Journal of Banking & Finance 64 (2016) 52–70

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

A test of efficiency for the S&P 500 index option market using the generalized spectrum method ☆

Henry H. Huang [a,*], Kent Wang [b], Zhanglong Wang [c]

[a] Department of Finance, National Central University, Taiwan, ROC
[b] UQ Business School, University of Queensland, Brisbane, Australia
[c] Guosen Securities Co., Ltd., China

ARTICLE INFO

Article history:
Received 20 August 2014
Accepted 20 November 2015
Available online 24 December 2015

JEL Classification:
G13
G14
C22
C14

Keywords:
Model-Free Forward Variance
Spectral density test
Index jump
Market efficiency

ABSTRACT

This paper examines the efficiency of the S&P 500 options market by testing the martingale properties of the Model-Free Forward Variance (MFFV) time series using the Generalized Spectral Test (GST). Based on a sample from January 1, 1996 to May 31, 2010, our tests show robust evidence that the S&P 500 options market is not efficient. By examining the subsamples before and after the 2008 financial crisis, we find this options market inefficiency is mainly driven by the outbreak of the subprime crisis. Our diagnostic tests further indicate that this inefficiency is due to the skewness-in-mean effect of forward variance. Specifically, the skewness-in-mean effect is weakened once we account for the S&P 500 index jump effects. Hence, we can establish a link between jumps and options market inefficiency. Finally, we find that the lagged skewness of the forward variance can help forecasting the forward variance both in-sample and out-of-sample. The economic significance of this forecasting ability is further highlighted by the performance of a trading strategy based on forward variance. In sum, out study provides robust evidence and a trading implication on testing the S&P 500 options market efficiency.

1. Introduction

One of the most enduring issues in finance and economics is whether the financial markets are efficient or satisfy the no-arbitrage condition. The violation of the no-arbitrage condition, which is referred by Longstaff (1995) as the "martingale restriction", implies the existence of arbitrage opportunities and/or market frictions. Inefficiency of financial markets is to some extent analogous to the predictability of asset returns and is still a subject of ongoing debates and intensive empirical research.[1] Testing for the martingale property and understanding the predictability of options market are thus an issue fundamentally important to all market participants.

While the literature on testing market efficiency has mainly focused on the stock market, efficiency of the options market has received less attention and the empirical evidence has been inconclusive. Harvey and Whaley (1992) demonstrate the forecastability of the implied volatility from the S&P 100 index option. By using a trading strategy based on daily out-of-sample volatility projections, they show that this strategy does not yield abnormal returns even when transaction costs are considered. Their results suggest that the predictable time-varying volatility is consistent with market efficiency. Jiang and Tian (2012) (JT (2012), hereafter) empirically test the random-walk hypothesis by examining the martingale property of the Model-Free Forward Variance (MFFV), and their results support the random-walk hypothesis, suggesting that the options market is efficient. [2] However, several papers have provided indirect evidence on options market inefficiency. For example, Cremers and Weinbaum (2010) show that deviations from put-call parity exist and can be used to predict both negative and positive future stock abnormal returns. Xing et al. (2010) conclude that volatility skew, defined as the difference between the implied

☆ We thank the editor (Carol Alexander) and two anonymous referees for their insightful and helpful comments. Henry H. Huang wishes to thank the financial support from Ministry of Science and Technology in Taiwan with project number [NSC 101-2628-H-008-002-].
* Corresponding author.
E-mail addresses: hongming@ncu.edu.tw (H.H. Huang), k.wang@business.uq.edu.au (K. Wang), carl57@yahoo.cn (Z. Wang).
[1] See for example, Ang and Bekaert (2007), Campbell and Thompson (2008), Goyal and Welch (2008), Rapach et al. (2010), and Sekkel (2011).

[2] In the random-walk hypothesis, the asset price change is serially uncorrelated with past information as defined in Campbell et al. (1997).

http://dx.doi.org/10.1016/j.jbankfin.2015.11.007
0378-4266/© 2015 Elsevier B.V. All rights reserved.

volatilities of out-of-the-money puts and at-the-money calls, has significant predictive power for future cross-sectional equity returns. Their results are consistent with the notion that informed traders with negative news prefer to trade out-of-the-money put options, and that the options market is slow in incorporating the information embedded in volatility smirks. Du (2012) also provides empirical evidence that trading activities in the option markets indeed provide useful information on future stock market movements.

In this paper we test the option markets efficiency hypothesis and provide direct rather than indirect evidence. We re-examine the hypothesis of options market efficiency based on the martingale property of MFFV, as in JT (2012). In an efficient market, the time-series of forward variance form a martingale process. The martingale property implies that any change in forward variance must be driven by the arrival of unexpected news and is thus orthogonal to past information. Through linear orthogonality tests, JT (2012) find that the forward-variance series satisfy the martingale property, and changes in forward variance follow a random walk. They thus conclude the informational efficiency of the S&P 500 index options market.

The conclusion of informational efficiency in the options market in JT (2012) may result from their linear econometric specifications. For example, the orthogonality tests of JT (2012) only use several special examples of variables in the conditional information set, such as lagged forward variance, lagged spot variance, and lagged variance spreads. These variables may not contain sufficient information content about the dynamic behavior of forward variance. Although their final auto-regressive regression test utilizes a larger number of potentially more informative conditioning variables, it is based on a finite-dimensional information set and therefore may neglect some dependence structure in the conditional mean at omitted lags. Furthermore, there exists a gap between a Martingale Difference Sequence (MDS) and a White Noise (WN) process for forward variance changes. However, the orthogonality tests in JT (2012) only can test WN, not MDS. [3] This is indeed one of the problems we attempt to overcome in this current study. In addition, JT (2012) use a sample prior to the 2008 financial crisis, and given the magnitude and deep influence of the crisis, we cannot ignore such a big event in testing the option-market efficiency hypothesis.

Hong (1999) and Hong and Lee (2003) propose the Generalized Spectral Test (GST) that can distinguish MDS and WN by accounting for non-linearity in economic and financial data. As pointed out by Hong and Lee (2003), traditional methods, such as Box and Pierce Q-statistic test and Ljung and Box (LB) test, only investigate serially uncorrelated relationships rather than a martingale difference. GST can capture all pairwise dependencies in the conditional mean over various lags, requires no moment conditions, and allows conditional heteroscedasticity of unknown form. Given such favorable properties and power over standard orthogonality tests, GST has been used in a wide range of empirical studies. For instance, using GST, Hong and Lee (2003) find significant predictability for future exchange rate returns, Hong et al. (2004) specify the best spot interest-rate model in terms of out-of-sample forecasting performance, McPherson and Palardy (2007) note evidence against the martingale difference hypothesis in five out of nine international daily stock index returns, Yang et al. (2008) test the martingale property of the exchange market and find that martingale behavior cannot be rejected for Euro exchange rates with major currencies, and Hong et al. (2012) examine the predictability of corporate bond index returns and present evidence of significant serial and cross-serial dependence with a complex non-linear structure in daily investment-grade and high-yield bond returns. However, while GST has been extensively used for testing the expectations hypothesis in the equity, interest rate, and foreign currency markets, no such studies have been carried out in the options market.

In this paper we combine the GST with regression-based methods to test the martingale restriction. Using the S&P 500 index options data and the JT (2012) method to construct the Model-Free Forward Variance, we find that during the period from January 1996 to October 2010, the short-maturity forward variance series are rejected as being MDS by GST, while the long-maturity groups cannot be rejected as being MDS. We further examine this issue by dividing our sample into the periods before the subprime crisis and after the subprime crisis. We find that the options market inefficiency is only observed after the subprime crisis. This finding is consistent with JT (2012), who find that the options market is efficient based on the sample before the subprime crisis. To dig into the sample after the subprime crisis, we further use the diagnostic tests of GST and uncover that the skewness-in-mean effect of daily changes in forward variance is the factor that causes the violation of the martingale restriction for the short-maturity groups. The regression-based analysis subsequently confirms the results of GST and demonstrates a negative correlation between the daily change in forward variance and the third-order conditional moment.

To ensure the robustness of the results, we consider a new forward variance measure based on Bakshi et al. (2003) and Du and Kapadia (2012), which is more resilient to the price jumps of the underlying assets, to re-examine the options market efficiency in our full sample. Although the series of short-maturity forward variance are closer to a martingale compared to the original forward variance using this new measure, they are still rejected by GST to be MDS. Moreover, the skewness-in-mean effect still exists in the new forward variance series. For the long-maturity forward variance groups, we confirm the conclusion of JT (2012) that the options market is efficient, even under the more powerful GST. However, in contrast to JT (2012), we reject the hypothesis that the short-maturity forward variance groups are MDS. The underlying reason is that the random-walk test employed by JT (2012) may not be general enough in testing the non-linear dependence of the forward variance series. Hence, we try to overcome this problem by using a more general testing approach.

As suggested by JT (2012), the options market inefficiency maybe due to the effect of market illiquidity. The trading volume of long-maturity options is less than the trading volume of short-maturity options, and this deterioration of market liquidity may results in options market inefficiency. To rule out this possibility, we control the market liquidity, as proxied by trading volumn, in each maturity group and conduct the regression analysis again. We find that market illiquidity does not affect our original conclusion.

In the literature, skewness in stock returns has long been related to price jumps.[4] Jumps in the prices of underlying assets may contribute to the skewness-in-mean effect and thus lead to informational inefficiency, as the jump information cannot be instantaneously and completely incorporated by the market. To address this issue, we use the Barndorff-Nielsen and Shephard (2006) method based on 5-min high frequency data to identify the S&P500 index jumps and find that the skewness-in-mean effect is weakened when we control for the jump effect. This finding has

[3] As Hong and Lee (2003) explain, there exists a non-trivial gap between MDS and WN. The former implies the latter, but not vice versa. A non-linear time series can have zero autocorrelation, but a non-zero mean conditional on its past history. For example, the non-linar moving average process $Y_t = be_{t-1}e_{t-2} + e_t$ and the bi-linear autoregressive process $Y_t = be_{t-1}Y_{t-2} + e_t$, where e_t is i.i.d. These are WN processes, but there exist predictable non-linearities in mean and they are not MDS.

[4] See for example, Rubinstein (1973), Kraus and Litzenberger (1976), Kraus and Litzenberger (1983), Merton (1976), and Das and Sundaram (1997).

far reaching implications for understanding the relationship between jumps and skewness, as well as for analyzing the role that jump risk plays in market friction and return prediction. Our findings are also consistent with previous literature which finds that tail risk measures do predict future excess return (see Kelly (2014) and Bollerselv and Todorov (2011)).

To further illustrate the applications of our testing results, we create a forecasting model that uses lagged third-order moments of forward variance to forecast one-day-ahead forward variance. Under the null hypothesis that the options market is efficient, the time series of forward variance should be captured by the random-walk model. By comparing with the forecasting performance of a random-walk model, our third-order-moment based model significantly outperforms. This suggests that lagged skewness of forward variance indeed contains useful information for predicting future forward variance, and once again confirms that the options market is inefficient.

To motivate the economic importance of our empirical findings and the practical application of our forecasting model, we construct a trading strategy based on our third-order moment based forecasting model. Compared with the performance of a trading strategy based on the random-walk model, our third-order-moment based trading strategy significantly outperforms after the financial crisis in 2008, especially in the short-maturity groups. This finding provides S&P 500 index option investors a profitable trading strategy and presents an economically prominent contribution of our paper.

We organize this paper as follows. Section 2 provides an explanation of our empirical methodologies, including the GST method and data processing procedures. Section 3 presents and discusses the main test results. Section 4 offers robustness checks and further investigations. Concluding remarks are given in Section 5. The detailed introductions of GST and Barndorff-Nielsen and Shephard (2006) jump detection method are presented in the Appendix A.

2. Methodology and data description

2.1. Methodology

2.1.1. Forward variance and the martingale restriction

Forward variance in the options market is an analogous concept to the forward interest rate in the bond market. According to JT (2012), forward variance measures the market's projection of variance over a certain time period in the future:

$$v(t;T_1,T_j) \equiv \frac{1}{T_j - T_1} E_t^{\mathbb{Q}}\left[\int_{T_1}^{T_j}\left(\frac{dS_\tau}{S_\tau}\right)^2 |\Omega_t\right] \quad (1)$$

where t is the current date, j denotes the corresponding maturity group, $t \leqslant T_1 \leqslant T_j, S_\tau$ is the price of the underlying asset on day τ, T_1 and T_j are two future dates, the probability measure $\mathbb{Q}$ is the risk-neutral probability measure, and Ω_t is the information set on day t.

If the options market is efficient, then Eq. (2) should be held for any time t and any pairs of future dates T_1 and T_j, or Eq. (3) equivalently:

$$v(t;T_1,T_j) = E_t^{\mathbb{Q}}[v(s;T_1,T_j)|\Omega_t] \quad (2)$$

$$E_t^{\mathbb{Q}}[v(s;T_1,T_j) - v(t;T_1,T_j)|\Omega_t] = 0 \quad (3)$$

For all $t < s \leqslant T_1$. This equation shows the change in forward variance is a MDS and orthogonal to the conditional information set Ω_t. Therefore, in subsequent empirical tests, we focus on the daily change in forward variance:

$$\triangle v(t;T_1,T_j) \equiv v(t;T_1,T_j) - v(t-1;T_1,T_j) \quad (4)$$

There are two problems in the martingale restriction tests based on the forward variance $v(t;T_1,T_j)$ of the options market. First, the martingale restriction in Eq. (2) holds under the risk-neutral probability measure, while empirical tests are performed under the objective probability measure. The difference between the two probability measures is the variance risk premium.We overcome this problem by following the argument of JT (2012), who present that the effect of the variance risk premium is negligible at daily frequency. Fama (1991) made a similar argument on the validity of the event study methodology, stating that asset price changes over short time intervals are mainly driven by information shocks and are immune from the risk premium effect. Based on this argument, our test mitigates the contamination of variance risk premium.

Second, even if the time series of daily change in forward variance are WN, they do not necessarily satisfy the martingale restriction in Eq. (3). From the perspective of the non-linear time series, there exists a non-trivial gap between MDS and WN. Hong and Lee (2003) give a non-linear example to show that MDS implies WN, but not vice versa. The random-walk test can easily neglect these non-linear dependent structures and lagged higher-order information. However, these problems can be solved using GST.

2.1.2. Martingale difference sequence test based on GST

The early literature on testing MDS is based on linear measures of dependence and a finite-dimensional conditioning set, such as sample autocorrelations (Box–Pierce–Ljung's portmanteau tests) and the variance ratio test (Lo and Mackinlay, 1988). More recently, the literature has presented non-linear measures of dependence and infinite-dimensional conditioning sets (see Durlauf (1991), Hong (1999), Deo (2000), Hong and Lee (2003), Dominguez and Lobato (2003), and Escanciano and Mayoral (2010). In this study we use Hong and Lee (2003)'s GST to investigate serial dependence of forward variance changes with the belief that it has at least two benefits: first, it can capture any type of pairwise serial dependence over various lags, including those that could be missed by the power spectrum and higher-order spectrum like the bi-spectrum; second, it provides a class of easy-to-interpret diagnostic tests to gauge the possible sources of the deviation of the martingale restriction. Hong (1999) proposes a class of test statistics $M(m,l)$:

$$M(m,l) \equiv \left[\int \sum_{j=1}^{n-1} k^2(j/p)(n-j) \times |\widehat{\sigma}_j^{(m,l)}(u,v)|^2 dW_1(u)dW_2(v) - \widehat{C}_0^{(m,l)} \sum_{j=1}^{n-1} k^4(j/p)\right] \div \left[\widehat{D}_0^{(m,l)} \sum_{j=1}^{n-2} k^4(j/p)\right]^{1/2} \quad (5)$$

The centering and standardization factors can be defined as:

$$\widehat{C}_0^{(m,l)} \equiv \int \hat{\sigma}_0^{(m,m)}(u,-u)dW_1(u) \int \hat{\sigma}_0^{(l,l)}(v,-v)dW_2(v)$$

$$\widehat{D}_0^{(m,l)} \equiv 2\int |\hat{\sigma}_0^{(m,m)}(u,u')|^2 dW_1(u)dW_1(u') \int |\hat{\sigma}_0^{(l,l)}(u,u')|^2 dW_2(v)dW_2(v').$$

where $\mathbf{i} \equiv \sqrt{-1}, u, v \in (-\infty,\infty), k(\cdot)$ is a kernel function, p is a bandwidth, and $W_1(\cdot)$ and $W_2(\cdot)$ are positive non-decreasing weighting functions that set weights on zero equally. In additon, $\hat{\sigma}_j(u,v) = \hat{\varphi}_j(u,v) - \hat{\varphi}_j(u,0)\hat{\varphi}_j(0,v)$ is the empirical generalized covariance, where $\hat{\varphi}_j(u,v) \equiv (n-|j|)^{-1}\sum_{t=|j|+1}^{n} e^{\mathbf{i}(uY_t + vY_{t-|j|})}$ is the empirical pairwise characteristic function for $j = 0, \pm 1, \ldots$. $\hat{\sigma}_j^{(m,l)}$ is the partial derivative of the empirical generalized covariance, and $\hat{\sigma}_j^{(m,l)}(u,v) = \partial^{(m+l)}\hat{\sigma}_j(u,v)/\partial^m u \partial^l v$ for $m,l \geqslant 0$. When sample size is large, the test statistic $M(m,l)$ is asymptotically distributed as

$\mathbf{N}(0,1)$. When the sample size is finite and greater than 100, as Hong and Lee (2003) suggest, the wild bootstrap procedure with 300 bootstrap replications is adequate for calculating the p-value for the GST test.[5] The detailed description of GST is left to the Appendix A.

Following Hong and Lee (2003), we choose (m,l) = (1,0) to check if there exists any type of serial dependence. Once generic serial dependence is discovered using $M(1,0)$, we go further to choose $(m,l) = (1,\ l)$ for $l = 1,\ 2,\ 3,\ 4$, testing if $Cov(Y_t, Y_{t-j}^l) = 0$ for all $j > 0$. These settings verify whether correlation, ARCH-in-mean, skewness-in-mean, and kurtosis-in-mean effects are present. Table 1 lists the relevant spectral derivative tests and dependence types that can be detected.

In this study we first use GST to check whether the time series of daily changes in forward variance is a MDS. Our null hypothesis is that the daily change in forward variance on one day is not predictable by any prior information of the daily change in forward variance. If the null hypothesis is rejected, we then apply the diagnostic tests to determine the possible source of rejection. Finally, we use the traditional auto-regression methods to check whether the results obtained are consistent.

2.2. Data processing procedure

We pool and utilize several data sources for our empirical tests. Daily closing data for the SPX options are sourced from the Chicago Board Options Exchange (CBOE) and covers the period from January 1996 to October 2010.[6] Daily Treasury bill yields (our proxy for the risk-free rates) are obtained from the Federal Reserve Bulletin. We only retain options that have strictly positive bid quotes and whose bid prices are strictly smaller than the ask prices. The implied volatilities of options with zero bid quotes are always greater than 1 in the database, and thus are not included. Following the method in JT (2012), option quotes equal or less than $\frac{3}{8}$ and options with less than a week remaining to maturity are excluded from the sample. For market illiquidity and microstructure concerns, we exclude in-the-money options. Finally, options violating the boundary, monotonicity, and convexity conditions are also removed from the sample. This is to ensure that we only use a subset of all available option prices that are absent of arbitrage opportunities (e.g. Aï-Sahalia and Duarte (2003) and Carr and Madan (2005)).

We use the forward index level to identify the in-the-money and out-of-money options. Following previous research (for example, Jiang and Tian (2007) and JT (2012)), the forward index level F_0 is calculated from at-the-money options using the put-call parity:

$$F_0 = K_* + \exp(-rT)[C_t(K_*, T) - P_t(K_*, T)] \tag{6}$$

where $C_t(K_*, T)$ $(P_t(K_*, T))$ is the call (put) option price with strike price K_* and maturity T at time t. According to the CBOE procedure, the at-the-money strike price is determined as the strike price (K_*) at which the difference between the call and the put prices is the smallest. Given the forward index value, we select the out-of-money options by choosing call options with a strike price greater than K_* and put options with a strike price less than or equal to K_*. We also define the daily closing option prices as the midpoint of the closing bid and ask quotes. We calculate the spot variance by Eq. (7) and select a pair of arbitrary spot variances to get one forward variance by the transformation function (8):

$$\begin{aligned} v(t,T) &\equiv \frac{1}{T-t} E_t^{\mathbb{Q}}\left[\int_{T_1}^{T_2} \left(\frac{dS_\tau}{S_\tau}\right)^2 |\Omega_t\right] \\ &= \frac{2\exp[r(T-t)]}{T-t}\left[\int_0^{F_0} \frac{P_t(K,T)}{K^2}dK + \int_{F_0}^{\infty} \frac{C_t(K,T)}{K^2}dK\right] \end{aligned} \tag{7}$$

$$v(t;T_1,T_j) = \frac{(T_j - t)v(t,T_j) - (T_1 - t)v(t,T_1)}{T_j - T_1} \tag{8}$$

To calculate the spot variances as defined in Eq. (7), we need option prices across all strike prices over an infinite range $[0, +\infty]$. However, option prices across all strike prices over a finite range exist at discrete increments. As shown in Jiang and Tian (2005) and Jiang and Tian (2007), this problem can be solved using a curve-fitting method with appropriate extrapolation, and in a related study, Carr and Wu (2009) use both linear interpolation and flat extrapolation methods. We use the latter method to calculate MFFV with both out-of-money call and put options. Between the maximal and minimal available strike prices, we first linearly interpolate implied volatilities at different levels of moneyness, defined as $money = log(K/K_*)$. For values of *money* below the lowest available moneyness level, we use the implied volatility at the lowest strike price. For *money* above the highest available moneyness, we use the implied volatility at the highest strike price. Using this interpolation and extrapolation procedure, we generate a fine grid of implied volatility points with a strike range of +/- 3 standard deviations from the at-the-money strike price K_*. Given the fine grid of implied volatility quotes, we then compute the option prices that are not listed by the Black–Scholes–Merton (BSM) formula. After we calculate the spot variances, we transform two spot variances into one forward variance using Eq. (8). Table 2 shows the summary statistics of spot variances.

Consider the summary statistics of spot variances reported in Panel A of Table 2. The mean spot variances vary in a narrow range from 0.0549 to 0.0560 across the five maturity groups. Spot variance tends to be higher for options with longer maturity, suggesting a general upward-sloping variance term structure, which is consistent with the findings of JT (2012) on the term structure of model-free spot volatilities. However, the mean value here is relatively higher given our analysis includes the 2008 financial crisis period in which all maturities of spot variance were at historically high levels. The spot variance is not normally distributed and has an obvious high peak and fat tails since the skewness and kurtosis are extremely high across all the maturities. However, the kurtosis decreases rapidly with maturities. To further illustrate these leptokurtosis and fat-tail phenomena, we plot the time series of spot variances in Fig. 1. The spot variance in the first maturity group is higher than those in the other maturity groups during the 2008

Table 1
Summary of Generalized Spectral Tests (GST).

Test	Statistic $M(m,l)$	Weights (W_1, W_2)	Test function $\sigma_j^{(m,l)}(u,v)$	Notation
MDS	$M(1,0)$	(δ, W_0)	$cov(Y_t, e^{\mathbf{i}vY_{t-1}})$	M_1
Correlation	$M(1,1)$	(δ,δ)	$cov(Y_t, Y_{t-j})$	M_2
ARCH-in-mean	$M(1,2)$	(δ,δ)	$cov(Y_t, Y_{t-j}^2)$	M_3
Skewness-in-mean	$M(1,3)$	(δ,δ)	$cov(Y_t, Y_{t-j}^3)$	M_4
Kurtosis-in-mean	$M(1,4)$	(δ,δ)	$cov(Y_t, Y_{t-j}^4)$	M_5

Note: $\{Y_t\}_{t=1}^n$ is a time-series and $\mathbf{i} = \sqrt{-1}$. When $l = 0$, we set $W_2(\cdot) = W_0(\cdot)$, where $W_0(\cdot)$ is the $\mathbf{N}(0,1)$ CDF. When $l > 0$, we set $W_2(\cdot) = \delta(\cdot)$, where $\delta(\cdot)$ is the Dirac delta function.

[5] Please refer to section II.B in Hong and Lee (2003) for the simulation results of GST's finite sample performance.

[6] Note that our full sample size (from January 1996 to October 2010) used in the GST is 3732. The sample size before financial crisis (January 1996 to January 2008) is 3,019, and the sample size after the financial crisis (February 2008 to October 2010) is 713. The adequacy of sample sizes used in both full sample test and subsample tests for applying GST is verified by the simulation results of Hong and Lee (2003), and our simulation results upon request.

Table 2
Summary statistics of the JT (2012) model-free implied variances: January, 1996-October, 2010.

Panel A: spot variance					
	T_1	T_2	T_3	T_4	T_5
Mean_Days	22	52	89	152	237
Median_Days	22	52	84	144	234
Mean	0.0549	0.055	0.0557	0.0559	0.056
Median	0.0414	0.0441	0.0464	0.0479	0.0488
Std	0.0582	0.0494	0.0456	0.042	0.038
Skewness	4.5282	3.7337	3.3995	3.2087	2.835
Kurtosis	32.5108	23.3908	19.8805	18.3761	14.8926
AR(1)	0.9609	0.9785	0.9838	0.9870	0.9899
AR(2)	0.9301	0.9609	0.9706	0.9765	0.9819
AR(3)	0.9099	0.9499	0.962	0.9690	0.9757
AR(4)	0.8891	0.9363	0.9505	0.9595	0.9681
AR(5)	0.8837	0.9292	0.9435	0.9534	0.9627
Panel B: forward variance					
	Group1: T_1T_2	Group 2: T_1T_3	Group 3: T_1T_4	Group 4: T_1T_5	
Mean	0.0557	0.0562	0.0564	0.0567	
Median	0.0467	0.0481	0.0489	0.0492	
Std	0.0453	0.0427	0.0398	0.0364	
Skewness	3.2425	3.0507	2.8873	2.5658	
Kurtosis	18.6947	16.7189	15.5554	12.7199	
AR(1)	0.9824	0.9875	0.899	0.9919	
AR(2)	0.9689	0.9771	0.9813	0.9851	
AR(3)	0.9596	0.9697	0.9748	0.9797	
AR(4)	0.9487	0.9604	0.9672	0.9736	
AR(5)	0.9399	0.9526	0.9608	0.968	

Note: Panels A and B report the descriptive statistics of the daily model-free implied spot and forward variances for different maturity groups. The first maturity group of the spot variance consists of options with the nearest maturity month; the second maturity group consists of options with the next maturity month, and so on. Mean and median numbers of days to maturity are also reported for each maturity group. The table also reports the first five autocorrelations of the daily model-free implied variance.

financial crisis, indicating that the shorter maturity groups have fatter tails of volatility than the longer maturity groups.

Panel B of Table 2 presents the summary statistics of forward variance. These summary statistics are similar to the corresponding summary statistics of spot variance in Panel A, but with two notable exceptions. First, as reported in JT (2012), the forward variance tends to have stronger autocorrelation than spot variance, especially at shorter maturities. At the same time, forward variance has very similar levels of autocorrelation across maturity groups, while spot variance has very different levels of autocorrelation across maturity groups. For example, the first-order autocorrelation varies in a very narrow range between 0.982 and 0.991 for forward variance across maturity groups, but varies in a relatively wider range between 0.961 and 0.989 for spot variance. Furthermore, the kurtosis of spot variance is much larger than those of forward variance, especially for the short maturity groups. For example, the kurtosis of the spot variance in maturity Group 1 is approximately 33, nearly twice the kurtosis of forward variance in Group 1. The kurtosis for spot variance also decays rapidly across maturity groups while the kurtosis for forward variance is more stable.

3. Empirical results: testing the martingale restriction on forward variance

3.1. Unconditional tests

Following Longstaff (2000) and JT (2012), we first conduct an unconditional test with the null hypothesis that daily changes in forward variance have a zero mean. As aforementioned, this martingale condition is only true under the risk-neutral probability measure. The expected daily change in forward variance may not be zero under the objective probability measure. This unconditional test is thus not a true model-free test, but a joint test of the martingale restriction and the variance risk premium effect. JT (2012) argue that the effect of the variance risk premium is negligible at daily frequency based on the argument of Fama (1991). Therefore, our test mitigates the joint hypothesis problem in market efficiency tests. We perform the unconditional test separately for the four forward-maturity groups defined in Table 2. Again, each group is defined by a forward period (i.e. (T_1, T_j) for $j = 2, 3, 4, 5$), spanning a future period between two option maturity dates. A time series of forward variance is then constructed from option prices using Eq. (8).

Table 3 reports results of the unconditional test. As expected, the average daily changes in forward variance are close to zeros in all four maturity groups. They range from $-1.34 * 10^{-4}$ to $2.86 * 10^{-5}$, with two positive means and two negative means. We cannot reject the hypothesis of a zero mean for each group. Therefore, the unconditional tests provide preliminary empirical support for the martingale restriction and the negligible impact of a variance risk premium.

3.2. Generalized Spectral Test (GST)

As noted earlier, the $M(1,0)$ test is suitable to evaluate the presence of MDS, whereas various derivative tests $M(1,l)$ with $l \geqslant 0$ are informative in revealing the nature of departure from a MDS. Following Hong and Lee (2003), we use the wild bootstrap method to implement GST to test the null hypothesis that MFFV is MDS. Let $\widehat{M}_i$ be the M_i statistic based on the time series of the forward variance $\{Y_t\}_{t=1}^n$. Let $\left\{Y_t^b\right\}_{t=1}^n$ be a bootstrapped sample of $\{Y_t\}_{t=1}^n$, and let M_i^b be the M_i statistic based on $\left\{Y_t^b\right\}_{t=1}^n$. We then calculate the bootstrapped p-value of $\widehat{M}_i$ approximated by $p_i^B = B^{-1}\sum_{b=1}^{B}\mathbf{I}(\widehat{M}_i^b \geqslant \widehat{M}_i)$, where $\mathbf{I}(\cdot)$ is the indicator function and b is the number of bootstrap replications. We set the number of bootstrap replications to 5000.[7]

The statistic $M(1,l)$ with $l \geqslant 0$ involves the choice of a bandwidth p, which may have a significant effect on the power of the test. We first choose a preliminary bandwidth $\bar{p}$ for the range 6–15 to ensure the robustness. Table 4 reports the values of $M(1,l)$ together with the bootstrapped p-values for the four maturity groups of the daily changes in forward variance, using the medium preliminary lag order $\bar{p} = 10$.[8] We use the 10% significant level in our paper, and for all tests, we use the Daniell kernel to maximize the asymptotic power of $M(1,l)$ over a class of kernels.

In Table 4, the results are different from group to group. $M(1,0)$ rejects the null hypothesis of MDS for the first two maturity groups of daily change in forward variance, while the last two groups are not rejected under the $M(1,0)$ test. All the p-values of the correlation test (i.e. $M(1,1)$) are larger than 10%, indicating that all four maturity groups of the forward variance are WN series. These results indicate that the WN series are not necessarily MDS. The ARCH-in-mean effect $M(1,2)$ and the kurtosis-in-mean effect $M(1,4)$ are insignificant for all groups, which shows that the violation of a martingale cannot be explained by the second- and fourth-order conditional moments. However, the

[7] We generate a bootstrap sample according to the formula $Y_t^b = aY_t$ with probability $p = a/\sqrt{5}$ and $Y_t^b = (1-a)Y_t$ with probability $1-p$, where $a = (1+\sqrt{5})/2$. According to Hong and Lee (2003), the test statistic of GST can be implemented by bootstraping, and the performance is adequate when the number of boostrap replications is over 300, and the sample size is larger than 100.

[8] The results for other values of $\bar{p}$ are quite similar and not reported here.

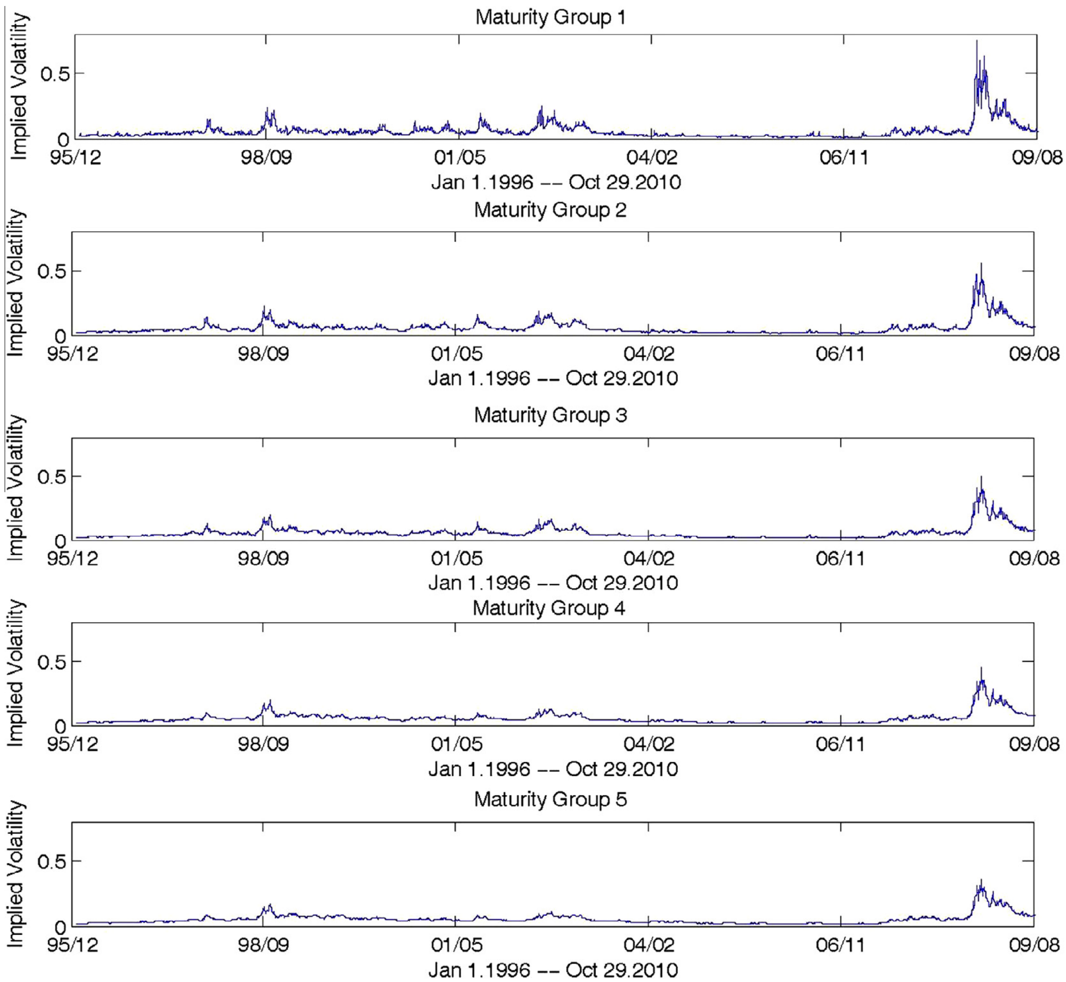


Fig. 1. Spot variances across maturities from Jan. 1996 to Oct. 2010. *Note:* This figure plots the JT (2012) model-free spot variance of five maturity groups. The time horizon is from Jan. 1996 to Oct. 2010.

Table 3
Summary statistics of daily change in forward variance: unconditional test.

Group	N	Mean (10^{-5})	Std. (10^{-5})	T-stat
T_1T_2	3732	−13.4470	12.6289	−1.0648
T_1T_3	3732	−2.1959	10.3112	−0.2130
T_1T_4	3732	2.8613	8.3983	0.3407
T_1T_5	3732	0.7011	7.1474	0.0981

Note: ***, **, and * indicate the coefficient is significantly different from 0 at the 1%, 5%, and 10% levels respectively.

skewness-in-mean effect $M(1,3)$ is significant with p-values less than 5% for the first two groups, indicating that the third-order conditional moments of the daily change in forward variance can help to predict future forward variance in the first two groups. Furthermore, the correlation between the lagged third-order conditional moments and the daily change in forward variance is linear. Thus, regressing the daily change in forward variance against the lagged third-order conditional moments should produce significant coefficients.

3.3. Forecasting forward variance with lagged moments

We follow JT (2012) and conduct a series of linear orthogonality tests to check our findings from GST. We regress the one-day-ahead forward variance on the daily changes of the first to the fourth lagged moments of forward variance. This is done for two

Table 4
Summary statistics of p-values of GST.

	T_1T_2		T_1T_3		T_1T_4		T_1T_5	
	Statistic	p-Values	Statistic	p-Values	Statistic	p-Values	Statistic	p-Values
M (1,0)	46.701**	(0.024)	32.137*	(0.056)	7.162	(0.636)	7.289	(0.564)
M (1,1)	34.467	(0.224)	53.095	(0.118)	30.161	(0.291)	26.235	(0.314)
M (1,2)	13.158	(0.976)	10.578	(0.972)	24.392	(0.892)	16.547	(0.900)
M (1,3)	95.613**	(0.034)	99.675**	(0.012)	77.580	(0.186)	72.316	(0.160)
M (1,4)	47.885	(0.658)	28.808	(0.908)	33.132	(0.914)	25.111	(0.886)
N	3732		3732		3732		3732	

Note: We compute bootstrapped p-values using 5000 bootstrap replications, with preliminary lag order 10. N denotes the number of observations. ***, **, and * indicate the coefficient is significantly different from 0 at the 1%, 5%, and 10% levels, respectively.

reasons: First, to confirm the results of the GST by applying more reliable explanatory variables in the conditional information set. Second, to determine whether the lagged third-order moment can be used to forecast daily changes in forward variance.

There are three steps to conduct this regression analysis. First, daily changes in forward variance are regressed against a number of lagged daily changes in forward variance with first to fourth orders using the following auto-regressive regression:

$$\triangle v(t;T_1,T_j) = \alpha + \sum_{i=1}^{k}[\phi_i^1 \triangle v(t-i;T_1,T_j) + \phi_i^2 \triangle v^2(t-i;T_1,T_j) + \phi_i^3 \triangle v^3(t-i;T_1,T_j) + \phi_i^4 \triangle v^4(t-i;T_1,T_j)] + \epsilon_t \quad (9)$$

where k is the number of lags, j = 2, 3, 4, 5, ϕ_i^1 is the serial correlation coefficient, ϕ_i^2 is the ARCH-in-mean coefficient that captures the lagged second-momment effect, ϕ_i^3 is the skewness-in-mean coefficient that captures the lagged third-momment effect, and ϕ_i^4 is the kurtosis-in-mean coefficient that captures the lagged fourth-momment effect. The daily change in forward variance on day t is not only regressed against lagged daily changes in prior days from $t-1$ to $t-k$, but also against the higher orders of the lagged daily changes in prior days.

In the second step, we regress daily changes of forward variance on the conditional moments with lags up to order four. This is done separately in each group of forward variance as follows:

$$\triangle v(t;T_1,T_j) = \alpha + \sum_{i=1}^{k}\phi_i^1 \triangle v(t-i;T_1,T_j) + \epsilon_t \quad (10)$$

$$\triangle v(t;T_1,T_j) = \alpha + \sum_{i=1}^{k}\phi_i^2 \triangle v^2(t-i;T_1,T_j) + \epsilon_t \quad (11)$$

$$\triangle v(t;T_1,T_j) = \alpha + \sum_{i=1}^{k}\phi_i^3 \triangle v^3(t-i;T_1,T_j) + \epsilon_t \quad (12)$$

$$\triangle v(t;T_1,T_j) = \alpha + \sum_{i=1}^{k}\phi_i^4 \triangle v^4(t-i;T_1,T_j) + \epsilon_t \quad (13)$$

The final step depends on the results of the first two steps. If some higher-order conditional moments are significant in both initial regressions, we include them as regressors in the final step and re-estimate the regression. For example, if the third-order and the fourth-order conditional moments are significant in both two steps, then we regress the daily change of the forward variance against the third-order and the fourth-order conditional moments.

According to the GST results, the third-order conditional moments of lagged change in forward variance may provide information to predict the daily change in forward variance in the first two maturity groups. However, the lagged daily changes as well as the lagged second-order and the fourth-order moments of the changes in forward variance will not help forecasting the daily change in forward variance on day t for all four maturity groups. If the results of the auto-regression test are the same as those of the GST, then some coefficients should be significantly different from zero for the first and second groups in all three steps of regression, while other coefficients should be zero in any one of the three regression steps. In other words, except for some lagged third-order moments of daily changes for Group 1 and Group 2, the daily change on day t should be uncorrelated with any of the lagged daily changes (raised up to the power of four) before day t for all four groups.

In conducting the auto-regression analysis, we first determine the maximum lagged order k. JT (2012) set k to be 25, which is the lag period of just over one month. However, they also illustrate that the autocorrelation diminish rapidly in both magnitude and statistical significance beyond the first two lags for all four groups. As a result, in our paper we only consider the prior information within one trading week and set the maximum lagged order to be 5. This is also consistent with GST in which we set the lag order p to be 10, as we only consider the information two weeks prior and put more weight on the first week under the Daniell kernel function.

We conduct these three-step predictive regressions as a way to test the robustness of the significant coefficients. We report the results of the first step in Table 5 and the results of the second step in Table 6. As shown in Table 5, all the coefficients of the last two groups are not significant. These results are consistent with GST, but for the first two maturity groups, all the absolute values of the coefficients of lagged daily change in forward variance are smaller than 0.08, and none of them are statistically significant at any conventional significance level. This is consistent with the finding of JT (2012). However, the coefficients of some higher order lags are not negligible. For the lagged second moment, the coefficients of the second lags in Group 1 and the fifth lags in both Group 1 and Group 2 are significant at the 5% level. For the lagged third moments, the coefficient of second lags is significant at the 1% level for Group 1 and at the 5% level for Group 2. Finally, we reject the hypothesis that the coefficient of the second lags of fourth moments is zero at the 1% significance level and that the coefficient of the fifth lags is zero at the 5% significance level. We then perform the second-step regressions to ensure robustness.

As mentioned above, the significance may change with different regressors. In Panel A of Table 6, we only regress the daily change in forward variance against the five lagged daily changes. The coefficient of the third lags is statistically significant at the 1% level, which is true for all four maturity groups. Contrary to Table 5, none of the correlation coefficients are statistically significant for all four groups. Furthermore, in contrast to the results in Table 5, none of the ARCH-in-mean coefficients in Panel B of Table 6 is significant. This indicates that the significance of some lagged ARCH-in-mean coefficients for Group 1 and Group 2 in Table 5 is not robust.

For maturity Group 1, the second lagged skewness-in-mean coefficient is significant at the 1% level in both regressions. For maturity Group 2, the significance of the second lagged skew-in-mean coefficient still remains. In addition, the fifth lagged kurtosis-in-mean coefficient is also significant at the 5% level. On the other hand, the significance of the kurtosis-in-mean coefficient for maturity Group 1 shifts from the second lag in Table 5 to the third and the fifth lags in Panel D of Table 6. The third lagged kurtosis-in-mean coefficient is significant using the second-step regression, but not in the first-step regression. In summary, the results in Table 5 and Table 6 show that not only the third-order conditional moments of daily changes in forward variance provide useful information to predict the daily change in forward variance, but the fourth-order conditional moments do so as well. This result deviates slightly from the result of GST.

To explain the difference between the two methods of analysis and to ensure robustness of the results, we perform the third-step regression. In particular, we use the following auto-regressive regression (for $j = 2, 3, 4, 5$):

$$\triangle v(t;T_1,T_j) = \alpha + \sum_{i=1}^{k}[\phi_i^3 \triangle v^3(t-i;T_1,T_j) + \phi_i^4 \triangle v^4(t-i;T_1,T_j)] + \epsilon_t. \quad (14)$$

We keep k equal to 5 to determine whether the significance of the skewness-in-mean and kurtosis-in-mean coefficients will change. The results are shown in Table 7. While the second lagged skewness-in-mean coefficient remains significant, the kurtosis-in-mean coefficients are no longer significant. These results indicate

Table 5
Auto-regression of daily change in forward variance with lagged first to fourth moments.

Coefficient	Maturity group 1		Maturity group 2		Maturity group 3		Maturity group 4	
	Estimate	Std.	Estimate	Std.	Estimate	Std.	Estimate	Std.
α	−0.0002	(0.0001)	−0.0001	(0.0001)	−0.0001	(0.0001)	−0.0001	(0.0001)
ϕ_1^1	−0.0206	(0.0488)	0.0025	(0.0449)	−0.0266	(0.0611)	0.0003	(0.0550)
ϕ_2^1	0.0747	(0.0531)	0.0448	(0.0626)	0.0092	(0.0657)	0.0342	(0.0688)
ϕ_3^1	0.0634	(0.0540)	0.1191	(0.0632)	0.0319	(0.0482)	0.1064	(0.0632)
ϕ_4^1	0.0052	(0.0588)	−0.0320	(0.0631)	0.0543	(0.0527)	−0.0259	(0.0629)
ϕ_5^1	−0.0179	(0.0552)	−0.0149	(0.0664)	−0.0296	(0.0628)	−0.0193	(0.0585)
ϕ_1^2	−4.0239	(3.1974)	−2.0448	(2.9073)	−1.2625	(3.9894)	−3.9248	(6.4050)
ϕ_2^2	8.7703**	(4.4162)	3.5552	(4.5284)	1.3215	(5.3364)	2.4338	(6.9254)
ϕ_3^2	−5.4903	(3.5652)	−4.7213	(4.5614)	0.2950	(4.8884)	1.1738	(6.0463)
ϕ_4^2	2.5567	(4.0614)	5.5298	(5.7251)	0.7118	(6.1732)	3.7484	(8.1934)
ϕ_5^2	6.6470**	(3.1010)	10.903**	(4.6812)	11.881	(6.4832)	13.787	(7.9016)
ϕ_1^3	−2.3576	(24.208)	−46.080	(31.174)	−27.615	(80.027)	−86.504	(108.485)
ϕ_2^3	−74.944***	(20.305)	−79.521**	(37.431)	−74.767	(43.928)	−104.617	(92.465)
ϕ_3^3	−12.865	(23.592)	−22.472	(30.904)	41.682	(67.690)	−49.154	(108.558)
ϕ_4^3	−24.925	(24.079)	−23.360	(31.445)	−88.3628	(58.364)	8.7268	(92.633)
ϕ_5^3	−26.506	(28.486)	−64.779	(39.785)	−19.276	(76.107)	−67.967	(101.126)
ϕ_1^4	466.097	(439.197)	502.898	(698.347)	368.156	(2117.50)	586.874	(4530.66)
ϕ_2^4	−1116.94**	(566.369)	−1551.55	(1151.74)	−1603.99	(2216.78)	−4508.96	(3790.63)
ϕ_3^4	748.3573	(516.916)	1376.353	(887.689)	−336.869	(2008.83)	1472.305	(3351.74)
ϕ_4^4	−561.014	(451.958)	−1170.57	(1030.43)	832.4185	(1598.54)	−1664.5	(4067.59)
ϕ_5^4	−595.593	(395.893)	−1739.6	(1286.99)	−4328.98	(2805.78)	−8146.06	(5194.24)
N	3723		3723		3723		3723	
R^2	0.1353		0.1246		0.0866		0.0760	
DW	1.993		1.959		2.000		1.993	

Note: This table reports the OLS results of the regression for the original forward variances as in (9). The standard errors of the coefficients' estimates are computed following a robust procedure taking into account the heteroscedastic and auto-correlated errors structure and reported in the brackets next to the coefficient estimates. ***, **, and * indicate the coefficient is significantly different from 0 at the 1%, 5%, and 10% levels, respectively. DW denotes the Durbin–Watson statistic. We use lines to separate the lagged orders of different moments.

that the lagged third-order moments of daily changes in forward variance contain the information of the lagged fourth-order moments of daily changes in forward variance. From this, we identify that the lagged cubic power of daily changes in forward variance is an important source of information for forecasting forward variance, which is consistent with the findings of GST.

In summary, when we combine the results of GST and regression-based methods, we find evidence against options market efficiency. For the long-maturity forward-variance groups, we reach the same conclusion as JT (2012) even by using GST. However, for the short-maturity groups, we reject the hypothesis that the forward variance is a MDS. We therefore get a different conclusion on the options market efficiency compared to JT (2012), and this is largely due to our more sophisticated approach in testing the non-linear dependence of the time series of forward variance.

4. Robustness checks and discussions

Why do the regression results above appear inconsistent with the efficient market hypothesis? We construct a new measure of spot variance and conduct additional tests to investigate the source of the skewness-in-mean effect and test the robustness of our findings. One possible source of the skewness-in-mean effect is the jump of underlying asset prices. As shown in Du and Kapadia (2012) (henceforth DK (2012)), the spot variance calculated by Eq. (7) is significantly biased when jumps contribute to 20% of the total variance. Therefore, the jumps may threaten the robustness of our previous conclusion of informational inefficiency based on MFFV. Furthermore, the efficient market hypothesis postulates that security prices accurately and instantaneously incorporate all available information. However, if the jumps are large and sudden, then the information cannot be instantaneously and completely incorporated by the market. Therefore, the observable jump may exhibit forecastability for the excess stock return in the next period, suggesting the presence of the options-market inefficiency.

4.1. Robust spot variance measure

We therefore measure the spot variance using the method of DK (2012). Let $v(t,T)$ be the annualized spot variance of the log-return, $v(t,T) = \frac{1}{T}(E^{\mathbb{Q}}(ln(S_T/S_0))^2 - \mu_{0,T}^2)$, where $\mu_{0,T}^2 = E^{\mathbb{Q}}(ln(S_T/S_0))$. They use the following equation to calculate $E^{\mathbb{Q}}\left[e^{-rT}(ln(S_T/S_0))^2\right]$:

$$E^{\mathbb{Q}}\left[e^{-rT}(\ln(S_T/S_0))^2\right] = \int_{K>S_0} \frac{2(1 - ln(K/S_0))}{K^2} C(S_0;K,T)dK + \int_{K<S_0} \frac{2(1 - ln(S_0/K))}{K^2} P(S_0;K,T)dK \quad (15)$$

where r is the constant risk-free rate; S_0 is the underlying asset price at time 0; S_T is the underlying asset price at time T; $C(S_0;K,T)$ is the call option with strike price K; $P(S_0;K,T)$ is the put option of strike price K, and T is the remaining time to expire. Once we have $\mu_{0,T}^2$ and $E^{\mathbb{Q}}\left[e^{-rT}(ln(S_T/S_0))^2\right]$, we can calculate the spot variance $v(t,T)$. The DK (2012) variance is a more accurate measure of the spot variance as it correctly accounts for jump risk. Once we have the spot variance, we use the no-arbitrage argument as that in deriving forward interest rate from the spot interest rate to caculate forward variance. We report the descriptive statistics of the DK (2012) spot variance measure and forward variance measure in Table 8.

The DK (2012) spot variance has the same upward-sloping term structure pattern as the JT (2012) spot variance. The

Table 6
Regression of daily change in forward variance with lagged first to fourth moments.

Coefficient	Maturity Group 1		Maturity Group 2		Maturity Group 3		Maturity Group 4	
	Estimate	Std.	Estimate	Std.	Estimate	Std.	Estimate	Std.
	Panel A: regression of daily change in forward variance with lagged first moment							
α	−0.0001	(0.0001)	−0.0001	(0.0001)	1.51E−05	(8.56E−05)	1.17E−05	(7.10E−05)
ϕ_1^1	−0.0583	(0.0526)	−0.0688	(0.0515)	−0.0518	(0.0519)	−0.0735	(0.0537)
ϕ_2^1	−0.0769	(0.0639)	−0.0736	(0.0535)	−0.0831	(0.0545)	−0.06136	(0.0484)
ϕ_3^1	0.1080**	(0.0436)	0.1261***	(0.0402)	0.0898**	(0.0404)	0.0802**	(0.0373)
ϕ_4^1	−0.0575	(0.0824)	−0.0496	(0.0725)	−0.0578	(0.0809)	−0.0265	(0.0722)
ϕ_5^1	−0.0270	(0.1077)	−0.0604	(0.0971)	−0.0624	(0.0885)	−0.0307	(0.0796)
R^2	0.024		0.033		0.024		0.016	
DW	2.011		2.000		2.011		2.017	
	Panel B: regression of daily change in forward variance with lagged second moment							
α	−4.55E−05	(0.0001)	3.59E−05	(8.22E−05)	1.88E−05	(6.95E−05)	3.85E−05	6.02E−05
ϕ_1^2	0.5980	(1.8162)	0.3395	(2.5825)	−0.5248	(2.9846)	−1.91751	(3.9534)
ϕ_2^2	−0.6672	(2.657)	−2.7050	(2.8914)	−4.9005	(3.0027)	−5.4576	(3.9329)
ϕ_3^2	1.9594	(1.6308)	3.1573	(1.9949)	3.3054	(3.3014)	4.5620	(2.8024)
ϕ_4^2	−1.5570	(1.7197)	−0.1523	(2.8087)	2.7695	(4.7068)	0.2639	(4.4924)
ϕ_5^2	−0.9021	(1.9947)	−1.838	(2.4880)	−0.3885	(2.0867)	−0.9436	(3.7055)
R^2	0.011		0.014		0.024		0.028	
DW	2.121		2.127		2.118		2.127	
	Panel C: regression of daily change in forward variance with lagged third moment							
α	−5.82E−05	(0.0001)	1.65E−05	(0.0001)	4.31E−05	(8.29E−05	2.77E−05	6.99E−05)
ϕ_1^3	−9.2419	(16.2366)	41.4816	(26.2908)	−42.3925	(54.4389)	−100.010	(85.1218)
ϕ_2^3	−50.7403***	(14.3694)	−56.6302***	(20.2869)	−42.3925	(54.4389)	−100.010	(85.1218)
ϕ_3^3	35.2535***	(13.1997)	25.8707	(18.3391)	55.5891	(37.6553)	42.5797	(42.6739)
ϕ_4^3	−31.2098	(22.7426)	−41.1561	(32.7427)	−84.5217	(52.1636)	−65.9821	(83.5206)
ϕ_5^3	6.3686	(41.1452)	−33.0859	(69.6682)	−68.0233	(61.1473)	−28.2094	(120.064)
R^2	0.058		0.05		0.056		0.028	
DW	2.039		1.991		2.041		2.022	
	Panel D: regression of daily change in forward variance with lagged fourth moment							
α	−1.92E−05	(0.0001)	3.73E−05	(9.25E−05)	3.43E−05	(8.32E−05)	3.81E−05	(6.53E−05)
ϕ_1^4	115.853	(126.742)	245.721	(591.745)	−35.494	(232.563)	−522.358	(2405.846)
ϕ_2^4	−426.246	(267.061)	−839.318	(557.502)	358.409	(228.994)	−2801.696	(1835.332)
ϕ_3^4	346.390**	(167.728)	499.887	(305.405)	368.114	(228.362)	2154.727	(1337.115)
ϕ_4^4	2.965	(195.243)	128.791	(517.413)	356.402	(229.789)	871.891	(1999.215)
ϕ_5^4	−417.850***	(117.189)	−1045.512**	(411.435)	−357.962	(232.889)	−2164.940	(2327.351)
R^2	0.043		0.039		0.035		0.037	
DW	2.089		2.096		2.099		2.113	

Note: This table reports the OLS results of the regression for the original forward variances as in (10)–(13). The standard errors of the coefficients' estimates are computed following a robust procedure taking into account the heteroscedastic and autocorrelated errors structure and reported in the brackets next to the coefficient estimates. ***, **, and * indicate the coefficient is significantly different from 0 at the 1%, 5%, and 10% levels, respectively. DW denotes the Durbin–Watson statistic.

autocorrelations of the DK (2012) spot variance and forward variance in each group are also similar to the original spot and forward variances. Generally speaking, all the statistics of the DK (2012) variance measure are similar to those previously reported, except for the mean and median. Compared to Table 2, the mean (median) of all groups of the DK (2012) spot variance and forward variance in all groups are higher than those of the JT (2012) spot variances. This is because the DK (2012) spot variance accounts for the jump risk more accurately. To further illustrate the difference between two spot-variance measures, we calculate the "relative-difference percentage" as the fraction of the difference between the DK (2012) spot variance and the JT (2012) spot variance, divided by the DK (2012) spot variance. The time series of the difference percentages in the first and second groups are plotted in Fig. 2.

The relative difference percentages are mostly positive with few negative percentages. The magnitudes of negative percentages are small, less than 2%, while the maximal positive difference percentages reach as high as 10% during the 1997 Asian financial crisis and 2008 financial crisis. This is consistent with the findings of DK (2012) on the bias of original spot variance measures.

To examine the martingale properties of the DK (2012) forward variance, we perform the same unconditional tests for all four maturity groups. The unconditional test results and GST results are shown in Table 9 and Table 10, respectively. The average daily changes of the DK (2012) forward variance in Table 9 are again close to zero in all four maturity groups, indicating that the daily variance risk premium is negligible for the new forward variance. The GST results also reject the martingale hypothesis for the first two maturity groups and the diagnostic tests show that there also exists a skewness-in-mean effect. However, the p-value of the skewness-in-mean coefficient is much larger than that in the previous results (almost close to 0.1), which suggests the skewness-in-mean effect is weakened in the DK (2012) forward variance time series, and that the time series of DK (2012) forward variance is much closer to a martingale.

In the next step, we perform the same three-step regression analysis, but only for the first two groups.[9] This is sufficient for our analysis as the coefficients of third-order conditional moments

[9] The regression results are available upon request.

Table 7
Regression of daily change in forward variance with lagged third and fourth moments.

Coefficient	Maturity group 1		Maturity group 2	
	Estimate	Std.	Estimate	Std.
α	−6.41E−06	(0.0001)	−41.4581	(25.9765)
ϕ_1^3	−4.8255	(25.4116)	−41.4581	(29.9227)
ϕ_2^3	−64.5474**	(25.4116)	−57.3948**	(26.3755)
ϕ_3^3	21.0103	(19.8439)	16.8440	(17.1401)
ϕ_4^3	−34.6092	(24.9995)	−36.5071	(35.5229)
ϕ_5^3	−6.7766	(36.5750)	−45.9082	(48.2054)
ϕ_1^4	48.7196	(174.724)	218.785	(372.005)
ϕ_2^4	−5.3785	(281.965)	−650.559	(584.498)
ϕ_3^4	332.649	(248.322)	814.375	(496.238)
ϕ_4^4	−432.187	(301.964)	−85.6781	(669.942)
ϕ_5^4	−249.137	(251.71)	−945.065	(664.700)
R^2	0.083		0.082	
DW	2.048		1.973	

Note: This table reports the OLS results of the regression for the original forward variances as in (14). The standard errors of the coefficients' estimates are computed following a robust procedure taking into account the heteroscedastic and auto-correlated error structure and reported in the brackets next to the coefficient estimates. ***, **, and * indicate the coefficient is significantly different from 0 at the 1%, 5%, and 10% levels, respectively. DW denotes the Durbin–Watson statistic. We use lines to separate the lagged orders of different moments.

Table 8
Summary statistics of the DK (2012) model-free implied variances: January 1996-october, 2010.

Panel A: spot variance					
	T_1	T_2	T_3	T_4	T_5
Mean_Days	22	52	89	152	237
Median_Days	22	52	84	144	234
N	3733	3733	3733	3733	3733
Mean	0.0541	0.0568	0.0581	0.0594	0.0599
Median	0.0404	0.0449	0.0478	0.0502	0.0516
Std	0.0584	0.0530	0.0498	0.0470	0.0437
Skewness	4.5711	3.8886	3.5544	3.3124	2.9243
Kurtosis	32.5146	24.9972	21.3597	19.2805	15.5937
AR(1)	0.9634	0.9781	0.9833	0.9867	0.9895
AR(2)	0.9343	0.9599	0.9698	0.9758	0.9668
AR(3)	0.9165	0.9487	0.9608	0.9681	0.9746
AR(4)	0.8974	0.9346	0.9487	0.9583	0.9668
AR(5)	0.8906	0.9273	0.9414	0.9522	0.9610
Panel B: forward variance					
	T_1T_2	T_1T_3	T_1T_4	T_1T_5	
Mean	0.0588	0.0599	0.0605	0.0606	
Median	0.0485	0.0502	0.0517	0.0524	
Std	0.0507	0.0479	0.0453	0.0427	
Skewness	3.4728	3.2508	3.0019	2.6848	
Kurtosis	20.9164	18.5014	16.6468	13.5898	
AR(1)	0.9810	0.9869	0.9893	0.9912	
AR(2)	0.9664	0.9761	0.9802	0.9840	
AR(3)	0.9564	0.9683	0.9733	0.9781	
AR(4)	0.9444	0.9582	0.9657	0.9715	
AR(5)	0.9354	0.9499	0.9595	0.9657	

Note: Panels A and B report the descriptive statistics of the Du and Kapadia's (2012) daily model-free implied spot and forward variances for different maturity groups.

are significant in both the first two steps of the analysis. The results of the first-step regression show that only the coefficient of the second lagged third-order moment is statistically significant in Group 1, while there are three statistically significant coefficients in Group 2 (the second and fifth lagged third-order moments and the fifth lagged fourth-order moments). Although the statistical significance of the second lagged third-order moments does not change, the magnitude is smaller than its corresponding values in the regressions of the JT (2012) forward variances. This means that the skewness-in-mean effect is weaken in the DK (2012) forward variance series.

In the second-step regression, only the third-order conditional moments are correlated with the daily changes in forward variance for both Group 1 and Group 2. Therefore, when combining the results of the first- and second-step regressions, we consider the third-order conditional moments as being robust and providing useful information for forecasting the daily changes in forward variance.

4.2. Jump effects

The skewness-in-mean effect exists for both forward variance measures, and in this section, we examine the impact of underlying asset price jumps on the violation of orthogonality condition.[10] There are various methods available to identify jumps in a nonparametric model-free context, such as Barndorff-Nielsen and Shephard (2006), Huang and Tauchen (2005), and Lee and Mykland (2008). This paper uses the Barndorff-Nielsen and Shephard (2006) (henceforth BNS (2006)) method to detect whether there are jumps on date t based on 5-min high frequency data of S&P 500 index.[11]

Using a dummy variable for the jump days based on the BNS (2006) testing method, we modify and re-estimate regression (12) by adding cross terms of the lagged third-order moment variables and the jump dummy variables for the first and second maturity groups:

$$\triangle v(t;T_1,T_j) = \alpha + \sum_{i=1}^{k}(\phi_i^3 + \gamma_i D_{t-i}) \cdot \triangle v^3(t-i;T_1,T_j) + \epsilon_t \qquad (16)$$

where $j = 2$ and 3, and D_{t-i} is a dummy variable that is equal to 1 if day $t-i$ has a jump and zero otherwise. Coefficients of the cross terms γ_i capture the impact of any jump in the underlying asset price. If the underlying asset price jump is related to the future daily change in forward variance, we expect the cross term coefficients γ_i to be significantly non-zero. Table 11 presents the results of the dummy variable regressions (16). Panel A shows the results of the JT (2012) forward variance measure and Panel B shows the results of the DK (2012) forward variance measure.

For both the maturity Groups 1 and 2 of the JT (2012) forward variance measure, the significance of the lagged third moments in forecasting the forward variance weakens once we control for the jump effect,[12] and lagged jump information contributes significantly to the prediction of forward variance.[13] We can find similar results for the DK (2012) forward variance.[14] Therefore, the jump of the S&P 500 index is an important factor contributing to the inefficiency of the options market, but it may not be the only source as the lagged third moments information still helps forecasting the forward variance even if the jump effect is controlled.

[10] Although the calculation of the forward variance here is robust to jumps according to DK (2012), the time-series property (i.e. martingale property) of the forward variance may still be affected by the jumps of underlying asset prices.

[11] The BNS (2006) method is used in many preceding works such as Wright and Zhou (2009), Zhang, Zhou, and Zhu (2009), Guo et al. (2014) etc. Based on simulation results in Huang and Tauchen (2005), when the variance of microstructure noise is not large, the testing power of the BNS method on 5-min data is adequate compared to others which are adjusted for microstructure noise. Therefore, in our paper, we take 5-min data as our sample. See Appendix A for the detailed introduction of this testing procedure.

[12] We compare the significance of the coefficients in Panel C of Table 6 with Panel A of Table 11. For Group 1, ϕ_2^3 and ϕ_3^3 are significant at the 1% significance level in Table 6 while ϕ_2^3 is no longer significant and ϕ_3^3 is only significant at the 5% significance level when considering the jump effect. Similar comparision also can be found in Group 2.

[13] γ_2 is statistically significant for Group 1 and γ_1 and γ_2 are significant for Group 2.

[14] One exception is that none of the jump dummy variable coefficients are significant for Group 1.

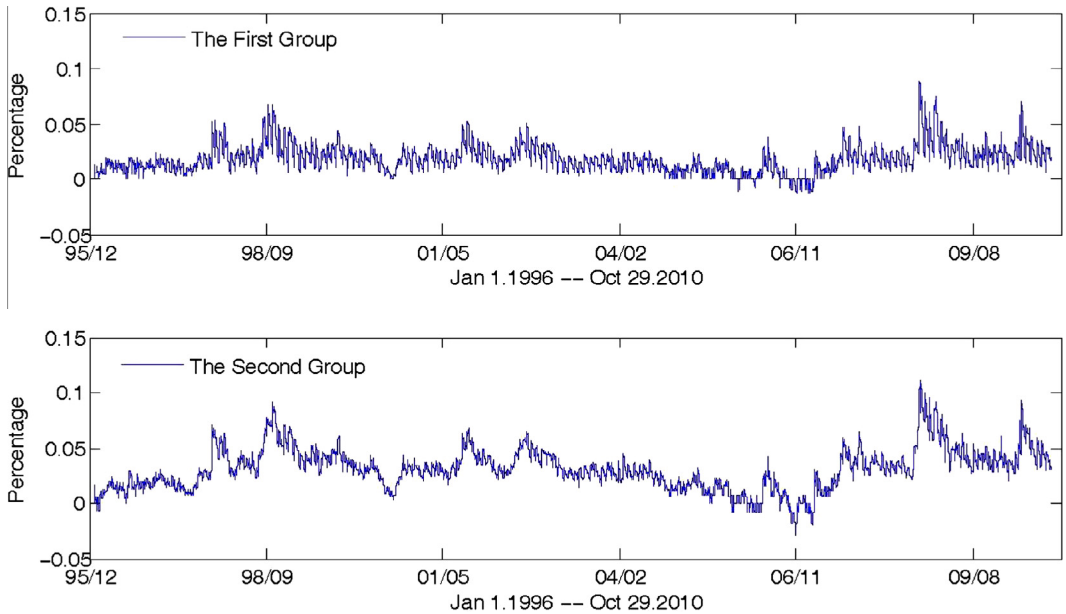


Fig. 2. The relative difference between DK (2012) spot variances and JT (2012) spot variances of the first two maturity groups during Jan. 1996 – Oct. 2010. *Note:* This figure plots the relative difference between DK (2012) spot variance and JT (2012) spot variance. The positive number indicates the DK (2012) spot variance has a larger magnitude, and vice versa. Since we focus on the information inefficiency of the options market, we only report the results of the first two maturity groups. The time horizon is from Jan. 1996 to Oct. 2010.

Table 9
Summary statistics of daily change in DK (2012) forward variance: unconditional test.

Group	N	Mean (10^{-5})	Std. (10^{-5})	T-stat
T_1T_2	3732	−13.5681	13.9414	−0.9745
T_1T_3	3732	−6.4042	11.1435	−0.5800
T_1T_4	3732	2.6606	9.1820	0.2898
T_1T_5	3732	2.1242	7.8510	0.2706

Note: This table reports the mean and standard error of new daily changes in forward variance. The standard errors are computed following a robust procedure taking into account the heteroscedastic and autocorrelated error structure and reported in the brackets next to the coefficient estimates. *N* denotes the number of observations. ***, **, and * indicate the coefficient is significantly different from 0 at the 1%, 5%, and 10% levels, respectively.

There are two possible channels through which underlying asset price jumps cause inefficiency in the options market. The first channel is the forecastability of the jump risk on the excess stock return. Jacquier and Okou (2012) argue that the realized jump risk can predict the excess stock return in the short-horizon, while jumps lack predicting power for the medium/long-term excess stock return. If the jump can predict the excess return, we can develop trading strategies for the corresponding options, which lead to options market inefficiency. This channel can also explain why only short-maturity forward variances violate the martingale restriction. The maturity of the first two daily changes in forward variance groups is less than six months, which is consistent with the predictable horizon of the realized jump as in Guo et al. (2014). Another channel is the jump-induced hedging errors, which are investigated by Li (2012). If the price of the underlying asset jumps, then the delta-hedging investor cannot execute a complete hedge and hence many non-perfect hedging strategies will affect the family of options based on the same underlying assets. In other words, the prices of the options cannot completely reflect the jump information of the underlying asset, and this provides one particular case of market inefficiency.

4.3. Controlling for the effect of market illiquidity

To ensure our results are robust to the effect of market illiquidity, we re-examine options market efficiency by controlling the effect of market illiquidity. To examine the impact of market illiquidity, we use trading volume as a proxy for market liquidity and separate trading days in our sample period into high-volume days and low-volume days. For each maturity group, we aggregate trading volumes for all options in the group on each trading day and use the total volume as our proxy for liquidity. We consider trends and seasonality in classifying high-volume vs. low-volume days to cope with the fact that trading volume tends to rise over time and

Table 10
Summary statistics of *p*-values of generalized spectral test for DK (2012) forward variance.

	T_1T_2		T_1T_3		T_1T_4		T_1T_5	
	Statistic	*p*-Values	Statistic	*p*-Values	Statistic	*p*-Values	Statistic	*p*-Values
M (1,0)	56.170*	0.080	63.327*	0.078	5.162	0.736	6.305	0.665
M (1,1)	36.705	0.338	53.796	0.114	38.102	0.251	27.364	0.402
M (1,2)	13.042	0.958	12.699	0.960	20.425	0.858	16.808	0.856
M (1,3)	102.032*	0.092	99.243*	0.080	82.897	0.144	60.169	0.282
M (1,4)	63.949	0.318	38.443	0.656	39.534	0.674	23.352	0.832
N	3732		3732		3,732		3732	

Note: We compute bootstrapped *p*-values using 5000 bootstrap replications, with preliminary lag order 10. *N* denotes the number of observations. ***, **, and * indicate the coefficient is significantly different from 0 at the 1%, 5%, and 10% levels, respectively.

Table 11
The effect of underlying asset jumps.

Coefficient	Maturity group 1		Maturity group 2	
	Estimate	Std.	Estimate	Std.
Panel A: the jump effect on the JT (2012) forward variance				
α	0.0000	(0.0001)	3.42E−05	(0.0001)
ϕ_1^3	−13.3115	(34.3412)	−41.6488	(38.8909)
ϕ_2^3	−38.2869	(26.3182)	-57.9232^{**}	(28.0024)
ϕ_3^3	52.6076^{**}	(20.6222)	56.6388	(34.7494)
ϕ_4^3	−10.1873	(33.9611)	−23.6922	(39.0953)
ϕ_5^3	−5.9315	(52.4980)	−38.7141	(73.1999)
γ_1	26.9468	(35.3419)	143.8965^{***}	(39.5499)
γ_2	-58.9899^{**}	(26.9689)	-212.5595^{***}	(27.4561)
γ_3	−7.8454	(41.5115)	49.67109	(79.5704)
γ_4	−42.4282	(37.1735)	−148.657	(92.2489)
γ_5	61.1579	(44.2565)	72.0626	(49.2093)
N	3,611		3,611	
R^2	0.091		0.112	
DW	1.992		1.952	
Panel B: the jump effect on the DK (2012) forward variance				
α	−8.24E−05	(0.0001)	−1.83E−05	(0.0001)
ϕ_1^3	−4.5372	(26.5836)	−27.3340	(27.1875)
ϕ_2^3	−28.6612	(18.5471)	-38.3620^{**}	(16.7209)
ϕ_3^3	40.7604^{***}	(13.9284)	36.2640	(22.1394)
ϕ_4^3	−6.3949	(25.4869)	−16.0508	(27.1222)
ϕ_5^3	0.9845	(35.1102)	−25.2730	(49.7993)
γ_1	6.4870	(27.2585)	84.1760^{***}	(27.1319)
γ_2	−21.3702	(19.0089)	-116.7098^{***}	(16.6267)
γ_3	−15.4117	(25.6033)	22.2526	(51.9462)
γ_4	−22.6768	(25.4045)	−93.3187	(57.1905)
γ_5	29.7550	(31.8327)	108.9215	(38.5240)
N	3,611		3,611	
R^2	0.096		0.132	
DW	1.982		1.939	

Note: This table reports the OLS results of the regression as in (16). The regressions are run with 5 lags. The standard errors of the coefficients' estimates are computed following a robust procedure taking into account the heteroscedastic and autocorrelated error structure and reported in the brackets next to the coefficient estimates. ***, **, and * indicate the coefficient is significantly different from 0 at the 1%, 5%, and 10% levels, respectively. DW denotes the Durbin–Watson statistic.

Table 12
The effect of market illiquidity.

Coefficient	Maturity group 1		Maturity group 2	
	Estimate	Std.	Estimate	Std.
Panel A: the market illiquidity effect on the JT (2012) forward variance				
α	0.0000	(0.0001)	3.42E−05	(0.0001)
ϕ_1^3	−10.1720	(38.8853)	−73.240	(33.181)
ϕ_2^3	-46.0132^{***}	(17.2102)	-55.690^{**}	(22.646)
ϕ_3^3	3.6217	(23.917)	−18.332	(24.590)
ϕ_4^3	6.0790	(31.420)	−10.997	(52.488)
ϕ_5^3	−6.2024	(48.425)	−57.992	(72.189)
η_1	27.051	(42.338)	98.064	(67.288)
η_2	−35.833	(34.347)	−51.915	(91.102)
η_3	62.384	(47.193)	123.44	(84.675)
η_4	−79.270	(60.608)	−127.71	(84.859)
η_5	3.681	(34.547)	32.911	(51.682)
N	3,611		3,611	
R^2	0.099		0.107	
DW	2.033		1.932	
Panel B: the market illiquidity effect on the DK (2012) forward variance				
α	−8.24E−05	(0.0001)	−1.83E−05	(0.0001)
ϕ_1^3	−17.1166	(11.8001)	-40.4089^{***}	(13.1841)
ϕ_2^3	-32.5056^{***}	(10.5436)	-38.1553^{**}	(18.7739)
ϕ_3^3	10.2063	(10.0250)	−12.3784	(15.5878)
ϕ_4^3	−14.3312	(15.6001)	−6.7852	(33.8601)
ϕ_5^3	3.9045	(22.7850)	−35.6168	(45.4338)
η_1	0.1997	(0.1233)	62.0330	(42.7359)
η_2	−0.0286	(0.1468)	−42.4089	(55.4022)
η_3	0.1307	(0.1404)	49.1278	(30.1198)
η_4	−0.1546	(0.1583)	−95.7071	(56.3180)
η_5	−0.0868	(0.0694)	21.5958	(30.8499)
N	3611		3611	
R^2	0.098		0.122	
DW	2.076		1.955	

Note: This table reports the OLS results of the regression as in (17). The regressions are run with 5 lags. The standard errors of the coefficients' estimates are computed following a robust procedure taking into account the heteroscedastic and autocorrelated error structure and reported in the brackets next to the coefficient estimates. ***, **, and * indicate the coefficient is significantly different from 0 at the 1%, 5%, and 10% levels, respectively. DW denotes the Durbin–Watson statistic.

exhibit seasonality (e.g. day-of-the-week effect). We first detrend the daily trading volume over time for each maturity group. We then classify a trading day as a high-volume (low-volume) day if the trading volume on that day is higher (lower) than the median daily trading volume over the past month. By specifying a dummy variable η_i for low-volume days, we modify and re-run regression (16) by adding cross terms of the lag variables with the low-volume dummy:

$$\triangle v(t; T_1, T_j) = \alpha + \sum_{i=1}^{k} (\phi_i^3 + \eta_i D_{t-i}) \cdot \triangle v^3(t-i; T_1, T_j) + \epsilon_t \tag{17}$$

where D_{t-i} is the low-volume dummy variable that equals 1 if day $t-i$ is a low-volume day and zero otherwise. In this dummy variable regression, the coefficient ϕ_i^3 is the effect of the lagged daily change in forward variance on high-volume days while the coefficient $(\phi_i^3 + \eta_i)$ is the corresponding effect on low-volume days. The coefficient of the cross terms (η_i) captures the difference between high- and low-volume days and represents the impact of market illiquidity. If market illiquidity does not affect our previous results, then we expect the cross term coefficient (η_i) to be zero. From Table 12, we find that even when market illiqudity is controlled, the two-period lagged third moment of forward variance is still negatively and significantly associated with current forward variance in two short maturity groups. This finding is consistent with the one without controlling for market illiqudity, suggesting that our conclusion is not affected by market illiquidity.

4.4. The effect of the 2008 financial crisis

There are a few studies in the literature examining market structure and efficiency changes after the 2008 finanical crisis with most of them finding that the market structure and information costs changed after the financial crisis. For example, Angelini et al. (2011) examine the interbank interest-rate markets around the globe after the crisis and find the interest-rate spreads became exceptionally large and volatile after the 2008 crisis. The literature also reveals a surge of stock return predictability studies since 2008. This might indicate markets may go under consisderable stucutural change after the event of the 2008 financial crisis.

To further dig into the origin of options market inefficiency, we take January 1, 2008 as the date to divdide our sample into the period before financial crisis and the period after the financial crisis. We apply GST and use both JT (2012) and DK (2012) forward varaince measures to examine options market efficiency before and after the financial crisis. As shown in Table 13, we find that before the financial crisis, the hypothesis that options market is efficient is not rejected at any significance level and in both foward varaince measures. This finding is in line with JT (2012), where they use S&P options data before the financial crisis as a research sample.

However, in examining the sample after the financial crisis, we find that the hypothesis that the options market is efficient is rejected at the 5% significance level in short maturity groups and in both foward variance measures. This result confirms that the inefficiency of the S&P options market during 1996–2010 is mainly driven by the inefficiency of the period after the financial crisis. The outbreak of the financial crisis indeed brought some structural changes to trading patterns of option market participants, resulting in options market inefficiency.

4.5. Discussions

Yang et al. (2010) find that skewness of returns distribution is a negative pricing factor on a stock's estimated excess return. To examine this corresponding issue, we also run a series of regressions that regress the implied forward variance on the third moment (skewness) of lagged implied forward variance. The result as shown in Table 7 indicates that the lagged skewness of implied forward variance, especially the second-order lag in the short-maturity group, is also negatively associated with the current implied forward variance changes. The asset-pricing theory suggests that lagged skewness of forward variance is indeed a negative pricing factor on current implied forward variance. We link an economic explanation of this finding to the jump effect. Intuitively, jump risk is averted by risk averse investors, and assets whose returns are resilient to jump risk might be preferred by investors who are willing to pay a premium for such assets. This might cause negative correlations between jump risk and future risk premiums. It is generally believed in the literature that jump risk and skewness risk are inter-related in such a way that skewness risk of asset prices is partly driven by the jump risk. Guo et al. (2014) among other current studies suggest realized jump risk predicts excess stock market returns up to 12 months, but such predictability diminishes as we increase the horizon from one month to 12 months. Once this effect is projected onto the options market, it might suggest that short-term forward variances changes more than long-term forward variances, causing the forward variance to move negatively with the skewness effect.

We also provide empirical evidence that confirms this linkage in Table 11. We argue that past daily small jumps in the S&P500 index indeed contain useful information that causes the forward variance distribution to skew, indirectly and negatively affecting the current forward variance. Similiar to Yang et al. (2010), we document a negative relationship between the skewness-in-mean effect and the option implied forward variance spreads in the current study, although the economic explanation stands different.

5. Forecasting performance and trading strategy

5.1. Forecasting performance

We now examine whether forward variance can be forecasted based on the lagged third-order moments. Table 14 offers further evidence that the forward variances violate the martingale restriction. We compare the forecasting performances of two forward-variance models: the "random-walk model" and "third-order-moment based model" in out-of-sample forecasting. We use the last daily change in forward variance as the predictor of the next

Table 13
Summary statistics of p-values of GST before and after financial crisis.

	T_1T_2		T_1T_3		T_1T_4		T_1T_5	
	Statistic	p-Values	Statistic	p-Values	Statistic	p-Values	Statistic	p-Values
	Panel A: GST test results before 2008 (using JT (2012) forward variance)							
M (1,0)	15.82636	0.539	21.3981	0.210	9.0256	0.546	8.1735	0.521
M (1,1)	7.619281	0.658	7.562679	0.278	11.29164	0.328	9.628671	0.358
M (1,2)	61.58407	0.186	5.698009	0.6	13.8763	0.576	1.514269	0.97
M (1,3)	49.91985	0.376	17.37301	0.234	32.06433	0.484	18.66409	0.43
M (1,4)	82.09488	0.136	7.193949	0.556	23.05881	0.566	8.459168	0.74
N	3732		3732		3732		3732	
	Panel B: GST test results before 2008 (using DK (2012) forward variance)							
M (1,0)	14.2715	0.690	19.7881	0.193	6.2547	0.445	6.0713	0.501
M (1,1)	7.588099	0.77	12.62408	0.304	12.43643	0.334	13.64129	0.212
M (1,2)	58.52827	0.136	15.42608	0.504	14.20708	0.51	5.116835	0.902
M (1,3)	49.52808	0.434	30.35216	0.384	23.86599	0.544	13.55536	0.558
M (1,4)	81.41104	0.06	26.38818	0.444	25.26951	0.446	4.675217	0.832
N	3732		3732		3,732		3732	
	Panel C: GST test results after 2008 (using JT (2012) forward variance)							
M (1,0)	56.187**	0.020	52.181**	0.036	17.126	0.361	18.174	0.312
M (1,1)	10.51644	0.302	36.90894	0.228	5.875266	0.34	5.760777	0.488
M (1,2)	0.181315	0.988	6.425705	0.964	2.963678	0.8	2.623107	0.724
M (1,3)	20.69635**	0.042	97.939**	0.046	15.48029	0.3	16.67291	0.152
M (1,4)	7.946577	0.59	14.76172	0.93	4.362112	0.724	3.406964	0.702
N	3,732		3,732		3,732		3,732	
	Panel D: GST test results after 2008 (using DK (2012) forward variance)							
M (1,0)	59.251**	0.018	60.089**	0.047	22.766	0.218	20.350	0.275
M (1,1)	11.2702	0.196	15.07947	0.046	9.986026	0.208	7.500918	0.198
M (1,2)	1.680719	0.918	1.593202	0.868	3.796476	0.632	2.936266	0.72
M (1,3)	23.6269**	0.034	21.010**	0.042	13.25688	0.258	12.03743	0.252
M (1,4)	15.06312	0.468	7.354095	0.45	7.177261	0.578	3.859952	0.742
N	3732		3732		3732		3732	

Note: We compute bootstrapped p-values using 5000 bootstrap replications, with preliminary lag order 10. N denotes the number of observations. ***, **, and * indicate the coefficient is significantly different from 0 at the 1%, 5%, and 10% levels, respectively.

Table 14
Comparisons of out-of-sample forecasting errors.

	Model1 vs. Model2	MSE1/MSE2	MDM test statistic	95 Percent CV
	Panel A: JT (2012) forward variance			
Group 1	Third-order moment vs. Random walk	0.6953	4.6243**	1.6457
Group 2	Third-order moment vs. Random walk	0.6913	4.7688**	1.6457
	Panel B: DK (2012) forward variance			
Group 1	Third-order moment vs. Random walk	0.6999	4.4251**	1.6457
Group 2	Third-order moment vs. Random walk	0.6882	4.5904**	1.6457

Note: This table reports the results of one-day-ahead, non-nested forecast comparisons. We first divide the full forward variance series into two equal parts. The first part (Jan. 1996 to June, 2003) is considered as the initial estimation sample. We then forecast the one-day ahead forward variance through the rolling method in the second part (July 2003 to Oct. 2010). For each group, two models are compared. Model 1 always uses lagged third-order moment as a predictive variable; Model 2 uses the last daily change in forward variance as the predictor. The column labeled "MSE1/MSE2" reports the ratio of the root-mean-squared forecasting error of Model 1 to Model 2. The modified Diebold-Mariano Test statistic, altered to test for forecast encompassing (Harvey et al., 1998), appears in the column labeled as "MDM Test Statistic." The column labeled "95 Percent CV" reports the 95% critical value for this statistic based on t_{n-1} distribution, where n is the number of out-of-sample forecasting observations. The null hypothesis is that Model 2 encompassing Model 1. The ***, ** and * indicate the coefficient significantly different from 0 at 1%, 5% and 10% levels respectively. The specification of the models are as follows:
Model 1: $\Delta f v_t = \alpha + \beta_1 \Delta f v_{t-1}^3 + \beta_2 \Delta f v_{t-2}^3 + \beta_3 \Delta f v_{t-3}^3 + \beta_4 \Delta f v_{t-4}^3 + \beta_5 \Delta f v_{t-5}^3 + \epsilon_t$.
Model 2: $\Delta f v_t = \Delta f v_{t-1} + \epsilon_t$.
where ϵ_t is i.i.d. error term.

period in the random-walk model and use lagged third-order moments of daily change in forward variance to forecast the forward variance in the next period. The model specifications are given as follows:

$$\Delta f v_t = \Delta f v_{t-1} + \epsilon_t \qquad \text{random-walk model}$$

$$\Delta f v_t = \alpha + \beta_1 (\Delta f v_{t-1})^3 + \beta_2 (\Delta f v_{t-2})^3 + \beta_3 (\Delta f v_{t-3})^3 + \beta_4 (\Delta f v_{t-4})^3 + \beta_5 (\Delta f v_{t-5})^3 + \epsilon_t \qquad \text{third-order-moment based model}$$

where $\Delta f v_t$ is the forward variance change at time t and ϵ_t is i.i.d. error term. The forecasting study is conducted as follows: we first divide the full forward variance series into two equal parts. The first part (January 1996 to June, 2003) is considered as the initial estimation sample. We then forecast the one-day ahead forward variance through the rolling method in the second part (July 2003 to October 2010). We use two statistics to gauge the out-of-sample forecasting power and conduct the out-of-sample

Table 15
Forecast daily change in forward variance using the lagged third-order moments with different horizons.

	Lag(1)		Lag(2)		Lag(3)		Lag(4)		Lag(5)		R^2
	Coefficient	Std.	Coefficient	Std.	Coefficient	Std.	Coefficient	Std.	Coefficient	Std.	
	Panel A: forecasting 2-days-ahead daily change in Forward Variance										
G1	−59.621***	(15.094)	41.358***	(9.907)	−48.378***	(18.223)	7.783	(38.684)	−32.824	(21.423)	0.069
G2	−55.681**	(22.914)	45.171***	(18.704)	−51.645	(31.997)	−27.485	(68.109)	−25.930	(37.188)	0.041
NG1	−43.161***	(6.029)	26.830***	(4.261)	−34.132***	(8.644)	8.472	(19.659)	−24.098	(14.920)	0.094
NG2	−35.288**	(13.942)	26.209**	(12.081)	−36.254	(22.688)	−17.007	(44.978)	−15.231	(20.120)	0.044
	Panel B: forecasting 3-days-ahead daily change in forward variance										
G1	33.763**	(13.844)	−16.147	(10.811)	−14.654	(37.181)	−17.254	(22.122)	−37.043	(22.916)	0.044
G2	41.521**	(18.793)	−29.959	(20.096)	−47.083	(57.826)	−20.649	(45.354)	−50.323	(44.956)	0.036
NG1	22.359***	(7.002)	−9.904	(6.235)	−5.899	(20.208)	−12.410	(11.011)	−23.809	(14.820)	0.053
NG2	24.233**	(12.075)	−19.950	(12.365)	−30.825	(37.874)	−13.661	(26.228)	−35.791	(26.111)	0.043
	Panel C: forecasting 4-days-ahead daily change in forward variance										
G1	−14.836	(11.096)	−32.638	(30.623)	−10.323	(20.559)	−46.8305**	(19.038)	10.920	(12.025)	0.033
G2	−34.575	(21.068)	−63.389	(53.777)	−20.856	(43.292)	−60.076	(42.480)	−0.602	(23.114)	0.028
NG1	−9.611	(6.659)	−17.671	(16.982)	−7.559	(10.247)	−30.151	(19.133)	7.357	(6.612)	0.039
NG2	−20.886	(12.906)	−41.066	(35.219)	−16.337	(25.069)	−42.464	(26.952)	−4.958	(13.438)	0.036
	Panel D: forecasting 5-days-ahead daily change in forward variance										
G1	−27.903	(28.348)	−3.251	(20.776)	−40.768	(25.304)	16.074	(10.181)	13.604	(14.873)	0.032
G2	−53.927	(53.966)	−6.669	(43.456)	−51.354	(41.385)	10.604	(20.914)	24.684	(16.533)	0.024
NG1	−14.596	(15.391)	−3.213	(10.564)	−25.821	(16.059)	10.417	(6.398)	9.766	(8.583)	0.039
NG2	−32.731	(34.756)	−4.142	(25.292)	−35.230	(24.291)	4.335	(11.436)	16.365	(11.027)	0.030

Note: This table presents the results of forecasting the change of forward variance using lagged third order moments within one week horizon. G1 and G2 denote the first and the second maturity groups of JT (2012) forward variance. NG1 and NG2 denote the first and the second maturity groups of DK (2012) forward variance. The standard errors of the coefficients estimates are computed following a robust procedure taking into account of the heteroscedastic and autocorrelated error structure and reported in the brackets next to the coefficient estimates. The ***, ** and * indicate the coefficient significantly different from 0 at 1%, 5% and 10 % levels respectively.

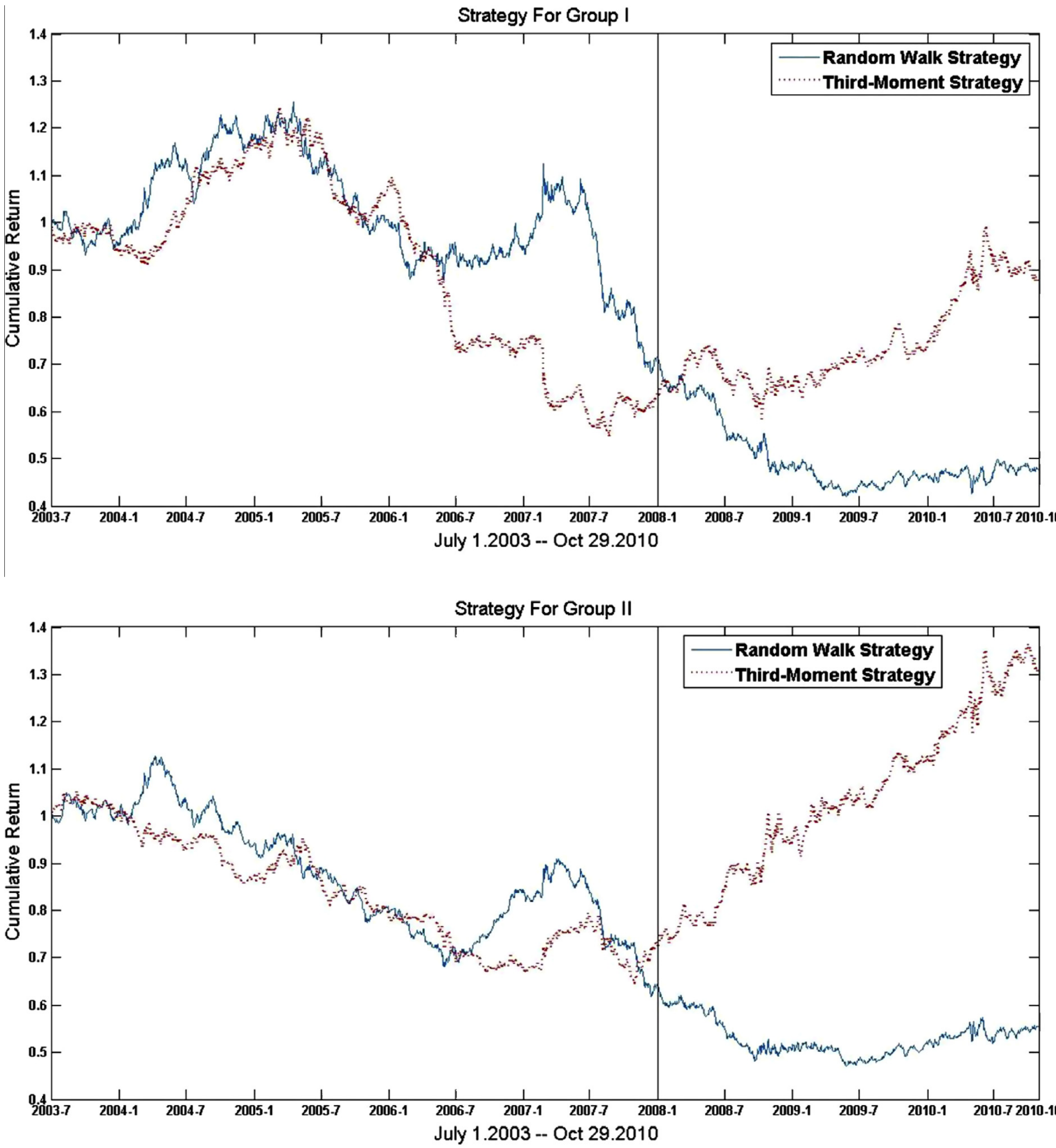


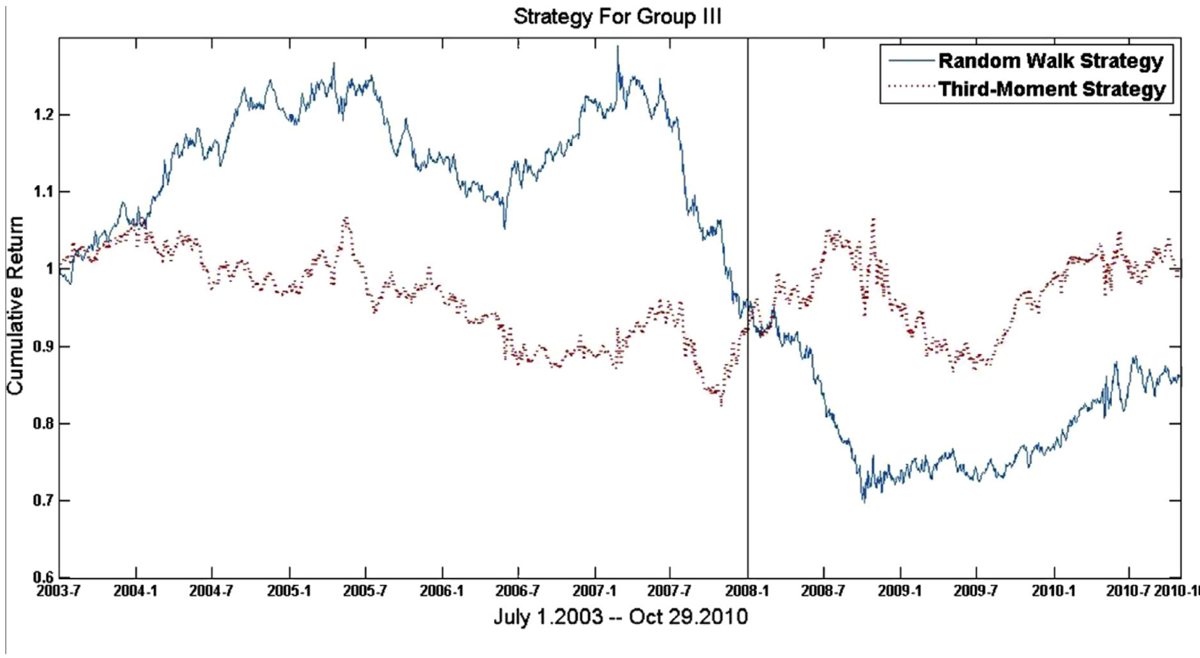
Fig. 3. The cumulative returns of trading strategies based on the random-walk model and the third-moment based model. *Note:* This figure plots the cumulative returns of trading strategies based on random-walk model and the third-moment-based model. The estimation period ranges from Jan. 1996 to June 2003. The trading period ranges from July 2003 to Oct. 2010. We assume the onset of the financial crisis is in Jan. 2008.

forecasting study in the first two maturity groups using both JT (2012) and DK (2012) forward variance measures.

MSEa/MSEb[15] is the ratio of the mean squared forecasting error of the forecasting model (i.e. third-order-moment based model) to that of the benchmark model (i.e. random-walk model). This ratio in Table 14 shows that the third-order-moment based model outperforms the benchmark random-walk model with ratios significantly less than 1 for all four groups. In addition, the modified Diebold-Mariano (MDM) test is used to test the null hypothesis that the benchmark model incorporates all the information about the forward variance in the next period against the alternative hypothesis that lagged third-order moments provide additional information. Given two time series of forecast errors (e_{1t}, e_{2t}), and a specified loss function $g(e)$, the null hypothesis of MDM test is $E[g(e_{1t}) - g(e_{2t})] = 0$. Specifically, the MDM test is based on the sample mean of d_t, where $d_t = g(e_{1t}) - g(e_{2t})$, and this test statistic follows a t-distribution. All the MDM statistics estimates are around 4.5, with a 5% critical value at 1.6457. This clearly indicates the rejection of the null hypothesis.

We finally investigate the forecasting horizons of the third-order moment-based model. Table 15 shows that the predicting power of the lagged third-order moments diminishes quickly with increasing forecasting horizons. For most of the forward variance

[15] For n observations, $MSEa/MSEb = \Sigma e_a^2/\Sigma e_b^2$, where e_a is the forecasting error of model a, and e_b is the forecasting error of model b.



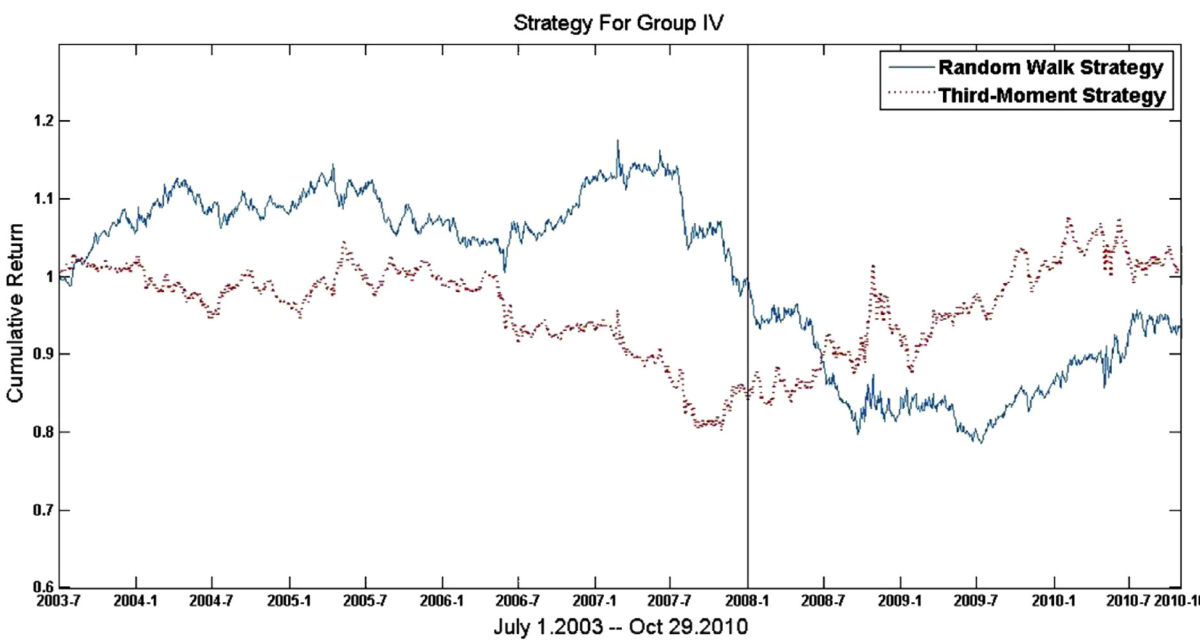


Fig. 3 (*continued*)

series, forecastability based on the lagged third-order moments disappears when the horizons are longer than three days except for Group 1 of the JT (2012) forward variance series. The number of significant lagged third-order moments is also diminishing from Panels A to D, which is a further evidence indicating decreasing forecastability of third-order moments.

5.2. Constructing trading strategies

We further discuss how to contruct a trading strategy based on the lagged third-moment based model. From Eq. (7), the spot variance of options can be regarded as a portfolio composed of a basket of put options with different strike prices and a basket of call options with different strike prices. From Eq. (8), the forward variance can be constructed by buying long-maturity spot variance and selling short-maturity variance. In sum, the forward variance by construction is a portfolio of a long position in long-maturity options and a short position in short-maturity options. This portfolio hereafter is called the forward-varaince mimicking portfolio.[16]

In our prediction model, we respectively use one-period lagged forward variance (random-walk model) and lagged third-order moments (third-order moment-based model) to predict current foward varaince. Similarly, we follow the essence of our prediction models to construct the trading strategy. If the predicted current forward variance increases (i.e. foward variance change is positive), we long the forward variance mimicking portfolio at market opening and sell this portfolio at market closing. Conversely, if the

[16] The detailed steps of constructing the mimicking portfolio are given in Appendix A.

predicted current forward variance decreases (i.e. forward varaince change decreases), we short the forward variance mimicking portfolio at the market opening and long this portfolio back at market closing.[17]

Similar to the steps of examining the performance of forecasting models, we first divide the full forward variance series into two equal parts. The first part is considered as the initial estimation sample. We then use this estimation result to calculate forward variance and construct trading strategies. We take the trading performance of the random-walk model as the benchmark. Under the null hypothesis that the S&P 500 options market is efficient, the trading performance of the third-order moment-based model should not outperform the one of random-walk model. In Fig. 3 we observe that the trading performance of third-order moment-based model doesn't outperform the one of random-walk model in four different maturity groups before financial crisis. However, the trading performance of third-order-moment based model significantly outperforms the one of random-walk model in four different maturity groups after the financial crisis. This outperformance is especially significant in short-maturity groups. These results are consistent with our findings that the options market is not efficient, and we can use the information of the lagged third-order moments of forward variance to predict future forward varaince. From this prediction, we can construct a trading strategy that significantly outperforms the one based on the random-walk model. This exercise of a trading strategy demonestrates that our results are not only of statistical significance, but also of economic importance.

6. Concluding remarks

In this study, we use GST of Hong (1999) to test the martingale restriction of the forward variance in the S&P 500 index options market. Contrary to previous studies such as JT (2012), we find that third-order conditional moments can predict future changes in the forward variance, especially in the short-maturity groups. We also consider an alternative variance measure that accounts for jump effects and find similar results. The negative correlation between the daily changes of forward variance and its third-order conditional moments also can be explained by the jump effect in the S&P 500 index, suggesting that the price jumps of the index has a significant predicting power on forward variances of index options.

We also discover that the last two maturity groups of forward variance, which are less actively traded, exhibit the martingale properties. This evidence indicates that the illiquidity effect as suggested in JT (2012) would not be a cause of market inefficiency in relatively long-maturity groups.

By using the information of lagged third-order moments of forward variance, we conduct an out-of-sample forecasting and find that information of lagged third-order moments indeed has a strong predicting power on future forward variance. Our findings confirm the "skewness-in-mean" effect of forward variance and provide an effective prediction indicator for the forward variances of S&P 500 index options.

We further use the information of model predictions and construct a forward variance mimicking portfolio to form a trading strategy. We find that the trading performance based on our proposed prediction model significantly outperform the one based on the random-walk model in short-maturity groups after the recent financial crisis. This finding provides an economically significant support of our testing results as well as a profitable application of our testing results.

[17] In this strategy, the transaction cost is not considered, and we assume that short selling is allowed.

Our study thus provides robust and consistent evidence about the inefficiency of S&P 500 index options market. Specifically, our study presents an effective forecasting structure for the forward variances of S&P 500 index options. Our proposed profitable forward variance-based trading strategy illustrates a pratical application of our empirical results and offers a new venue for future research about contructing profitable trading strategies in the S&P 500 options market.

Appendix A

In this Appendix A we first describe the details of Hong (1999) Generalized Spectral Test. Second, we introduce BNS (2006) jump testing method. Third, we outline the detailed steps of constructing the forward variance mimicking portoflio.

A.1. Hong (1999) Generalized Spectral Test (GST)

The basic idea of GST is to transform a strictly stationary series $\{Y_t\}$ and consider the spectrum of the transformed series. Suppose that $\{Y_t\}$ has marginal characteristic function $\varphi(u) \equiv E\left[e^{\mathbf{i}uY_t}\right]$ and a pairwise joint characteristic function $\varphi_j(u,v) \equiv E\left[e^{\mathbf{i}(uY_t+vY_{t-|j|})}\right]$, where $\mathbf{i} \equiv \sqrt{-1}, u, v \in (-\infty,\infty)$, and $j = 0, \pm 1, \ldots$. Define the covariance function between the transformed variables $e^{\mathbf{i}uY_t}$ and $e^{\mathbf{i}(uY_t+vY_{t-|j|})}$ as:

$$\sigma_j(u,v) \equiv cov(e^{\mathbf{i}uY_t}, e^{\mathbf{i}(uY_t+vY_{t-|j|})}) \tag{A.1}$$

Straightforward algebra yields $\sigma_j(u,v) = \varphi_j(u,v) - \varphi(u)\varphi(v)$. Because $\varphi_j(u,v) = \varphi(u)\varphi(v)$ for all u, v if and only if Y_t and $Y_{t-|j|}$ are independent, $\sigma_j(u,v)$ can capture any type of pairwise serial dependence over various lags, including those with zero autocorrelation.

When $\sup_{u,v\in(-\infty,\infty)}\sum_{j=-\infty}^{\infty}|\sigma_j(u,v)| < \infty$, the Fourier transform of $\sigma_j(u,v)$ exists:

$$f(w,u,v) \equiv \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\sigma_j(u,v)e^{-\mathbf{i}jw}, \quad w \in [-\pi,\pi]. \tag{A.2}$$

Like $\sigma_j(u,v), f(w,u,v)$ can capture all pairwise serial dependencies in $\{Y_t\}$ over various lags. It requires no moment condition. When $var(Y_t) = \sigma^2$ exists, the power spectrum $H(w)$ of $\{Y_t\}$ can be obtained by differentiating $f(w,u,v)$ with respect to (u,v) at $(0,0)$:

$$H(w) \equiv \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma(j)e^{-\mathbf{i}jw} = -\frac{\partial^2 f(w,u,v)}{\partial u \partial v}|_{(u,v)=(0,0)}, \tag{A.3}$$

where $\gamma(j) \equiv cov(Y_t, Y_{t-|j|})$. For this reason, we call $f(w,u,v)$ a "generalized spectral density", although it does not have the mathematical properties of a probability density.

Hong (1999, Theorem 1) shows that $f(w,u,v)$ can be consistently estimated by

$$\hat{f}_n(w,u,v) \equiv \frac{1}{2\pi}\sum_{j=1-n}^{n-1}(1-|j|/n)^{\frac{1}{2}}k(j/p)\hat{\sigma}_j(u,v)e^{-\mathbf{i}jw}, \tag{A.4}$$

where $\widehat{\sigma}_j(u,v) \equiv \hat{\varphi}_j(u,v) - \hat{\varphi}_j(u,0)\hat{\varphi}_j(0,v)$ is the empirical generalized covariance, $\hat{\varphi}_j(u,v) \equiv (n-|j|)^{-1}\sum_{t=|j|+1}^{n} e^{\mathbf{i}(uY_t+vY_{t-|j|})}$ is the empirical pairwise empirical characteristic function, $p \equiv p_n$ is a bandwidth or lag order, and $k(\cdot)$ is a kernel function or "lag window". Commonly used kernels include Bartlett, Daniell, Parzen and Quadratic-Spectral kernels. The factor $(1-|j|/n)^{\frac{1}{2}}$ modifies the variance of $\hat{\sigma}_j(u, v)$. It could be replaced by 1, but it gives better finite sample performance for the tests based on $\hat{f}_n(w, u, v)$.

When $\{Y_t\}$ is IID, $f(w, u, v)$ becomes a "flat" generalized spectrum:

$$f_0(w, u, v) \equiv \frac{1}{2\pi}\sigma_0(u, v), \quad w \in [-\pi, \pi]. \tag{A.5}$$

Any deviation of $f(w, u, v)$ from the flat spectrum $f_0(w, u, v)$ is evidence of serial dependence. Thus, to detect serial dependence, we can compare $\hat{f}_n(w, u, v)$ with the estimator:

$$\hat{f}_0(w, u, v) \equiv \frac{1}{2\pi}\hat{\sigma}_0(u, v), \quad w \in [-\pi, \pi]. \tag{A.6}$$

Once the existence of generic serial dependence is detected, one may like to further explore the nature of serial dependence. Different types of serial dependence have different implications for predictability of Y_t. If $\{Y_t\}$ is MDS, for example, then serial dependnce in higher moments will not help predict the conditional mean of Y_t.

To explore the nature of serial dependence, one can compare the derivative estimators:

$$\hat{f}_n^{(0,m,l)}(w, u, v) \equiv \frac{1}{2\pi}\sum_{j=1-n}^{n-1}(1 - |j|/n)^{\frac{1}{2}}k(j/p)\hat{\sigma}_j^{(m,l)}(u, v)e^{-\mathbf{i}jw}, \tag{A.7}$$

$$\hat{f}_n^{(0,m,l)}(w, u, v) \equiv \frac{1}{2\pi}\hat{\sigma}_0^{(m,l)}(u, v), \tag{A.8}$$

where $\hat{\sigma}_j^{(m,l)} \equiv \partial^{m+l}\hat{\sigma}_j(u, v)/\partial^m u \partial^l v$ for $m, l \geqslant 0$. Just as the characteristic function can be differentiated to generate various moments, generalized spectral derivatives can capture various specific aspects of serial dependence, thus providing information on possible types of serial dependence.

Hong (1999) proposes a class of tests based on the quadratic norm:

$$\begin{aligned} Q\left(\hat{f}_n^{(0,m,l)}, \hat{f}_0^{(0,m,l)}\right) &\equiv \int_{-\pi}^{\pi} |\hat{f}_n^{(0,m,l)}(w, u, v) - \hat{f}_0^{(0,m,l)}(w, u, v)|^2 dw dW_1(u) dW_2(v) \\ &= \frac{2}{\pi}\int \sum_{j=1}^{n-1} k^2(j/p)(1 - j/n)\left|\hat{\sigma}_j^{(m,l)}(u, v)\right|^2 dW_1(u) dW_2(v) \end{aligned}$$

Here, the second equality follows Parseval's identity, and the unspecified integrals are taken over the support of $W_1(\cdot)$ and $W_2(\cdot)$, which are positive non-decreasing weighting functions that set weight about zero equally. An example of $W_1(\cdot)$ and $W_2(\cdot)$ is the $\mathbf{N}(0, 1)$ CDF, which is commonly used in the characteristic function literature. The test statistic is a standardized version of the quadratic form:

$$\begin{aligned} M(m, l) \equiv &\left[\int \sum_{j=1}^{n-1} k^2(j/p)(n - j) \times |\hat{\sigma}_j^{(m,l)}(u, v)|^2 dW_1(u) dW_2(v) \right. \\ &\left. - \hat{C}_0^{(m,l)}\sum_{j=1}^{n-1} k^4(j/p)\right] \div \left[\hat{D}_0^{(m,l)}\sum_{j=1}^{n-2} k^4(j/p)\right]^{1/2} \end{aligned} \tag{A.9}$$

The centering and standardization factors can be defined as:

$$\hat{C}_0^{(m,l)} \equiv \int \hat{\sigma}_0^{(m,m)}(u, -u) dW_1(u) \int \hat{\sigma}_0^{(l,l)}(v, -v) dW_2(v)$$

$$\hat{D}_0^{(m,l)} \equiv 2\int |\hat{\sigma}_0^{(m,m)}(u, u')|^2 dW_1(u) dW_1(u') \int |\hat{\sigma}_0^{(l,l)}(v, v')|^2 dW_2(v) dW_2(v').$$

Given (m, l), $M(m, l)$ is asymptotically one-sided $\mathbf{N}(0, 1)$ under the null hypothesis of serial independence. For a kernel $k(\cdot)$with unbounded support, $M(m, l)$ employs all $n - 1$ lags in the sample. This is desirable when the alternative has persistent serial dependence. Non-uniform kernels, such as the Daniell kernel $k(z) = sin(\pi z)/\pi z, z \in (-\infty, \infty)$, usually weight down higher order lags. This is expected to enhance good power of the tests in an empirical study, because economic agents normally discount past information. This is particularly true in stock or options markets, where investors digest information relatively fast. Hong (1999) finds that the Daniell kernel maximizes the power of $M(m, l)$ over a class of kernels that include Parzen and Quadratic-Spectral kernels for spectral density estimation.

Following this procedure, once we have the value of testing statistics (5), and we know the fact that $M(m, l)$ is asymptotically one-sided $\mathbf{N}(0, 1)$ under the null hypothesis of serial independence, then we can test the serial independence of time series. By changing the specification of (m, l), we can test a variety of serial dependences as given in Table 1. When the sample size is finite and greater than 100, as suggested in Hong and Lee (2003), the wild bootstrap procedure with 300 bootstrap replications is adequate for calculating the p-value for the GST test.

A.2. BNS (2006) Jump testing method

In applying the BNS (2006) method, assume the log of stock price $s_t = logS_t$ follows a general jump diffusion model:

$$ds_t = \mu_t d_t + \sigma_t dB_t + J_t dq_t, \tag{A.10}$$

where μ_t and σ_t are the drift and diffusion terms, B_t is the standard Brownian motion, dq_t is Poisson jump process, and J_t is the jump size following a normal distribution with mean μ_J and standard deviation σ_J. Intraday return can be expressed as $r_{t,i} = s_{t,j\cdot\tau} - s_{t,(j-1)\cdot\tau}$, where τ is the sampling interval and $r_{t,j\cdot\tau}$ denotes the j-th intraday return in day t.

The RV and BV can be written as:

$$RV_t \equiv \sum_{i=1}^{M} r_{t,i}^2, \tag{A.11}$$

$$BV_t \equiv \frac{\pi}{2}\frac{M}{2}\sum_{i=2}^{M}|r_{t,i}||r_{t,i-1}|, \tag{A.12}$$

where M is the number of intraday observations. Similar to Barndorff-Nielsen and Shephard, 2004, we choose the following jump test statistics:

$$ZJ_t \equiv \frac{\frac{RV_t - BV_t}{RV_t}}{\sqrt{\left(\frac{\pi^2}{4} + \pi - 5\right)\frac{1}{M}\max\left(1, \frac{TP_t}{BV_t^2}\right)}}, \tag{A.13}$$

where TP_t is the tripower quarticity as defined in Barndorff-Nielsen and Shephard, 2004:

$$TP_t = \frac{M}{M - 2}\cdot\frac{M}{[\Gamma(7/6)\Gamma(1/2)]^3}\cdot\sum_{i=3}^{M}|r_{t,i}|^{4/3}|r_{t,i-1}|^{4/3}|r_{t,i-2}|^{4/3}, \tag{A.14}$$

where $\Gamma(\cdot)$ is the Gamma function. The test statistic TP_t is asymptotically normally distributed. If this test statistics exceed the 1% critical value of normal distribution, we reject the null hypothesis that there is no jump within day t.

A.3. Constructing forward-variance mimicking portfolio

By Eqs. (7) and (8), the spot variance is equivalent to a portfolio of call options and put options with different strike prices. The forward variance is equivalent to a portfolio of buying long-maturity spot variance and selling a short-maturity spot variance. Combining the above two definitions, forward variance can be regarded as a portfolio of long positions in long-maturity call and put options and short positions in short-maturity call and put options. In detail, in Eq. (7), spot variance is defined as:

$$v(t,T) = \frac{2\exp[r(T-t)]}{T-t}\left[\int_0^{F_0}\frac{P_t(K,T)}{K^2}dK + \int_{F_0}^{\infty}\frac{C_t(K,T)}{K^2}dK\right]$$

In practice, the options' exercise price is discrete, and the intergrals in (7) can be expressed in discrete form. For simplicity, we ignore $\left(\frac{2\exp[r(T-t)]}{T-t}\right)$ and the spot variance can be expressed as:

$$v(t,T) \approx \sum_{i=1}^{n}\frac{P_t(K_i,T)}{K_i^2}\mathbf{I}(K_i < F_0)\Delta K + \sum_{j=1}^{m}\frac{C_t(K_j,T)}{K_j^2}\mathbf{I}(K_j > F_0)\Delta K \tag{A.15}$$

Here, $\mathbf{I}(K_i < F_0)$ is the indicator function, and it equals 1 when $K_i < F_0$, and zero otherwise. There are n strike prices $\{K_i\}_{j=1}^{n}$ less than the forward index level F_0, and m strike prices $\{K_j\}_{j=1}^{m}$ greater than the forward index level.

By Eq. (8), the definition of forward variance can be given as:

$$v(t;T_1,T_j) = \left(\frac{T_j-t}{T_j-T_1}\right)v(t,T_j) - \left(\frac{T_1-t}{T_j-T_1}\right)v(t,T_1)$$

For simplicity, we ignore the linear weight in this definition, and use (A.15) to define the forward variance in discrete form, which is shown as:

$$v(t;T_1,T_j) \approx \underbrace{\left[\sum_{i=1}^{n}\frac{P_t(K_i,T_j)}{K_i^2}\mathbf{I}(K_i<F_0)\Delta K + \sum_{j=1}^{m}\frac{C_t(K_j,T_j)}{K_j^2}\mathbf{I}(K_j>F_0)\Delta K\right]}_{\text{a portfolio of call and put options with maturity } T_j \text{ in different exercise prices}} - \underbrace{\left[\sum_{i=1}^{n}\frac{P_t(K_i,T_1)}{K_i^2}\mathbf{I}(K_i<F_0)\Delta K + \sum_{j=1}^{m}\frac{C_t(K_j,T_1)}{K_j^2}\mathbf{I}(K_j>F_0)\Delta K\right]}_{\text{a portfolio of call and put options with maturity } T_1 \text{ in different exercise prices}} \tag{A.16}$$

We then use this portfolio to mimic forward variance, and contruct the the trading strategy as given in Section 5.2.

References

Aï-Sahalia, Y., Duarte, J., 2003. Nonparametric option pricing under shape restrictions. Journal of Econometrics 116, 9–47.

Ang, A., Bekaert, G., 2007. Stock return predictability: is it there? Review of Financial Studies 20, 651–707.

Angelini, P., Nobili, A., Picillo, C., 2011. The interbank market after August 2007: what has changed, and why? Journal of Money, Credit and Banking 43, 923–958.

Bakshi, G., Kapadia, N., Madan, D., 2003. Stock return characteristics, skew laws and the differential pricing of individual equity options. Review of Financial Studies 16, 101–143.

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

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

Bollerselv, T.M., Todorov, V., 2011. Tails, and risk premia. Journal of Finance 66, 2165–2211.

Campbell, J.Y., Thompson, S.B., 2008. Predicting excess stock returns out of sample: can anything beat the historical average? Review of Financial Studies 21, 1509–1531.

Campbell, J.Y., Lo, A.W., MacKinlay, C.A., Adamek, P., Viceira, L.M., 1997. The Econometrics of Financial Markets. Princeton University Press, Princeton, NJ.

Carr, P., Madan, D.B., 2005. A note on sufficient conditions for no arbitrage. Finance Research Letters 2, 125–130.

Carr, P., Wu, L.R., 2009. Variance risk premiums. Review of Financial Studies 22, 1311–1341.

Cremers, M., Weinbaum, D., 2010. Deviations from put-call parity and stock return predictability. Journal of Financial and Quantitative Analysis 45, 335–367.

Das, S., Sundaram, R., 1997. Taming the Skew: Higher Order Moments in Modeling Asset Price Processes in Finance, NBER working paper No.5967.

Deo, R.S., 2000. Spectral tests of the Martingale hypothesis under conditional heteroskedasticity. Journal of Econometrics 99, 291–315.

Dominguez, M.A., Lobato, I.N., 2003. Testing the martingale difference hypothesis. Econometric Reviews 22, 351–377.

Du, J., 2012. Essays on Variance Risk, Jump Risk and Options Information (Doctoral Dissertations) 1911–2013. Paper 405. <http://scholarworks.umass.edu/dissertations1/405>

Du, J., Kapadia, N., 2012. The Tail in the Volatility Index, Working paper, U. Massachusetts, Amherst.

Durlauf, S., 1991. Spectral based testing of the martingale hypothesis. Journal of Econometrics 50, 355–376.

Escanciano, J.D., Mayoral, S., 2010. Data-driven smooth tests for the martingale difference hypothesis. Computational Statistics & Data Analysis 54, 1983–1998.

Fama, E., 1991. Efficient capital markets II. The Journal of Finance 46, 1575–1617.

Goyal, A., Welch, I., 2008. A Comprehensive look at the empirical performance of equity premium prediction. Review of Financial Studies 21, 1455–1508.

Guo, H., Wang, K., Zhang, H., 2014. Good Jumps, Bad Jumps, and Conditional Equity Premium, University of Cincinnati Working Paper.

Harvey, C.R., Whaley, R.E., 1992. Market volatility prediction and the efficiency of the S&P 100 index option market. Journal of Financial Economics 31, 43–73.

Harvey, D., Stephen, L., Newbold, P., 1998. Tests for forecast encompassing. Journal of Business and Economic Statistics 16, 254–259.

Hong, Y., 1999. Hypothesis testing in time series via the empirical characteristic function: a generalized spectral density approach. Journal of American Statistical Association 84, 1201–1220.

Hong, Y., Lee, T.H., 2003. Inference on predictability of foreign exchange rates via generalized spectrum and nonlinear time series models. Review of Economics and Statistics 85, 1048–1062.

Hong, Y., Li, H.T., Zhao, F., 2004. Out-of-sample performance of discrete-time spot interest rate models. Journal of Business and Economic Statistics 22, 457–473.

Hong, Y., Lin, H., Wu, C., 2012. Are corporate bond market returns predictable? Journal of Banking and Finance 36, 2216–2232.

Huang, X., Tauchen, G., 2005. The relative contribution of jumps to total price variance. Journal of Financial Econometrics 3, 456–499.

Jacquier, E., Okou, C., 2012. Segregating Continuous Volatility from Jumps in Long-Run Risk-Return Trade-Offs, Working paper, HEC Montreal and MIT.

Jiang, G.J., Tian, Y.S., 2005. The model-free implied volatility and its information content. Review of Financial Studies 18, 1305–1342.

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

Jiang, G.J., Tian, Y.S., 2012. A random walk down the options market. Journal of Futures Markets 32, 505–535.

Kelly, B., 2014. Tail risk and asset prices. Review of Financial Studies. http://dx.doi.org/10.1093/rfs/hhu039.

Kraus, A., Litzenberger, R., 1976. Skewness preference and the valuation of risk assets. Journal of Finance 31, 1085–1100.

Kraus, A., Litzenberger, R., 1983. On the distributional conditions for a consumption-oriented three moment CAPM. Journal of Finance 38, 1381–1391.

Lee, S.S., Mykland, P.A., 2008. Jumps in financial markets: a new nonparametric test and jump dynamics. Review of Financial Studies 21, 2535–2563.

Li, J., 2012. Measuring Stock Price Jumps: A Delta-Hedging Perspective, Working paper, Princeton University.

Lo, A.W., Mackinlay, A.C., 1988. Stock market prices do not follow random walks: evidence from a simple specification test. Review of Financial Studies 1, 41–66.

Longstaff, F.A., 1995. Option pricing and the martingale restriction. Review of Financial Studies 8, 1091–1124.

Longstaff, F.A., 2000. The term structure of very short-term rates: new evidence for the expectations hypothesis. Journal of Financial Economics 58, 397–415.

Merton, R.C., 1976. Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics 3, 125–144.

McPherson, M.Q., Palardy, J., 2007. Are international stock returns predictable?: an examination of linear and non-linear predictability using generalized spectral tests. Journal of International Financial Markets, Institutions and Money 17, 452–464.

Rapach, D.E., Strauss, J.K., Zhou, G.F., 2010. Out-of-sample equity premium prediction: combination forecasts and links to the real economy. Review of Financial Studies 23, 821–862.

Rubinstein, M.E., 1973. The fundamental theorem of parameter-preference security valuation. Journal of Financial and Quantitative Analysis 8, 61–69.

Sekkel, R., 2011. International evidence on bond risk premia. Journal of Banking and Finance 35, 174–181.

Wright, J.H., Zhou, H., 2009. Bond risk premia and realized jump risk. Journal of Banking and Finance 33 (12), 2333–2345.

Xing, Y., Zhang, X., Zhao, R., 2010. What does individual option volatility smirk tell us about future equity returns? Journal of Financial and Quantitative Analysis 45, 641–662.

Yang, J., Su, X.J., Kolari, J.W., 2008. Do Euro exchange rates follow a martingale? some out-of-sample evidence. Journal of Banking and Finance 32, 729–740.

Yang, J., Zhou, Y., Wang, Z., 2010. Conditional coskewness in stock and bond markets: time-series evidence. Management Science 56, 2031–2049.

Zhang, B., Zhou, H., Zhu, H., 2009. Explaining credit default swap spreads with the equity volatility and jump risks of individual firms. Review of Financial Studies 22, 5099–5131.