Statistics and Its Interface Volume 11 (2018) 153–168

On the surprising explanatory power of higher realized moments in practice

Keren Shen*, Jianfeng Yao, and Wai Keung Li,

Realized moments of higher order computed from intraday returns have been introduced in recent years. The literature indicates that realized skewness is an important factor in explaining future asset returns. However, the literature mainly focuses on the whole market, as well as the monthly or weekly scale. In this paper, we conduct an extensive empirical analysis to investigate the forecasting abilities of realized skewness and realized kurtosis towards an individual stock's future return and variance in the daily scale. It is found that realized kurtosis possesses significant forecasting power for the stock's future variance and in contrast with the existing literature, realized skewness is lack of explanatory power of future daily returns for individual stocks in the short term.

AMS 2000 subject classifications: Primary 62P20; secondary 91B84.
Keywords and phrases: High-frequency, Realized variance, Realized kurtosis, Linear regression, Trading volume.

1. INTRODUCTION

It is well known that excess kurtosis and negative skewness are stylized facts of stock return distributions. Realized moments of higher order computed from intraday returns have been introduced in recent years. This article conducts an extensive empirical analysis to investigate properties of realized skewness and realized kurtosis, especially the forecasting abilities of two higher order moments towards the stock's future return and variance. It is found that realized kurtosis possesses significant forecasting power for the stock's future variance.

Neuberger [15] first considers *realized skewness* of the asset price returns. Amaya et al. [1] further define *realized kurtosis*. These realized moments are constructed by the empirical sum of the corresponding powers of returns, which we call the naive estimator. However, the naive estimator is consistent only in the absence of microstructure noise, which must be handled with more sophisticated approaches.

Based on the pre-averaging method in Jacod et al. [12] for constructing realized variance, Liu et al. [13] introduce the pre-averaging estimator for realized skewness and kurtosis. In addition, they prove the consistency of the estimators in the presence of microstructure noise. They also find that realized skewness of the market price has significant forecasting power for the one-month-ahead excess equity market returns, by evidence from both in-sample and out-of-sample analysis. Furthermore, Choi and Lee [8] find that the relationship between realized skewness and subsequent stock returns depends on the impact of information releases. The relationship is negative when there is no high-impact information release, but it becomes positive if such releases appear. In addition, Schneider et al [16] show that the beta and volatility based low risk anomalies are driven by return skewness, both theoretically and empirically. In Amaya et al. [1], the authors investigate whether realized skewness and realized kurtosis are informative for next week's stock returns. They find that realized skewness has a significant negative effect on future stock returns. The authors also demonstrate the significance in the economic sense that buying stocks with the lowest realized skewness and selling stocks with the highest realized skewness generates a profit significantly. In addition, realized kurtosis exhibits a positive relationship with the weekly returns, even though the evidence is not always robust and statistically significant. Continuing the exploration along this line, we investigate whether higher realized moments have explaining power on daily returns and variances of future assets, estimated by realized variance.

In the empirical study, we show that in contrast with Amaya et al. [1] and Liu et al. [13], realized skewness shows not enough explanatory power for future daily returns. On the other hand, realized kurtosis, which is able to reflect the price jump size, shows strong evidence of forecasting power for future realized variances. We conduct regression analysis towards 70 randomly selected stocks from different industries and with capitalization sizes. Fifty four out of the 70 stocks are shown to have this property. Moreover, we find that the square root of realized kurtosis has an even better forecasting ability for future realized variances.

In addition, we compare the forecasting ability of realized kurtosis with other well-known variables, which may help in predicting asset's volatility, namely trading volume and signed daily return. In Chan and Fong [7], the authors conduct regression analysis of realized volatility against trading volume, trading frequency, average trading size and or-

arXiv: 1604.07969
*Corresponding author.

der imbalance. Trading volume is comprised of trading frequency and average trading size, while order imbalance is the difference between the number of trades initiated by buyers and sellers. The authors find that daily trading volume and trading frequency give equally good predictions on realized volatility, while average trading size and order imbalance add little explaining power. Therefore, in our empirical analysis, we only include daily trading volume as a possible covariate. Furthermore, signed returns are also informative for the volatility, especially the negative returns, which is usually interpreted as the *leverage effect*. The effect is first discussed by Black [3] and Christie [9], and is due to the fact that a negative return leads to an increase in the debt-to-equity ratio, resulting in an increase in the future volatility of the return [4]. In this article, we include both positive and negative daily returns as covariates. We find that in the presence of trading volume and signed daily returns, realized kurtosis generally still has great predictive power.

In summary, the main findings of the paper are the following:

- Realized skewness and kurtosis appear to be highly asymmetric and fat-tailed.
- Realized skewness lacks explanatory power for the future daily returns for individual stocks in the short term.
- Realized kurtosis exhibits significant forecasting power for the future realized variance.
- Realized kurtosis incorporates some information contained in trading volume.
- Some nonlinear relationships may exist between realized kurtosis and future daily volatility.

The rest of the paper is organized as follows. Section 2 reviews the estimators of higher realized moments. In Section 3, we examine the forecasting ability of higher realized moments for the future daily returns and return variances, for a chosen stock. The robustness of the result in Section 3 is checked in Section 4. Section 5 concludes the paper.

2. METHODOLOGY

2.1 Model setup

Define an adapted process $\{X_t, t \geq 0\}$ on some probability space $(\Omega, \mathcal{F}, P)$ as follows:

$$X_t = \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + \sum_{s \leq t} \Delta_s X, \tag{1}$$

where $\{\mu_s, 0 \leq s \leq t\}$ is an adapted locally bounded process, $\{\sigma_s, 0 \leq s \leq t\}$ is a càdlàg volatility process, and $\Delta_s X = X_s - X_{s-}$ is the jump of X at time s. Assume that the jump of X arrives through a finite jump process, for example, the compound Poisson process. The *quadratic variation* for the T-th day is defined as:

$$[X, X]_T = \int_{T-1}^{T} \sigma_s^2 ds + \sum_{T-1 \leq s \leq T} (\Delta_s X)^2.$$

2.2 Naive estimator

Let the grid of observation times of the T-th day be given by $\mathcal{G} = \{t_0, t_1, \cdots, t_n\}$, which satisfies that

$$T - 1 = t_0 < t_1 < \cdots < t_n = T.$$

For simplicity, we assume that the observation point is equidistant, which is frequently used in the literature, i.e. $t_i - t_{i-1} \equiv \delta$ for any $1 \leq i \leq n$. *Realized variance* ($rvar$) for the T-th day is defined as:

$$rvar := \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})^2. \tag{2}$$

In the absence of microstructure noise, when n goes to infinity,

$$rvar \rightarrow_p [X, X]_T. \tag{3}$$

Similarly we define *realized skewness* ($rskew$) and *realized kurtosis* ($rkurt$) for the T-th day as:

$$rskew := \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})^3, \tag{4}$$

and

$$rkurt := \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})^4. \tag{5}$$

In addition, realized skewness and realized kurtosis can be normalized as:

$$nrskew := \frac{rskew}{rvar^{3/2}}, \tag{6}$$

$$nrkurt := \frac{rkurt}{rvar^2}. \tag{7}$$

When n goes to infinity, normalized realized skewness and kurtosis have the following limits in probability [13]:

$$nrskew \rightarrow_p \frac{\sum_{T-1 \leq s \leq T} (\Delta_s X)^3}{(\int_{T-1}^{T} \sigma_s^2 ds + \sum_{T-1 \leq s \leq T} (\Delta_s X)^2)^{3/2}}, \tag{8}$$

$$nrkurt \rightarrow_p \frac{\sum_{T-1 \leq s \leq T} (\Delta_s X)^4}{(\int_{T-1}^{T} \sigma_s^2 ds + \sum_{T-1 \leq s \leq T} (\Delta_s X)^2)^2}. \tag{9}$$

We call the above realized moments the naive ones, for example, the naive realized skewness.

2.3 Pre-averaging estimator

In practice, it is commonly admitted that microstructure noise is inherent in the high-frequency price process so that we are not able to observe directly X_{t_i}, but Y_{t_i}, a noisy version of X_{t_i} at times $i = 0, \cdots, n$. In this paper, we assume that

$$Y_{t_i} = X_{t_i} + \epsilon_{t_i}, \tag{10}$$

where ϵ_{t_i} are i.i.d. microstructure noise with mean zero and variance η^2, and ϵ_{t_i} and X_{t_i} are independent of each other.

To reduce the effect of microstructure noise, the pre-averaging method [13] is used within blocks of length k_n. In the i-th block, the pre-averaging return is constructed as

$$\Delta^n_{i,k_n} Y(g) = \sum_{j=1}^{k_n} g(\frac{j}{k_n})(Y_{t_{i+j}} - Y_{t_{i+j-1}}), \tag{11}$$

and

$$\Delta^n_{i,k_n} \bar{Y}(g) = \sum_{j=1}^{k_n} (g(\frac{j}{k_n}) - g(\frac{j-1}{k_n}))^2 (Y_{t_{i+j}} - Y_{t_{i+j-1}})^2, \tag{12}$$

with a non-negative piece-wise differentiable Lipschitz function g, satisfying $g(x) = 0$ when $x \notin (0,1)$ and $\bar{g}(p) = \int_0^1 g^p(x)dx > 0$. From the empirical analysis, we use $g(x) = \min\{x, 1-x\}$ for $0 \le x \le 1$, which is often used in the literature. As a result, the pre-averaging realized measures are constructed as follows:

$$rvar := \frac{1}{\bar{g}(2)} \left(\frac{1}{k_n} \sum_{i=1}^{n-k_n} (\Delta^n_{i,k_n} Y(g))^2 - \frac{1}{2k_n} \sum_{i=1}^{n-k_n} (\Delta^n_{i,k_n} \bar{Y}(g)) \right), \tag{13}$$

$$rskew := \frac{1}{\bar{g}(3)} \left(\frac{1}{k_n} \sum_{i=1}^{n-k_n} (\Delta^n_{i,k_n} Y(g))^3 \right), \tag{14}$$

$$rkurt := \frac{1}{\bar{g}(4)} \left(\frac{1}{k_n} \sum_{i=1}^{n-k_n} (\Delta^n_{i,k_n} Y(g))^4 \right), \tag{15}$$

and

$$nrskew := \frac{rskew}{rvar^{3/2}}, \quad nrkurt := \frac{rkurt}{rvar^2}. \tag{16}$$

In the presence of microstructure noise, following [13], the above pre-averaging estimators have the following limits in probability when $k_n, n \to \infty$ and $k_n/n \to 0$:

$$nrskew \to_p \frac{\sum_{T-1 \le s \le T} (\Delta_s X)^3}{(\int_{T-1}^T \sigma_s^2 ds + \sum_{T-1 \le s \le T} (\Delta_s X)^2)^{3/2}}, \tag{17}$$

$$nrkurt \to_p \frac{\sum_{T-1 \le s \le T} (\Delta_s X)^4}{(\int_{T-1}^T \sigma_s^2 ds + \sum_{T-1 \le s \le T} (\Delta_s X)^2)^2}. \tag{18}$$

3. EMPIRICAL DATA ANALYSIS

Some existing literatures indicate that realized skewness is an important factor in explaining future asset returns. However, these literatures mainly focus on the whole market and on the monthly or weekly scale [13, 1]. In this section, we test the cross-sectional forecasting performance of realized skewness for the individual stock and the daily scale. Furthermore, we examine whether higher moments, namely realized skewness and realized kurtosis, have any explaining power for the variances of the stock prices.

3.1 Data and exploratory analysis

Our empirical analysis is based on the transaction prices from Wharton Research Database (WRDS) for International Business Machines (IBM). The sample period starts on January 2, 2005 and ends on December 31, 2013 and the daily transaction records start from 09:30 to 16:00. We have a total of 2265 days for the stock. We only report the results for the pre-averaging estimator of the stock to save space, as the conclusion is generally the same for the naive estimator, which is computed by the summation of the corresponding power of the 5-minute log-returns. We check the robustness of the result of this section by exploring other stocks in Section 4. We find that the properties we see in this section generally apply.

We first conduct data cleaning with the procedures introduced in Brownlees and Gallo [5] and Barndorff-Nielsen et al. [2]. The steps are as follows:

1. Delete entries with a time stamp outside 9:30 - 16:00 when the exchange is open.
2. Delete entries with a transaction price equal to zero.
3. If multiple transactions have the same time stamp, use the median price.
4. Delete entries with prices which are outliers. Let $\{p_i\}_{i=1}^N$ be an ordered tick-by-tick price series. We call the i-th price an outlier if $|p_i - \bar{p}_i(m)| > 3s_i(m)$, where $\bar{p}_i(m)$ and $s_i(m)$ denote the sample mean and sample standard deviation of a neighborhood of m observations around i, respectively. For the beginning prices which may not have enough left hand side neighbors, we get $m-i$ neighbors from $i+1$ to $m+1$. Similar procedures are used for the ending prices. We take $m = 5$ here.

Daily returns are computed as the difference of the logrithm of the closing prices for the current and previous days. Realized moments are estimated by the pre-averaging method. We take $\Delta_n := 1/n = 1$ minute, $k_n = 10$ and $g(x) = \min(x, 1-x)$.

The descriptive statistics for IBM daily returns ($dret$), realized variance ($rvar$), realized skewness ($rskew$), realized kurtosis ($rkurt$), normalized realized skewness ($nrskew$) and normalized realized kurtosis ($nrkurt$) are shown in Table 1. In addition, their plots are shown in Figures 1 – 6.

In Table 1, we find that the daily returns exhibit a slightly negative skewness, and that realized skewness also shows some negative skewness, while all other measures show positive skewness. All measures show larger kurtosis than that of the normal distribution, i.e. fat tails. In addition, from Figures 1 to 6, we can see that realized variance exhibits the clear volatility clustering phenomenon. It seems that realized skewness and realized kurtosis show a similar pattern. Meanwhile, normalized realized skewness and normalized kurtosis seem more random and behave like white noise.

Table 1. Descriptive statistics for the daily return and realized moments

	$dret$	$rvar$	$rskew$	$rkurt$	$nrskew$	$nrkurt$
Maximum	1.11×10^{-1}	5.85×10^{-3}	7.94×10^{-5}	5.64×10^{-5}	1.45	2.61
Minimum	-8.78×10^{-2}	8.40×10^{-6}	-4.67×10^{-4}	5.02×10^{-12}	-1.26	0.04
Mean	2.88×10^{-4}	1.42×10^{-4}	-1.96×10^{-7}	3.58×10^{-8}	0.01	0.09
SD	1.40×10^{-2}	3.10×10^{-4}	1.01×10^{-5}	1.19×10^{-6}	0.13	0.09
Skewness	-0.17	9.00	-43.3	46.4	0.53	18.5
Kurtosis	9.23	117	2004	2188	14.7	487

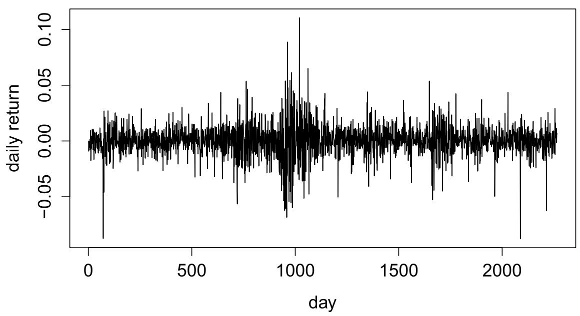

Figure 1. Daily log-returns of International Business Machines.

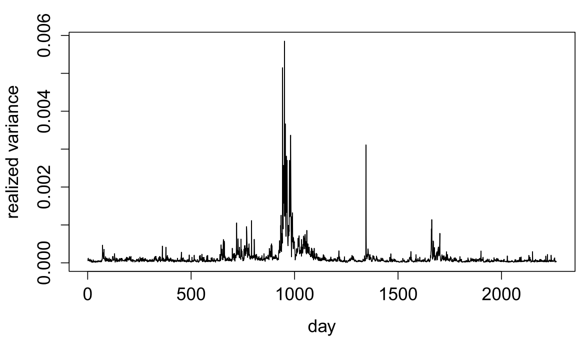

Figure 2. Daily realized variance of International Business Machines.

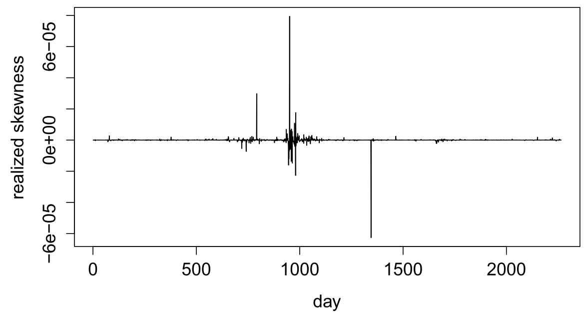

Figure 3. Daily realized skewness of International Business Machines.

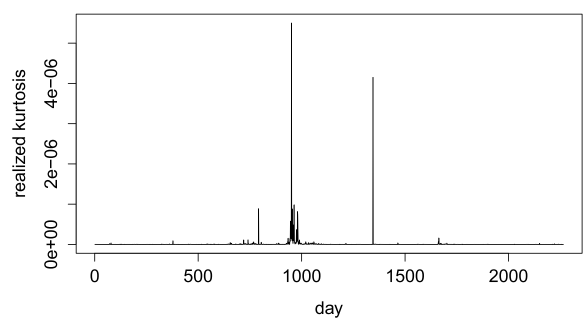

Figure 4. Daily realized kurtosis of International Business Machines.

We fit simple time series to the pre-averaging realized variance, realized skewness, realized kurtosis, etc. From the Ljung-Box test (Table 2), we conclude that normalized realized skewness and normalized realized kurtosis can be treated as white noise.

3.2 Predicting daily returns

As mentioned earlier, realized skewness has been thought to have explaining power for the future daily returns of the equity market. Now, we employ the regression models to investigate whether the conclusion holds for the individual stock and the daily horizon. Here, we regress daily returns with respect to the previous day's realized variance, realized skewness and realized kurtosis (and their normalized counterparts).

We employ the following regression models:

(19) $$dret_{t+1} = \alpha_0 + \alpha_1 rvar_t + \alpha_2 rskew_t + \alpha_3 rkurt_t + \epsilon_{t+1},$$

and

(20) $$dret_{t+1} = \alpha_0 + \alpha_1 rvar_t + \alpha_2 nrskew_t + \alpha_3 nrkurt_t + \epsilon_{t+1}.$$

The above equations are predictive regressive models for forecasting one-day ahead daily returns with different realized measures. Tables 3 and 4 show the result of the regression models.

From Tables 3 and 4, one may find that there is no linear relationship with daily returns and previous-day realized variance, normalized realized skewness and normalized realized kurtosis. Meanwhile, the coefficients of realized skewness and kurtosis are significant. Realized skewness shows a

Table 2. Ljung-Box test of the series with lag 10

	rvar	*rskew*	*rkurt*	*nrskew*	*nrkurt*
p-value	$< 2.2 \times 10^{-16}$	9.9×10^{-10}	$< 2.2 \times 10^{-16}$	0.49	0.80

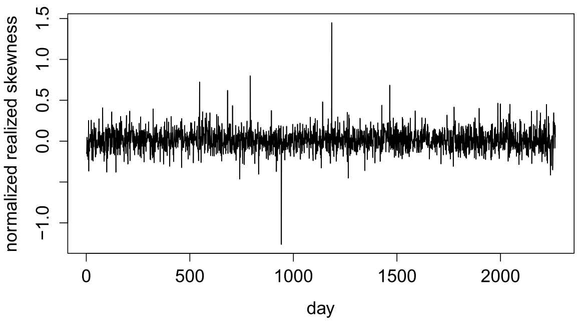

Figure 5. Daily normalized realized skewness of International Business Machines.

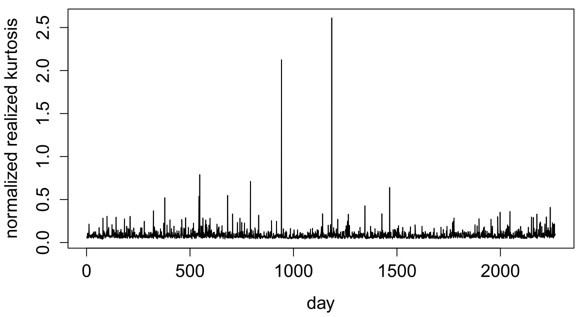

Figure 6. Daily normalized realized kurtosis of International Business Machines.

positive effect on daily returns. This finding is different from some literature that realized skewness has a negative effect for daily returns. There is a possibility that daily scale and individual aspects greatly make a difference in explaining this. To confirm the above findings, the daily return is regressed against the previous-day's realized skewness solely for those 70 stocks mentioned in Section 4 by the following regression equation:

$$dret_{t+1} = \alpha_0 + \alpha_1 rskew_t + \epsilon_{t+1}. \tag{21}$$

It is found that realized skewness has significant explanatory power for the future daily returns in only 17 out of 70 cases. For those stocks with large capitalization size, realized skewness shows significant effects with about 20% chance; while for medium and small sizes, the chances become 10% and 45%, respectively. As a result, the explanatory power may be associated with the capitalization size of the stock that realized skewness shows little effect on the future daily return for stocks with relatively large cap size. This may be because the stocks with a smaller cap size are relatively

Table 3. Regression model (19) for daily return with previous day realized moments

	$\alpha_1(rvar)$	$\alpha_2(rskew)$	$\alpha_3(rkurt)$	R^2	F test
Estimate	-1.86	4.54×10^2	4.15×10^3	0.01	
SE	1.16	1.09×10^2	9.65×10^2		
p-value	0.11	2.94×10^{-5}	1.75×10^{-5}		3.20×10^{-4}
		$***$	$***$		

Table 4. Regression model (20) for daily return with previous day realized moments

	$\alpha_1(rvar)$	$\alpha_2(nrskew)$	$\alpha_3(nrkurt)$	R^2	F test
Estimate	0.33	5.84×10^{-4}	1.93×10^{-3}	0.0003	
SE	1.01	2.53×10^{-3}	3.75×10^{-3}		
p-value	0.74	0.81	0.61		0.91

less liquid and they absorb the incoming news at a slower rate. From above analysis, we conclude that for individual stocks, especially those with large capitalization sizes, realized skewness does not have enough forecasting power for the one-day ahead daily returns.

3.3 Predicting the variance

3.3.1 Regression analysis

In this subsection, we regress realized variance against the previous day return, realized skewness, realized kurtosis and so on. We would like to determine the variables which exist to explain the realized variances. We employ the following regression models.

$$rvar_{t+d} = \alpha_0 + \alpha_1 dret_t + \alpha_2 rskew_t + \alpha_3 rkurt_t + \epsilon_{t+d}, \tag{22}$$

and

$$rvar_{t+d} = \alpha_0 + \alpha_1 dret_t + \alpha_2 nrskew_t + \alpha_3 nrkurt_t + \epsilon_{t+d}, \tag{23}$$

where $d = 1$ for this subsection. The above equations are predictive regression models for forecasting one-day ahead realized variance with different realized measures. Tables 5 and 6 show the result of the regression models.

Tables 5 and 6 show that realized kurtosis is extremely significant in explaining future realized variances, while the daily return also has a significant effect, indicating a possible leverage effect. We will explore this effect more in later subsections. The coefficients estimated are positive for realized kurtosis, suggesting that larger price jumps lead to larger price fluctuations in the near future. The coefficient of realized skewness is also significant. We see that other

Table 5. Regression model (22) for realized variance with previous day realized moments, $d = 1$

	$\alpha_1(dret)$	$\alpha_2(rskew)$	$\alpha_3(rkurt)$	R^2	F test
Estimate	-3.39×10^{-3}	5.82	69.0	0.03	
SE	5.10×10^{-4}	2.30	19.4		
p-value	3.75×10^{-11}	0.01	3.83×10^{-4}		7.03×10^{-14}
	* * *	*	* * *		

Table 6. Regression model (23) for realized variance with previous day realized moments, $d = 1$

	$\alpha_1(dret)$	$\alpha_2(nrskew)$	$\alpha_3(nrkurt)$	R^2	F test
Estimate	-3.18×10^{-3}	-3.58×10^{-5}	1.27×10^{-5}	0.02	
SE	5.82×10^{-4}	6.51×10^{-5}	8.31×10^{-5}		
p-value	5.13×10^{-8}	0.58	0.88		2.11×10^{-9}
	* * *				

Table 7. Regression model (22) for realized variance with previous day realized moments, $d = 2$

	$\alpha_1(dret)$	$\alpha_2(rskew)$	$\alpha_3(rkurt)$	R^2	F test
Estimate	-6.11×10^{-3}	1.63×10^{1}	1.57×10^{2}	0.08	
SE	8.90×10^{-4}	3.05	2.55×10^{1}		
p-value	1.18×10^{-11}	1.05×10^{-7}	1.11×10^{-9}		$< 2.2 \times 10^{-16}$
	* * *	* * *	* * *		

Table 8. Regression model (22) for realized variance with previous day realized moments, $d = 5$

	$\alpha_1(dret)$	$\alpha_2(rskew)$	$\alpha_3(rkurt)$	R^2	F test
Estimate	-4.23×10^{-3}	9.87	1.09×10^{2}	0.04	
SE	9.06×10^{-4}	3.10	2.60×10^{1}		
p-value	3.39×10^{-6}	1.49×10^{-3}	3.19×10^{-5}		7.32×10^{-10}
	* * *	**	* * *		

realized measures show no forecasting power; for example, normalized realized kurtosis. This may be because the size of the jumps has been normalized "out."

To confirm this explanation of realized kurtosis on one-day-ahead realized variance, we regress realized variance against the previous-day's realized kurtosis, solely for those 70 stocks mentioned in Section 4, by the following regression equation:

$$rvar_{t+1} = \alpha_0 + \alpha_1 rkurt_t + \epsilon_{t+1}. \tag{24}$$

We find that realized kurtosis has significant explanatory power for the future realized variance in 65 out of 70 cases. For those stocks with a large capitalization size, realized kurtosis shows a significant effect with about 90% chance; while for medium and small sizes, the chance becomes 90% and 100%, respectively. In consequence, realized kurtosis has explanatory power for the one-day ahead volatility of stocks for different capitalization sizes.

3.3.2 Regression analysis with longer horizon

We have seen that realized kurtosis has forecasting power for the one-day ahead realized variances and we explore whether the same conclusion holds for a longer forecasting horizon. As a result, the same regression model (22) with the longer horizon is shown here, with horizons of 2-days, 5-days and 22-days, i.e. $d = 2, 5, 22$, corresponding to a trading period of two days, one week and one month, respectively. The results are shown in Tables 7 – 9.

When the horizon for prediction becomes longer and longer, the predicting power of realized kurtosis on realized variance becomes less and less, which is really natural. We find that when $d = 2$, realized kurtosis has an extremely significant effect on the regressand; and when $d = 22$, the effect becomes very significant. Moreover, the coefficient for realized kurtosis is always positive, which is the same in the case when $d = 1$. In addition, when the time horizon is long ($d = 22$), realized skewness shows no explaining power for the future realized variance, while the daily return still has some effect on the future realized variance.

Additionally, from Tables 5, 7 and 8, it is odd to see that the R-Square of shortest-horizon is the smallest. We have done the same experiment for those stocks mentioned in Section 4. It is found that for those stocks that realized kurtosis shows significant explanatory power for future realized variance (65 out of 70), in most situations (61 out of 65), the

Table 9. Regression model (22) for realized variance with previous day realized moments, $d = 22$

	$\alpha_1(dret)$	$\alpha_2(rskew)$	$\alpha_3(rkurt)$	R^2	F test
Estimate	-2.12×10^{-3}	4.35	8.10×10^{1}	0.04	
SE	9.08×10^{-4}	3.11	2.60×10^{1}		
p-value	0.02	0.16	1.92×10^{-3}		6.44×10^{-10}
	*		**		

Table 10. Regression model (25) for realized variance with previous day trading volume

	$\alpha_1(tvol)$	R^2	F test
Estimate	4.81×10^{-11}	0.17	
SE	2.05×10^{-12}		
p-value	$< 2.2 \times 10^{-16}$		$< 2.2 \times 10^{-16}$
	* * *		

Table 11. Regression model (26) for realized variance with previous day trading volume and realized kurtosis

	$\alpha_1(tvol)$	$\alpha_2(rkurt)$	R^2	F test
Estimate	4.05×10^{-11}	1.96×10^{1}	0.18	
SE	2.05×10^{-12}	5.25		
p-value	$< 2.2 \times 10^{-16}$	1.99×10^{-4}		$< 2.2 \times 10^{-16}$
	* * *	* * *		

R-Square shows normal pattern, i.e. the predicting power of realized kurtosis decays when the horizon of prediction in lengthened. As a result, we think that the abnormal pattern appears for the stock IBM may be because there is a strong spurious relationship between realized kurtosis and two-day ahead realized variance or it is just bad luck.

3.3.3 Adding other covariates

In practice, there exist some other covariates used to explain and/or forecast the price volatility, for example, trading volume of the stock within a period, and negative daily returns. Trading volume is a covariate used to explain the volatility in the finance field, which tends to be larger in the case of higher volatility. In addition, negative daily returns reflect the so-called leverage effect. In this subsection, we regress realized variance against the previous day realized kurtosis, positive and negative daily returns and trading volume. We find that realized kurtosis still exhibits significant explaining power in the presence of other covariates.

Let $tvol$ denote trading volume, $dret^{+}$ positive daily return, and $dret^{-}$ negative daily return. We employ the following regression models for analysis:

$$rvar_{t+1} = \alpha_0 + \alpha_1 tvol_t + \epsilon_{t+1}, \tag{25}$$

and

$$rvar_{t+1} = \alpha_0 + \alpha_1 tvol_t + \alpha_2 rkurt_t + \epsilon_{t+1}. \tag{26}$$

The estimated results are shown in Tables 10 and 11.

We observe from Table 10 that the previous day's trading volume has a significant positive relationship with realized variance. The appearance of a large trading volume is probable due to the new information released to the market. As a result, the volatility of the stock becomes larger in this situation. When we add realized kurtosis to the regression model, it is found from Table 11 that in the presence of trading volume, realized kurtosis exhibits an extremely significant effect on the future realized variance, and trading volume is still significant. The reason may be that realized kurtosis already contains some of the information contained in the trading volume. One possible explanation is that realized kurtosis measures the jumps within the day, which may correspond to the large trading volume for one particular trade. Consequently, these two measures may have a relationship with each other. Moreover, the addition of the realized kurtosis improves the R^2 from 0.17 to 0.18.

We next consider the effect of positive and negative daily returns towards realized variance. We employ the following equations:

$$rvar_{t+1} = \alpha_0 + \alpha_1 dret_t^{+} + \alpha_2 dret_t^{-} + \epsilon_{t+1}, \tag{27}$$

and

$$rvar_{t+1} = \alpha_0 + \alpha_1 dret_t^{+} + \alpha_2 dret_t^{-} + \alpha_3 rkurt_t + \epsilon_{t+1}. \tag{28}$$

The results are shown in Tables 12 and 13.

It is found from Table 12 that both positive and negative daily returns have significant effect on realized variance. When we add realized kurtosis in the regression equation,

Table 12. Regression model (27) for realized variance with previous day signed daily returns

	$\alpha_1(dret^+)$	$\alpha_2(dret^-)$	R^2	F test
Estimate	1.03×10^{-2}	-1.67×10^{-2}	0.20	
SE	7.84×10^{-4}	7.74×10^{-4}		
p-value	$< 2.2 \times 10^{-16}$	$< 2.2 \times 10^{-16}$		$< 2.2 \times 10^{-16}$
	***	***		

Table 13. Regression model (28) for realized variance with previous day signed daily returns and realized kurtosis

	$\alpha_1(dret^+)$	$\alpha_2(dret^-)$	$\alpha_3(rkurt)$	R^2	F test
Estimate	1.02×10^{-2}	-1.65×10^{-2}	1.34×10^{1}	0.21	
SE	7.83×10^{-4}	7.77×10^{-4}	5.17		
p-value	$< 2.2 \times 10^{-16}$	$< 2.2 \times 10^{-16}$	9.37×10^{-3}		$< 2.2 \times 10^{-16}$
	***	***	**		

Table 14. The comparison of the in-sample and out-of-sample prediction performance

In-sample analysis: Regression model (30)				
	$\alpha_1(rvar)$	$\alpha_2(rkurt)$	R^2	F test
Estimate	7.45×10^{-1}	-5.59×10^{1}	0.466	
SE	1.80×10^{-2}	4.64		
p-value	$< 2.2 \times 10^{-16}$	$< 2.2 \times 10^{-16}$		$< 2.2 \times 10^{-16}$
	***	***		

we can see from Table 13 that realized kurtosis is very significant for realized variance. However, R^2 does not greatly improve, from 0.20 to 0.21.

In addition, we employ the following regression models to see whether realized kurtosis still maintains some explaining power when the first lag of realized variance is included:

$$rvar_{t+1} = \alpha_0 + \alpha_1 rvar_t + \epsilon_{t+1}, \tag{29}$$

and

$$rvar_{t+1} = \alpha_0 + \alpha_1 rvar_t + \alpha_2 rkurt_t + \epsilon_{t+1}. \tag{30}$$

We see from Table 14 that in Equation (30), realized kurtosis shows an extremely significant effect in the presence of realized variance. Additionally, the model, including realized kurtosis, has less error in its forecasting. Therefore, realized kurtosis indicates some additional information besides the past history of realized variance.

Finally, we combine all the pertinent covariates together in regression model (31).

$$\begin{aligned} rvar_{t+1} &= \alpha_0 + \alpha_1 rkurt_t + \alpha_2 tvol_t + \alpha_3 dret_t^+ + \alpha_4 dret_t^- \\ &+ \alpha_5 rvar_t + \epsilon_{t+1}. \end{aligned} \tag{31}$$

We can see from Table 15 that realized kurtosis still exhibits an extremely significant effect on the future realized variance.

3.3.4 Out-of-sample forecasting

We investigate whether adding realized kurtosis into the regressive models improves the out-of-sample forecasting accuracy. We focus on the following three regression models:

$$rvar_{t+1} = \alpha_0 + \alpha_1 dret_t + \alpha_2 rskew_t + \epsilon_{t+1}, \tag{32}$$

$$rvar_{t+1} = \alpha_0 + \alpha_1 tvol_t + \epsilon_{t+1}, \tag{33}$$

$$rvar_{t+1} = \alpha_0 + \alpha_1 dret_t^+ + \alpha_2 dret_t^- + \epsilon_{t+1}, \tag{34}$$

and

$$rvar_{t+1} = \alpha_0 + \alpha_1 rvar_t + \epsilon_{t+1}. \tag{35}$$

After adding the realized kurtosis, the regression models become:

$$rvar_{t+1} = \alpha_0 + \alpha_1 dret_t + \alpha_2 rskew_t + \alpha_3 rkurt_t + \epsilon_{t+1}, \tag{36}$$

$$rvar_{t+1} = \alpha_0 + \alpha_1 tvol_t + \alpha_2 rkurt_t + \epsilon_{t+1}, \tag{37}$$

$$rvar_{t+1} = \alpha_0 + \alpha_1 dret_t^+ + \alpha_2 dret_t^- + \alpha_3 rkurt_t + \epsilon_{t+1}, \tag{38}$$

and

$$rvar_{t+1} = \alpha_0 + \alpha_1 rvar_t + \alpha_2 rkurt_t + \epsilon_{t+1}. \tag{39}$$

We compare the out-of-sample prediction performance of Model (32) against Model (36), Model (33) against Model (37), Model (34) against Model (38), and Model (35) against Model (39). We use two metrics to do the comparison, the normalized mean square error (MSE) and the Clark and McCraken (CM) test. The normalized MSE

Table 15. Regression model (31) for realized variance with all covariates

	$\alpha_1(rkurt)$	$\alpha_2(tvol)$	$\alpha_3(dret^+)$	$\alpha_4(dret^-)$	$\alpha_5(rvar)$	R^2	F test
Estimate	-5.23$\times10^1$	-4.05$\times10^{-14}$	2.26$\times10^{-3}$	-7.39$\times10^{-3}$	6.65$\times10^{-1}$	0.50	
SE	4.60	2.12$\times10^{-12}$	7.34$\times10^{-4}$	7.41$\times10^{-4}$	2.09$\times10^{-2}$		
p-value	$< 2\times10^{-16}$	0.98	2.09$\times10^{-3}$	$< 2\times10^{-16}$	$< 2\times10^{-16}$		$< 2.2\times10^{-16}$
	$***$		**	$***$	$***$		

Table 16. The comparison of the out-of-sample prediction performance

	MSE_1	MSE_2	CM statistic	0.90	0.95	0.99
			Model (32) and Model (36)			
(36) versus (32)	3.44	3.46	1.74	0.322	0.489	0.908
			Model (33) and Model (37)			
(37) versus (33)	1.37	1.39	1.68	0.322	0.489	0.908
			Model (34) and Model (38)			
(38) versus (34)	5.50	5.61	1.96	0.322	0.489	0.908
			Model (35) and Model (39)			
(39) versus (35)	0.58	0.68	1.94	0.322	0.489	0.908

is defined as:

$$MSE = \frac{\sum(\text{predicted realized variance} - \text{true value})^2}{\sum(\text{true value})^2}.$$

Moreover, the CM statistic refers to the Clark and McCracken [10] Encompassing test, which compares the out-of-sample prediction ability of nested models. The larger the CM statistic, the better the latter model is. The result is shown in Table 16. The column with MSE_1 exhibits the MSE's of Models (36), (37), (38) and (39), while the column with MSE_2 for Models (32), (33), (34) and (35). The last three columns give the 90th, 95th and 99th percentiles of the distribution of the statistic derived under the null, which is from Clark and McCraken (2001) and can be treated as the critical values. The sample period is the first 2000 days, from January 3rd, 2005 to December 11th, 2012, and the forecast period is the next 200 days, from December 12th, 2012 to September 27th, 2013.

From Table 16, we see that the models including realized kurtosis have less or equal MSE than the ones without realized kurtosis. For instance, in the comparison of Model (34) and Model (38), Model (38) with realized kurtosis has an MSE of 5.50, which is smaller than 5.61 of Model (34). Consequently, realized kurtosis helps produce less forecasting errors. Furthermore, the out-of-sample performance of Models (36) - (39) is significantly better than Models (32) - (35), respectively, as the CM statistics are all larger than 0.908, the 99th percentile.

3.4 Conclusions

From the above in-sample and out-of-sample analysis, we conclude that realized kurtosis does have some forecasting power for the future daily volatility within a short time period. In addition, the in-sample improvement is larger than the out-of-sample one with realized kurtosis, which suggests the presence of nonlinear relationships and linear regression may not be well suited. Moreover, we see that the R^2 for the regression model is small, which indicates that in practice, it is difficult to do the prediction very precisely. That's why the addition of realized kurtosis does not seem to be generating great improvement in the out-of-sample analysis. For those stocks covered in Section 4, it is found that R^2 for the regression model changes from case to case and the maximum value of R^2 is 0.36, which is quite high. In the literature, we also witness low R^2 in the regression analysis for this field, for example, Bollerslev and Zhou [4] and Liu et al. [13]. However, the significance of the variable, realized kurtosis, is still valid in this case.

Furthermore, we use the stepwise regression with the Akaike information criterion (AIC) to choose the best regressors for the one-day ahead realized variance. We use positive daily return, negative daily return, realized skewness, realized kurtosis and trading volume as regressors for initiation. The resulting regressors are positive daily return and realized kurtosis, which also indicates the significance of realized kurtosis. We conduct the same procedure to those 70 stocks in Section 4.

4. ROBUSTNESS OF THE RESULT

From the above empirical analysis, we find that realized kurtosis possesses explaining power for the future daily volatility of the stock IBM, which is a proxy of realized variance. In this section, we explore the performance of realized kurtosis for more stocks and give some explanations for the explaining power.

To check whether the performance of realized kurtosis is robust across different stocks, we employ 70 stocks from NYSE, where 20 of S&P 600 (small cap), 20 of S&P 400 (mid cap) and 30 of S&P 500 (large cap). The stocks are the following:

1. Small: AVD, CLW, DEL, DKS, EGO, EGP, ENS, FMC, GEO, HCSG, IART, MATW, NNN, POWL, PKE, PZZA, RAD, SONC, TE and TRN.
2. Medium: ALGN, ASH, BKH, CBSH, CIM, DF, DSX, DV, JLL, KBR, KMT, LAMR, LII, MDC, MPW, NEU, OIS, RRD, TDW and WSM.
3. Large: AIG, AXP, BA, C, CAT, CVX, DD, DIS, GE, GS, HD, HON, IBM, JNJ, JPM, KO, MCD, MMM, MRK, NKE, PFE, PG, SBUX,T, TRV, UNH, UTX, V, VZ and WMT.

The details of these stock variables can be found in Table A.1 in the appendix. These stocks are from different sectors of the industry, namely technology, healthcare, industrial goods, consumer goods, basic materials, utilities, financial, services and so on, thus providing sufficient samples of the stocks in the USA market. Note that the tables in this Section are all very long, so we put them in the appendix, making it easy to read. The sampling period starts from January 2nd, 2009 and ends on December 31st, 2013.

To explore the forecasting power of realized kurtosis on the daily volatility, we adopt the regression models in Section 3, namely Equations (22), (26), (28) and (31). In addition, we use the stepwise method with the AIC to select adequate covariates for the regression of the future realized variance.

In Table A.2 of the appendix, we show the result for the performance of realized kurtosis. The first column exhibits the stock variables. The second to fifth columns show the significance of the coefficient for realized kurtosis in Equation (22), (26), (28) and (31), respectively. For example, the second entry in the first row, $dret + rskew$, indicates that the second column shows the forecasting performance of realized kurtosis in the presence of daily return and realized skewness. The numbers shown in the table stand for the different levels of significance. "0" stands for not significant, with p-value bigger than 0.1; "0.5" stands for marginally significant, with p-value between 0.05 and 0.1; "1" stands for significant, with p-value between 0.01 and 0.05; "2" stands for very significant, with p-value between 0.001 and 0.01; and "3" stands for extremely significant, with p-value less than 0.001. The last column indicates the result of the covariate selection. The potential covariates include realized skewness, realized kurtosis, trading volume and positive and negative daily returns. The number "1" in some entries of the last column means that no covariate is selected.

It is found from Table A.2 that in general realized kurtosis always shows significant explaining power when trading volume and negative return are absent. When trading volume or negative return are added, realized kurtosis sometimes exhibits no significant power in forecasting the volatility. However, for the small capitalization group in the covariate selection, 16 of 20 stocks include the realized kurtosis. For the medium size group, 13 of 20 stocks include realized kurtosis as a covariate. For large companies, 25 of 30 stocks include realized kurtosis. In other words, the performance of realized kurtosis is stable and is significant in a majority of cases, with respect to stocks with different sizes and different sectors.

In the above regression analysis, we use realized kurtosis as a covariate for realized variance. However, the orders of the two variables are not the same from Equation (18). As a result, we take the square root of realized kurtosis to make it of the same order with realized variance, and then conduct the same regression analysis. The result is shown in Table A.3. Remember that all $rkurt$'s in Table A.3 stand for the square root of realized kurtosis.

It is found that the square root of realized kurtosis performs better in explaining the future daily volatility, as shown by Columns 2 to 5 of Table A.3. When all possible covariates are included, the square root of realized kurtosis shows a significant effect on the future realized variance in 65 of the 70 cases. Furthermore, for the small capitalization group in the covariate selection, 20 of 20 stocks include realized kurtosis. For the medium size group, 19 of 20 stocks include realized kurtosis as a covariate. For large companies, 27 of 30 stocks include realized kurtosis. We see that the performance of the square root of realized kurtosis is even better than that of realized kurtosis, for the firms with all sizes. In addition, it seems that the square root of realized kurtosis, signed returns, and trading volume are able to account for almost all of the explaining power for the future daily volatility, in general.

From Equation (18), realized variance converges to the sum of two parts, the integrated variance and the sum of the square of jumps. Additionally, realized kurtosis converges to the sum of the fourth power of price jumps. As a consequence, it is natural to wonder if realized kurtosis possesses some explaining power for realized variance, as they both have a jump component. However, this seems not to be the case. In Huang and Tauchen [11], the authors separate the two components of realized variance to check for the contribution of the jump component. In the empirical study, they find that the jump component only accounts for 7% of stock market price variance, which indicates that it is the continuous component that dominates. As a result, it is worth considering why the previous day's price jump affects the continuous price fluctuation. In addition, even if the jump is really important for daily volatility, we know that the jump of the price always corresponds to the unexpected arrival of new information, so that it is unnatural that the previous-day's jump has strong forecasting power on future jumps. Nevertheless, we separate out the continuous component of realized variance, estimated by the bipower variation. We conduct all the regression models with respect to the bipower variation to see whether the explanatory power of the square

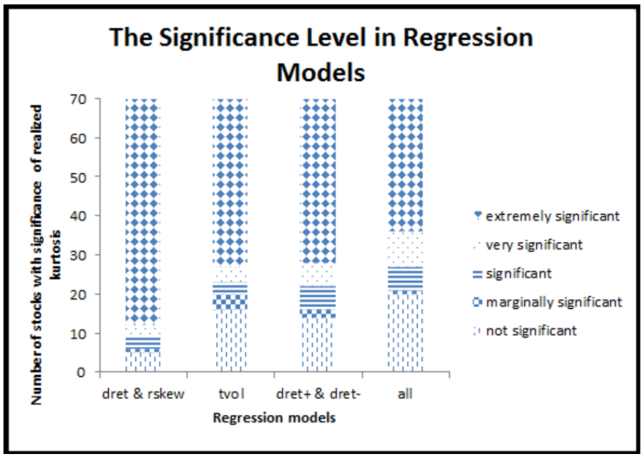

Figure 7. The forecasting power of realized kurtosis towards realized variance in different regression models.

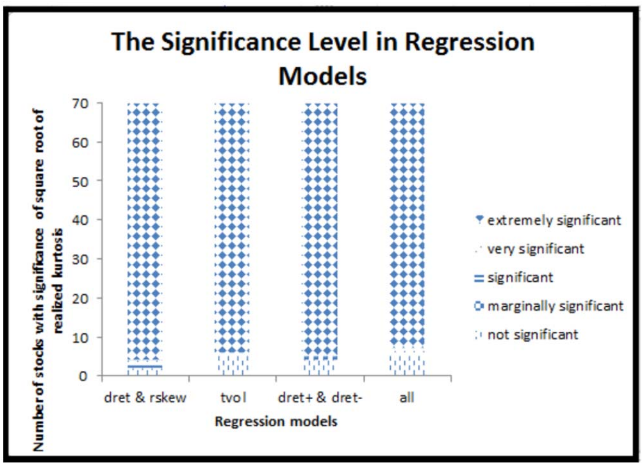

Figure 8. The forecasting power of the square root of realized kurtosis towards realized variance in different regression models.

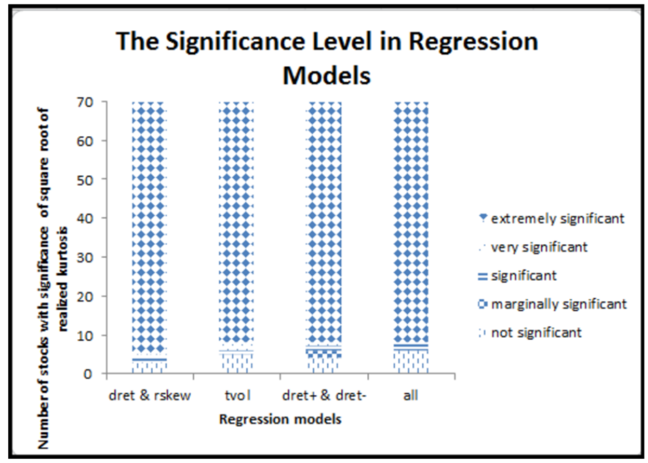

Figure 9. The forecasting power of the square root of realized kurtosis towards bipower variation in different regression models.

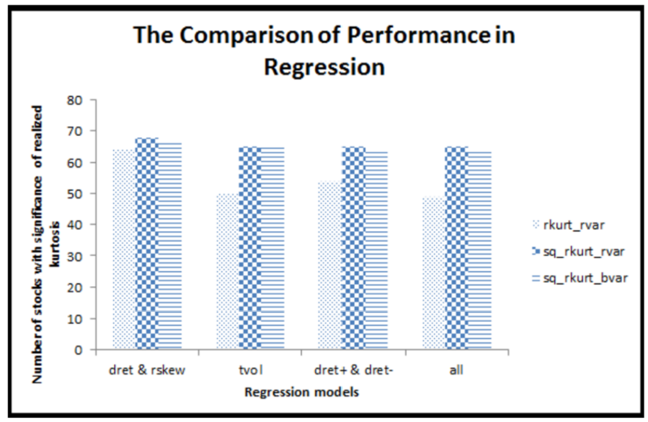

Figure 10. The comparison of the forecasting power for different regression models.

root of realized kurtosis remains. The result is shown in Table A.4.

It is found that on average, the square root of realized kurtosis remains powerful in predicting the future daily bipower variation, shown by Columns 2 to 5 of Table A.4. The performance is comparable with that under realized variance, which is also shown in the covariate selection column. For stocks with small capitalization, 20 of 20 stocks include realized kurtosis. For the medium size, 20 of 20 stocks include realized kurtosis as a covariate. For large size companies, 27 of 30 stocks include realized kurtosis.

In summary, we use the following graphs to illustrate the overall performance of the regressors. In Figures 7-9, each bar represents the forecasting performance in one regression model involving the variables shown below the bar. Different filling patterns stand for the level of significance for realized kurtosis or the square root of realized kurtosis, shown by the legend, for example, horizontal line means realized kurtosis is significant, with p-value between 0.01 and 0.05. Figure 10 compares the performance of realized kurtosis in forecasting realized variance, the square root of realized kurtosis in forecasting realized variance, and the square root of realized kurtosis in forecasting the bipower variation.

From the graphs, we see that the (square root of) realized kurtosis always performs better in the absence of trading volume. Nevertheless, the overall performance of the (square root of) realized kurtosis is satisfactory. The worst case is when we add all variables in the regression models for realized kurtosis, where about 90% of the stocks still indicate the significance of realized kurtosis in forecasting the future realized variance. From Figure 10, we see that the performance of the square root of realized kurtosis forecasting realized variance is always the best among the three no matter what the regression model is.

It appears that realized kurtosis, whether taking the square root or not, measures the price jump size to some extent. It contains information for the future daily volatility, possibly due to the following reasons. Firstly, Merton [14] points out that the continuous part of the stock volatility may be due to the change in the economic anticipation and the temporary imbalance between supply and demand. Moreover, sometimes the price jumps are incidental, which correspond to the newly arrived news to the market. When the market is unable to digest the news efficiently, the news effect aggregates and the jump should have some forecasting power for the future volatility. This is also indicated by the result that the long horizon forecasting performance of realized kurtosis becomes worse, since the news has been digested gradually by the market after a long time. Additionally, sometimes the price jump is artificial, i.e. the price jump is due to manipulation by some large financial institutions. In this case, the market will fluctuate corresponding to the reaction of the public, and partially the follow-up actions by the institutions. Consequently, the future volatility has some relationship with previous price jumps. Certainly, some other undiscovered reasons remain.

5. CONCLUSIONS

In this paper, we analyze whether higher realized moments have explaining power for future daily returns or realized variance. We find that realized skewness does not provide enough evidence of the effect on daily returns, in contrast with the literature. On the other hand, realized kurtosis exhibits significant forecasting power for the future realized variance in a short period. Furthermore, the square root of realized kurtosis shows even better forecasting ability. In addition, the two high-order realized moments appear to be highly asymmetrical and fat-tailed, which may have great practical importance for financial asset returns and can be studied further.

However, it is found that in the regression analysis, the R^2 is relatively low (the largest R^2 detected is 0.36 among all stocks), even though the effect of realized kurtosis is significant, which indicates that the prediction cannot be very precise when linear regression models are used. This phenomenon also suggests that some nonlinear regression models may be used to fit the relationships between realized variance and realized kurtosis, which will be pursued in the future.

ACKNOWLEDGEMENTS

We are grateful to the referee and co-editor for their careful reading and useful comments. The second author's research is partially supported by HK GRF Grant (17332416). The third author's research is partially supported by HK GRF Grant (17303315).

APPENDIX

Table A.1. Information of the selected stocks. In the Sector column, Tech stands for Technology, Bmat for Basic materials, Util for Utilities, Fin for Financial, Serv for Services, Igood for Industrial goods, Cgood for Consumer goods, and Heal for Healthcare.

Stock	Firm	Sector
AVD	American Vanguard Corp	Bmat
CLW	Clearwater Paper Corp	Cgood
DEL	Deltic Timber Corp	Igood
DKS	Dick's Sporting Goods	Serv
EGP	Eastgroup Properties Inc	Fin
EGO	Eldorado Gold	Bmat
ENS	EnerSys	Igood
FMC	FMC	Bmat
GEO	The GEO Group	Fin
HCSG	Healthcare Services Group	Serv
IART	Integra Life Sciences Holdings Corp	Heal
MATW	Matthews International Corp	Serv
NNN	National Retail Properties	Fin
PKE	Park Electrochemical Corp	Tech
POWL	Powell Industries Inc	Igood
PZZA	Papa John's International Inc	Serv
RAD	Rite Aid	Serv
SONC	Sonic Corp	Serv
TE	TECO Energy	Util
TRN	Trinity Industries	Serv
ALGN	Align Technology Inc	Heal
ASH	Ashland Inc	Bmat
BKH	Black Hills Corp	Util
CBSH	Commerce Bancshare	Fin
CIM	Chimera Investment	Fin
DF	Dean Foods	Cgood
DSX	Diana Shipping	Serv
DV	DeVry Education	Serv
JLL	Jones Lang LaSalle	Fin
KBR	KBR Inc	Serv
KMT	Kennametal Inc	Igood
LAMR	Lamar Advertising Corp	Fin
LII	Lennox International Inc	Igood
MDC	MDC Holdings Inc	Igood
MPW	Medical Properties Trust	Fin
NEU	Newmarket corp	Bmat
OIS	Oil states international inc	Bmat
RRD	R.R. Donnelley & Sons Company	Serv
TDW	Tidewater	Bmat
WSM	Williams-Sonoma Inc	Serv
AIG	American International Group	Fin
AXP	American Express	Fin
BA	Boeing	Igood
C	Citigroup	Fin
CAT	Catepillar	Igood
CVX	Chevron	Bmat
DD	E.I. du Pont de Nemours	Bmat
DIS	Walt Disney	Serv
GE	General Electric	Igood
GS	Goldman Sachs	Fin
HD	Home Depot	Serv
HON	Honeywell International	Igood
IBM	International Business Machines	Tech
JNJ	Johnson & Johnson	Heal
JPM	JPMorgan	Fin
KO	Coca-Cola	Cgood
MCD	McDonald's	Serv
MMM	3M	Igood
MRK	Merch & Co	Heal
NKE	NIKE	Cgood
PFE	Pfizer	Heal
PG	Procter & Gamble	Cgood
SBUX	Starbucks Corp	Serv
T	AT&T	Tech
TRV	Travelers	Fin
UNH	UnitedHealth	Heal
UTX	United Technologies	Igood
V	Visa	Fin
VZ	Verizon Communications	Tech
WMT	Wal-Mart Stores	Serv

Table A.2. The performance of realized kurtosis towards realized variance. The second to fifth columns show the significance of the coefficient for realized kurtosis in Equations (22), (26), (28) and (31), respectively. The numbers shown in the table stand for the different level of significance. "0" for p-value bigger than 0.1; "0.5" for p-value between 0.05 and 0.1; "1" for p-value between 0.01 and 0.05; "2" for p-value between 0.001 and 0.01; and "3" for p-value less than 0.001. The last column indicates the result of the covariate selection. The potential covariates include realized skewness, realized kurtosis, trading volume and positive and negative daily returns. The number "1" in some entries of the last column means that no covariate is selected.

Stock	$dret$ & $rskew$	$tvol$	$dret^+$ & $dret^-$	all	covariate selection
AVD	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
CLW	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
DEL	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
DKS	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
EGO	3	1	0	0	$dret^+ + dret^- + tvol + rkurt$
EGP	3	3	3	3	$tvol$
ENS	3	1	1	0	$dret^+ + dret^- + tvol + rskew$
FMC	3	3	3	2	$dret^+ + dret^- + tvol + rskew + rkurt$
GEO	1	0.5	0.5	0.5	$dret^+ + dret^- + rskew$
HCSG	3	2	3	1	$dret^- + tvol + rkurt$
IART	0	2	0	0	$dret^+ + dret^-$
MATW	3	3	2	2	$dret^+ + dret^- + tvol + rkurt$
NNN	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
PKE	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
POWL	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
PZZA	3	3	3	3	$dret^- + tvol + rskew + rkurt$
RAD	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
SONC	3	3	3	3	$dret^+ + dret^- + rkurt$
TE	2	3	2	2	$dret^+ + dret^- + tvol + rskew + rkurt$
TRN	3	0.5	3	1	$dret^+ + dret^- + tvol + rskew + rkurt$
ALGN	2	0.5	0	0	$dret^+ + dret^-$
ASH	3	0	3	0	$dret^+ + dret^- + tvol + rskew + rkurt$
BKH	3	3	0	0	$dret^+ + dret^- + tvol + rskew$
CBSH	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
CIM	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
DF	3	1	2	1	$tvol + rkurt$
DSX	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
DV	2	0	0	0	$dret^- + tvol + rskew$
JLL	3	3	3	2	$dret^+ + dret^- + tvol + rskew + rkurt$
KBR	3	0.5	0	0	$dret^+ + dret^- + tvol + rskew + rkurt$
KMT	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
LAMR	3	0	0	0	$dret^+ + dret^- + tvol + rskew + rkurt$
LII	3	0	0	0	$dret^+ + dret^- + tvol + rskew$
MDC	0	0	0	0	$dret^+ + dret^- + tvol + rskew$
MPW	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
NEU	3	0	0	0	$dret^+ + dret^- + tvol$
OIS	3	3	3	3	$dret^+ + dret^- + rkurt$
RRD	3	0	1	2	$dret^+ + dret^- + tvol + rskew + rkurt$
TDW	1	0	0	0	$dret^+ + dret^- + tvol + rskew$
WSM	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
AIG	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
AXP	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
BA	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
C	3	3	2	2	$dret^+ + dret^- + rskew + rkurt$
CAT	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
CVX	0	0	0.5	0	$dret^+ + dret^- + tvol + rskew$
DD	3	2	3	2	$dret^+ + dret^- + tvol + rskew + rkurt$
DIS	3	0	1	0	$dret^+ + dret^- + tvol + rskew + rkurt$
GE	3	0	3	0	$dret^- + tvol + rskew$
GS	1	0	0	0	$dret^+ + dret^- + tvol + rskew$
HD	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
HON	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
IBM	3	3	3	3	$dret^- + rkurt$
JNJ	3	3	2	2	$dret^+ + dret^- + tvol + rskew + rkurt$
JPM	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
KO	3	3	3	3	$dret^- + tvol + rkurt$
MCD	3	3	3	3	$dret^- + rskew + rkurt$
MMM	0.5	0	0	0	$tvol$
MRK	3	3	1	1	$dret^+ + dret^- + rskew + rkurt$
NKE	3	3	3	3	$dret^- + tvol + rkurt$
PFE	3	0	1	0	$dret^- + rskew + rkurt + tvol$
PG	0	0	0	0	$tvol$
SBUX	3	3	3	3	$dret^- + rskew + rkurt + tvol$
T	3	3	3	2	$dret^+ + dret^- + tvol + rskew + rkurt$
TRV	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
UNH	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
UTX	3	2	2	1	$dret^- + tvol + rskew + rkurt$
V	0	0	1	1	$dret^+ + dret^- + tvol + rskew + rkurt$
VZ	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
WMT	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$

Table A.3. The performance of the square root of realized kurtosis towards realized variance

Stock	$dret$ & $rskew$	$tvol$	$dret^+$ & $dret^-$	all	covariate selection
AVD	3	3	3	3	$dret^+ + dret^- + rkurt + tvol$
CLW	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
DEL	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
DKS	3	3	3	3	$dret^- + rkurt$
EGO	3	3	3	3	$dret^+ + dret^- + rkurt$
EGP	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
ENS	3	3	3	3	$dret^+ + dret^- + rkurt$
FMC	3	3	3	3	$dret^+ + dret^- + rkurt$
GEO	3	3	3	3	$dret^+ + dret^- + rkurt$
HCSG	3	3	3	3	$dret^- + rkurt$
IART	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
MATW	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
NNN	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
PKE	3	3	3	3	$dret^- + tvol + rskew + rkurt$
POWL	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
PZZA	3	3	3	3	$dret^- + tvol + rskew + rkurt$
RAD	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
SONC	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
TE	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
TRN	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
ALGN	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
ASH	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
BKH	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
CBSH	3	2	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
CIM	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
DF	3	3	3	3	$tvol + rkurt$
DSX	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
DV	3	3	3	2	$dret^- + tvol + rskew + rkurt$
JLL	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
KBR	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
KMT	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
LAMR	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
LII	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
MDC	0	0	0	0	$dret^+ + dret^- + tvol + rskew$
MPW	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
NEU	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
OIS	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
RRD	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
TDW	3	3	3	2	$dret^+ + dret^- + tvol + rskew + rkurt$
WSM	3	3	3	3	$dret^+ + dret^- + rkurt$
AIG	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
AXP	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
BA	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
C	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
CAT	3	3	3	3	$dret^- + tvol + rskew + rkurt$
CVX	0	0	0.5	0	$dret^+ + dret^- + tvol + rskew$
DD	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
DIS	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
GE	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
GS	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
HD	3	3	3	3	$dret^- + tvol + rkurt$
HON	3	3	3	3	$dret^- + tvol + rskew + rkurt$
IBM	3	3	3	3	$dret^- + rkurt$
JNJ	3	3	3	3	$dret^- + tvol + rskew + rkurt$
JPM	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
KO	3	3	3	3	$tvol + rkurt$
MCD	3	3	3	3	$dret^- + rskew + rkurt$
MMM	2	0	0	0	$tvol$
MRK	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
NKE	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
PFE	3	3	3	3	$dret^- + tvol + rskew + rkurt$
PG	1	0	0	0	$rskew + rkurt$
SBUX	3	3	3	3	$dret^- + rkurt$
T	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
TRV	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
UNH	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
UTX	3	3	3	3	$dret^- + tvol + rskew + rkurt$
V	3	0	0	0	$dret^+ + dret^- + tvol$
VZ	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
WMT	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$

Table A.4. The performance of the square root of realized kurtosis with respect to the bipower variation

Stock	$dret$ & $rskew$	$tvol$	$dret^+$ & $dret^-$	all	covariate selection
AVD	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
CLW	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
DEL	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
DKS	3	3	3	3	$dret^- + tvol + rskew + rkurt$
EGO	3	3	3	3	$dret^+ + dret^- + rkurt$
EGP	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
ENS	3	3	3	3	$dret^+ + dret^- + rkurt$
FMC	3	3	3	3	$dret^+ + dret^- + rkurt$
GEO	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
HCSG	3	3	3	3	$dret^- + rkurt$
IART	3	3	3	3	$dret^- + rkurt$
MATW	3	3	3	3	$dret^- + tvol + rkurt$
NNN	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
PKE	3	3	3	3	$dret^- + tvol + rskew + rkurt$
POWL	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
PZZA	3	3	3	3	$dret^- + tvol + rskew + rkurt$
RAD	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
SONC	3	3	3	3	$dret^+ + dret^- + rkurt$
TE	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
TRN	3	3	3	3	$dret^- + tvol + rskew + rkurt$
ALGN	3	2	1	1	$dret^- + rkurt$
ASH	3	3	3	3	$dret^- + tvol + rskew + rkurt$
BKH	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
CBSH	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
CIM	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
DF	3	3	3	3	$tvol + rkurt$
DSX	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
DV	3	3	3	3	$dret^- + tvol + rkurt$
JLL	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
KBR	3	3	3	3	$dret^- + tvol + rskew + rkurt$
KMT	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
LAMR	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
LII	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
MDC	0	0	0	0	$dret^+ + dret^- + tvol + rskew + rkurt$
MPW	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
NEU	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
OIS	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
RRD	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
TDW	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
WSM	3	3	3	3	$dret^+ + dret^- + rkurt$
AIG	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
AXP	3	3	3	3	$dret^+ + dret^- + tvol + rkurt$
BA	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
C	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
CAT	3	3	3	3	$dret^- + tvol + rskew + rkurt$
CVX	0	0	0	0	1
DD	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
DIS	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
GE	3	3	3	3	$dret^- + tvol + rkurt$
GS	2	0	0.5	0	$tvol + rskew + rkurt$
HD	3	3	3	3	$dret^- + tvol + rkurt$
HON	3	3	3	3	$dret^- + tvol + rskew + rkurt$
IBM	3	3	3	3	$dret^- + rkurt$
JNJ	3	2	2	1	$rkurt$
JPM	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
KO	3	3	3	3	$tvol + rkurt$
MCD	3	3	3	3	$dret^- + rskew + rkurt$
MMM	1	0	0	0	$tvol$
MRK	3	3	3	3	$dret^+ + dret^- + rskew + rkurt$
NKE	3	3	3	3	$dret^- + rskew + rkurt$
PFE	3	3	3	3	$dret^- + tvol + rskew + rkurt$
PG	0	0	0	0	1
SBUX	3	3	3	3	$dret^- + rkurt$
T	3	3	3	3	$dret^- + rskew + rkurt$
TRV	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
UNH	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
UTX	3	3	3	3	$dret^- + tvol + rskew + rkurt$
V	3	1	0.5	0	$dret^+ + dret^- + tvol + rskew + rkurt$
VZ	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$
WMT	3	3	3	3	$dret^+ + dret^- + tvol + rskew + rkurt$

Received 18 April 2017

REFERENCES

[1] Amaya, D., Christoffersen, P., Jacobs, K., and Vasquez, A. (2013). Do realized skewness and kurtosis predict the cross-section of equity returns? Available at SSRN 1898735. MR2597033

[2] Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008). Realized kernels in practice: Trades and quotes. *Econometrics Journal*, 4, 1–32. MR0431445

[3] Black, F. (1976). Studies in stock price volatility changes. *Proceedings of the 1976 Business and Economics Section, American Statistical Association*, 177–181. MR2275997

[4] Bollerslev, T., and Zhou, H. (2006). Volatility puzzles: a simple framework for gauging return-volatility regressions. *Journal of Econometrics*, 131(1), 123–150. MR2307498

[5] Brownlees, C. T., and Gallo, G. M. (2006). Financial econometric analysis at ultra-high frequency: Data handling concerns. *Computational Statistics & Data Analysis*, 51(4), 2232–2245.

[6] Buchatskaya, E., and Khrapov, S. (2013). Realized skewness predicts jump variation. Working paper.

[7] Chan, C. C., and Fong, W. M. (2006). Realized volatility and transactions. *Journal of Banking & Finance*, 30(7), 2063–2085.

[8] Choi, Y, and Lee, S. S. (2006). Realized skewness and future stock returns: The role of information.

[9] Christie, A. (1982). The stochastic behavior of common stock variances-values, leverage and interest rate effects. *Journal of Financial Economics*, 10, 407–432. MR1864213

[10] Clark, T. E., and McCracken, M. W. (2001). Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics*, 105(1), 85–110.

[11] Huang, X., and Tauchen, G. (2005). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics*, 3(4), 456–499. MR2662350

[12] Jacod, J., Podolskij, M., and Vetter, M. (2010). Limit theorems for moving averages of discretized processes plus noise. *The Annals of Statistics*, 38(3), 1478–1545.

[13] Liu, J., Wang, K., and Liu, Z. (2013). Measuring the realized skewness in noisy semi-martingale with jumps using high frequency data. *Working paper.*

[14] Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1), 125–144.

[15] Neuberger, A. (2012). Realized skewness. *Review of Financial Studies*, 25(11), 3423–3455.

[16] Schneider, P., Wagner, C., and Zechner, J. (2016). Low risk anomalies?

Keren Shen
Department of Statistics and Actuarial Science
The University of Hong Kong
Pokfulam
Hong Kong
E-mail address: rayshenkr@hku.hk

Jianfeng Yao
Department of Statistics and Actuarial Science
The University of Hong Kong
Pokfulam
Hong Kong
E-mail address: jeffyao@hku.hk

Wai Keung Li
Department of Statistics and Actuarial Science
The University of Hong Kong
Pokfulam
Hong Kong
E-mail address: hrntlwk@hku.hk