that inherent in the total $VRP$.

The difference between the components of the realized (implied) variance is also considered in several studies where the main focus is on equity risk premium predictability. For instance, by decomposing the realized jump volatility into negative jumps and positive jumps, Guo, Wang, and Zhou (2019) construct the signed jump risk as the difference between the two jump risk components. In addition, the signed jump risk is found to contain information about future market returns that is independent of that captured by the $VRP$. With an alternative decomposition approach, Bollerslev, Li, and Zhao (2017) derive the signed jump variation as the difference between the good and bad realized volatility and show that this difference significantly predicts the variation in future returns. Under the risk-neutral measure, Huang and Li (2019) identify

a significant relationship between future stock returns and implied variance asymmetry defined as the difference between the upside and downside semi-variances derived from out-of-the-money (OTM) options.

Against this backdrop, the main contributions of this paper are threefold. First, by decomposing the implied and realized variances into upside and downside components, we investigate the presence of fractional co-integration in each pair of semi-variances and establish new evidence for the long-run relationship between different variance components. Second, we show that the common long-memory component in the fractionally co-integrated system, as part of the conditional variance of market returns, is intimately associated with the vol-of-vol driving the variance premium. Moreover, that common component plays a key role in connecting the strength of the fractional co-integration between variances and the stock return predictability afforded by the variance differences. The dominant predictive power of the downside $VRP$ documented in the literature can therefore be reasonably explained by the long-run equilibrium relation between the downside implied and realized semi-variances. Third, to alleviate the impact on our empirical findings of the limited availability of observed strikes in the construction of implied variances, we employ a simulation study to verify the claims outlined above.

Following the procedures of Barndorff-Nielsen, Kinnebrock, and Shephard (2010) and Andersen, Bondarenko, and Gonzalez-Perez (2015), we obtain upside and downside semi-variances of the realized ($RV$) and implied ($IV$) variances, which we refer to as $RV^U$, $RV^D$, $IV^U$ and $IV^D$. The difference between $IV^U$ $\left(IV^D\right)$ and $RV^U$ $\left(RV^D\right)$ is defined as the upside (downside) variance risk premium, henceforth $VRP^U$ $\left(VRP^D\right)$. The signed jump ($SJ$) and implied variance asymmetry ($IVA$) are respectively measured as the difference between $RV^U$ and $RV^D$ and the difference between $IV^U$ and $IV^D$. Using a semi-parametric approach, we demonstrate the existence of a fractionally co-integrating relation in each pair of the semi-variances listed above. In addition, our results reveal the role of $IV^D$ as a long-run upward biased forecast of future $RV^D$. This evidence is also consistent with the existing literature which documents that investors dislike fluctuations

in variances induced by the extreme downside returns and are willing to pay a premium to insure against increases in bad uncertainty.

To explain the association between fractional co-integration detected in pairs of variances and the return predictability implicit in their differences, we extract the common long-memory components in variances using the permanent-transitory decomposition proposed by Gonzalo and Granger (1995). This procedure is implemented within a parametric fractionally co-integrated VAR framework. We show that the common long-memory component represents the vol-of-vol factor that constitutes the unique source of the variance premium. As emphasized by Bollerslev, Tauchen, and Zhou (2009), the persistence of the vol-of-vol serves as a positive determinant of the long-horizon return predictability. Given that finding, an increase in the persistence of the common permanent component not only magnifies the co-fractional relation between variances but also results in better return predictions implied by their differences. The superior performance of the $VRP^D$, in terms of predictive power for future returns, can thus be explained by the stronger fractional co-integration between $IV^D$ and $RV^D$. However, the linkage between the fractional co-integration and return predictability does not hold for the $SJ$ and $IVA$ that affect variations in equity premium through a channel other than the vol-of-vol. Given that the $SJ$ and $IVA$ are found to be weakly associated with the long-run future excess returns, they are considered as relatively short-lived return predictors.

Based on a wide range and fine partition of strikes, the results of our simulation study are generally in line with the empirical findings of an association between the degree of co-fractional relation found in pairs of variances and the return predictability of their differences. With a list of option quotes for OTM calls and puts that are sufficiently deep in the tails, the $VRP$ and its components clearly outperform the $SJ$ and $IVA$ in predicting future returns across various horizons where the $SJ$ is notably the worst predictor.

The rest of this paper is organized as follows. Section 2 defines the implied and realized semi-variances together with the variance differences considered in the present study. The origin

of our data is introduced in Section 3 and the empirical results are reported in Section 4. Section 5 provides the results of a simulation study designed to alleviate the adverse implementation issues encountered in the empirical application. Section 6 concludes.

# 2 Measuring Variance Differences

This section introduces the methods of decomposing the implied and realized variances into their upside and downside components. These form the basis for the construction of the $VRP^U$, $VRP^D$, $SJ$ and $IVA$.

## 2.1 Realized Variance Components

We adopt the realized variance estimator proposed by Barndorff-Nielsen and Shephard (2002), which is equal to the sum of intraday squared returns

$$rv_t = \sum_{j=1}^{M} r_{t,j}^2 \tag{1}$$

where $r_{t,j}$ stands for intraday returns within each 5-minute interval. In addition, we include the squared overnight return, computed as the squared close-to-open logarithmic price change, to the $rv_t$ obtained over the trading day. We then follow Barndorff-Nielsen, Kinnebrock, and Shephard (2010) in deriving upside and downside semi-variances, which separately identify the contributions of positive and negative intraday price increments.

$$rv_t^D = \sum_{j=1}^{M} r_{t,j}^2 \mathbf{1}_{(r_{t,j} \leq 0)} \tag{2}$$

$$rv_t^U = \sum_{j=1}^{M} r_{t,j}^2 \mathbf{1}_{(r_{t,j} > 0)} \tag{3}$$

Given that $rv_t = rv_t^U + rv_t^D$, the upside and downside components provide a complete decomposition of the realized variance. Using a jump-diffusion process for the stock price, Barndorff-Nielsen,

Kinnebrock, and Shephard (2010) demonstrate that, as $M \to \infty$, $rv_t^U$ and $rv_t^D$ each converge separately to one-half of the integrated variance plus the sum of positive and negative quadratic jumps.

## 2.2 Implied Variance Components

For the construction of the risk-neutral expectation of the one-month forward return variation, we follow Bollerslev, Tauchen, and Zhou (2009) and Bollerslev et al. (2013) among others, by employing the CBOE $VIX$ index as a measure of the total implied volatility proxy for the US S&P 500 contract.

Based on the idea of fair value of future volatility developed by Demeterfi et al. (1999), the $VIX$ index is given as

$$VIX_t^{CBOE} = \sqrt{\frac{2}{\tau}\sum_i \frac{\Delta K_i}{K_i^2} e^{r\tau} Q(\tau, K_i) - \frac{1}{\tau}\left(\frac{F_0}{K_0} - 1\right)^2} \tag{4}$$

where $\tau = 30/365$ is the option maturity measured in annual units, $r$ is the annualized risk-free interest rate, $K_i$ is the strike price of the $i$th OTM option in the calculation[1], $K_0$ is the first strike price below the forward price $F_0$ at maturity $\tau$ ($K_0 \leq F_0$), $Q(\tau, K_i)$ is the midpoint of the latest available bid and ask prices of the OTM option at strike $K_i$, and $\Delta K_i$ stands for the strike price interval given as $\Delta K_i = (K_{i+1} - K_{i-1})/2$.

As shown by Jiang and Tian (2007), the $VIX$ is conceptually equivalent to the model-free implied volatility developed by Britten-Jones and Neuberger (2000) as follows

$$MFIV_t = \sqrt{\frac{2}{\tau} e^{r\tau} \left[\int_0^{F_t} \frac{P(t,\tau;K)}{K^2} dK + \int_{F_t}^{\infty} \frac{C(t,\tau;K)}{K^2} dK\right]} \tag{5}$$

[1] At the boundaries of strike prices, $\Delta K_i$ is adjusted as the difference between the two highest (or lowest) strike prices. In addition, at the strike price $K_0$ the option price $Q_i(\tau, K_i)$ is modified to be the average of call and put prices. The CBOE computes the $VIX$ from an interpolation of two volatility indices with respect to two different maturities: $\tau_t^l$ and $\tau_t^u$. The $VIX$ index is finally obtained by taking a weighted average of these two $VIX$ measures based on $\tau_t^l$ and $\tau_t^u$.

where $P(t,\tau;K)$ and $C(t,\tau;K)$ are the respective mid-quotes for European put and call options with strike price $K$ and maturity $\tau$. By construction, only OTM options (call if $K > F_t$ and put otherwise) are taken into account.

Following the methodology of Andersen, Bondarenko, and Gonzalez-Perez (2015), we respectively define the upside and downside implied volatilities by

$$iv_t^U = \sqrt{\frac{2}{\tau}e^{r\tau}\int_{F_t}^{\infty}\frac{C(t,\tau;K)}{K^2}dK} \tag{6}$$

$$iv_t^D = \sqrt{\frac{2}{\tau}e^{r\tau}\int_{0}^{F_t}\frac{P(t,\tau;K)}{K^2}dK} \tag{7}$$

Similar to the construction of the $VIX$ index, all the risk-neutral measures are derived from options across two nearest maturities (less than 30 days and greater than 30 days); and the 30-day implied volatilities are computed by interpolating between the two separate maturities. We then transform the annualized implied volatility measures into monthly variance units as follows

$$VIX_t^2 = \frac{30}{365}\left(VIX_t^{CBOE}\right)^2 \tag{8}$$

$$IV_t^U = \frac{30}{365}\left(iv_t^U\right)^2 \tag{9}$$

$$IV_t^D = \frac{30}{365}\left(iv_t^D\right)^2 \tag{10}$$

## 2.3 Construction of Variance Differences

Given that the $VIX_t^2$ and its components in equations (8 to 10) are based on one-month maturity option prices, we compute the $RV_t$ and its components over the remaining one month of the options as follows

$$RV_t = \frac{1}{\tau}\sum_{i=1}^{\tau} rv_{t+i} \tag{11}$$

$$RV_t^U = \frac{1}{\tau}\sum_{i=1}^{\tau} rv_{t+i}^U \tag{12}$$

$$RV_t^D = \frac{1}{\tau}\sum_{i=1}^{\tau} rv_{t+i}^D \tag{13}$$

where $rv_t$, $rv_t^U$ and $rv_t^D$ are daily measures derived from intraday returns and $\tau$ denotes the number of trading days during one month. To remove the impact of overlapping data on the results of fractional co-integration, throughout the present paper we follow Kellard, Dunis, and Sarantis (2010) by constructing a monthly dataset from the daily version. Specifically, we select the data for the implied variances only from the trading day following the final day employed in the computation of the previous realized variances.

Following Bollerslev, Tauchen, and Zhou (2009) and Bollerslev et al. (2014), among others, we define the $VRP$ and its components by

$$VRP_t = VIX_t^2 - RV_{t-1} \tag{14}$$

$$VRP_t^U = IV_t^U - RV_{t-1}^U \tag{15}$$

$$VRP_t^D = IV_t^D - RV_{t-1}^D \tag{16}$$

Such constructions have the advantages of being model-free and directly observable at time $t$. The forward-looking $VRP$ based on the statistical expectation of the $RV$ $(E_t\left(RV_{t,t+1}\right))$ is also considered in the literature, where $E_t\left(RV_{t,t+1}\right)$ is approximated using various parametric modelling methods designed to forecast the $RV$. As noted by Bollerslev, Tauchen, and Zhou (2009), under the assumption that the $RV$ follows a martingale difference or exhibits strong long-memory dynamics, the two measures of the $VRP$ are in line with each other.

The signed jump $(SJ)$ proposed by Bollerslev, Li, and Zhao (2017) is computed as the difference between the semi-variances $RV_t^U$ and $RV_t^D$, which eliminates the variation caused by

the continuous part but retains the variation arising from jumps[2].

$$SJ_t = RV^U_{t-1} - RV^D_{t-1} \tag{17}$$

Finally, we construct the implied variance asymmetry ($IVA$) introduced by Huang and Li (2019). This is the difference between upside and downside implied semi-variances as given by

$$IVA_t = IV^U_t - IV^D_t \tag{18}$$

# 3 Data

This section gives a description of our data and the standard summary statistics of the return and variance series considered in the paper. To construct $rv_t$ and its components for the period January 02, 2003 to December 30, 2016 with a total of 3526 observations, we use the five-minute high-frequency data of the aggregate S&P 500 (SPX) composite index obtained from Tick Data Inc.. Risk-neutral variance measures are derived from daily data of European put and call options on the SPX, as collected from Optionmetrics via the WRDS system. To represent risk-free rates, daily one-month and three-month Treasury-bill yields[3] are extracted from the Federal Reserve Bulletin.

To address the issue of bid-ask bounce, we compute the option price as the average of bid and ask quotes. Options with a time to maturity of fewer than seven days are filtered out, and those with Black–Scholes implied volatilities below zero or above 100% are also excluded. As discussed in Section 2.3, we convert the daily dataset into a monthly version to circumvent the problem of overlapping data. Thereby, the sample comprises 168 observations. Following Andersen, Bondarenko, and Gonzalez-Perez (2015) and Yao and Izzeldin (2018), we consider a

[2] In deriving the $SJ_t$, we consider the $RV^U_{t-1}$ and $RV^D_{t-1}$ as opposed to $RV^U_t$ and $RV^D_t$ in order to make the $SJ_t$ directly observable at time $t$.

[3] Following Jiang and Tian (2007), the risk-free rate is linearly interpolated between these two yields. However, when the maturity is shorter (longer) than one (three) month, the one-month (three-month) yield is adopted.

replicated $VIX$ index, $RX$, using the exact CBOE procedure every day. The $RX$ provides an equivalent of the $VIX$ using SPX option prices from the Optionmetrics data set[4].

Summary statistics for the excess return and variance series are reported in Table 1. The excess return is derived by subtracting the monthly risk-free rate from the SPX log-return. In accord with the existing literature, the implied and realized variances together with their components are fairly persistent as suggested by the AR(1) parameters. As persistence is largely reduced when the variance differences are considered, this is indicative of the presence of fractional co-integration between the variances. In addition, the features of the $VRP$ and its components are consistent with those documented in the literature: i) the total, downside and upside variance premium are on average positive; ii) compared with the $VRP$ and $VRP^D$, the $VRP^U$ exhibits a much lower mean value and more negative skewness; iii) the $VRP^D$ serves as the main component of the $VRP$.

# 4 Empirical Results

We begin this section by investigating the fractional co-integrating relations between the variances, prior to seeking in the fractionally co-integrated system the components that determine the variance premium. This is followed by an analysis of the risk-return relations across various frequencies and the pattern of return predictability afforded by variance differences.

## 4.1 Fractional Co-integrating Relations

Two series $y_t$ and $x_t$ are said to be fractionally co-integrated if both are integrated of order $d_x$ and a linear combination $u_t = y_t - \theta x_t$ is $I(d_u)$ where $d_u < d_x$. In contrast to conventional co-integration, $d_x$ can be a fractional number rather than an integer. In the next subsection, we examine the

[4]The CBOE calculates the $VIX$ index using option prices updated every five minutes. However, the Optionmetrics database includes the last daily bid-ask quote only, which might not correspond to the data published by CBOE for their final end-of-day computation. Constructed from the exact same dataset, all the risk-neutral variance measures considered in this paper are directly comparable to each other.

presence of fractional co-integration between the variances and provide information on the nature of long-run relations that may exist.

We first estimate the order of fractional integration $d$ by exploiting the exact local Whittle estimator of Shimotsu and Phillips (2005) where the bandwidth parameter is set as $m_1 = T^{0.5}$. This is commonly adopted in the literature[5], and for more details see Appendix A. The results, which are outlined in the first column of Panel A in Table 2, indicate that all the variance series are fractionally integrated. We then proceed to test for the equality of the fractional difference parameters, which is a necessary condition for the existence of fractional co-integration. However, as stressed by Robinson and Yajima (2002), the test for homogeneity of $\widehat{d}$ could deliver misleading conclusions if the co-integration is not accounted for. Hence, we first estimate the co-integrating rank for each pair of variances using the exact local Whittle rank test by Nielsen and Shimotsu (2007) as introduced in Appendix B.

Panel A of Table 2 presents the results of the co-integration rank test, where the rank $\widehat{r}$ is estimated by minimizing $L(u)$ in equation (31) of Appendix B. Given that $L(1) < L(0)$ for each pair of variances, we obtain $\widehat{r} = 1$ for all the scenarios considered. Moreover, a large value of the test statistic $\widehat{T}_0$ is evidence against the null hypothesis of the equality of $\widehat{d}$. In comparing the $\widehat{T}_0$ reported in Panel A with the 95% critical value of the $\chi^2(1)$ distribution (3.841), we accept the null and conclude that there exists a fractional co-integrating relation for $(RX_t^2,\, RV_t)$, $\left(IV_t^U,\, RV_t^U\right)$, $\left(IV_t^D,\, RV_t^D\right)$, $\left(RV_t^U,\, RV_t^D\right)$ and $\left(IV_t^U,\, IV_t^D\right)$. The average $\widehat{d}$ is taken as $\frac{1}{2}(\widehat{d}_{y_t} + \widehat{d}_{x_t})$ which we denote by Ave.$\widehat{d}$.

To further evaluate the long-run relationship between the variances, we consider the regression

$$y_t = c + \theta x_t + u_t \tag{19}$$

where $y_t$ and $x_t$ represent the pair of variances under analysis. Since the OLS fails to produce

[5]Other parameters $m_1 = T^{0.6}$ and $m_1 = T^{0.7}$ are also attempted, which leads to qualitatively similar results. To conserve space, we only report the results based on $m_1 = T^{0.5}$.

consistent estimates of the co-integrating vector for equation (19), where both the regressors and errors have long memory (Bandi and Perron (2006)), we adopt the fully modified narrow-band least squares (FMNBLS) proposed by Nielsen and Frederiksen (2011). This approach uses the local Whittle estimator to obtain the integration orders of $y_t$ and $x_t$ and then derives the estimator of the co-integrating relation before the integration order of the residuals is estimated. Consistent with the earlier discussion, we allow the bandwidth parameter to be $m_1 = T^{0.5}$ and then follow Nielsen and Frederiksen (2011) in setting values for the other parameters, $m_0$, $m_2$ and $m_3$. As the FMNBLS estimates of the fractional difference parameters $\widehat{d}$ are close to those presented in Panel A of Table 2, they are not reported.

As noted by Bandi and Perron (2006) and Nielsen and Frederiksen (2011), a negative intercept $c$ might point to the presence of a long-run risk premium and long-run unbiasedness is supported if $\theta = 1$ irrespective of the existence of the intercept estimate. Results in Panel B of Table 2 support the long-run unbiasedness hypothesis, $\theta = 1$, for $(RX_t^2,\, RV_t)$, $\left(IV_t^U,\, RV_t^U\right)$ and $\left(IV_t^U,\, IV_t^D\right)$. As suggested by the slope estimate for $\left(IV_t^D,\, RV_t^D\right)$, which is significantly below unity, in the long run the $IV_t^D$ tends to overestimate the $RV_t^D$. Our results provide evidence that investors dislike, and are willing to pay a premium to hedge against, downside (bad) variance risk, and that they respond more to downside than to upside (good) uncertainty. The point estimate $\widehat{\theta}$ for $\left(RV_t^U,\, RV_t^D\right)$, which is slightly in excess of unity, indicates that under the physical measure, upside returns may occur more frequently than downside returns.

Turning to the memory estimates of the residuals $\widehat{d}_u$, Panel B of Table 2 supports the notion of $I(d) - I(0)$ co-integration for all the pairs of variances; i.e., $\widehat{d}_u$ is not significantly different from zero. The degree of fractional co-integration is measured by $\left(Ave.\widehat{d} - \widehat{d}_u\right)$ as reported in the last column of Panel B: this is the reduction in the integration order of $u_t$ as compared to $y_t$ and $x_t$. The strongest fractional co-integration is found for $\left(IV_t^D,\, RV_t^D\right)$ and the degree of that long-run relation becomes lower for $(RX_t^2,\, RV_t)$ and $\left(IV_t^U,\, RV_t^U\right)$. The stronger co-movements between $IV_t^D$ and $RV_t^D$ suggest that they are more closely tied together by a highly persistent

common stochastic trend toward which they converge in the long run. Given that the $IV_t^D$ and $RV_t^D$ are both associated with dramatic market declines, our observations indicate that the well documented long-run equilibrium relationship between $RX_t^2$ and $RV_t$ may be non-trivially ascribed to investors' fear of a market crash, which is described by the relation between $IV_t^D$ and $RV_t^D$. In addition, the two components of the implied variance co-move with each other with the degree of fractional co-integration slightly lower than $\left(IV_t^D,\, RV_t^D\right)$. The $\left(RV_t^U,\, RV_t^D\right)$ exhibits the weakest co-integrating relation across the pairs of variances considered.

## 4.2 the Common Long-Memory Component

This subsection attempts to establish an intuitive argument for the linkage between the fractional co-integration found in the pairs of variances and the return predictability suggested by the variance differences. As indicated by Bollerslev, Tauchen, and Zhou (2009) and by Bollerslev, Sizova, and Tauchen (2012), the vol-of-vol factor (the fundamental uncertainty) in the generalized LRR model solely determines the temporal variation in $VRP$, imparting it with forecasting power for returns in cases where the vol-of-vol serves as the driving force of the variation in the equity premium. In an extended model presented by Feunou, Jahan-Parvar, and Okou (2017), decompositions of the $VRP$ are theoretically motivated and asymmetry in the upside and downside vol-of-vol factors is shown to drive the dynamics in the components of $VRP$. In the work of Conrad and Loch (2015), a long-term volatility component from the GARCH-MIDAS model is found to represent the vol-of-vol factor, and the new $VRP$ measure based on that component displays higher predictive power for future market returns. In the light of these findings, we propose to treat the common long-memory component of variances that are fractionally co-integrated as the factor intimately associated with the vol-of-vol that determines the variance premium.

In what follows, we employ the permanent-transitory (PT) decomposition of Gonzalo and Granger (1995) to identify the common permanent component of each pair of variances. To facilitate the PT decomposition, we follow Dolatabadi et al. (2018) by constructing a fully specified

framework which accommodates long-run and short-run effects in a system containing fractionally co-integrated variables. Such a framework is afforded by the fractionally co-integrated vector autoregressive (FCVAR) model of Johansen (2008) and Johansen and Nielsen (2012) as follows

$$\Delta^d(X_t-\mu)=\alpha\beta'\Delta^{d-b}L_b(X_t-\mu)+\sum_{c=1}^{k}\Gamma_c\Delta^dL_b^c(X_t-\mu)+\varepsilon_t \tag{20}$$

where $X_t$ denotes the pair of variances under analysis, $\Delta^d$ and $L_b$ respectively represent the fractional difference operator and the fractional lag operator[6]. As noted by Johansen and Nielsen (2015), the level parameter $\mu$ serves to reduce the effects of pre-sample observations. The error correction term is expressed as $\beta'\Delta^{d-b}L_bX_t$, where $\beta$ is of order $(2\times1)$. The linear combination $\beta'X_t$ is integrated of order $(d-b)$, suggesting that the co-integrating combination reduces the integration order of $X_t$ by $b$. The parameter $b$ therefore measures the degree of fractional co-integration in the FCVAR framework. The matrix $\alpha$ is of order $(2\times1)$ and contains parameters representing the speed of adjustment towards long-run equilibrium. The short-run dynamics are measured by the lag coefficients $(\Gamma_1,\ldots,\Gamma_k)$.

With the estimated parameters of the FCVAR, we then obtain the PT decomposition as follows

$$X_t=A_1LM_t+A_2SM_t \tag{22}$$

where the common long-memory component of $X_t$ is given by $LM_t=\alpha_\perp'X_t$ with $\alpha_\perp'\alpha=0$ and the transitory component is constituted by $SM_t=\beta'X_t$. Additional details relating to the construction of $A_1$ and $A_2$ may be found in Gonzalo and Granger (1995). For a given value of the transitory component, a high value of $LM_t$ increases the magnitude of the forecast of the conditional variance,

[6] $L_b=1-\Delta^b$ and $\Delta^d=(1-L)^d$ where

$$\begin{aligned}(1-L)^d &= \sum_{i=0}^{\infty}\theta_i(d)L^i,\ \text{with}\ \theta_i(d)=(-1)^i\binom{d}{i}=\frac{\Gamma(-d+i)}{\Gamma(-d)\Gamma(i+1)} \\ &= 1-dL+\frac{d(d-1)}{2!}L^2-\frac{d(d-1)(d-2)}{3!}L^3+\cdots\end{aligned} \tag{21}$$

which renders the $LM_t$ closely associated with the vol-of-vol discussed in Bollerslev, Tauchen, and Zhou (2009).

Panel A of Table 3 presents parameter estimates of the FCVAR for different pairs of variances[7]. Consistent with earlier findings given by the FMNBLS estimator, the results in Table 3 show that (i) $\left(IV_t^D,\, RV_t^D\right)$ possesses the greatest magnitude of co-integration, having the highest value of $\widehat{b}$ and (ii) that the long-run relation becomes weaker for $(RX_t^2,\, RV_t)$ and $\left(IV_t^U,\, RV_t^U\right)$. Notably, the memory estimates $\widehat{d}$ within the FCVAR framework do not accord with those based on the local Whittle estimator. This may be explained by the fact that the estimates $\widehat{d}$, $\widehat{b}$ and $\widehat{\beta}$ are highly sensitive to the presence of short-run dynamics. Such dynamics are not parameterized in the local Whittle estimator and the FMNBLS estimator. We then extract the common permanent component of each pair of variances and examine whether that common component works to represent the vol-of-vol factor that constitutes the source of the variance premium.

We follow Conrad and Loch (2015) by exploiting the model in Bollerslev, Sizova, and Tauchen (2012) where the $VRP$ is written as an affine function of the vol-of-vol factor as follows

$$\begin{aligned} IV_t &= c + RV_t + b^{(LM)} LM_t \qquad (23) \\ \text{alternatively, } IV_t - RV_t &= c + b^{(LM)} LM_t \end{aligned}$$

where $b^{(LM)} > 0$. Given that we construct the $VRP$ and its components under the assumption that the $RV_t$ follows a martingale difference, we first regress $IV_t$ against a constant, $RV_{t-1}$ and $LM_t$, and then regress the ex-post $VRP$ measured by $(IV_t - RV_t)$ against a constant, $RV_{t-1}$ and $LM_t$. If the $LM_t$ can be thought of as the factor that determines the $VRP$, the two slope coefficients in the first regression should be significant but where only $LM_t$ is significant in the second regression.

Panel B of Table 3 reports the results for the first regression where the $(IV_t,\, RV_t)$ is replaced by $(RX_t^2,\, RV_t)$, $\left(IV_t^U,\, RV_t^U\right)$ and $\left(IV_t^D,\, RV_t^D\right)$, and the corresponding common permanent component

[7] Lag $k$ is selected based on the Bayesian information criterion (BIC) and the white noise test on the residuals. For more details, see Nielsen and Shibaev (2018).

is denoted by $LM_t$, $LM_t^U$ and $LM_t^D$. For each scenario considered, we show that both the realized measure of variance and the common long-memory component exert significantly positive effects on the implied variance counterparts. In addition, results for the second regression presented in Panel C suggest that only the permanent components play a significant role in driving the ex-post $VRP$. Our results thus present strong evidence that the common long-memory component of the total (upside/downside) implied-realized variances can be regarded as a factor closely associated with the fundamental uncertainty, i.e. the total (upside/downside) vol-of-vol.

In the work of Bollerslev, Tauchen, and Zhou (2009), the degree of persistence in the vol-of-vol factor plays a critical role in determining the return predictability over longer horizons. We consider the common permanent component ($LM_t$) of the fractionally co-integrated system as a proxy of the vol-of-vol and infer the association between the $LM_t$ and return predictability as follows. An increase in the persistence of the $LM_t$ would suggest a greater reduction in the integration order of $X_t$ via the co-integrating combination, and thus a higher magnitude of fractional co-integrating relationship. By the same logic as in Bollerslev, Tauchen, and Zhou (2009), a higher degree of that relationship then leads to a higher degree of long-horizon return predictability.

Thus, the $VRP_t^D$ is expected to display higher predictive power for future returns with the stronger co-fractional relation between $\left(IV_t^D,\, RV_t^D\right)$. On the other hand, the weaker fractional co-integration found for $(RX_t^2,\, RV_t)$ and $\left(IV_t^U,\, RV_t^U\right)$ may adversely affect the return predictability implied by their differences, i.e. $VRP_t$ and $VRP_t^U$. Such hypotheses will be further examined in the following subsection. Here, we do not seek the proxy of the vol-of-vol factor for the differences between $\left(RV_t^U,\, RV_t^D\right)$ and $\left(IV_t^U,\, IV_t^D\right)$, i.e. $SJ_t$ and $IVA_t$, since the jump risk affects the equity premium via the impact of conditional volatility rather than the vol-of-vol factor as documented in Guo, Wang, and Zhou (2019).

### 4.3 Return Predictability

Before evaluating the return predictability implied by the variance differences, we first investigate the risk-return relations over different frequencies. We adopt the time-domain band-pass filtering procedure considered in Bollerslev et al. (2013) to decompose each of the variables (i.e. variances and future excess returns) into their long-run, intermediate, and short-run components. In applying the local Whittle estimator in Section 4.1, we set the bandwidth parameter $m_1 = T^{0.5}$, which corresponds to a low frequency around $\underline{\omega} \approx 0.4845$, or periods around 12 months. Given that the return predictability of the $VRP$ achieves its maximum at the quarterly horizon in the study of Bollerslev, Tauchen, and Zhou (2009), we isolate the high-frequency component with periodicities of less than 3 months with $\overline{\omega} \approx 2.0933$. Finally, we extract an intermediate-frequency component located in the band $\underline{\omega} < \omega < \overline{\omega}$ for periods between 3 and 12 months.

Table 4 outlines correlations between the extracted components of the variance differences and the corresponding components of excess returns. These show that the $VRP_t$, $VRP_t^U$ and $VRP_t^D$ all exhibit non-trivial association with excess returns over different frequency bands, where the greatest correlations occur at the low frequencies. It is evident that $VRP_t^D$ dominates its rivals, $VRP_t$ and $VRP_t^U$, in demonstrating higher risk-return relations, so indicating better predictive power for future returns. However, neither of $SJ_t$ and $IVA_t$ are significantly correlated with excess returns at intermediate and low frequencies. The significantly negative risk-return relation for $SJ_t$ in the short run is in agreement with the results of Bollerslev, Li, and Zhao (2017). The long-run correlations between variance differences and returns are also depicted in Figure 1, where $VRP_t$, $VRP_t^U$ and $VRP_t^D$ generally display strong co-movements with returns. However, no clear pattern of long-run co-movements is evident for $SJ_t$ and $IVA_t$.

Our analysis of return predictability is based on linear regressions of SPX excess returns on a

set of lagged predictors as follows

$$\frac{1}{h}\sum_{j=1}^{h} r_{t+j} = b_0 + b_1 V_t + e_{t+h,t} \tag{24}$$

where $V_t$ denotes a number of variance differences including $VRP_t$, $VRP_t^U$, $VRP_t^D$, $SJ_t$ and $IVA_t$ and the issue of overlapping return observations is accommodated using the heteroskedasticity and serial correlation consistent standard errors. Values of the adjusted $R^2$s are plotted in Figure 2. In line with the existing literature, we show a hump-shaped pattern as a function of the time horizons for $VRP_t$ and $VRP_t^U$, where the $R^2$ peaks at the quarterly horizon. Compared with the $VRP_t^U$, the $VRP_t$ yields higher adjusted $R^2$ values over various horizons. This echoes the higher correlation between $VRP_t$ and excess returns across frequencies reported in Table 4. Moreover, the superior predictive power for future returns is observed for $VRP_t^D$ which achieves a maximum of return predictability at the six-month horizon.

The striking performance of the $VRP_t^D$ may be explained by the stronger co-fractional relation between $\left(IV_t^D,\, RV_t^D\right)$ uncovered in Section 4.1. The higher degree of that relation accompanies higher persistence in the common long-memory component of $\left(IV_t^D,\, RV_t^D\right)$, which prolongs the return predictability of $VRP_t^D$. The ranking of the magnitude of co-fractional relation between $(RX_t^2,\, RV_t)$, $\left(IV_t^U,\, RV_t^U\right)$ and $\left(IV_t^D,\, RV_t^D\right)$ is identical to that obtained for the degree of return predictability inherent in their differences. This confirms our hypothesis for the positive linkage between the degree of fractional co-integration between variances and the magnitude of return predictability afforded by their differences. Turning to the results of $SJ_t$ and $IVA_t$, the values of their $R^2$ appear negligible compared with those implied by the $VRP$ and its components. In addition, their predictive power for future returns is initially fairly low, before climbing to a maximum around $h = 3$ then dissipating quickly in expanding the forecasting horizon. This may be attributable to the weak association of $SJ_t$ and $IVA_t$ with the future returns at frequencies in excess of 3 months.

# 5 Simulation

A common challenge in the derivation of implied variances is that only a limited range and coarse partition of strikes are actually traded in the market. This violates the definition of the model-free implied variance of Britten-Jones and Neuberger (2000) where an integral of option prices is used over an infinite range of strikes. The approximation errors induced by the discrete set of strikes may result in the mis-estimation of both $VRP$ and its predictive power for future returns. In practice, such implementation issues are fixed via the use of interpolation and extrapolation scheme as in Jiang and Tian (2007), Andersen and Bondarenko (2007) and Kilic and Shaliastovich (2018), among others. However, little consensus exist regarding the most appropriate extrapolation method for the construction of unbiased estimates of implied variances. Against this background, the following subsection undertakes a simulation study, which exploits a sufficiently fine partition and wide range of strikes to test the empirical findings discussed above.

Our simulations are based upon the jump-diffusion model employed in Duan and Yeh (2010). The strike price increment is fixed as $\Delta K = 1$ and the moneyness range $K/F_t$ is set from 0.85 to 1.15. That level of moneyness is in line with the work of Kilic and Shaliastovich (2018) and is considered sufficiently deep in the tails to allow for the observed volatility smile or skew. Details in terms of the model specification and the values of parameters are provided in Appendix C. For simplicity, we assume no dividends and a zero interest rate. Sampling at a frequency of once every 5 minutes is considered, which corresponds to 78 intraday return observations. We simulate 4500 daily observations and construct a monthly dataset as in Section 4, which results in 150 non-overlapping observations for each series of return and variances.

Table 5 summarizes the estimates of the orders of fractional integration and the results of the co-integrating regressions for each pair of variances. Panel A presents clear Monte Carlo evidence for the existence of fractional co-integration between variances with $\widehat{r} = 1$. Panel B of the table supports the long-run unbiasedness between $(VIX_t^2,\ RV_t)$ and reveals an upward bias in the $IV_t^D$

as a predictor of $RV_t^D$. This in turn supports the empirical findings reported earlier. Unlike the empirical study, the $IV_t^U$ underestimates $RV_t^U$ with the slope estimate significantly greater than unity. This could be due to the high volatility setting considered in the simulation where investors dislike downside variance risk and prefer variation in upside variance risk which improves the likelihood of substantial gains. The results in Panel B suggest that extreme downside returns occur more frequently than extreme upside returns, under both the risk-neutral and physical measures. The conflict with empirical evidence for the long-run unbiasedness between $\left(IV_t^D,\, RV_t^D\right)$ may well be explained by the inclusion of deep OTM put options in our simulation which are absent in the market. As indicated by $\left(Ave.\widehat{d} - \widehat{d}_u\right)$, the pairs of variances yield the same pattern as the empirical study for the strength of fractional co-integration, where the highest magnitude is observed for $\left(IV_t^D,\, RV_t^D\right)$.

We finish this section by reporting in Table 6 the results for our Monte Carlo equivalent to the return regression in equation (24). The superiority of $VRP_t^D$ in terms of return predictions is clearly evident once $h > 6$. Although the $VRP_t^U$ outperforms $VRP_t$ in yielding higher values of adjusted $R^2$ for $h < 12$, its predictive power for future returns is exceeded by the $VRP_t$ as the forecasting horizon grows. In addition, our simulation supports the empirical evidence of the inferior performances of the $SJ_t$ and $IVA_t$ in predicting returns, especially for intermediate and long horizons after $h > 9$. The nature of the $SJ_t$ as a short-lived return predictor is discussed in Bollerslev, Li, and Zhao (2017), who show that the $SJ_t$ only predicts returns at weekly horizons and raise concerns regarding the role of $SJ_t$ as a priced risk factor.

Taken as a whole, the simulation study, that attenuates the approximation errors induced by the limited availability of strikes, supports the empirical findings concerning the positive relationship between the strength of fractional co-integration and the degree of long-horizon return predictability. We therefore conclude that such linkage is not due to the implementation errors of implied variances that are commonly encountered in the empirical study.

# 6 Conclusion

Our study provides new evidence of the presence of fractional co-integration between the upside (downside) components of the SPX implied and realized variances. The strength of such long-run relationship is found to enhance the degree of return predictability suggested by the corresponding variance differences. We explain such linkage by the common long-memory component of the fractionally co-integrated system, which is intimately associated with the volatility-of-volatility factor determining the variation in the variance premium. Specifically, an increase in the degree of persistence of the common long-memory component results in higher fractional co-integration between variances in magnitude, which leads to better return predictability over longer horizons that are inherent in their variance differences. Hence, the better performance of the downside variance premium as a return predictor is attributable to the stronger co-fractional relation between the downside implied and realized variances.

Despite the presence of the non-trivial fractional co-integration between the upside and downside components of the implied (realized) variances, their difference is only weakly associated with future returns at low frequencies and the predictive power dissipates over longer horizons. To circumvent the issue of measurement errors underlying the construction of implied variances, we base a simulation study on a wide range and fine partition of strikes. Our simulation results generally support the empirical findings illustrated above.

# References


Andersen, T. G., and O. Bondarenko (2007): "Construction and interpretation of model-free implied volatility," in *Volatility as an Asset Class*, ed. by I. Nelken, pp. 141–181. Risk Books, London, UK.

Andersen, T. G., O. Bondarenko, and M. T. Gonzalez-Perez (2015): "Exploring Return Dynamics via Corridor Implied Volatility," *Review of Financial Studies*, 28(10), 2902–2945.

Bali, T. G., and H. Zhou (2016): "Risk, Uncertainty, and Expected Returns," *Journal of Financial and Quantitative Analysis*, 51(3), 707–735.

Bandi, F. M., and B. Perron (2006): "Long memory and the relation between implied and realized volatility," *Journal of Financial Econometrics*, 4(4), 636–670.

Bansal, R., and A. Yaron (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *The journal of Finance*, 59(4), 1481–1509.

Barndorff-Nielsen, O. E., S. Kinnebrock, and N. Shephard (2010): "Measuring downside risk-realised semivariance," in *Volatility and Time Series Econo- metrics: Essays in Honor of Robert F. Engle*, ed. by T. Bollersley, J. R. Russell, and M. Watson, pp. 117–136. Operational Research Society, New York.

Barndorff-Nielsen, O. E., and N. Shephard (2002): "Econometric analysis of realized volatility and its use in estimating stochastic volatility models," *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 64, 253–280.

Barndorff-Nielsen, O. E., and A. E. Veraart (2013): "Stochastic Volatility of Volatility and Variance Risk Premia," *Journal of Financial Econometrics*, 11(1), 1–46.

Bollerslev, T., S. Z. Li, and B. Zhao (2017): "Good Volatility , Bad Volatility , and the Cross Section of Stock Returns," *Journal of Financial and Quantitative Analysis*, pp. 1–57.

Bollerslev, T., J. Marrone, L. Xu, and H. Zhou (2014): "Stock Return Predictability and Variance Risk Premia: Statistical Inference and International Evidence," *Journal of Financial and Quantitative Analysis*, 49(03), 633–661.

Bollerslev, T., D. Osterrieder, N. Sizova, and G. Tauchen (2013): "Risk and return: Long-run relations, fractional cointegration, and return predictability," *Journal of Financial Economics*, 108(2), 409–424.

Bollerslev, T., N. Sizova, and G. Tauchen (2012): "Volatility in equilibrium: Asymmetries and dynamic dependencies," *Review of Finance*, 16(1), 31–80.

Bollerslev, T., G. Tauchen, and H. Zhou (2009): "Expected stock returns and variance risk premia," *Review of Financial Studies*, 22(11), 4463–4492.

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

Conrad, C., and K. Loch (2015): "The variance risk premium and fundamental uncertainty," *Economics Letters*, 132, 56–60.

Demeterfi, K., E. Derman, M. Kamal, and J. Zou (1999): "More Than You Ever Wanted To Know About Volatility Swaps," *Working Paper*.

Dolatabadi, S., P. K. Narayan, M. Ø. Nielsen, and K. Xu (2018): "Economic significance of commodity return forecasts from the fractionally cointegrated VAR model," *Journal of Futures Markets*, 38(2), 219–242.

Drechsler, I., and A. Yaron (2010): "What's Vol got to do with it," *Review of Financial Studies*, 24(1), 1–45.

Duan, J.-C., and J.-G. Simonato (1998): "Empirical martingale simulation for asset prices," *Management Science*, 44, 1218–1233.

Duan, J.-C., and C.-Y. Yeh (2010): "Jump and volatility risk premiums implied by VIX," *Journal of Economic Dynamics and Control*, 34(11), 2232–2244.

Feunou, B., M. R. Jahan-Parvar, and C. Okou (2017): "Downside Variance Risk Premium," *Journal of Financial Econometrics*, 16(3), 341–383.

Gonzalo, J., and C. W. Granger (1995): "Estimation of Common Long-Memory Components in Cointegrated Systems," *Journal of Business and Economic Statistics*, 13(1), 27–35.

Guo, H., K. Wang, and H. Zhou (2019): "Good Jumps, Bad Jumps, and Conditional Equity Premium," *Working Paper*.

Huang, T., and J. Li (2019): "Option-Implied variance asymmetry and the cross-section of stock returns," *Journal of Banking and Finance*, 101, 21–36.

Jiang, G. J., and Y. S. Tian (2007): "Extracting Model-Free Volatility from Option Prices: An examination of the VIX index," *Journal of Derivatives*, 14(3), 35–60.

Johansen, S. (2008): "A representation theory for a class of vector autoregressive models for fractional processes," *Econometric Theory*, 24(03), 651–676.

Johansen, S., and M. Ø. Nielsen (2012): "Likelihood inference for a fractionally cointegrated vector autoregressive model," *Econometrica*, 80, 2667–2732.

Johansen, S., and M. Ø. Nielsen (2015): "The Role of Initial Values in Conditional Sum-of-Squares Estimation of Nonstationary Fractional Time Series Models," *Econometric Theory*, 32(5), 1–45.

Kellard, N., C. Dunis, and N. Sarantis (2010): "Foreign exchange, fractional cointegration and the implied-realized volatility relation," *Journal of Banking and Finance*, 34(4), 882–891.

Kilic, M., and I. Shaliastovich (2018): "Good and Bad Variance Premia and Expected Returns," *Management Science*, 65(6), 2522–2544.

Nielsen, M. Ø. (2007): "Local Whittle Analysis of Stationary Fractional Cointegration and the Implied-Realized Volatility Relation," *Journal of Business and Economic Statistics*, 25(4), 427–446.

Nielsen, M. Ø., and P. Frederiksen (2011): "Fully modified narrow-band least squares estimation of weak fractional cointegration," *The Econometrics Journal*, 14(1), 77–120.

Nielsen, M. Ø., and S. S. Shibaev (2018): "Forecasting daily political opinion polls using the fractionally cointegrated vector auto-regressive model," *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 181(1), 3–33.

Nielsen, M. Ø., and K. Shimotsu (2007): "Determining the cointegrating rank in nonstationary fractional systems by the exact local Whittle approach," *Journal of Econometrics*, 141(2), 574–596.

Robinson, P., and Y. Yajima (2002): "Determination of cointegrating rank in fractional systems," *Journal of Econometrics*, 106(2), 217–241.

Shimotsu, K., and P. C. B. Phillips (2005): "Exact local Whittle estimation of fractional integration," *The Annals of Statistics*, 33(4), 1890–1933.

Yao, X., and M. Izzeldin (2018): "Forecasting using alternative measures of model-free option-implied volatility," *Journal of Futures Markets*, 38(2), 199–218.

# Data Citation

[data set] Federal Reserve Bulletin, 2016; Federal Reserve Bulletin, January 02 2003 to December 30 2016.

[data set] OptionMetrics Database, 2016; OptionMetrics, January 02 2003 to December 30 2016.

[data set] Tick Data Inc., 2016; Tick Data, January 02 2003 to December 30 2016.

# Appendix

## A Exact Local Whittle Estimator

To define the local Whittle estimator, we assume that a process $X_t$ integrated of order $d$ has the spectral density $f(\lambda)$ given by

$$f(\lambda) \sim G\lambda^{-2d}, \text{ as } \lambda \to 0_+ \tag{25}$$

where $G \in (0, \infty)$ is a finite and nonzero matrix with strictly positive diagonal elements. The discrete Fourier transform (DFT), $\omega_x(\lambda_j)$, and the periodogram, $I_x(\lambda_j)$ of $X_t$, $t = 1, \cdots, T$ at the fundamental frequencies can be written as

$$\begin{aligned} \omega_x(\lambda_j) &= (2\pi T)^{-1/2} \Sigma_{t=1}^{T} X_t e^{it\lambda_j}, \ \lambda_j = \frac{2\pi j}{T}, \ j = 1, ..., m_1 < \frac{T}{2} \\ I_x(\lambda_j) &= |\omega_x(\lambda_j)|^2 \end{aligned} \tag{26}$$

Shimotsu and Phillips (2005) provide a procedure which can be applied in the stationary and nonstationary regions to estimate $(G, d)$ by minimizing the objective function

$$Q_{m_1}(G, d) = \frac{1}{m_1} \Sigma_{j=1}^{m_1} [\log(G\lambda_j^{-2d}) + \frac{1}{G} I_{\Delta^d x}(\lambda_j)] \tag{27}$$

where $I_{\Delta^d x}(\lambda_j)$ is the periodogram of $\Delta^d X_t$. Concentrating $Q_{m_1}(G,d)$ with respect to $G$, we obtain the exact local Whittle (ELW) estimator given by

$$\widetilde{d} = \underset{d\in[\Delta_1,\Delta_2]}{\arg\min} R(d) \tag{28}$$

with

$$\begin{aligned} R(d) &= \log \widehat{G}(d) - 2d\frac{1}{m_1}\Sigma_{j=1}^{m_1}\log\lambda_j \\ \widehat{G}(d) &= \frac{1}{m_1}\Sigma_{j=1}^{m_1} I_{\Delta^d x}(\lambda_j) \end{aligned} \tag{29}$$

where $\Delta_1$ and $\Delta_2$ are the lower and upper bounds of the admissible values of $d$.

## B Exact Local Whittle Rank Test

Assume there is a $p$-vector fractional process $X_t$ where each element is fractionally integrated of order $d_1, ..., d_p$. The work of Nielsen and Shimotsu (2007) builds on the assumption of equal integration orders and thus $d_1, ..., d_p$ are represented by $d_*$, where $\widetilde{d}_* = \frac{1}{p}\sum_{a=1}^{p}\widehat{d}_a$ with each $\widehat{d}_a$ given by equation (28). The consistent estimator of the spectral density at the origin is given by

$$\widehat{G}(d_*) = \frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}[I_{\Delta(L;d_*,...,d_*)x}(\lambda_j)] \tag{30}$$

where $\lambda_j = \frac{2\pi j}{T}$ and $I_{\Delta(L;d_*,...,d_*)x}(\lambda_j)$ is the periodogram of $(\Delta^{d_*}X_{1t,...,}\Delta^{d_*}X_{pt})'$. Let $\widehat{\delta}_a$ be the $a$th eigenvalues of $\widehat{G}(d_*)$ and the co-integration rank $r$ can be determined by adopting the procedure of Robinson and Yajima (2002)

$$\begin{aligned} \widehat{r} &= \underset{u=0,...,p-1}{\arg\min} L(u) \\ L(u) &= v(T)(p-u) - \sum_{a=1}^{p-u}\widehat{\delta}_a \end{aligned} \tag{31}$$

where $v(T)$ should be positive and should meet the assumption as follows

$$v(T) + \frac{1}{m^{1/2}v(T)} \to 0 \tag{32}$$

Nielsen and Shimotsu (2007) show that a higher rank estimate is more likely to be selected when a larger $v(T)$ is applied. In order to obtain a more conservative estimate of $r$, we choose to employ a small $v(T) = m^{-0.4}$.

Once the presence of the co-fractional relation has been investigated by the rank estimation, we can examine the equality of the orders of fractional integration by testing the null $H_0 : d_a = d_*$, $a = 1, ...p$. The test statistic is given by

$$\widehat{T}_0 = m(S\widehat{d})'(S\frac{1}{4}\widehat{D}^{-1}(\widehat{G} \odot \widehat{G})\widehat{D}^{-1}S' + h(T)^2 I_{p-1})^{-1}(S\widehat{d}) \tag{33}$$

where $\odot$ represents the Hadamard product, $S = [I_{p-1}, -\iota]'$, $\iota$ is the $(p-1)$-vector of ones, and $\widehat{D} = diag(\widehat{G}_{11}, \cdots \widehat{G}_{pp})$. The memory estimates $\widehat{d}$ of variables in the vector are derived by the univariate exact local Whittle estimator (Shimotsu and Phillips (2005)). When variables are not fractionally co-integrated, $\widehat{T}_0 \to_d \chi^2(p-1)$, while $\widehat{T}_0 \to_p 0$ when they are co-integrated.

## C Simulation Design

We follow Duan and Yeh (2010) by employing a stochastic volatility model with jumps in simulating the asset price and the latent stochastic volatility as follows

$$\begin{aligned} d\ln S_t &= [r - q + \delta_s V_t - \frac{V_t}{2}]dt + \sqrt{V_t}dW_t + J_t dN_t - \lambda\mu_J dt \\ dV_t &= \kappa(\theta - V_t)dt + \upsilon V_t^{\gamma} dB_t \end{aligned} \tag{34}$$

where $W_t$ and $B_t$ are Wiener processes correlated with the coefficient $\rho$; $N_t$ represents a Poisson process with intensity $\lambda$; $J_t$ is an independent random variable following $N(\mu_J, \sigma_J^2)$. The price, $S_t$, and volatility, $V_t$, processes are dependent through the correlated diffusive terms–$W_t$ and $B_t$. The other parameters, $r$, $q$ and $\delta_s$ are the risk-free rate, the dividend yield and the asset risk

premium, respectively. The option pricing is conducted using the corresponding model under the risk-neutral probability measure given by

$$\begin{aligned} d\ln S_t &= [r-q-\frac{V_t}{2}+\lambda^*(\mu_J^*+1-e^{\mu_J^*+\sigma_J^2})]dt+\sqrt{V_t}dW_t^* \\ &\quad +J_t^*dN_t^*-\lambda^*\mu_J^*dt \\ dV_t &= (\kappa\theta-\kappa^*V_t)dt+\upsilon V_t^{\gamma}dB_t^* \end{aligned} \tag{35}$$

where $\kappa^*=\kappa+\delta_V$ and $B_t^*=B_t+\delta_V/\upsilon\int_0^t V_s^{1-\gamma}ds$ with $\delta_V$ being the volatility risk premium. Other notations with the superscript $*$ have similar features to their counterparts under the physical measure in equation (34) and the standard deviation of $J_t^*$ remains unchanged as $\sigma_J$. To compute option prices, we adopt the empirical martingales simulation (EMS) method developed by Duan and Simonato (1998) given that there is no closed-form option pricing formula for equation (35).

The initial stock price ($S_0$) and latent stochastic volatility ($V_0$) are set respectively as 1000 and 0.08. Values of the parameters are set similar to those considered in Duan and Yeh (2010).

| $\kappa$ | $\theta$ | $\lambda$ | $\mu_J(\%)$ | $\sigma_J(\%)$ | $\upsilon$ | $\rho$ | $\gamma$ | $\delta_s$ | $\kappa^*$ | $\phi^*(\%)$ | $\delta_V$ | $\delta_J(\%)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2.500 | 0.080 | 55.000 | 0.300 | 0.500 | 1.400 | -0.800 | 0.900 | 0.420 | -13.000 | 0.035 | -15.500 | -0.059 |

**Table 1**

This table reports standard summary statistics for the monthly excess returns and the variance series considered in this study. In order to measure the return variation during the overnight period, we add the squared overnight returns, computed as the squared close-to-open logarithmic price change, to the realized variance obtained over the trading day. To overcome the issue of the overlapping data, the risk-neutral measures are taken only from the subsequent trading day after the final day employed in the derivation of the previous realized measure. The sample covers the period of Jan 2003 to Dec 2016, with a total of 168 non-overlapping observations.

| | Mean (%) | StDev. (%) | Skew. | Kurt. | AR(1) |
|---|---|---|---|---|---|
| **Excess Return** | | | | | |
| $r_t$ | 0.180 | 2.759 | -0.013 | 2.737 | 0.102 |
| **Variance** | | | | | |
| $RV_t$ | 12.933 | 28.537 | 7.186 | 64.905 | 0.668 |
| $RX_t^2$ | 19.232 | 22.748 | 4.402 | 28.348 | 0.744 |
| $VRP_t$ | 6.298 | 23.622 | -5.646 | 65.262 | 0.312 |
| $RV_t^U$ | 6.521 | 14.648 | 7.196 | 64.851 | 0.668 |
| $IV_t^U$ | 9.354 | 10.652 | 3.600 | 20.327 | 0.649 |
| $VRP_t^U$ | 2.833 | 13.413 | -5.881 | 60.784 | 0.365 |
| $RV_t^D$ | 6.412 | 13.901 | 7.161 | 64.767 | 0.667 |
| $IV_t^D$ | 11.587 | 15.554 | 3.786 | 19.453 | 0.660 |
| $VRP_t^D$ | 5.175 | 11.798 | 0.693 | 21.620 | 0.269 |
| $SJ_t$ | 0.109 | 1.106 | 2.979 | 19.218 | 0.336 |
| $IVA_t$ | -2.233 | 10.076 | -3.252 | 16.598 | 0.137 |

**Table 2**

Fractional Integration and Co-integration: an empirical analysis based on data from Jan 2003 to Dec 2016. Panel A reports the exact local Whittle estimates of fractional integration by Shimotsu and Phillips (2005) and the rank statistics from Nielsen and Shimotsu (2007). Ave.$\widehat{d}$ is the simple average of $\widehat{d}$ of each pair of variances; $\widehat{T}_0$ denotes the statistic of the modified Wald test for the null hypothesis of equality of $\widehat{d}$, which is calculated with $h(T) = 1/log(T)$. The value of the function $L(u)$ is calculated by employing a small $v(T) = m^{-0.4}$; $r = 1$ when $L(1) < L(0)$, suggesting the presence of co-fractional relation. Panel B reports the FMNBLS estimates. Numbers in parenthesis are asymptotic standard errors.

| | Panel A: $\widehat{d}$ estimates and co-integrating rank analysis $m_1 = T^{0.5}$, $m = T^{0.45}$, $v(T) = m^{-0.4}$ | | | | | | Panel B: FMNBLS estimates $m_0 = T^{0.4}$, $m_1 = T^{0.5}$, $m_2 = T^{0.8}$, $m_3 = m_0$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\widehat{d}$ | Ave.$\widehat{d}$ | $\widehat{T}_0$ | $L(0)$ | $L(1)$ | $\widehat{r}$ | $\widehat{c}$ | $\widehat{\theta}$ | $\widehat{d}_u$ | Ave.$\widehat{d} - \widehat{d}_u$ |
| $RV_t$ $(y_t)$ | 0.269 | | | | | | -7.370 | 1.056 | -0.060 | 0.387 |
| $RX_t^2$ $(x_t)$ | 0.383 | 0.326 | 0.585 | -1.204 | -1.531 | 1 | | (0.078) | (0.144) | |
| $RV_t^U$ $(y_t)$ | 0.270 | | | | | | -3.427 | 1.064 | 0.022 | 0.367 |
| $IV_t^U$ $(x_t)$ | 0.508 | 0.389 | 2.378 | -1.204 | -1.507 | 1 | | (0.069) | (0.144) | |
| $RV_t^D$ $(y_t)$ | 0.269 | | | | | | -3.236 | 0.833 | -0.127 | 0.425 |
| $IV_t^D$ $(x_t)$ | 0.328 | 0.298 | 0.157 | -1.204 | -1.533 | 1 | | (0.052) | (0.144) | |
| $RV_t^U$ $(y_t)$ | 0.270 | | | | | | -0.278 | 1.060 | 0.054 | 0.215 |
| $RV_t^D$ $(x_t)$ | 0.269 | 0.269 | 0.000 | -1.204 | -1.602 | 1 | | (0.009) | (0.144) | |
| $IV_t^U$ $(y_t)$ | 0.508 | | | | | | -1.911 | 0.972 | -0.003 | 0.420 |
| $IV_t^D$ $(x_t)$ | 0.328 | 0.418 | 1.727 | -1.204 | -1.572 | 1 | | (0.094) | (0.144) | |

**Table 3**

The Variance Premium and Fundamental Uncertainty: an empirical analysis based on data from Jan 2003 to Dec 2016. Panel A reports the FCVAR estimates and the $P$ values are presented in parentheses; Panel B demonstrates the estimation results for two regressions where $VRP_t = RX_t^2 - RV_t$, $VRP_t^U = IV_t^U - RV_t^U$ and $VRP_t^D = IV_t^D - RV_t^D$. The numbers in parentheses are $P$ values based on Newey-West standard errors.

| Panel A: FCVAR estimates | | | | | | Panel B: Regression I | | | | Panel C: Regression II | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $X_t = \left(RV_t,\, RX_t^2\right)$ | | | | | | $RX_t^2 = c + b^{(RV)}RV_{t-1} + b^{(LM)}LM_t + \xi_t$ | | | | $VRP_t = c + b^{(RV)}RV_{t-1} + b^{(LM)}LM_t + \xi_t$ | | | |
| $\widehat{d}$ | $\widehat{b}$ | $\widehat{\alpha}$ | $\widehat{\beta}$ | $\widehat{k}$ | $\widehat{\alpha}'_\perp$ | $\widehat{c}$ | $\widehat{b}^{(RV)}$ | $\widehat{b}^{(LM)}$ | adj. $R^2$ | $\widehat{c}$ | $\widehat{b}^{(RV)}$ | $\widehat{b}^{(LM)}$ | adj. $R^2$ |
| 0.352 (0.000) | 0.352 (0.000) | $\begin{pmatrix} -0.246 \\ 0.337 \end{pmatrix}$ | $\begin{pmatrix} 1 \\ -0.938 \end{pmatrix}$ | 5 | $\begin{pmatrix} 0.578 & 0.422 \end{pmatrix}$ | 7.474 (0.000) | 0.343 (0.000) | 0.466 (0.000) | 74.580 | 3.692 (0.001) | −0.085 (0.166) | 0.221 (0.000) | 3.066 |
| $X_t = \left(RV_t^U,\, IV_t^U\right)$ | | | | | | $IV_t^U = c + b^{(RV)}RV_{t-1}^U + b^{(LM)}LM_t^U + \xi_t$ | | | | $VRP_t^U = c + b^{(RV)}RV_{t-1}^U + b^{(LM)}LM_t^U + \xi_t$ | | | |
| $\widehat{d}$ | $\widehat{b}$ | $\widehat{\alpha}$ | $\widehat{\beta}'$ | $\widehat{k}$ | $\widehat{\alpha}'_\perp$ | $\widehat{c}$ | $\widehat{b}^{(RV)}$ | $\widehat{b}^{(LM)}$ | adj. $R^2$ | $\widehat{c}$ | $\widehat{b}^{(RV)}$ | $\widehat{b}^{(LM)}$ | adj. $R^2$ |
| 0.231 (0.049) | 0.231 (0.069) | $\begin{pmatrix} -0.302 \\ 0.430 \end{pmatrix}$ | $\begin{pmatrix} 1 \\ -0.738 \end{pmatrix}$ | 3 | $\begin{pmatrix} 0.588 & 0.412 \end{pmatrix}$ | 4.188 (0.006) | 0.196 (0.008) | 0.503 (0.000) | 58.024 | 1.716 (0.079) | −0.222 (0.101) | 0.295 (0.018) | 8.520 |
| $X_t = \left(RV_t^D,\, IV_t^D\right)$ | | | | | | $IV_t^D = c + b^{(RV)}RV_{t-1}^D + b^{(LM)}LM_t^D + \xi_t$ | | | | $VRP_t^D = c + b^{(RV)}RV_{t-1}^D + b^{(LM)}LM_t^D + \xi_t$ | | | |
| $\widehat{d}$ | $\widehat{b}$ | $\widehat{\alpha}$ | $\widehat{\beta}'$ | $\widehat{k}$ | $\widehat{\alpha}'_\perp$ | $\widehat{c}$ | $\widehat{b}^{(RV)}$ | $\widehat{b}^{(LM)}$ | adj. $R^2$ | $\widehat{c}$ | $\widehat{b}^{(RV)}$ | $\widehat{b}^{(LM)}$ | adj. $R^2$ |
| 0.435 (0.000) | 0.435 (0.000) | $\begin{pmatrix} 0.751 \\ -0.364 \end{pmatrix}$ | $\begin{pmatrix} 1 \\ -1.392 \end{pmatrix}$ | 4 | $\begin{pmatrix} 0.327 & 0.673 \end{pmatrix}$ | 2.595 (0.000) | 0.461 (0.000) | 0.606 (0.000) | 68.662 | 1.354 (0.122) | −0.246 (0.204) | 0.542 (0.000) | 17.487 |

**Table 4**

Risk-Return Relations across Frequencies. The table demonstrates the correlations between the short-(periodicities of 3 months and shorter), intermediate-(between 3 months and 12 months) and long-run (greater than 12 months) components of the future excess returns and the same components of various variance differences. The $P$-values reported in parentheses are based on heteroscedasticity and autocorrelation consistent (HAC) covariance matrix estimator with optimal lag selection.

| | $r_{t+1}^{high}$ | $r_{t+1}^{medium}$ | $r_{t+1}^{low}$ | | $r_{t+1}^{high}$ | $r_{t+1}^{medium}$ | $r_{t+1}^{low}$ |
|---|---|---|---|---|---|---|---|
| $VRP_t^{high}$ | 0.270 | | | $SJ_t^{high}$ | -0.144 | | |
| | (0.000) | | | | (0.022) | | |
| $VRP_t^{medium}$ | | 0.236 | | $SJ_t^{medium}$ | | 0.108 | |
| | | (0.002) | | | | (0.164) | |
| $VRP_t^{low}$ | | | 0.523 | $SJ_t^{low}$ | | | 0.061 |
| | | | (0.000) | | | | (0.432) |
| $VRP_t^{U^{high}}$ | 0.214 | | | $IVA_t^{high}$ | -0.064 | | |
| | (0.005) | | | | (0.414) | | |
| $VRP_t^{U^{medium}}$ | | 0.175 | | $IVA_t^{medium}$ | | -0.107 | |
| | | (0.024) | | | | (0.168) | |
| $VRP_t^{U^{low}}$ | | | 0.501 | $IVA_t^{low}$ | | | 0.075 |
| | | | (0.000) | | | | (0.338) |
| $VRP_t^{D^{high}}$ | 0.287 | | | | | | |
| | (0.000) | | | | | | |
| $VRP_t^{D^{medium}}$ | | 0.292 | | | | | |
| | | (0.000) | | | | | |
| $VRP_t^{D^{low}}$ | | | 0.588 | | | | |
| | | | (0.000) | | | | |

**Table 5**

Fractional Integration and Co-integration: a simulation study. Panel A reports the exact local Whittle estimates of fractional integration by Shimotsu and Phillips (2005) and the rank statistics from Nielsen and Shimotsu (2007). Ave.$\widehat{d}$ is the simple average of $\widehat{d}$ of each pair of variances; $\widehat{T}_0$ denotes the statistic of the modified Wald test for the null hypothesis of equality of $\widehat{d}$, which is calculated with $h(T) = 1/log(T)$. The value of the function $L(u)$ is calculated by employing a small $v(T) = m^{-0.4}$; $r = 1$ when $L(1) < L(0)$, suggesting the presence of co-fractional relation. Panel B reports the FMNBLS estimates. Numbers in parenthesis are asymptotic standard errors.

| | Panel A: $\widehat{d}$ estimates and co-integrating rank analysis $m_1 = T^{0.5}$, $m = T^{0.45}$, $v(T) = m^{-0.4}$ | | | | | | Panel B: FMNBLS estimates $m_0 = T^{0.4}$, $m_1 = T^{0.5}$, $m_2 = T^{0.8}$, $m_3 = m_0$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\widehat{d}$ | Ave.$\widehat{d}$ | $\widehat{T}_0$ | $L(0)$ | $L(1)$ | $\widehat{r}$ | $\widehat{c}$ | $\widehat{\theta}$ | $\widehat{d}_u$ | Ave.$\widehat{d} - \widehat{d}_u$ |
| $RV_t$ $(y_t)$ | 0.420 | | | | | | -34.797 | 1.002 | -0.051 | 0.451 |
| $VIX_t^2$ $(x_t)$ | 0.381 | 0.401 | 0.127 | -1.356 | -1.613 | 1 | | (0.025) | (0.107) | |
| $RV_t^U$ $(y_t)$ | 0.420 | | | | | | -17.704 | 1.135 | 0.002 | 0.399 |
| $IV_t^U$ $(x_t)$ | 0.383 | 0.401 | 0.121 | -1.356 | -1.621 | 1 | | (0.040) | (0.107) | |
| $RV_t^D$ $(y_t)$ | 0.417 | | | | | | -18.956 | 0.883 | -0.072 | 0.468 |
| $IV_t^D$ $(x_t)$ | 0.376 | 0.397 | 0.138 | -1.356 | -1.601 | 1 | | (0.012) | (0.107) | |
| $RV_t^U$ $(y_t)$ | 0.420 | | | | | | 0.037 | 0.960 | 0.204 | 0.215 |
| $RV_t^D$ $(x_t)$ | 0.417 | 0.419 | 0.001 | -1.356 | -1.676 | 1 | | (0.002) | (0.107) | |
| $IV_t^U$ $(y_t)$ | 0.383 | | | | | | -1.658 | 0.766 | 0.077 | 0.303 |
| $IV_t^D$ $(x_t)$ | 0.376 | 0.380 | 0.005 | -1.356 | -1.678 | 1 | | (0.002) | (0.107) | |

**Table 6**

Return predictability over various horizons: a simulation study. This table reports the adjusted $R^2 s(\%)$ implied by the return regression $\frac{1}{h}\sum_{j=1}^{h} r_{t+j} = b_0 + b_1 V_t + e_{t+h,t}$ with $h$ ranging from 1 month to 24 months and $V_t$ defined as various variance differences considered in the paper. All the regressions are based on monthly observations using the Newey-West standard errors.

| Variance Differences | Horizon | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 |
| $VRP_t$ | 1.148 | 0.331 | 0.475 | 1.335 | 0.851 | 0.716 | 0.947 | 0.903 | 0.742 |
| $VRP_t^U$ | 1.410 | 0.588 | 0.506 | 1.396 | 0.916 | 0.677 | 0.814 | 0.840 | 0.721 |
| $VRP_t^D$ | 1.025 | 0.194 | 0.558 | 1.446 | 0.955 | 0.876 | 1.198 | 1.122 | 0.911 |
| | | | | | | | | | |
| $SJ_t$ | 0.674 | 0.083 | 0.040 | 0.461 | 0.166 | 0.148 | 0.036 | 0.019 | 0.005 |
| $IVA_t$ | 0.100 | 0.113 | 0.165 | 0.943 | 0.170 | 0.176 | 0.172 | 0.119 | 0.035 |

**Figure 1**

Long-run co-movements between the excess returns and variance differences: an empirical analysis based on data from Jan 2003 to Dec 2016. The figure plots the long-run components (greater than 12 months) of the returns and various variance differences considered in the paper.

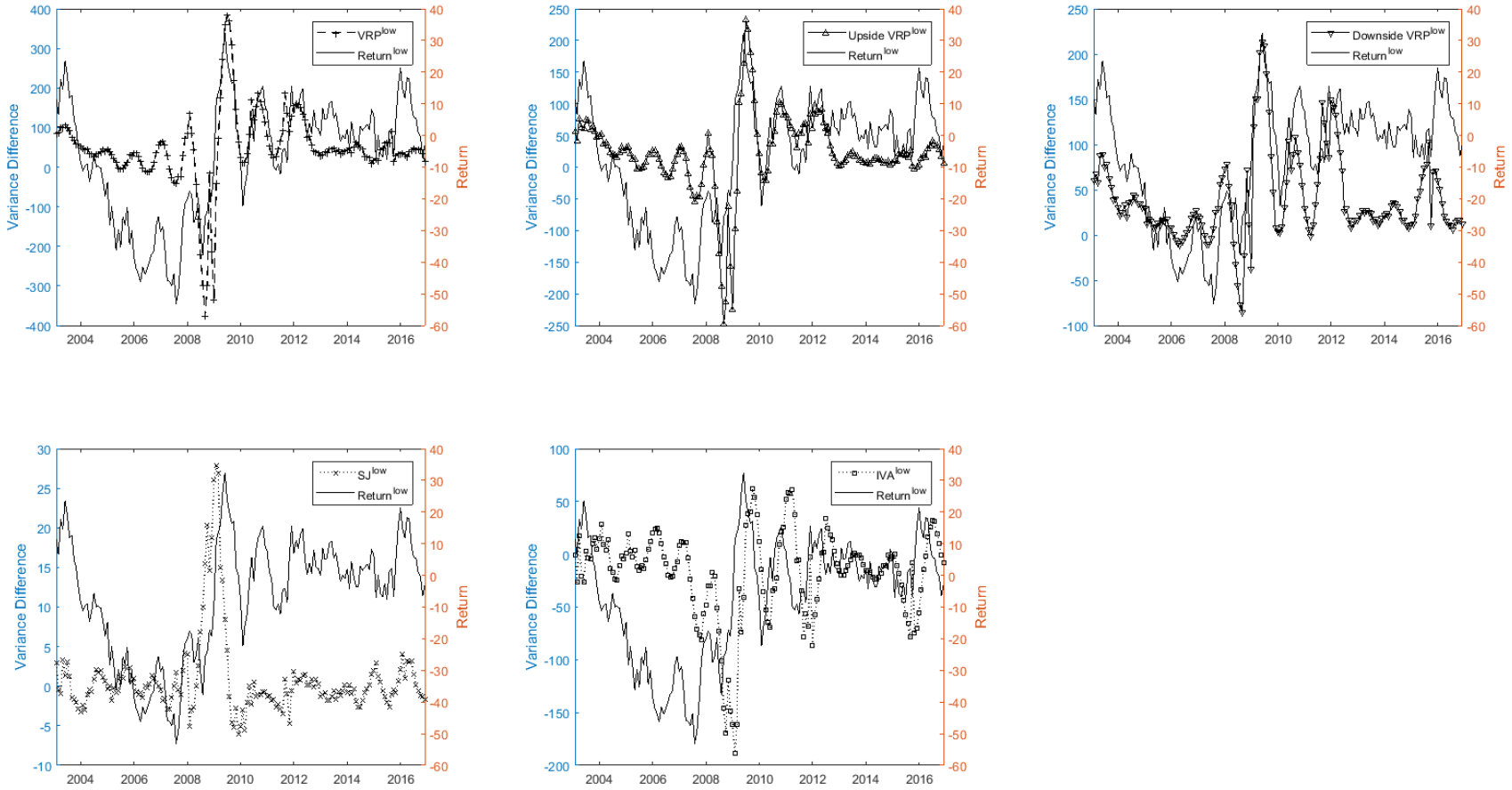

**Figure 2**

Return predictability: an empirical analysis based on data from Jan 2003 to Dec 2016. The figure plots the adjusted $R^2$ implied by the regression $\frac{1}{h}\sum_{j=1}^{h} r_{t+j} = b_0 + b_1 V_t + e_{t+h,t}$ with $h$ ranging from 1 month to 24 months and $V_t$ defined as various variance differences considered in the paper. The sample size is $T = 168$.