Journal of Econometrics 194 (2016) 153–186

Contents lists available at ScienceDirect

Journal of Econometrics

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

A nonparametric test of a strong leverage hypothesis☆

Oliver Linton[a], Yoon-Jae Whang[b], Yu-Min Yen[c,*]

[a] Department of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge CB3 9DD, United Kingdom

[b] Department of Economics, Seoul National University, Seoul 151-742, Republic of Korea

[c] Department of International Business, National Chengchi University, 64, Sec. 2, Zhi-nan Rd., Wenshan, Taipei 116, Taiwan

ARTICLE INFO

Article history:
Received 22 July 2013
Received in revised form
30 January 2016
Accepted 10 February 2016
Available online 9 June 2016

JEL classification:
C14
C15

Keywords:
Distribution function
Leverage effect
Gaussian process

ABSTRACT

The so-called leverage hypothesis is that negative shocks to prices/returns affect volatility more than equal positive shocks. Whether this is attributable to changing financial leverage is still subject to dispute but the terminology is in wide use. There are many tests of the leverage hypothesis using discrete time data. These typically involve fitting of a general parametric or semiparametric model to conditional volatility and then testing the implied restrictions on parameters or curves. We propose an alternative way of testing this hypothesis using realized volatility as an alternative direct nonparametric measure. Our null hypothesis is of conditional distributional dominance and so is much stronger than the usual hypotheses considered previously. We implement our test on individual stocks and a stock index using intraday data over a long span. We find only very weak evidence against our hypothesis.

1. Introduction

The so-called leverage hypothesis, Black (1976) and Christie (1982), is essentially that negative shocks to stock prices affect their volatility more than equal magnitude positive shocks. Whether this is attributable to changing financial leverage or is a result of the volatility feedback effect (French et al., 1987; Campbell and Hentschel, 1992), is still subject to dispute (Engle and Ng, 1993; Figlewski and Wang, 2000; Bekaert and Wu, 2000; Bollerslev et al., 2006 and Dufour et al., 2012), but the terminology is in wide use. There are many statistical tests of the leverage hypothesis using discrete time data. These typically involve fitting of a general parametric or semiparametric model to conditional volatility and then testing the implied restrictions on parameters or curves, see for example Nelson (1991), Engle and Ng (1993), Linton and Mammen (2005), and Rodriguez and Ruiz (2012). Most authors have found the parameters governing asymmetric volatility response in daily individual stock returns and in indexes to be statistically significant.

A theoretical justification of the leverage effect is given in Christie (1982) inside a continuous time model, and recently there has been an important literature on measuring leverage effects in high frequency data. Aït-Sahalia et al. (2013) investigate the leverage effect "puzzle" within the continuous time framework. The puzzle is that natural estimators of the leverage effect based on high frequency data are usually very small and insignificant. They take apart the sources of this finding and interpret it as bias due to microstructure noise issues, and they propose a solution to this based on a bias correction. Empirically their method seems to uncover a stronger leverage effect. Wang and Mykland (2014) propose a nonparametric estimator of a class of leverage parameters inside a very general class of continuous time stochastic processes. They propose an estimator that is quite simple and easily studied and provide its limiting properties. They extend the theory to allow for measurement error and more sophisticated estimators of volatility and leverage. Their modified procedure is consistent and asymptotically mixed normal in this case too, although the rate of convergence is slower. They provide the means to conduct inference about the leverage parameter, although their application is more toward prediction of volatility. They demonstrate the value added that their leverage effect has in this purpose.

Bandi and Renò (2012) propose a nonparametric method for estimating the leverage effect in a continuous time stochastic volatility with jumps model. They use a flexible function of the

☆ We thank Valentina Corradi, Jean-Marie Dufour, and Haim Levy for helpful comments. R software for carrying out the conditional dominance test is available from the web site www.oliverlinton.me.uk.
* Corresponding author.
E-mail addresses: obl20@cam.ac.uk (O. Linton), whang@snu.ac.kr (Y.-J. Whang), yyu_min@nccu.edu.tw (Y.-M. Yen).

http://dx.doi.org/10.1016/j.jeconom.2016.02.018
0304-4076/

state of the firm, which is associated with intraday returns and spot variances to measure the leverage effect. They prove consistency of the functional estimates as the number of observations diverges to infinity and the interval for estimating the intradaily spot variances approaches zero. Asymptotic properties of their estimators also depend on behaviors of the jump components in the price process. Using the proposed estimators with intradaily asset returns and estimated spot variances as inputs, they find that the leverage effect is time varying and its magnitude increases with the variance level.

Our focus is on the low frequency (daily) volatility and return relationship. We propose a way of testing the leverage hypothesis nonparametrically without requiring a specific parametric or semiparametric model. Consequently our test statistics do not need an estimated quantity for measuring the leverage effect as an input. This is a major difference between our approach and the aforementioned methods proposed by Aït-Sahalia et al. (2013), Bandi and Renò (2012) and Wang and Mykland (2014), which all rely on using the estimated leverage effect parameter for statistical inference. Our inference is robust to the model choices that many previous studies have adopted. In fact, we test a "strong leverage" hypothesis. Our null hypothesis is that the conditional distribution of volatility given negative returns and past volatility stochastically dominates in the first order sense the distribution of volatility given positive returns and past volatility. This hypothesis is stronger in some sense than those considered previously since we refer to the distribution rather than just the mean of the outcome.[1] If our null hypothesis is satisfied then any investor who values volatility negatively would prefer the distribution of volatility that arises after positive shocks to returns to the distribution that arises after negative shocks (Levy, 2006). A further advantage of formulating our hypothesis in terms of distributions is that the tests are less sensitive to the existence of moments. A lot of informal evidence around the leverage effect is reported based on cross correlations between squared returns and lags and leads of returns, see for example Bouchaud et al. (2001). As Mikosch and Starica (2000) have shown, the asymptotic behavior of sample correlograms can be badly affected by heavy tails, which themselves have been widely documented in daily stock returns. Therefore, confidence intervals and hypothesis tests under these circumstances need to be evaluated with care. Our distribution theory builds on the work of Linton et al. (2005) who considered tests of unconditional stochastic dominance for time series data. Linton et al. (2010) consider conditional dominance tests but inside specific semiparametric models. We allow for a general stationary and mixing process for both returns and volatility and impose some smoothness conditions needed for our asymptotic approximations, but otherwise our test is model-free. We obtain the limiting distribution of our test statistic: it is a functional of a Gaussian process. Since the limit distribution depends in a complicated way on nuisance parameters, we propose an inference method based on subsampling (Politis and Romano, 1994). Our test is consistent against a general class of alternatives.

A key part of our methodology is volatility, and we work with ex post volatility that is estimated from high frequency data. Our asymptotic framework requires $n \to \infty$ and $T \to \infty$, where n denotes the number of high-frequency intra-period returns used to compute the realized variance in every period, and T denotes the number of low-frequency time-periods used in the estimation of the test statistic. We derive the limiting distribution of the estimated coefficients under this double asymptotic framework.[2] We find that under fairly strong conditions on n and T, the estimates are $\sqrt{T}$-consistent and have the standard distribution as when there is no measurement-error. However, if the above condition is not satisfied, there is an asymptotic bias that would invalidate this approximation. In that case, we find that under weaker conditions on n and T, a bias-corrected estimator has the standard limiting distribution. This improvement is particularly relevant in the empirical case we examine where n is quite modest. The above is an important methodological contribution to the extant literature on high-frequency volatility estimation. Most work has currently been about just estimating that quantity itself and using it to compare discrete time models in settings where the noise is small. Our approach is concerned with small sample issues when using estimated realized volatility as regressors in the estimation of parameters associated with the unobserved quadratic variation. This involves a useful extension of the existing asymptotic results for realized volatility[3] concerned with the uniformity of the estimation error. We establish the properties of the parameter estimates and propose a bias correction in the case where the estimation error is large. Our methodology sits between discrete time econometrics and continuous time econometrics, since we use concepts from both literatures. If the volatility measure we use can be interpreted as an unbiased estimator of ex ante volatility, then our hypothesis can be interpreted inside the typical discrete time framework.

We apply our testing methodology to stock returns. We focus on whether there is a leverage effect between daily volatility and daily lagged returns on the S&P500 (cash) index and on individual stocks. The stocks we consider are five constituents of the Dow Jones Industrial Average. The sample period covers 1993 to the end of 2009, which includes several very volatile episodes as well as some more tranquil ones. In our main empirical analysis, we measure daily volatility using realized volatility (computed from one minute and five minute intraday transactions data) and a realized intraday range estimator, which only requires daily high and low prices. These data are widely available both for indexes and individual stocks. Dufour et al. (2012) in their study of S&P500 futures data used also the VIX index of implied volatility but this type of traded volatility instruments are not available for individual stocks for the long time span we consider. We find little evidence against the strong leverage effect in these data. We also carry out several robustness checks, including using different volatility estimators, different sample periods, different conditioning values, and both with and without an explicit bias correction method. Our main conclusions survive in all cases. In addition, we also conduct intensive simulations to investigate how our testing methodology performs. We find our proposed test statistics work well on detecting the conditional leverage effect under various situations (see Appendix E). Finally, we compare our results with those obtained from three alternative approaches, which include: (1) Estimating HAR-RV type models with the leverage effect, and two newly developed methods for estimating the leverage effect parameter proposed by (2) Wang and Mykland (2014) and (3) Aït-Sahalia et al. (2013). Due to the generality of our approach we can cannnot explicitly quantify the magnitude of the leverage effect,

[1] Although Wang and Mykland (2014) also allow for the leverage effect to be defined through any (given) function F of volatility.

[2] Corradi and Distaso (2006) use realized variance estimators to test for the correct specification of the functional form of the volatility process within the class of eigenfunction stochastic volatility models. The procedure is based on the comparison of the moments of realized volatility measures with the corresponding ones of integrated volatility implied by the model under the null hypothesis. They allow for measurement error in the realized variance and consider an asymptotic framework similar to ours.

[3] See Barndorff-Nielsen and Shephard (2002).

whereas these alternative parametric approaches can. We carefully conduct these alternative approaches with different settings and show the results in Appendix D. Overall the results indicate that for the S&P500 (cash) index and five DJIA constituents, negative shocks to returns affect volatility significantly more than equal positive shocks. These results lend support to our conclusions in the main text.

2. Hypotheses of interest

We suppose that we observe a stationary and weakly dependent process $\{y_t, x_t, r_t\}_{t=1}^T$, where $x_t \in \mathbb{R}^{d_x}$ for some d_x, $y_t \in \mathbb{R}$ and $r_t \in \mathbb{R}$. Let:

$$F^+(y|x) = \Pr(y_t \le y \mid r_{t-1} \ge 0, x_t = x);$$
$$F^-(y|x) = \Pr(y_t \le y \mid r_{t-1} < 0, x_t = x).$$

We consider the hypotheses:

$$H_0 : F^+(y|x) \ge F^-(y|x) \quad \text{a.s. for all } (y,x) \in \mathcal{Y} \times \mathcal{X}$$
$$H_1 : F^+(y|x) < F^-(y|x) \quad \text{for some } (y,x) \in \mathcal{Y} \times \mathcal{X},$$

where $\mathcal{Y} \subset \mathbb{R}$ denotes the support of y_t and $\mathcal{X} \subset \mathbb{R}^{d_x}$ denotes the support of x_t.

A leading example is where $y_t = \sigma_t^2$ and $x_t = \sigma_{t-1}^2$, in which case the hypothesis is that bad news on returns ($r_{t-1} < 0$) lead to a bigger effect on the conditional distribution of future volatility (σ_t^2) than good news ($r_{t-1} \ge 0$) whatever the current level of volatility (σ_{t-1}^2). In this case, we can take $\mathcal{Y} = \mathcal{X} \subset \mathbb{R}_+$. Suppose that σ_t^2 was generated from a Glosten et al. (1993) process, henceforth GJR, i.e.,

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \gamma_+ r_{t-1}^2 1(r_{t-1} \ge 0) + \gamma_- r_{t-1}^2 1(r_{t-1} < 0). \quad (1)$$

The case where $\gamma_- > \gamma_+$ corresponds to the presence of a leverage effect. In this case the distribution $F^-(y|x)$ first order dominates $F^+(y|x)$ for all x. The same dominance relation can be found inside the Nelson (1991) model taking $y_t = \ln\sigma_t^2$ and $x_t = \ln\sigma_{t-1}^2$.

We allow for a more general formulation than $y_t = \sigma_t^2$ and $x_t = \sigma_{t-1}^2$ for practical reasons. In view of the possible strong dependence in volatility we might consider conditioning on a vector of past volatilities $x_t = (\sigma_{t-1}^2, \ldots, \sigma_{t-p}^2)^\top \in \mathbb{R}^p$ instead of just on σ_{t-1}^2. In practice, however, this is likely to work poorly for large p because of the curse of dimensionality. We consider a compromise approach in which we condition on a lower dimensional transform of a vector of lagged volatilities. Specifically, let $h : \mathbb{R}^p \to \mathbb{R}^{d_x}$ for $d_x < p \le \infty$ be a measurable function and replace σ_{t-1}^2 by $x_t = h(\sigma_{t-1}^2, \ldots, \sigma_{t-p}^2)$. For example, $h(x_1, \ldots, x_p) = \sum_{j=1}^p c_j x_j$ for known $c_1, \ldots, c_p$. In this case we consider the conditional distributions $F^+(y|x) = \Pr(\sigma_t^2 \le y \mid r_{t-1} \ge 0, h(\sigma_{t-1}^2, \ldots, \sigma_{t-p}^2) = x)$ and $F^-(y|x) = \Pr(\sigma_t^2 \le y \mid r_{t-1} < 0, h(\sigma_{t-1}^2, \ldots, \sigma_{t-p}^2) = x)$.

We next rewrite the null hypothesis in a way that we will use for testing. Letting

$$\pi_0^+(x) = \Pr(r_{t-1} \ge 0 | x_t = x);$$
$$\pi_0^-(x) = \Pr(r_{t-1} < 0 | x_t = x),$$

we can write the above hypotheses by the conditional moment inequalities:

$$H_0 : E\left[1(y_t \le y)\left(\frac{1(r_{t-1} < 0)}{\pi_0^-(x_t)} - \frac{1(r_{t-1} \ge 0)}{\pi_0^+(x_t)}\right)\middle| x_t = x\right] \le 0$$
$$\text{a.s. for all } (y,x) \in \mathcal{Y} \times \mathcal{X},$$
$$H_1 : E\left[1(y_t \le y)\left(\frac{1(r_{t-1} < 0)}{\pi_0^-(x_t)} - \frac{1(r_{t-1} \ge 0)}{\pi_0^+(x_t)}\right)\middle| x_t = x\right] > 0$$
$$\text{for some } (y,x) \in \mathcal{Y} \times \mathcal{X},$$

or equivalently:

$$H_0 : E\left[1(y_t \le y)\left\{\pi_0^+(x_t) - 1(r_{t-1} \ge 0)\right\}\middle| x_t = x\right] \le 0$$
$$\text{a.s. for all } (y,x) \in \mathcal{Y} \times \mathcal{X}$$
$$H_1 : E\left[1(y_t \le y)\left\{\pi_0^+(x_t) - 1(r_{t-1} \ge 0)\right\}\middle| x_t = x\right] > 0$$
$$\text{for some } (y,x) \in \mathcal{Y} \times \mathcal{X},$$

assuming $\pi_0^+(x) = 1 - \pi_0^-(x) > 0$ for all x. It is well known that the hypotheses of H_0 and H_1 can be equivalently stated using the unconditional moment inequalities:

$$H_0 : E\left[1(y_t \le y)g(x_t)\left\{\pi_0^+(x_t) - 1(r_{t-1} \ge 0)\right\}\right] \le 0$$
$$\text{for all } (y,g) \in \mathcal{Y} \times \mathcal{G}, \quad (2)$$
$$H_1 : E\left[1(y_t \le y)g(x_t)\left\{\pi_0^+(x_t) - 1(r_{t-1} \ge 0)\right\}\right] > 0$$
$$\text{for some } (y,g) \in \mathcal{Y} \times \mathcal{G}, \quad (3)$$

where g is an instrument that depends on the conditioning variable x_t and $\mathcal{G}$ is the collection of instruments, see, e.g., Andrews and Shi (2013) and the references therein. In this paper, we take

$$\mathcal{G} = \left\{g_{a,b} : g_{a,b}(x) = \Pi_{i=1}^{d_x} 1(a_i < x_i \le b_i) \text{ for some } a, b \in \mathcal{X}\right\},$$

see Andrews and Shi (2013) for more examples of instruments. We will use the relation (2) to generate a test statistic.

We emphasize the null hypothesis of a leverage effect. One might instead consider the conditional independence hypothesis, i.e.,

$$\sigma_t^2 \text{ is independent of sign}(r_{t-1}) \text{ given } \sigma_{t-1}^2. \quad (4)$$

This hypothesis would be consistent with a GARCH(1, 1) process for σ_t^2, namely, $\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \gamma r_{t-1}^2$ for positive parameters ω, γ, β. The GJR process (1) is incompatible with this hypothesis. In fact, the GJR process is incompatible with (4) whenever $\gamma_- \ne \gamma_+$. In the general nonparametric setting we are in, the alternative hypothesis to the independence hypothesis (4) contains many processes that do not represent what we think a leverage effect should be, so that rejection of that null hypothesis would not provide any information about leverage. Perhaps we could adjust the critical region to only look in the direction of the leverage alternatives, and while this is straightforward in the parametric case, it is not so in the nonparametric case. This is why we do not pursue this hypothesis here further. Instead, one might take the null hypothesis to be the absence of a leverage effect, i.e., to take H_1 as the null hypothesis. This is an approach advocated by Davidson and Duclos (2012) in the context of stochastic dominance testing. We remark that in this wider stochastic dominance literature, the dominant paradigm is the null hypothesis of dominance. In our specific case, we think that there is a lot of evidence around the leverage effect to suggest that it might make a reasonable working hypothesis. Furthermore, our null hypothesis is a strong leverage effect and imposes much stronger restrictions on the data generating process than the usual parametric hypotheses. We think this is a reasonable approach to take from a scientific point of view, i.e., impose strong restrictions on the data and subject them to rigorous testing.

3. Test statistic

We next define empirical versions of the moment inequalities and estimated daily volatility. In practice, the volatility σ_t^2 is unobserved and needs to be estimated using high frequency data. Let $\hat{y}_t$ and $\hat{x}_t$ be estimators of y_t and x_t, respectively, based on the estimated volatility $\hat{\sigma}_t^2$, see Section 5.1 for examples of $\hat{\sigma}_t^2$.

Let $\hat{\pi}^+$ be nonparametric kernel estimators of π_0^+, i.e.,

$$\hat{\pi}^+(x) = \frac{\sum_{t=2}^{T} 1(r_{t-1} \geq 0) K_h\left(x - \hat{x}_t\right)}{\sum_{t=2}^{T} K_h\left(x - \hat{x}_t\right)},$$

where $K : \mathbb{R}^{d_x} \to \mathbb{R}$ is a kernel function and $K_h(\cdot) = K(\cdot/h)/h$ and h is a bandwidth parameter satisfying the assumptions below. Now the hypothesis can be tested based on the following statistic

$$\bar{m}_T(y, g, \pi) = \frac{1}{T} \sum_{t=2}^{T} 1(y_t \leq y) g(x_t) \{\pi(x_t) - 1(r_{t-1} \geq 0)\}.$$

Let $\hat{\bar{m}}_T(y, g, \pi)$ be $\bar{m}_T(y, g, \pi)$ with y_t, x_t replaced by $\hat{y}_t, \hat{x}_t$, respectively. We consider a Kolmogorov Smirnov-type (KS) test statistic defined by

$$S_T = \sup_{(y,g) \in \mathcal{Y} \times \mathcal{G}} \left\{ \sqrt{T} \hat{\bar{m}}_T(y, g, \hat{\pi}^+) \right\}.$$

The statistic is relatively easy to compute, it requires some choices regarding kernel, bandwidth and function class $\mathcal{G}$, but these will be discussed below.

We suppose that we observe a process $\{y_t, x_t, r_t\}_{t=1}^T$. In practice, we have only an estimate $\widehat{\sigma}_t^2$ of σ_t^2 computed from high frequency data. Barndorff-Nielsen and Shephard (2002) have shown that the realized volatility consistently estimates integrated volatility σ_t^2 at rate $n_t^{-1/2}$, where n_t is the number of high frequency observations within day t. We may assume that n_t is (effectively) very large relative to T such that we may safely ignore the fact that volatility is estimated.[4] In some contexts this may not be a good assumption. Ghosh and Linton (2013) worked with monthly volatility estimates computed from daily data. They developed a bias correction method suitable for a special class of moment condition asset pricing models that takes account of estimation error in volatility. We extend their analysis to the current situation. This type of analysis is rather difficult to conduct in this context because of the lack of smoothness (indicator functions), but we obtain a formula for the bias that arrives from the high frequency estimation. We note that the volatility measures we consider are widely used in empirical studies, see for example French et al. (1987), and typically are used without any adjustment for estimation error.

4. Estimation of volatility

We describe here our general approach to estimating volatility from high frequency data and how it fits in with the low frequency testing strategy. We suppose that we observe high frequency returns $r_{t_j}, j = 1, \ldots, n_t$ for each t (where $t = 1, \ldots, T$). We suppose that they are generated by the following (sequence of) discrete-time model(s)

$$r_{t_j} = n_t^{-1} \mu_{t_j} + n_t^{-1/2} \sigma_{t_j} \eta_{t_j}, \tag{5}$$

where η_{t_j} is stationary and ergodic and furthermore: η_{t_j} and $\eta_{t_j}^2 - 1$ are martingale difference sequences with respect to $\mathcal{F}_{t_{j-1}}$, where $\mathcal{F}_{t_{j-1}}$ contains all information up to time t_{j-1}, including μ_{t_j}, σ_{t_j}. The stochastic processes $\{\mu_{t_j}, \sigma_{t_j}\}_{j=1,t=1}^{n_t,T}$ are not assumed to be independent of the process $\{\eta_{t_j}\}_{j=1,t=1}^{n_t,T}$, i.e., we allow for leverage and volatility feedback effects in intraday returns. In particular, η_{t_j} can affect $\sigma_{s_{j+k}}$ for $s = t, k \geqslant 1$ and $s > t, k \geqslant 0$. We do not assume Gaussianity for the innovation process, so that the conditional distribution of returns can be heavy tailed. This framework is broadly consistent with observed returns being the discretized approximation to the continuously compounded return $r_{t_j}^* = p_{t_j}^* - p_{t_{j-1}}^*$, where the true underlying efficient log-price p^* follows the continuous time diffusion

$$dp_t^* = \mu(p_t^*)dt + \sigma(p_t^*)dW_t, \tag{6}$$

for functions $\mu(.)$, $\sigma(.)$, and Brownian motion W. Clearly, if $\mu(.) \equiv 0$ and $\sigma(.) = \sigma$ (a constant), we have $p_t^* = \sigma W_t$, so that r_{t_j} are independent and normally distributed and $r_{t_j} = r_{t_j}^*$ so that $\eta_{t_j} \sim N(0, 1)$ and are i.i.d. More generally, one can show (under some conditions) that, with probability one, $r_{t_j} = r_{t_j}^* + o(n_t^{-\rho})$ for some $\rho > 1$ (Euler, Milstein approximations; see, e.g., Gonçalves and Meddahi (2009), Mykland and Lan (2009)).[5] The process (5) is consistent with a stochastic volatility process as in Gonçalves and Meddahi (2009, section 4): in their case, high frequency returns are mutually independent but heterogeneous, conditioning on the drift and volatility function. Our process can be also seen as an example of the discrete time approximations developed in Nelson (1990) where we replaced his generic sequence h by the specific one n_t^{-1}. For example, compare (5) with his expression 2.22 (with $c = 0$) ${}_h r_{(k-1)h:kh} = {}_h \sigma_{kh} \times {}_h Z_{kh}$, where ${}_h Z_{kh} \sim N(0, h)$ and ${}_h \sigma_{(k+1)h}^2$ has some particular dynamic specification. Under appropriate conditions, special cases of our process can be shown to converge to a stochastic volatility process (and our estimator below would converge to the quadratic variation of that limiting diffusion process).

We do not explicitly allow for microstructure measurement error in the observed prices, although we consider some procedures designed to take account of this issue in the application. We also do not explicitly allow for jumps in our stochastic process: we are treating large observations through the traditional discrete time lens, whereby a heavy tailed distribution for η_{t_j} would lead frequently to large values of r_{t_j}, which captures some aspects of the continuous time notion of a jump.

We next define the ex-post measure of volatility for day t in this framework. We shall assume that the following probability limit exists (and uniformly so in t)

$$\sigma_t^2 \equiv \operatorname*{p\,lim}_{n_t \to \infty} \overline{\sigma}_t^2, \quad \text{where } \overline{\sigma}_t^2 = \frac{1}{n_t} \sum_{j=1}^{n_t} \sigma_{t_j}^2, \tag{7}$$

where σ_t^2 can be stochastic; we further suppose that the convergence in (7) occurs very fast so that the error term from replacing σ_t^2 by $\overline{\sigma}_t^2$ is negligible. If the underlying model were the diffusion process (6), then $\sigma_t^2 = \int_0^1 \sigma^2(t-1+s)ds$ is the integrated variance and the approximation in (7) is indeed good.

We consider several different volatility estimates in our empirical work, including those designed to accommodate "high frequency jumps" and market microstructure. However, in our theoretical treatment we concentrate on the following realized variance estimator computed from the intra period returns

$$\widehat{\sigma}_t^2 = \sum_{j=1}^{n_t} r_{t_j}^2. \tag{8}$$

[4] See Corradi et al. (2012) for some results in this direction.

[5] We will use the continuous time theory to justify some of our methodology. We recognize that the approximation we make here in principle could affect our results, but remark that the more complicated higher order approximations to the discrete time process may not necessarily work better in practice.

Other more complicated estimators can be treated similarly. In the diffusion case, the theory of quadratic variation implies that the realized variance provides a consistent nonparametric measure of the integrated variance (see, e.g., Andersen et al. (2003) and Barndorff-Nielsen and Shephard (2002)): $p\lim_{n_t\to\infty}\widehat{\sigma}_t^2 = \sigma_t^2$, where the convergence is uniform in probability (over $t = 1,\ldots,T$). Also, Jacod (1994), Jacod and Protter (1998), and Barndorff-Nielsen and Shephard (2002) develop the following asymptotic distribution theory for realized variance as an estimator of the integrated variance $n_t^{1/2}(\widehat{\sigma}_t^2-\sigma_t^2) \Rightarrow \sqrt{2}(\int_0^1\sigma^2(t-1+s)dB(t-1+s))$ as $n_t \to \infty$, where B is a Brownian motion that is independent of W in Eq. (6) and the convergence is in law stable as a process. This result implies that $n_t^{1/2}(\widehat{\sigma}_t^2-\sigma_t^2) \Longrightarrow MN(0, 2\int_0^1\sigma^4(t-1+s)ds)$, where MN denotes a mixed Gaussian distribution. Barndorff-Nielsen and Shephard (2002) showed that the above result can be used in practice as the integrated quarticity $IQ \equiv \int_0^1\sigma^4(t-1+s)ds$ can be consistently estimated using $(1/3)RQ_t$, where

$$RQ_t = n_t\sum_{j=1}^{n_t} r_{t_j}^4. \tag{9}$$

It further follows that $(1.5RQ_t^{-1}n_t)^{1/2}(\widehat{\sigma}_t^2-\sigma_t^2) \Longrightarrow N(0,1)$. This is a nonparametric result as it does not require the specification of the form of the drift, $\mu(.)$, or the diffusion, $\sigma(.)$, in Eq. (6).

The integrated quarticity will play an important role in our bias correction procedure below under our model assumptions. Under the model (5), we have by the martingale CLT (Hall and Heyde, 1980, Corollary 3.1),

$$n_t^{1/2}(\widehat{\sigma}_t^2-\overline{\sigma}_t^2) \Longrightarrow MN(0, v_t), \tag{10}$$

where $v_t = p\lim_{n_t\to\infty}\frac{1}{n_t}\sum_{j=1}^{n_t}\sigma_{t_j}^4\vartheta_{t_j}$, where $\vartheta_{t_j} = E[\eta_{t_j}^4|\mathcal{F}_{t_{j-1}}]-1$. Under some additional conditions, $RQ_t \to \lim_{n_t\to\infty}\frac{1}{n_t}\sum_{j=1}^{n_t}\sigma_{t_j}^4E[\eta_{t_j}^4|\mathcal{F}_{t_{j-1}}]$. Under stronger conditions (for example, suppose that $\sigma_{t_j}^2$ are deterministic or stochastic but independent of the process $\{\eta_{t_j}\}$), then, $v_t = \sigma_t^4(\kappa-1)$, where $\kappa = E[\eta_{t_j}^4]$. If η_{t_j} were standard Gaussian, which follows from (6), then $\kappa = 3$. Furthermore, in this case we will have $((\kappa-1)RQ_t/\kappa n_t)^{-1/2}(\widehat{\sigma}_t^2-\sigma_t^2) \Longrightarrow N(0,1)$. We do not assume Gaussianity for our innovation process, although we do make this assumption to define a simple bias correction method in the next section.

4.1. Bias correction term

We define a bias correction method that is designed to capture the leading consequence of estimating σ_t^2 from the high frequency returns data. We shall define this based on the stronger assumptions we discussed at the end of the previous section, namely i.i.d. innovations with known κ. In practice we shall implement this with $\kappa = 3$, which is consistent with Gaussian innovations.

Define kernel estimators of the unknown covariate density and conditional cdfs:

$$\widehat{f}(x) = \frac{1}{T}\sum_{t=2}^{T}K_h\left(x-\hat{x}_t\right)$$

$$\widehat{F}(y|x) = \frac{\sum_{t=2}^{T}1(\widehat{y}_t\le y)K_h\left(x-\hat{x}_t\right)}{\sum_{t=2}^{T}K_h\left(x-\hat{x}_t\right)}$$

$$\widehat{F}^{+}(y|x) = \frac{\sum_{t=2}^{T}1(\widehat{y}_t\le y)1(r_{t-1}\ge 0)K_h\left(x-\hat{x}_t\right)}{\sum_{t=2}^{T}K_h\left(x-\hat{x}_t\right)},$$

and let $\widehat{f}''(x) = \partial^2\widehat{f}(x)/\partial x^2$ and likewise with $\widehat{F}''(y|x)$ and $\widehat{F}^{+''}(y|x)$. We define the following bias correction quantity

$$\begin{aligned}\widehat{\Delta}_T(y,g) = &\left(\frac{\kappa-1}{\kappa T}\sum_{t=2}^{T}\frac{RQ_t}{n_t}\right)\\ &\times\frac{1}{T}\sum_{t=1}^{T}\left(\left[\widehat{F}''(y|x_t)-\widehat{F}^{+''}(y|x_t)\right]\right.\\ &\left.+\left[\widehat{F}(y|x_t)-\widehat{F}^{+}(y|x_t)\right]\frac{\widehat{f}''(x_t)}{\widehat{f}(x_t)}\right)g\left(x_t\right)\widehat{\pi}_0^{+}(x_t).\end{aligned}$$

We implicitly assume that η_{t_j} are i.i.d. and in fact we shall further use the specific value of $\kappa = 3$, which is implied by the diffusion model.[6]

Then define the bias corrected test statistic

$$S_T^{bc} = \sup_{(y,g)\in\mathcal{Y}\times\mathcal{G}}\sqrt{T}\left\{\hat{\hat{m}}_T(y,g,\hat{\pi}^{+})-\hat{\Delta}_T(y,g)\right\}.$$

5. Asymptotic theory

5.1. The null distribution

Let $\mathcal{X}_\varepsilon$ be an ε-neighborhood of $\mathcal{X}$ for some $\varepsilon > 0$. For some constant $B < \infty$, let

$$\Pi = \left\{\pi : \|\pi(\cdot)\|_{q,\mathcal{X}_\varepsilon}\le B\right\}, \tag{11}$$

where q is an integer that satisfies $q > d_x/2$. For nonnegative integers k, λ and ω with $\omega \ge \lambda$, we define the following class of kernels:

$$\begin{aligned}\mathcal{K}_{k,\lambda,\omega} = \Big\{&K(\cdot):\mathbb{R}^k\to\mathbb{R}:\int K(x)dx = 1,\ \int x^\mu K(x)dx = 0\\ &\forall 1\le|\mu|\le\omega-\lambda-1,\ \int|x^\mu K(x)|dx<\infty\ \forall|\mu| = \omega-\lambda,\\ &D^\mu K(x)\to 0 \text{ as } \|x\|\to\infty\ \forall\mu \text{ with } |\mu|<\lambda,\\ &\sup_{x\in\mathbb{R}^k}|D^{\mu+e_j}K(x)|(\|x\|\vee 1)<\infty\ \forall\mu \text{ with } |\mu|\le\lambda\ \forall j = 1,\ldots,k,\\ &\text{and } K(\cdot) \text{ is zero outside a bounded set in } \mathbb{R}^k,\\ &\text{where } e_j \text{ denotes the } j\text{th elementary } k\text{-vector.}\Big\}.\end{aligned}$$

Assumption A. 1. (i) $\{(y_t, x_t, r_t) : t \ge 1\}$ is a sequence of strictly stationary strong mixing random variables with mixing numbers of size $-2(4d_x+5)(d_x+2)$. (ii) $\mathcal{X}$ is an open bounded subset of $\mathbb{R}^{d_x}$ with minimally smooth boundary.

2. (i) The distribution of x_t is absolutely continuous with respect to Lebesgue measure with density $f(x)$. (ii) $\inf_{x\in\mathcal{X}_\varepsilon}f(x) > 0$, $D^\mu f(x)$ exists and is continuous on $\mathbb{R}^{d_x}$ and $\sup_{x\in\mathcal{X}_\varepsilon}|D^\mu f(x)| < \infty$ $\forall\mu$ with $|\mu| \le \max\{\omega, q\}$, where ω is a positive integer that also appears in the other assumptions below. (iii) The conditional distribution $F(y|x)$ of y_t given $x_t = x$ has bounded density $f(y|x)$ for almost all $x \in \mathbb{R}^{d_x}$.

[6] An alternative strategy is to estimate κ from the data, but this requires consistent estimation of spot volatility, which would require further conditions, Kristensen (2010) for example.

3. $D^\mu\left[\pi_0^+(x)f(x)\right]$ exists and are continuous on $\mathbb{R}^{d_x}$ and $\sup_{x\in\mathcal{X}_\varepsilon}|D^\mu\left[\pi_0^+(x)f(x)\right]|<\infty\ \forall\mu$ with $|\mu|\le\max\{\omega,q\}$.
4. $K(\cdot)\in\mathcal{K}_{d_x,0,\omega}\bigcap\mathcal{K}_{d_x,q,q}$.
5. The bandwidth parameter h satisfies $T^{\min\{\frac{1}{2(d_x+q)},\frac{1}{4d_x}\}}h\to\infty$ and $T^{\frac{1}{2\omega}}h\to 0$.

Assumption A1 requires that x_t lies in an open bounded set with minimally smooth boundary. Examples of sets with minimally smooth boundaries include open bounded sets that are convex or whose boundaries are C^1-embedded in $\mathbb{R}^{d_x}$. Finite unions of aforementioned type whose closures are disjoint also have minimally smooth boundaries. The boundedness assumption is not restrictive, because, if needed, we can transform the values of x_t into a compact interval, say $[0,1]^{d_x}$, via strictly increasing transformation. Assumptions A2 and 3 impose smoothness on f and π_0^+. They are needed to ensure that the realization of $\hat{\pi}^+$ is smooth with probability tending to one and therefore the stochastic equicontinuity condition of a stochastic process $\{\bar{\nu}_T(\cdot,\cdot,\cdot):T\ge 1\}$ that appears in our proof can be verified. The use of higher-order kernel $K(\cdot)$ in Assumption A4 is due to the need to establish T^κ convergence of the kernel estimators $\hat{f}$, $\hat{\pi}^+(x)$ (see (21) and (22) in the Appendix) for some sufficiently large $\kappa\ge 1/4$. Assumption A5 imposes some conditions on the rate of convergence of bandwidth to zero. The conditions are compatible if ω is sufficiently large. These conditions can be relaxed, if needed, to allow for data-dependent methods of choosing bandwidth parameters, e.g. cross-validation or plug-in procedures.

Assumption B. 1. The process μ_{t_j} is uniformly bounded. There exists a small $\epsilon>0$ such that with probability one, for large enough T and some constant M,

$$\max_{1<t<T}\frac{1}{n_t}\sum_{j=1}^{n_t}\sigma_{t_j}^4\le MT^\epsilon.$$

2. $n=\min_{1\le t\le T}n_t=O(T^\gamma)\le\max_{1\le t\le T}n_t=O(T^\gamma)$ for some $\gamma>\max\{2/(k-1),2\epsilon\}$, where k and ϵ are as in Assumptions B1 and 4, respectively.
3. For $\theta\ge 2$, there exists a probability limit σ_t^θ for each t such that

$$\max_{1\le t\le T}\left|\frac{1}{n_t}\sum_{j=1}^{n_t}\sigma_{t_j}^\theta-\sigma_t^\theta\right|=O_p(n^{-\lambda})$$

for some $\frac{1}{2}(1-\frac{\epsilon}{\gamma})<\lambda<1$.
4. The process $\{\eta_{t_j}\}_{j=1,t=1}^{n_t,T}$ is stationary and ergodic and has finite kth moment for some large $k>6$.
5. The process $\{\eta_{t_j}\}_{j=1,t=1}^{n_t,T}$ is i.i.d. with mean zero, variance one, and finite kth moment for some large $k>6$. Let $\kappa=E\eta_{t_j}^4$.

Condition B1 controls the behavior of the volatility process over long time spans. One possibility is to require that the process $\sigma_{t_j}^2$ is uniformly bounded over all t and all j and all sample paths, but this is a little strong. Instead, we shall control the rate of growth of the maximum value this process can achieve over many periods. Let $m_t=\sum_{j=1}^{n_t}\sigma_{t_j}^4/n_t$ denote the intraperiod second moment of volatilities. Suppose, for example, that the stochastic process m_t was stationary and Gaussian, then $\max_{1\le t\le T}m_t$ would grow to infinity at a logarithmic rate. We shall allow instead this process to grow at an algebraic rate that is much faster than logarithmic. Over the sample period 1927–2010, daily excess market returns are highly leptokurtic with the degree of excess kurtosis being 20.9. The evidence of very fat tails in the distribution of returns highlights the importance of this assumption. We are not assuming Gaussianity of η_{t_j} and we are not exploiting the structure of an underlying continuous time model so we need to make some strong assumptions like B3. Note that this assumption is similar to Assumption H of Gonçalves and Meddahi (2009) with the added feature that we need to control this error uniformly over the low frequency long span aspect of our data. This assumption is consistent with many sample schemes. For example, suppose that $\sigma_{t_j}^\theta$ were i.i.d., then we can argue that $\frac{1}{\sqrt{n_t}}\sum_{j=1}^{n_t}\sigma_{t_j}^\theta-E\sigma_{t_j}^\theta$ is asymptotically normal for each t. In that case we would have to control the growth of $\max_{1\le t\le T}Z_t/\sqrt{n_t}$, which can easily be shown to be of order $(\log T)^{1/2}/\sqrt{n}$, and so given our assumptions on the relative magnitude of T and n, this term is of small order in probability.[7] Since Assumption B4 is with regard to the standardized return series, η_{t_j}, it is not so strong as requiring that returns themselves have many moments.

We now derive the asymptotic of the test statistic under the null hypothesis. Define the empirical process in $(y,g)\in\mathbb{R}\times\mathcal{G}$

$$\nu_T(y,g)=\sqrt{T}\left\{\xi_T(y,g)-E\xi_T(y,g)\right\},\tag{12}$$

where

$$\xi_T(y,g)=\frac{1}{T}\sum_{t=1}^{T}\{1(y_t\le y)-F(y|x_t)\}g(x_t)\times\left\{\pi_0^+(x_t)-1(r_{t-1}\ge 0)\right\}.\tag{13}$$

Let $\nu(y,g)$ be a mean zero Gaussian process with covariance function given by

$$C((y_1,g_1),(y_2,g_2))=\lim_{T\to\infty}\mathrm{cov}\left(\nu_T(y_1,g_1),\nu_T(y_2,g_2)\right).$$

The limiting null distribution of our test statistic is given in the following theorem.

Theorem 1. *Suppose that Assumptions A and B1–4 hold and $n^x/T\to\infty$, where $x>1/\gamma$ and $\gamma>2\epsilon+\frac{1}{2}$. Then, under the null hypothesis H_0,*

$$S_T\Rightarrow\begin{cases}\sup_{(y,g)\in\mathcal{B}}[\nu(y,g)] & \text{if } \mathcal{B}\neq\emptyset\\ -\infty & \text{if } \mathcal{B}=\emptyset,\end{cases}$$

where $\mathcal{B}=\{(y,g)\in\mathcal{Y}\times\mathcal{G}:E[1(y_t\le y)g(x_t)\{\pi_0^+(x_t)-1(r_{t-1}\ge 0)\}]=0\}$.

Theorem 1 shows that our test statistic has a non-degenerate limiting distribution on the boundary of the null hypothesis, i.e. the case where the "contact set" (i.e., the subset of $\mathcal{Y}\times\mathcal{G}$ where the null hypothesis (2) holds with equality) is non-empty. Since the distribution depends on the true data generating process, we cannot tabulate it once and for all. We suggest estimating the critical values by a subsampling procedure.

When the conditions on N and T in Theorem 1 are not satisfied, then the bias term may not vanish asymptotically and we need to consider the bias corrected test statistic S_T^{bc}. Theorem 2 requires a weaker condition on N and T than Theorem 1.

Theorem 2. *Suppose that Assumptions A and B1–5 hold and $n^x/T\to\infty$, where $x>1/\gamma$ and $\gamma>\max\{2\epsilon+\frac{1}{3},\frac{1+\epsilon}{1+2\lambda}\}$. Then, under the null hypothesis H_0,*

$$S_T^{bc}\Rightarrow\begin{cases}\sup_{(y,g)\in\mathcal{B}}[\nu(y,g)] & \text{if } \mathcal{B}\neq\emptyset\\ -\infty & \text{if } \mathcal{B}=\emptyset,\end{cases}$$

where $\mathcal{B}=\{(y,g)\in\mathcal{Y}\times\mathcal{G}:E[1(y_t\le y)g(x_t)\{\pi_0^+(x_t)-1(r_{t-1}\ge 0)\}]=0\}$.

[7] If σ_{t_j} were purely deterministic, then provided the underlying function $\sigma^2(.)$ is Hölder continuous of order α

$$\left|\frac{1}{n_t}\sum_{j=1}^{n_t}\sigma_{t_j}^2-\int_0^1\sigma^2(u)du\right|\le Cn_t^{-\alpha}.$$

5.2. Critical values and consistency

We first define the subsampling procedure. Let S_T^* denote either S_T or S_T^{bc}. With some abuse of notation, the test statistic S_T^* can be re-written as a function of the data $\{W_t : t = 1, \ldots, T\}$:

$$S_T^* = \sqrt{T}\tau_T(W_1, \ldots, W_T),$$

where $\tau_T(W_1, \ldots, W_T)$ is given by either $\sup_{(y,g)\in\mathcal{Y}\times\mathcal{G}} \hat{\hat{m}}_T(y, g, \hat{\pi}^+)$ or $\sup_{(y,g)\in\mathcal{Y}\times\mathcal{G}} \left\{\hat{\hat{m}}_T(y, g, \hat{\pi}^+) - \hat{\Delta}_T(y, g)\right\}$. Let

$$G_T(\cdot) = \Pr\left(\sqrt{T}\tau_T(W_1, \ldots, W_T) \le \cdot\right) \tag{14}$$

denote the distribution function of S_T^*. Let $\tau_{T,b,t}$ be equal to the statistic τ_b evaluated at the subsample $\{W_t, \ldots, W_{t+b-1}\}$ of size b, i.e.,

$$\tau_{T,b,t} = \tau(W_t, W_{t+1}, \ldots, W_{t+b-1}) \quad \text{for } t = 1, \ldots, T - b + 1.$$

We note that each subsample of size b (taken *without replacement* from the original data) is indeed a sample of size b from the true sampling distribution of the original data. Hence, it is clear that one can approximate the sampling distribution of S_T^* using the distribution of the values of $\tau_{T,b,t}$ computed over $T - b + 1$ different subsamples of size b. That is, we approximate the sampling distribution G_T of S_T^* by

$$\hat{G}_{T,b}(\cdot) = \frac{1}{T-b+1}\sum_{t=1}^{T-b+1} 1\left(\sqrt{b}\tau_{T,b,t} \le \cdot\right).$$

Let $g_{T,b}(1-\alpha)$ denote the $(1-\alpha)$th sample quantile of $\hat{G}_{T,b}(\cdot)$, i.e.,

$$g_{T,b}(1-\alpha) = \inf\{w : \hat{G}_{T,b}(w) \ge 1 - \alpha\}.$$

We call it the *subsample critical value* of significance level α. Thus, we reject the null hypothesis at the significance level α if $S_T^* > g_{T,b}(1-\alpha)$. The computation of this critical value is not particularly onerous, although it depends on how big b is. The subsampling method has been proposed in Politis and Romano (1994) and is thoroughly reviewed in Politis et al. (1999). It works in many cases where the standard bootstrap fails: in heavy tailed distributions, in unit root cases, in cases where the parameter is on the boundary of its space, etc.

We now show that our subsampling procedure works under a very weak condition on b. In many practical situations, the choice of b will be data-dependent, see Linton et al. (2005, Section 5.2) for some methodology for choosing b. To accommodate such possibilities, we assume that $b = \hat{b}_T$ is a data-dependent sequence satisfying

Assumption C. $\Pr[l_T \le \hat{b}_T \le u_T] \to 1$ where l_T and u_T are integers satisfying $1 \le l_T \le u_T \le T$, $l_T \to \infty$ and $u_T/T \to 0$ as $T \to \infty$.

The following theorem shows that our test based on the subsample critical value has asymptotically correct size:

Theorem 3. *Suppose Assumptions A, B, and C hold. Then, under the null hypothesis H_0,*

$$\lim_{T\to\infty} \Pr[S_T^* > g_{T,\hat{b}_T}(1-\alpha)] \le \alpha,$$

with equality holding if $\mathcal{B} \neq \emptyset$, where $\mathcal{B}$ is defined in Theorem 1.

Theorem 2 shows that our test based on the subsampling critical values has asymptotically valid size under the null hypothesis and has asymptotically exact size on the boundary of the null hypothesis. Under additional regularity conditions, we can extend this pointwise result to establish that our test has asymptotically correct size uniformly over the distributions under the null hypothesis, using the arguments of Andrews and Shi (2013) and Linton et al. (2010). For brevity, we do not discuss the details of this issue in this paper.

We next establish that the test S_T based on the subsampling critical values is consistent against the fixed alternative H_1.

Theorem 4. *Suppose that Assumptions A, B, and C hold. Then, under the alternative hypothesis H_1,*

$$\lim_{T\to\infty} \Pr[S_T^* > g_{T,\hat{b}_T}(1-\alpha)] = 1.$$

6. Empirical results

In this section we test the leverage hypothesis with data of the S&P500 (cash) index and five constituents of the Dow Jones Industrial Average (DJIA): Microsoft (MSFT), IBM (IBM), General Electronic (GE), Procter& Gamble (PG) and 3M (MMM). The sample period of the data for the test spans from Jan-04-1993 to Dec-31-2009 (4283 trading days). We focus on whether there is a strong leverage effect between daily volatility and lagged return. We first introduce estimators for estimating the daily volatility and then detail how we calculate the test statistics with the estimated daily volatility and present the empirical results. We also conduct intensive simulations to show the test statistics perform well in various situations. Detailed simulation results can be found in Appendix E.

6.1. Estimating the daily volatility

In our benchmark empirical analysis, we consider two methods to estimate the daily volatility. The first one is the realized variance estimator of (8), $\hat{\sigma}_t^2 = \sum_{j=1}^{n_t} r_{t_j}^2$, where r_{t_j} is the jth intraday log return on day t, n_t is the total number of intraday log return observations on day t. We use 5 min intraday log return data on day t to evaluate the realized variance $\hat{\sigma}_t^2$ and use a conventional notation $RV_t^{5\text{ min}}$ to denote such an evaluation. The second estimator we consider for estimating the daily volatility is the squared intraday range RG_t^2 (Garman and Klass, 1980; Parkinson, 1980):

$$RG_t^2 = \frac{IG^2}{4\log 2},$$

$$IG = \max_{t-1\le\tau<t} \log P(\tau) - \min_{t-1\le\tau<t} \log P(\tau),$$

where $P(\tau)$ is the intraday asset price at time stamp τ on day t, $t - 1 \le \tau < t$. The constant $4\log 2$ is an adjustment factor to scale IG^2 in order to obtain an unbiased estimate and together with other mild regular conditions, RG_t^2 will be a conditionally unbiased estimator for the daily volatility.

The two daily volatility estimators are evaluated by using data of different sampling frequencies. We would like to see whether the daily volatility estimated from using high frequency intraday or low frequency daily data can affect the test results. In addition, the two estimators are easy to use and they have comparable performances as other volatility estimators. When daily data are used, it is known that the squared intraday range estimator is more robust to microstructure noise than the squared daily return on capturing daily volatility dynamics. When high frequency intraday data are used, Liu et al. (2013) show that over a wide range class of assets and long sample period, there is strong evidence that more

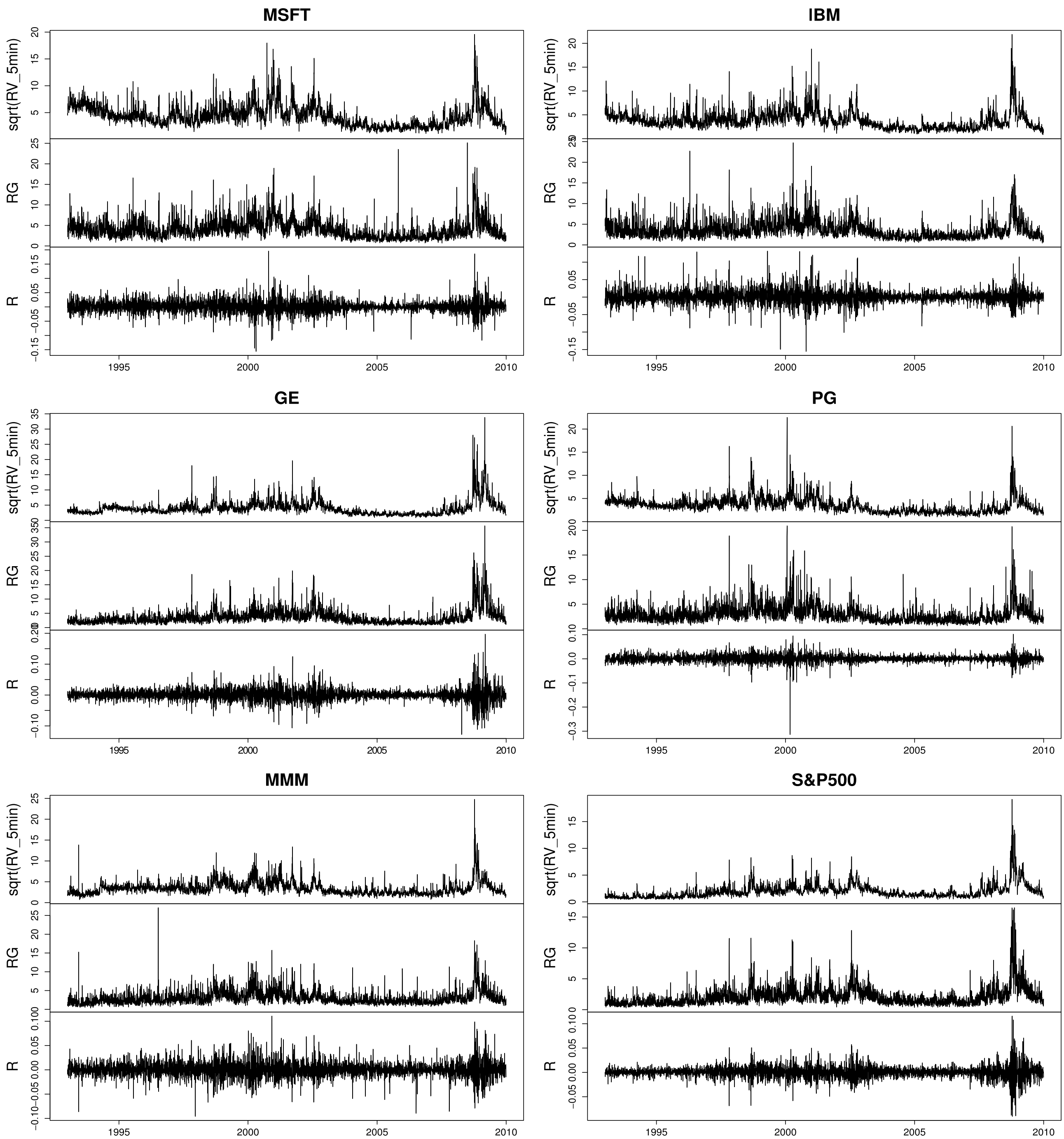

Fig. 1. Time series plots of daily $\sqrt{RV_t^{5\text{ min}}}$, RG_t and r_t^{daily} for the S&P500 index and five constituents from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). The quantities of $\sqrt{RV_t^{5\text{ min}}}$ and RG_t shown here are scaled by 252 (annualized). The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

sophisticated daily volatility estimators do not outperform the realized variance estimator (with 5-min data) in terms of accuracy of daily volatility estimations.

A detailed discussion of how we filter the high frequency data before they are used for the estimations can be found in Appendix B. Fig. 1 shows time series plots of: $\sqrt{RV_t^{5\text{ min}}}$, RG_t and the daily return r_t^{daily} for the S&P500 index and the five constituents of DJIA. Since the values of the daily estimated volatilities are very small, we scale them by 252 (annualized) before we make the plots. From the figure, during the 2008 financial crisis period all the five constituents and S&P500 index show huge fluctuations in daily returns and volatilities, while in other periods these fluctuations are relatively mild.

Table 1 shows summary statistics of the daily returns and estimated daily volatilities (scaled by 252). From the table, for the five DJIA constituents, it can be seen that RG_t has a lower mean value but a higher standard deviation than $\sqrt{RV_t^{5\text{ min}}}$, which suggests that the intraday range estimator might have a downward bias. The intraday range estimator also has a lower first order autocorrelation (denoted by ACF(1)) than the square root of the realized variance. Comparing means of $\sqrt{RV_t^{5\text{ min}}}$ among the five constituents, on average MSFT is the most volatile stock and MMM is the least volatile stock. Comparing means of RG_t, however, on

Table 1

The table shows summary statistics of daily 5-min realized volatility $\sqrt{RV_t^{5\text{ min}}}$, intraday range estimator RG_t and stock return r_t^{daily} for S&P500 index and five stocks from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). We scale $\sqrt{RV_t^{5\text{ min}}}$ and RG_t by 252 (annualized) before we calculate the statistics. The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

	Min.	Mean	Max.	Std.	Skew.	Kurt.	ACF(1)
MSFT							
$\sqrt{RV_t^{5\text{ min}}}$	0.8891	4.4160	19.5800	1.9999	1.3869	4.0803	0.8020
RG_t	0.7091	4.0450	25.1400	2.2855	1.9047	7.1083	0.5422
r_t^{daily}	−0.1560	0.0004	0.1957	0.0221	0.2258	5.8996	−0.0333
IBM							
$\sqrt{RV_t^{5\text{ min}}}$	0.8271	3.8270	21.8800	1.8541	2.0632	8.7406	0.7626
RG_t	0.4431	3.5910	24.7300	2.1427	2.1492	8.4899	0.5345
r_t^{daily}	−0.1554	0.0007	0.1316	0.0198	0.3747	6.3703	−0.0328
GE							
$\sqrt{RV_t^{5\text{ min}}}$	0.8364	3.9390	33.8200	2.3898	3.5067	21.6421	0.8158
RG_t	0.4919	3.5610	35.6600	2.6436	3.2342	17.1989	0.6692
r_t^{daily}	−0.1279	0.0004	0.1970	0.0196	0.3141	8.7333	−0.0127
PG							
$\sqrt{RV_t^{5\text{ min}}}$	0.8274	3.5010	22.5000	1.6252	2.2713	12.7597	0.7711
RG_t	0.5403	2.9810	20.9300	1.7922	2.7397	14.1404	0.5479
r_t^{daily}	−0.3138	0.0005	0.1021	0.0160	−1.8188	38.2772	−0.0546
MMM							
$\sqrt{RV_t^{5\text{ min}}}$	0.7142	3.4760	24.7700	1.6041	2.6075	15.3994	0.7498
RG_t	0.3507	3.0780	27.0600	1.8462	2.5815	13.8397	0.5137
r_t^{daily}	−0.0959	0.0004	0.1107	0.0159	0.1628	4.3600	−0.0348
S&P500							
$\sqrt{RV_t^{5\text{ min}}}$	0.4155	1.9620	19.1600	1.3363	3.2613	19.9993	0.8385
RG_t	0.2685	2.0920	16.5000	1.5470	3.1255	16.7762	0.6363
r_t^{daily}	−0.0904	0.0003	0.1158	0.0122	−0.0024	9.1608	−0.0668

average PG is the least volatile stock. As for the S&P500 index, it can be seen that on average it is less volatile than the five DJIA constituents during the sample period. However, the two estimated daily volatilities of the S&P500 index overall have higher sample skewness and kurtosis than the five DJIA constituents (except for GE), which indicates that shapes of unconditional distributions of the two estimated daily volatilities are more right-skewed and fat-tailed for the S&P500 index.

6.2. Conducting the leverage hypothesis test

With the estimated daily volatility, we then empirically construct the test statistic to test the conditional leverage hypothesis. We use the test statistic S_T described in Section 3 to test the null hypothesis of (2). However, for the lagged return, the lagged period j we consider is a more general case of $j \geq 1$ rather than just $j = 1$. Thus we will replace r_{t-1} with r_{t-j} on constructing the test statistic. Also for distinguishing different evaluations of the test statistic from using different lagged returns and for ease of notation, we use S_{T_j} and $\hat{\hat{m}}_{T_j}$ to denote the corresponding quantities of S_T and $\hat{\hat{m}}_T$ when the lagged j period return r_{t-j} is used.

Let

$$\hat{\tau}^1_{T_j}\left(y, g, \hat{\pi}^+\right) = \frac{1}{T}\sum_{t>j}^{T} 1\left\{\hat{y}_t \leq y\right\} 1\left\{r_{t-j} < 0\right\} g\left(\hat{x}_t\right) \hat{\pi}^+\left(\hat{x}_t\right),$$

$$\hat{\tau}^2_{T_j}\left(y, g, \hat{\pi}^+\right) = \frac{1}{T}\sum_{t>j}^{T_j} 1\left\{\hat{y}_t \leq y\right\} 1\left\{r_{t-j} \geq 0\right\} g\left(\hat{x}_t\right)\left(1 - \hat{\pi}^+\left(\hat{x}_t\right)\right),$$

where $1\{.\}$ is the indicator function and $\hat{\pi}^+$ is the nonparametric kernel estimator of π_0^+ described in Section 3. Then the function $\hat{\hat{m}}_{T_j}$ can be expressed as

$$\hat{\hat{m}}_{T_j}\left(y, g, \hat{\pi}^+\right) = \hat{\tau}^1_{T_j}\left(y, g, \hat{\pi}^+\right) - \hat{\tau}^2_{T_j}\left(y, g, \hat{\pi}^+\right).$$

To practically evaluate the full sample statistic S_{T_j}, we apply the following procedures. First, we set the instrument function $g\left(\hat{x}_t\right) = 1\left\{\underline{\sigma} < \hat{x}_t \leq \overline{\sigma}\right\}$, where $\underline{\sigma}$ and $\overline{\sigma}$ are lower and upper bounds for x_t. We fix the lower bound $\underline{\sigma} = \min_{t=j+1,\ldots,T} \ln\left(0.9 \times \exp(\hat{x}_t)\right)$ for reducing computational burden. We then seek a combination of $(y, \overline{\sigma})$ that maximize the function $\hat{\hat{m}}_{T_j}$, where:

$$y \in \left[\min_{t=j+1,\ldots,T} \ln\left(0.9 \times \exp(\hat{y}_t)\right), \max_{t=j+1,\ldots,T} \ln\left(1.1 \times \exp(\hat{y}_t)\right)\right]$$

$$\overline{\sigma} \in \left[\min_{t=j+1,\ldots,T} \ln\left(0.9 \times \exp(\hat{x}_t)\right), \max_{t=j+1,\ldots,T} \ln\left(1.1 \times \exp(\hat{x}_t)\right)\right],$$

and $\sqrt{T}$ times the maximum value of $\hat{\hat{m}}_{T_j}$ is our evaluation of S_{T_j}.

Since the estimated daily volatility is nonnegative, to avoid biased estimations caused by the boundary issue when using the Nadaraya–Watson estimator in constructing the test statistic, we make a log transformation for the estimated daily volatility

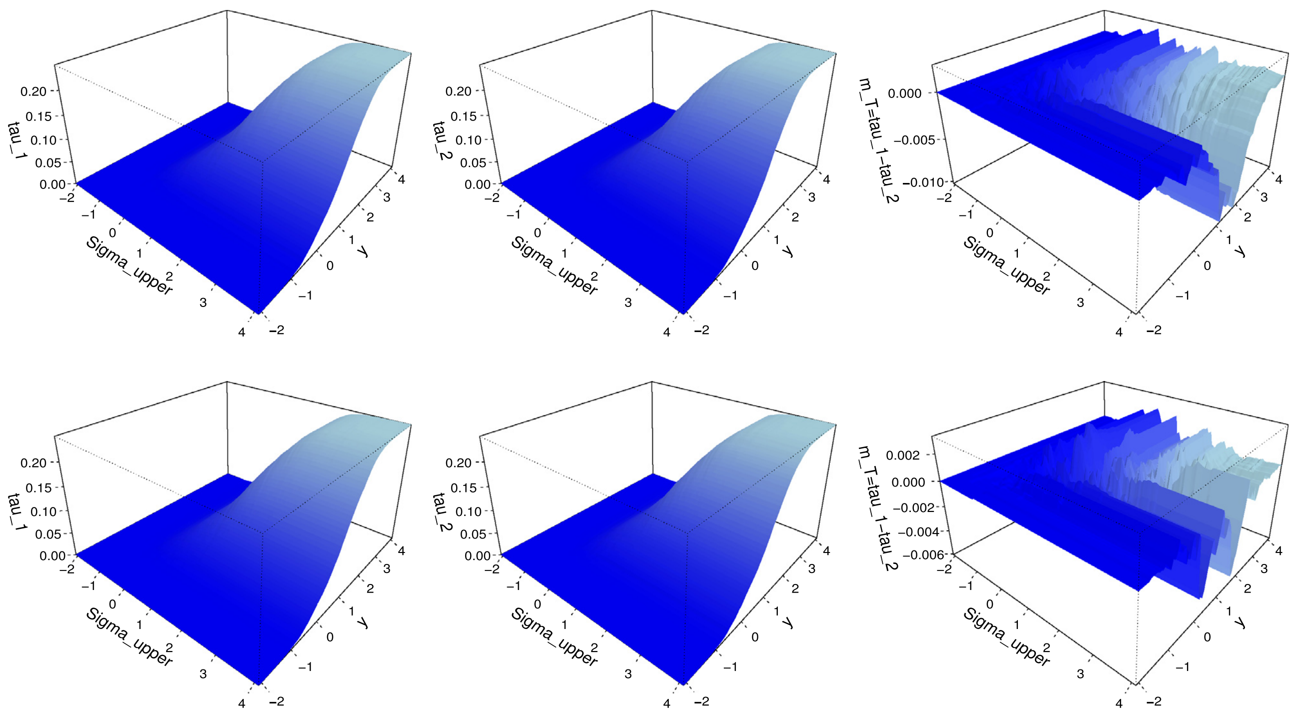

Fig. 2. Plots of surfaces $\hat{\tau}^2_{T_j}(y, g, \hat{\pi}^+)$ (left) $\hat{\tau}^1_{T_j}(y, g, \hat{\pi}^+)$ (middle) and $\hat{\bar{m}}_{T_j}(y, g, \hat{\pi}^+)$ (right) of MSFT. Note that $\hat{\bar{m}}_{T_j}(y, g, \hat{\pi}^+) = \hat{\tau}^1_{T_j}(y, g, \hat{\pi}^+) - \hat{\tau}^2_{T_j}(y, g, \hat{\pi}^+)$. Upper: $j = 1$. Bottom: $j = 5$. Here $RV_t^{5\text{ min}}$ is the estimated daily volatility.

before we empirically calculate S_{T_j}. We use the two estimated daily volatilities $RV_t^{5\text{ min}}$ and RG_t^2 and daily return r_t^{daily} in the test. Accordingly, in the function $\hat{\bar{m}}_{T_j}$, we set $\hat{y}_t = \ln RV_t^{5\text{ min}}$ (or $\ln RG_t^2$) and $\hat{x}_t = \ln RV_{t-1}^{5\text{ min}}$ (or $\ln RG_{t-1}^2$) and replace r_{t-j} with r_{t-j}^{daily}.

For constructing the approximated sample distribution for S_{T_j}, we apply the subsampling scheme described in Section 5.2 to obtain the subsampling distribution of the test statistic. To illustrate how the test statistic behaves, we use MSFT as an example. Fig. 2 shows plots of $\hat{\tau}^1_{T_j}(y, g, \hat{\pi}^+)$, $\hat{\tau}^2_{T_j}(y, g, \hat{\pi}^+)$, and $\hat{\bar{m}}_{T_j}(y, g, \hat{\pi}^+)$ against $(y, \overline{\sigma})$ when one and five period lagged returns are used ($j = 1$ and 5). Here $RV_t^{5\text{ min}}$ is used as the estimated daily volatility. From the figure, the functions $\hat{\tau}^1_{T_j}(y, g, \hat{\pi}^+)$ and $\hat{\tau}^2_{T_j}(y, g, \hat{\pi}^+)$ are smooth and monotonically increasing with y and $\overline{\sigma}$, and visually surfaces of the two functions look almost the same. It also can be seen that surfaces of $\hat{\bar{m}}_{T_j}(y, g, \hat{\pi}^+)$ are not everywhere negative. Searching for their maximum values, we find S_{T_1} and S_{T_5} are around 0.1782 and 0.2144 respectively.

With the same configurations as in Fig. 2, we show the empirical critical values over different subsample sizes in Fig. 3. Overall the subsample $(1-\alpha)$th quantile at different significant levels α decreases as the subsample size b increases. The empirical critical values also become more concentrated as the subsample size becomes large. This is expected, since as the subsample size approaches to the full sample size, the approximated sample distribution will converge to a point mass. Fig. 4 shows the corresponding empirical p-values over different subsample sizes. The empirical p-values show decreasing trends as b increases, which is consistent with the property that the approximated sample distribution becomes more concentrated as b approaches to the full sample size. Over different subsample sizes, the empirical p-value ranges from 0.8633 to 0.5357 for $j = 1$ and from 0.8728 to 0.4974 for $j = 5$, which lends support to the claim that the leverage effect exists in MSFT.

As for more empirical results, Tables 2 and 3 show the subsample critical values for the five DJIA constituents and S&P500 index when different estimated daily volatilities are used. We focus on the cases when the subsample sizes $b = 500$, 1000 and 2000 and lag lengths of returns $j = 1$ and 5. In each table we also report the full sample test statistic S_{T_j}. Table 4 shows their corresponding empirical p-values. From the Tables, it can be seen that only in a few cases when $RV_t^{5\text{ min}}$ is used, the conditional leverage hypothesis can be rejected at $\alpha = 0.1$. For example, in Table 2, IBM has $S_{T_5} = 0.4143$ and the null hypothesis of (2) can be rejected at $\alpha = 0.1$ when $b = 1000$ and 2000. Another example is the S&P500 index. In Table 2, it can be seen that S_{T_5} of the S&P500 index is around 0.2650 and the corresponding empirical p-values are around 0.0810 when $b = 2000$. For the remaining cases, the results shown here are consistent: At $\alpha = 0.1$, the null hypothesis of (2) cannot be rejected for $j = 1$ and 5, which suggests that controlling for the lagged volatility, the strong leverage effect may still exist in the five DJIA constituents and S&P500 index.

6.3. Robustness checks

We then conduct some robustness checks for our empirical results shown in the previous section. Due to similar results from using the different estimated daily volatilities and lag lengths of returns, in the following we will only focus on results from using $RV_t^{5\text{ min}}$ and $j = 1$.

In Section 6.2 we have shown that the leverage effect may exist in the five DJIA constituents and S&P500 index over the sample period from Jan-04-1993 to Dec-31-2009. To see whether our results still hold within other time periods, we divide the sample period into two subperiods: From Jan-04-1993 to Nov-30-2001 (2248 trading days) and Dec-03-2001 to Dec-31-2009 (2035 trading days), and conduct the conditional test for each of them. The division of sample period is based on the U.S. recessions identified by the NBER: The first subperiod spans the 2001 recession ending in Nov-2001, while the second subperiod spans the 2007–2009 recession ending in June-2009. Due to a smaller sample size in each subperiod, here we set $b = 500$ and 1000 for the subsampling scheme. We show the results in Table 5.

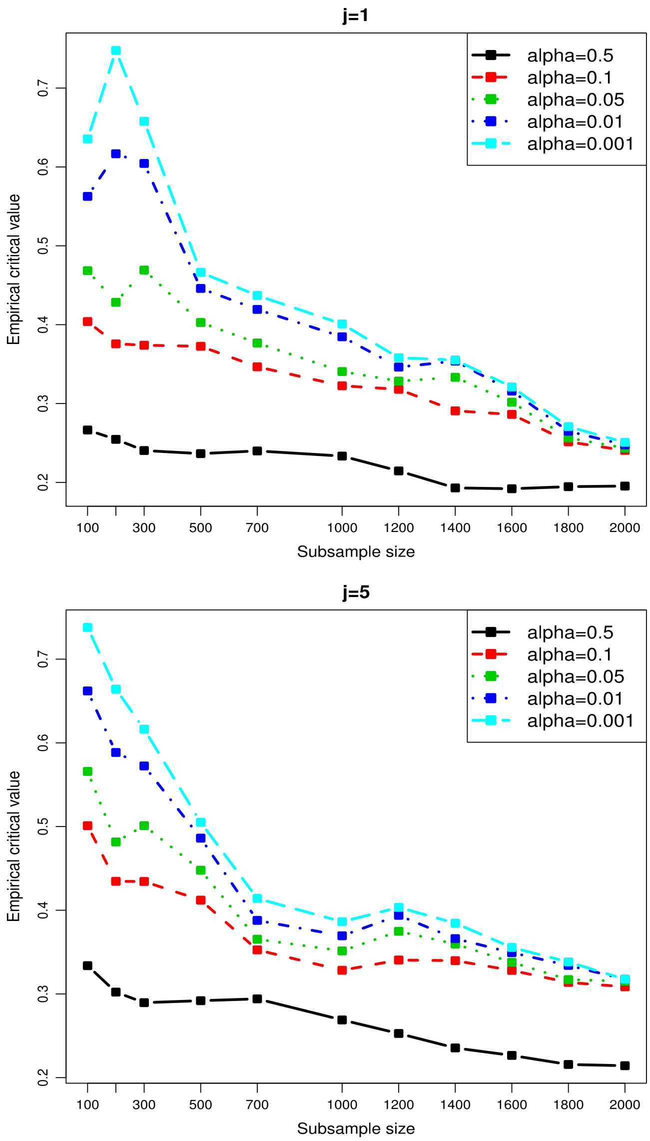

Fig. 3. The subsample critical values of MSFT. Upper: $j = 1$. Bottom: $j = 5$. Here $RV_t^{5\text{ min}}$ is the estimated daily volatility.

It can be seen that for all cases, the leverage hypothesis cannot be rejected at the significant level 0.1. The results shown here are consistent with those for the whole sample period, and for the five DJIA constituents and S&P500 index, we conclude that the leverage effect may still exist within the two subperiods.

In Section 6.2 we simply set x_t equal to log of the lagged one period daily volatility σ_{t-1}^2 instead of considering a more general functional form of the lagged daily volatility. To see whether using a more complicated functional form of $h\left(\sigma_{t-1}^2, \ldots, \sigma_{t-p}^2\right)$ affects the test results, here we specify it as

$$h_1\left(\sigma_{t-1}^2, \ldots, \sigma_{t-p}^2\right) = \frac{2}{p+1} \sum_{i=0}^{p-1} \left(\frac{p-1}{p+1}\right)^i \sigma_{t-1-i}^2, \tag{15}$$

which is an approximation for the exponential weighted moving average (EWMA) of σ_{t-1}^2:

$$\text{EWMA}\left(\sigma_{t-1}^2\right) = \frac{2}{p+1}\sigma_{t-1}^2 + \left(1 - \frac{2}{p+1}\right) EWMA\left(\sigma_{t-2}^2\right).$$

Here the term $2(p+1)^{-1}$ is the smoothing ratio. We set the parameter p equal to 5, 10 and 22, which correspond to using daily

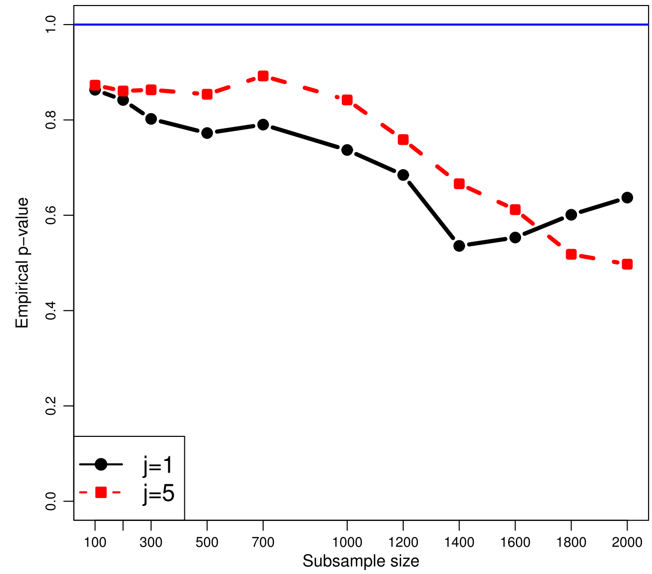

Fig. 4. The subsample empirical p-values for S_{T_j} of MSFT. Here $RV_t^{5\text{ min}}$ is the estimated daily volatility.

observations in previous one week, two weeks and one month respectively. For evaluating (15), we replace $\sigma_{t-i}^2, i = 1, \ldots, p$ with daily $RV_{t-i}^{5\text{ min}}, i = 1, \ldots, p$ in (15). Upper panel of Table 6 shows results of the conditional leverage hypothesis test conditioning on $\ln\left[h_1\left(RV_{t-1}^{5\text{ min}}, \ldots, RV_{t-p}^{5\text{ min}}\right)\right]$ and one day lagged return R_{t-1} with $b = 2000$ for the subsampling scheme. It can be seen that over the whole cases the null hypothesis cannot be rejected at a moderately significant level and there is evidence for the presence of the leverage effect in these assets.

We can also specify $h\left(\sigma_{t-1}^2, \ldots, \sigma_{t-p}^2\right)$ as a predictive equation for σ_t^2 conditioning on $\sigma_{t-1}^2, \ldots, \sigma_{t-p}^2$:

$$h_2\left(\sigma_{t-1}^2, \ldots, \sigma_{t-p}^2\right) = \mathbb{E}_{t-1}\left(\sigma_t^2 | \sigma_{t-1}^2, \ldots, \sigma_{t-p}^2\right).$$

With the realized variances at hand, a practical way to constructing the predictive equation is to estimate the HAR-RV (heterogeneous autoregressive realized variance) regression

$$\hat{\sigma}_t^2 = \alpha_D + \beta_{RD}\hat{\sigma}_{t-1}^2 + \beta_{RW}\hat{\sigma}_{t-1,week}^2 + \beta_{RM}\hat{\sigma}_{t-1,month}^2 + \varepsilon_t, \tag{16}$$

where $\hat{\sigma}_t^2$ is the realized variance estimator of (8) evaluated on day t and

$$\hat{\sigma}_{t-1,week}^2 = \frac{1}{5}\sum_{i=0}^{4}\hat{\sigma}_{t-1-i}^2,$$

$$\hat{\sigma}_{t-1,month}^2 = \frac{1}{22}\sum_{i=0}^{21}\hat{\sigma}_{t-1-i}^2,$$

are called normalized weekly and monthly realized variances. We run the HAR-RV regression (16) in the fashion of real time forecasting with $RV^{5\text{ min}}$ as the inputs. The predictive regression at each period t is estimated by using an expanding window scheme with initial window length equal to 22. At time $t-1$, the real time projected realized variance for time t, denote by $\hat{RV}_t^{5\text{ min}}$, is a function of $RV_{t-1}^{5\text{ min}}, \ldots, RV_{t-22}^{5\text{ min}}$. The bottom panel of Table 6 shows the conditional leverage hypothesis test conditioning on $\ln\left(\hat{RV}_t^{5\text{ min}}\right)$ and one day lagged return r_{t-1}^{daily} with $b = 2000$ for the subsampling scheme. The values of the test statistic range from 0.0839 to 0.1804 and the empirical p-values are all far larger than the frequently used rejection levels. The evidence shown here supports that the leverage effect may exist even when we include information for predictions of future volatility.

Table 2
The table shows the subsample critical value for the conditional leverage hypothesis test at four different levels of α for the S&P500 index and five constituents from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). We set lag lengths of daily returns $j = 1$ and 5 and subsample sizes $b = 500$, 1000 and 2000. Daily realized variances are estimated with 5-min log returns. The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

		$j=1$					$j=5$			
		α					α			
	b	0.1	0.05	0.01	0.001		0.1	0.05	0.01	0.001
MSFT	500	0.3726	0.4028	0.4460	0.4664		0.4121	0.4479	0.4862	0.5050
$S_{T_1} = 0.1782$	1000	0.3225	0.3406	0.3847	0.4010	$S_{T_5} = 0.2144$	0.3283	0.3514	0.3695	0.3863
	2000	0.2406	0.2436	0.2476	0.2509		0.3086	0.3153	0.3175	0.3176
IBM	500	0.3579	0.3935	0.4510	0.5504		0.4770	0.5440	0.6272	0.6748
$S_{T_1} = 0.1198$	1000	0.2826	0.3398	0.3838	0.4161	$S_{T_5} = 0.4143$	0.4048	0.4577	0.4926	0.5162
	2000	0.2114	0.2243	0.2387	0.2431		0.3746	0.4368	0.4654	0.4729
GE	500	0.3515	0.3830	0.4182	0.4370		0.3919	0.4244	0.4738	0.4908
$S_{T_1} = 0.1455$	1000	0.2805	0.3059	0.3916	0.4102	$S_{T_5} = 0.1499$	0.3621	0.4026	0.434	0.4441
	2000	0.2362	0.2567	0.2880	0.3178		0.2217	0.2295	0.2398	0.2496
PG	500	0.2917	0.3342	0.3788	0.4290		0.3383	0.3671	0.4867	0.5439
$S_{T_1} = 0.0877$	1000	0.2467	0.2749	0.3039	0.3161	$S_{T_5} = 0.2141$	0.2857	0.3227	0.4087	0.4407
	2000	0.1518	0.1533	0.1573	0.1621		0.2501	0.2629	0.2850	0.2936
MMM	500	0.4279	0.5167	0.6368	0.6806		0.3664	0.4061	0.461	0.4862
$S_{T_1} = 0.2370$	1000	0.4500	0.5115	0.5790	0.6060	$S_{T_5} = 0.1296$	0.2894	0.3267	0.4127	0.4217
	2000	0.2658	0.2763	0.2937	0.3251		0.1998	0.2093	0.2197	0.2274
S&P500	500	0.2472	0.2583	0.2806	0.3161		0.4081	0.5176	0.5861	0.6113
$S_{T_1} = 0.0972$	1000	0.1864	0.2468	0.3026	0.3111	$S_{T_5} = 0.2650$	0.4039	0.4290	0.4621	0.4798
	2000	0.1440	0.1556	0.1716	0.1785		0.2493	0.3084	0.3420	0.3522

Table 3
The table shows the subsample critical value for the conditional leverage hypothesis test at four different levels of α for the S&P500 index and five constituents from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). We set lag lengths of daily returns $j = 1$ and 5 and subsample sizes $b = 500$, 1000 and 2000. Daily realized variances are estimated with the intraday range estimator. The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

		$j=1$					$j=5$			
		α					α			
	b	0.1	0.05	0.01	0.001		0.1	0.05	0.01	0.001
MSFT	500	0.5382	0.6402	0.8059	0.8513		0.4649	0.5369	0.6658	0.6975
$S_{T_1} = 0.2205$	1000	0.5372	0.5664	0.6056	0.6218	$S_{T_5} = 0.1772$	0.3803	0.4632	0.5490	0.5733
	2000	0.3657	0.3737	0.3985	0.4141		0.3041	0.3114	0.3256	0.3388
IBM	500	0.3545	0.3731	0.4074	0.4382		0.3964	0.4257	0.4820	0.5438
$S_{T_1} = 0.2107$	1000	0.3330	0.3601	0.3870	0.4057	$S_{T_5} = 0.2276$	0.3398	0.3603	0.3965	0.4231
	2000	0.2965	0.3100	0.3239	0.3388		0.2347	0.2458	0.2710	0.2771
GE	500	0.5823	0.6293	0.7944	0.8420		0.3956	0.4368	0.5355	0.5775
$S_{T_1} = 0.3599$	1000	0.4854	0.5248	0.5687	0.5966	$S_{T_5} = 0.1376$	0.3092	0.3243	0.3483	0.3682
	2000	0.3610	0.3745	0.3988	0.4132		0.2364	0.2573	0.3135	0.3247
PG	500	0.3455	0.3761	0.4436	0.4899		0.4305	0.4744	0.5451	0.5823
$S_{T_1} = 0.1645$	1000	0.2088	0.2195	0.2379	0.2469	$S_{T_5} = 0.1756$	0.4286	0.4597	0.5058	0.5480
	2000	0.1984	0.2035	0.2159	0.2213		0.2030	0.2165	0.2256	0.2377
MMM	500	0.4316	0.5011	0.5837	0.6309		0.5387	0.5803	0.6262	0.6724
$S_{T_1} = 0.2300$	1000	0.4101	0.4566	0.5483	0.5787	$S_{T_5} = 0.2052$	0.4159	0.4939	0.5826	0.6133
	2000	0.2891	0.3065	0.3442	0.3631		0.2638	0.2837	0.3472	0.3649
S&P500	500	0.3795	0.4227	0.4978	0.5236		0.4599	0.5801	0.6729	0.7097
$S_{T_1} = 0.1124$	1000	0.3398	0.3682	0.4071	0.4158	$S_{T_5} = 0.2208$	0.4008	0.4728	0.6082	0.6311
	2000	0.2442	0.2594	0.2879	0.3090		0.2428	0.3144	0.4016	0.4073

We then test the conditional leverage hypothesis when different lagged j period returns are used. So far we only consider either $j = 1$ or 5. In Table 7 we show the full sample statistics and the corresponding empirical p-values of the conditional test when j varies at nine different levels with $b = 2000$ for the subsampling scheme. As the rejection level $\alpha = 0.1$, except for the cases of statistical significances mentioned in Section 6.2, we only have four additional cases of rejections here (MSFT with $j = 15$, IBM with $j = 4$ and 10, and GE with $j = 2$), and over the other cases the conditional leverage hypothesis test still cannot be rejected at $\alpha = 0.1$.

Finally, in Table 8 we show the subsample critical value of the bias-corrected test statistic S_T^{bc} described in Section 4.1 for the conditional leverage hypothesis test at four different levels of α for the S&P500 index and five DJIA constituents when $RV_t^{5\text{ min}}$ is used. Here we use $S_{T_j}^{bc}$ to denote the bias-corrected test statistic when r_{t-j}^{daily} is used for the evaluation. In Table 9 we show the corresponding empirical p-values. To illustrate how the bias-corrected test statistic behaves, in Fig. 5 we show plots of surfaces $\hat{\hat{m}}_{T_j}\left(y, g, \hat{\pi}^+\right)$, $\hat{\triangle}\left(y, g\right)$ and $\hat{\hat{m}}_{T_j}\left(y, g, \hat{\pi}^+\right) - \hat{\triangle}\left(y, g\right)$ of MSFT.

From Table 8, we find the bias correction quantity is small and has very little effects on the full sample test statistic and the subsampling distribution. Consequently the corresponding empirical p-values in Table 9 are very similar as the case of using the uncorrected test statistic S_{T_j} shown in Table 4.

In Appendix D we further consider three alternative methods on detecting the leverage effect in the assets: (1) Estimating HAR-RV type models with the leverage effect, and two recently methods for estimating the quantity for measuring the leverage effect proposed by (2) Wang and Mykland (2014) and (3) Aït-Sahalia et al.

Table 4
The table shows the empirical p-values of the conditional test statistic S_{T_j} for the S&P500 index and five constituents from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM) when different realized variances are used ($RV_t^{5\text{ min}}$ and RG_t^2). We set lag lengths of daily returns $j = 1$ and 5 and subsample sizes $b = 500$, 1000 and 2000. The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

		$RV_t^{5\text{ min}}$			RG_t^2		
		b			b		
	j	500	1000	2000	500	1000	2000
MSFT	1	0.7725	0.7369	0.6370	0.8007	0.9132	0.9991
	5	0.8541	0.8420	0.4982	0.8864	0.9190	0.9904
IBM	1	0.9577	0.6660	0.5140	0.5973	0.5600	0.4072
	5	0.1422	0.0892	0.0989	0.6506	0.5463	0.1497
GE	1	0.9334	0.9041	0.8840	0.4255	0.2710	0.1046
	5	0.9165	0.9604	0.9807	0.9902	1.0000	0.9006
PG	1	0.9968	0.9881	0.9912	0.7540	0.4406	0.4623
	5	0.6599	0.5061	0.4645	0.8166	0.5241	0.2408
MMM	1	0.3853	0.2731	0.1454	0.7611	0.6778	0.7732
	5	0.9020	0.8669	0.7557	0.7666	0.7071	0.5854
S&P500	1	0.9131	0.7351	0.4658	0.8475	0.8544	0.8109
	5	0.4556	0.2351	0.0810	0.5510	0.4166	0.1554

(2013). We find results from these parametric methods support our analysis.

6.4. Using a more sophisticated volatility estimator for the test

For robustly estimating the conditional volatility when microstructure noises are presented, in addition to using coarser sampled intraday high frequency data, recently several sophisticated methods are developed, for example, the *pre-averaging based estimation of quadratic variation* (Jacod et al., 2009; Hautsch and Podolskij, 2013). The pre-averaging method relies on using weighted averaged high frequency returns (pre-averaging returns) as inputs for the conditional volatility estimations. The ith pre-averaging return on day t is defined as

$$\bar{r}_{t_i} = \sum_{j=1}^{k_{n_t}} g\left(\frac{j}{k_{n_t}}\right) r_{t_{i+j}}.$$

Here k_{n_t} is a window length for the weighted averaged return, which is a function of total number of intraday return observations on day t, n_t. The non-zero real-valued function $g : [0, 1] \to R$ is a weight for the weighted averaged return and is required to be continuous, piecewise continuously differentiable such that its first derivative is piecewise Lipschitz and $g(0) = g(1) = 0$. Let $V(r, 2)_t^{n_t} = \sum_{i=0}^{n_t - k_{n_t}} \bar{r}_{t_i}^2$, $\psi_1 = \int_0^1 g'(u)\,du$ and $\psi_2 = \int_0^1 g^2(u)\,du$. Jacod et al. (2009) show that under some regular conditions, the following estimator:

$$C_t^{n_t} = \frac{1}{\sqrt{n_t}\theta\psi_2} V(r, 2)_t^{n_t} - \frac{\psi_1}{n_t 2\theta^2 \psi_2}\hat{\sigma}_t^2$$

can consistently estimate the daily volatility on day t when the microstructure noises are presented. Here $\hat{\sigma}_t$ is the realized variance estimator of (8) and θ is the pre-averaging parameter, which is a constant and satisfies $k_{n_t} n_t^{-\frac{1}{2}} = \theta + o\left(n_t^{-\frac{1}{4}}\right)$. For adjusting the estimation bias caused by using finite sample, they further propose the following modified estimator for $C_t^{n_t}$:

$$C_{t,a}^{n_t} = \left(1 - \frac{\psi_1^{k_{n_t}}}{n_t 2\theta^2 \psi_2^{k_{n_t}}}\right)^{-1} \times \left(\frac{\sqrt{n_t}}{\left(n_t - k_{n_t} + 2\right)\theta \psi_2^{k_{n_t}}} V(r, 2)_t^{n_t} - \frac{\psi_1^{k_{n_t}}}{n_t 2\theta^2 \psi_2^{k_{n_t}}}\hat{\sigma}_t^2\right),$$

where

$$\psi_1^{k_{n_t}} = k_{n_t} \sum_{j=1}^{k_{n_t}} \left(g\left(\frac{j+1}{k_{n_t}}\right) - g\left(\frac{j}{k_{n_t}}\right)\right)^2,$$

$$\psi_2^{k_{n_t}} = \frac{1}{k_{n_t}} \sum_{j=1}^{k_{n_t}-1} g^2\left(\frac{j}{k_{n_t}}\right),$$

are finite sample analogues of ψ_1 and ψ_2. It can be shown that as $k_{n_t} \to \infty$, $\psi_1^{k_{n_t}} \to \psi_1$ and $\psi_2^{k_{n_t}} \to \psi_2$.

In the following we calculate $C_{t,a}^{n_t}$ and use it as an input to evaluate the conditional leverage hypothesis test statistic as we do in Section 6.2 and examine whether the results presented in Section 6.2 still hold when the sophisticated daily conditional volatility estimator is used to tackle the problem of bias estimation caused by the microstructure noises. However, here we only focus on the cases of the five DJIA constituents. There are two reasons. One is that the microstructure noises mainly occur in a situation when the security is really traded, and the S&P500 cash index is not a traded security in the secondary market. The second reason is that the finest high frequency data of the S&P500 cash index available to us are 1 min data, but to empirically obtain a reliable evaluation of $C_{t,a}^{n_t}$, typically much higher frequency data are needed (Hautsch and Podolskij, 2013).[8]

The daily $C_{t,a}^{n_t}$ for conducting the conditional leverage hypothesis is obtained with 5 s intraday log returns and the preaveraging parameter $\theta = 0.6$. In Appendix C, we give a detailed discussion on how we evaluate the daily $C_{t,a}^{n_t}$ and its empirical properties. In Table 10 we show the subsample critical values and the full sample test statistic S_{T_j} for the five DJIA constituents when the daily $C_{t,a}^{n_t}$ are used. In Table 11 we show their corresponding empirical p-values. From the Tables, we can see that the results presented here are similar as those shown in Section 6.2. In Section 6.2 when $RV_t^{5\text{ min}}$ is used, the conditional leverage hypothesis can be rejected in a few cases at the significant level $\alpha = 0.1$, but here the same hypothesis can only be rejected in one case at the significant level $\alpha = 0.15$ (for MMM, the p-value equals 0.1462 when lagged period of return $j = 5$ and subsample size $b = 2000$). By using a more sophisticated daily conditional volatility estimator against the microstructure noises, the evidence shown here supports the claim that after controlling for the lagged volatility, the strong leverage effect may still exist in the five DJIA constituents.

7. Conclusion and extensions

We have found there is almost no evidence against the strong leverage effect in daily stock returns both at the individual stock level and the index level. The null hypothesis we consider is quite

8 Using the pre-averaging based estimation of quadratic variation often relies on appropriately dealing with very high frequency data to guarantee that the estimations are accurate. If the lower frequency data are used for the more sophisticated volatility estimators, their performances may not be as good as one expects. For example, in Hautsch and Podolskij (2013), the authors demonstrate that as the data frequency becomes lower, the pre-averaging estimator seems to result in downward bias, as other estimators do.

Table 5
The table shows value of test statistic S_{T_j}, empirical p-value and the subsample critical values for the conditional leverage hypothesis test at four different levels of α for SP500 index and five constituents from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). Daily realized variances are with 5-min log returns. We set lag lengths of stock returns $j = 1$ and subsample sizes $b = 500$ and 1000. The sample period is divided into two subperiods: (1) from Jan-04-1993 to Nov-30-2001 (2248 trading days). (2) from Dec-03-2001 to Dec-31-2009 (2035 trading days).

					α			
	Period	S_{T_1}	b	p-value	0.1	0.05	0.01	0.0001
MSFT	Jan. 93–Nov. 01	0.1233	500	0.9514	0.3348	0.3828	0.4243	0.4641
			1000	0.9592	0.2392	0.2483	0.2630	0.2761
	Dec. 01–Dec. 09	0.2138	500	0.7240	0.4019	0.4277	0.4557	0.4668
			1000	0.8050	0.3537	0.3706	0.3932	0.4126
IBM	Jan. 93–Nov. 01	0.1320	500	1.0000	0.3963	0.4222	0.4762	0.5545
			1000	1.0000	0.3506	0.3714	0.3984	0.4241
	Dec. 01–Dec. 09	0.0678	500	0.9974	0.2457	0.2578	0.2725	0.2850
			1000	0.9942	0.1139	0.1263	0.1421	0.1611
GE	Jan. 93–Nov. 01	0.2455	500	0.4991	0.3822	0.4040	0.4263	0.4480
			1000	0.1177	0.2531	0.2990	0.3464	0.3640
	Dec. 01–Dec. 09	0.2088	500	0.5085	0.3062	0.3260	0.3563	0.3683
			1000	0.7064	0.2845	0.2886	0.2961	0.3034
PG	Jan. 93–Nov. 01	0.1026	500	0.9874	0.2948	0.3263	0.3948	0.4375
			1000	0.9384	0.2824	0.2957	0.3100	0.3179
	Dec. 01–Dec. 09	0.1094	500	0.9512	0.2422	0.2544	0.2798	0.2869
			1000	0.9469	0.1606	0.1694	0.1826	0.2032
MMM	Jan. 93–Nov. 01	0.2859	500	0.5512	0.5228	0.596	0.6485	0.6852
			1000	0.4564	0.5226	0.5667	0.5894	0.6080
	Dec. 01–Dec. 09	0.1882	500	0.4076	0.2470	0.2830	0.3102	0.3240
			1000	0.3002	0.2122	0.2216	0.2314	0.2389
S&P500	Jan. 93-Nov. 01	0.1192	500	0.9766	0.2516	0.2642	0.3023	0.3194
			1000	0.7230	0.1696	0.1757	0.1921	0.2281
	Dec. 01–Dec. 09	0.0894	500	0.8503	0.2002	0.2212	0.2524	0.2640
			1000	0.5685	0.1262	0.1367	0.1489	0.1518

Table 6
The table shows the empirical p-values, the test statistic S_{T_j}, and the subsample critical values for the conditional leverage hypothesis test at four different levels of α for the S&P500 index and five constituents from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). In addition to the lagged one day return R_{t-1}, the test is also conditioning on more general forms of $\ln\left[h\left(\sigma^2_{t-1}, \ldots, \sigma^2_{t-p}\right)\right]$. Upper panel shows results of the test when $h(.)$ is a finite approximation for the exponential moving average in (15). Bottom panel shows results of the test when $h(.)$ is a real time forecast of the daily realized variance from the HAR-RV model in (16). Daily realized variances are with 5-min log returns. We set lag lengths of stock returns $j = 1$ and subsample size $b = 2000$. The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

	Finite Exponential Moving Average						
				α			
	p	S_{T_1}	p-value	0.1	0.05	0.01	0.001
MSFT	5	0.1653	0.5814	0.2533	0.2559	0.2691	0.2733
	10	0.1687	0.4194	0.2490	0.2509	0.2612	0.2630
	22	0.1572	0.4567	0.2249	0.2267	0.2454	0.2819
IBM	5	0.1152	0.5871	0.1771	0.1887	0.2054	0.2097
	10	0.0907	0.9680	0.1550	0.2016	0.2250	0.2252
	22	0.0824	0.9759	0.1624	0.1742	0.1770	0.1871
GE	5	0.1242	0.6121	0.1694	0.1724	0.1786	0.1857
	10	0.1565	0.6335	0.2254	0.2303	0.2379	0.2423
	22	0.1413	0.7500	0.2161	0.2211	0.2414	0.2658
PG	5	0.1044	0.7014	0.1588	0.1665	0.1761	0.1822
	10	0.0656	0.7636	0.1151	0.1364	0.1548	0.1598
	22	0.0853	0.7723	0.1339	0.1405	0.1478	0.1725
MMM	5	0.1307	0.7036	0.2267	0.2435	0.2566	0.2798
	10	0.2129	0.1261	0.2170	0.2252	0.2426	0.2708
	22	0.1280	0.5468	0.1535	0.1691	0.2147	0.2374
S&P500	5	0.1199	0.5131	0.1733	0.1893	0.2109	0.2122
	10	0.0783	0.9996	0.1545	0.1629	0.1767	0.1828
	22	0.0788	0.9974	0.1911	0.1962	0.2118	0.2175

	Real Time Forecast from the HAR-RV model						
				α			
	p	S_{T_1}	p-value	0.1	0.05	0.01	0.001
MSFT	22	0.1804	0.4825	0.3059	0.3066	0.3187	0.3224
IBM	22	0.0839	0.9991	0.1747	0.1787	0.1952	0.2066
GE	22	0.1356	0.6020	0.1832	0.1899	0.2104	0.2214
PG	22	0.0984	0.8577	0.1584	0.1659	0.1716	0.1849
MMM	22	0.1154	0.5849	0.1785	0.1843	0.2165	0.2230
S&P500	22	0.0867	0.8082	0.1431	0.1503	0.1518	0.1531

Table 7
The table shows the conditional test statistic S_{T_j} and corresponding empirical p-values for the S&P500 index and five constituents from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). Daily realized variances are with 5-min log returns. We set lag lengths of daily returns at nine different levels and subsample sizes $b = 2000$. The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

		j								
		1	2	3	4	5	8	10	15	20
MSFT	S_{T_j}	0.1782	0.2831	0.2229	0.3719	0.2144	0.3518	0.2352	0.5078	0.2894
	p-value	0.6370	0.6362	0.8713	0.1944	0.4982	0.1217	0.7294	0.0530	0.6883
IBM	S_{T_j}	0.1198	0.1124	0.2346	0.2358	0.4143	0.3343	0.3677	0.4351	0.2374
	p-value	0.5140	0.9961	0.2767	0.0871	0.0989	0.3853	0.0477	0.1187	0.3415
GE	S_{T_j}	0.1455	0.3078	0.1763	0.3105	0.1499	0.2630	0.1222	0.2558	0.2906
	p-value	0.8840	0.0946	0.4834	0.2968	0.9807	0.1182	0.8284	0.2456	0.0451
PG	S_{T_j}	0.0877	0.4160	0.2542	0.3344	0.2141	0.2195	0.2421	0.5243	0.4307
	p-value	0.9912	0.1703	0.1095	0.9295	0.4645	0.4615	0.1427	0.4750	0.3665
MMM	S_{T_j}	0.2370	0.2317	0.3109	0.2038	0.1296	0.7474	0.3204	0.1761	0.2231
	p-value	0.1454	0.2430	0.3187	0.6624	0.7557	0.3454	0.7036	0.9435	0.5306
S&P500	S_{T_j}	0.0972	0.2983	0.1952	0.1944	0.2650	0.3486	0.2493	0.2301	0.0984
	p-value	0.4658	0.6690	0.5582	0.1773	0.0810	0.2088	0.1506	0.4203	0.9829

Table 8
The table shows the subsample critical value of the bias-corrected test statistic S_T^{bc} for the conditional leverage hypothesis test at four different levels of α for S&P500 index and five stocks from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). We set lag lengths of daily returns $j = 1$ and 5 and subsample sizes $b = 500$, 1000 and 2000. Daily realized variances are estimated with 5-min log returns. The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

		$j = 1$					$j = 5$			
		α					α			
	b	0.1	0.05	0.01	0.001		0.1	0.05	0.01	0.001
MSFT	500	0.3739	0.4037	0.4465	0.4703		0.4119	0.4474	0.4877	0.5062
$S_{T_1}^{bc} = 0.1806$	1000	0.3206	0.3409	0.3832	0.4065	$S_{T_5}^{bc} = 0.2183$	0.3295	0.3506	0.3714	0.3894
	2000	0.2412	0.2449	0.2502	0.2545		0.3064	0.3134	0.3220	0.3255
IBM	500	0.3575	0.3944	0.453	0.5497		0.4754	0.5444	0.6273	0.6776
$S_{T_1}^{bc} = 0.1218$	1000	0.2838	0.3389	0.3875	0.4167	$S_{T_5}^{bc} = 0.4225$	0.4023	0.4572	0.4935	0.5141
	2000	0.2127	0.2261	0.2386	0.2464		0.3825	0.4372	0.4672	0.4813
GE	500	0.3536	0.383	0.4187	0.4472		0.3923	0.4248	0.4752	0.4953
$S_{T_1}^{bc} = 0.1474$	1000	0.2797	0.3074	0.3934	0.4112	$S_{T_5}^{bc} = 0.1528$	0.3631	0.4012	0.4350	0.4526
	2000	0.2350	0.2582	0.2885	0.3181		0.2215	0.2295	0.2409	0.2524
PG	500	0.2925	0.3338	0.3811	0.4317		0.3381	0.3672	0.4854	0.5375
$S_{T_1}^{bc} = 0.0893$	1000	0.2469	0.2743	0.305	0.3195	$S_{T_5}^{bc} = 0.2168$	0.2874	0.3223	0.4103	0.4505
	2000	0.1526	0.1549	0.1585	0.1634		0.2506	0.2623	0.2862	0.2987
MMM	500	0.4292	0.5183	0.6341	0.6782		0.3657	0.4086	0.4633	0.487
$S_{T_1}^{bc} = 0.2413$	1000	0.4512	0.5103	0.5799	0.6064	$S_{T_5}^{bc} = 0.1317$	0.2876	0.3283	0.4109	0.4298
	2000	0.2638	0.2765	0.2946	0.3300		0.2001	0.2102	0.2213	0.2297
S&P500	500	0.2472	0.2595	0.2805	0.3168		0.4072	0.5154	0.5876	0.6157
$S_{T_1}^{bc} = 0.0990$	1000	0.1868	0.2503	0.3023	0.3129	$S_{T_5}^{bc} = 0.2699$	0.4043	0.4297	0.4649	0.4848
	2000	0.1434	0.1553	0.1721	0.1781		0.2500	0.3046	0.3426	0.3576

strong, namely first order distributional dominance. Therefore, it is quite powerful that the data do not reject this hypothesis. Investors care not just about the level of volatility but also about the volatility of volatility and indeed about its entire conditional distribution, which is why this result may be of value. Our empirical evidence is robust along a number of directions. For example, our results still hold for subperiods and for different specifications on the functional form of the aggregator of lagged volatilities. In addition, several recently developed alternative methods are used on our data and their results also support our findings.

On the theoretical side, we have considered stationary processes, but this can be relaxed along the lines of Dahlhaus (1997) and Dahlhaus and Subba Rao (2006). We may also weaken the restrictions on the amount of dependence to be consistent with some evidence on the time series properties of realized volatility, see for example Andersen et al. (2011), although to allow long memory processes would be technically challenging. One could use the local polynomial regression estimator instead of the Nadaraya–Watson estimator in evaluating the test statistics, and they would share similar asymptotic properties.

Our testing methodology may be useful for other applications, i.e., for other choices of y_t, x_t and binary variable r_t. The case of first order dominance between the conditional distributions $F^-(y|x), F^+(y|x)$ may be quite strong in other datasets or applications, and one might consider the weaker concept of second order or third order dominance. The theory for these tests follows along the lines considered here and in Linton et al. (2005).

Acknowledgments

The first author thanks the ERC-NAMSEF for financial support. The second author thanks the National Research Foundation of Korea (NRF-2011-342-B00004 and NRF-2012S1A3A2033467) and Seoul National University for financial supports.

Appendix A. Proofs

For simplicity, we shall focus on the case where $\sigma_t^2 = y_t$ and $\sigma_{t-1}^2 = x_t$. Let $\hat{\delta}_t = \hat{\sigma}_t^2 - \sigma_t^2$, where $\hat{\sigma}_t^2$ is the estimator of σ_t^2 defined in (9). Lemma 1 is a modification of Ghosh and Linton (2013, Lemma 1), where we replace the exponential inequality of Bosq for mixing processes by the exponential inequality of de la

Table 9
The table shows the empirical p-values of the conditional test statistic $S^{bc}_{T_j}$ for S&P500 index and five stocks from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). We use daily realized variances estimated with 5-min log returns. We set lag lengths of daily returns $j = 1$ and 5 and subsample sizes $b = 500$, 1000 and 2000. The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

		b		
	j	500	1000	2000
MSFT	1	0.7571	0.7262	0.6191
	5	0.8438	0.8295	0.4461
IBM	1	0.9564	0.6623	0.4996
	5	0.1364	0.0810	0.0871
GE	1	0.9323	0.8928	0.8713
	5	0.9078	0.9558	0.9676
PG	1	0.9934	0.9833	0.9869
	5	0.6430	0.4781	0.4269
MMM	1	0.3679	0.2588	0.1388
	5	0.8938	0.8590	0.7233
S&P500	1	0.9046	0.7129	0.4571
	5	0.4363	0.2323	0.0749

Peña (1999, Theorem 1.2A) for martingale difference sequences, which we repeat here for convenience.

THEOREM. *Let $(X_j, \mathcal{F}_j), j = 1, 2, \ldots, n$, be a martingale difference sequence with $E(X_j|\mathcal{F}_{j-1}) = 0$ and $\sigma_j^2 = E(X_j^2|\mathcal{F}_{j-1})$ and let $V_n^2 = \sum_{j=1}^n \sigma_j^2$. Furthermore, suppose that for some c*

$$\Pr\left[|X_j| \le c|\mathcal{F}_{j-1}\right] = 1.$$

Then

$$\Pr\left[\sum_{j=1}^n X_j \ge x, V_n^2 \le y\right] \le \exp\left(-\frac{x^2}{2(y+cx)}\right). \tag{17}$$

Lemma 1. *Suppose that Assumptions B1–4 hold. Then, we have*

(a) $\max_{1\le t\le T}\left|(\hat{\sigma}_t^2 - \sigma_t^2) - \frac{1}{n_t}\sum_{j=1}^{n_t}\sigma_{t_j}^2\left(\eta_{t_j}^2 - 1\right)\right| = O_p(T^{-\lambda\gamma})$,

(b) $T^{\alpha}\max_{1\le t\le T}\left|\hat{\sigma}_t^2 - \sigma_t^2\right| = o_p(1)$ *for $\alpha < \gamma/2 - \varepsilon$.*

Table 11
The table shows the empirical p-values of the conditional test statistic S_{T_j} for S&P500 index and five stocks from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). Daily $C^{n_t}_{t,a}$ is estimated with 5-sec log returns and preaveraging parameter $\theta = 0.6$. We set lag lengths of daily returns $j = 1$ and 5 and subsample sizes $b = 500$, 1000 and 2000. The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

		b		
	j	500	1000	2000
MSFT	1	0.8306	0.9099	0.3590
	5	1.0000	0.9960	0.9978
IBM	1	0.9773	0.9875	0.8884
	5	0.7762	0.6833	0.4829
GE	1	0.6385	0.6163	0.4842
	5	0.9849	0.9967	0.9032
PG	1	0.8655	0.8739	0.8599
	5	0.5624	0.4598	0.2097
MMM	1	0.7878	0.6075	0.4882
	5	0.7413	0.5219	0.1462

Proof of Lemma 1. From (5) we obtain

$$\sum_{j=1}^{n_t} r_{t_j}^2 = \frac{1}{n_t^2}\sum_{j=1}^{n_t}\mu_{t_j}^2 + \frac{1}{n_t}\sum_{j=1}^{n_t}\sigma_{t_j}^2\eta_{t_j}^2 + \frac{2}{n_t^{3/2}}\sum_{j=1}^{n_t}\mu_{t_j}\sigma_{t_j}\eta_{t_j}.$$

Therefore,

$$\begin{aligned}\hat{\sigma}_t^2 - \sigma_t^2 &= \underbrace{\left(\frac{1}{n_t}\sum_{j=1}^{n_t}\sigma_{t_j}^2 - \sigma_t^2\right)}_{O_p(n^{-\lambda})} + \underbrace{\frac{1}{n_t}\sum_{j=1}^{n_t}\sigma_{t_j}^2(\eta_{t_j}^2 - 1)}_{O_p\left(\sqrt{\frac{T^{\epsilon}}{n}}\right)} \\ &\quad + \underbrace{\frac{1}{n_t^2}\sum_{j=1}^{n_t}\mu_{t_j}^2}_{O_p(n^{-1})} + \underbrace{\frac{2}{n_t^{3/2}}\sum_{j=1}^{n_t}\mu_{t_j}\sigma_{t_j}\eta_{t_j}}_{O_p(n^{-1})} \\ &= J_{1t} + J_{2t} + J_{3t} + J_{4t}.\end{aligned} \tag{18}$$

We have $\max_{1\le t\le T}|J_{1t}| = O_p(n^{-\lambda})$ by Assumption B3. We have $\max_{1\le t\le T}|J_{3t}| = O_p(n^{-1})$ under our conditions. Furthermore, $\max_{1\le t\le T}|J_{4t}| = O_p(n^{-1})$ by a similar argument to the sequel

Table 10
The table shows the subsample critical value for the conditional leverage hypothesis test at four different levels of α for S&P500 index and five stocks from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). We set lag lengths of daily returns $j = 1$ and 5 and subsample sizes $b = 500$, 1000 and 2000. Daily $C^{n_t}_{t,a}$ is estimated with 5-sec log returns and preaveraging parameter $\theta = 0.6$. The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

		$j = 1$					$j = 5$			
		α					α			
	b	0.1	0.05	0.01	0.001		0.1	0.05	0.01	0.001
MSFT	500	0.2555	0.2845	0.3245	0.3511		0.3549	0.3751	0.4024	0.4294
$S_{T_1} = 0.1343$	1000	0.2319	0.2434	0.2759	0.2922	$S_{T_5} = 0.1342$	0.2843	0.3250	0.3766	0.3956
	2000	0.1569	0.1768	0.1914	0.1975		0.2329	0.2434	0.2592	0.2697
IBM	500	0.3229	0.3531	0.5082	0.5464		0.3859	0.4189	0.5106	0.5444
$S_{T_1} = 0.1124$	1000	0.3372	0.3687	0.396	0.4203	$S_{T_5} = 0.2063$	0.3810	0.4000	0.4238	0.4373
	2000	0.2120	0.2215	0.2455	0.2531		0.2636	0.2725	0.2823	0.2895
GE	500	0.3904	0.4296	0.5016	0.5518		0.3678	0.3977	0.4423	0.4596
$S_{T_1} = 0.2389$	1000	0.3762	0.3903	0.4147	0.4278	$S_{T_5} = 0.1354$	0.3440	0.3785	0.3977	0.4153
	2000	0.2992	0.3007	0.3218	0.3319		0.2051	0.2233	0.2417	0.2525
PG	500	0.3165	0.3396	0.3678	0.4078		0.3399	0.3617	0.3991	0.4263
$S_{T_1} = 0.1485$	1000	0.2424	0.2524	0.2717	0.2787	$S_{T_5} = 0.2334$	0.2909	0.3018	0.3230	0.3366
	2000	0.2354	0.2434	0.2580	0.2703		0.3071	0.3199	0.3352	0.3451
MMM	500	0.3940	0.4233	0.4942	0.5338		0.3121	0.3386	0.4056	0.4343
$S_{T_1} = 0.1781$	1000	0.3176	0.3437	0.3696	0.4111	$S_{T_5} = 0.1876$	0.2471	0.2662	0.2870	0.3043
	2000	0.2060	0.2184	0.2395	0.2429		0.1931	0.2070	0.2439	0.2618

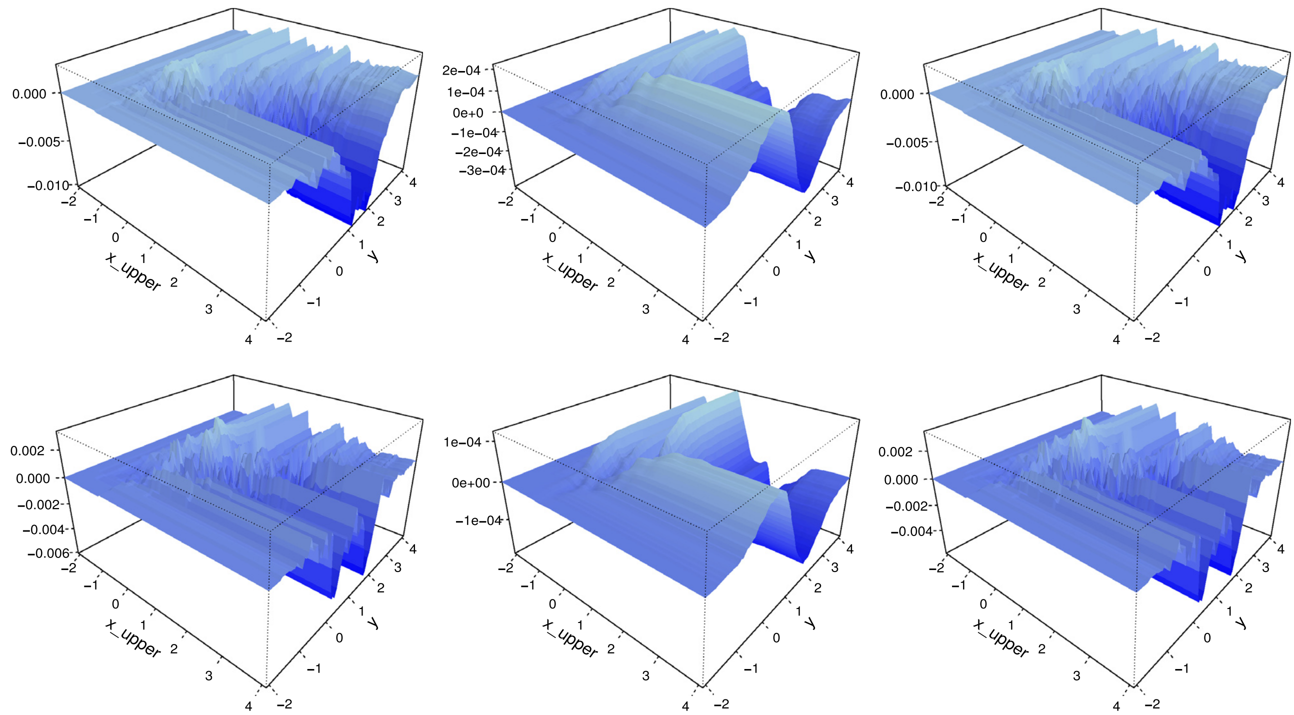

Fig. 5. Plots of surfaces $\hat{\hat{m}}_{T_j}\left(y, g, \hat{\pi}^+\right)$ (left), $\hat{\triangle}\left(y, g\right)$ (middle) and $\hat{\hat{m}}_{T_j}\left(y, g, \hat{\pi}^+\right) - \hat{\triangle}\left(y, g\right)$ (right) of MSFT. Upper: $j = 1$. Bottom: $j = 5$. Here $RV_t^{5\ \min}$ is the estimated daily volatility.

under our conditions. Thus, J_{3t} and J_{4t} are smaller than J_{1t} since $\lambda < 1$. This establishes Lemma 1(a).

Consider next the term J_{2t}. This is a sum of martingale differences, which satisfies for each t

$$\frac{1}{n_t}\sum_{j=1}^{n_t} E\left[\sigma_{t_j}^2(\eta_{t_j}^2 - 1)|\mathcal{F}_{t_{j-1}}\right] = 0 \tag{19}$$

$$\begin{aligned}
&\frac{1}{n_t^2}\sum_{j=1}^{n_t} E\left[\sigma_{t_j}^4(\eta_{t_j}^2 - 1)^2|\mathcal{F}_{t_{j-1}}\right] \\
&= \frac{1}{n_t}\left(\frac{1}{n_t}\sum_{j=1}^{n_t} \sigma_{t_j}^4 E\left[(\eta_{t_j}^2 - 1)^2|\mathcal{F}_{t_{j-1}}\right]\right) \\
&\leq \frac{C}{n_t}\frac{1}{n_t}\sum_{j=1}^{n_t}\sigma_{t_j}^4 \\
&\leqslant \frac{C}{n}MT^{\epsilon} \\
&= O\left(\frac{T^{\epsilon}}{n}\right) = o(1), \quad \text{since } \gamma > \epsilon.
\end{aligned} \tag{20}$$

Note that the second-to-last line derives from Assumption B1. Therefore, $J_{2t} = O_p\left(\sqrt{\frac{T^{\epsilon}}{n}}\right)$ uniformly over t.

Let $X_j(t) = \sigma_{t_j}^2(\eta_{t_j}^2 - 1)$ and write $X_j(t) = X_j^+(t) + X_j^-(t)$, where:

$$\begin{aligned}
X_j^+(t) &= X_j(t)1\left(|X_j(t)| \leq \sqrt{\frac{n}{\log n}}\right) \\
&\quad - E\left[X_j(t)1\left(|X_j(t)| \leq \sqrt{\frac{n}{\log n}}\right)|\mathcal{F}_{j-1}\right] \\
X_j^-(t) &= X_j(t)1\left(|X_j(t)| > \sqrt{\frac{n}{\log n}}\right) \\
&\quad - E\left[X_j(t)1\left(|X_j(t)| > \sqrt{\frac{n}{\log n}}\right)|\mathcal{F}_{j-1}\right].
\end{aligned}$$

Then we have

$$\begin{aligned}
\max_{1\leq t\leq T}\left|\frac{1}{n_t}\sum_{j=1}^{n_t}\sigma_{t_j}^2(\eta_{t_j}^2 - 1)\right| &\leq \max_{1\leq t\leq T}\left|\frac{1}{n_t}\sum_{j=1}^{n_t}X_j^+(t)\right| \\
&\quad + \max_{1\leq t\leq T}\left|\frac{1}{n_t}\sum_{j=1}^{n_t}X_j^-(t)\right| \\
&= I_1 + I_2.
\end{aligned}$$

For I_1 we can apply (17). Write $\sigma_j^{+2}(t) = E[X_j^{+2}(t)|\mathcal{F}_{j-1}]$ and $V_{n_t}^{2+} = \sum_{j=1}^{n_t}\sigma_j^{+2}(t)$. Then with probability one $\max_{1\leq t\leq T} V_{n_t}^{2+} \leq CnT^{\epsilon}$ for some constant C. By the Bonferroni inequality and (17)

$$\begin{aligned}
&\Pr\left[\max_{1\leq t\leq T}\left|\sum_{j=1}^{n_t}X_j^+(t)\right| \geq cT^{\epsilon/2}\sqrt{n\log n}\right] \\
&\leq 2T\Pr\left[\left|\sum_{j=1}^{n_t}X_j^+(t)\right| \geq cT^{\epsilon/2}\sqrt{n\log n},\ V_{n_t}^{2+} \leq CnT^{\varepsilon}\right] \\
&\leq 2T\exp\left(-\frac{c^2T^{\epsilon}n\log n}{2(CnT^{\varepsilon} + \sqrt{\frac{n}{\log n}}cT^{\varepsilon/2}\sqrt{n\log n})}\right) \\
&= 2T\exp\left(-\frac{c^2\log n}{2(C + cT^{-\epsilon/2})}\right) \\
&= 2Tn^{-\rho}
\end{aligned}$$

for some $\rho > 0$. Choosing ρ large enough ensures this term is small, which means that $I_1 = O_P\left(\sqrt{\frac{T^{\epsilon}\log n}{n}}\right)$. Regarding I_2, we have

$$\begin{aligned}
&\Pr\left[\max_{1\leq t\leq T}\left|\sum_{j=1}^{n_t}X_j^-(t)\right| \geq cT^{\epsilon/2}\sqrt{n\log n}\right] \\
&\leq \Pr\left[\max_{1\leq t\leq T}\max_{1\leq j\leq n_t}\left|\sigma_{t_j}^2(\eta_{t_j}^2 - 1)\right| > c\sqrt{\frac{n}{\log n}}\right]
\end{aligned}$$

$$\leq Tnc' \frac{(\log n)^{k/2} E\left[(\eta_{t_j}^2 - 1)^k\right]}{n^{k/2}}$$
$$= o(1)$$

by the Markov inequality applied to the stationary process $\eta_{t_j}^2 - 1$. The last equality holds provided k is taken large enough.

Thus, it follows that

$$\max_{1\leq t\leq T} |\widehat{\sigma}_t^2 - \sigma_t^2| = O_p\left(\sqrt{\frac{T^{\epsilon}\log n}{n}}\right).$$

So, provided $\alpha < \frac{\gamma}{2} - \epsilon$, the result of Lemma 1(b) follows. ■

Define

$$\bar{m}_T^{\delta}(y, g, \pi) = \frac{1}{T}\sum_{t=2}^{T} 1(y_t \leq y - \delta_t) g(x_t + \delta_{t-1}) \times \{\pi(x_t + \delta_{t-1}) - 1(r_{t-1} \geq 0)\},$$

where $\delta = (\delta_1, \ldots, \delta_T)^{\top}$ denotes a vector of constants. Notice that $\bar{m}_T^{\hat{\delta}}(y, g, \pi) = \hat{\bar{m}}_T(y, g, \pi)$ and $\bar{m}_T^{0}(y, g, \pi) = \bar{m}_T(y, g, \pi)$.

Consider the empirical process $\bar{\nu}_T^{\delta}(y, g, \pi)$ indexed by $(y, g, \pi) \in \mathbb{R} \times \mathcal{G} \times \Pi$

$$\bar{\nu}_T^{\delta}(y, g, \pi) = \sqrt{T}\left\{\bar{m}_T^{\delta}(y, g, \pi) - E\bar{m}_T^{\delta}(y, g, \pi)\right\}.$$

The following Lemma establishes stochastic equicontinuity of $\{\bar{\nu}_T(\cdot, \cdot, \cdot) : T \geq 1\}$.

Lemma 2. *Suppose that Assumption A holds. Then, for each $\varepsilon > 0$ and $\eta > 0$, there exists $\Delta > 0$ such that*

$$\overline{\lim_{T\to\infty}} \overset{*}{\Pr}\left[\sup_{\rho((y_1,g_1,\pi_1),(y_2,g_2,\pi_2))<\Delta} \left|\bar{\nu}_T^0(y_1, g_1, \pi_1) - \bar{\nu}_T^0(y_2, g_2, \pi_2)\right| > \eta\right] < \varepsilon,$$

where:

$$\rho((y_1, g_1, \pi_1), (y_2, g_2, \pi_2)) = \rho_a((y_1, g_1), (y_2, g_2)) \vee \rho_b(\pi_1, \pi_2)$$
$$\rho_a((y_1, g_1), (y_2, g_2)) = \left(\int_{\mathcal{Y}}\int_{\mathcal{Y}} \{[1(w_1 \leq y_1) - F(y_1|w_2)] g_1(w_2) - [1(w_1 \leq y_2) - F(y_2|w_2)] g_2(w_2)\}^2 dw_1 dw_2\right)^{1/2}$$
$$\rho_b(\pi_1, \pi_2) = \left(\int (\pi_1(x) - \pi_2(x))^2 f(x)dx\right)^{1/2}.$$

Proof of Lemma 2. The result of Lemma 2 follows from the stochastic equicontinuity result of Andrews (1989, Theorem 7) that is applicable to classes of functions that are products of smooth functions from an infinite dimensional class and a Type IV class of uniformly bounded functions. It suffices to verify Assumption E of the latter paper. (Assumption) E(i) holds by taking W_{aTt}, W_{bTt}, $\tau_a(\cdot)$, $\tau_b(\cdot)$, $m_a(W_{aTt}, \tau_a)$, and $m_b(W_{bTb}, \tau_b)$ to be x_t, (y_t, x_t), $\pi(\cdot)$, $1(\cdot \leq y)g(\cdot)$, $\pi(x_t)$ and $1(y_t \leq y)g(x_t)$, respectively. E(ii) holds by Assumption A1(ii) with $\mathcal{W}_a^*$ given by $\mathcal{X}$. E(iii) follows from Assumptions A2–3 and the definition of Π in (11). E(iv) is irrelevant to our case. E(v) holds since $\{1(\cdot \leq y)g(\cdot) : y \in \mathcal{Y}, g \in \mathcal{G}\}$ is a type IV class of uniformly bounded functions with index $p = 2$, constant $\psi = 1/2$, and dimension $d = d_x + 1$. Finally, E(vi) holds by Assumption A1(i). ■

Proof of Theorem 1. To prove Theorem 1, we first establish the following results:

$$T^{1/4}\sup_{x\in\mathcal{X}} \left|\hat{\pi}^+(x) - \pi_0^+(x)\right| \xrightarrow{p} 0, \tag{21}$$

$$T^{1/4}\sup_{x\in\mathbb{R}^{d_x}} \left|\hat{f}(x) - f(x)\right| \xrightarrow{p} 0, \tag{22}$$

$$\sup_{x\in\mathcal{X}_\varepsilon} \left|D^{\mu}\hat{\pi}^+(x) - D^{\mu}\pi_0^+(x)\right| \xrightarrow{p} 0 \quad \forall \mu \text{ with } 1 \leq \mu < q, \tag{23}$$

$$\sup_{(y,g)\in\mathcal{Y}\times\mathcal{G}} \left|\bar{\nu}_T^{\hat{\delta}}(y, g, \hat{\pi}^+) - \bar{\nu}_T^0(y, g, \pi_0^+)\right| \xrightarrow{p} 0, \tag{24}$$

$$\sup_{(y,g)\in\mathcal{Y}\times\mathcal{G}} \left|\sqrt{T}E\bar{m}_T^{\delta}(y, g, \pi)\big|_{\pi=\hat{\pi}^+,\delta=\hat{\delta}} - \sqrt{T}\hat{R}_T(y, g)\right| \xrightarrow{p} 0, \tag{25}$$

where:

$$\hat{R}_T(y, g) = R_T(y, g) + \Delta_T(y, g) \quad \text{and}$$
$$R_T(y, g) = \frac{1}{T}\sum_{t=2}^{T}\left\{1(r_{t-1} \geq 0) - \pi_0^+(x_t)\right\} F(y|x_t) g(x_t) + \int \left[F(y|x) - F^+(y|x)\right] g(x)\pi_0^+(x) f(x)dx,$$

$$\Delta_T(y, g) = \left(\frac{1}{2T}\sum_{t=2}^{T}\hat{\delta}_t^2\right)\int \left[F''(y|x) - F^{+\prime\prime}(y|x)\right] \times g(x)\pi_0^+(x) f(x)dx + \left(\frac{1}{2T}\sum_{t=2}^{T}\hat{\delta}_{t-1}^2\right)\int \left[F(y|x) - F^+(y|x)\right] \times g(x)\pi_0^+(x) f''(x)dx. \tag{26}$$

Eqs. (21)–(23) can be established using Theorem 2 of Andrews (1995) by verifying its Assumptions NP1–NP5 and NP9. Notice that NP1–NP3 are implied by Assumptions A1–3 with $\eta = \beta = \infty$ and $|\lambda| = \mu$ and Y_t, X_t, $f_t(x)$, and $g(x)$ given by $1(r_{t-1} \geq 0)$, x_t, $f(x)$, and $\pi_0^+(x)$, respectively. NP4 (a) and (c) holds by Assumption A4 with $\Omega = \hat{\Omega} = 1$ and NP4(b) is not relevant in our case, see Comment 5 to Theorem 1 of Andrews (1994). NP5 is implied by Assumption A5. Finally, NP9 holds by Assumption B using Lemma 1. This establishes (21)–(23).

Eq. (24) holds because, for each $\eta > 0$ and some $\Delta > 0$, we have

$$\Pr\left(\sup_{(y,g)\in\mathcal{Y}\times\mathcal{G}} \left|\bar{\nu}_T^{\hat{\delta}}(y, g, \hat{\pi}^+) - \bar{\nu}_T^0(y, g, \pi_0^+)\right| > \eta\right)$$
$$\leq \Pr\left(\sup_{\rho((y_1,g_1,\pi_1),(y_2,g_2,\pi_2))<\Delta} \left|\bar{\nu}_T^0(y_1, g_1, \pi_1) - \bar{\nu}_T^0(y_2, g_2, \pi_2)\right| > \eta\right) + \Pr\left(\rho_b(\hat{\pi}^+, \pi_0^+) \geq \Delta\right) + \Pr\left(\max_{1\leq t\leq T}\hat{\delta}_t \geq \Delta\right). \tag{27}$$

The right hand side of (27) can be made arbitrary small for T sufficiently large using Lemma 1, Lemma 2, and the facts that $\Pr\left(\hat{\pi}^+ \in \Pi\right) \to 1$ and $\rho_b(\hat{\pi}^+, \pi_0^+) \xrightarrow{p} 0$ that follow from (23) and Assumption A3.

To establish (25), write

$$\sqrt{T}E\bar{m}_T^{\delta}(y, g, \pi)\big|_{\pi=\hat{\pi}^+,\delta=\hat{\delta}} = \sqrt{T}E1(y_t \leq y - \delta_t) g(x_t + \delta_{t-1}) \times \{\pi(x_t + \delta_{t-1}) - 1(r_{t-1} \geq 0)\}\big|_{\pi=\hat{\pi}^+,\delta=\hat{\delta}}$$

$$= \frac{1}{\sqrt{T}} \sum_{t=2}^{T} \int F(y - \hat{\delta}_t | x) g(x) \left[\hat{\pi}^+(x) - \pi_0^+(x)\right] f(x - \hat{\delta}_{t-1}) dx$$

$$+ \frac{1}{\sqrt{T}} \sum_{t=2}^{T} \int \left[F(y - \hat{\delta}_t | x) - F^+(y - \hat{\delta}_t | x)\right]$$

$$\times g(x) \pi_0^+(x) f(x - \hat{\delta}_{t-1}) dx$$

$$\equiv M_{1T} + M_{2T}, \tag{28}$$

where the second equality holds by rearranging terms and applying law of iterated expectations and change of variables.

We first consider the second term M_{2T} of (28). We have

$$M_{2T} = \frac{1}{\sqrt{T}} \sum_{t=2}^{T} \int \left[F(y|x) - F^+(y|x)\right] g(x) \pi_0^+(x) f(x) dx$$

$$- \left(\frac{1}{\sqrt{T}} \sum_{t=2}^{T} \hat{\delta}_t\right) \int \left[F'(y|x) - F^{+\prime}(y|x)\right] g(x) \pi_0^+(x) f(x) dx$$

$$- \left(\frac{1}{\sqrt{T}} \sum_{t=2}^{T} \hat{\delta}_{t-1}\right) \int \left[F(y|x) - F^+(y|x)\right] g(x) \pi_0^+(x) f'(x) dx$$

$$+ \sqrt{T} \Delta_T(y, g) + \sqrt{T} \Lambda_T(y, g), \tag{29}$$

where $\Lambda_T(y, g)$ denotes the remainder term of the Taylor expansion.

Recall that in the decomposition (18), the leading term is given by J_{2t}. Therefore, we have

$$\frac{1}{\sqrt{T}} \sum_{t=2}^{T} \hat{\delta}_t = O_p\left(\frac{1}{\sqrt{T}} \sum_{t=2}^{T} \left\{\frac{1}{n_t} \sum_{j=1}^{n_t} \sigma_{t_j}^2 \left(\eta_{t_j}^2 - 1\right)\right\}\right)$$

$$= O_p\left(\sqrt{\frac{T^\epsilon}{n}}\right) = o_p(1), \tag{30}$$

where the second equality follows from (19)–(20) and the last equality holds since $\gamma > \epsilon$. Similarly, $T^{-1/2} \sum_{t=2}^{T} \hat{\delta}_{t-1} = o_p(1)$. Therefore, the second and third terms of the right hand sides of (29) are $o_p(1)$ uniformly over (y, g) since the integral terms are bounded under Assumption A. On the other hand, by the proof of Lemma 1(b), we have

$$\frac{1}{T} \sum_{t=2}^{T} \hat{\delta}_t^2 = O_p\left(\frac{T^\epsilon \log n}{n}\right) = o_p(T^{-1/2}), \tag{31}$$

provided $\gamma > 2\epsilon + 1/2$. (30) and (31) imply that $\frac{1}{T} \sum_{t=2}^{T} \hat{\delta}_t^2 = o_p(T^{-1/2})$ and hence $\sqrt{T} \Delta_T(y, g) = o_p(1)$ uniformly over (y, g). Similarly, we have

$$\Lambda_T(y, g) = O_p\left(\frac{1}{T} \sum_{t=2}^{T} \hat{\delta}_t^3\right)$$

$$= O_p\left(\left(\frac{T^\epsilon \log n}{n}\right)^{3/2}\right) = o_p(T^{-1/2}), \tag{32}$$

uniformly over (y, g), provided $\gamma > 2\epsilon + 1/3$. Combining these results, we have: uniformly in (y, g),

$$M_{2T} = \frac{1}{\sqrt{T}} \sum_{t=2}^{T} \int \left[F(y|x) - F^+(y|x)\right] g(x) \pi_0^+(x) f(x) dx$$

$$+ o_p(1). \tag{33}$$

Next, consider the first term M_{1T} of (28). By rearranging terms and a mean value expansion, we have

$$M_{1T} = \sqrt{T} \int F(y|x) g(x) \left[\hat{\pi}^+(x) - \pi_0^+(x)\right] \hat{f}(x) dx$$

$$+ \sqrt{T} \int F(y|x) g(x) \left[\hat{\pi}^+(x) - \pi_0^+(x)\right] \left[f(x) - \hat{f}(x)\right] dx$$

$$+ \frac{1}{\sqrt{T}} \sum_{t=2}^{T} \hat{\delta}_t \left(\int F'(y_t^*|x) g(x) \left[\hat{\pi}^+(x) - \pi_0^+(x)\right] f(x) dx\right)$$

$$+ \frac{1}{\sqrt{T}} \sum_{t=2}^{T} \hat{\delta}_{t-1} \left(\int F(y|x) g(x) \left[\hat{\pi}^+(x) - \pi_0^+(x)\right] f'(x_t^*) dx\right),$$

$$\equiv A_{1T} + A_{2T} + A_{3T} + A_{4T}, \quad \text{say}, \tag{34}$$

where y_t^* and x_t^* lie between $y - \hat{\delta}_t$ and y and $x - \hat{\delta}_{t-1}$ and x, respectively. The term A_{2T} is asymptotically negligible because it is bounded uniformly over $(y, g) \in \mathcal{Y} \times \mathcal{G}$ by

$$T^{1/4} \sup_{x \in \mathcal{X}} \left|\hat{\pi}^+(x) - \pi_0^+(x)\right| \times T^{1/4} \sup_{x \in \mathbb{R}^{dx}} \left|f(x) - \hat{f}(x)\right| \xrightarrow{p} 0 \tag{35}$$

using (21) and (22). The terms A_{3T} and A_{4T} are also $o_p(1)$ uniformly over (y, g) using an argument similar to (30).

Now, consider A_{1T}. Write

$$\left[\hat{\pi}^+(x) - \pi_0^+(x)\right] \hat{f}(x)$$

$$= \frac{1}{T} \sum_t K_h\left(x - \hat{x}_t\right) \left\{1(r_{t-1} \geq 0) - \pi_0^+(x_t)\right\}$$

$$+ \frac{1}{T} \sum_t K_h\left(x - \hat{x}_t\right) \left\{\pi_0^+(x_t) - \pi_0^+(x)\right\}.$$

Notice that

$$A_{1T} = \sqrt{T} \int F(y|x) g(x)$$

$$\times \left[\frac{1}{T} \sum_t K_h\left(x - \hat{x}_t\right) \left\{1(r_{t-1} \geq 0) - \pi_0^+(x_t)\right\}\right] dx$$

$$= \frac{1}{\sqrt{T}} \sum_t \left\{1(r_{t-1} \geq 0) - \pi_0^+(x_t)\right\} \left[F(y|x_t) g(x_t)\right.$$

$$\left. + \int \left\{F(y|x_t + uh) g(x_t + uh) - F(y|x_t) g(x_t)\right\} K(u) du\right]$$

$$+ o_p(1)$$

$$= \frac{1}{\sqrt{T}} \sum_t \left\{1(r_{t-1} \geq 0) - \pi_0^+(x_t)\right\} F(y|x_t) g(x_t) + o_p(1), \tag{36}$$

uniformly over $(y, g) \in \mathcal{Y} \times \mathcal{G}$, where the second equality holds by Lemma 1(b) and a change of variables and the last equality (36) holds by the following arguments: For each $(y, g) \in \mathcal{Y} \times \mathcal{G}$ and some $\delta > 0$, we have

$$E\left(\frac{1}{\sqrt{T}} \sum_t \left\{1(r_{t-1} \geq 0) - \pi_0^+(x_t)\right\}\right.$$

$$\left. \times \int \left\{F(y|x_t + uh) g(x_t + uh) - F(y|x_t) g(x_t)\right\} K(u) du\right)^2$$

$$\leq C\left(E\left|\int \left\{F(y|x_t + uh) g(x_t + uh)\right.\right.\right.$$

$$\left.\left.\left. - F(y|x_t) g(x_t)\right\} K(u) du\right|^{2+\delta}\right)^{2/(2+\delta)}$$

$$\to 0,$$

where the inequality follows by the moment inequality for sums of strong mixing random variables (see Lemma 3.1 of Dehling and Philipp (2002)) and Assumption A1(i), and convergence to zero holds by the fact that $F(y|\cdot)g(\cdot)$ is a bounded function using a well known convergence result for convolutions of functions in an L_p-space (with $p = 2+\delta$) (see Theorem 8.14(a) of Folland (1984)). Using a stochastic equicontinuity argument as in Lemma 1, we can show that the convergence to zero holds uniformly over $(y,g) \in \mathcal{Y} \times \mathcal{G}$. This establishes (36).

Furthermore, we have

$$\begin{aligned}
&\sup_{(y,g)\in\mathcal{Y}\times\mathcal{G}} \left| \sqrt{T} \int F(y|x)g(x) \left[\frac{1}{T}\sum_t K_h(x-x_t) \right.\right. \\
&\quad \left.\left. \times \left\{ \pi_0^+(x_t) - \pi_0^+(x) \right\} \right] dx \right| \\
&= \sup_{(y,g)\in\mathcal{Y}\times\mathcal{G}} \left| \frac{1}{\sqrt{T}} \sum_t \int F(y|x_t+uh)g(x_t+uh) \right. \\
&\quad \left. \times \left\{ \pi_0^+(x_t) - \pi_0^+(x_t+uh) \right\} K(u)du \right| \\
&\le O_p(T^{1/2}h^{\omega}) \xrightarrow{p} 0, \qquad (37)
\end{aligned}$$

where the equality holds by a change of variables, the inequality holds by an ω-term Taylor expansion using Assumptions A3 and 4, and the last convergence to zero holds by Assumption A5. Combining (34)–(37), we have

$$\begin{aligned}
M_{1T} &= \frac{1}{\sqrt{T}} \sum_t \left\{ 1(r_{t-1} \ge 0) - \pi_0^+(x_t) \right\} \\
&\quad \times F(y|x_t)g(x_t) + o_p(1). \qquad (38)
\end{aligned}$$

Now, the result (25) is established by (33) and (38).

We now establish Theorem 1. We have

$$\begin{aligned}
&\sqrt{T}\hat{\bar{m}}_T(y,g,\hat{\pi}^+) \\
&= \bar{\nu}_T^{\hat{\delta}}(y,g,\hat{\pi}^+) + \sqrt{T}E\bar{m}_T^{\delta}(y,g,\pi)\big|_{\pi=\hat{\pi}^+,\delta=\hat{\delta}} \\
&= \bar{\nu}_T^{0}(y,g,\pi_0^+) + \sqrt{T}E\bar{m}_T^{\delta}(y,g,\pi)\big|_{\pi=\hat{\pi}^+,\delta=\hat{\delta}} + o_p(1) \qquad (39) \\
&= \nu_T(y,g) + \sqrt{T} \int \left[F(y|x) - F^+(y|x) \right] \\
&\quad \times g(x)\,\pi_0^+(x)f(x)dx + o_p(1), \qquad (40)
\end{aligned}$$

where the second equation holds by (24) and the last equality holds by (25).

Under the null hypothesis, we have $\int \left[F(y|x) - F^+(y|x)\right] g(x)\, \pi_0^+(x)f(x)dx = 0$ for all $(y,g) \in \mathcal{B}$, while $\int \left[F(y|x) - F^+(y|x)\right] g(x)\,\pi_0^+(x)f(x)dx < 0$ if $(y,g) \notin \mathcal{B}$. Furthermore, we can show that

$$\nu_T(\cdot,\cdot) \Rightarrow \nu(\cdot,\cdot) \qquad (41)$$

with the sample paths of $\nu(\cdot,\cdot)$ uniformly continuous with respect to pseudometric ρ_a on $\mathcal{Y}\times\mathcal{G}$ with probability one. The latter holds by a standard argument because Lemma 1 implies that the pseudometric space $(\mathcal{Y}\times\mathcal{G}, \rho_a)$ is totally bounded, $\{\nu_T(\cdot,\cdot) : T \ge 1\}$ is stochastically equicontinuous, and finite dimensional convergence in distribution holds using a CLT for bounded strong mixing random variables (see Corollary 5.1 of Hall and Heyde (1980)). Therefore, using the same arguments as those in Linton et al. (2005, Proof of Theorem 1) and continuous mapping theorem, Theorem 1 is now established as desired. ■

Proof of Theorem 2. If the conditions of Theorem 1 do not hold, then $\sqrt{T}\Delta_T(y,g)$ term is no longer $o_p(1)$. Instead, we have

$$\begin{aligned}
\sqrt{T}\hat{\bar{m}}_T(y,g,\hat{\pi}^+) &= \nu_T(y,g) + \sqrt{T}\int \left[F(y|x) - F^+(y|x)\right] \\
&\quad \times g(x)\,\pi_0^+(x)f(x)dx + \sqrt{T}\Delta_T(y,g) + o_p(1).
\end{aligned}$$

Therefore, it suffices to show that

$$\sup_{y,g} \left| \hat{\Delta}_T(y,g) - \Delta_T(y,g) \right| = o_p(T^{-1/2}). \qquad (42)$$

Let

$$\begin{aligned}
H_T(y,g) &= \frac{1}{T}\sum_{t=1}^{T} \left(\left[\hat{F}''(y|x_t) - \hat{F}^{+''}(y|x_t) \right] \right. \\
&\quad \left. + \left[\hat{F}(y|x_t) - \hat{F}^{+}(y|x_t) \right] \frac{\hat{f}''(x_t)}{\hat{f}(x_t)} \right) g(x_t)\,\hat{\pi}_0^+(x_t), \\
H(y,g) &= \int \left(\left[F''(y|x) - F^{+''}(y|x) \right] \right. \\
&\quad \left. + \left[F(y|x) - F^{+}(y|x) \right] \frac{f''(x)}{f(x)} \right) g(x)\,\pi_0^+(x)f(x)dx. \qquad (43)
\end{aligned}$$

By rearranging terms, we have

$$\begin{aligned}
\hat{\Delta}_T(y,g) - \Delta_T(y,g) &= \left(\frac{\kappa-1}{\kappa T}\sum_{t=2}^{T} \frac{RQ_t}{n_t} - \frac{1}{T}\sum_{t=2}^{T}\hat{\delta}_t^2 \right) H_T(y,g) \\
&\quad + (H_T(y,g) - H(y,g)) \left(\frac{1}{T}\sum_{t=2}^{T} \hat{\delta}_t^2 \right) \\
&\quad - \left[\frac{1}{T}\left(\hat{\delta}_T^2 - \hat{\delta}_1^2 \right) \right] O(1). \qquad (44)
\end{aligned}$$

Consider the first term on the right hand side of (44). By Chebyshev inequality and Assumption B1, we can show that

$$\begin{aligned}
&\frac{1}{T}\sum_{t=2}^{T} \left(\frac{1}{n_t}\sum_{j=1}^{n_t} \sigma_{t_j}^2(\eta_{t_j}^2 - 1) \right)^2 \\
&\quad - (\kappa-1)\frac{1}{T}\sum_{t=2}^{T}\frac{1}{n_t^2}\sum_{j=1}^{n_t}\sigma_{t_j}^4 = o_p(T^{-1/2}), \qquad (45)
\end{aligned}$$

provided $\gamma > \epsilon$. Also, in the decomposition (18), J_{2t} and J_{1t} are the first and second leading terms, respectively, that satisfy

$$\frac{1}{T}\sum_{t=2}^{T} J_{1t}J_{2t} = O_p\left(n^{-\lambda}\sqrt{\frac{T^{\epsilon}}{n}} \right) = o_p(T^{-1/2}), \qquad (46)$$

provided $\gamma > (1+\epsilon)/(1+2\lambda)$. (45)–(46) imply that

$$\frac{1}{T}\sum_{t=2}^{T}\hat{\delta}_t^2 = (\kappa-1)\frac{1}{T}\sum_{t=2}^{T}\frac{1}{n_t^2}\sum_{j=1}^{n_t}\sigma_{t_j}^4 + o_p(T^{-1/2}), \qquad (47)$$

Similarly, we have

$$\begin{aligned}
\frac{1}{T}\sum_{t=2}^{T}\frac{RQ_t}{n_t} &= \frac{1}{T}\sum_{t=2}^{T}\sum_{j=1}^{n_t}\left(\frac{1}{n_t}\mu_{t_j} + \frac{1}{\sqrt{n_t}}\sigma_{t_j}\eta_{t_j} \right)^4 \\
&= \frac{1}{T}\sum_{t=2}^{T}\frac{1}{n_t^2}\sum_{j=1}^{n_t}\sigma_{t_j}^4\eta_{t_j}^4 + O_p(n^{-\frac{5}{2}}) \\
&= \kappa\frac{1}{T}\sum_{t=2}^{T}\frac{1}{n_t^2}\sum_{j=1}^{n_t}\sigma_{t_j}^4 + o_p(T^{-1/2}), \qquad (48)
\end{aligned}$$

where the second equality holds by Assumption B4 and the last equality holds by Assumptions B1, 2 and 4, provided $\gamma > 1/5$. From (47)–(48), we now have

$$\frac{\kappa-1}{\kappa T}\sum_{t=2}^{T}\frac{RQ_t}{n_t}-\frac{1}{T}\sum_{t=2}^{T}\hat{\delta}_t^2=o_p(T^{-1/2}). \tag{49}$$

Therefore, from (44), we have

$$\begin{aligned}\hat{\Delta}_T(y,g)-\Delta_T(y,g)&=o_p(T^{-1/2})+o_p(T^{-1/4})O_p\left(\frac{T^{\epsilon}\log n}{n}\right)\\&=o_p(T^{-1/2}),\end{aligned} \tag{50}$$

uniformly in (y, g), where the first term on the right hand side of (50) follows from (49), the second term follows from Assumption B4, the uniform $T^{-1/4}$—consistency of the kernel estimators similar to (35), the stochastic equicontinuity argument and (31) and the last equality holds since $\gamma > 2\epsilon + \frac{1}{4}$. This establishes Theorem 2. ■

Proof of Theorem 3. The proof is similar to the proof of Theorem 2 of Linton et al. (2005). ■

Proof of Theorem 4. The proof is similar to the proof of Theorem 3 of Linton et al. (2005). ■

Appendix B. Data descriptions and constructions

The data for estimating the daily volatility come from different sources. For the realized variance of the S&P500 cash index, we use intraday high frequency data provided by `tickdata.com`, which consist of 1-minute and 5-minute index prices in regular trading time. The squared intraday range RG_t^2 of the S&P500 index is evaluated by using data of the highest and lowest trading prices of a day, which come from `yahoo finance`. For the five DJIA constituents, their RG_t^2 are evaluated with daily highest and lowest price data from the CRSP, and their RV_t are evaluated with the intraday trade price data from the TAQ database.

The raw data of the high frequency observations from the TAQ database contain noises. In order to obtain more accurate estimations of the realized variances, we adopt the following procedures, which are suggested by Barndorff-Nielsen et al. (2009), to clean the high frequency price data of the five DJIA constituents:

1. We keep the data points between 09:30 AM to 16:00 PM (regular trading time), and delete data points with a time stamp outside this time interval.
2. We delete the data points whose prices are zero.
3. We keep data points which the trade occurred on AMEX (A), NYSE (N), NASD (T/Q), and delete the rest data points.
4. Data points which are corrected traded are deleted (their Correction Indicator is not zero, CORR!= 0).
5. Data points whose trades are not in abnormal sale condition are kept. (the entries in the column COND which do not have a letter code, or have the letter "F" or "E".)
6. If multiple trades have the same time stamp, we use their median price.
7. We delete the data point in which the absolute difference between its price and median of 50 neighborhood observations is larger than five times mean absolute deviation from the median.

Note that the above rule 7 is to replace rule T4 in Barndorff-Nielsen et al. (2009) for cleaning outliers in the high frequency trade data. The rule T4 uses the quote data to discipline the trade data: If the trade prices are above the ask plus bid–ask spread or below the bid minus the bid–ask spread, they will be deleted. However, it can be shown that such rule in practice is rarely activated. Since our raw data are the trade data, in order to more efficiently implementing the cleaning procedures without using the quote data, we use a more viable rule such as rule 7 for dealing with the outliers.

The time unit of the cleaned intraday price data is one second, but the data points are not equally-spaced. We then transform the cleaned data to equally-spaced data by using the last-tick method. We set the time intervals for the equally-spaced price data equal to some frequently used choices: 3 s 5 s and 10 s for evaluating the modified pre-averaging based estimation of quadratic variation $C_{t,a}^{n_t}$, 1 min for evaluating quantities for measuring the leverage effect of Wang and Mykland (2014) in Appendix D.2 and Aït-Sahalia et al. (2013) in Appendix D.3 and 5 min for evaluating the benchmark 5-min realized variance $RV_t^{5\text{ min}}$ and the bi-power variation BV_t and tri-power variation TV_t in Appendix D.1. Finally the equal-spaced price data are used to calculate intraday log returns for obtaining different conditional daily volatility estimations and quantities for measuring the leverage effect.

Appendix C. Evaluating $C_{t,a}^{n_t}$

In the following we give a detailed discussion on how we evaluate the modified pre-averaging based estimation of quadratic variation $C_{t,a}^{n_t}$ in Jacod et al. (2009) and Hautsch and Podolskij (2013) and its empirical properties. We first evaluate the daily $C_{t,a}^{n_t}$ with intraday log returns sampled from different time intervals. We set the window length for the weighted averaged return $k_{n_t} = \lceil\theta\sqrt{n_t}\rceil$. In Fig. 6, we show θ-signature plots of daily $\sqrt{CV_{t,a}^{n_t}}$ for the five DJIA constituents when 3, 5 and 10-second intraday log returns are used for the evaluations. The θ-signature plots express time series averages of daily $\sqrt{C_{t,a}^{n_t}}$ when different $\theta \in [0.1, 2]$ are used. For the purpose of comparisons, in each plot we also show time series average of daily $\sqrt{RV_t^{5\text{ min}}}$ and time series average of daily RG_t (the horizontal lines).

The choice of the pre-averaging parameter θ in $C_{t,a}^{n_t}$ can have effects on the quality of the estimation. Too small θ results in too short window length k_{n_t}, which will not effectively reduce the market microstructure noise and estimation of the conditional volatility will be biased. On contrary, too high θ results in too long window length k_{n_t}, which will "oversmooth" the estimators and the resulting estimation of conditional volatility will also be biased. As can be seen in Fig. 6, the θ-signature plots consistently indicate that as θ increases, time series averages of daily $\sqrt{C_{t,a}^{n_t}}$ decreases. When θ is small, on average daily $\sqrt{C_{t,a}^{n_t}}$ has a much higher value than do daily $\sqrt{RV_t^{5\text{ min}}}$ and RG_t. When θ is becoming large, the average of daily $\sqrt{C_{t,a}^{n_t}}$ is declining to a level lower than the average of daily $\sqrt{RV_t^{5\text{ min}}}$. It is worth noting that such a decline becomes stable when $\theta \geq 0.6$. This suggests that in our cases, too small (large) θ might result in an upward (a downward) biased estimation for the daily conditional volatility. In addition, given the same θ, daily $\sqrt{C_{t,a}^{n_t}}$ evaluated with higher frequency data on average is higher than those evaluated with lower frequency data, which suggests that as higher data frequency is used, a higher value of θ is needed to reduce the estimation bias.

Hautsch and Podolskij (2013) indicate that to optimally choose θ, one should consider bias and efficiency trade-off, local signal-to-noise ratio and asset liquidity. However, these criteria are time varying and taking all of them into account perhaps will make the choice problem too complicated to solve in practice. Thus they suggest a simple rule that an optimal choice of θ should be the smallest possible value at which the θ-signature plots tend to

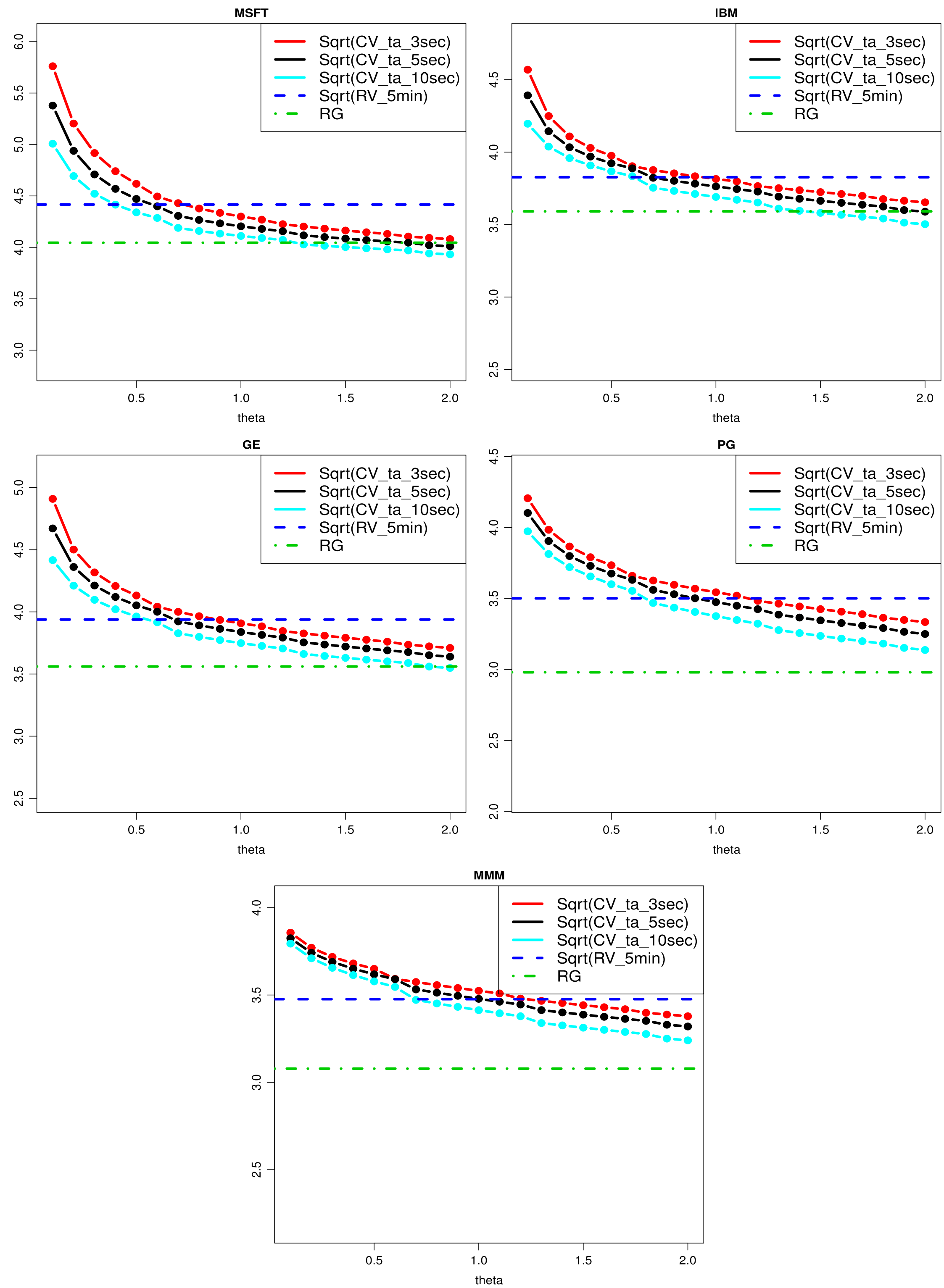

Fig. 6. θ-signature plots of daily $\sqrt{CV_t^{n_t a}}$ for the five constituents from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). The θ-signature plots show time series averages of daily $\sqrt{C_t^{n_t a}}$ when different θ are used. The daily $C_t^{n_t a}$ is evaluated with 3, 5, and 10-second intraday high frequency log return data and the pre-averaging parameter $\theta \in [0.1, 2]$. The horizontal lines in each plot are time series average of daily $\sqrt{RV_t^{5\text{ min}}}$ and time series average of daily RG_t. The quantities of $\sqrt{CV_t^{n_t a}}$, $\sqrt{RV_t^{5\text{ min}}}$ and RG_t shown here are scaled by 252 (annualized). The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

stabilize and the divergence between different sampling schemes tends to vanish. Note that the rule aims to find a "global" optimal θ rather than a "local" optimal θ based on the aforementioned time varying criteria. Through an intensive empirical analysis, they find that an optimal choice of the pre-averaging parameter θ is between 0.3 and 0.6.

Inspecting the θ-signature plots in Fig. 6, we find setting $\theta = 0.6$ seems to satisfy the above simple rule. We therefore choose

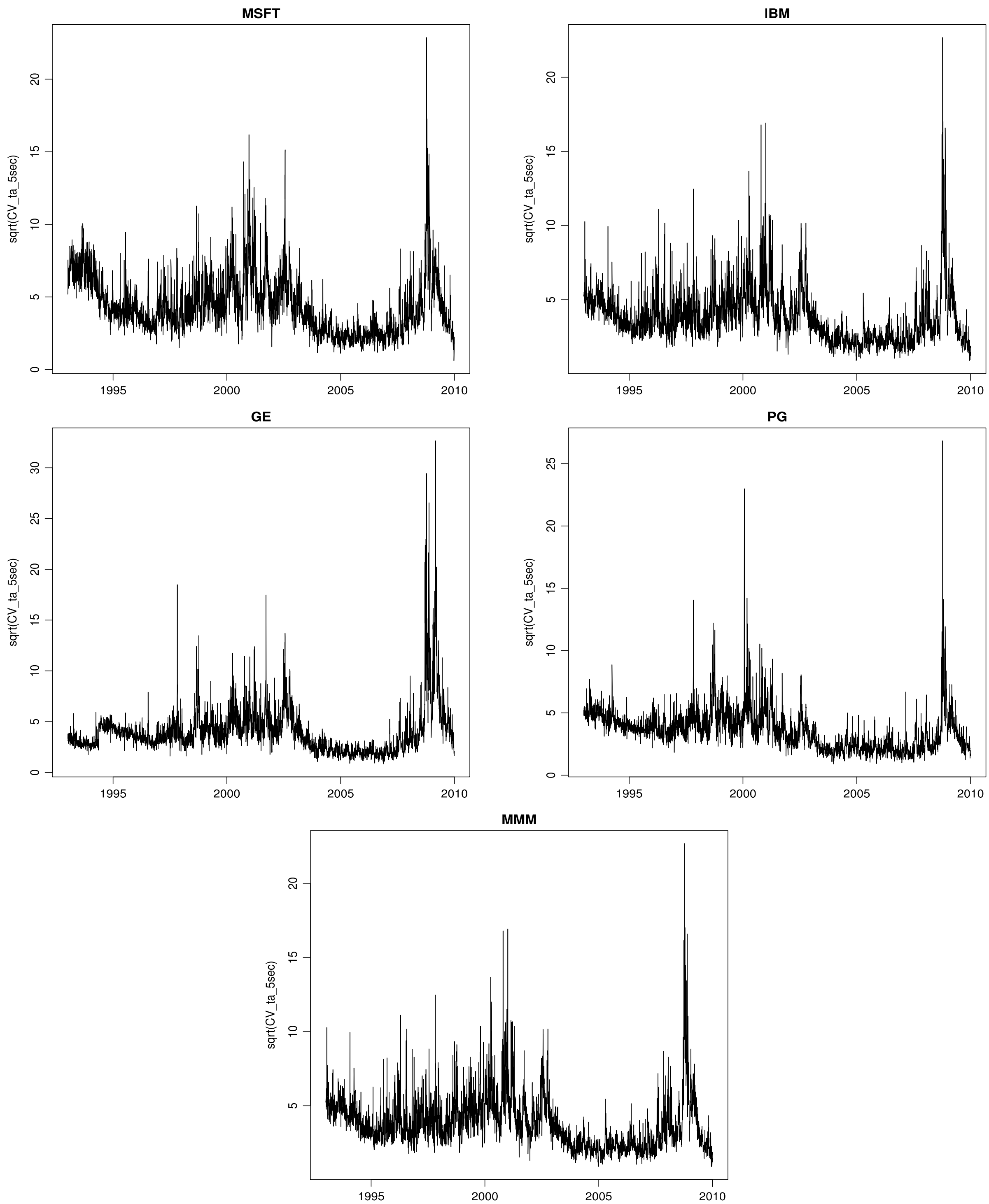

Fig. 7. Time series plots of daily $\sqrt{CV_t^{n_t a}}$ with $\theta = 0.6$ for the five constituents from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). $C_t^{n_t a}$ is evaluated with 5-second intraday high frequency log return data. The quantities of $\sqrt{CV_t^{n_t a}}$ shown here are scaled by 252 (annualized). The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

to use daily $C_{t,a}^{n_t}$ evaluated with $\theta = 0.6$ as inputs to evaluate the conditional leverage hypothesis test statistic. For testing the conditional leverage hypothesis, we focus on the cases when the daily $C_{t,a}^{n_t}$ is evaluated with 5 s log return data. We show time series plots of the daily $\sqrt{C_{t,a}^{n_t}}$ for the five DJIA constituents in Fig. 7.

Table 12 shows summary statistics of the daily $\sqrt{C_{t,a}^{n_t}}$. Comparing summary statistics in this table with those in Table 1, we can see that except for MSFT, the daily $\sqrt{C_{t,a}^{n_t}}$ on average is higher than the daily $\sqrt{RV_t^{5\text{ min}}}$ for the other four DJIA constituents. In addition the

Table 12

The table shows summary statistics of daily $\sqrt{C^{n_t}_{t,a}}$ evaluated with 5 s intraday log returns for five constituents of the Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). For the evaluations we set the preaveraging parameter $\theta = 0.6$. We scale $\sqrt{C^{n_t}_{t,a}}$ by 252 before we calculate the statistics. The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

	Min.	Mean	Max.	Std.	Skew.	Kurt.	ACF(1)
MSFT	0.6092	4.3990	22.8700	1.9869	1.3998	4.3696	0.8361
IBM	0.8872	3.8880	22.6800	1.8290	2.0358	8.8230	0.8078
GE	0.8317	4.0010	32.6500	2.3803	3.5180	21.6187	0.8452
PG	0.8988	3.6320	26.8300	1.6242	2.3189	17.8697	0.7984
MMM	0.7658	3.5900	25.2400	1.5804	2.5842	16.3450	0.7860

daily $\sqrt{C^{n_t}_{t,a}}$ also has a lower standard deviation and a higher first order autocorrelation (ACF(1)) than does the daily $\sqrt{RV_t^{5\text{ min}}}$.

Appendix D. Alternative methods

In this appendix we consider three alternative parametric methods on detecting the leverage effect to check whether our results in the main text still hold. The first method is based on estimating HAR-RV type models with the leverage effect. The other two methods are recently developed by Wang and Mykland (2014) and Aït-Sahalia et al. (2013), which are based on continuous time finance and high frequency data estimations. We find results from these methods lend further supports to our previous empirical analysis.

D.1. HAR-RV type model with the Leverage effect

We first use a modified HAR-RV model to test whether the level effect exists. Following Corsi and Renò (2012), we incorporate the leverage effect into the HAR-RV model:

$$\begin{aligned}\hat{\sigma}_t^2 &= \alpha_D + \beta_{RD}\hat{\sigma}_{t-1}^2 + \beta_{RW}\hat{\sigma}_{t-1,week}^2 + \beta_{RM}\hat{\sigma}_{t-1,month}^2 \\ &\quad + \gamma_+\left(r_{t-1}^{daily}\right)^2 \times 1\left\{r_{t-1}^{daily} \geq 0\right\} + \gamma_-\left(r_{t-1}^{daily}\right)^2 \\ &\quad \times 1\left\{r_{t-1}^{daily} < 0\right\} + \varepsilon_t, \end{aligned} \tag{51}$$

where $\hat{\sigma}_t^2$ is the realized variance of (8) and $\hat{\sigma}_{t-1,week}^2$ and $\hat{\sigma}_{t-1,month}^2$ are normalized weekly and monthly realized variances described in Section 6.3. If $\gamma_- > \gamma_+$, we can say there is a leverage effect. Thus the null hypothesis to test whether the leverage effect presents or not can be specified as $H_0 : \gamma_- \leq \gamma_+$ and if it is rejected, we have evidence to say that the leverage effect may exist.

The inputs for fitting the regressions $(\hat{\sigma}_t^2)$ are all estimated with 5-min intraday log returns. In addition to $\left(r_{t-1}^{daily}\right)^2$, $\left|r_{t-1}^{daily}\right|$ is also used in the regressions. The OLS estimation results of the linear regression (51) are shown in Table 13. In the parenthesis under the estimated coefficients are the t-statistics obtained with Newey–West standard errors. In the table, it can be seen that the in-sample fittings have moderately high adjusted R^2 (all above 0.45), which is one of the most documented empirical features of the HAR-RV type model. The estimated γ_- are all positive and most of them are statistically significant. The estimated γ_+, however, are only statistically significant in a few cases and some of them are even negative. When $\left(r_{t-1}^{daily}\right)^2$ is used, the estimated γ_- ranges from 0.0127 (PG) to 0.2228 (GE). For the case of $\left|r_{t-1}^{daily}\right|$, the same estimates range from 0.0037 (IBM) to 0.0133 (GE). The results suggest that in our cases, realized variance of GE reacts most when its lagged return receives a negative impact.

We show results of testing $H_0 : \gamma_- \leq \gamma_+$ in the last two columns of Table 13. Except for PG and MMM, the t test statistics are all above 2.76, which suggests the hypothesis can be rejected at the significant level 0.003. For PG and MMM, however, the hypothesis can still be well rejected at the significant level of 0.046. The evidence shown here indicates that the hypothesis at least can be rejected at a moderate significant level. To sum, for the five DJIA constituents and the S&P500 index, there is evidence to say that negative shocks to their asset returns have more impacts on their realized variances than do equal positive shocks, and the leverage effect may exist in their price process.

D.2. Wang and Mykland (2014)

We then use the quadratic co-variation approach proposed by Wang and Mykland (2014) (henceforth WM) to verify the existence of the leverage effect between intraday log returns and volatilities. Consider the following data generating process for the log price $X_t := \log P_t$ and volatility σ_t:

$$dX_t = \mu_t dt + \sigma_t dW_t, \tag{52}$$

$$d\sigma_t = a_t dt + f_t dW_t + g_t dB_t, \tag{53}$$

where W_t and B_t are two mutually independent Brownian motions. WM propose to use the quadratic co-variation between X_t and $F\left(\sigma_t^2\right)$ as a quantitative measure for the contemporaneous leverage effect:

$$\left\langle X, F\left(\sigma^2\right)\right\rangle_T = 2\int_0^T F'\left(\sigma_t^2\right)\sigma_t^2 f_t dt. \tag{54}$$

The function $F(.)$ is twice differentiable and monotonic on $(0, \infty)$, and in the following we will assume either $F(x) = x$ or $F(x) = 1/2\log(x)$.

Recall that in this paper we define the leverage effect as negative shocks to prices/returns affect volatility more than equal positive shocks, which is somehow different from the contemporaneous leverage effect defined in (54). While (54) only evaluates covariation between X_t and $F\left(\sigma_t^2\right)$, it does not tell whether the negative or positive shocks have more effects on the volatility. One way to link (54) and our definition of the leverage effect is to require the parameter f_t to be negative: If $f_t < 0$, negative (positive) shocks to the log returns increase (decrease) volatility.

Suppose within the time interval $[0, T]$, the log price process X_t is observed at equally spaced time stamps; i.e., r_t is observed every $\triangle t_{n,i+1} = T/n$ units of time. To empirically estimate $\left\langle X, F\left(\sigma^2\right)\right\rangle_T$, we first divide the observed $X_t's$ into different blocks. Suppose the number of such blocks is K_n and each block contains $M_n = \lfloor c\sqrt{n}\rfloor$ observations, where c is some constant. WM propose to use

$$\widehat{\left\langle X, F\left(\sigma^2\right)\right\rangle}_T = 2\sum_{i=0}^{K_n-2}\left(X_{\tau_{n,i+1}} - X_{\tau_{n,i}}\right)\left(F\left(\hat{\sigma}^2_{\tau_{n,i+1}}\right) - F\left(\hat{\sigma}^2_{\tau_{n,i}}\right)\right)$$

to estimate $\left\langle X, F\left(\sigma^2\right)\right\rangle_T$, where $\tau_{n,i}$, $i = 0, \ldots, K_n - 1$ is the lower bound of the ith block, and

$$\hat{\sigma}^2_{\tau_{n,i+1}} = \frac{n}{M_n \times T}\sum_{t_{n,j}\in\left(\tau_{n,i}, \tau_{n,i+1}\right]}\left(X_{t_{n,j+1}} - X_{t_{n,j}}\right)^2$$

Table 13
The table shows the OLS estimation results of the HAR-RV model augmented with the terms for the leverage effect (51) and results of testing the hypothesis $H_0 : \gamma_- \le \gamma_+$. Daily data are used for the OLS fittings. Daily realized variances are estimated with 5 min log returns. In the parenthesis are t-statistics obtained from using Newey–West standard errors with 18 lag periods. The cases considered here are S&P500 index and five stocks from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

		α_D	β_{RD}	β_{RW}	β_{RM}	γ_+	γ_-	Adj R^2	$H_0 : \gamma_- \le \gamma_+$ t-statistic	p-value
MSFT	$\left(r_{t-1}^{daily}\right)^2$	0.0000	0.3202	0.1995	0.3465	0.0127	0.0644	0.5925	2.76	0.0029
		(4.2696)	(5.4763)	(3.5118)	(3.6269)	(1.6222)	(3.2397)			
	$\left\lvert r_{t-1}^{daily}\right\rvert$	0.0000	0.3076	0.1884	0.3467	0.0015	0.0057	0.6004	5.72	0.0000
		(0.4198)	(5.3052)	(3.2045)	(3.5082)	(2.0771)	(6.1779)			
IBM	$\left(r_{t-1}^{daily}\right)^2$	0.0000	0.2128	0.2122	0.4306	−0.0008	0.0513	0.4906	3.09	0.0010
		(3.9354)	(2.0372)	(3.9484)	(2.8959)	(−0.0992)	(2.8537)			
	$\left\lvert r_{t-1}^{daily}\right\rvert$	0.0000	0.2117	0.2056	0.4251	−0.0006	0.0037	0.4906	4.60	0.0000
		(1.2266)	(2.0603)	(3.7706)	(2.8827)	(−1.1340)	(4.1343)			
GE	$\left(r_{t-1}^{daily}\right)^2$	0.0000	0.3058	0.1335	0.3446	0.0015	0.2228	0.5809	3.43	0.0040
		(3.0457)	(4.3681)	(2.1598)	(6.0780)	(0.0860)	(3.6753)			
	$\left\lvert r_{t-1}^{daily}\right\rvert$	−0.0000	0.3178	0.1293	0.3360	0.0025	0.0133	0.5682	3.73	0.0001
		(−1.6384)	(4.3389)	(2.1617)	(5.8487)	(1.2616)	(4.2635)			
PG	$\left(r_{t-1}^{daily}\right)^2$	0.0000	0.3643	0.3362	0.1853	−0.0148	0.0127	0.4504	1.71	0.0438
		(4.2570)	(5.6873)	(5.5528)	(1.8452)	(−0.8230)	(0.8580)			
	$\left\lvert r_{t-1}^{daily}\right\rvert$	0.0000	0.3194	0.3257	0.1725	0.0010	0.0052	0.4702	2.72	0.0033
		(0.8859)	(4.8528)	(6.0125)	(1.6034)	(0.9804)	(2.1908)			
MMM	$\left(r_{t-1}^{daily}\right)^2$	0.0000	0.0826	0.1864	0.5745	0.0087	0.1333	0.5219	1.69	0.0452
		(2.6976)	(0.8075)	(2.8080)	(4.1842)	(0.8651)	(1.7520)			
	$\left\lvert r_{t-1}^{daily}\right\rvert$	−0.0000	0.0863	0.1675	0.5839	0.0013	0.0058	0.5157	2.12	0.0170
		(−0.0963)	(0.8581)	(2.2617)	(3.8594)	(1.8297)	(2.2099)			
S&P500	$\left(r_{t-1}^{daily}\right)^2$	0.0000	0.1778	0.1375	0.4057	0.0368	0.2100	0.6841	3.88	0.0001
		(2.8481)	(1.8672)	(3.1106)	(3.5751)	(1.5492)	(3.7249)			
	$\left\lvert r_{t-1}^{daily}\right\rvert$	−0.0000	0.1751	0.1203	0.4852	0.0014	0.0074	0.6402	3.46	0.0003
		(−1.6242)	(1.5078)	(1.8701)	(3.2776)	(1.4557)	(3.0290)			

is an estimate for the integrate variance within the block $\left(\tau_{n,i}, \tau_{n,i+1}\right]$. They show that

$$n^{\frac{1}{4}}\left(\langle X, \widehat{F(\sigma^2)}\rangle_T - \langle X, F(\sigma^2)\rangle_T\right)$$

converges in law to $Z \times B(c, T)$, where Z is a standard normal random variable and independent of any information up to time T and

$$B(c,T) = \sqrt{\frac{16}{c}\int_0^T \left(F'\left(\sigma_t^2\right)\right)^2 \sigma_t^6 dt + cT\int_0^T \left(F'\left(\sigma_t^2\right)\right)^2 \sigma_t^4\left(\frac{44}{3}f_t^2 + \frac{22}{3}g_t^2\right)dt}.$$

For empirically estimating $B(c, T)$, we can use $\sqrt{G_n^1 + G_n^2}$, where

$$G_n^1 = 2\sqrt{n}\sum_i \left(X_{\tau_{n,i+1}} - X_{\tau_{n,i}}\right)^2 \left(F\left(\hat{\sigma}^2_{\tau_{n,i+1}}\right) - F\left(\hat{\sigma}^2_{\tau_{n,i}}\right)\right)^2,$$

$$G_n^2 = 2\frac{M_n T}{\sqrt{n}}\sum_i \hat{\sigma}^2_{\tau_{n,i}}\left(F\left(\hat{\sigma}^2_{\tau_{n,i+1}}\right) - F\left(\hat{\sigma}^2_{\tau_{n,i}}\right)\right)^2.$$

It can be shown that $G_n^1 + G_n^2$ converges in probability to $B(c, T)^2$. We also can use

$$\tilde{\sigma}^2_{\tau_{n,i+1}} = \frac{n}{M_n \times T}\sum_{t_{n,j}\in\left(\tau_{n,i},\tau_{n,i+1}\right]}\left(X_{t_{n,j+1}} - \overline{\triangle X_{\tau_{n,i+1}}}\right)^2$$

to replace $\hat{\sigma}^2_{\tau_{n,i+1}}$ in the estimation, where $\overline{\triangle X_{\tau_{n,i+1}}} = 1/M_n \left(X_{\tau_{n,i+1}} - X_{\tau_{n,i}}\right)$ is an average of log return changes within the time interval $\left(\tau_{n,i}, \tau_{n,i+1}\right]$. Let $\langle X, \widetilde{F(\sigma^2)}\rangle_T$ denote the leverage estimate when $\tilde{\sigma}^2_{\tau_{n,i+1}}$ is used. As shown by WM, $\hat{\sigma}^2_{\tau_{n,i+1}}$ and $\tilde{\sigma}^2_{\tau_{n,i+1}}$ are asymptotic equivalent, and therefore $\langle X, \widehat{F(\sigma^2)}\rangle_T$ and $\langle X, \widetilde{F(\sigma^2)}\rangle_T$ are also asymptotic equivalent. Furthermore, the following two test statistics can be used to detect local leverage effect,

$$L_1 := \frac{n^{\frac{1}{4}}\left(\langle X, \widehat{F(\sigma^2)}\rangle_T - \langle X, F(\sigma^2)\rangle_T\right)}{\sqrt{G_n^1 + G_n^2},}, \quad (55)$$

$$L_2 := \frac{n^{\frac{1}{4}}\left(\langle X, \widetilde{F(\sigma^2)}\rangle_T - \langle X, F(\sigma^2)\rangle_T\right)}{\sqrt{G_n^1 + G_n^2},}. \quad (56)$$

It can be shown that L_1 and L_2 both converge in law to standard normal.

We use 1-min equally spaced data,[9] and set T = one day and $M_n = 30$ for the estimations. Fig. 8 shows time series plots of daily standardized quadratic co-variations, which are just daily L_1 and L_2 with $\langle X, F(\sigma^2)\rangle_T = 0$. Table 14 reports some statistics of these standardized quadratic co-variations. It can be seen that except MMM, all the rest four DJIA constituents and S&P500 index on average have negative L_1 and L_2. Without considering signs of the estimates, we compare absolute values of the daily L_1 and L_2

9 The finest high frequency data we have for the S&P 500 index are 1 min data. Indeed we have finer high frequency data for the five DJIA constituents. However, to unify the usage of the data for the estimations, we also use 1 min data to evaluate these quantities for the five DJIA constituents here and in Appendix D.3.

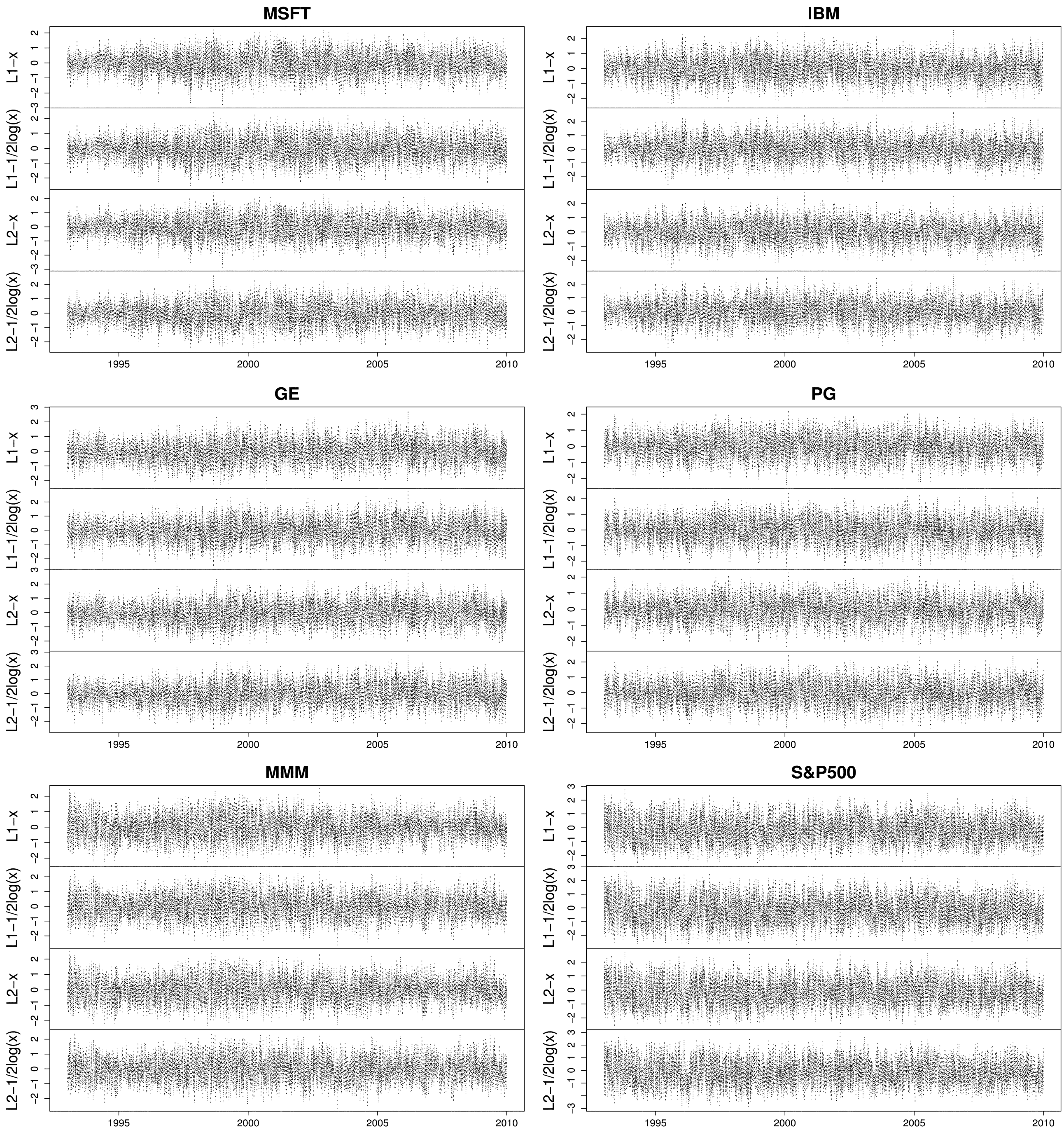

Fig. 8. Time series plots of daily standardized quadratic co-variations between intraday log return X_t and function of spot variance $F\left(\sigma_t^2\right)$ obtained by using the method in Wang and Mykland (2014). The standardized quadratic co-variations are defined as L_1 and L_2 in (55) and (56) with $\langle X, F\left(\sigma^2\right)\rangle_T = 0$. We use 1-min equally spaced data, and set $T =$ one day and $M_n = 30$ for the estimations. We assume $F(x) = x$ and $F(x) = 1/2\log(x)$. The cases considered here are the S&P500 index and five constituents from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

with two critical values 1.96 and 2.58 (corresponding to significant levels $\alpha = 0.05$ and 0.01 under standard normal). We find very few days have significant L_1 and L_2: Among the 4283 days, the number of significant days ranges from 18 to 257 for $\alpha = 0.05$ and 0 to 13 for $\alpha = 0.01$. If considering significant negativity only, the numbers of significant days range from 60 to 392 for $\alpha = 0.05$ and 0 to 42 days for $\alpha = 0.01$ (with one-sided critical values equal to -1.64 and -2.33 under standard normal). The S&P500 index on average has lower standardized quadratic co-variations and more days of statistically significant test statistics than the five DJIA constituents. Overall, the results suggest that the leverage effect may still exist in the five DJIA constituents and S&P500 index when log returns and volatilities are estimated at the intraday level, but only in certain periods, the leverage effect is strong enough to be detected by the quadratic co-variation approach.

D.3. Aït-Sahalia et al. (2013)

We finally consider the method in Aït-Sahalia et al. (2013) (henceforth AFL). Consider the following CIR process for the

Table 14
The table shows minimum, mean and maximum values of daily standardized quadratic co-variations between intraday log price X_t and function of spot variance $F\left(\sigma_t^2\right)$ obtained by using the method in Wang and Mykland (2014), and number of significant days when comparing the daily standardized quadratic co-variations with different critical values. The standardized quadratic co-variations are defined as L_1 and L_2 in (55) and (56) with $\langle X, F\left(\sigma^2\right)\rangle_T = 0$. We use 1-min equally spaced data, and set $T =$ one day and $M_n = 30$ for the estimations. We assume $F(x) = x$ and $F(x) = 1/2\log(x)$. The cases considered here are the S&P500 index and five constituents from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

	MSFT				IBM			
	$L_1 - x$	$L_1 - \frac{1}{2}\log(x)$	$L_2 - x$	$L_2 - \frac{1}{2}\log(x)$	$L_1 - x$	$L_1 - \frac{1}{2}\log(x)$	$L_2 - x$	$L_2 - \frac{1}{2}\log(x)$
Min.	−2.807	−2.561	−2.906	−2.520	−2.421	−2.691	−2.509	−2.709
Mean	−0.059	−0.065	−0.058	−0.065	−0.014	−0.017	−0.015	−0.019
Max.	2.224	2.498	2.429	2.716	2.584	2.728	2.763	2.728
$\lvert.\rvert \geq 1.96$	29	46	32	47	39	51	40	54
$. \leq -1.64$	85	102	88	98	66	94	70	99
$\lvert.\rvert \geq 2.58$	2	0	1	1	1	2	1	3
$. \leq -2.33$	2	3	3	5	1	3	1	3

	GE				PG			
	$L_1 - x$	$L_1 - \frac{1}{2}\log(x)$	$L_2 - x$	$L_2 - \frac{1}{2}\log(x)$	$L_1 - x$	$L_1 - \frac{1}{2}\log(x)$	$L_2 - x$	$L_2 - \frac{1}{2}\log(x)$
Min.	−2.322	−2.613	−2.51	−2.601	−2.416	−2.354	−2.426	−2.403
Mean	−0.038	−0.039	−0.042	−0.042	−0.037	−0.038	−0.041	−0.043
Max.	2.815	2.765	2.790	2.851	2.263	2.475	2.260	2.501
$\lvert.\rvert \geq 1.96$	42	63	42	70	18	35	20	37
$. \leq -1.64$	83	108	86	109	60	95	66	101
$\lvert.\rvert \geq 2.58$	1	2	1	2	0	0	0	0
$. \leq -2.33$	0	1	1	2	1	2	1	2

	MMM				S&P500			
	$L_1 - x$	$L_1 - \frac{1}{2}\log(x)$	$L_2 - x$	$L_2 - \frac{1}{2}\log(x)$	$L_1 - x$	$L_1 - \frac{1}{2}\log(x)$	$L_2 - x$	$L_2 - \frac{1}{2}\log(x)$
Min.	−2.362	−2.638	−2.383	−2.646	−2.605	−2.735	−2.620	−2.987
Mean	0.001	0.003	0.000	0.003	−0.191	−0.186	−0.186	−0.186
Max.	2.496	2.471	2.504	2.434	2.850	2.785	2.781	2.969
$\lvert.\rvert \geq 1.96$	43	64	50	71	148	244	157	257
$. \leq -1.64$	88	99	87	101	308	379	307	392
$\lvert.\rvert \geq 2.58$	0	1	0	1	2	8	4	13
$. \leq -2.33$	1	6	1	6	15	39	19	42

squared volatility:

$$dv_t = \alpha_v\left(\theta - v_t\right) + \kappa\sqrt{v_t}\,dB_t,$$

where $v_t := \sigma_t^2$, and $2\alpha_v\theta > \kappa^2$. For the log price process, AFL assume it follows the same process as in (52) except now $\mathbb{E}\left(dB_t dW_t\right) = f dt$, where f is a constant. It can be shown that the parameter f is the limit of correlation between $v_{t+l} - v_t$ and $X_{t+l} - X_t$ when the time interval l approaches zero, i.e.,

$$f = \lim_{l\to 0} Corr\left(v_{t+l} - v_t, X_{t+l} - X_t\right). \tag{57}$$

AFL use the limit correlation above as a measure for the leverage effect, which is not exactly the same as the leverage effects defined in Wang and Mykland (2014) and our paper. To link the limit correlation (57) to our definition of the leverage effect, again we may assume f to be negative.

Let $\hat{f}_k$ and f_k be the sample and true correlations between the difference of the estimated integrated variances and difference of log prices at time $t + k\triangle$ and t. Here $\triangle$ is the time unit and we follow AFL to assume it as one day for our cases.[10] Under some regular conditions, AFL show that f_k and f satisfy the following linear relationship:

$$f_k = f + b \times k + o\left(k\triangle\right),$$

which provides an easily-implementable way to identify the limit correlation f: That is, running a linear regression of f_k (or $\hat{f}_k$ if f_k is unknown) on the intercept term and k, and the estimated intercept term can be used as an estimate of f. Furthermore, AFL propose the following data driven procedures for practically estimate the linear regression:

1. For each $k = 1, \ldots, K$, we calculate $\hat{f}_k$, and rank the $\hat{f}_k^{\prime}s$ for $k = k_0, \ldots, \lfloor K/2 \rfloor$. Then we take the three smallest values of these ranked $\hat{f}_k^{\prime}s$. Let $k_{(1)}, k_{(2)}$, and $k_{(3)}$ be the corresponding indices $k's$ that the three smallest $\hat{f}_k^{\prime}s$ have. Let $\underline{k}^* = \max\left(k_{(1)}, k_{(2)}, k_{(3)}\right)$.
2. Regressing $\hat{f}_k$ on k with $k = \underline{k}^*, \ldots, \underline{k}^* + m$, where $m = a_0, \ldots, K - \underline{k}^*$. Let m^* denote the value of m that the regression yields the highest (unadjusted) R^2. Then let $\overline{k}^* = \underline{k}^* + m^*$. The estimated intercept term of the regression with the data $\left\{k, \hat{f}_k\right\}_{k=\underline{k}^*}^{\overline{k}^*}$ is the final estimate of f.

As shown in Aït-Sahalia et al. (2013), the sample correlation $\hat{f}_k$ is a bias estimation for the true f_k. To improve performance of the data driven approach above, we can replace the sample correlation $\hat{f}_k$ with the following bias corrected estimation in the above data driven procedures:

$$\hat{f}_k^{bc} = \gamma\frac{2\sqrt{k^2 - k/3}}{2k - 1}\hat{f}_k, \tag{58}$$

where

$$\gamma = \left(1 - \frac{4\triangle^2\,\mathbb{E}\left(v_t^2\right)}{nVar\left(RV_{t+k\triangle} - RV_t\right)}\right)^{-\frac{1}{2}}.$$

Here the parameter n is the number of log return observations used to estimate the realized variance. For estimating $\triangle^2\,\mathbb{E}\left(v_t^2\right)$, we can first evaluate the realized quarticity:

$$QV_t = \frac{n_t}{3}\sum_{j=1}^{n_t} r_{t_j}^4,$$

[10] Note that in Aït-Sahalia et al. (2013), the basic time unit for t is one year, so $\triangle =$ one day $= 1/252$.

Table 15
The table shows estimated $f = \lim_{l\to 0} Corr(v_{t+l} - v_t, X_{t+l} - X_t)$ by using the data driven method in Aït-Sahalia et al. (2013). Here $X_t := \log P_t$ and $v_t := \sigma_t^2$. The estimates are based on sample correlation and bias corrected sample correlation (denoted by $\hat{f}_k$ and $\hat{f}_k^{bc}$) between daily returns r_t^{daily} and the difference of the realized variances RV_t^{1min}. We also report the upper and lower bounds $\overline{k}^*$ and $\underline{k}^*$ for $\left\{k, \hat{f}_k\right\}_{k=\underline{k}^*}^{\overline{k}^*}$ used for estimating the regression, and the (unadjusted) R^2 of the regression. The cases considered here are the S&P500 index and five constituents from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

	Uncorrected $\hat{f}_k$					
	IBM	MSFT	GE	PG	MMM	S&P500
f	−0.24	−0.31	−0.48	−0.32	−0.28	−0.36
$\overline{k}^*$	15	21	44	29	17	17
$\underline{k}^*$	55	53	87	91	54	94
R^2	0.97	0.92	0.96	0.98	0.94	0.89
	Bias Corrected $\hat{f}_k^{bc}$					
	IBM	MSFT	GE	PG	MMM	S&P500
f	−0.25	−0.32	−0.48	−0.34	−0.30	−0.37
$\overline{k}^*$	15	20	31	29	17	15
$\underline{k}^*$	55	53	88	88	54	93
R^2	0.97	0.93	0.97	0.98	0.95	0.91

and then calculate its sample mean. In the following we use

$$\hat{\gamma} = \left(1 - \frac{4 \times \text{sample mean of } QV_t}{\text{sample mean of } n_t \times \text{sample variance of } RV_t}\right)^{-\frac{1}{2}} \quad (59)$$

to replace γ for $\hat{f}_k^{bc}$.

All the following empirical analyses are in daily basis: We use 1-min log returns to estimate the daily realized variance and quarticities in $\hat{\gamma}$ and evaluate $\hat{f}_k$ with daily return r_t^{daily} and the estimated realized variance. Fig. 9 plots $\hat{f}_k$ and $\hat{f}_k^{bc}$ against k for the five DJIA constituents and S&P500 index. For the cases, the two estimated correlations are all negative over k, and the bias corrected correlations are constantly lower than the uncorrected ones. Except for small k, the two correlation estimations follow extremely similar patterns within each case. Comparing the results over different cases, however, the patterns of the correlation estimations are somehow different, but they often have a higher value as k equals one and then suddenly drop to a lower value as k deviates from one. It also can be seen that as $k > 100$, the estimations gradually become stable and all of them steadily move either up or down as k becomes large. Overall, the estimated correlations vary substantially as the time interval k changes, no matter whether they are bias corrected or not; and the smaller the time interval k, the more possible that we will get a higher estimated correlation.

To estimate f, we may use $\hat{f}_1$ or $\hat{f}_1^{bc}$. As shown above, however, it is very likely that we get a higher estimated correlation with $k = 1$ than with $k > 1$, and the high $\hat{f}_1$ or $\hat{f}_1^{bc}$ perhaps is an upward bias estimate for the limit correlation f. To obtain a more accurately estimated f, we use the data driven procedures introduced above. We set $K = 252$, $k_0 = 6$ and $a_0 = 11$ for the data driven procedures. In Table 15 we report the final estimate of f based on the estimated intercept term from running linear regressions of $\hat{f}_k$ (or $\hat{f}_k^{bc}$) on $k \in \left[\underline{k}^*, \overline{k}^*\right]$, and the unadjusted R^2 of the regression. The R^2 is very high for all of the six cases, and $\underline{k}^*$ and $\overline{k}^*$ in each case are similar when either $\hat{f}_k$ or $\hat{f}_k^{bc}$ is used. For all the six cases, the estimated f is moderately negative, with range from −0.24 (IBM) to −0.48 (GE).[11] The negativity of the estimated f implies that on average a positive shock to the log price has a smaller impact on the volatility than does an equal negative shock, and the leverage effect may exist. It also can be seen that the estimated f from using $\hat{f}_k^{bc}$ and $\hat{f}_k$ are qualitatively similar, but the former is slightly lower than the latter.

Appendix E. Simulation results

In this section we show simulation results from sample paths generated by two models: the GJR model (Glosten et al., 1993) and the Heston model (Heston, 1993). The GJR model is a discrete time model for the conditional volatility, which has a form of Eq. (1):

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \gamma_+ r_{t-1}^2 1(r_{t-1} \geq 0) + \gamma_- r_{t-1}^2 1(r_{t-1} < 0).$$

The asset return $r_t = \varepsilon_t\sigma_t$ and ε_t is a serially independent sequence with zero mean, variance one, and symmetric density with finite kurtosis. Here we assume $\varepsilon_t \sim N(0, 1)$ and are i.i.d. Following Rodriguez and Ruiz (2012), we set $\omega = 0.035$, $\beta = 0.83$, $\gamma_- + \gamma_+ = 0.1$. To see how the test statistic of the conditional leverage hypothesis performs, we vary $\bar{\gamma} = \gamma_- - \gamma_+$ at five different levels: 0, −0.01, −0.02, −0.04 and −0.07. Notice that $\bar{\gamma} = 0$ corresponds to the least favorable configuration under the null and $\bar{\gamma} < 0$ corresponds to the alternative. The estimated conditional volatility used as an input for the test is $\hat{\sigma}_t^2 = \sigma_t^2 + \varepsilon_t^{noise}$, which may be viewed as some hypothetical volatility estimator obtained from using intraday data. We assume $\varepsilon_t^{noise} \sim N\left(0, \Delta_t \times Var\left(\sigma_t^2\right)\right)$ and are i.i.d. and $\Delta_t = 0$, 1/78 and 1/13.

It is necessary to restrict the parameters in order to ensure positivity of σ_t^2 and $\hat{\sigma}_t^2$. To ensure that σ_t^2 is positive here, Hentschel (1995) showed that $\omega > 0$, $\gamma_- + \gamma_+ \geq 0$, $\beta \geq 0$ and $\bar{\gamma} \geq 0$ should hold. Also to ensure the model is stationary, $\gamma_- - \gamma_+ < 2(1 - \gamma_- - \gamma_+ - \beta)$ should hold. In the simulations, the cases of $\bar{\gamma} < 0$ violate the necessary condition for the positivity of σ_t^2. To avoid nonpositive σ_t^2 and $\hat{\sigma}_t^2$, we have carefully checked the generated sample paths and eliminated the unqualified ones before we perform the conditional leverage hypothesis test.

In Tables 16–18 we show simulation results of the conditional leverage test for the GJR model when lengths of sample paths equal to $T = 500$, 2000 and 3500. Each scenario is simulated 1000 times. We set lag lengths of daily return $j = 1$ and subsample sizes $b = 25$, 50 and 150 for $T = 500$, $b = 50$, 100 and 300 for $T = 2000$ and $b = 75$, 150 and 450 for $T = 3500$. From the tables, it can be seen that the rejection frequency decreases with $\bar{\gamma} = \gamma_- - \gamma_+$, which indicates that power of the test statistic rises as the parameter of the leverage effect $\bar{\gamma}$ deviates from the null. The power of the test also increases with the sample size, as can be seen by comparing results of $\bar{\gamma} > 0$ in the three tables. When $T = 2000$ and the parameter of $\bar{\gamma} = -0.04$, at the significant level $\alpha = 0.01$, the rejection frequencies are around 0.86 to 0.98 (depending on different subsample sizes) and when $T = 3500$ the rejection frequencies are around 0.986 to 1. Even for $\bar{\gamma}$ at a moderately low level −0.02,

[11] One thing worth noting is that, the estimated f of MSFT and the S&P500 index shown here are different from those shown in Aït-Sahalia et al. (2013). It is perhaps because we use different sample periods and different realized variance estimators. In Aït-Sahalia et al. (2013), they use the pre-averaging approach (Jacod et al., 2009; Hautsch and Podolskij, 2013) to estimate the integrated variances. With our estimated 1-min realized variance, we re-estimate the f of MSFT with the data driven method over the same sample period as theirs (Jan-2005–June-2007). The estimated f from using $\hat{f}_k^{bc}$ ($\hat{f}_k$) is −0.90 (−0.87) and $\left[\underline{k}^*, \overline{k}^*\right] = [125, 165]$ ([125, 165]). As for the S&P500 index (sample period is from Jan-2004 to Dec-2007), the estimated f from using $\hat{f}_k^{bc}$ ($\hat{f}_k$) is −0.62 (−0.60), and $\left[\underline{k}^*, \overline{k}^*\right] = [22, 101]$ ([24, 101]).

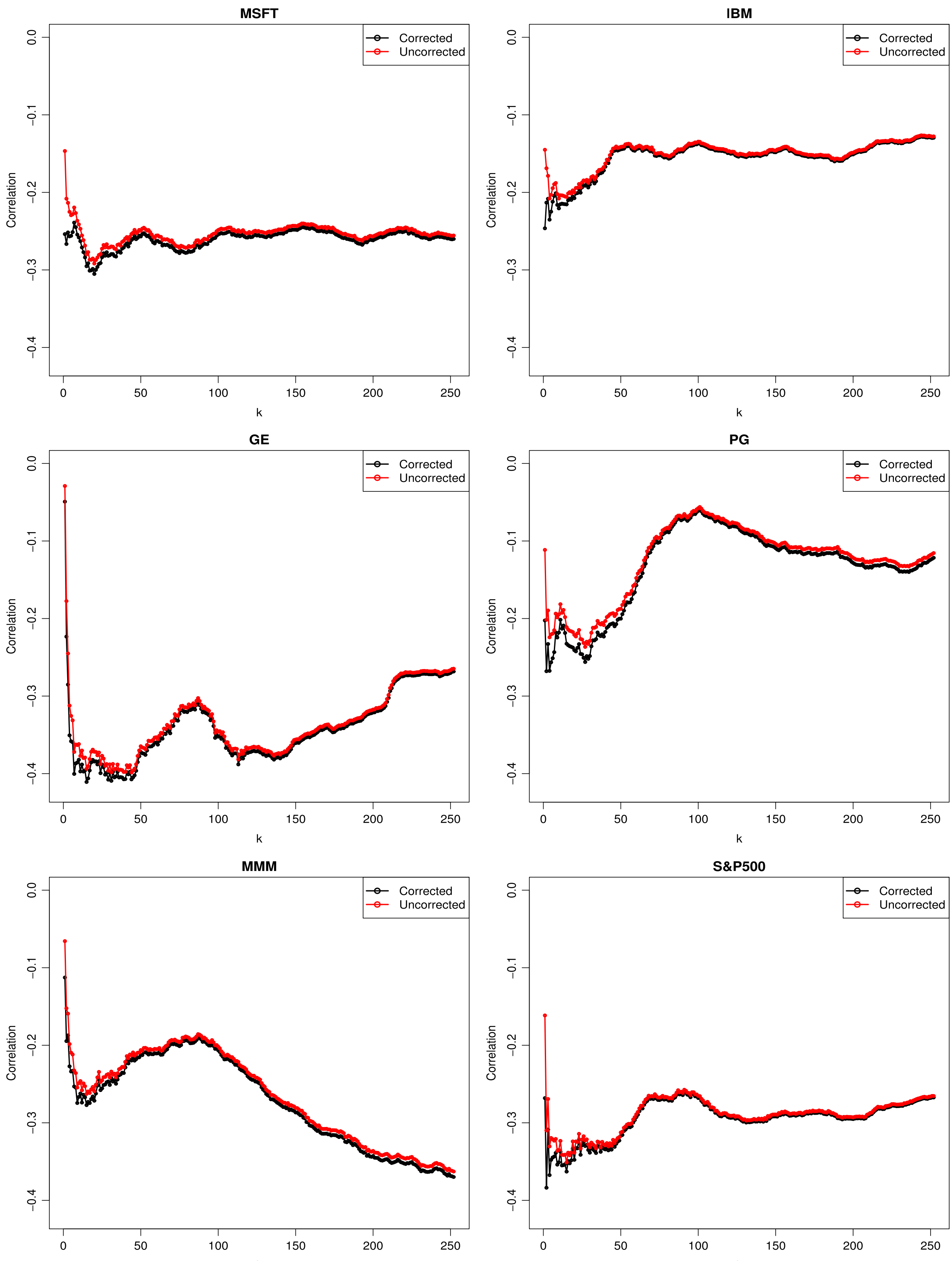

Fig. 9. Plots of sample correlation between returns and difference of the daily realized variances, $\hat{f}_k$ and bias corrected sample correlation $\hat{f}_k^{bc}$ in (58) against k. The bias corrected sample correlation $\hat{f}_k^{bc}$ is obtained by using method in Aït-Sahalia et al. (2013). We replace γ in (58) with $\hat{\gamma}$ in (59) which is estimated with mean of the realized quarticities and sample variance of the realized variances. We use 1-min equally spaced log returns to estimate the daily realized variances and quarticities. The cases considered here are the S&P500 index and five constituents from Dow Jones Industrial Averages: Microsoft (MSFT), International Business Machines Corporation (IBM), General Electric (GE), Procter & Gamble (PG) and 3M (MMM). The sample period is from Jan-04-1993 to Dec-31-2009 (4283 trading days).

as the sample size increases to 3500, the rejection frequency can be around 0.6 and 0.8 when $\alpha = 0.05$ and around 0.81 to 0.9 when $\alpha = 0.1$. But at the least favorable configuration of $\bar{\gamma} = 0$, size of the test is overall slightly lower than the corresponding significant level. When different subsample sizes are used, the power and size of the test are still stable. When sample size is low ($T = 500$) and when the alternative is not strong ($\bar{\gamma} = -0.01$ and -0.02.), using the noisy volatility (when $\Delta_t > 0$) seems to have a sightly higher chance to reject the null than using the true volatility (when $\Delta_t = 0$). But when the alternative becomes strong ($\bar{\gamma} = -0.04$

Table 16
The table shows the rejection frequencies of the conditional leverage hypothesis test at three different levels of α for the samples generated with the GJR model of Eq. (1). The length of the sample path $T = 500$ and each scenario is simulated 1000 times. Notice that $\bar{\gamma} = 0$ corresponds to the least favorable configuration under the null and $\bar{\gamma} < 0$ corresponds to the alternative. The estimated conditional volatility used as an input for the test is $\hat{\sigma}_t^2 = \sigma_t^2 + \varepsilon_t^{noise}$ and $\varepsilon_t^{noise} \sim N\left(0, \Delta_t \times Var\left(\sigma_t^2\right)\right)$ and are i.i.d. We set lag lengths of daily returns $j = 1$.

		$\Delta_t = 0$			$\Delta_t = \frac{1}{78}$			$\Delta_t = \frac{1}{13}$		
		α			α			α		
$\bar{\gamma}$	b	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1
0	25	0.003	0.012	0.023	0.006	0.021	0.036	0.008	0.029	0.071
	50	0.003	0.014	0.025	0.006	0.025	0.044	0.008	0.035	0.081
	150	0.005	0.038	0.058	0.016	0.048	0.082	0.024	0.056	0.097
−0.01	25	0.008	0.056	0.126	0.016	0.062	0.134	0.020	0.062	0.140
	50	0.020	0.066	0.130	0.020	0.070	0.146	0.024	0.064	0.132
	150	0.050	0.094	0.144	0.038	0.084	0.152	0.050	0.110	0.162
−0.02	25	0.030	0.124	0.222	0.036	0.137	0.248	0.034	0.124	0.282
	50	0.033	0.119	0.231	0.044	0.118	0.244	0.047	0.139	0.272
	150	0.086	0.160	0.240	0.064	0.160	0.256	0.086	0.145	0.229
−0.04	25	0.224	0.520	0.718	0.234	0.506	0.696	0.220	0.514	0.678
	50	0.260	0.482	0.644	0.248	0.472	0.638	0.216	0.444	0.624
	150	0.310	0.468	0.578	0.302	0.434	0.572	0.272	0.424	0.540
−0.07	25	0.961	0.997	1.000	0.957	0.996	1.000	0.888	0.981	0.996
	50	0.932	0.990	1.000	0.937	0.986	0.997	0.859	0.964	0.991
	150	0.897	0.956	0.976	0.845	0.918	0.954	0.763	0.886	0.947

Table 17
The table shows the rejection frequencies of the conditional leverage hypothesis test at three different levels of α for the samples generated with the GJR model of Eq. (1). The length of the sample path $T = 2000$ and each scenario is simulated 1000 times. Notice that $\bar{\gamma} = 0$ corresponds to the least favorable configuration under the null and $\bar{\gamma} < 0$ corresponds to the alternative. The estimated conditional volatility used as an input for the test is $\hat{\sigma}_t^2 = \sigma_t^2 + \varepsilon_t^{noise}$ and $\varepsilon_t^{noise} \sim N\left(0, \Delta_t \times Var\left(\sigma_t^2\right)\right)$ and are i.i.d. We set lag lengths of daily returns $j = 1$.

		$\Delta_t = 0$			$\Delta_t = \frac{1}{78}$			$\Delta_t = \frac{1}{13}$		
		α			α			α		
$\bar{\gamma}$	b	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1
0	50	0.002	0.008	0.034	0.000	0.014	0.040	0.004	0.018	0.056
	100	0.002	0.023	0.042	0.000	0.017	0.051	0.005	0.032	0.073
	300	0.010	0.026	0.056	0.006	0.028	0.067	0.010	0.037	0.081
−0.01	50	0.032	0.134	0.250	0.038	0.144	0.270	0.024	0.120	0.230
	100	0.042	0.146	0.278	0.054	0.138	0.264	0.036	0.126	0.254
	300	0.074	0.162	0.294	0.046	0.146	0.250	0.058	0.136	0.220
−0.02	50	0.212	0.478	0.668	0.202	0.468	0.659	0.154	0.408	0.584
	100	0.215	0.488	0.667	0.198	0.469	0.626	0.178	0.405	0.578
	300	0.268	0.467	0.616	0.247	0.466	0.612	0.169	0.324	0.471
−0.04	50	0.978	0.998	1.000	0.980	0.998	1.000	0.928	0.994	1.000
	100	0.974	1.000	1.000	0.960	0.996	1.000	0.886	0.984	0.998
	300	0.958	0.996	1.000	0.936	0.990	0.998	0.864	0.956	0.988
−0.07	50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	100	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	300	1.000	1.000	1.000	1.000	1.000	1.000	0.998	1.000	1.000

and −0.07) or sample size increases ($T = 2000$ and 3500), the true volatility performs better than the noisy volatility.

The second model used to generate the sample paths is the Heston model, which is a continuous time model and has the following form:

$$dp_t^* = \left(\mu - \frac{1}{2}v_t\right) dt + \sqrt{v_t}\left(\rho dW_{1t} + \sqrt{1-\rho^2} dW_{2t}\right), \quad (60)$$
$$dv_t = \alpha_v \left(\theta_v - v_t\right) dt + \kappa \sqrt{v_t} dW_{1t}.$$

Here p_t^* is the true underlying efficient log-price and v_t is the squared spot volatility. W_{1t} and W_{2t} are two mutually independent Brownian motions. Following Aït-Sahalia et al. (2013), we set $\mu = 0.05$, $\alpha_v = 5$, $\theta_v = 0.1$, $\kappa = 0.5$ and vary the parameter ρ at five different levels: 0.8, 0.3, 0, −0.3, and −0.8. For simulating the process, we use a simple Euler discretization scheme with the step size calibrated to one second. One simulated trading day is assumed to be 6.5 h in length and so there are 23,400 s per trading day. Notice that the parameters are in annual basis and so the step size $dt = 1/(23{,}400 \times 252)$. To avoid negative v_t in the simulated sample path, we replace the negative v_t with 0. Each scenario is simulated 1000 times. The observed log price is $p_t = p_t^* + \varepsilon_t^{noise}$, where $\varepsilon_t^{noise} \sim N\left(0, (0.0005)^2\right)$ are market microstructure noises and they are i.i.d. We use v_t and p_t to calculate four different daily conditional volatility estimators: the daily integrated variance $IV = \sum_{i=1}^{23{,}400} v_i \times dt$, daily RG^2, daily $RV^{5\text{ min}}$ and daily pre-averaging estimator C_a^n. For calculating the daily pre-averaging estimator C_a^n, we set the number of observations equal to 23,400 and the pre-averaging parameter $\theta = 0.5$.

As mentioned in (57) in Appendix D.3, it can be shown that:

$$\rho = \lim_{l \to 0} Corr\left(p_{t+l}^* - p_t^*, v_{t+l} - v_t\right).$$

The parameter ρ captures the degree of contemporaneous comovement between instant changes of p_t^* and v_t, which is the "leverage effect" defined in Aït-Sahalia et al. (2013). The definition is not exactly the same as the conditional leverage effect defined in the paper. The goals of the simulations here are to see whether ρ can result in the conditional leverage effect defined in the paper, and if so, how different values of ρ affect performances of the conditional leverage test statistic.

Table 18
The table shows the rejection frequencies of the conditional leverage hypothesis test at three different levels of α for the samples generated with the GJR model of Eq. (1). The length of the sample path $T = 3500$ and each scenario is simulated 1000 times. Notice that $\bar{\gamma} = 0$ corresponds to the least favorable configuration under the null and $\bar{\gamma} < 0$ corresponds to the alternative. The estimated conditional volatility used as an input for the test is $\hat{\sigma}_t^2 = \sigma_t^2 + \varepsilon_t^{noise}$ and $\varepsilon_t^{noise} \sim N\left(0, \Delta_t \times Var\left(\sigma_t^2\right)\right)$ and are i.i.d. We set lag lengths of daily returns $j = 1$.

		$\Delta_t = 0$			$\Delta_t = \frac{1}{78}$			$\Delta_t = \frac{1}{13}$		
		α			α			α		
$\bar{\gamma}$	b	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1
0	75	0.004	0.014	0.026	0.004	0.020	0.038	0.000	0.030	0.070
	150	0.004	0.016	0.036	0.002	0.020	0.060	0.004	0.026	0.064
	450	0.006	0.030	0.066	0.012	0.030	0.062	0.016	0.040	0.078
−0.01	75	0.062	0.220	0.344	0.052	0.216	0.362	0.048	0.198	0.358
	150	0.062	0.228	0.358	0.074	0.224	0.348	0.042	0.188	0.322
	450	0.120	0.230	0.350	0.110	0.230	0.364	0.080	0.194	0.306
−0.02	75	0.530	0.816	0.918	0.464	0.792	0.904	0.392	0.734	0.854
	150	0.550	0.798	0.912	0.502	0.790	0.888	0.418	0.654	0.808
	450	0.578	0.766	0.882	0.536	0.732	0.858	0.400	0.612	0.768
−0.04	75	1.000	1.000	1.000	1.000	1.000	1.000	0.998	1.000	1.000
	150	0.996	1.000	1.000	0.998	1.000	1.000	0.988	1.000	1.000
	450	1.000	1.000	1.000	0.998	1.000	1.000	0.986	0.998	1.000
−0.07	75	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	150	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	450	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 19
The table shows averages of $\hat{\bar{\gamma}} = \hat{\gamma}_- - \hat{\gamma}_+$ and averages of its t-statistics from estimating the linear regression of Eq. (61) with daily IV, RG^2, $RV^{5\text{ min}}$ and pre-averaging estimator C_a^n as the daily estimated volatility. The daily IV, RG^2, $RV^{5\text{ min}}$ and C_a^n are calculated by using samples $p_t = p_t^* + \epsilon_t$, where p_t^* is the true underlying efficient log-price generated from the Heston model of Eq. (61) and $\epsilon_t \sim N\left(0, (0.0005)^2\right)$ are i.i.d. market microstructure noises.

		IV			RG^2			$RV^{5\text{ min}}$			C_a^n		
		T			T			T			T		
ρ	Avg.	500	2000	3500	500	2000	3500	500	2000	3500	500	2000	3500
0.8	$\hat{\bar{\gamma}}$	−0.039	−0.037	−0.037	−0.056	−0.048	−0.048	−0.041	−0.039	−0.039	−0.040	−0.038	−0.038
	t-stat.	−8.944	−15.190	−18.981	−1.371	−2.037	−2.613	−2.867	−4.840	−6.125	−6.000	−9.949	−12.514
0.3	$\hat{\bar{\gamma}}$	−0.015	−0.014	−0.014	−0.026	−0.018	−0.018	−0.016	−0.014	−0.013	−0.015	−0.014	−0.014
	t-stat.	−3.520	−6.244	−8.031	−0.609	−0.758	−0.956	−1.102	−1.763	−2.203	−2.240	−3.889	−4.961
0	$\hat{\bar{\gamma}}$	0.000	0.000	0.000	−0.006	0.000	0.000	−0.001	0.000	0.000	0.000	0.000	0.000
	t-stat.	−0.024	−0.011	−0.006	−0.139	0.018	0.027	−0.054	0.052	0.033	−0.043	−0.025	−0.020
−0.3	$\hat{\bar{\gamma}}$	0.015	0.014	0.014	0.015	0.019	0.018	0.015	0.015	0.014	0.015	0.014	0.014
	t-stat.	3.477	6.255	8.097	0.339	0.802	0.948	1.024	1.855	2.320	2.211	3.787	4.856
−0.8	$\hat{\bar{\gamma}}$	0.040	0.037	0.037	0.050	0.050	0.049	0.043	0.040	0.039	0.041	0.037	0.037
	t-stat.	9.118	15.362	19.112	1.254	2.164	2.717	2.899	4.893	6.145	6.037	9.841	12.338

To know whether ρ can result in the conditional leverage effect, we first estimate the following linear regression:

$$\hat{\sigma}_t^2 = \alpha_D + \beta_{RD}\hat{\sigma}_{t-1}^2 + \gamma_+\left(r_{t-1}^{daily}\right)^2 1\left\{r_{t-1}^{daily} > 0\right\} + \gamma_-\left(r_{t-1}^{daily}\right)^2 1\left\{r_{t-1}^{daily} \le 0\right\} + \varepsilon_t \tag{61}$$

by using the daily IV, RG^2, $RV^{5\text{ min}}$ and C_a^n as $\hat{\sigma}_t^2$. The lengths of sample paths are $T = 500$, 2000 and 3500 days. The linear regression of (61) admits the form of the GJR model. Let $\hat{\gamma}_-$ and $\hat{\gamma}_+$ denote estimated values of $\hat{\gamma}_-$ and $\hat{\gamma}_+$. Table 19 shows average values of $\hat{\bar{\gamma}} = \hat{\gamma}_- - \hat{\gamma}_+$ and average values of the corresponding t-statistics ($H_0 : \bar{\gamma} = 0$) from the 1000 simulations. A positive and statistically significant $\hat{\bar{\gamma}}$ suggests that the conditional leverage effect may exist. From the table, it can be seen that as ρ is negative (positive), on average $\hat{\bar{\gamma}}$ is positive (negative). A lower (higher) value of ρ corresponds to a higher (lower) average value of $\hat{\bar{\gamma}}$. When $\rho \neq 0$, average magnitude of the t-statistics increases with the sample length, suggesting that evidence for $\hat{\bar{\gamma}} \neq 0$ becomes stronger as more samples are used in the linear regression estimation. The average magnitude of the t-statistics also varies substantially across different conditional volatility estimators. When $\rho \neq 0$, on average the more sophisticated conditional volatility estimators (IV and C_a^n) generate larger t-statistics than the less sophisticated conditional ones (RG^2 and $RV^{5\text{ min}}$). Finally, when $\rho = 0$, on average $\hat{\bar{\gamma}}$ is nearly zero and statistically insignificant. The results indicate that different values of ρ not only cause different comovements of instant changes of the efficient log price p_t^* and squared spot volatility v_t at the intraday level, but also result in the conditional leverage effect at the daily level.

Tables 20–22 show simulation results of the conditional leverage test for the Heston model when lengths of sample paths $T = 500$, 2000 and 3500 days. Again we set the lag length of daily return $j = 1$, the subsample sizes $b = 25$, 50 and 150 for $T = 500$ days, $b = 50$, 100 and 300 days for $T = 2000$ days and $b = 75$, 150 and 450 for $T = 3500$ days. Here we will only report results of $\rho > 0$ and $\rho = 0$, since in Table 19 we empirically show that $\rho < 0$ corresponds to the case of $\bar{\gamma} > 0$, which is in the interior of the null hypothesis.

From the tables, it can be seen that the rejection frequency increases with ρ. When $\rho > 0$, the rejection frequency also increases with the sample length T, no matter which conditional volatility estimator is used. The results are consistent with those shown in Table 19, in which a higher value of ρ on average implies a lower value of $\bar{\gamma}$. Also from Table 19, when $\rho = 0$, on average $\hat{\bar{\gamma}}$ is nearly equal to zero and statistically insignificant. It is similar as the case

Table 20
The table shows the rejection frequencies of the conditional leverage hypothesis test at three different levels of α for the samples generated with the Heston model of Eq. (60). The length of the sample path $T = 500$ days. Each scenario is simulated 1000 times. The observed log price is $p_t = p_t^* + \varepsilon_t^{noise}$, where $\varepsilon_t^{noise} \sim N\left(0, (0.0005)^2\right)$ are i.i.d. market microstructure noises. IV denotes the daily integrated variance. We use p_t to calculate daily RG^2, $RV^{5\text{ min}}$ and the pre-averaging estimator C_a^n and use them as the inputs for the test. For calculating the C_a^n, we set number of observations equal to 23,400 and the preaveraging parameter $\theta = 0.5$. We set lag length of daily return $j = 1$.

		IV			RG^2			$RV^{5\text{ min}}$			C_a^n		
		α			α			α			α		
ρ	b	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1
0.8	25	0.080	0.192	0.353	0.204	0.368	0.548	0.132	0.308	0.464	0.061	0.216	0.368
	50	0.060	0.165	0.332	0.173	0.261	0.425	0.101	0.253	0.380	0.069	0.164	0.341
	150	0.165	0.269	0.364	0.160	0.225	0.336	0.148	0.260	0.357	0.156	0.257	0.353
0.3	25	0.000	0.008	0.028	0.048	0.141	0.232	0.012	0.040	0.100	0.000	0.004	0.052
	50	0.000	0.012	0.045	0.027	0.084	0.168	0.004	0.024	0.096	0.001	0.025	0.061
	150	0.049	0.113	0.160	0.032	0.056	0.141	0.029	0.056	0.125	0.036	0.096	0.153
0	25	0.000	0.000	0.000	0.024	0.072	0.137	0.000	0.024	0.075	0.000	0.000	0.011
	50	0.000	0.000	0.005	0.016	0.041	0.100	0.000	0.021	0.064	0.000	0.000	0.016
	150	0.004	0.020	0.041	0.029	0.048	0.088	0.013	0.028	0.077	0.004	0.032	0.065

Table 21
The table shows the rejection frequencies of the conditional leverage hypothesis test at three different levels of α for the samples generated with the Heston model of Eq. (60). The length of the sample path $T = 2000$ days. Each scenario is simulated 1000 times. The observed log price is $p_t = p_t^* + \varepsilon_t^{noise}$, where $\varepsilon_t^{noise} \sim N\left(0, (0.0005)^2\right)$ are i.i.d. market microstructure noises. IV denotes the daily integrated variance. We use p_t to calculate daily RG^2, $RV^{5\text{ min}}$ and the pre-averaging estimator C_a^n and use them as the inputs for the test. For calculating the C_a^n, we set number of observations equal to 23,400 and the preaveraging parameter $\theta = 0.5$. We set lag length of daily return $j = 1$.

		IV			RG^2			$RV^{5\text{ min}}$			C_a^n		
		α			α			α			α		
ρ	b	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1
0.8	50	0.556	0.917	0.980	0.548	0.832	0.908	0.597	0.856	0.940	0.541	0.889	0.964
	100	0.627	0.900	0.981	0.517	0.764	0.873	0.573	0.833	0.920	0.577	0.888	0.953
	300	0.772	0.908	0.961	0.420	0.637	0.760	0.612	0.797	0.904	0.676	0.869	0.932
0.3	50	0.004	0.084	0.196	0.040	0.152	0.268	0.016	0.132	0.284	0.004	0.048	0.152
	100	0.041	0.137	0.272	0.029	0.112	0.220	0.029	0.132	0.280	0.013	0.097	0.236
	300	0.157	0.281	0.473	0.036	0.109	0.204	0.104	0.193	0.333	0.105	0.213	0.348
0	50	0.000	0.000	0.000	0.004	0.032	0.068	0.000	0.000	0.020	0.000	0.000	0.004
	100	0.000	0.000	0.002	0.004	0.033	0.061	0.000	0.005	0.020	0.000	0.005	0.005
	300	0.000	0.008	0.033	0.017	0.037	0.064	0.017	0.036	0.065	0.004	0.009	0.056

Table 22
The table shows the rejection frequencies of the conditional leverage hypothesis test at three different levels of α for the samples generated with the Heston model of Eq. (60). The length of the sample path $T = 3500$ days. Each scenario is simulated 1000 times. The observed log price is $p_t = p_t^* + \varepsilon_t^{noise}$, where $\varepsilon_t^{noise} \sim N\left(0, (0.0005)^2\right)$ are i.i.d. market microstructure noises. IV denotes the daily integrated variance. We use p_t to calculate daily RG^2, $RV^{5\text{ min}}$ and the pre-averaging estimator C_a^n and use them as the inputs for the test. For calculating the C_a^n, we set number of observations equal to 23,400 and the preaveraging parameter $\theta = 0.5$. We set lag length of daily return $j = 1$.

		IV			RG^2			$RV^{5\text{ min}}$			C_a^n		
		α			α			α			α		
ρ	b	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1
0.8	75	0.960	1.000	1.000	0.873	0.940	1.000	0.913	0.987	1.000	0.953	1.000	1.000
	150	0.960	1.000	1.000	0.820	0.920	0.967	0.853	0.987	1.000	0.933	1.000	1.000
	450	0.987	1.000	1.000	0.773	0.907	0.960	0.893	0.973	1.000	0.980	1.000	1.000
0.3	75	0.060	0.333	0.567	0.107	0.293	0.387	0.093	0.367	0.627	0.027	0.167	0.440
	150	0.153	0.447	0.693	0.133	0.267	0.347	0.073	0.367	0.660	0.080	0.307	0.580
	450	0.347	0.660	0.847	0.133	0.213	0.320	0.240	0.460	0.673	0.173	0.433	0.667
0	75	0.000	0.026	0.060	0.005	0.033	0.067	0.002	0.025	0.061	0.000	0.025	0.061
	150	0.005	0.025	0.064	0.000	0.027	0.047	0.004	0.026	0.063	0.004	0.027	0.064
	450	0.007	0.026	0.066	0.007	0.020	0.047	0.007	0.024	0.064	0.007	0.025	0.064

of the least favorable configuration. In this situation, rejection frequency is overall lower than the corresponding significant level except in a few cases when sample length $T = 500$. For performances of different conditional volatility estimators, when $\rho > 0$ and T is small ($T = 500$), in some cases using the less sophisticated conditional volatility estimator (RG^2 or $RV^{5\text{ min}}$) seems to have a slightly higher chance to reject the null than using the more sophisticated one (IV or C_a^n). However, when $\rho > 0$ and T becomes larger ($T = 2000$ and 3500), rejection frequency of using the more sophisticated conditional volatility estimator increases more than using the less sophisticated one. This phenomenon is similar as in the case of the GJR model. For the performance of the bias corrected test statistic, we focus on the cases of $T = 500$, 2000 and 3500 days with subsample size $b = 150$, 300 and 450 days and report their results in Table 23. We find that using the bias corrected test statistic can increase the rejection frequency as $\rho > 0$. When $\rho = 0$, the bias corrected test statistic can make the rejection frequency closer to the corresponding significant level than the test statistic without the bias correction.

We finally investigate how different data sampling frequencies affect performances of the test statistic. As volatilities are estimated with discretely sampled data, the estimation errors will exist even when there is no market microstructure noise. The estimation errors rise as data frequency decreases. However, when the market microstructure noise presents, using coarser sampling data may reduce the estimation error. It is hard to distinguish

Table 23
The table shows the rejection frequencies of using the bias corrected test statistic in the conditional leverage hypothesis test at three different levels of α for the samples generated with the Heston model of Eq. (60). The lengths of the sample paths are 500, 2000 and 3500 days. Each scenario is simulated 1000 times. The observed log price is $p_t = p_t^* + \varepsilon_t^{noise}$, where $\varepsilon_t^{noise} \sim N\left(0, (0.0005)^2\right)$ are i.i.d. market microstructure noises. IV denotes the daily integrated variance. We use p_t to calculate daily RG^2, $RV^{5\text{ min}}$ and the pre-averaging estimator C_a^n and use them as the inputs for the test. For calculating the daily pre-averaging estimator, we set number of observations equal to 23,400 and the preaveraging parameter $\theta = 0.5$. We set lag length of daily return $j = 1$.

	IV			RG^2			$RV^{5\text{ min}}$			C_a^n		
	$T = 500, b = 150$											
	α			α			α			α		
ρ	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1
0.8	0.164	0.268	0.360	0.160	0.225	0.332	0.147	0.261	0.344	0.156	0.252	0.348
0.3	0.048	0.111	0.153	0.022	0.046	0.126	0.028	0.056	0.120	0.036	0.090	0.147
0	0.004	0.020	0.041	0.018	0.025	0.072	0.012	0.020	0.067	0.003	0.023	0.044
	$T = 2000, b = 300$											
	α			α			α			α		
ρ	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1
0.8	0.763	0.907	0.956	0.420	0.628	0.748	0.608	0.792	0.904	0.675	0.860	0.931
0.3	0.147	0.273	0.464	0.036	0.096	0.201	0.100	0.183	0.328	0.104	0.207	0.343
0	0.000	0.008	0.032	0.017	0.023	0.057	0.016	0.022	0.056	0.004	0.008	0.031
	$T = 3500, b = 450$											
	α			α			α			α		
ρ	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1
0.8	1.000	1.000	1.000	0.883	0.947	1.000	0.953	1.000	1.000	1.000	1.000	1.000
0.3	0.387	0.690	0.883	0.143	0.243	0.340	0.261	0.482	0.693	0.203	0.463	0.707
0	0.009	0.042	0.075	0.007	0.030	0.053	0.008	0.040	0.070	0.008	0.040	0.072

Table 24
The table shows the rejection frequencies of the conditional leverage hypothesis test at three different levels of α for the samples generated with the Heston model of Eq. (60). The length of the sample path $T = 2000$ days. Each scenario is simulated 1000 times. The observed log price is the true underlying efficient log-price: $p_t = p_t^*$. We use p_t to calculate daily $RV^{1\text{ s}}$, $RV^{1\text{ min}}$, $RV^{5\text{ min}}$ and $RV^{30\text{ min}}$ as the inputs for the test. We set lag length of daily return $j = 1$.

		RV^{1sec}			RV^{1min}			$RV^{5\text{ min}}$			RV^{30min}		
		α			α			α			α		
ρ	b	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1
0.8	50	0.556	0.914	0.979	0.541	0.892	0.976	0.588	0.904	0.968	0.652	0.868	0.952
	100	0.612	0.907	0.979	0.596	0.901	0.965	0.573	0.869	0.960	0.605	0.825	0.928
	300	0.736	0.901	0.973	0.704	0.844	0.940	0.665	0.828	0.909	0.621	0.784	0.857
0.3	50	0.004	0.056	0.192	0.016	0.096	0.224	0.048	0.113	0.236	0.092	0.260	0.392
	100	0.020	0.129	0.288	0.044	0.141	0.264	0.041	0.156	0.245	0.088	0.220	0.377
	300	0.113	0.284	0.465	0.101	0.220	0.360	0.100	0.208	0.316	0.121	0.232	0.356
0	50	0.000	0.000	0.000	0.000	0.004	0.008	0.000	0.000	0.021	0.016	0.044	0.072
	100	0.000	0.000	0.005	0.004	0.004	0.008	0.000	0.000	0.021	0.004	0.040	0.081
	300	0.004	0.009	0.034	0.008	0.016	0.038	0.000	0.016	0.043	0.020	0.040	0.081

Table 25
The table shows the rejection frequencies of the conditional leverage hypothesis test at three different levels of α for the samples generated with the Heston model of Eq. (60). The length of the sample path $T = 3500$ days. Each scenario is simulated 1000 times. The observed log price is the true underlying efficient log-price: $p_t = p_t^*$. We use p_t to calculate daily $RV^{1\text{ s}}$, $RV^{1\text{ min}}$, $RV^{5\text{ min}}$ and $RV^{30\text{ min}}$ as the inputs for the test. We set lag length of daily return $j = 1$.

ρ	b	$RV^{1\text{ s}}$			$RV^{1\text{ min}}$			$RV^{5\text{ min}}$			$RV^{30\text{ min}}$		
		α			α			α			α		
		0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1
0.8	75	0.973	1.000	1.000	0.947	1.000	1.000	0.927	0.993	1.000	0.893	0.960	0.987
	150	0.973	1.000	1.000	0.973	0.993	1.000	0.907	0.980	0.993	0.860	0.967	0.980
	450	0.983	0.997	1.000	0.983	0.993	1.000	0.933	0.973	0.987	0.840	0.940	0.973
0.3	75	0.073	0.300	0.527	0.033	0.333	0.520	0.040	0.260	0.460	0.120	0.340	0.540
	150	0.107	0.387	0.687	0.080	0.413	0.607	0.087	0.313	0.513	0.093	0.347	0.527
	450	0.313	0.633	0.800	0.247	0.480	0.667	0.207	0.373	0.513	0.113	0.313	0.480
0	75	0.000	0.025	0.060	0.001	0.025	0.061	0.001	0.024	0.062	0.002	0.030	0.067
	150	0.005	0.026	0.064	0.005	0.027	0.063	0.004	0.027	0.062	0.006	0.027	0.066
	450	0.007	0.025	0.067	0.008	0.026	0.066	0.007	0.026	0.067	0.007	0.030	0.067

how large the estimation errors are due to the discretization errors or due to the market microstructure noise. We have shown how the test statistic performs when the microstructure noise exists in previous discussions. Here we focus on the case when there is *no* market microstructure noise. So the observed log price is the true underlying efficient log-price: $p_t = p_t^*$. The volatility estimators considered are the daily realized variances calculated with the observed log price p_t sampled from four different time intervals: 1 s, 1 min, 5 min and 30 min. Under the settings, different testing results will highlight the effects of using data sampled from different frequencies. We show results of $T = 2000$ and 3500 days in Tables 24 and 25. We can see that when $\rho > 0$, using volatilities estimated with lower frequency data (5-min and 30-min) results in a lower rejection frequency than using those estimated with higher frequency data (1-sec and 1-min). The rejection frequency also increases as the sample length T increases from 2000 to 3500. The results suggest that the estimation errors due to a lower sampling frequency and sample length T both affect performance of the test statistic.

References

Aït-Sahalia, Y., Fan, J., Li, Y., 2013. The leverage effect puzzle: Disentangling sources of bias at high frequency. J. Financ. Econ. 109, 224–249.
Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71, 579–625.
Andersen, T.G., Bollerslev, T., Christoffersen, P.F., Diebold, F.X., 2011. Financial Risk Measurement for Financial Risk Management. In: Handbook of the Economics of Finance.
Andrews, D.W.K., (1989) Stochastic Equicontinuity for Semiparametric Models, Unpublished manuscript, Cowles Foundation, Yale University.
Andrews, D.W.K., 1994. Asymptotics for semiparametric econometric models via stochastic equicontinuity. Econometrica 62, 43–72.
Andrews, D.W.K., 1995. Nonparametric Kernel estimation for semiparametric models. Econometric Theory 11, 560–596.
Andrews, D.W.K, Shi, X., 2013. Inference based on conditional moment inequalities. Econometrica 81, 609–666.
Bandi, F.M., Renò, R., 2012. Time-varying leverage effects. J. Econometrics 169, 94–113.
Barndorff-Nielsen, O.E., Shephard, N., 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. J. Roy. Statist. Soc. Ser. B 64, 253–280.
Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2009. Realized kernels in practice: Trades and quotes. Econom. J. 12, 1–32.
Bekaert, G., Wu, G., 2000. Asymmetric volatility and risk in equity markets. Rev. Financ. Stud. 13, 1–42.
Black, F., 1976. Studies of stock price volatility changes. In: Proceedings of the 1976 Meetings of the American Statistical Association. pp. 171–181.
Bollerslev, T., Litvinova, J., Tauchen, G., 2006. Leverage and volatility feedback effects in high-frequency data. J. Financ. Econom. 4, 353–384.
Bouchaud, J.P., Matacz, A., Potters, M., 2001. Leverage effect in financial markets: The retarded volatility model. Phys. Rev. Lett. 87, 22.
Campbell, J.Y., Hentschel, L., 1992. No news is good news: An asymmetric model of changing volatility in stock returns. J. Financ. Econ. 31, 281–318.
Christie, A.A., 1982. The stochastic behavior of common stock variances: Value, leverage and interest rate effects. J. Financ. Econ. 10, 407–432.
Corradi, V., Distaso, W., 2006. Semi-parametric comparison of stochastic volatility models using realized measures. Rev. Econom. Stud. 73 (3), 635–667.
Corradi, V., Distaso, W., Fernandes, M., 2012. International market links and volatility transmission. J. Econometrics 170, 117–141.
Corsi, F., Renò, R., 2012. Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. J. Bus. Econom. Statist. 30, 368–380.
Dahlhaus, R., 1997. Fitting time series models to nonstationary processes. Ann. Statist. 25, 1–37.
Dahlhaus, R., Subba Rao, S., 2006. Statistical inference for time-varying ARCH processes. Ann. Statist. 34, 1075–1114.
Davidson, R., Duclos, J.-Y., 2012. Testing for restricted stochastic dominance. Econometric Rev. 32, 84–125.
Dehling, H., Philipp, W., 2002. Empirical process techniques for dependent data. In: Dehling, H., Mikosch, T., Sørensen, M. (Eds.), Empirical Process Techniques for Dependent Data. Cambridge University Press.
de la Peña, V.H., 1999. A general class of exponential inequalities for martingales and ratios. Ann. Probab. 27, 537–564.
Dufour, J.-M., Garcia, R., Taamouti, A., 2012. Measuring high-frequency causality between returns, realized volatility, and implied volatility. J. Financ. Econom. 10, 124–163.
Engle, R.F., Ng, V.K., 1993. Measuring and testing the impact of news on volatility. J. Financ. 48, 1749–1778.
Figlewski, S., Wang, X., (2000) Is the ''Leverage Effect'' a Leverage Effect? Manuscript, NYU.
Folland, G.B., 1984. Real Analysis: Modern Techniques and their Applications. John Wiley, New York.
French, K.R., Schwert, G.W., Stambaugh, R.F., 1987. Expected stock returns and volatility. J. Financ. Econ. 19, 3–29.
Garman, M.B., Klass, M.J., 1980. On the estimation of security price volatilities from historical data. J. Bus. 53, 67–78.
Ghosh, A., Linton, O., (2013) Consistent Estimation of the Risk-Return Tradeoff in the Presence of Measurement Error, Working paper, LSE.
Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. J. Financ. 48, 1779–1801.
Gonçalves, S., Meddahi, N., 2009. Bootstrapping realized volatility. Econometrica 77, 283–306.
Hall, P., Heyde, C.C., 1980. Martingale Limit Theory and Its Aplication. Academic Press, New York.
Hautsch, N., Podolskij, M., 2013. Preaveraging-based estimation of quadratic variation in the presence of noise and jumps: Theory, implementation, and empirical evidence. J. Bus. Econom. Statist. 31, 165–183.
Hentschel, L., 1995. All in the family: Nesting symmetric and asymmetric garch models. J. Financ. Econ. 39, 71–104.
Heston, S.L., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev. Financ. Stud. 6, 327–343.
Jacod, J., (1994) Limit of Random Measures Associated with the Increments of a Brownian Semimartingale, Preprint number 120, Laboratoire de Probabilitiés, Université Pierre et Marie Curie, Paris.
Jacod, J., Protter, P., 1998. Asymptotic error distributions for the Euler method for stochastic differential equations. Ann. Probab. 26, 267–307.
Jacod, J., Li, Y., Mykland, P.A., Podolskij, M., Vetter, M., 2009. Microstructure noise in the continuous case: The pre-averaging approach. Stochastic Process. Appl. 119, 2249–2276.
Kristensen, D., 2010. Nonparametric filtering of the realized spot volatility: A kernel-based approach. Econometric Theory 26, 1469–4360.
Levy, H., 2006. Stochastic Dominance: Investment Decision Making under Uncertainty. second ed.. Springer, Berlin.
Linton, O., Maasoumi, E., Whang, Y.-J., 2005. Consistent testing for stochastic dominance under general sampling schemes. Rev. Econom. Stud. 72, 735–765.
Linton, O., Mammen, E., 2005. Estimating semiparametric ARCH(∞) models by Kernel smoothing methods. Econometrica 73, 771–836.
Linton, O., Song, K., Whang, Y.-J., 2010. An improved bootstrap test of stochastic dominance. J. Econometrics 154, 186–202.
Liu, L., Patton, A.J., Sheppard, K., 2013. Does Anything Beat 5-Minute RV? A Comparison of Realized Measures Across Multiple Asset Classes. In: Economics Working Papers Series, vol. 645. Department of Economics, University of Oxford.
Mikosch, T., Starica, C., 2000. Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process. Ann. Statist. 28, 1427–1451.
Mykland, P.A., Lan, Zhang, 2009. Inference for continuous semimartingales observed at high frequency. Econometrica 77 (5), 1403–1445.
Nelson, D.B., 1990. ARCH models as diffusion approximations. J. Econometrics 45, 7–38.
Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: A new approach. Econometrica 59, 347–370.
Parkinson, M., 1980. The extreme value method for estimating the variance of the rate of return. J. Bus. 53, 61–65.
Politis, D.N., Romano, J.P., 1994. Large sample confidence regions based on subsamples under minimal assumptions. Ann. Statist. 22, 2031–2050.
Politis, D.N., Romano, J.P., Wolf, M., 1999. Subsampling. Springer-Verlag, New York.
Rodriguez, M.J., Ruiz, E., 2012. Revisiting several popular GARCH models with leverage effect: Differences and similarities. J. Financ. Econom. 10, 637–668.
Wang, D.C., Mykland, P.A., 2014. The estimation of leverage effect with high frequency data. J. Amer. Statist. Assoc. 109, 197–215.