UNIVERSITY OF CAMBRIDGE
Faculty of Economics

Institute for New Economic Thinking

Cambridge-INET Institute

Cambridge-INET Working Paper Series No: 2016/26

Cambridge Working Paper Economics: 1671

A COUPLED COMPONENT GARCH MODEL FOR INTRADAY AND OVERNIGHT VOLATILITY

Oliver Linton (University of Cambridge)

Jianbin Wu (School of Management, Xiamen)

We propose a semi-parametric coupled component GARCH model for intraday and overnight volatility that allows the two periods to have different properties. To capture the very heavy tails of overnight returns, we adopt a dynamic conditional score model with t innovations. We propose a several step estimation procedure that captures the nonparametric slowly moving components by kernel estimation and the dynamic parameters by t maximum likelihood. We establish the consistency and asymptotic normality of our estimation procedures. We extend the modelling to the multivariate case. We apply our model to the study of the component stocks of the Dow Jones industrial average over the period 1991-2016. We show that actually overnight volatility has increased in importance during this period. In addition, our model provides better intraday volatility forecast since it takes account of the full dynamic consequences of the overnight shock and previous ones.

A coupled component GARCH model for intraday and overnight volatility*

Oliver Linton[†]
University of Cambridge

Jianbin Wu[‡]
Xiamen University

November 24, 2016

Abstract

We propose a semi-parametric coupled component GARCH model for intraday and overnight volatility that allows the two periods to have different properties. To capture the very heavy tails of overnight returns, we adopt a dynamic conditional score model with t innovations. We propose a several step estimation procedure that captures the nonparametric slowly moving components by kernel estimation and the dynamic parameters by t maximum likelihood. We establish the consistency and asymptotic normality of our estimation procedures. We extend the modelling to the multivariate case. We apply our model to the study of the component stocks of the Dow Jones industrial average over the period 1991-2016. We show that actually overnight volatility has increased in importance during this period. In addition, our model provides better intraday volatility forecast since it takes account of the full dynamic consequences of the overnight shock and previous ones.

1 Introduction

The balance between intraday and overnight returns is of considerable interest as it sheds light on many issues in finance: the efficient markets hypothesis, the calendar time versus trading time models, the process by which information is impacted into stock prices, the relative merits of auction

*We would like to thank Hanhai Tang and Chen Wang for helpful comments.

[†]Faculty of Economics, Austin Robinson Building, Sidgwick Avenue, Cambridge, CB3 9DD. Email: obl20@cam.ac.uk.Thanks to the Cambridge INET for financial support.

[‡]School of Management, Xiamen, China. Email:jianbin.wu@xmu.edu.cn

versus continuous trading, the effect of high frequency trading on market quality, and the globalization and connectedness of international markets. We propose a time series model for intraday and overnight returns that respects their temporal ordering and permits them to have different properties. In particular, we propose a volatility model for each return series that has a long run component that slowly evolves over time, and is treated nonparametrically, and a parametric dynamic volatility component that allows for short run deviations from the long run process, which depend on previous intraday and overnight shocks. We adopt a dynamic conditional score model, Harvey (2013) and Harvey and Luati (2014), that links the news impact curves of the innovations to the shock distributions, which we assume to be t-distributions with unknown degrees of freedom (which may differ between day and night). In practice, the overnight return distribution is more heavy tailed than the intraday return, and in fact very heavily tailed. Our model allows for a difference in tail thickness in the conditional distributions. The short run dynamic process allows for leverage effects and separates the overnight shock from the intraday shock. Our model extends Blanc, Chicheportiche, and Bouchaud (2014) who consider an asymmetric ARCH(∞) process with t shocks. We also introduce a multivariate model that allows for time varying correlations.

We apply our model to the study of the component stocks of the Dow Jones industrial average over the period 1991-2016, a period which saw several substantial institutional changes. There are several purposes for our application. First, many authors have argued that the introduction of computerized trading and the increased prevalence of high frequency trading strategies in the period post 2005 has lead to an increase in volatility, see Linton, O'Hara, and Zigrand (2013). A direct comparison of volatility before and after would be problematic here because of the Global Financial Crisis (GFC), which raised volatility during the same period that High Frequency Trading (HFT) was becoming more prevalent. However, this hypothesis would suggest that the ratio of intraday to overnight volatility should have increased during this period because trading is not taking place during the market close period. We would like to evaluate whether this has occurred. One could just compare the daily return variance from the intraday segment with the daily return variance from the overnight segment, as many studies such as French and Roll (1986) have done. However, this would ignore both fast and slow variation in volatility through business cycle and other causal factors. Also, overnight raw returns are very heavy tailed and so sample variances are not very accurate. We use our dynamic two component model that allows for both fast and slow dynamic components to volatility, as is now common practice. Our model also allows dynamic feedback between overnight and intraday volatility, which is of interest in itself. Our model generates heavy tails in observed returns and parameter estimates that are robust to this phenomenon. Our model therefore allows us to compare the long run components of volatility over this period without over

reliance on Gaussian-type theory. We show that for the Dow Jones stocks actually the long run component of overnight volatility has increased in importance during this period relative to the long run component of intraday volatility. We provide a formal test statistic that quantifies the strength of this effect. This seems to be hard to reconcile with the view that trading has increased volatility. We also document the short run dynamic processes. Notably, we find, unlike Blanc, Chicheportiche, and Bouchaud (2014), that overnight returns significantly affect future intraday volatility. We also find overnight return shocks to have t-distributions with degrees of freedom roughly equal to three, which emphasizes the potential fragility of Gaussian-based estimation routines that earlier work has been based on. We also estimate the multivariate model and document that there has been an upward trend in the long run component of contemporary overnight correlation between stocks as well as in the long run component of contemporary intraday correlation between stocks. However, the trend development for the overnight correlations started later than for intraday, and started happening only after 2005, whereas the intraday correlations appear to be slowly increasing more or less from the beginning of the period.

A second practical purpose for our model is to improve forecasts of intraday volatility or close to close volatility. Our model allows us to condition on the opening price to forecast intraday volatility or to update the close to close volatility forecast and also to take account of the full dynamic consequences of the overnight shock and previous ones. We compare forecast performance of our model with a procedure based only on close to close returns and find in most cases superior performance.

Since the seminal work of Engle (1982) and Bollerslev (1986), there is a large literature on GARCH models. Our work is closely related to the multiplicative GARCH literature, which decomposes the volatility dynamic into short-run and long-run components; see, e.g., Engle and Lee (1999), Engle and Rangel (2008), Hafner and Linton (2010), Rangel and Engle (2012), and Han and Kristensen (2015). We are also related to the dynamic conditional score GARCH models, proposed by Creal, Koopman, and Lucas (2012), Harvey (2013), and Harvey and Luati (2014), which allow for heavy tailed overnight innovations. Furthermore, our proposed time varying correlation GARCH model contributes to the literature on high-dimensional GARCH; see, e.g., Bollerslev (1990), Engle and Kroner (1995), Engle (2002), Engle, Shephard, and Sheppard (2007), and Francq and Zakoian (2014).

Our work also adds to the study of overnight returns. A large number of studies have found substantial differences between the overnight and intraday returns. Overnight returns are shown to be higher than intraday returns in e.g., Cooper, Cliff, and Gulen (2008), Kelly and Clark (2011), Berkman, Koch, Tuttle, and Zhang (2012), Lachance (2015), while lower in Harris (1986) and Aretz and Bartram (2015). Overnight returns are in general less volatile than intraday returns (Lockwood

and Linn, 1990, French and Roll, 1986, and Aretz and Bartram, 2015), but more leptokurtic (Ng and Masulis, 1995 and Blanc, Chicheportiche, and Bouchaud, 2014). Several theoretical models have been developed to explain these differences; see, e.g., Slezak (1994), and Hong and Wang (2000). Another main stream in the literature focuses on the predictability of overnight information on future stock returns or volatilities, including Lin, Engle, and Ito (1994), Gallo (2001), Branch and Ma (2006), Tsiakas (2008), Kang and Babbs (2012), Blanc, Chicheportiche, and Bouchaud (2014), and Fuertes, Kalotychou, and Todorovic (2015). We propose a time series model to better incorporate these features of overnight returns, and to reinvestigate these empirical questions. [1]

2 Model and Properties

We let r_t^D denote intraday returns and r_t^N denote overnight returns on day t. We take the ordering that night precedes day so that $r_t^D = \ln(P_t^C/P_t^O)$ and $r_t^N = \ln(P_t^O/P_{t-1}^C)$, where P_t^O denotes the opening price on day t and P_t^C denotes the closing price on day t. Our model allows intraday returns to depend on overnight returns with the same t, but overnight returns just depend on lagged variables. Suppose that

$$\begin{pmatrix} 1 & \delta \\ 0 & 1 \end{pmatrix} \begin{pmatrix} r_t^D \\ r_t^N \end{pmatrix} = \begin{pmatrix} \mu_D \\ \mu_N \end{pmatrix} + \Pi \begin{pmatrix} r_{t-1}^D \\ r_{t-1}^N \end{pmatrix} + \begin{pmatrix} u_t^D \\ u_t^N \end{pmatrix}, \tag{1}$$

where u_t^D and u_t^N are conditional mean zero shocks. Under the EMH, $\delta = 0$ and $\Pi = 0$, but we allow these coefficients to be nonzero to pick up small short run effects such as due to microstructure, Scholes and Williams (1977).

We further suppose that the error process has conditional heteroskedasticity, both long run and short run effects. Specifically, we suppose that

$$u_t = \begin{pmatrix} \exp(\lambda_t^D)\exp(\sigma^D(t/T)) & 0 \\ 0 & \exp(\lambda_t^N)\exp(\sigma^N(t/T)) \end{pmatrix} \begin{pmatrix} \varepsilon_t^D \\ \varepsilon_t^N \end{pmatrix}, \tag{2}$$

where: ε_t^D and ε_t^N are i.i.d. mean zero shocks from t distributions with v_D and v_N degrees of freedom, respectively, while $\sigma^D(\cdot)$ and $\sigma^N(\cdot)$ are unknown but smooth functions that will represent the slowly varying (long-run) scale of the process, and T is the number of observations. Suppose that for $j = D, N$,

$$\sigma^j(s) = \sum\nolimits_{i=1}^{\infty} \theta_i^j \psi_i^j(s), \quad s \in [0,1] \tag{3}$$

[1] Some other interesting works include, e.g., Longstaff (1995), Bollerslev, Li, and Todorov (2016), Hansen and Lunde (2005), Dai, Li, Liu, and Wang (2015),Lou, Polk, and Skouras (2015), and Fong and Martens (2002).

for some orthonormal basis $\{\psi_i^j(s)\}_{i=1}^{\infty}$ with $\int_0^1 \psi_i^j(s)\,ds = 0$ and

$$\int \psi_i^j(s)\,\psi_k^j(s)\,ds = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k. \end{cases}$$

We suppose that $\sigma^D(\cdot)$ and $\sigma^N(\cdot)$ integrate to zero to achieve identification. Note that we choose this different normalization from Hafner and Linton (2010) and consequently we do not have to restrict the parameters of the short run dynamic processes. In the following, j is always used to denote D, N without mentioning.

Regarding the short run dynamic part of (2), we adopt a dynamic conditional score approach, Creal, Koopman, and Lucas (2012) and Harvey and Luati (2014). Let $e_t^j = \exp(-\sigma^j(t/T))u_t^j$, and the conditional score function is defined as

$$m_t^j = \frac{(1+v_j)(e_t^j)^2}{v_j \exp(2\lambda_t^j) + (e_t^j)^2} - 1, \quad v_j > 0.$$

We suppose that λ_t^D and λ_t^N are linear combinations of past values of the shocks determined by the conditional score function

$$\begin{aligned} \lambda_t^D &= \omega_D(1-\beta_D) + \beta_D\lambda_{t-1}^D + \gamma_D m_{t-1}^D + \rho_D m_t^N \\ &\quad + \gamma_D^*(m_{t-1}^D + 1)\text{sign}(e_{t-1}^D) + \rho_D^*(m_t^N + 1)\text{sign}(e_t^N) \end{aligned} \tag{4}$$

$$\begin{aligned} \lambda_t^N &= \omega_N(1-\beta_N) + \beta_N\lambda_{t-1}^N + \gamma_N m_{t-1}^N + \rho_N m_{t-1}^D \\ &\quad + \rho_N^*(m_{t-1}^D + 1)\text{sign}(e_{t-1}^D) + \gamma_N^*(m_{t-1}^N + 1)\text{sign}(e_{t-1}^N). \end{aligned} \tag{5}$$

This gives two dynamic processes for the short run scale of the overnight and intraday return. We allow the overnight shock to affect the intraday scale through the parameter ρ_D, and we allow for leverage effects through the parameters $\gamma_D^*, \rho_D^*, \rho_N^*$, and γ_N^*.[2] Let

$$\phi = (\omega_D, \beta_D, \gamma_D, \gamma_D^*, \rho_D, \rho_D^*, v_D, \omega_N, \beta_N, \gamma_N, \gamma_N^*, \rho_N, \rho_N^*, v_N)^\intercal \in \mathbb{R}^{14}$$

be the finite dimensional parameters of interest and θ be the "parameters" in the functions $\sigma^D(\cdot)$ and $\sigma^N(\cdot)$, so that θ is infinite-dimensional.

Harvey (2013) argues that the quadratic innovations that feature in GARCH models naturally fit with the Gaussian distribution for the shock, but once one allows heavier tail distributions like the t-distribution, it is anomalous to focus on quadratic innovations, and indeed this focus leads to a lack of robustness because large shocks are fed substantially into the volatility update. He

[2]The shock variable m_t^j can be expressed as $m_t^j = (v_j+1)b_t^j - 1$, where b_t^j has a beta distribution, $beta\,(1/2, v_j/2)$.

argues it is more natural to link the shock to volatility to the distribution of the rescaled return shock, which in the case of the t distribution has the advantage that large shocks are automatically downweighted, and in such a way driven by the shape of the error distribution. This type of argument is similar to the argument in limited dependent variable models such as binary choice where a linear function of covariates is connected to the observed outcome by a link function determined by the distributional assumption. The DCS model has the incidental advantage that there are analytic expressions for moments, autocorrelation functions, multistep forecasts, and their mean square errors. Our semiparametric model is also tractable in a number of dimensions. For example, we may obtain the dynamic intraday value at risk conditional on overnight returns and past information as follows

$$VaR_t^D(\alpha) = \mu_t^D + s_t^D t_\alpha(v_D),$$

$$\mu_t^D = \mu_D - \delta r_t^N - \Pi_{11} r_{t-1}^D - \Pi_{12} r_{t-1}^N$$
$$s_t^D = \exp(\lambda_t^D)\exp(\sigma^D(t/T)),$$

where $t_\alpha(v)$ is the α quantile of the t-distribution with degrees of freedom v. Likewise, we can find the overnight value at risk conditional on the closing price.

3 Estimation

We first outline our estimation strategy. Taking unconditional expectations of the absolute errors we have for $j = N, D$,

$$E\left(\left|u_t^j\right|\right) = E(\left|\varepsilon_t^j\right|)E\left(\exp\left(\lambda_t^j\right)\right)\exp(\sigma^j(t/T)) = c^j(\phi) \times \exp(\sigma^j(t/T)),$$

where c^j is a constant that depends in a complicated way on the parameter vector ϕ. Therefore, we can estimate $\sigma^N(s), \sigma^D(s)$ as follows. Suppose that we know δ, μ, Π (in practice these can be replaced by root-T consistent estimators). Although we define the sieve expansion of σ, we use kernel technology to estimate the nonparametric part. Let $K(u)$ be a kernel with support $[-1, 1]$ and h a bandwidth, and let $K_h(.) = K(./h)/h$. Then let

$$\widetilde{\sigma}^j(s) = \log\left(\frac{1}{T}\sum_{t=1}^{T} K_h(s - t/T)\left|u_t^j\right|\right), \tag{6}$$

for any $s \in (0, 1)$. In fact, we employ a boundary modification for $s \in [0, h] \cup [1 - h, 1]$, whereby K is replaced by a boundary kernel, which is a function of two arguments $K(u, c)$, where the parameter

c controls the support of the kernel; thus left boundary kernel $K(u,c)$ with $c = s/h$ has support $[-1,c]$ and satisfies $\int_{-1}^{c} K(u,c)du = 1$, $\int_{-1}^{c} uK(u,c)du = 0$, and $\int_{-1}^{c} u^2K(u,c)du < \infty$. Similarly for the right boundary. The purpose of the boundary modification is to ensure that the bias property holds throughout $[0,1]$. For identification, we rescale $\widetilde{\sigma}^j(t/T)$ as

$$\widetilde{\sigma}^j(t/T) = \widetilde{\sigma}^j(t/T) - \frac{1}{T}\sum_{t=1}^{T}\widetilde{\sigma}^j(t/T). \tag{7}$$

Let $\widetilde{e}_t^N = \exp(-\widetilde{\sigma}^N(t/T))u_t^N$ and $\widetilde{e}_t^D = \exp(-\widetilde{\sigma}^D(t/T))u_t^D$, and let $\widetilde{\theta}$ denote $\{\widetilde{\sigma}^j(s),\ s \in [0,1],\ j = N, D\}$. Define the global log-likelihood function for ϕ (apart from an unnecessary constant and conditional on the estimated values of θ)

$$\begin{aligned} l_T(\phi;\widetilde{\theta}) &= \frac{1}{T}\sum_{t=1}^{T}\left(l_t^N(\phi;\widetilde{\theta}) + l_t^D(\phi;\widetilde{\theta})\right), \\ l_t^j(\phi;\widetilde{\theta}) &= -\lambda_t^j(\phi;\widetilde{\theta}) - \frac{v_j+1}{2}\ln\left(1 + \frac{(\widetilde{e}_t^j)^2}{v_j\exp(2\lambda_t^j(\phi;\widetilde{\theta}))}\right) + \ln\Gamma\left(\frac{v_j+1}{2}\right) - \frac{1}{2}\ln v_j - \ln\Gamma\left(\frac{v_j}{2}\right), \end{aligned} \tag{8}$$

where $\lambda_t^j(\phi;\widetilde{\theta})$ are defined in (4) and (5). For practical purposes, $\lambda_{1|0}^j$ may be set equal to the unconditional mean, $\lambda_{1|0}^j = \omega_j$. We estimate ϕ by maximizing $l_T(\phi;\widetilde{\theta})$ with respect to ϕ. Let $\widetilde{\phi}$ denote these estimates.

Given estimates of ϕ and the preliminary estimates of $\sigma^D(\cdot), \sigma^N(\cdot)$, we calculate

$$\widetilde{\eta}_t^N = \exp(-\widetilde{\lambda}_t^N)u_t^N\ ;\ \widetilde{\eta}_t^D = \exp(-\widetilde{\lambda}_t^D)u_t^D,$$

where $\widetilde{\lambda}_t^j = \lambda_t^j(\widetilde{\phi};\widetilde{\theta})$. We then update the estimates of $\sigma^D(\cdot), \sigma^N(\cdot)$ using the local likelihood function in Severini and Wong (1992) given $\widetilde{\eta}_t^j$ and $\widetilde{v}_j$, i.e., minimize the objective function

$$\widetilde{L}_T^j(\gamma;\widetilde{\lambda}^j, s) = \frac{1}{T}\sum_{t=1}^{T}K_h(s - t/T)\left[\gamma + \frac{\widetilde{v}_j+1}{2}\ln\left(1 + \frac{(\widetilde{\eta}_t^j\exp(-\gamma))^2}{\widetilde{v}_j}\right)\right] \tag{9}$$

with respect to $\gamma \in \mathbb{R}$, for $j = D, N$ separately, where $\widetilde{\lambda}^j = (\widetilde{\lambda}_1^j, \ldots, \widetilde{\lambda}_T^j)^\intercal$. Likewise here we use a boundary kernel for $s \in [0,h] \cup [1-h,1]$. In practice we use Newton-Raphson iterations making use of the derivatives of the objective functions, which are given in (19).

To summarize, the estimation algorithm is as follows.

Algorithm

Step 1. Estimate δ, μ^j, Π by least squares and $\widetilde{\sigma}^j(u)$, $u \in [0,1]$, $j = N, D$ from (6) and (7)

Step 2. Estimate ϕ by optimizing $l_T(\phi; \widetilde{\theta})$with respect to ϕ (by Newton-Raphson) to give $\widetilde{\phi}$.

Step 3. Given the initial estimates $\widetilde{\theta}$ and $\widetilde{\phi}$, we replace λ_t^j with $\widetilde{\lambda}_t^j = \lambda_t^j(\widetilde{\phi}; \widetilde{\theta})$. Then let $\widehat{\sigma}^j(s)$ optimize $\widetilde{L}_T^j(\sigma^j(s); \widetilde{\lambda}, s)$ with respect to $\sigma^j(s)$. For identification, we rescale $\widehat{\sigma}^j(t/T) = \widehat{\sigma}^j(t/T) - \frac{1}{T}\sum_{t=1}^T \widehat{\sigma}^j(t/T)$.

Step 4. Repeat Steps 2-3 to update $\widehat{\theta}$ and $\widehat{\phi}$ until convergence. We define convergence in terms of the distance measure

$$\Delta_r = \sum_{j=D,N} \int \left[\widehat{\sigma}^{j,[r]}(u) - \widehat{\sigma}^{j,[r-1]}(u)\right]^2 du + \left(\widehat{\phi}^{[r]} - \widehat{\phi}^{[r-1]}\right)^{\intercal} \left(\widehat{\phi}^{[r]} - \widehat{\phi}^{[r-1]}\right),$$

that is, we stop when $\Delta_r \leq \epsilon$ for some prespecified small ϵ.

4 Large Sample Properties of Estimators

In this section we give the asymptotic distribution theory of the estimators considered above. Let $h_t^j = \lambda_t^j + \sigma^j(t/T)$, and let:

$$A_t = \begin{bmatrix} 1 & a_t^{DN} \\ 0 & 1 \end{bmatrix}, \quad B_{t-1} = \begin{bmatrix} \left(\beta_D + a_{t-1}^{DD}\right) & 0 \\ a_{t-1}^{ND} & \left(\beta_N + a_{t-1}^{NN}\right) \end{bmatrix},$$

$$\begin{aligned} a_{t-1}^{DD} &= -2\left(\gamma_D + \gamma_D^* \text{sign}(u_{t-1}^D)\right)(v_D + 1)\, b_{t-1}^D \left(1 - b_{t-1}^D\right) \\ a_t^{DN} &= -2\left(\rho_D + \rho_D^* \text{sign}(u_t^N)\right)(v_N + 1)\, b_t^N \left(1 - b_t^N\right) \\ a_{t-1}^{NN} &= -2\left(\gamma_N + \gamma_N^* \text{sign}(u_{t-1}^N)\right)(v_N + 1)\, b_{t-1}^N \left(1 - b_{t-1}^N\right) \\ a_{t-1}^{ND} &= -2\left(\rho_N + \rho_N^* \text{sign}(u_{t-1}^D)\right)(v_D + 1)\, b_{t-1}^D \left(1 - b_{t-1}^D\right) \end{aligned}$$

$$b_t^D = \frac{(e_t^D)^2}{v_D \exp(2\lambda_t^D) + (e_t^D)^2} \quad ; \quad b_t^N = \frac{(e_t^N)^2}{v_D \exp(2\lambda_t^N) + (e_t^N)^2}.$$

We use the maximum row sum matrix norm, $\|\cdot\|_\infty$, defined by

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

Assumptions A

1. $\|E(A_t \otimes A_t)\|_\infty < \infty, \|EB_tEA_t\|_\infty < 1, \|E(B_{t-1}A_{t-1} \otimes B_{t-1}A_{t-1})\|_\infty < \|EB_tEA_t\|_\infty$ *,and the top-Lyapunov exponent of the sequence of* A_tB_{t-1} *is strictly negative. The top Lyapunov exponent is defined as Theorem 4.26 of* Douc, Moulines, and Stoffer (2014).

2. $|\beta_j| < 1$.

3. h_t^j *starts from infinite past. The parameter* ϕ_0 *is an interior point of* $\Phi \subset \mathbb{R}^{14}$, *where* Φ *is the parameter space of* ϕ_0.

4. *The function* σ^j *is twice continuously differentiable on* $[0, 1]$.

5. $E|u_t^j|^{2+\delta} < \infty$ *for some* $\delta > 0$.

6. *The kernel function* K *is bounded, symmetric about zero with compact support, that is* $K(v) = 0$ for all $|v| > C_1$ with some $C_1 < \infty$. *Moreover, it is Lipschitz, that is* $|K(v) - K(v')| \leq L|v - v'|$ *for some* $L < \infty$ *and all* $v, v' \in \mathbb{R}$.

7. $h(T) \to 0$,*as* $T \to \infty$ *such that* $T^{1/2-\delta}h \to \infty$ *for some small* $\delta > 0$.

Assumptions A3-A7 are used to derive the properties of $\widetilde{\sigma}^j(s)$, in line with Vogt and Linton (2014) and Vogt et al. (2012). But we only require that $E|u_t^j|^{2+\delta} < \infty$, since we use $\widetilde{\sigma}^j(s) = \log\left(\frac{1}{T}\sum_{t=1}^T K_h(s - t/T)\left|u_t^j\right|\right)$. This is in line with the fact that the fourth-order moment of overnight returns often does not exist. The mixing condition in Vogt and Linton (2014) is replaced by Assumption A2, because of our tight model structure. Assumption A1 is required to derive the stationarity of score functions, where $\|E(A_t \otimes A_t)\|_\infty < \infty$ can be verified easily, since b_t^N in A_t follows a beta distribution.

The first result gives the uniform convergence rate of the initial estimator $\widetilde{\sigma}^j(s)$. The proof mainly follows Theorem 3 in Vogt and Linton (2014).

Lemma 1 *Suppose that Assumptions A2-A7 hold. Then,*

$$\sup_{s\in[0,1]} \left|\widetilde{\sigma}^j(s) - \sigma_0^j(s)\right| = O_p\left(h^2 + \sqrt{\frac{\log T}{Th}}\right).$$

Proof. See 8.2.1.

Theorem 1 *Suppose that Assumptions A1-A4 hold. Then, for each* k *and* i, *for* $k \in \{1, \ldots, \infty\}$ *and* $i \in \{1, \ldots, 14\}$, *we have*

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^T E\left[\frac{\partial l_t(\theta_0, \phi_0)}{\partial \theta_k}\frac{\partial l_t(\theta_0, \phi_0)}{\partial \phi_i}\right] = 0.$$

The proof of Theorem 1 is provided in section 8.2.2. Theorem 1 implies that the score function with respect to θ and the score function with respect to ϕ are orthogonal. Therefore, the particular form of the variance in Theorem 2 follows.

Let

$$I(\phi_0) = E\left[\frac{\partial l_t(\theta_0,\phi_0)}{\partial \phi}\frac{\partial l_t(\theta_0,\phi_0)}{\partial \phi^{\intercal}}\right].$$

Theorem 2 *Suppose that Assumptions A1-A7 hold. Then*

$$\sqrt{T}\left(\widehat{\phi}-\phi_0\right) \xrightarrow{D} N\left(0, I(\phi_0)^{-1}\right).$$

Theorem 3 *Suppose that Assumptions A1-A7 hold. Then*

$$\sqrt{Th}\begin{pmatrix} \widehat{\sigma}^D(t/T) \\ \widehat{\sigma}^N(t/T) \end{pmatrix} - \begin{pmatrix} \sigma_0^D(t/T) \\ \sigma_0^N(t/T) \end{pmatrix} \Longrightarrow N\left(0, ||K||_2^2 \begin{pmatrix} \frac{(v_D+3)}{2v_D} & 0 \\ 0 & \frac{(v_N+3)}{2v_N} \end{pmatrix}\right). \tag{10}$$

Theorem 2 and Theorem 3 give the consistency and asymptotic normality of $\widehat{\phi}$ and $\widehat{\sigma}^j(s)$. The form of the limiting variance in (10) is consistent with the known Fisher information for the estimation of scale parameter of a t-distribution with known location and degrees of freedom (these quantities are estimated at a faster rate), which makes this part of the procedure also efficient in the sense considered in Tibshirani (1984).

The proofs of Theorem 2 and 3 are provided in the Appendix. The information matrix, $I(\phi_0)$, can be computed explicitly, as in Appendix 2. We can conduct inference with Theorem 2 and Theorem 3 using plug-in estimates of the unknown quantities.

5 A Multivariate model

We next consider an extension to a multivariate model. We keep a similar structure to the univariate model except that we allow the slowly moving component to be matrix valued. Suppose that

$$r_t = \begin{pmatrix} r_t^D \\ r_t^N \end{pmatrix};\ \mu = \begin{pmatrix} \mu_D \\ \mu_N \end{pmatrix},$$

where r_t^D, r_t^D are $n \times 1$ vectors containing all the intraday and overnight returns respectively, and let

$$\mathbf{D}r_t = \mu + \Pi r_{t-1} + u_t,$$

where u_t^D and u_t^N are mean zero shocks, while

$$\mathbf{D} = \begin{pmatrix} I_n & \text{diag}\,(\Delta) \\ 0 & I_n \end{pmatrix}; \quad \Pi = \begin{pmatrix} \text{diag}(\Pi_{11}) & \text{diag}(\Pi_{12}) \\ \text{diag}(\Pi_{21}) & \text{diag}(\Pi_{22}) \end{pmatrix},$$

and $\Delta, \Pi_{11}, \Pi_{12}, \Pi_{21}$,and Π_{22} are $n \times 1$ vectors. We further suppose that

$$u_t = \begin{pmatrix} \Sigma^D(\frac{t}{T})^{\frac{1}{2}}\text{diag}\left(\exp(\lambda_t^D)\right) & 0 \\ 0 & \Sigma^N(\frac{t}{T})^{\frac{1}{2}}\text{diag}\left(\exp(\lambda_t^N)\right) \end{pmatrix} \begin{pmatrix} \varepsilon_t^D \\ \varepsilon_t^N \end{pmatrix},$$

where: ε_{it}^j is i.i.d. shocks from univariate t distributions with v_{ij} degrees of freedom, while λ_t^j are $n \times 1$ vectors.

We assume that $\Sigma^D(.)$ and $\Sigma^N(.)$ are smooth matrix functions but are otherwise unknown. We can write these covariance matrices in terms of the correlation matrices and the variances as follows

$$\Sigma^j(s) = \text{diag}\left(\exp(\sigma^j(s))\right) R^j(s)\text{diag}\left(\exp(\sigma^j(s))\right), \tag{11}$$

with $\text{diag}\,(\exp(\sigma^j(s)))$ being the volatility matrix and $R^j(s)$ being the correlation matrix. For identification, we still assume $\int_0^1 \sigma_i^j(s)ds = 0$, for $i \in \{1, \ldots n\}$.

As with the univariate model, define $e_t^j = \text{diag}\left(\exp(\lambda_t^j)\right)\varepsilon_t^j$, and suppose that

$$m_{it}^j = \frac{(1+v_{ij})(e_{it}^j)^2}{v_{ij}\exp(2\lambda_{it}^j) + (e_{it}^j)^2} - 1,$$

$$\begin{aligned} \lambda_{it}^D &= \omega_{iD}(1-\beta_{iD}) + \beta_{iD}\lambda_{it-1}^D + \gamma_{iD}m_{it-1}^D + \rho_{iD}m_{it}^N \\ &\quad + \gamma_{iD}^*(m_{it-1}^D + 1)\text{sign}(u_{it-1}^D) + \rho_{iD}^*(m_{it}^N + 1)\text{sign}(u_{it}^N) \\ \lambda_{it}^N &= \omega_{iN}(1-\beta_{iN}) + \beta_{iN}\lambda_{it-1}^N + \gamma_{iN}m_{it-1}^N + \rho_{iN}m_{it-1}^D \\ &\quad + \rho_{iN}^*(m_{it-1}^D + 1)\text{sign}(u_{it-1}^D) + \gamma_{iN}^*(m_{it-1}^N + 1)\text{sign}(u_{it-1}^N). \end{aligned}$$

For each i define the parameter vector $\phi_i = (\omega_{iD}, \beta_{iD}, \gamma_{iD}, \gamma_{iD}^*, \rho_{iD}, \rho_{iD}^*, v_{iD}, \omega_{iN}, \beta_{iN}, \gamma_{iN}, \gamma_{iN}^*, \rho_{iN}, \rho_{iN}^*, v_{iN})^\intercal \in \mathbb{R}^{14}$ and let $\phi = (\phi_1^\intercal, \ldots, \phi_n^\intercal)^\intercal$ denote all the dynamic parameters.

Define ι_i the vector with the ith element 1 and all others 0, so that $\varepsilon_{it}^j = \iota_i^\intercal\text{diag}\left(\exp(-\lambda_t^j)\right)\left(\Sigma^j(\frac{t}{T})\right)^{-1/2} u_t^j$. The global log-likelihood function is

$$l_T(\phi; \Sigma(.)) = \frac{1}{T}\sum_{t=1}^{T}\left(l_t^N + l_t^D\right)$$

$$l_t^j(\phi;\Sigma^j(.)) = \sum_{i=1}^{n}\left(-\lambda_t^j - \frac{v_{ij}+1}{2}\ln\left(1+\frac{(\iota_i^\intercal \mathrm{diag}\left(\exp(-\lambda_t^j-\sigma^j(t/T))\right)\left(\Sigma^j\left(\frac{t}{T}\right)\right)^{-1/2}u_t^j)^2}{v_{ij}}\right)\right)$$
$$-\frac{1}{2}\log\left|\Sigma^j\left(\frac{t}{T}\right)\right| + \sum_{i=1}^{n}\left(\ln\Gamma\left(\frac{v_{ij}+1}{2}\right) - \frac{1}{2}\ln v_{ij} - \ln\Gamma\left(\frac{v_{ij}}{2}\right)\right).$$

We first define an initial estimator for $\Sigma^j(\frac{t}{T})$ and then obtain an estimator of ϕ, and then we update them. Suppose that we know Δ, Π and μ. To give an estimator of $\Sigma^j(\frac{t}{T})$ robust to heavy tails, we estimate the volatility parameter

$$\widetilde{\sigma}_i^j(s) = \log\left(\frac{1}{T}\sum_{t=1}^{T}K_h(s-t/T)\left|u_{it}^j\right|\right).$$

Supposing that the heavy tails issue is less severe in the estimation of correlation, which seems reasonable; we estimate the correlation parameter by standard procedures

$$\widetilde{R}_{ik}^j(s) = \frac{\sum_{t=1}^{T}K_h(s-t/T)u_{ik}^j u_{ik}^j}{\sqrt{\sum_{t=1}^{T}K_h(s-\frac{t}{T})u_{it}^j u_{it}^j \sum_{t=1}^{T}K_h(s-\frac{t}{T})u_{kt}^j u_{kt}^j}} \tag{12}$$

for $s \in (0,1)$, and boundary modification as before. For identification, we rescale $\widetilde{\sigma}^j(t/T)$ as

$$\widetilde{\sigma}_i^j(t/T) = \widetilde{\sigma}_i^j(t/T) - \frac{1}{T}\sum_{t=1}^{T}\widetilde{\sigma}_i^j(t/T),$$

and compute

$$\widetilde{\Sigma}^j(s) = \mathrm{diag}\left(\exp(\widetilde{\sigma}^j(s))\right)\widetilde{R}^j(s)\mathrm{diag}\left(\exp(\widetilde{\sigma}^j(s))\right). \tag{13}$$

Letting $\widetilde{e}_t^j = \widetilde{\Sigma}^j(\frac{t}{T})^{-1/2}u_t^j$, we obtain $\widetilde{\phi}_i$ by maximizing the univariate log-likelihood function of $\widetilde{e}_{it}^j$ in (8) for each i.

To update the estimator for each $\Sigma^j(\frac{t}{T})$, denote $\Theta = \Sigma^{j-1/2}$. We first obtain $\widehat{\Theta}$ with the local likelihood function given $\widetilde{\lambda}_t^j$ and $\widetilde{v}_j$,i.e., maximize the local objective function

$$L_T^j(\Theta;\widetilde{\lambda},s) = \frac{1}{T}\sum_{t=1}^{T}K_h(s-t/T)\left[\log|\Theta| - \sum_{i=1}^{n}\left(\frac{\widetilde{v}_{ij}+1}{2}\ln\left(1+\frac{(\iota_i^\intercal \mathrm{diag}\left(\exp(-\widetilde{\lambda}_t^j)\right)\Theta u_t^j)^2}{\widetilde{v}_{ij}}\right)\right)\right]$$

with respect to vech(Θ), and let $\widehat{\Sigma}^j(t/T) = \widehat{\Theta}^{-2}$. The derivatives of the objective function are given in (29) and (30). Then we rescale $\widehat{\Sigma}^j(s)$ with the same procedure of $\widetilde{\Sigma}^j(s)$. Likewise, define

$\widehat{e}_t^j = \widehat{\Sigma}^j(s)^{-1/2}u_t^j$ and obtain $\widehat{\phi}_i$ by maximizing the univariate log-likelihood function of $\widehat{e}_{it}^j$. One can iterate this procedure by updating Θ with the local likelihood using the new $\widehat{\phi}_i$ and $\widehat{\lambda}^j$ and so on.

Our multivariate model can be considered as a GARCH model with a slowly moving correlation matrix. Assuming diagonality on the short run component λ_t^j enables us to estimate the model easily and fast. Especially, the computation time of the initial estimator is only of order n, with n being the number of assets considered; it is thus feasible even with quite large n. The extension to models with non-diagonal short run components is possible, but only feasible with small n.

6 Application

We investigate 28 components of the Dow Jones industrial average index over the period 1991-11-12 to 2016-04-13. The 28 stocks are: MMM, AXP, AAPL.O, BA, CAT, CVX, CSCO.O, KO, DD, XOM, GE, HD, IBM, INTC.O, JNJ, JPM, MCD, MRK, MSFT.O, NKE, PFE, PG, TRV, UNH, UTX, VZ, WMT, DIS. GS and V are excluded since they did not officially go public until 1999 and 2008, respectively. The data is obtained from Thompson Reuters Eikon, and has been adjusted for corporate actions. We define overnight returns as the log price change between the close of one trading day to the opening of the next trading day. We do not incorporate weekend and holiday effects into our model, since they are not the focus of this paper, and our model is already rather complicated in terms of both model specification and estimation. In addition, although the weekend effect is documented by e.g. French (1980) and Rogalski (1984), and further supported by Cho, Linton, and Whang (2007) with a stochastic dominance approach, many studies suggest the disappearance of the weekend effect, e.g. Mehdian and Perry (2001) and Steeley (2001). Especially, Sullivan, Timmermann, and White (2001) claim that calendar effects are the result of data-snooping.

6.1 Overnight Returns

Many studies find significant higher overnight returns, e.g. Cooper, Cliff, and Gulen (2008) and Berkman, Koch, Tuttle, and Zhang (2012). Cooper, Cliff, and Gulen (2008) even suggest US equity premium is solely due to overnight returns during the research period from 1993 to 2006. To investigate this, Fig. 1 plots the cumulative returns for these 28 stocks . For AAPL, CAT, CSCO, HD, INTC, JPM, PFE, UTX and WMT, positive cumulative returns indeed mainly come from overnight periods, but for MMM, KO, DD,XOM, JNJ, MCD, MRK, NKE, PG, TRV, VZ and DIS, positive cumulative returns mainly come from intraday periods. There is no clear dominance of positive overnight returns from this figure.

Furthermore, Berkman, Koch, Tuttle, and Zhang (2012) find significant positive mean overnight returns of +10 basis points per day, along with -7 basis points for the intraday returns from 3000 largest U.S. stocks. With the same procedure of Berkman, Koch, Tuttle, and Zhang (2012), we first compute the cross-sectional mean(or median) returns for each day, then compute the time series mean and the standard derivation of these cross-sectional mean(or median). The mean intraday return is 0.0218% with the standard deviation 0.0096, while the mean overnight return is 0.0164% with the standard deviation 0.0060. The mean of the cross-sectional median for intraday return is 0.0084% with the standard deviation 0.0092, and its overnight counterpart is 0.0113% with the standard deviation 0.0054. Still, we do not observe this significant overnight anomaly from the 28 Dow Jones stocks.

Table 1 gives the summary statistics for intraday and overnight returns. Compared with intraday returns, overnight returns exhibit more negative skewness and leptokurtosis. More specific, 10 of those 28 stocks have negative intraday skewness, while 26 of 28 stocks have negative overnight skewness. The largest sample kurtosis for overnight returns is extremely high, 884.7702, suggesting the non-existence of the population kurtosis.

6.2 Univariate Model Estimates

We multiply returns by 100 to give more readable coefficients. Table 2 reports the estimates and their robust standard errors in the mean equations. Π_{ij} refers to the element of the ith row jth column in the coefficient matrix Π. For the prediction of intraday returns, 12 of the 28 stocks have significant Π_{11} which are all negative, 7 of 28 stocks have significant δ which are all positive. This suggests that both overnight and intraday returns tend to have reversal effect in their subsequent intraday return, in line with Branch and Ma, 2006 and Berkman et al., 2012. However, we do not find clear patterns for predicting overnight returns. The constant terms, μ_D and μ_N, are positive for most Dow Jones stocks.

Table 3 gives the estimates of the variance equations. The parameters β_D and β_N are significantly different from 1, ρ_D, γ_D, ρ_N and γ_N are positive and significant. In addition, we find significant leverage effects, negative and significant ρ_D^*, γ_D^*, ρ_N^* and γ_N^*, suggesting higher volatility after negative returns.

We also concern about the difference between the overnight and intraday parameters. Table 4 reports the Wald tests with the null hypothesis that the intraday and overnight parameters are equal within each stock. The parameter, ω_D, determining the unconditional short-run scale, is significantly larger than ω_N. The overnight degree-of-freedom parameter is around 3, significantly smaller than

the intraday one, around 8. Both are in line with the descriptive statistics in Table 1 and previous studies that overnight returns are more leptokurtic but less volatile. With other pairs of intraday and overnight parameters, $\beta_j, \gamma, \rho_j, \gamma_j^*, \rho_j^*$, the null hypothesis is seldom rejected. However, the joint null hypothesis, $(\beta_D, \gamma, \rho_D, \gamma_D^*, \rho_D^*) = (\beta_N, \gamma, \rho_N, \gamma_N^*, \rho_N^*)$, is rejected by many stocks. It is noteworthy that the null hypothesis $H_0 : \gamma_N = \rho_D$ is not rejected by our data, which is inconsistent with Blanc, Chicheportiche, and Bouchaud (2014). They suggest that past overnight returns affect weakly the future intraday volatilities, except for the very next one, but impact substantially future overnight volatilities. This inconsistency is probably because the dynamic conditional score model shrinks the impact of extreme overnight observations. After this shrinkage, overnight innovations become closer to the intraday innovations.

Fig. 2 displays the ratios of the overnight to intraday variances. The stocks all exhibit upward trends over the 24-year period considered here, and many of them had peaks around August 2011, corresponding to the August 2011 stock markets fall event. Fig. 3 depicts the long-run intraday and overnight components, $\sigma^D(t/T)$ and $\sigma^N(t/T)$, and their 95% point-wise confidence intervals. Most stocks arrived at their first peaks around 10 March 2000, corresponding to the Dot-com bubble event, while some arrived around September 2011, the 9-11 attacks. The intraday components reached the second peaks during the financial crisis in September 2008, while overnight components still went up and reached their highest points during the 2010 Flash crash. Roughly speaking, the intraday components were larger than the overnight ones before the first peaks, but smaller after financial crisis in September 2008. But remember that the long-run components are constructed with rescaling $\int_0^1 \sigma(s)ds = 0$. In general, the intraday variances are still larger.

We test the constancy of the ratio of overnight to intraday variance through the null hypothesis

$$H_0 : \exp\left(\sigma_0^N(\cdot)\right) = \rho \exp\left(\sigma_0^D(\cdot)\right) \tag{14}$$

for some $\rho \in \mathbb{R}_+$ versus the general alternative. By Theorem 3 and the delta method, $\exp(\widehat{\sigma}^D(s))$ and $\exp(\widehat{\sigma}^N(s))$ converge jointly to a normal distribution, and are asymptotically mutually independent. It follows that

$$\widehat{\tau}(s) = \sqrt{Th}\left(\exp\left(\widehat{\sigma}^N(s)\right) - \widehat{\rho}\exp\left(\widehat{\sigma}^D(s)\right)\right) \Longrightarrow N\left(0, \rho^2 V_s^D + V_s^N\right)$$

$$\widehat{\rho} = \frac{1}{T}\sum_{t=1}^{T} \frac{\exp\left(\widehat{\sigma}^N(t/T)\right)}{\exp\left(\widehat{\sigma}^D(t/T)\right)},$$

where $V_s^j = \exp\left(2\sigma_0^j(s)\right)\frac{(v_j+3)}{2v_j}||K||_2^2$, for $j = D, N$.

Fig. 4 displays the test statistics $\widehat{\tau}(s)$ and the 95 % point-wise confidence intervals for $s \in [0, 1]$. Consistent with the results above, the equal ratio null hypothesis is mostly rejected before the first peaks (in 2000) and after the second peaks (in 2010).

The Ljung-Box tests on the absolute and the squared standardised residuals are used to verify whether the coupled component GARCH model is adequate to capture the heteroscedasticity, shown in Table 6. With the absolute form, strong heteroskedasticity exists in both intraday and overnight returns, but disappears in the standardised residuals, saying that our model captures the heteroscedasticity well. On the other hand, we are sometimes unable to detect the heteroscedasticity in overnight returns with squared values. In general, the use of the absolute form is more robust when the distribution is heavy tailed.

Fig. 6 displays the quantile-quantile(Q-Q) plots of the intraday innovations, comparing with the student t distribution with $\hat{\nu_D}$ degrees of freedom. The points in the Q-Q plots approximately lie on a line, saying that the intraday innovations closely approximate the t distribution. Fig. 7 displays the Q-Q plots of the overnight innovations. Many stocks have several outliers in the lower left corners. Our model only partly captures the negative skewness and leptokurtosis of overnight innovations.

We also want to compare our coupled component GARCH model with its one component version for the open to close return to see the improvement in volatility forecast from using overnight returns. We construct 10 rolling windows, each containing 5652 in-sample and 50 out-of-sample observations. In each rolling window, the parameters in the short-run variances are estimated with the in-sample data once and stay the same during the one-step out-of-sample forecast. In the one-step ahead forecast of the long-run covariance matrices, the single-side weight function is used. For instance, to forecast the long-run covariance matrix of period τ ($s = \tau/T$), we set the two-side weight function $K_h(s - t/T) = 0$, for $t >= \tau$, and then rescale $K_h(s - t/T)$ to get a sum of 1. Table 5 reports Giacomini and White (2006) model pair-wise comparison tests with the out-of-sample quasi Gaussian and student t log-likelihood loss functions. For most stocks, the coupled component GARCH model dominates the one component model. Some dominances are statistically significant. We omit the comparison for overnight variance forecast between the one component and the coupled component model, since it is not plausible to estimate a GARCH model with overnight returns alone.

As a robustness check, we investigate the ratio of VIX to the Rogers and Satchell (1991)(RS) volatility. The idea is that the VIX measures one month ahead volatility, total volatility including presumably intraday and overnight, whereas the aggregated RS volatility only includes intraday volatility. Therefore, the ratio reflects the variability of intraday to overnight to some extent, although it is quite noisy. Fig. 5 presents: (1)the RS volatility on daily Dow Jones stocks, (2)one month ahead volatility from RS, $sqrt(\sum_{i=1}^{2} 2(volrs_{t+i}.^2))$, (3)VIX, and (4) the ratio of VIX to the one month ahead

RS volatility. The reported RS volatility is the average RS volatility across the 28 stocks. The ratio of VIX to the one month ahead RS volatility shows a upward trend during the research period, in line with our previous finding that the ratio of overnight to intraday volatility is increasing.

6.3 Multivariate Model

Fig. 9 and Fig. 8 present the long-run correlations between intraday returns and between overnight returns, respectively. Each subplot presents the 27 time series of long-run correlations between that stock and the rest of stocks. The correlations exhibited an obvious upward trend during our research period of 1998-2014, and were typically high in 2008 financial crisis.

Fig. 10 displays the eigenvalues of the dynamic covariance matrices, as well as their proportions (the eigenvalues divided by the sum of eigenvalues). The dynamic of eigenvalues reinforces the previous remark that the stock markets had high risk in the 11 September attacks and in the 2008 financial crisis. The largest eigenvalue represents a strong common component, saying that a large proportion of the market financial risk can be explained by one single factor. More specific, the largest eigenvalue proportion increased substantially between 1998 and 2016. The second and third largest eigenvalues still counted for a considerable amount of proportion in the volatile period from 2000 to 2002, but became rather insignificant in the volatile period from 2008 to 2011. The largest intraday eigenvalue proportion reached the peak in 2008, while the largest overnight one remained high until 2011. Remarkably, the largest eigenvalue explained nearly 50% of intraday risk in the 2008 financial crisis, and 70% of overnight risk in the August 2011 stock markets fall period.

Table 7 provides the estimates of the multivariate model. Compared to the univariate models, the average β_D decreases to 0.8796 from 0.9515, and the average β_N decreases to 0.8926 from 0.9553. Together with the increase of γ_j and ρ_j, it says that more weights are given to information in the most recent days by considering correlations.

One concern is that our initial correlation estimator is based on the Pearson product moment correlation. This Pearson estimator may perform poorly due to the heavy tails of overnight innovations. So we now investigate the estimates by using a robust correlation estimator in the initial step. More specific, we first compute the pairwise Kendall tau

$$\hat{\tau}_{k,l}(s) = \frac{\sum_{i=1}^{T}\sum_{j=i}^{T-1} K_h(s-i/T)K_h(s-j/T)\left(I\left\{(u_{i,k}-u_{j,k})(u_{i,l}-u_{j,l})>0\right\} - I\left\{(u_{i,k}-u_{j,k})(u_{i,l}-u_{j,l})<0\right\}\right)}{\sum_{i=1}^{T}\sum_{j=i}^{T-1} K_h(s-i/T)K_h(s-j/T)\left(I\left\{(u_{i,k}-u_{j,k})(u_{i,l}-u_{j,l})>0\right\} + I\left\{(u_{i,k}-u_{j,k})(u_{i,l}-u_{j,l})<0\right\}\right)}$$

Applying the relation between Kendall tau and the linear correlation coefficient for the elliptical

distribution suggested by Lindskog, Mcneil, and Schmock (2003) and Battey and Linton (2014), we obtain the robust linear correlation estimator,

$$\hat{\rho}_{k,l}(s) = \sin\left(\frac{\pi}{2}\hat{\tau}_{k,l}(s)\right).$$

In some cases, the matrix of pairwise correlations must be adjusted to ensure that the resulting matrix is positive definite, although we did not encounter this problem here.

Fig. 3 plots the largest eigenvalue proportions of the estimated covariance matrices to see the difference of using robust (in black) and using non-robust (in red) correlation estimators in the initial step. We use solid lines for the initial estimators, and dash lines for the updated estimators. The updated estimators are obtained in the final estimation step, the one we report in previous parts. Despite the large difference of the initial estimators, especially for the overnight returns, the updated estimators are roughly the same. Like the eigenvalues, the updated covariances themselves also remain unchanged for different initial estimators. To save space, we omit the plot of covariances.

7 Conclusion

We have introduced a new coupled component GARCH model for intraday and overnight volatility. This model is able to capture the heavy tails of overnight returns. For each component, we further specify a non-parametric long run smoothly evolving component with a parametric short term fluctuates. The large sample properties of the estimators are provided for the univariate model.

The empirical results show that the ratio of overnight to intraday volatility has increased during previous 20 years when accounting for slowly changing and rapidly changing components. This is contrary to what is often argued with regard to the change in market structure and the predatory practices of certain traders. The information in overnight returns is valuable for updating the forecast of the close to close volatility. In the multivariate model we found that (slowly moving) correlations between assets have increased during our sample period

References

Aretz, Kevin and Söhnke M Bartram (2015), "Making money while you sleep? anomalies in international day and night returns." *Available at SSRN 2670841*.

Battey, Heather and Oliver Linton (2014), "Nonparametric estimation of multivariate elliptic densities via finite mixture sieves." *Journal of Multivariate Analysis*, 123, 43–67.

Berkman, Henk, Paul D Koch, Laura Tuttle, and Ying Jenny Zhang (2012), "Paying attention: overnight returns and the hidden cost of buying at the open." *Journal of Financial and Quantitative Analysis*, 47, 715–741.

Blanc, Pierre, Rémy Chicheportiche, and Jean-Philippe Bouchaud (2014), "The fine structure of volatility feedback ii: Overnight and intra-day effects." *Physica A: Statistical Mechanics and its Applications*, 402, 58–75.

Bollerslev, Tim (1986), "Generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics*, 31, 307–327.

Bollerslev, Tim (1990), "Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model." *The Review of Economics and Statistics*, 72, 498–505.

Bollerslev, Tim, Sophia Zhengzi Li, and Viktor Todorov (2016), "Roughing up beta: Continuous versus discontinuous betas and the cross section of expected stock returns." *Journal of Financial Economics*, 120, 464–490.

Branch, Ben S and Aixin Ma (2006), "The overnight return, one more anomaly." Working paper. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=937997.

Carrasco, Marine and Xiaohong Chen (2002), "Mixing and moment properties of various garch and stochastic volatility models." *Econometric Theory*, 18, 17–39.

Cho, Young-Hyun, Oliver Linton, and Yoon-Jae Whang (2007), "Are there monday effects in stock returns: A stochastic dominance approach." *Journal of Empirical Finance*, 14, 736–755.

Cooper, Michael J, Michael T Cliff, and Huseyin Gulen (2008), "Return differences between trading and non-trading hours: Like night and day." Working paper. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1004081.

Creal, Drew, Siem Jan Koopman, and André Lucas (2012), "A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations." *Journal of Business & Economic Statistics*.

Dai, Min, Peifan Li, Hong Liu, and Yajun Wang (2015), "Portfolio choice with market closure and implications for liquidity premia." *Management Science*, 62, 368–386.

Douc, Randal, Eric Moulines, and David Stoffer (2014), *Nonlinear time series: Theory, methods and applications with R examples*. CRC Press.

Engle, Robert (2002), "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models." *Journal of Business & Economic Statistics*, 20, 339–350.

Engle, Robert F (1982), "Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation." *Econometrica: Journal of the Econometric Society*, 987–1007.

Engle, Robert F and Kenneth F Kroner (1995), "Multivariate simultaneous generalized arch." *Econometric theory*, 11, 122–150.

Engle, Robert F and Gary Lee (1999), "A long-run and short-run component model of stock return volatility." *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive WJ Granger*, 475–497.

Engle, Robert F and Jose Gonzalo Rangel (2008), "The spline-garch model for low-frequency volatility and its global macroeconomic causes." *Review of Financial Studies*, 21, 1187–1222.

Engle, Robert F, Neil Shephard, and Kevin Sheppard (2007), "Fitting and testing vast dimensional time-varying covariance models." Working paper. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1293629.

Fong, Kingsley and Martin Martens (2002), "Overnight futures trading: now even australia and us have common trading hours." *Journal of International Financial Markets, Institutions and Money*, 12, 167–182.

Francq, Christian and Jean-Michel Zakoian (2014), "Estimating multivariate garch and stochastic correlation models equation by equation." Working paper. Available at: http://perso.univ-lille3.fr/ cfrancq/Christian-Francq/statistics-econometrics-papers/EbEE.pdf.

French, Kenneth R (1980), "Stock returns and the weekend effect." *Journal of Financial Economics*, 8, 55–69.

French, Kenneth R and Richard Roll (1986), "Stock return variances: The arrival of information and the reaction of traders." *Journal of Financial Economics*, 17, 5–26.

Fuertes, Ana-Maria, Elena Kalotychou, and Natasa Todorovic (2015), "Daily volume, intraday and overnight returns for volatility prediction: profitability or accuracy?" *Review of Quantitative Finance and Accounting*, 45, 251–278.

Gallo, Giampiero M (2001), "Modelling the impact of overnight surprises on intra-daily volatility." *Australian Economic Papers*, 40, 567–580.

Giacomini, Raffaella and Halbert White (2006), "Tests of conditional predictive ability." *Econometrica*, 74, 1545–1578.

Hafner, Christian M and Oliver Linton (2010), "Efficient estimation of a multivariate multiplicative volatility model." *Journal of Econometrics*, 159, 55–73.

Han, Heejoon and Dennis Kristensen (2015), "Semiparametric multiplicative garch-x model: Adopting economic variables to explain volatility." Working paper. Available at: http://creates.au.dk/fileadmin/site_files/filer_oekonomi/subsites/creates/Diverse_2015/SoFiE_2015/Papers/47_Semiparametric_Multiplicative_GARCH-X_Model_-_Adopting_Economic_Variables_To_Explain_Volatility.pdf.

Hansen, Peter Reinhard and Asger Lunde (2005), "A realized variance for the whole day based on intermittent high-frequency data." *Journal of Financial Econometrics*, 3, 525–554.

Harris, Lawrence (1986), "A transaction data study of weekly and intradaily patterns in stock returns." *Journal of Financial Economics*, 16, 99–117.

Harvey, Andrew and Alessandra Luati (2014), "Filtering with heavy tails." *Journal of the American Statistical Association*, 109, 1112–1122.

Harvey, Andrew C (2013), *Dynamic models for volatility and heavy tails: with applications to financial and economic time series*, volume 52. Cambridge University Press.

Hong, Harrison and Jiang Wang (2000), "Trading and returns under periodic market closures." *The Journal of Finance*, 55, 297–354.

Kang, Long and Simon H Babbs (2012), "Modeling overnight and daytime returns using a multivariate generalized autoregressive conditional heteroskedasticity copula model." *The Journal of Risk*, 14, 35–63.

Kelly, Michael A and Steven P Clark (2011), "Returns in trading versus non-trading hours: The difference is day and night." *Journal of Asset Management*, 12, 132–145.

Kristensen, Dennis (2009), "Uniform convergence rates of kernel estimators with heterogeneous dependent data." *Econometric Theory*, 25, 1433–1445.

Lachance, Marie-Eve (2015), "Night trading: Lower risk but higher returns?" Working paper. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2633476.

Lin, Wen-Ling, Robert F Engle, and Takatoshi Ito (1994), "Do bulls and bears move across borders? international transmission of stock returns and volatility." *Review of Financial Studies*, 7, 507–538.

Lindskog, Filip, Alexander Mcneil, and Uwe Schmock (2003), "Kendalls tau for elliptical distributions." In *Credit Risk*, 149–156, Springer.

Linton, Oliver, Maureen O'Hara, and Jean-Pierre Zigrand (2013), "The regulatory challenge of high-frequency markets." 207–230, Risk Books, London.

Lockwood, Larry J and Scott C Linn (1990), "An examination of stock market return volatility during overnight and intraday periods, 1964–1989." *The Journal of Finance*, 45, 591–601.

Longstaff, Francis A (1995), "How much can marketability affect security values?" *The Journal of Finance*, 50, 1767–1774.

Lou, Dong, Christopher Polk, and Spyros Skouras (2015), "A tug of war: Overnight versus intraday expected returns." Working paper. Available at: http://personal.lse.ac.uk/loud/ Overnight-Mom.pdf.

Mehdian, Seyed and Mark J Perry (2001), "The reversal of the monday effect: new evidence from us equity markets." *Journal of Business Finance & Accounting*, 28, 1043–1065.

Ng, Victor and Ronald W Masulis (1995), "Overnight and daytime stock return dynamics on the london stock exchange." *Journal of Business and Economic Statistics*, 13, 365–378.

Rangel, José Gonzalo and Robert F Engle (2012), "The factor–spline–garch model for high and low frequency correlations." *Journal of Business & Economic Statistics*, 30, 109–124.

Rogalski, Richard J (1984), "New findings regarding day-of-the-week returns over trading and non-trading periods: a note." *The Journal of Finance*, 39, 1603–1614.

Rogers, L Christopher G and Stephen E Satchell (1991), "Estimating variance from high, low and closing prices." *The Annals of Applied Probability*, 504–512.

Scholes, Myron and Joseph Williams (1977), "Estimating betas from nonsynchronous data." *Journal of Financial Economics*, 5, 309–327.

Severini, Thomas A and Wing Hung Wong (1992), "Profile likelihood and conditionally parametric models." *The Annals of Statistics*, 1768–1802.

Slezak, Steve L (1994), "A theory of the dynamics of security returns around market closures." *The Journal of Finance*, 49, 1163–1211.

Steeley, James M (2001), "A note on information seasonality and the disappearance of the weekend effect in the uk stock market." *Journal of Banking & Finance*, 25, 1941–1956.

Sullivan, Ryan, Allan Timmermann, and Halbert White (2001), "Dangers of data mining: The case of calendar effects in stock returns." *Journal of Econometrics*, 105, 249–286.

Tasaki, Hiroyuki (2009), "Convergence rates of approximate sums of riemann integrals." *Journal of Approximation Theory*, 161, 477–490.

Tibshirani, RJ (1984), *Local likelihood estimation*. Ph.D. thesis, Stanford University.

Tsiakas, Ilias (2008), "Overnight information and stochastic volatility: A study of european and us stock exchanges." *Journal of Banking & Finance*, 32, 251–268.

Vogt, Michael and Oliver Linton (2014), "Nonparametric estimation of a periodic sequence in the presence of a smooth trend." *Biometrika*, 101, 121–140.

Vogt, Michael et al. (2012), "Nonparametric regression for locally stationary time series." *The Annals of Statistics*, 40, 2601–2633.

8 Appendix

8.1 Appendix 1: some further properties

We consider the prediction problem. Note that λ_t^D is not the conditional predictor given the day before, but given the updated information set, and the linear predictor of λ_t^D given only $\mathcal{F}_{t-1}$ is

$$E\left[\lambda_t^D|\mathcal{F}_{t-1}\right] = \omega_D(1-\beta_D) + \beta_D\lambda_{t-1}^D + \gamma_D m_{t-1}^D + \gamma_D^*(m_{t-1}^D+1)\text{sign}(e_t^D) \neq \lambda_t^D.$$

The expectation of $\exp(2\lambda_t^D)$ given $\mathcal{F}_{t-1}$ is

$$E\left[\exp(2\lambda_t^D)|\mathcal{F}_{t-1}\right] = \Lambda_t E\left[\exp(2\rho_D m_t^N + 2\rho_D^*(m_t^N+1)\text{sign}(e_t^N))|\mathcal{F}_{t-1}\right],$$

where $\Lambda_t = \exp(2\omega_D(1-\beta_D) + 2\beta_D\lambda_{t-1}^D + 2\gamma_D m_{t-1}^D + 2\gamma_D^*(m_{t-1}^D+1)\text{sign}(e_{t-1}^D)$. We can express $E\left[\exp(2\rho_D m_t^N + 2\rho_D^*(m_t^N+1)\text{sign}(e_t^N))|\mathcal{F}_{t-1}\right]$ as

$$\frac{1}{2}\exp(-2\rho_D)\left\{E\left[\exp((2\rho_D+2\rho_D^*)(v_N+1)b_t^N) + \exp((2\rho_D-2\rho_D^*)(v_N+1)b_t^N|\mathcal{F}_{t-1}\right]\right\}.$$

Since b_t^N follows a beta $(1/2, v_N/2)$ distribution,

$$E\left[\exp((2\rho_D+2\rho_D^*)(v_N+1)b_t^N)\right] = {}_1F_1(1/2, 1/2+v_N/2, (2\rho_D+2\rho_D^*)(v_N+1)),$$

where ${}_1F_1$ is the Kummer's function

$${}_1F_1(\alpha,\beta,c) = 1 + \sum_{k=0}^{\infty}\left(\prod_{r=1}^{k-1}\frac{\alpha+r}{\beta+r}\right)\frac{c^k}{k!}, \quad \alpha,\beta>0.$$

Hence, we have

$$\begin{aligned} E\left[\exp(2\lambda_t^D)|\mathcal{F}_{t-1}\right] &= \frac{1}{2}\exp(-2\rho_D)\Lambda_t\,{}_1F_1(1/2, 1/2+v_N/2, (2\rho_D+2\rho_D^*)(v_N+1)) \\ &\quad + \frac{1}{2}\exp(-2\rho_D)\Lambda_t\,{}_1F_1(1/2, 1/2+v_N/2, (2\rho_D-2\rho_D^*)(v_N+1)). \end{aligned}$$

Note that we have 2 different information sets, $\mathcal{F}_{t-1}$ and $\mathcal{F}_{t-1}\cup\{r_t^N\}$, where $\mathcal{F}_{t-1}$ is sigma field generated by $\{r_{t-1}^D, r_{t-1}^N, r_{t-2}^D, r_{t-2}^N, \ldots\}$ and $\mathcal{F}_{t-1}\cup\{r_t^N\}$ is sigma field generated by $\{r_t^N, r_{t-1}^D, r_{t-1}^N, r_{t-2}^D, r_{t-2}^N, \ldots\}$.

	intraday				overnight			
	mean	std.dev.	skew	kurt	mean	std.dev.	skew	kurt
MMM	0.0003	0.0127	0.0040	7.1842	0.0000	0.0073	-0.6030	20.3340
AXP	0.0003	0.0192	-0.0473	9.7455	0.0001	0.0109	-1.3220	30.8063
AAPL	-0.0004	0.0242	0.1903	6.3090	0.0010	0.0180	-7.2410	288.9583
BA	0.0002	0.0159	-0.0091	6.3725	0.0001	0.0103	-2.4955	59.8270
CAT	-0.0001	0.0174	0.0190	5.7938	0.0005	0.0110	-0.8610	18.7038
CVX	0.0001	0.0137	0.0810	10.6472	0.0002	0.0074	-0.9206	13.1831
CSCO	-0.0001	0.0229	0.0227	10.6339	0.0008	0.0141	-0.6753	22.9465
KO	0.0006	0.0126	0.0474	8.3962	-0.0003	0.0071	-0.3592	12.6505
DD	0.0003	0.0157	0.0273	7.1028	-0.0002	0.0091	-0.3597	16.1463
XOM	0.0004	0.0131	0.0861	10.8546	-0.0002	0.0071	-0.9341	14.7103
GE	-0.0001	0.0157	-0.0019	11.2698	0.0004	0.0101	0.1864	30.3115
HD	0.0001	0.0172	0.2999	6.9232	0.0004	0.0108	-3.0134	81.4299
IBM	0.0003	0.0150	0.0487	7.3880	-0.0001	0.0100	-0.7826	39.9142
INTC	0.0001	0.0204	0.1913	7.2270	0.0005	0.0141	-2.6159	51.1115
JNJ	0.0003	0.0120	0.0789	6.4480	0.0000	0.0071	-3.0637	79.0347
JPM	0.0000	0.0210	0.4353	14.1952	0.0003	0.0124	0.1249	19.7644
MCD	0.0005	0.0137	-0.0941	8.1253	-0.0001	0.0083	-0.4635	15.1683
MRK	0.0004	0.0149	-0.0381	7.6085	-0.0002	0.0096	-6.2306	172.5119
MSFT	0.0003	0.0171	0.1528	5.6204	0.0003	0.0109	-0.5826	32.3796
NKE	0.0006	0.0177	0.1389	9.5287	-0.0001	0.0104	-2.0513	50.5907
PFE	-0.0001	0.0152	-0.0349	5.6494	0.0004	0.0097	-2.0215	40.9346
PG	0.0009	0.0124	-0.0109	9.0587	-0.0006	0.0085	-18.5373	884.7702
TRV	0.0002	0.0163	-0.1116	16.1477	0.0001	0.0087	-1.7067	65.3085
UNH	0.0003	0.0203	-0.1451	13.8989	0.0004	0.0117	-2.9957	64.2616
UTX	0.0001	0.0145	-0.2758	9.4068	0.0004	0.0084	-1.6782	38.7102
VZ	0.0001	0.0142	0.4456	7.6871	0.0000	0.0079	-0.4672	15.4566
WMT	-0.0000	0.0147	0.1238	7.9677	0.0003	0.0084	-0.7025	16.3198
DIS	0.0004	0.0161	0.1634	7.1087	0.0000	0.0110	-1.0887	47.8462

This table gives the summary statistics for the intraday and overnight returns.

Table 1: Summary statistics for intraday and overnight returns

	δ	μ_D	μ_N	Π_{11}	Π_{12}	Π_{21}	Π_{22}
MMM	-0.0094	0.0304	0.0035	-0.0201	-0.0079	-0.0245	-0.0186
	(0.0331)	(0.0163)	(0.0093)	(0.0180)	(0.0297)	(0.0108)	(0.0201)
AXP	-0.0482	0.0302	0.0112	-0.0530	0.0055	-0.0112	-0.0463
	(0.0449)	(0.0245)	(0.0139)	(0.0218)	(0.0444)	(0.0141)	(0.0259)
AAPL	0.0473	-0.0468	0.1009	-0.0609	0.0794	-0.0137	0.0113
	(0.0433)	(0.0310)	(0.0233)	(0.0172)	(0.0262)	(0.0188)	(0.0220)
BA	-0.0284	0.0149	0.0120	-0.0048	0.0162	-0.0100	0.0239
	(0.0313)	(0.0203)	(0.0132)	(0.0197)	(0.0413)	(0.0144)	(0.0184)
CAT	0.0142	-0.0095	0.0531	-0.0013	-0.0159	0.0155	-0.0132
	(0.0342)	(0.0223)	(0.0140)	(0.0185)	(0.0260)	(0.0113)	(0.0179)
CVX	-0.0847	0.0130	0.0163	-0.0515	-0.0449	-0.0059	-0.0316
	(0.0435)	(0.0176)	(0.0095)	(0.0211)	(0.0452)	(0.0137)	(0.0226)
CSCO	0.0319	-0.0120	0.0832	-0.0615	0.0081	0.0268	-0.0205
	(0.0322)	(0.0294)	(0.0181)	(0.0203)	(0.0303)	(0.0105)	(0.0181)
KO	0.0516	0.0571	-0.0250	-0.0188	0.0354	-0.0134	0.0537
	(0.0372)	(0.0162)	(0.0091)	(0.0193)	(0.0353)	(0.0123)	(0.0171)
DD	0.0627	0.0343	-0.0174	0.0025	-0.0230	-0.0094	-0.0083
	(0.0370)	(0.0201)	(0.0117)	(0.0204)	(0.0315)	(0.0114)	(0.0172)
XOM	-0.0482	0.0475	-0.0148	-0.0898	-0.0144	-0.0287	-0.0222
	(0.0465)	(0.0166)	(0.0090)	(0.0220)	(0.0447)	(0.0147)	(0.0236)
GE	0.1066	-0.0125	0.0382	-0.0276	0.0234	-0.0020	0.0159
	(0.0493)	(0.0201)	(0.0130)	(0.0283)	(0.0429)	(0.0173)	(0.0289)
HD	0.0306	0.0143	0.0362	0.0217	-0.0563	0.0208	-0.0054
	(0.0414)	(0.0220)	(0.0139)	(0.0191)	(0.0288)	(0.0120)	(0.0197)
IBM	-0.0067	0.0365	-0.0061	-0.0375	0.0609	0.0070	-0.0477
	(0.0297)	(0.0191)	(0.0127)	(0.0177)	(0.0254)	(0.0112)	(0.0190)
INTC	0.0335	0.0034	0.0475	-0.0542	0.0742	0.0061	-0.0515
	(0.0277)	(0.0260)	(0.0180)	(0.0180)	(0.0307)	(0.0121)	(0.0192)
JNJ	0.0866	0.0346	0.0013	-0.0327	0.0016	0.0254	0.0271
	(0.0309)	(0.0154)	(0.0092)	(0.0181)	(0.0334)	(0.0122)	(0.0200)
JPM	-0.0020	0.0037	0.0308	-0.0701	0.0065	0.0195	-0.0601
	(0.0493)	(0.0268)	(0.0158)	(0.0288)	(0.0462)	(0.0161)	(0.0237)
MCD	0.1500	0.0546	-0.0089	-0.0169	0.0221	-0.0216	0.0235
	(0.0329)	(0.0175)	(0.0106)	(0.0200)	(0.0305)	(0.0112)	(0.0177)
MRK	0.0032	0.0352	-0.0202	-0.0084	-0.0151	0.0165	-0.0022
	(0.0292)	(0.0190)	(0.0123)	(0.0197)	(0.0271)	(0.0114)	(0.0187)
MSFT	-0.0200	0.0268	0.0277	-0.0536	0.0461	0.0033	-0.0247
	(0.0322)	(0.0218)	(0.0139)	(0.0182)	(0.0265)	(0.0107)	(0.0153)
NKE	0.0226	0.0616	-0.0059	0.0153	-0.0228	-0.0081	-0.0132
	(0.0339)	(0.0227)	(0.0133)	(0.0197)	(0.0287)	(0.0127)	(0.0160)
PFE	0.1283	-0.0078	0.0355	0.0014	-0.0094	-0.0024	0.0222
	(0.0285)	(0.0195)	(0.0124)	(0.0168)	(0.0283)	(0.0114)	(0.0187)
PG	0.0929	0.0986	-0.0530	-0.0612	0.0729	-0.0268	-0.0029
	(0.0275)	(0.0160)	(0.0112)	(0.0216)	(0.0289)	(0.0134)	(0.0121)
TRV	0.1594	0.0264	0.0070	-0.0423	-0.0529	-0.0194	-0.0312
	(0.0686)	(0.0208)	(0.0111)	(0.0299)	(0.0460)	(0.0191)	(0.0220)
UNH	-0.0305	0.0332	0.0353	0.0229	-0.0370	0.0058	-0.0084
	(0.0593)	(0.0261)	(0.0150)	(0.0205)	(0.0400)	(0.0145)	(0.0214)
UTX	-0.0565	0.0096	0.0395	-0.0216	-0.0386	0.0032	-0.0458
	(0.0698)	(0.0186)	(0.0107)	(0.0197)	(0.0323)	(0.0105)	(0.0181)
VZ	0.0621	0.0143	0.0012	-0.0405	-0.0316	-0.0129	0.0012
	(0.0429)	(0.0181)	(0.0101)	(0.0202)	(0.0376)	(0.0119)	(0.0188)
WMT	0.0852	-0.0054	0.0310	-0.0436	0.0315	-0.0053	0.0361
	(0.0347)	(0.0188)	(0.0107)	(0.0185)	(0.0341)	(0.0100)	(0.0175)
DIS	0.0376	0.0373	0.0024	-0.0236	0.0026	-0.0155	-0.0019
	(0.0290)	(0.0205)	(0.0140)	(0.0197)	(0.0369)	(0.0132)	(0.0234)
DJI	-0.0549	0.0328	-0.0020	-0.0615	0.1050	0.0034	-0.0282
	(0.0393)	(0.0208)	(0.0128)	(0.0205)	(0.0342)	(0.0130)	(0.0198)
average	0.0311	0.0223	0.0165	-0.0297	0.0042	-0.0029	-0.0086
pool est.	0.0299	0.0223	0.0162	-0.0313	0.0121	-0.0011	-0.0103
pool s.e.	(0.0034)	(0.0023)	(0.0015)	(0.0018)	(0.0032)	(0.0012)	(0.0019)
num of +	18	22	20	5	15	11	9
num of −	10	6	8	23	13	17	19
num of signi +	7	6	11	0	4	2	2
num of signi −	0	0	2	12	0	2	4

This table gives the estimates of the mean equations in the univariate coupled component models, and their asymptotic standard errors in parenthesis. 'pool est.' and 'pool s.e.' represent the MLE pool estimates and their standard errors.

Table 2: Estimates of the mean equations

	β_D	γ_D	ρ_D	γ_D^*	ρ_D^*	ν_D	ω_D
MMM	0.9716	0.0213	0.0273	-0.0189	-0.0107	6.8817	-0.0611
	(0.0055)	(0.0036)	(0.0042)	(0.0025)	(0.0025)	(0.5291)	(0.0266)
AXP	0.9710	0.0348	0.0298	-0.0162	-0.0112	9.6898	0.3035
	(0.0052)	(0.0040)	(0.0041)	(0.0025)	(0.0027)	(1.0451)	(0.0336)
AAPL	0.9177	0.0394	0.0381	-0.0157	-0.0109	8.9059	0.6216
	(0.0166)	(0.0056)	(0.0068)	(0.0032)	(0.0035)	(0.8917)	(0.0213)
BA	0.9587	0.0340	0.0261	-0.0110	-0.0116	8.6091	0.2315
	(0.0087)	(0.0045)	(0.0050)	(0.0026)	(0.0027)	(0.8269)	(0.0263)
CAT	0.9752	0.0236	0.0239	-0.0157	-0.0105	8.1879	0.3253
	(0.0050)	(0.0038)	(0.0042)	(0.0021)	(0.0022)	(0.7801)	(0.0298)
CVX	0.9626	0.0309	0.0256	-0.0177	-0.0137	14.2545	0.1014
	(0.0066)	(0.0039)	(0.0040)	(0.0025)	(0.0027)	(2.1566)	(0.0268)
CSCO	0.9527	0.0305	0.0314	-0.0187	-0.0183	11.8282	0.5052
	(0.0080)	(0.0041)	(0.0046)	(0.0029)	(0.0028)	(1.4338)	(0.0239)
KO	0.9616	0.0279	0.0262	-0.0163	-0.0143	8.9666	-0.0507
	(0.0082)	(0.0041)	(0.0048)	(0.0027)	(0.0026)	(0.9389)	(0.0258)
DD	0.9541	0.0351	0.0302	-0.0130	-0.0144	8.8106	0.1858
	(0.0102)	(0.0050)	(0.0053)	(0.0028)	(0.0029)	(0.8545)	(0.0258)
XOM	0.9605	0.0382	0.0263	-0.0095	-0.0157	12.8135	0.0350
	(0.0068)	(0.0043)	(0.0040)	(0.0026)	(0.0028)	(1.6619)	(0.0287)
GE	0.9590	0.0351	0.0351	-0.0182	-0.0152	11.7415	0.1273
	(0.0065)	(0.0041)	(0.0045)	(0.0025)	(0.0027)	(1.4348)	(0.0284)
HD	0.9559	0.0326	0.0345	-0.0245	-0.0176	8.7982	0.2666
	(0.0071)	(0.0041)	(0.0049)	(0.0030)	(0.0030)	(0.8661)	(0.0260)
IBM	0.9614	0.0331	0.0311	-0.0195	-0.0096	8.0274	0.1107
	(0.0076)	(0.0046)	(0.0051)	(0.0028)	(0.0027)	(0.7189)	(0.0276)
INTC	0.9638	0.0263	0.0221	-0.0116	-0.0125	13.8716	0.4828
	(0.0072)	(0.0034)	(0.0043)	(0.0023)	(0.0027)	(1.9630)	(0.0250)
JNJ	0.9440	0.0399	0.0394	-0.0226	-0.0131	9.2874	-0.1001
	(0.0084)	(0.0045)	(0.0051)	(0.0032)	(0.0032)	(0.9551)	(0.0255)
JPM	0.9739	0.0386	0.0370	-0.0187	-0.0102	8.9114	0.3347
	(0.0042)	(0.0042)	(0.0043)	(0.0025)	(0.0026)	(0.8937)	(0.0394)
MCD	0.7771	0.0677	0.0622	0.0017	0.0017	9.0474	0.2359
	(0.1108)	(0.0393)	(0.0097)	(0.0111)	(0.0102)	(1.0475)	(0.0945)
MRK	0.9391	0.0397	0.0400	-0.0153	-0.0105	7.6928	0.1438
	(0.0132)	(0.0054)	(0.0066)	(0.0032)	(0.0032)	(0.6331)	(0.0239)
MSFT	0.9366	0.0437	0.0495	-0.0111	-0.0073	11.3473	0.2991
	(0.0108)	(0.0047)	(0.0063)	(0.0030)	(0.0034)	(1.3919)	(0.0253)
NKE	0.9670	0.0315	0.0290	-0.0171	-0.0067	6.9838	0.2606
	(0.0069)	(0.0044)	(0.0051)	(0.0028)	(0.0026)	(0.5738)	(0.0292)
PFE	0.9640	0.0275	0.0322	-0.0144	-0.0077	11.3019	0.2051
	(0.0071)	(0.0038)	(0.0046)	(0.0025)	(0.0025)	(1.3785)	(0.0270)
PG	0.9480	0.0356	0.0332	-0.0163	-0.0108	8.7109	-0.0845
	(0.0096)	(0.0045)	(0.0051)	(0.0030)	(0.0032)	(0.8227)	(0.0246)
TRV	0.9660	0.0413	0.0333	-0.0119	-0.0101	8.2381	0.1086
	(0.0067)	(0.0050)	(0.0050)	(0.0028)	(0.0028)	(0.7880)	(0.0331)
UNH	0.9470	0.0405	0.0344	-0.0210	-0.0110	7.4073	0.3985
	(0.0099)	(0.0048)	(0.0058)	(0.0033)	(0.0033)	(0.6330)	(0.0254)
UTX	0.9596	0.0288	0.0336	-0.0182	-0.0163	9.3467	0.0982
	(0.0075)	(0.0044)	(0.0048)	(0.0027)	(0.0028)	(1.0045)	(0.0261)
VZ	0.9676	0.0278	0.0279	-0.0070	-0.0146	11.3300	0.0931
	(0.0062)	(0.0037)	(0.0042)	(0.0023)	(0.0026)	(1.3386)	(0.0280)
WMT	0.9742	0.0270	0.0238	-0.0097	-0.0085	8.0639	0.0663
	(0.0062)	(0.0037)	(0.0042)	(0.0023)	(0.0025)	(0.7403)	(0.0308)
DIS	0.9516	0.0344	0.0329	-0.0151	-0.0115	9.6181	0.1953
	(0.0116)	(0.0056)	(0.0051)	(0.0030)	(0.0031)	(1.0518)	(0.0298)
average	0.9515	0.0345	0.0327	-0.0151	-0.0115	9.5955	0.1943
pool est.	0.9861	0.0316	0.0314	-0.0059	-0.0022	8.0103	-0.8937
pool s.e.	(0.0004)	(0.0009)	(0.0006)	(0.0003)	(0.0003)	(0.1339)	(0.0374)

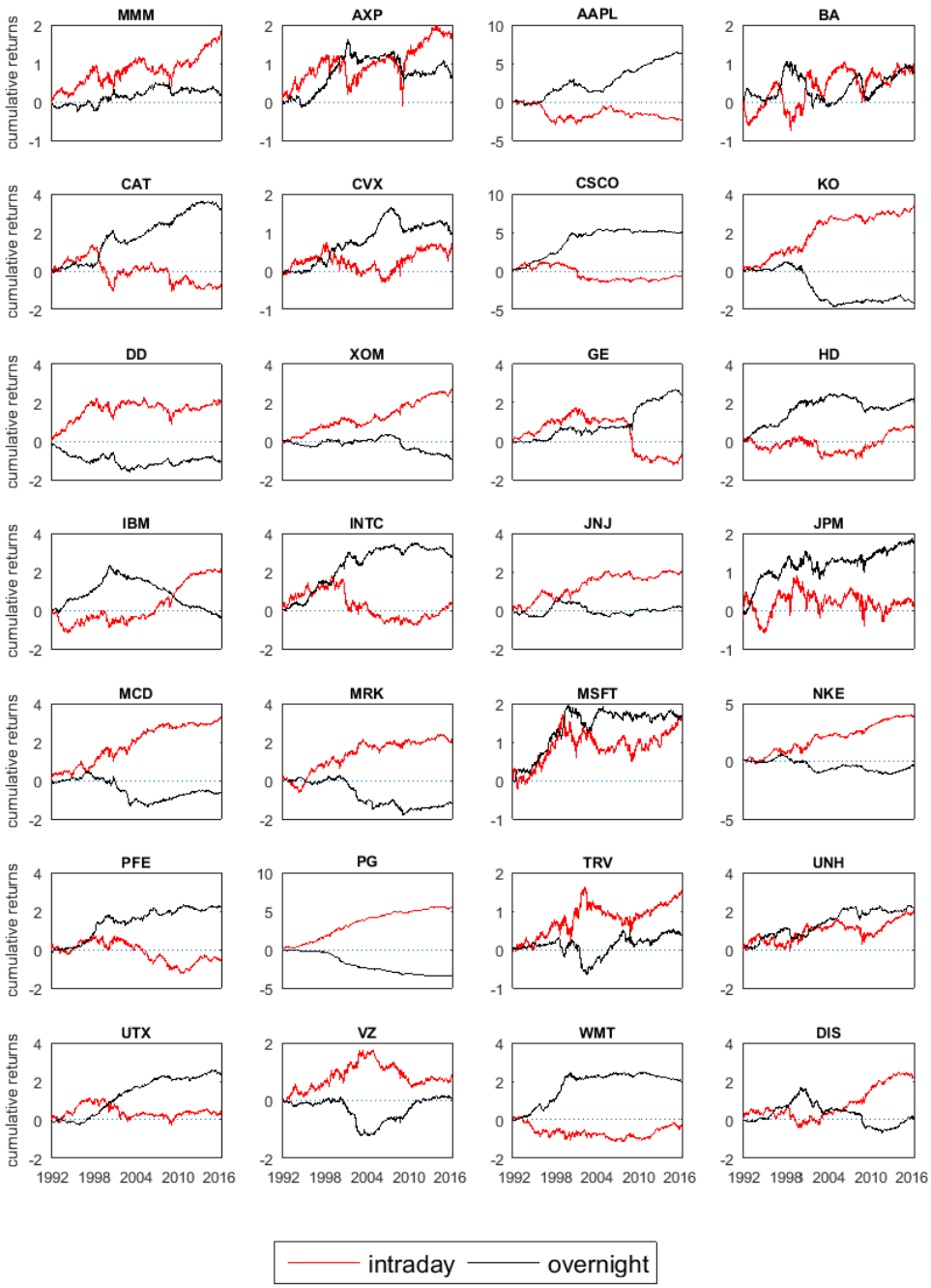
Continued on the next page.

Table 3: Estimates of variance equations in the univariate coupled component models

	β_N	γ_N	ρ_N	γ_N^*	ρ_N^*	ν_N	ω_N
MMM	0.9683	0.0344	0.0322	-0.0127	-0.0235	2.9893	-0.9498
	(0.0063)	(0.0054)	(0.0049)	(0.0033)	(0.0032)	(0.1333)	(0.0325)
AXP	0.9722	0.0387	0.0379	-0.0188	-0.0229	3.7613	-0.5535
	(0.0046)	(0.0055)	(0.0042)	(0.0032)	(0.0029)	(0.1895)	(0.0385)
AAPL	0.9234	0.0619	0.0629	-0.0154	-0.0146	2.6745	-0.3351
	(0.0117)	(0.0081)	(0.0064)	(0.0045)	(0.0040)	(0.0977)	(0.0278)
BA	0.9710	0.0206	0.0327	-0.0120	-0.0168	3.0865	-0.6114
	(0.0062)	(0.0051)	(0.0043)	(0.0026)	(0.0026)	(0.1367)	(0.0298)
CAT	0.9708	0.0305	0.0393	-0.0103	-0.0197	2.8998	-0.5496
	(0.0057)	(0.0058)	(0.0049)	(0.0030)	(0.0029)	(0.1229)	(0.0350)
CVX	0.9778	0.0197	0.0264	-0.0110	-0.0157	4.1452	-0.7624
	(0.0039)	(0.0039)	(0.0031)	(0.0023)	(0.0024)	(0.2379)	(0.0324)
CSCO	0.9631	0.0368	0.0344	-0.0140	-0.0212	3.5092	-0.3252
	(0.0065)	(0.0058)	(0.0044)	(0.0029)	(0.0030)	(0.1532)	(0.0306)
KO	0.9658	0.0409	0.0275	-0.0189	-0.0201	3.7388	-0.8462
	(0.0058)	(0.0053)	(0.0042)	(0.0031)	(0.0029)	(0.1821)	(0.0319)
DD	0.9566	0.0337	0.0425	-0.0175	-0.0159	3.5630	-0.6296
	(0.0084)	(0.0060)	(0.0050)	(0.0035)	(0.0030)	(0.1734)	(0.0290)
XOM	0.9729	0.0258	0.0310	-0.0122	-0.0140	5.0318	-0.7701
	(0.0052)	(0.0043)	(0.0037)	(0.0026)	(0.0025)	(0.3115)	(0.0322)
GE	0.9614	0.0401	0.0463	-0.0204	-0.0221	4.7818	-0.6118
	(0.0062)	(0.0056)	(0.0046)	(0.0031)	(0.0030)	(0.2773)	(0.0340)
HD	0.9537	0.0452	0.0373	-0.0220	-0.0259	3.1685	-0.6014
	(0.0076)	(0.0062)	(0.0048)	(0.0039)	(0.0036)	(0.1418)	(0.0299)
IBM	0.9632	0.0414	0.0389	-0.0121	-0.0218	2.7722	-0.7656
	(0.0064)	(0.0062)	(0.0048)	(0.0033)	(0.0030)	(0.1088)	(0.0330)
INTC	0.9662	0.0300	0.0305	-0.0131	-0.0155	2.8768	-0.4046
	(0.0083)	(0.0064)	(0.0047)	(0.0033)	(0.0027)	(0.1116)	(0.0295)
JNJ	0.9596	0.0323	0.0348	-0.0166	-0.0262	4.0828	-0.8547
	(0.0063)	(0.0050)	(0.0043)	(0.0030)	(0.0030)	(0.2189)	(0.0281)
JPM	0.9754	0.0350	0.0366	-0.0198	-0.0226	4.0152	-0.4103
	(0.0038)	(0.0046)	(0.0040)	(0.0027)	(0.0026)	(0.2098)	(0.0397)
MCD	0.8012	0.0488	0.0800	-0.0018	-0.0000	4.0211	-0.2824
	(0.0159)	(0.0477)	(0.0787)	(0.0082)	(0.0069)	(0.1751)	(0.0458)
MRK	0.9458	0.0398	0.0435	-0.0135	-0.0209	3.2409	-0.6822
	(0.0111)	(0.0064)	(0.0062)	(0.0036)	(0.0035)	(0.1456)	(0.0271)
MSFT	0.9433	0.0543	0.0452	-0.0098	-0.0115	2.8706	-0.6569
	(0.0102)	(0.0076)	(0.0053)	(0.0038)	(0.0035)	(0.1101)	(0.0293)
NKE	0.9543	0.0396	0.0456	-0.0147	-0.0177	2.3998	-0.8391
	(0.0095)	(0.0068)	(0.0061)	(0.0037)	(0.0038)	(0.0876)	(0.0301)
PFE	0.9619	0.0424	0.0364	-0.0112	-0.0118	3.4040	-0.6243
	(0.0074)	(0.0058)	(0.0045)	(0.0032)	(0.0031)	(0.1577)	(0.0324)
PG	0.9445	0.0292	0.0370	-0.0206	-0.0235	3.6016	-0.9050
	(0.0097)	(0.0058)	(0.0050)	(0.0035)	(0.0033)	(0.1717)	(0.0240)
TRV	0.9740	0.0398	0.0347	-0.0125	-0.0170	2.9686	-0.9225
	(0.0049)	(0.0060)	(0.0044)	(0.0031)	(0.0029)	(0.1268)	(0.0393)
UNH	0.9354	0.0521	0.0547	-0.0076	-0.0202	2.2533	-0.7491
	(0.0109)	(0.0077)	(0.0066)	(0.0045)	(0.0042)	(0.0836)	(0.0293)
UTX	0.9708	0.0303	0.0324	-0.0202	-0.0195	3.3906	-0.7534
	(0.0048)	(0.0048)	(0.0041)	(0.0030)	(0.0027)	(0.1623)	(0.0332)
VZ	0.9661	0.0302	0.0325	-0.0224	-0.0138	3.7265	-0.7555
	(0.0057)	(0.0052)	(0.0040)	(0.0031)	(0.0028)	(0.1886)	(0.0300)
WMT	0.9756	0.0240	0.0302	-0.0111	-0.0126	3.4965	-0.7485
	(0.0052)	(0.0047)	(0.0038)	(0.0026)	(0.0026)	(0.1641)	(0.0328)
DIS	0.9552	0.0369	0.0396	-0.0147	-0.0183	3.4211	-0.6600
	(0.0074)	(0.0073)	(0.0074)	(0.0034)	(0.0032)	(0.1618)	(0.0320)
Average	0.9553	0.0369	0.0394	-0.0145	-0.0180	3.4247	-0.6629
pool est.	0.9823	0.0426	0.0307	-0.0086	-0.0130	3.2019	-1.5104
pool s.e.	(0.0006)	(0.0010)	(0.0009)	(0.0005)	(0.0005)	(0.0265)	(0.0322)

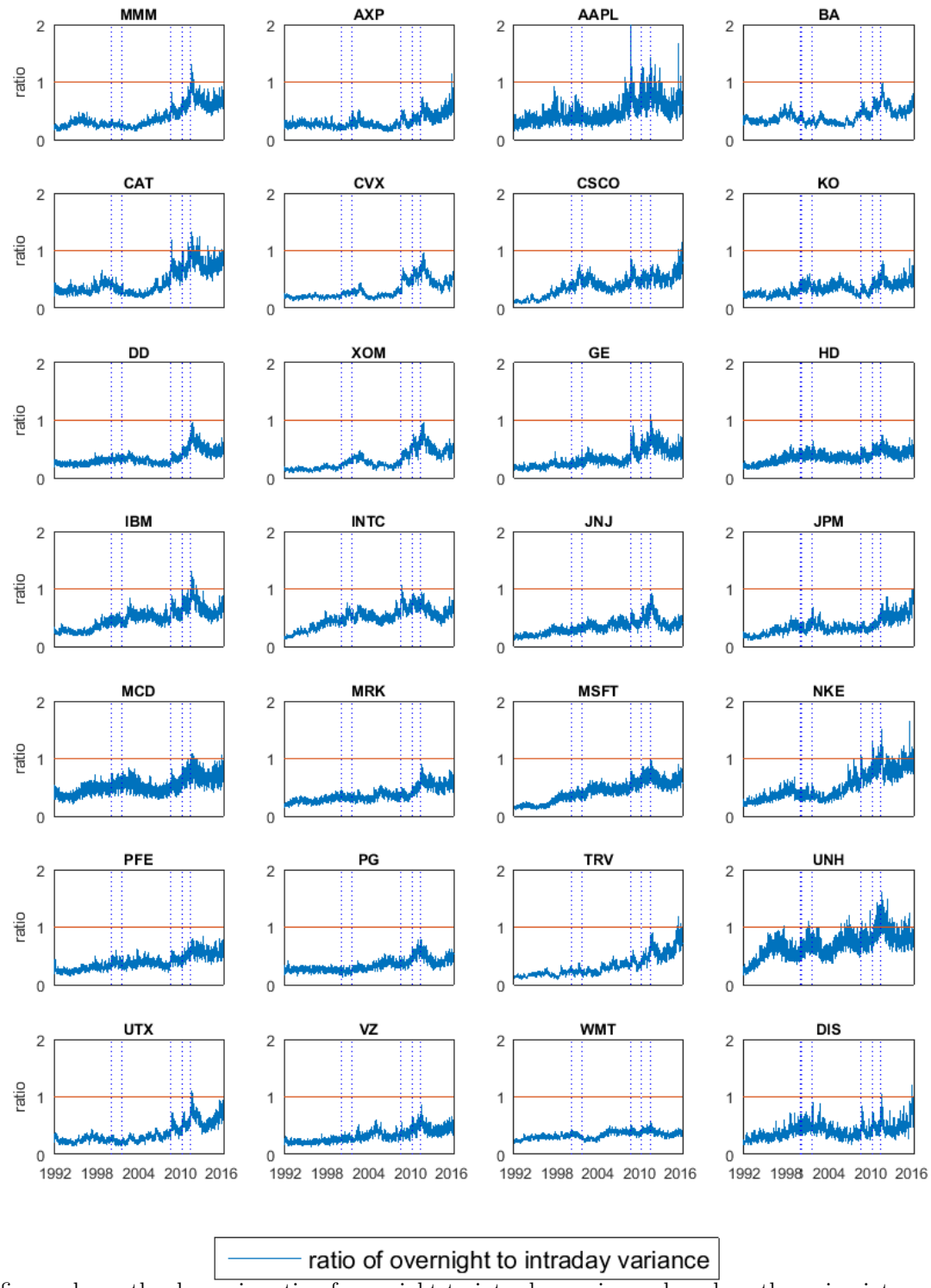
This table presents the estimates of the variance equations in the univariate coupled component models, and their asymptotic standard errors in parenthesis. 'pool est.' and 'pool s.e.' represent the MLE pool estimates and their standard errors.

Table 3: Estimates of variance equations in the univariate coupled component model model(cont.)



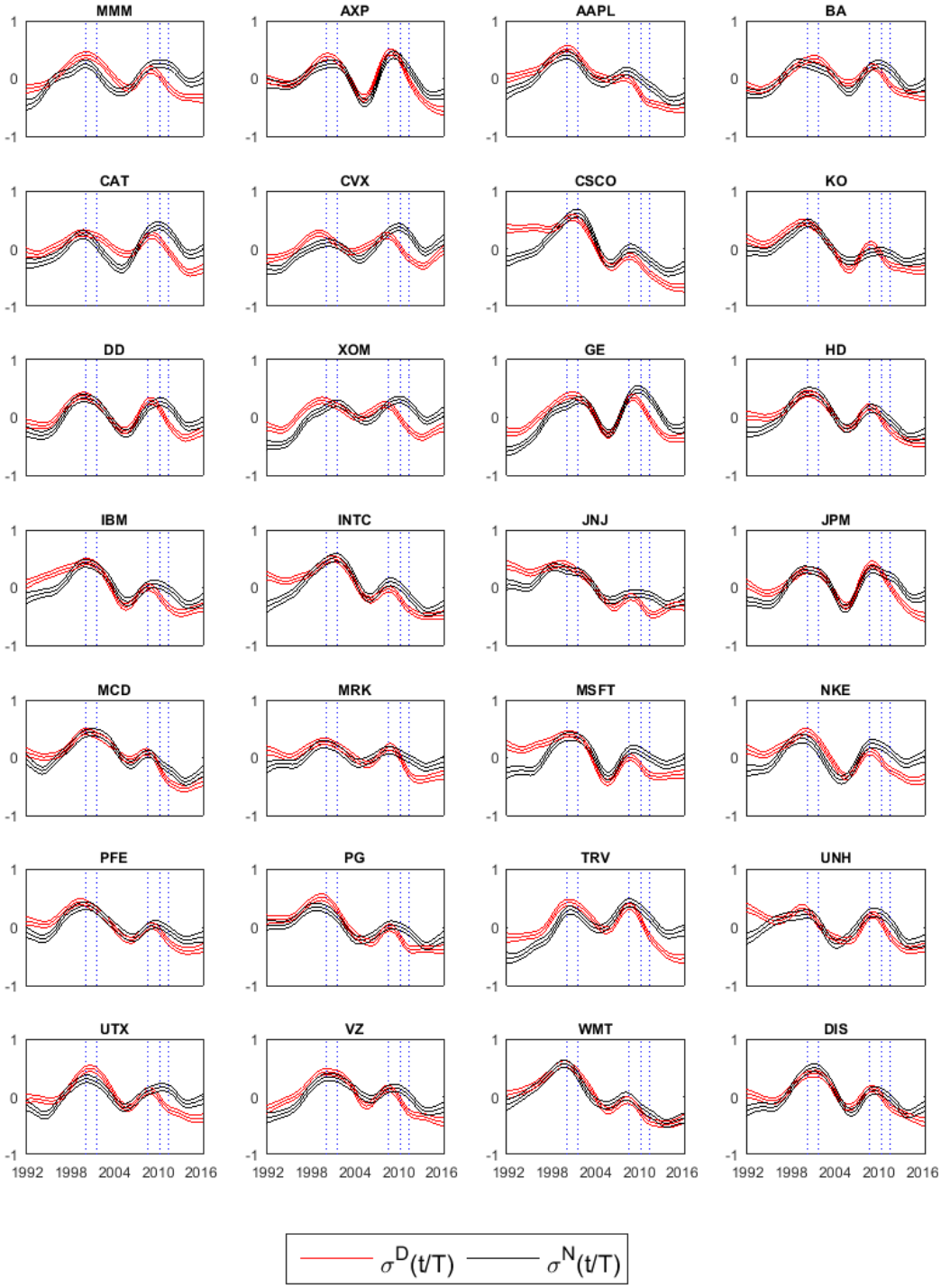
This figure shows the cumulative intraday (in red) and the cumulative overnight (in black) returns: one subplot for each stock.

Figure 1: Cumulative intraday and overnight returns



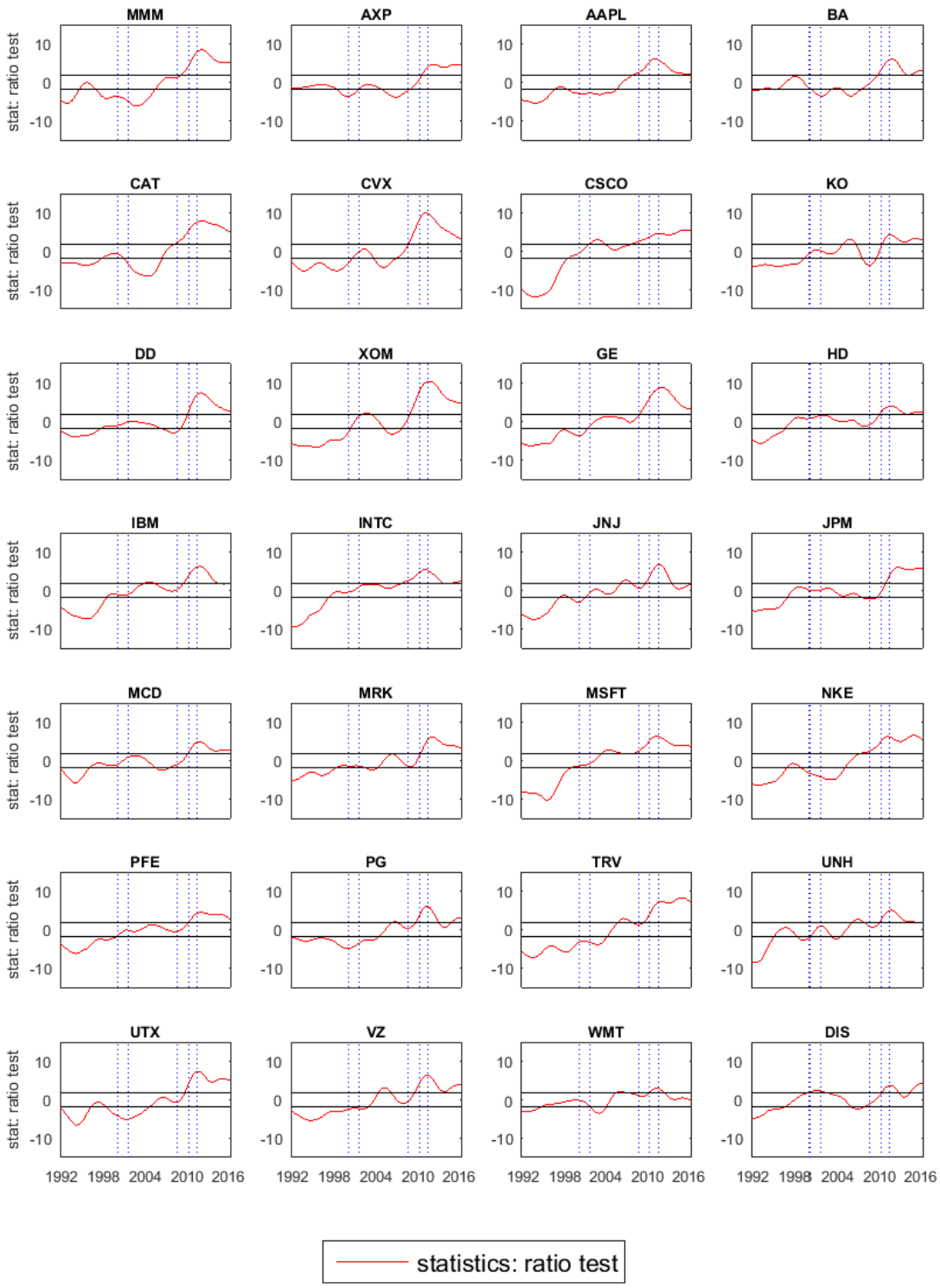
This figure shows the dynamic ratio of overnight to intraday variance, based on the univariate coupled component model: one subplot for each stock. The five dash vertical lines from left to right represent the dates: 10 March 2000(dot-com bubble), 11 September 2001(the September 11 attacks), 16 September 2008(financial crisis), 06 May 2010 (flash crash) and 01 August 2011 (August 2011 stock markets fall), respectively.

Figure 2: Ratios of overnight to intraday variance: univariate model

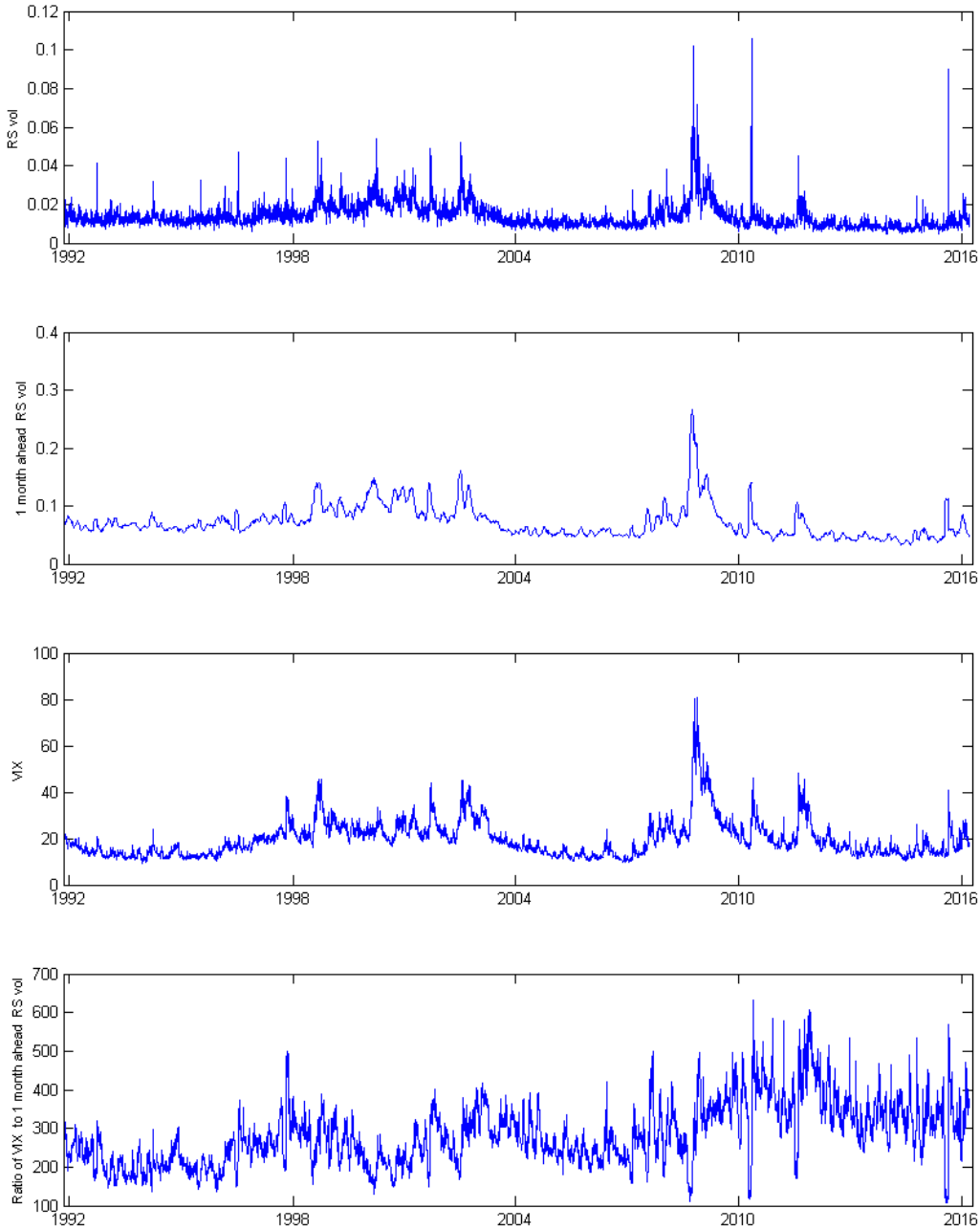


This figure shows the estimated intraday (in red) and overnight (in black) long run components,$\sigma^D(t/T)$ and $\sigma^N(t/T)$, based on the univariate coupled component model: one subplot for each stock. The five dash vertical lines from left to right represent the dates: 10 March 2000(dot-com bubble), 11 September 2001(the September 11 attacks), 16 September 2008(financial crisis), 06 May 2010 (flash crash) and 01 August 2011 (August 2011 stock markets fall), respectively.

Figure 3: Long run component σ: univariate model

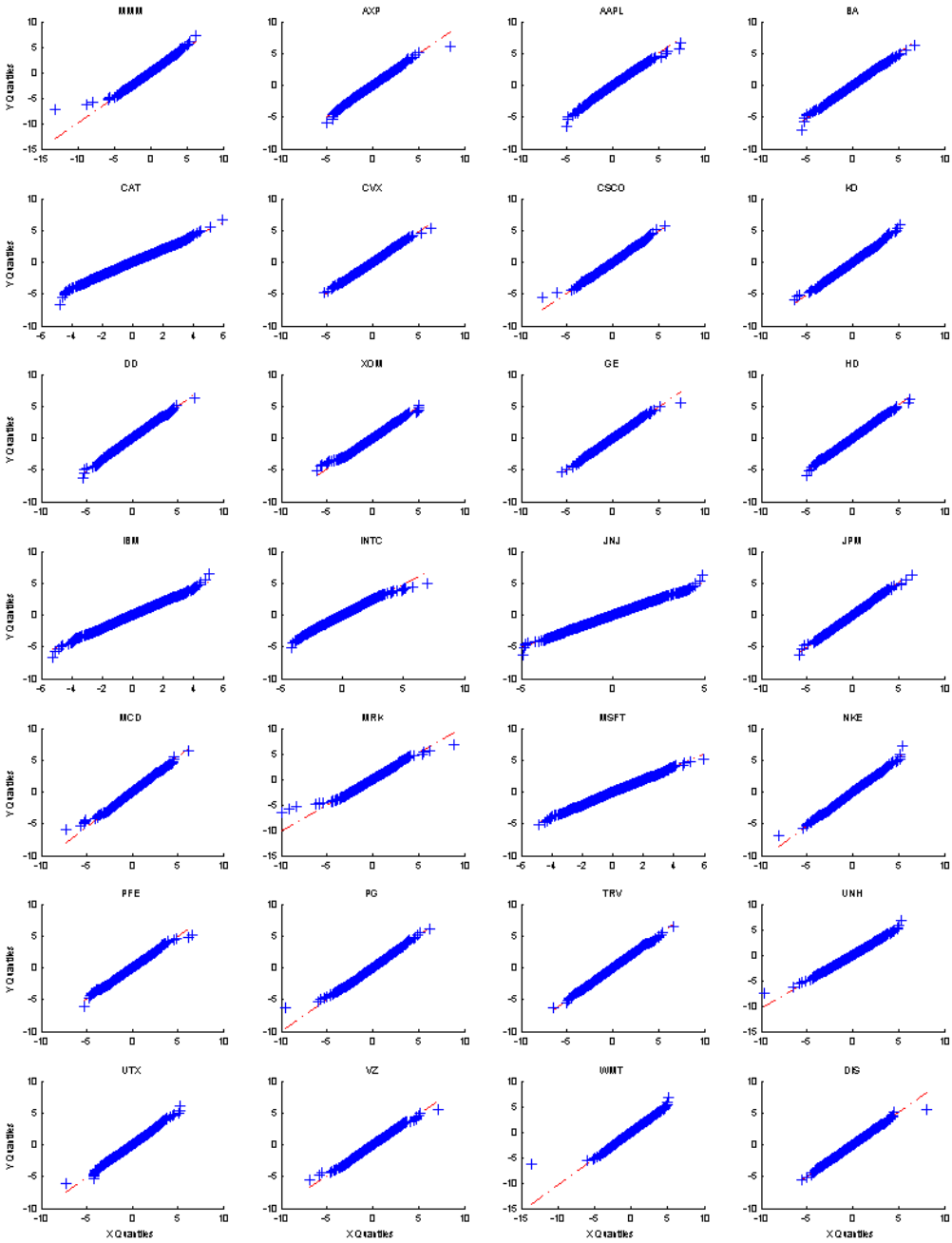


The red lines represent the statistics of the ratio tests, with the null hypothesis $H_0 : \exp\left(\sigma_0^N(t/T)\right) = \rho \exp\left(\sigma_0^D(t/T)\right)$. The black lines indicate the 95% confidence intervals of the statistics under the null. The five dash vertical lines from left to right represent the dates: 10 March 2000(dot-com bubble), 11 September 2001(the September 11 attacks), 16 September 2008(financial crisis), 06 May 2010 (flash crash) and 01 August 2011 (August 2011 stock markets fall), respectively.



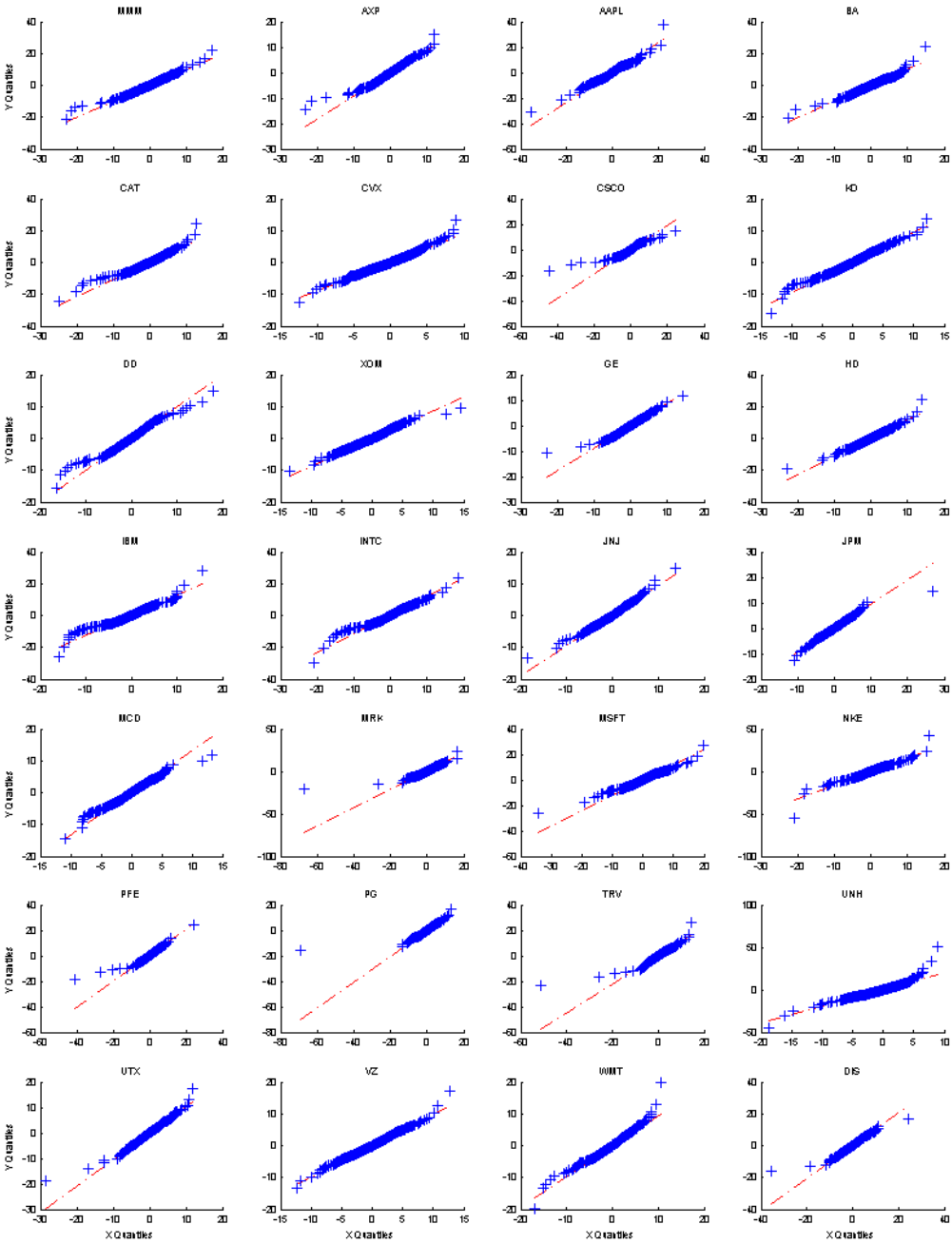
The figure shows the Rogers and Satschell(RS) volatility, the one-month ahead monthly RS volatility, VIX, and the ratio of VIX to the one-month ahead monthly RS volatility. The RS volatility is the average RS volatility across the 28 stocks.

Figure 5: RS, VIX and their ratio



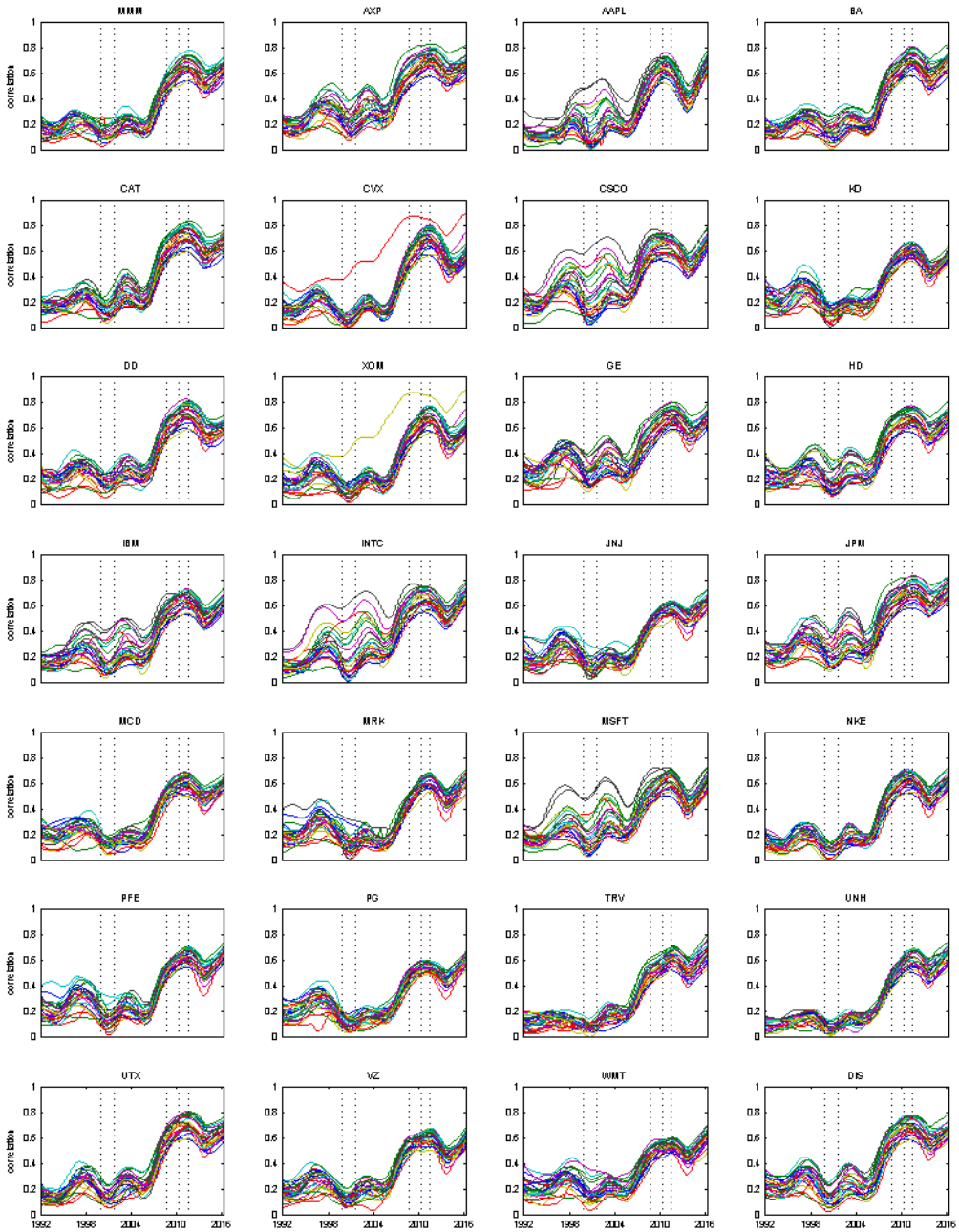
This figure displays Q-Q plots of the quantiles of the intraday innovations (X axis), versus the theoretical quantiles of the student t distribution with the $\hat{\nu}_D$ degrees of freedom (Y axis): one panel for each stock.

Figure 6: QQ plot of the intraday innovations



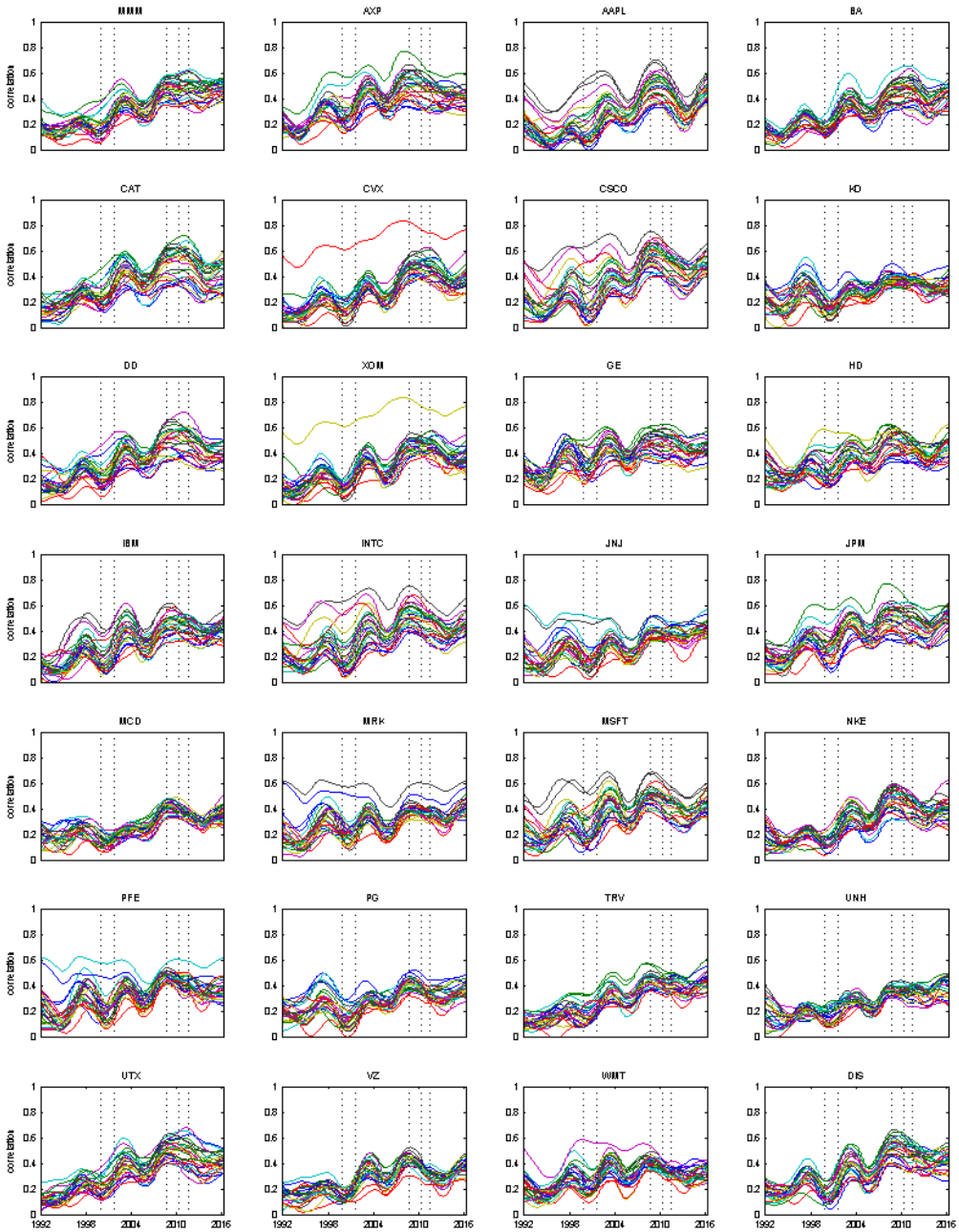
This figure displays Q-Q plots of the quantiles of the overnight innovations (X axis), versus the theoretical quantiles of the student t distribution with the $\hat{\nu}_D$ degrees of freedom (Y axis): one panel for each stock.

Figure 7: QQ plot of the overnight innovations



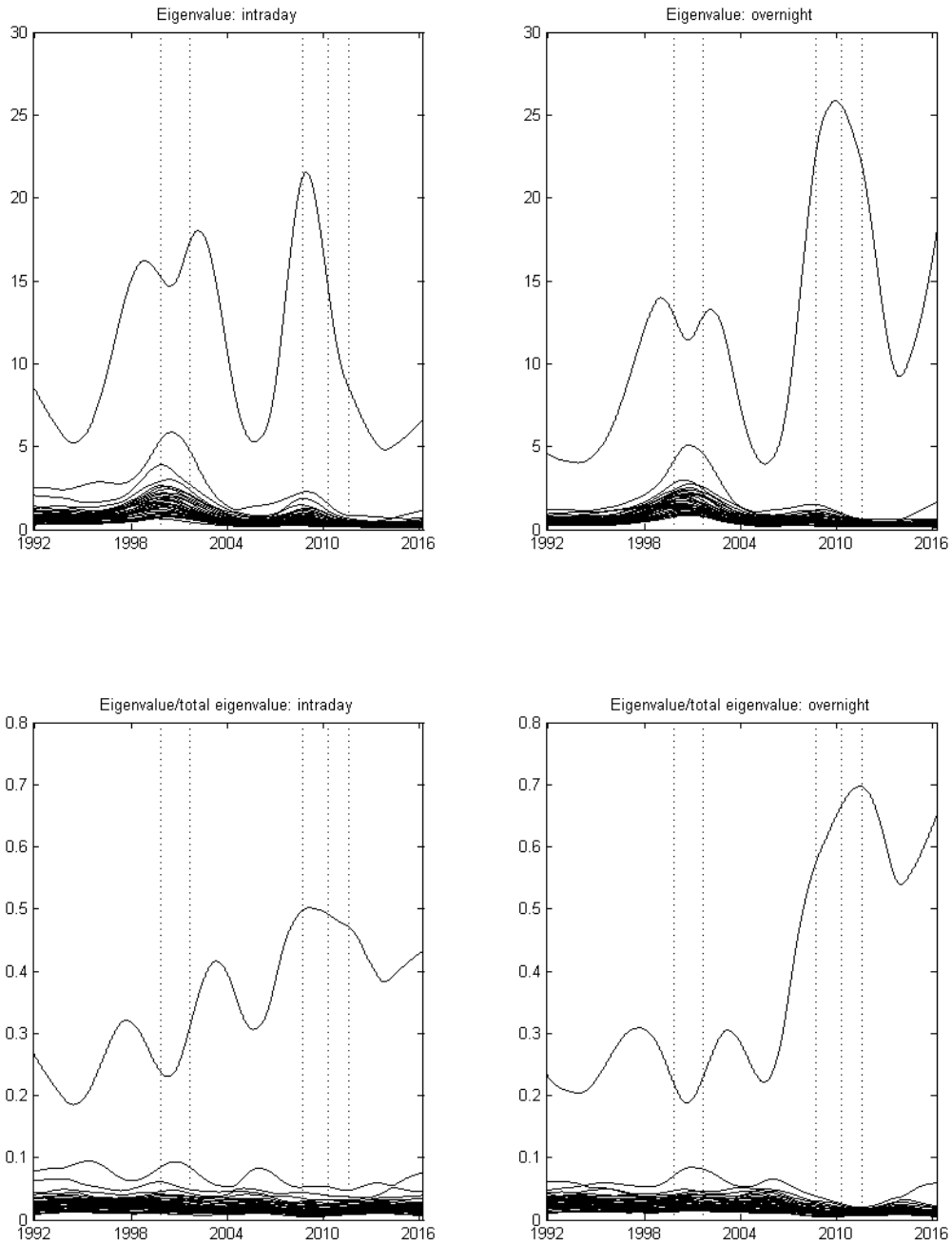
Each panel presents the long-run intraday correlations between that stock and the rest of stocks, implied by the multivariate coupled component model. The five dash vertical lines from left to right indicate the dates: 10 March 2000(dot-com bubble), 11 September 2001(the September 11 attacks), 16 September 2008(financial crisis), 06 May 2010 (flash crash) and 01 August 2011 (August 2011 stock markets fall), respectively.

Figure 8: Long-run overnight correlations



Each panel presents the 27 time series of long-run overnight correlations between that stock and the rest of stocks, implied by the multivariate coupled component model. The five dash vertical lines from left to right indicate the dates: 10 March 2000(dot-com bubble), 11 September 2001(the September 11 attacks), 16 September 2008(financial crisis), 06 May 2010 (flash crash) and 01 August 2011 (August 2011 stock markets fall), respectively.

Figure 9: Long-run intraday correlations



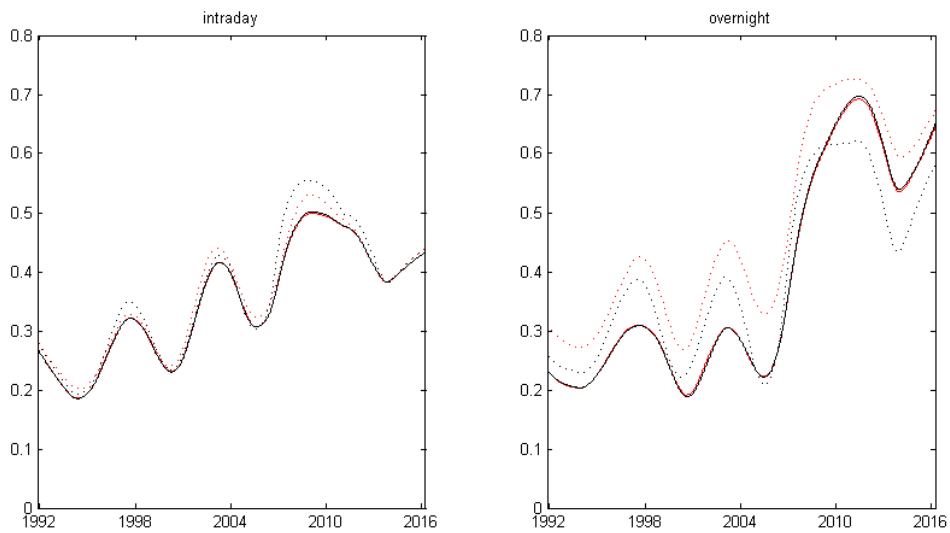
The upper panel plots the eigenvalues of the covariance matrices, and the lower panel plots the eigenvalues divided by the sum of eigenvalues. The five dash vertical lines from left to right indicate the dates: 10 March 2000(dot-com bubble), 11 September 2001(the September 11 attacks), 16 September 2008(financial crisis), 06 May 2010 (flash crash) and 01 August 2011 (August 2011 stock markets fall), respectively.

Figure 10: Eigenvalues of covariance matrices

	$\beta_D = \beta_N$	$\gamma_D = \gamma_N$	$\rho_D = \rho_N$	$\gamma_D^* = \gamma_N^*$	$\rho_D^* = \rho_N^*$	$\nu_D = \nu_N$	$\omega_D = \omega_N$	$\gamma_N = \rho_D$	$(\beta_D, \gamma_D, \rho_D, \gamma_D^*, \rho_D^*) = (\beta_N, \gamma_N, \rho_N, \gamma_N^*, \rho_N^*)$
MMM	0.4463	0.0404	0.4074	0.1300	0.0015	0.0000	0.0000	0.1158	0.0015
AXP	0.7387	0.5529	0.1337	0.5090	0.0031	0.0000	0.0000	0.0587	0.0004
AAPL	0.6655	0.0175	0.0047	0.9604	0.5031	0.0000	0.0000	0.0016	0.0000
BA	0.0412	0.0357	0.2624	0.7881	0.1553	0.0000	0.0000	0.2269	0.0222
CAT	0.2142	0.2870	0.0058	0.1250	0.0090	0.0000	0.0000	0.1444	0.0015
CVX	0.0010	0.0332	0.8557	0.0381	0.5759	0.0000	0.0000	0.1356	0.0106
CSCO	0.0811	0.3699	0.6042	0.2516	0.4850	0.0000	0.0000	0.3166	0.0169
KO	0.5108	0.0408	0.8287	0.5255	0.1401	0.0000	0.0000	0.0035	0.0122
DD	0.7178	0.8493	0.0533	0.2979	0.7134	0.0000	0.0000	0.5036	0.0953
XOM	0.0113	0.0358	0.3672	0.4597	0.6693	0.0000	0.0000	0.9082	0.0377
GE	0.5704	0.4586	0.0643	0.5741	0.0904	0.0000	0.0000	0.2692	0.0004
HD	0.7339	0.0834	0.6725	0.6071	0.0790	0.0000	0.0000	0.0616	0.1236
IBM	0.7507	0.2660	0.2244	0.0799	0.0027	0.0000	0.0000	0.0597	0.0032
INTC	0.7053	0.6005	0.1628	0.7053	0.4233	0.0000	0.0000	0.1713	0.1426
JNJ	0.0149	0.2568	0.4710	0.1707	0.0027	0.0000	0.0000	0.1627	0.0049
JPM	0.5514	0.5603	0.9484	0.7509	0.0008	0.0000	0.0000	0.6026	0.0012
MCD	0.8264	0.8265	0.8190	0.8200	0.8627	0.0000	0.0001	0.7875	0.9999
MRK	0.4594	0.9906	0.6508	0.7012	0.0283	0.0000	0.0000	0.9729	0.1458
MSFT	0.4468	0.2405	0.5895	0.7901	0.3990	0.0000	0.0000	0.5058	0.3634
NKE	0.0625	0.2967	0.0206	0.5988	0.0159	0.0000	0.0000	0.0646	0.0362
PFE	0.7098	0.0303	0.4915	0.4023	0.2893	0.0000	0.0000	0.0508	0.0377
PG	0.6715	0.3739	0.5645	0.3481	0.0069	0.0000	0.0000	0.4775	0.0481
TRV	0.0957	0.8364	0.8171	0.8814	0.0976	0.0000	0.0000	0.2369	0.0499
UNH	0.2059	0.1949	0.0139	0.0167	0.0901	0.0000	0.0000	0.0113	0.0028
UTX	0.0310	0.8085	0.8380	0.6094	0.4181	0.0000	0.0000	0.4670	0.0098
VZ	0.7429	0.6957	0.3921	0.0001	0.8273	0.0000	0.0000	0.6222	0.0015
WMT	0.7214	0.5910	0.2347	0.6712	0.2528	0.0000	0.0000	0.9611	0.4234
DIS	0.5415	0.7408	0.2892	0.9213	0.1324	0.0000	0.0000	0.4689	0.0731

This table presents the p-values of the Wald tests for several sets of null hypothesis: $H_0 : \beta_D = \beta_N$, $H_0 : \gamma_D = \gamma_N$, $H_0 : \rho_D = \rho_N$, $H_0 : \gamma_D^* = \gamma_N^*$, $H_0 : \rho_D^* = \rho_N^*$, $H_0 : \nu_D = \nu_N$, $H_0 : \omega_D = \omega_N$, $H_0 : \gamma_N = \rho_D$, and $H_0 : (\beta_D, \gamma_D, \rho_D, \gamma_D^*, \rho_D^*) = (\beta_N, \gamma_N, \rho_N, \gamma_N^*, \rho_N^*)$.

Table 4: Wald tests



Left(right) panel plots the largest eigenvalue proportion of estimated intraday(overnight) covariance matrix: red(black) lines for using non-robust(robust) correlation in the initial step. Solid(dash) lines are further used to indicate the initial(updated) estimators,

Figure 11: comparison of robust and non-robust initial correlation estimator

	student t log-likelihood				quasi Gaussian log-likelihood			
	$l_{cogarch}$	l_{garch}	GW stat.	p-val.	$l_{cogarch}$	l_{garch}	GW stat.	p-val.
MMM	1.1832	1.1963	6.3242	0.0119	1.1901	1.2092	6.2041	0.0127
AXP	1.3335	1.3473	2.7307	0.0984	1.3727	1.3974	2.2438	0.1342
AAPL	1.6569	1.6701	2.5099	0.1131	1.6642	1.6854	2.0921	0.1481
BA	1.4584	1.4715	4.7922	0.0286	1.4830	1.5048	5.5787	0.0182
CAT	1.5693	1.5761	1.9846	0.1589	1.5834	1.5915	1.4751	0.2245
CVX	1.6026	1.6058	0.2129	0.6445	1.6135	1.6178	0.2728	0.6015
CSCO	1.3807	1.3936	3.5761	0.0586	1.3832	1.4027	3.4340	0.0639
KO	1.1211	1.1347	2.3200	0.1277	1.1281	1.1483	2.6316	0.1048
DD	1.4821	1.4956	2.8619	0.0907	1.4996	1.5218	3.2895	0.0697
XOM	1.4198	1.4284	1.9184	0.1660	1.4304	1.4448	3.1887	0.0742
GE	1.3125	1.3328	1.8666	0.1719	1.3371	1.3638	0.7425	0.3889
HD	1.4018	1.4148	3.1614	0.0754	1.4179	1.4408	2.4230	0.1196
IBM	1.3442	1.3538	3.0191	0.0823	1.3586	1.3730	2.1378	0.1437
INTC	1.5674	1.5736	0.6431	0.4226	1.5913	1.6025	1.4440	0.2295
JNJ	1.1759	1.1807	0.2140	0.6436	1.1863	1.2141	1.0331	0.3094
JPM	1.3844	1.3902	0.4479	0.5033	1.3933	1.4047	0.5767	0.4476
MCD	1.1675	1.1772	0.6391	0.4240	1.2151	1.2332	0.6079	0.4356
MRK	1.4255	1.4298	0.4247	0.5146	1.4483	1.4694	1.0146	0.3138
MSFT	1.5373	1.5511	3.8529	0.0497	1.5612	1.5889	4.4249	0.0354
NKE	1.4760	1.4811	0.6902	0.4061	1.4872	1.5010	1.2014	0.2730
PFE	1.3581	1.3701	1.4091	0.2352	1.3698	1.3924	1.0637	0.3024
PG	1.0967	1.0992	0.1259	0.7227	1.1153	1.1345	1.0687	0.3012
TRV	1.1117	1.1185	0.5408	0.4621	1.1186	1.1280	0.8314	0.3619
UNH	1.6005	1.6018	0.1291	0.7193	1.5949	1.5961	0.0506	0.8219
UTX	1.3216	1.3336	2.1103	0.1463	1.3437	1.3566	0.7855	0.3755
VZ	1.1679	1.1791	3.5395	0.0599	1.1797	1.2149	1.9419	0.1635
WMT	1.2735	1.2720	0.0452	0.8316	1.4252	1.4429	1.3135	0.2518
DIS	1.3163	1.3222	0.4560	0.4995	1.3300	1.3439	0.9672	0.3254

The table presents the GW test of the null that the one-component and the coupled component model have equal expected loss, with minus the out-of-sample t log-likelihood or quasi Gaussian log-likelihood as the loss function. $l_{cogarch}$ represents the average loss value of the coupled component model, and l_{garch} represents the average loss value of the one component BETA-T-EGARCH model with open-close returns.

Table 5: GW tests: univariate model

	$\lvert\epsilon_d\rvert$	$\lvert\epsilon_n\rvert$	$\lvert r_d\rvert$	$\lvert r_n\rvert$	ϵ_d^2	ϵ_n^2	r_d^2	r_n^2
MMM	0.7102	0.2692	0.0000	0.0000	0.9611	0.9923	0.0000	0.0000
AXP	0.3953	0.1358	0.0000	0.0000	0.1828	0.9900	0.0000	0.0000
AAPL	0.7447	0.4793	0.0000	0.0000	0.0757	0.9997	0.0000	0.9558
BA	0.1846	0.2596	0.0000	0.0000	0.2914	0.9962	0.0000	0.0000
CAT	0.9292	0.2899	0.0000	0.0000	0.7909	0.9988	0.0000	0.0000
CVX	0.3064	0.5760	0.0000	0.0000	0.0637	0.5412	0.0000	0.0000
CSCO	0.0739	0.7701	0.0000	0.0000	0.0170	1.0000	0.0000	0.0000
KO	0.5040	0.7539	0.0000	0.0000	0.5222	0.9752	0.0000	0.0000
DD	0.6311	0.7901	0.0000	0.0000	0.0710	0.2432	0.0000	0.0000
XOM	0.4986	0.1337	0.0000	0.0000	0.3109	0.7145	0.0000	0.0000
GE	0.1884	0.8045	0.0000	0.0000	0.0138	0.9978	0.0000	0.0000
HD	0.3390	0.4095	0.0000	0.0000	0.1141	0.9977	0.0000	0.0120
IBM	0.4555	0.2208	0.0000	0.0000	0.7846	0.9603	0.0000	0.0000
INTC	0.3869	0.0606	0.0000	0.0000	0.8838	0.9853	0.0000	0.0003
JNJ	0.2579	0.4742	0.0000	0.0000	0.5205	0.9366	0.0000	0.0286
JPM	0.2962	0.4567	0.0000	0.0000	0.3807	1.0000	0.0000	0.0000
MCD	0.1996	0.3503	0.0000	0.0000	0.3556	0.9685	0.0000	0.0000
MRK	0.2003	0.7233	0.0000	0.0000	0.0000	1.0000	0.0000	1.0000
MSFT	0.5201	0.8494	0.0000	0.0000	0.7620	1.0000	0.0000	0.0000
NKE	0.0407	0.8972	0.0000	0.0000	0.0924	0.9980	0.0000	0.9857
PFE	0.1066	0.6689	0.0000	0.0000	0.0957	1.0000	0.0000	0.0025
PG	0.8185	0.8830	0.0000	0.0000	0.9378	1.0000	0.0000	1.0000
TRV	0.4188	0.3676	0.0000	0.0000	0.1089	0.9992	0.0000	0.0000
UNH	0.0045	0.4844	0.0000	0.0000	0.2356	0.9829	0.0000	0.0000
UTX	0.0096	0.1683	0.0000	0.0000	0.0018	0.9999	0.0000	0.0017
VZ	0.1302	0.2894	0.0000	0.0000	0.5510	0.0343	0.0000	0.0000
WMT	0.1704	0.5944	0.0000	0.0000	0.9867	0.9023	0.0000	0.0000
DIS	0.9257	0.1840	0.0000	0.0000	0.7475	1.0000	0.0000	0.0000

This table gives the p-values of the Ljung-Box Q-tests for absolute(squared) residuals and returns.

Table 6: Diagnostic checking for GARCH effects

	β_D	γ_D	ρ_D	γ_D^*	ρ_D^*	ν_D	ω_D
MMM	0.8504	0.0350	0.0788	-0.0094	0.0001	5.1024	-0.1047
	(0.0811)	(0.0097)	(0.0226)	(0.0053)	(0.0047)	(0.3138)	(0.0196)
AXP	0.9332	0.0442	0.0466	-0.0138	-0.0082	8.5370	-0.0540
	(0.0132)	(0.0054)	(0.0066)	(0.0033)	(0.0035)	(0.8093)	(0.0247)
AAPL	0.8056	0.0587	0.0646	-0.0169	0.0026	7.4693	-0.0421
	(0.0419)	(0.0080)	(0.0100)	(0.0047)	(0.0049)	(-0.6579)	(0.0190)
BA	0.7165	0.0636	0.0721	0.0000	0.0089	7.8943	-0.0268
	(0.0519)	(0.0079)	(0.0098)	(0.0050)	(0.0059)	(0.7117)	(0.0176)
CAT	0.9447	0.0271	0.0311	-0.0098	-0.0004	7.3387	-0.0276
	(0.0158)	(0.0047)	(0.0072)	(0.0026)	(0.0028)	(0.6330)	(0.0221)
CVX	0.9475	0.0331	0.0285	-0.0116	-0.0038	9.7347	-0.0069
	(0.0139)	(0.0050)	(0.0056)	(0.0029)	(0.0030)	(1.1058)	(0.0238)
CSCO	0.9282	0.0412	0.0412	-0.0152	-0.0134	8.3211	-0.0515
	(0.0154)	(0.0057)	(0.0067)	(0.0035)	(0.0031)	(0.7676)	(0.0229)
KO	0.8383	0.0477	0.0567	-0.0139	0.0001	8.9520	-0.0161
	(0.0471)	(0.0078)	(0.0101)	(0.0042)	(0.0049)	(0.9210)	(0.0189)
DD	0.7722	0.0523	0.0792	-0.0018	-0.0009	7.0406	-0.0625
	(0.0409)	(0.0072)	(0.0093)	(0.0049)	(0.0050)	(0.5741)	(0.0180)
XOM	0.9316	0.0417	0.0334	0.0029	-0.0031	11.2320	0.0081
	(0.0188)	(0.0063)	(0.0065)	(0.0032)	(0.0033)	(1.3386)	(0.0229)
GE	0.9905	0.0307	0.0366	-0.0093	-0.0065	7.4505	-0.1456
	(0.0015)	(0.0029)	(0.0048)	(0.0023)	(0.0021)	(1.0314)	(0.0300)
HD	0.8657	0.0573	0.0591	-0.0155	-0.0081	7.8864	-0.0342
	(0.0430)	(0.0098)	(0.0114)	(0.0042)	(0.0044)	(0.7215)	(0.0208)
IBM	0.9459	0.0324	0.0186	-0.0097	-0.0066	6.3196	-0.0476
	(0.0279)	(0.0046)	(0.0049)	(0.0021)	(0.0023)	(0.4762)	(0.0903)
INTC	0.9108	0.0322	0.0294	-0.0062	-0.0046	8.7912	-0.0289
	(0.0264)	(0.0059)	(0.0077)	(0.0032)	(0.0033)	(0.8483)	(0.0192)
JNJ	0.7695	0.0648	0.0760	-0.0137	0.0043	7.3108	-0.0319
	(0.0555)	(0.0087)	(0.0109)	(0.0051)	(0.0063)	(0.6096)	(0.0185)
JPM	0.9661	0.0407	0.0367	-0.0108	-0.0026	7.1034	-0.0644
	(0.0065)	(0.0047)	(0.0051)	(0.0027)	(0.0027)	(0.5958)	(0.0333)
MCD	0.7385	0.0545	0.0691	-0.0059	0.0033	7.1915	-0.0341
	(0.0757)	(0.0089)	(0.0121)	(0.0054)	(0.0057)	(0.5940)	(0.0174)
MRK	0.7964	0.0649	0.0799	-0.0090	-0.0066	6.8163	-0.0495
	(0.0443)	(0.0082)	(0.0110)	(0.0051)	(0.0052)	(0.5228)	(0.0190)
MSFT	0.8637	0.0554	0.0675	0.0003	0.0042	8.8889	-0.0320
	(0.0247)	(0.0063)	(0.0085)	(0.0040)	(0.0044)	(0.8994)	(0.0205)
NKE	0.9706	0.0200	0.0187	-0.0062	0.0005	5.8409	-0.0863
	(0.0069)	(0.0023)	(0.0037)	(0.0021)	(0.0016)	(-0.3850)	(0.0303)
PFE	0.9422	0.0316	0.0375	-0.0049	0.0014	8.5675	-0.0273
	(0.0147)	(0.0049)	(0.0064)	(0.0030)	(0.0032)	(0.8417)	(0.0230)
PG	0.7537	0.0369	0.0639	-0.0130	-0.0092	10.1869	0.0162
	(0.0910)	(0.0037)	(0.0080)	(0.0026)	(0.0025)	(1.4107)	(0.0305)
TRV	0.9248	0.0517	0.0502	-0.0078	-0.0002	7.3641	-0.0564
	(0.0154)	(0.0060)	(0.0075)	(0.0035)	(0.0035)	(0.6169)	(0.0248)
UNH	0.9465	0.0178	-0.0073	-0.0275	-0.0209	13.3228	0.2249
	(0.0098)	(0.0071)	(0.0091)	(0.0032)	(0.0036)	(3.6346)	(0.0761)
UTX	0.8583	0.0511	0.0591	-0.0072	-0.0071	8.4000	-0.0176
	(0.0331)	(0.0070)	(0.0092)	(0.0041)	(0.0042)	(0.8074)	(0.0201)
VZ	0.9098	0.0350	0.0528	0.0000	-0.0085	8.9866	-0.0281
	(0.0222)	(0.0059)	(0.0077)	(0.0034)	(0.0035)	(0.8905)	(0.0206)
WMT	0.9397	0.0355	0.0291	-0.0031	-0.0034	7.1862	-0.0342
	(0.0222)	(0.0063)	(0.0081)	(0.0031)	(0.0032)	(0.5837)	(0.0227)
DIS	0.8600	0.0279	0.0593	-0.0066	-0.0017	7.2250	-0.0372
	(0.0350)	(0.0050)	(0.0087)	(0.0040)	(0.0039)	(0.5754)	(0.0175)
Average	0.8793	0.0423	0.0489	-0.0088	-0.0032	8.0879	-0.0321

Continued on the next page.

Table 7: Estimates of variance equations in the multivariate coupled component model

	β_N	γ_N	ρ_N	γ_N^*	ρ_N^*	ν_N	ω_N
MMM	0.8804	0.0483	0.0383	0.0039	-0.0076	2.9933	-1.0124
	(0.0339)	(0.0076)	(0.0073)	(0.0046)	(0.0050)	(0.1300)	(0.0228)
AXP	0.9295	0.0595	0.0444	-0.0066	-0.0102	3.2091	-0.9629
	(0.0138)	(0.0081)	(0.0060)	(0.0042)	(0.0039)	(0.1363)	(0.0271)
AAPL	0.8619	0.0800	0.0798	-0.0005	-0.0141	2.3334	-1.2202
	(0.0196)	(0.0098)	(0.0081)	(0.0055)	(0.0051)	(0.0781)	(0.0241)
BA	0.7390	0.0558	0.0706	-0.0001	-0.0039	2.5584	-1.0268
	(0.0478)	(0.0111)	(0.0095)	(0.0064)	(0.0063)	(0.0950)	(0.0194)
CAT	0.9330	0.0449	0.0496	0.0055	-0.0104	2.4952	-1.0838
	(0.0157)	(0.0084)	(0.0071)	(0.0041)	(0.0039)	(0.0918)	(0.0267)
CVX	0.9692	0.0139	0.0232	-0.0016	-0.0069	3.4551	-0.8353
	(0.0080)	(0.0043)	(0.0038)	(0.0024)	(0.0024)	(0.1676)	(0.0244)
CSCO	0.9394	0.0547	0.0444	0.0003	-0.0138	2.7603	-1.1515
	(0.0137)	(0.0086)	(0.0062)	(0.0037)	(0.0039)	(0.1023)	(0.0283)
KO	0.8793	0.0422	0.0404	-0.0074	-0.0141	3.1641	-0.9414
	(0.0330)	(0.0077)	(0.0079)	(0.0045)	(0.0047)	(0.1354)	(0.0205)
DD	0.8112	0.0585	0.0546	-0.0072	0.0021	3.0657	-0.9734
	(0.0406)	(0.0098)	(0.0083)	(0.0055)	(0.0055)	(0.1278)	(0.0199)
XOM	0.9375	0.0329	0.0330	-0.0023	0.0007	4.0204	-0.8168
	(0.0158)	(0.0065)	(0.0051)	(0.0034)	(0.0034)	(0.2057)	(0.0228)
GE	0.9930	0.0345	0.0393	-0.0055	-0.0079	3.4604	-1.0347
	(0.0009)	(0.0046)	(0.0032)	(0.0023)	(0.0022)	(0.1365)	(0.0408)
HD	0.9001	0.0420	0.0540	-0.0070	-0.0183	2.7342	-1.0060
	(0.0231)	(0.0080)	(0.0080)	(0.0044)	(0.0045)	(0.1052)	(0.0225)
IBM	0.8994	0.0474	0.0490	-0.0028	-0.0150	2.3921	-1.1149
	(0.0642)	(0.0089)	(0.0067)	(0.0037)	(0.0038)	(0.0844)	(0.0833)
INTC	0.8185	0.0799	0.0611	0.0005	-0.0028	2.2863	-1.1842
	(0.0341)	(0.0112)	(0.0089)	(0.0061)	(0.0055)	(0.0772)	(0.0219)
JNJ	0.9383	0.0284	0.0334	-0.0031	-0.0172	3.3619	-0.8759
	(0.0158)	(0.0072)	(0.0066)	(0.0034)	(0.0035)	(0.1521)	(0.0226)
JPM	0.9683	0.0387	0.0391	-0.0103	-0.0093	3.3407	-0.9096
	(0.0063)	(0.0059)	(0.0048)	(0.0029)	(0.0028)	(0.1506)	(0.0350)
MCD	0.8857	0.0353	0.0441	-0.0046	-0.0117	2.9313	-0.9419
	(0.0349)	(0.0093)	(0.0088)	(0.0043)	(0.0043)	(0.1162)	(0.0201)
MRK	0.9148	0.0484	0.0430	-0.0056	-0.0100	2.7916	-1.0097
	(0.0205)	(0.0094)	(0.0081)	(0.0042)	(0.0042)	(0.1113)	(0.0240)
MSFT	0.8783	0.0725	0.0494	0.0040	-0.0022	2.4788	-1.1938
	(0.0251)	(0.0100)	(0.0075)	(0.0049)	(0.0047)	(0.0850)	(0.0233)
NKE	0.8054	0.0584	0.0628	-0.0084	-0.0159	2.2388	-1.1941
	(0.0842)	(0.0122)	(0.0128)	(0.0046)	(0.0062)	(0.0766)	(0.0240)
PFE	0.9505	0.0434	0.0375	-0.0001	0.0008	2.9088	-0.9905
	(0.0113)	(0.0068)	(0.0053)	(0.0036)	(0.0034)	(0.1175)	(0.0285)
PG	0.7464	0.0492	0.0486	-0.0188	-0.0070	3.1622	-0.9233
	(0.0823)	(0.0066)	(0.0060)	(0.0037)	(0.0032)	(0.1351)	(0.0259)
TRV	0.9478	0.0548	0.0408	-0.0060	-0.0065	2.7246	-0.9799
	(0.0105)	(0.0072)	(0.0060)	(0.0036)	(0.0037)	(0.1042)	(0.0300)
UNH	0.8028	0.0429	0.0800	0.0013	-0.0132	3.4955	-0.5698
	(0.0158)	(0.0480)	(0.0489)	(0.0079)	(0.0067)	(0.1653)	(0.0722)
UTX	0.9419	0.0323	0.0361	-0.0080	-0.0104	2.9737	-0.9300
	(0.0131)	(0.0058)	(0.0058)	(0.0033)	(0.0035)	(0.1238)	(0.0246)
VZ	0.9333	0.0429	0.0364	-0.0104	-0.0027	3.1833	-0.8965
	(0.0133)	(0.0071)	(0.0055)	(0.0035)	(0.0035)	(0.1345)	(0.0244)
WMT	0.9604	0.0249	0.0294	-0.0005	-0.0061	2.9159	-0.9766
	(0.0127)	(0.0054)	(0.0055)	(0.0028)	(0.0029)	(0.1159)	(0.0254)
DIS	0.8282	0.0800	0.0708	0.0010	0.0010	2.8703	-1.0500
	(0.0322)	(0.0108)	(0.0094)	(0.0056)	(0.0056)	(0.1147)	(0.0223)
Average	0.8926	0.0481	0.0476	-0.0036	-0.0083	2.9395	-0.9931

This table gives the estimates of the variance equations in the multivariate coupled component model, and their asymptotic standard errors in parenthesis. The last row shows the average estimated values.

Table 7: Estimates of variance equations in the multivariate coupled component model(cont.)

The conditional second order moments conditioning on different information sets are

$$
\begin{aligned}
\operatorname{var}\left[r_t^D|\mathcal{F}_{t-1}, r_t^N\right] &= \frac{v_D}{v_D-2}\exp(2\lambda_t^D + 2\sigma^D(t/T)) \\
\operatorname{var}\left[r_t^N|\mathcal{F}_{t-1}\right] &= \frac{v_N}{v_N-2}\exp(2\lambda_t^N + 2\sigma^N(t/T)) \\
\operatorname{var}\left[r_t^D|\mathcal{F}_{t-1}\right] &= \delta^2\frac{v_N}{v_N-2}\exp(2\lambda_t^N + 2\sigma^N(t/T)) + \frac{v_D}{v_D-2}\exp(2\sigma^D(t/T))E\left[\exp(2\lambda_t^D)|\mathcal{F}_{t-1}\right] \\
\operatorname{cov}\left(r_t^D, r_t^N|\mathcal{F}_{t-1}\right) &= -\delta\frac{v_N}{v_N-2}\exp(2\lambda_t^N + 2\sigma^N(t/T)) \\
\operatorname{corr}\left(r_t^D, r_t^N|\mathcal{F}_{t-1}\right) &= \frac{-\delta\sqrt{\frac{v_N}{v_N-2}}\exp(\lambda_t^N + \sigma^N(t/T))}{\sqrt{\delta^2\frac{v_N}{v_N-2}\exp(2\lambda_t^N + 2\sigma^N(t/T)) + \frac{v_D}{v_D-2}\exp(2\sigma^D(t/T))E\left[\exp(2\lambda_t^D)|\mathcal{F}_{t-1}\right]}} \\
\operatorname{var}\left(r_t^D + r_t^N|\mathcal{F}_{t-1}\right) &= \frac{v_N(1-\delta)^2}{v_N-2}\exp(2\lambda_t^N + 2\sigma^N(t/T)) + \frac{v_D}{v_D-2}\exp(2\sigma^D(t/T))E\left[\exp(2\lambda_t^D)|\mathcal{F}_{t-1}\right]
\end{aligned}
$$

For the unconditional second order moments, we first write the dynamic function of λ_t^j as

$$
\begin{aligned}
\lambda_t^D &= \beta_D^{t-1}\lambda_1^D + \omega_D(1-\beta_D)\sum\nolimits_{k=1}^{t-1}\beta_D^{k-1} + \gamma_D\sum\nolimits_{k=1}^{t-1}\beta_D^{k-1}m_{t-k}^D + \rho_D\sum\nolimits_{k=1}^{t-1}\beta_D^{k-1}m_{t-k+1}^N \\
&\quad + \gamma_D^*\sum\nolimits_{k=1}^{t-1}\beta_D^{k-1}(m_{t-k}^D+1)\operatorname{sign}(e_{t-k}^D) + \rho_D^*\sum\nolimits_{k=1}^{t-1}\beta_D^{k-1}(m_{t-k+1}^N+1)\operatorname{sign}(e_{t-k+1}^N) \\
\lambda_t^N &= \beta_N^{t-1}\lambda_1^N + \omega_N(1-\beta_N)\sum\nolimits_{k=1}^{t-1}\beta_N^{k-1} + \gamma_N\sum\nolimits_{k=1}^{t-1}m_{t-k}^N\beta_N^{k-1} + \rho_N\sum\nolimits_{k=1}^{t-1}\beta_N^{k-1}m_{t-k}^D \\
&\quad + \rho_N^*\sum\nolimits_{k=1}^{t-1}\beta_N^{k-1}(m_{t-k}^D+1)\operatorname{sign}(e_{t-k}^D) + \gamma_N^*\sum\nolimits_{k=1}^{t-1}\beta_N^{k-1}(m_{t-k}^N+1)\operatorname{sign}(e_{t-k}^N)
\end{aligned}
$$

When λ_t^j starts from infinite past,

$$
\begin{aligned}
\lambda_t^D &= \omega_D + \gamma_D\sum\nolimits_{k=1}^{\infty}\beta_D^{k-1}m_{t-k}^D + \rho_D\sum\nolimits_{k=1}^{\infty}\beta_D^{k-1}m_{t-k+1}^N \\
&\quad + \gamma_D^*\sum\nolimits_{k=1}^{\infty}\beta_D^{k-1}(m_{t-k}^D+1)\operatorname{sign}(e_{t-k}^D) + \rho_D^*\sum\nolimits_{k=1}^{\infty}\beta_D^{k-1}(m_{t-k+1}^N+1)\operatorname{sign}(e_{t-k+1}^N) \\
\lambda_t^N &= \omega_N + \gamma_N\sum\nolimits_{k=1}^{\infty}m_{t-k}^N\beta_N^{k-1} + \rho_N\sum\nolimits_{k=1}^{\infty}\beta_N^{k-1}m_{t-k}^D \\
&\quad + \rho_N^*\sum\nolimits_{k=1}^{\infty}\beta_N^{k-1}(m_{t-k}^D+1)\operatorname{sign}(e_{t-k}^D) + \gamma_N^*\sum\nolimits_{k=1}^{\infty}\beta_N^{k-1}(m_{t-k}^N+1)\operatorname{sign}(e_{t-k}^N).
\end{aligned} \tag{15}
$$

The unconditional second order moments $\operatorname{var}\left(u_t^D\right)$ and $\operatorname{var}\left(u_t^N\right)$ are

$$
\operatorname{var}\left(u_t^j\right) = \frac{v_j}{v_j-2}E\exp(2\lambda_t^j)E\exp\left(2\sigma^j(t/T)\right)
$$

with with

$$
\begin{aligned}
&E\exp\left(2\lambda_t^N\right)\\
&=\frac{1}{4}\exp\left(2\omega_N-\frac{2\left(\gamma_N+\rho_N\right)}{1-\beta_N}\right)\\
&\left[\prod_{k=0}^{\infty}{}_1F_1\left(\frac{1}{2},\frac{v^N+1}{2},2\left(\gamma_N+\gamma_N^*\right)\left(v^N+1\right)\beta_N^k\right)+\prod_{k=0}^{\infty}{}_1F_1\left(\frac{1}{2},\frac{v^N+1}{2},2\left(\gamma_N-\gamma_N^*\right)\left(v^N+1\right)\beta_N^k\right)\right]\\
&\left[\prod_{k=0}^{\infty}{}_1F_1\left(\frac{1}{2},\frac{v^D+1}{2},2\left(\rho_N+\rho_N^*\right)\left(v^D+1\right)\beta_N^k\right)+\prod_{k=0}^{\infty}{}_1F_1\left(\frac{1}{2},\frac{v^D+1}{2},2\left(\rho_N-\rho_N^*\right)\left(v^D+1\right)\beta_N^k\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
&E\exp\left(2\lambda_t^D\right)\\
&=\frac{1}{4}\exp\left(2\omega_D-\frac{2\left(\gamma_D+\rho_D\right)}{1-\beta_D}\right)\\
&\left[\prod_{k=0}^{\infty}{}_1F_1\left(\frac{1}{2},\frac{v^D+1}{2},2\left(\gamma_D+\gamma_D^*\right)\left(v^D+1\right)\beta_D^k\right)+\prod_{k=0}^{\infty}{}_1F_1\left(\frac{1}{2},\frac{v^D+1}{2},2\left(\gamma_D-\gamma_D^*\right)\left(v^D+1\right)\beta_D^k\right)\right]\\
&\left[\prod_{k=0}^{\infty}{}_1F_1\left(\frac{1}{2},\frac{v^N+1}{2},2\left(\rho_D+\rho_D^*\right)\left(v^N+1\right)\beta_D^k\right)+\prod_{k=0}^{\infty}{}_1F_1\left(\frac{1}{2},\frac{v^N+1}{2},2\left(\rho_D-\rho_D^*\right)\left(v^N+1\right)\beta_D^k\right)\right].
\end{aligned}
$$

8.2 Appendix 2: proof of asymptotics

8.2.1 Proof of Lemma 1

Denote $H^j(s)=\exp(\sigma^j(s))$. We drop the subscribe j here and have

$$
\begin{aligned}
|u_t| &= H(t/T)\,|e_t| = E\,|e_t|\,H(t/T)+H(t/T)\left(|e_t|-E\,|e_t|\right)\\
\frac{|u_t|}{E\,|e_t|} &= H(t/T)+\frac{H(t/T)}{E\,|e_t|}\left(|e_t|-E\,|e_t|\right)\\
&=: H(t/T)+\xi_t
\end{aligned}
$$

where $E\xi_t=0$. Suppose we know $E\,|e_t|$ .This gives a non-parametric regression function, so we can invoke Nadaraya-Waston estimator

$$
\widetilde{H}(s)^*=\frac{\sum_{t=1}^{T}K_h(s-t/T)\frac{|u_t|}{E|e_t|}}{\sum_{t=1}^{T}K_h(s-t/T)}.
$$

From Lemma 2, $\{e_t\}$ is a β mixing process with exponential decay, and ξ_t thereby is also a β mixing process with exponential decay. Invoking Theorem 3 in Vogt and Linton (2014), Theorem

4.1 in Vogt et al. (2012) or Kristensen (2009) yields

$$\sup_{s\in[C_1h,1-C_1h]}\left|\widetilde{H}(s)^{*}-H_0(s)\right|=O_p\left(\sqrt{\frac{\log T}{Th}}+h^2\right).$$

Denote $\widetilde{\sigma}(s)^{*}=\log(\widetilde{H}(s)^{*})$. Taylor expansion at $H_0(s)$ gives

$$\widetilde{\sigma}(s)^{*}=\sigma(s)+\left(\widetilde{H}(s)^{*}-H(s)\right)\log' H(s)+\frac{1}{2}\left(\widetilde{H}(s)^{*}-H(s)\right)^2\log'' H(s)^{**},$$

with $H(s)^{**}$ between $\widetilde{H}(s)^{*}$ and $H_0(s)$. Therefore,

$$\sup_{s\in[C_1h,1-C_1h]}\left|\widetilde{\sigma}(s)^{*}-\sigma_0(s)\right|=O_p\left(h^2+\sqrt{\frac{\log T}{Th}}\right).$$

For $s\in[0,h]\cup[1-h,1]$, we use a boundary kernel to ensure the bias property holds through $[0,1]$.

Until now we have obtained the property for the un-rescaled estimator $\widetilde{\sigma}(s)^{*}$. Next, we are going to show the convergence rate of the rescaled estimator $\widetilde{\sigma}(s)$. Recall that

$$\widetilde{\sigma}(s)=\widetilde{\sigma}(s)-\frac{1}{T}\sum_{t=1}^{T}\widetilde{\sigma}(\frac{t}{T}),$$

and we can rewrite $\widetilde{\sigma}(s)$ as

$$\widetilde{\sigma}(s)=\widetilde{\sigma}(s)^{*}-\frac{1}{T}\sum_{t=1}^{T}\widetilde{\sigma}(\frac{t}{T})^{*},$$

as $E\left|e_t\right|$ in $\widetilde{\sigma}(s)^*$ vanished by the rescaling. Plugging this into $\sup_{s\in[C_1h,1-C_1h]}\left|\widetilde{\sigma}(s)-\sigma_0(s)\right|$ gives

$$\begin{aligned}
&\sup_{s\in[0,1]}\left|\widetilde{\sigma}(s)-\sigma_0(s)\right| \\
&= \sup_{s\in[0,1]}\left|\widetilde{\sigma}(s)^* - \frac{1}{T}\sum_{t=1}^{T}\widetilde{\sigma}(\frac{t}{T})^* - \sigma_0(s)\right| \\
&= \sup_{s\in[0,1]}\left|\widetilde{\sigma}(s)^* - \frac{1}{T}\sum_{t=1}^{T}\widetilde{\sigma}(\frac{t}{T})^* - \sigma_0(s) - \frac{1}{T}\sum_{t=1}^{T}\sigma_0(\frac{t}{T}) + \frac{1}{T}\sum_{t=1}^{T}\sigma_0(\frac{t}{T})\right| \\
&\leq \sup_{s\in[0,1]}\left|\widetilde{\sigma}(s)^* - \sigma_0(s)\right| + \left|\frac{1}{T}\sum_{t=1}^{T}\left(\widetilde{\sigma}(\frac{t}{T})^* - \sigma_0(\frac{t}{T})\right)\right| + \left|\frac{1}{T}\sum_{t=1}^{T}\sigma_0(\frac{t}{T})\right| \\
&= O_p\left(h^2+\sqrt{\frac{\log T}{Th}}\right) + O_p\left(h^2+\sqrt{\frac{\log T}{Th}}\right) + \left|\frac{1}{T}\sum_{t=1}^{T}\sigma_0(\frac{t}{T})\right| \\
&= O_p\left(h^2+\sqrt{\frac{\log T}{Th}}\right) + \left|\frac{1}{T}\sum_{t=1}^{T}\sigma_0(\frac{t}{T})\right|.
\end{aligned}$$

We only have to work out the second term $\left|\frac{1}{T}\sum_{t=1}^{T}\sigma_0(\frac{t}{T})\right|$. According to Theorem 1.3 in Tasaki (2009),

$$\lim_{T\to\infty} T^2\left(\int_0^1 \sigma_0(s)ds - \frac{1}{2T}\sum_{t=1}^{T}\sigma_0(\frac{t}{T}) - \frac{1}{2T}\sum_{t=0}^{T-1}\sigma_0(\frac{t}{T})\right) = -\frac{1}{12}\left(\sigma_0'(1)-\sigma_0'(0)\right).$$

Since $\int_0^1\sigma_0(s)ds=0$ and $\sigma_0'(1)-\sigma_0'(0)$ is bounded by Assumption A4, it follows

$$\left|\frac{1}{2T}\sum_{t=1}^{T}\sigma_0(\frac{t}{T}) + \frac{1}{2T}\sum_{t=0}^{T-1}\sigma_0(\frac{t}{T})\right| = O(T^{-2})$$

and

$$\begin{aligned}
\left|\frac{1}{T}\sum_{t=1}^{T}\sigma_0(\frac{t}{T})\right| &\leq \left|\frac{1}{2T}\sum_{t=1}^{T}\sigma_0(\frac{t}{T}) + \frac{1}{2T}\sum_{t=0}^{T-1}\sigma_0(\frac{t}{T})\right| + \left|\frac{1}{2T}\sum_{t=1}^{T}\sigma_0(\frac{t}{T}) - \frac{1}{2T}\sum_{t=0}^{T-1}\sigma_0(\frac{t}{T})\right| \\
&= O(T^{-2}) + \frac{1}{2T}\left|\sigma_0(1)-\sigma_0(0)\right| \\
&= O(T^{-1}).
\end{aligned}$$

Therefore, the uniform convergence rate is

$$\sup_{s\in[0,1]} |\widetilde{\sigma}(s) - \sigma_0(s)| = O_p\left(h^2 + \sqrt{\frac{\log T}{Th}}\right) + O(T^{-1})$$
$$= O_p\left(h^2 + \sqrt{\frac{\log T}{Th}}\right).$$

8.2.2 Proof of Theorem 1

Let $\phi_i = \beta_D$ and θ_k be an element in function $\sigma^D(\cdot)$ (for simplicity, the subscript k is omitted in what follows). Recall that $h_t^j = \lambda_t^j + \sigma^j(t/T)$, and the log-likelihood function, without unnecessary constant, can be rewritten as a function of h_t^j

$$l_t^j = -h_t^j - \frac{v_j+1}{2}\ln\left(1 + \frac{(u_t^j)^2}{v_j \exp(2h_t^j)}\right) + \ln\Gamma\left(\frac{v_j+1}{2}\right) - \frac{1}{2}\ln v_j - \ln\Gamma\left(\frac{v_j}{2}\right)$$

with the score functions

$$\frac{\partial l_t}{\partial \theta} = \frac{\partial l_t^D}{\partial h_t^D}\frac{\partial h_t^D}{\partial \theta} + \frac{\partial l_t^N}{\partial h_t^D}\frac{\partial h_t^D}{\partial \theta} = m_t^D\frac{\partial h_t^D}{\partial \theta} + m_t^N\frac{\partial h_t^D}{\partial \theta}$$
$$\frac{\partial l_t}{\partial \beta_D} = \frac{\partial l_t^D}{\partial h_t^D}\frac{\partial h_t^D}{\partial \beta_D} = m_t^D\frac{\partial h_t^D}{\partial \beta_D} + m_t^N\frac{\partial h_t^D}{\partial \beta_D}.$$

Recall that $m_t^j = (v_j+1)b_t^j - 1$,with b_t^j independent and identically beta distributed, we have $E\left(m_t^N m_t^D\right) = 0$, $E\left(m_t^j\right)^2$ is time invariant, and $E\left(m_t^j\right)^2 < \infty$. Therefore, we can write

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} E\frac{\partial l_t}{\partial \theta}\frac{\partial l_t}{\partial \beta_D} = \lim_{T\to\infty}\frac{1}{T}E\left(m_t^D\right)^2\sum_{t=1}^{T}\frac{\partial h_t^D}{\partial \theta}\frac{\partial h_t^D}{\partial \beta_D}$$
$$+ \lim_{T\to\infty}\frac{1}{T}E\left(m_t^N\right)^2\sum_{t=1}^{T}E\frac{\partial h_t^N}{\partial \theta}\frac{\partial h_t^N}{\partial \beta_D}.$$

To prove the theorem, it then suffices to show that

$$\lim_{T\to\infty}\left\|\frac{1}{T}\sum_{t=1}^{T}E\begin{pmatrix}\frac{\partial h_t^D}{\partial \theta}\\ \frac{\partial h_t^N}{\partial \theta}\end{pmatrix}\begin{pmatrix}\frac{\partial h_t^D}{\partial \beta_D} & \frac{\partial h_t^N}{\partial \beta_D}\end{pmatrix}\right\|_\infty = 0.$$

By expressing λ_t^j as a function of φ and $\left\{\left(\sigma^j(\frac{t-i}{T}), u_{t-i}^j\right), i \geq 0\right\}$, we can write $\frac{\partial h_t^j}{\partial \theta}$ as

$$\begin{aligned}\begin{pmatrix}\frac{\partial h_t^D}{\partial \theta}\\ \frac{\partial h_t^N}{\partial \theta}\end{pmatrix} &= \sum_{k=0}^{T}\begin{pmatrix}\frac{\partial h_t^D}{\partial \sigma^D(\frac{t-k}{T})}\frac{\partial \sigma^D(\frac{t-k}{T})}{\partial \theta}\\ \frac{\partial h_t^N}{\partial \sigma^D(\frac{t-k}{T})}\frac{\partial \sigma^D(\frac{t-k}{T})}{\partial \theta}\end{pmatrix}\\ &= \sum_{k=0}^{T}\begin{pmatrix}\frac{\partial h_t^D}{\partial \sigma^D(\frac{t-k}{T})}\\ \frac{\partial h_t^N}{\partial \sigma^D(\frac{t-k}{T})}\end{pmatrix}\psi_i^D\left(\frac{t-k}{T}\right),\end{aligned}$$

when the limit exists. We obtain,

$$\begin{aligned}&\frac{1}{T}\sum_{t=1}^{T}E\begin{pmatrix}\frac{\partial h_t^D}{\partial \theta}\\ \frac{\partial h_t^N}{\partial \theta}\end{pmatrix}\begin{pmatrix}\frac{\partial h_t^D}{\partial \beta_D} & \frac{\partial h_t^N}{\partial \beta_D}\end{pmatrix}\\ &= \frac{1}{T}\sum_{k=0}^{T}\sum_{t=1}^{T}E\begin{pmatrix}\frac{\partial h_t^D}{\partial \sigma^D(\frac{t-k}{T})}\\ \frac{\partial h_t^N}{\partial \sigma^D(\frac{t-k}{T})}\end{pmatrix}\begin{pmatrix}\frac{\partial h_t^D}{\partial \beta_D} & \frac{\partial h_t^N}{\partial \beta_D}\end{pmatrix}\psi_i^D\left(\frac{t-k}{T}\right)\\ &= \frac{1}{T}\sum_{k=0}^{T}E\begin{pmatrix}\frac{\partial h_t^D}{\partial \sigma^D(\frac{t-k}{T})}\\ \frac{\partial h_t^N}{\partial \sigma^D(\frac{t-k}{T})}\end{pmatrix}\begin{pmatrix}\frac{\partial h_t^D}{\partial \beta_D} & \frac{\partial h_t^N}{\partial \beta_D}\end{pmatrix}\sum_{t=1}^{T}\psi_i^D\left(\frac{t-k}{T}\right).\end{aligned}$$

The second equality follows since $E\begin{pmatrix}\frac{\partial h_t^D}{\partial \sigma^j(\frac{t-k}{T})}\\ \frac{\partial h_t^N}{\partial \sigma^j(\frac{t-k}{T})}\end{pmatrix}\begin{pmatrix}\frac{\partial h_t^D}{\partial \beta_D} & \frac{\partial h_t^N}{\partial \beta_D}\end{pmatrix}$ is invariant across time t by Lemma 4. Taylor expansion of $\sum_{t=1}^{T}\psi_i^D\left(\frac{t-k}{T}\right)$ around $\sum_{t=1}^{T}\psi_i^D\left(\frac{t-k}{T}\right)$ gives

$$\begin{aligned}\frac{1}{T}\sum_t \psi_i^D\left(\frac{t-k}{T}\right) &= \frac{1}{T}\sum_t \psi_i^D\left(\frac{t}{T}\right) - \frac{1}{T}\frac{k}{T}\sum_t \psi_i^{D\prime}\left(\frac{t}{T}\right) + O\left(\frac{k}{T}\right)^2\\ &= O\left(\frac{1}{T}\right) + O\left(\frac{k}{T}\right) + O\left(\frac{k}{T}\right)^2\\ &= O\left(\frac{k}{T}\right)\end{aligned}$$

Hence, it suffices to show $\sum_{k=0}^{T}\left\|k\left(E\begin{pmatrix}\frac{\partial h_t^D}{\partial \sigma^D(\frac{t-k}{T})}\\ \frac{\partial h_t^N}{\partial \sigma^D(\frac{t-k}{T})}\end{pmatrix}\begin{pmatrix}\frac{\partial h_t^D}{\partial \beta_D} & \frac{\partial h_t^N}{\partial \beta_D}\end{pmatrix}\right)\right\|_\infty < \infty$, which is obtained by Lemma 3.

The proof with respect to v_D is similar, but the score function is slightly different. The score

functions of l_t^D and l_t^N with respect to v_D are

$$\begin{aligned}
\frac{\partial l_t^D}{\partial v_D} &= -\frac{1}{2}\ln\left(1+\frac{(u_t^D)^2}{v_D\exp(2h_t^D)}\right)+\frac{\partial}{\partial v_D}\left(\ln\Gamma\left(\frac{v_D+1}{2}\right)-\ln\Gamma\left(\frac{v_D}{2}\right)\right)-\frac{1}{2v_D} \\
&\quad+\frac{v_D+1}{2\left(1+\frac{(u_t^D)^2}{v_D\exp(2h_t^D)}\right)}\frac{(u_t^D)^2}{v_D^2\exp(2h_t^D)}\left(1+2v_D\frac{\partial h_t^D}{\partial v_D}\right)+\frac{\partial h_t^D}{\partial v_D} \\
\frac{\partial l_t^N}{\partial v_D} &= \frac{v_N+1}{2\left(1+\frac{(u_t^N)^2}{v_N\exp(2h_t^N)}\right)}\frac{(u_t^N)^2}{v_N^2\exp(2h_t^N)}\left(1+2v_N\frac{\partial h_t^N}{\partial v_D}\right)+\frac{\partial h_t^N}{\partial v_D}.
\end{aligned} \tag{16}$$

Then we can have

$$\begin{aligned}
&\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E\frac{\partial l_t^D}{\partial\theta}\frac{\partial l_t^D}{\partial v_D} \\
&=\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}Em_t^D\frac{\partial h_t^D}{\partial\theta}\left[\frac{\partial h_t^D}{\partial v_D}-\frac{1}{2}\ln\left(1+\frac{(u_t^D)^2}{v_D\exp(2h_t^D)}\right)\right] \\
&+\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}Em_t^D\frac{\partial h_t^D}{\partial\theta}\frac{\partial\ln\Gamma\left(\frac{v_D+1}{2}\right)-\frac{1}{2}\ln v_D-\ln\Gamma\left(\frac{v_D}{2}\right)}{\partial v_D} \\
&+\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}Em_t^D\frac{\partial h_t^D}{\partial\theta}\left[\frac{v_D+1}{2\left(1+\frac{(u_t^D)^2}{v_D\exp(2h_t^D)}\right)}\frac{(u_t^D)^2}{v_D^2\exp(2h_t^D)}\left(1+2v_D\frac{\partial h_t^D}{\partial v_D}\right)\right] \\
&=\lim_{T\to\infty}\frac{1}{2}E\left(m_t^D\left(-\ln\left(1+\frac{(\varepsilon_t^D)^2}{v_D}\right)+\frac{v_D+1}{2v_D+(\varepsilon_t^D)^2}\frac{(\varepsilon_t^D)^2}{v_D}\right)\right)\frac{1}{T}\sum_{t=1}^{T}E\frac{\partial h_t^D}{\partial\theta} \\
&+\lim_{T\to\infty}E\left(m_t^D\left(1+\frac{(v_D+1)(\varepsilon_t^D)^2}{2v_D+(\varepsilon_t^D)^2}\right)\right)\frac{1}{T}\sum_{t=1}^{T}E\frac{\partial h_t^D}{\partial v_D}\frac{\partial h_t^D}{\partial\theta}.
\end{aligned}$$

The first term vanishes by lemma 5. Then we can use the same procedure above to obtain $\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E\frac{\partial h_t^D}{\partial v_D}\frac{\partial h_t^D}{\partial\theta}=0$, and to finish the proof for v_D.

8.2.3 Proof of Theorem 2

Harvey (2013) gives the consistency and asymptotic normality of the MLE estimator in the beta-t-egarch model. The key point is the first three derivatives of l_t with respect to ϕ (except v_j) are linear combinations of $b_t^h(1-b_t)^k$, h, $k=0,1,2,\ldots$, with $b_t=\frac{(1+v)(e_t)^2}{v\exp(2\lambda_t)+(e_t)^2}$. Since b_t is beta distributed, these first three derivatives are all bounded. It is then straightforward to show that the score function satisfies a CLT, and its derivative converges to the information matrix by the ergodic theorem.

In the information matrix, we only provide the result for $E\frac{\partial l_t}{\partial \beta_D}\frac{\partial l_t}{\partial \beta_N}$, and without leverage effects $(\gamma_D^*, \gamma_N^*, \rho_D^*, \rho_N^* = 0)$. The generalization to other elements or to the case with leverage effects is straightforward. We have

$$E\frac{\partial l_t}{\partial \beta_D}\frac{\partial l_t}{\partial \beta_N} = E\left(m_t^D\right)^2 E\frac{\partial h_t^D}{\partial \beta_D}\frac{\partial h_t^D}{\partial \beta_N} + E\left(m_t^N\right)^2 E\frac{\partial h_t^N}{\partial \beta_D}\frac{\partial h_t^N}{\partial \beta_N},$$

with

$$\begin{pmatrix} \frac{\partial}{\partial \beta_D} h_t^D \\ \frac{\partial}{\partial \beta_D} h_t^N \end{pmatrix} = \sum_{i=1}^{\infty} A_t \left(\prod_{k=1}^{i-1} B_{t-k}A_{t-k}\right) \begin{pmatrix} \lambda_{t-i}^D - \omega_D \\ 0 \end{pmatrix} \tag{17}$$

$$\begin{pmatrix} \frac{\partial}{\partial \beta_N} h_t^D \\ \frac{\partial}{\partial \beta_N} h_t^N \end{pmatrix} = \sum_{j=1}^{\infty} A_t \left(\prod_{k=1}^{j-1} B_{t-k}A_{t-k}\right) \begin{pmatrix} 0 \\ \lambda_{t-1}^N - \omega_N \end{pmatrix}. \tag{18}$$

We can write $\lambda_{t-i}^D - \omega_D$ and $\lambda_{t-j}^N - \omega_N$

$$\lambda_{t-i}^D - \omega_D = \gamma_D \sum\nolimits_{k=1}^{j-i} \beta_D^{k-1} m_{t-k-i}^D + \rho_D \sum\nolimits_{k=1}^{j-i} \beta_D^{k-1} m_{t-k+1-i}^N + \beta_D^{j-i}\left(\lambda_{t-j}^D - \omega_D\right), \text{ when } i < j$$

$$\lambda_{t-j}^N - \omega_N = \gamma_N \sum\nolimits_{k=1}^{i-j} m_{t-k}^N \beta_N^{k-1} + \rho_N \sum\nolimits_{k=1}^{i-j} \beta_N^{k-1} m_{t-k}^D + \beta_N^{i-j}\left(\lambda_{t-i}^N - \omega_N\right), \text{ when } i > j.$$

When $i < j$, taking the expectation of the cross product of the ith term in (17) and the jth term in (18) gives

$$\begin{aligned} &\mathrm{vec} E A_t \prod_{k=1}^{i-1} B_{t-k}A_{t-k} \begin{pmatrix} \lambda_{t-i}^D - \omega_D \\ 0 \end{pmatrix} \begin{pmatrix} 0 & \lambda_{t-j}^N - \omega_N \end{pmatrix} \left(A_t \prod_{k=1}^{j-1} B_{t-k}A_{t-k}\right)^{\intercal} \\ &= E\left(A_t \otimes A_t\right) g^{i-1} \mathrm{vec}\left(E \begin{pmatrix} \lambda_{t-i}^D - \omega_D \\ 0 \end{pmatrix} \begin{pmatrix} 0 & \lambda_{t-j}^N - \omega_N \end{pmatrix} \left(A_t \prod_{k=i}^{j-1} B_{t-k}A_{t-k}\right)\right), \end{aligned}$$

with $g = \left(E\left(B_{t-1} \otimes B_{t-1}\right)\left(A_{t-1} \otimes A_{t-1}\right)\right)$, and

$$\begin{aligned} &E \begin{pmatrix} \lambda_{t-i}^D - \omega_D \\ 0 \end{pmatrix} \begin{pmatrix} 0 & \lambda_{t-j}^N - \omega_N \end{pmatrix} \left(A_t \prod_{k=i}^{j-1} B_{t-k}A_{t-k}\right) \\ &= E \begin{pmatrix} \gamma_D \sum_{k=1}^{j-i} \beta_D^{k-1} m_{t-k-i}^D + \rho_D \sum_{k=1}^{j-i} \beta_D^{k-1} m_{t-k+1-i}^N + \beta_D^{j-i}\left(\lambda_{t-j}^D - \omega_D\right) \\ 0 \end{pmatrix} \\ &\times \begin{pmatrix} 0 & \lambda_{t-j}^N - \omega_N \end{pmatrix} \left(A_t \prod_{k=i}^{j-1} B_{t-k}A_{t-k}\right) \\ &= E \begin{pmatrix} \beta_D^{j-i}\left(\lambda_{t-j}^D - \omega_D\right) \\ 0 \end{pmatrix} \begin{pmatrix} 0 & \lambda_{t-j}^N - \omega_N \end{pmatrix} \left(A_t \prod_{k=i}^{j-1} B_{t-k}A_{t-k}\right). \end{aligned}$$

The second equality follows since λ_{t-j}^N is independent of $\left\{\left(A_{t-s}, B_{t-s}, m_{t-s}^D, m_{t-s}^N\right),\ s \leq j\right\}$, and $E(\lambda_{t-j}^N - \omega_N) = 0$. Therefore,

$$\begin{aligned}
&\mathrm{vec} E A_t \prod_{k=1}^{i-1} B_{t-k} A_{t-k} \begin{pmatrix} \lambda_{t-i}^D - \omega_D \\ 0 \end{pmatrix} \begin{pmatrix} 0 & \lambda_{t-j}^N - \omega_N \end{pmatrix} \left(A_t \prod_{k=1}^{j-1} B_{t-k} A_{t-k} \right)^{\intercal} \\
&= E\left(A_t \otimes A_t\right) g^{i-1} \beta_D^{j-i} \left(\left(E A_t B_t\right) \otimes I\right)^{j-i} \mathrm{vec} \left(\begin{bmatrix} 0 & E\left(\lambda_{t-j}^D - \omega_D\right)\left(\lambda_{t-j}^N - \omega_N\right) \\ 0 & 0 \end{bmatrix} \right).
\end{aligned}$$

Similarly, when $i > j$,

$$\begin{aligned}
&\mathrm{vec} E A_t \prod_{k=1}^{i-1} B_{t-k} A_{t-k} \begin{pmatrix} \lambda_{t-i}^D - \omega_D \\ 0 \end{pmatrix} \begin{pmatrix} 0 & \lambda_{t-j}^N - \omega_N \end{pmatrix} \left(A_t \prod_{k=1}^{j-1} B_{t-k} A_{t-k} \right)^{\intercal} \\
&= E\left(A_t \otimes A_t\right) g^{j-1} {\beta_N}^{i-j-1} \gamma_N \rho_D \left(I \otimes \left(E B_t A_t\right)\right)^{i-j} \mathrm{vec} \begin{bmatrix} 0 & E m_{t-i}^N m_{t-i}^N \\ 0 & 0 \end{bmatrix} \\
&+ E\left(A_t \otimes A_t\right) g^{j-1} {\beta_N}^{i-j} \left(I \otimes \left(E B_t A_t\right)\right)^{i-j} \mathrm{vec} \begin{bmatrix} 0 & E\left(\lambda_{t-i}^N - \omega_N\right)\left(\lambda_{t-i}^D - \omega_D\right) \\ 0 & 0 \end{bmatrix},
\end{aligned}$$

as for any t, λ_{t-i}^D is independent of $\left\{\left(A_{t-s}, B_{t-s}, m_{t-s}^D\right),\ s \leq i\right\}$ and $\left\{m_{t-s+1}^N,\ s \leq i\right\}$. Finally, we obtain

$$\begin{aligned}
&\mathrm{vec} E \begin{pmatrix} \frac{\partial}{\partial \beta_D} h_t^D \\ \frac{\partial}{\partial \beta_D} h_t^N \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial \beta_N} h_t^D \\ \frac{\partial}{\partial \beta_N} h_t^N \end{pmatrix}^{\intercal} \\
&= \sum_{i=1}^{\infty} \sum_{j=i}^{\infty} E\left(A_t \otimes A_t\right) g^{i-1} \beta_D^{j-i} \left(\left(E A_t B_t\right) \otimes I\right)^{j-i} \mathrm{vec} \left(\begin{bmatrix} 0 & E\left(\lambda_{t-j}^D - \omega_D\right)\left(\lambda_{t-j}^N - \omega_N\right) \\ 0 & 0 \end{bmatrix} \right) \\
&+ \sum_{i=1}^{\infty} \sum_{j=1}^{i-1} E\left(A_t \otimes A_t\right) g^{j-1} {\beta_N}^{i-j-1} \gamma_N \rho_D \left(I \otimes E A_t B_t\right)^{i-j} \mathrm{vec} \left(\begin{bmatrix} 0 & E m_{t-i}^N m_{t-i}^N \\ 0 & 0 \end{bmatrix} \right) \\
&+ \sum_{i=1}^{\infty} \sum_{j=1}^{i-1} E\left(A_t \otimes A_t\right) g^{j-1} {\beta_N}^{i-j} \left(I \otimes E A_t B_t\right)^{i-j} \mathrm{vec} \left(\begin{bmatrix} 0 & E\left(\lambda_{t-i}^N - \omega_N\right)\left(\lambda_{t-i}^D - \omega_D\right) \\ 0 & 0 \end{bmatrix} \right) \\
&= E\left(A_t \otimes A_t\right) \left(I - g\right)^{-1} \left(I - \beta_D \left(\left(E A_t B_t\right) \otimes I\right)\right)^{-1} \mathrm{vec} \begin{bmatrix} 0 & E\left(\lambda_t^D - \omega_D\right)\left(\lambda_t^N - \omega_N\right) \\ 0 & 0 \end{bmatrix} \\
&+ \frac{\gamma_N \rho_D}{\beta_N} E\left(A_t \otimes A_t\right) \left(I - g\right)^{-1} \left(\left(I - \beta_N \left(I \otimes E A_t B_t\right)\right)^{-1} - I\right) \mathrm{vec} \begin{bmatrix} 0 & E m_t^N m_t^N \\ 0 & 0 \end{bmatrix} \\
&+ E\left(A_t \otimes A_t\right) \left(I - g\right)^{-1} \left(\left(I - \beta_N \left(I \otimes E A_t B_t\right)\right)^{-1} - I\right) \mathrm{vec} \begin{bmatrix} 0 & E\left(\lambda_t^N - \omega_N\right)\left(\lambda_t^D - \omega_D\right) \\ 0 & 0 \end{bmatrix},
\end{aligned}$$

with $Em_t^N m_t^N = \frac{2v_N}{v_N+3}$, and $E\left(\lambda_t^D - \omega_D\right)\left(\lambda_t^N - \omega_N\right) = \frac{\beta_D\gamma_N\rho_D}{1-\beta_N\beta_D}\frac{2v_N}{v_N+3} + \frac{\gamma_D\rho_N}{1-\beta_N\beta_D}\frac{2v_D}{v_D+3}$.

The consistency and asymptotic normality also hold for $\widehat{v}$ following Harvey (2013).

8.2.4 Proof of Theorem 3

Consider the local likelihood function given η_t^j and v_j, i.e., minimize the objective function

$$L_T^j(\sigma^j; s) = \frac{1}{T}\sum_{t=1}^{T} K_h(s - t/T)\left[\sigma^j + \frac{v_j+1}{2}\ln\left(1 + \frac{(\eta_t^j \exp(-\sigma^j))^2}{v_j}\right)\right]$$

with respect to ω, for $j = D, N$ separately. The first order and second order derivatives are

$$\begin{aligned}
\frac{\partial L_T^j(\sigma^j; s)}{\partial \sigma^j} &= \frac{1}{T}\sum_{t=1}^{T} K_h(s - t/T)\left[-(v_j+1)b_t^j(\sigma^j) + 1\right] \\
\frac{\partial^2 L_T^j(\sigma^j; s)}{\partial \sigma^{j2}} &= 2(v_j+1)\frac{1}{T}\sum_{t=1}^{T} K_h(s - t/T)\left[b_t^j(\sigma^j)\left(1 - b_t^j(\sigma^j)\right)\right] \qquad (19)
\end{aligned}$$

$$b_t^j(\sigma^j) = \frac{\frac{(\eta_t^j)^2}{v_j}}{\exp(2\sigma^j) + \frac{(\eta_t^j)^2}{v_j}}.$$

We have

$$\sqrt{Th}\left(\widehat{\sigma^j}(s) - \sigma_0^j(s)\right) = \left[\frac{1}{Th}\frac{\partial^2 L_T^j(\sigma_0^j; s)}{\partial \sigma^{j2}}\right]^{-1}\frac{1}{\sqrt{Th}}\frac{\partial L_T^j(\sigma_0^j; s)}{\partial \sigma^j} + o_p(1),$$

when the t distribution is correct (likelihood theory). This is asymptotically normal with mean zero and variance

$$\mathrm{var}\left[\frac{1}{\sqrt{Th}}\frac{\partial L_T^j(\sigma_0^j; s)}{\partial \sigma^j}\right] = ||K||_2^2 E\left[\left(1 - (v_j+1)b_t^j(\sigma_0^j(s))\right)^2\right]_{t/T=s}.$$

This follows because

$$E\left[\left(1 - (v_j+1)b_t^j(\sigma_0^j(s))\right)^2\right] = f(t/T)$$

for some smooth function f, and recall $\eta_t^j = \exp(\sigma^j(t/T))\varepsilon_t^j$. Denote $||K||_2^2 = \int K(s)^2 \mathrm{d}s$, since $\sum_{t=1}^T K^2(\frac{s-t/T}{h})\frac{1}{Th} \to \int K(s)^2\mathrm{d}s$, $h^2\sum_{t=1}^T K_h^2(s - t/T)\frac{1}{Th} = ||K||_2^2$. It follows that

$$\frac{h^2}{Th}\sum_{t=1}^{T} K_h^2(s - t/T)f(t/T) \to ||K||_2^2 f(s),$$

Therefore,

$$\sqrt{Th}\left(\widehat{\sigma^j}(s)-\sigma_0^j(s)\right) \Longrightarrow N\left(0, \frac{||K||_2^2}{E\left[\left(1-(v_j+1)b_t^j\right)^2\right]_{t/T=s}}\right)$$

Further, since b_t^j is distributed as $beta(\frac{1}{2}, \frac{v_j}{2})$, with

$$E\left[\left(1-(v_j+1)b_t^j\right)^2\right]_{t/T=s} = \frac{2v_j}{(v_j+3)}.$$

It thus follows that

$$\sqrt{Th}\left(\widehat{\sigma^j}(s-\sigma_0^j(s)\right) \Longrightarrow N\left(0, \sqrt{\frac{(v_j+3)}{2v_j}||K||_2^2}\right).$$

when the t distribution is correct (likelihood theory). This is asymptotically normal with mean zero and variance

8.2.5 Other Lemmas

Lemma 2 *If $|\beta_D| < 1$ and $|\beta_N| < 1$, e_t^j and λ_t^j are β-mixing with exponential decay.*

Proof. For simplicity, we consider the model without leverage effects

$$\begin{aligned}\lambda_t^D &= \omega_D(1-\beta_D) + \beta_D\lambda_{t-1}^D + \gamma_D m_{t-1}^D + \rho_D m_t^N \\ \lambda_t^N &= \omega_N(1-\beta_N) + \beta_N\lambda_{t-1}^N + \gamma_N m_{t-1}^N + \rho_N m_{t-1}^D.\end{aligned}$$

Let us write it as

$$\begin{pmatrix}\lambda_t^D \\ \lambda_t^N \\ m_t^D \\ m_t^N\end{pmatrix} = \begin{pmatrix}\beta_D & 0 & \gamma_D & 0 \\ 0 & \beta_N & \rho_N & \beta_N \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}\lambda_{t-1}^D \\ \lambda_{t-1}^N \\ m_{t-1}^D \\ m_{t-1}^N\end{pmatrix} + \begin{pmatrix}\rho_D m_t^N + \omega_D(1-\beta_D) \\ \omega_N(1-\beta_N) \\ m_t^D \\ m_t^N\end{pmatrix}.$$

Since m_t^N and m_t^D are i.i.d random variables and follow a beta distribution, we can easily find an integer $s \geq 1$ to satisfy $E\left|\begin{matrix}\rho_D m_t^N + \omega_D(1-\beta_D) \\ \omega_N(1-\beta_N) \\ m_t^D \\ m_t^N\end{matrix}\right|^s < \infty$ (Condition A_2 in Carrasco and Chen

(2002)). The largest eigenvalue of the matrix $\begin{vmatrix} \beta_D & 0 & \gamma_D & 0 \\ 0 & \beta_N & \rho_N & \beta_N \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{vmatrix}$ is smaller than 1 by assumption.

Define $X_t = \begin{pmatrix} \lambda_t^D & \lambda_t^N & m_t^D & m_t^N \end{pmatrix}^\intercal$. According to Proposition 2 in Carrasco and Chen (2002), the process X_t is Markov geometrically ergodic and $E\left|X_t\right|^s < \infty$. Moreover, if X_t is initialized from the invariant distribution, it is then strictly stationary and β-mixing with exponential decay. The process $\{e_t^j\}$ is a generalized hidden Markov model and β-mixing with a decay rate at least as fast as that of $\{\lambda_t^j\}$ by Proposition 4 in Carrasco and Chen (2002). The extension to the model with leverage effects is straightforward, by defining $X_t = \begin{pmatrix} \lambda_t^D & \lambda_t^N & m_t^D & m_t^N & \operatorname{sign}\left(e_t^D\right) & \operatorname{sign}\left(e_t^N\right) \end{pmatrix}^\intercal$. ■

Lemma 3 *Under Assumption A1-A4, it holds that*

$$\sum_k k \left\| E\left(\begin{bmatrix} \frac{\partial h_t^D}{\partial \sigma^D (t-k/T)} \\ \frac{\partial h_t^N}{\partial \sigma^D (t-k/T)} \end{bmatrix} \begin{pmatrix} \frac{\partial}{\partial \beta_D} h_t^D & \frac{\partial}{\partial \beta_D} h_t^N \end{pmatrix} \right) \right\|_\infty < \infty.$$

Proof. By (22) and (23), we have

$$E \begin{pmatrix} \frac{\partial h_{t+1}^D}{\partial \sigma^D (t+k/T)} \\ \frac{\partial h_{t+1}^N}{\partial \sigma^D (t+k/T)} \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial \beta_D} h_{t+1}^D & \frac{\partial}{\partial \beta_D} h_{t+1}^N \end{pmatrix} = E A_{t+1} \begin{pmatrix} a_t^{DD} \\ a_t^{ND} \end{pmatrix} \begin{pmatrix} \lambda_t^D - \omega_D & 0 \end{pmatrix} A_{t+1}^T; k = 1$$

$$E \begin{pmatrix} \frac{\partial h_{t+1}^D}{\partial \sigma^D (t+k/T)} \\ \frac{\partial h_{t+1}^N}{\partial \sigma^D (t+k/T)} \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial \beta_D} h_{t+1}^D & \frac{\partial}{\partial \beta_D} h_{t+1}^N \end{pmatrix} = E \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} \lambda_t^D - \omega_D & 0 \end{pmatrix} A_{t+1}^T = 0; k = 0.$$

When $k > 1$, it holds

$$
\begin{aligned}
&\operatorname{vec}\begin{pmatrix} \frac{\partial h_t^D}{\partial \sigma^D(t-k/T)} \\ \frac{\partial h_t^N}{\partial \sigma^D(t-k/T)} \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial \beta_D} h_t^D & \frac{\partial}{\partial \beta_D} h_t^N \end{pmatrix} \\
&= \operatorname{vec} A_t \left(\prod_{i=1}^{k-1} B_{t-i} A_{t-i} \right) \Lambda_{t-k} \begin{pmatrix} \lambda_{t-1}^D - \omega_D & 0 \end{pmatrix} A_t^T \\
&+ \operatorname{vec} A_t \left(\prod_{i=1}^{k-1} B_{t-i} A_{t-i} \right) \Lambda_{t-k} \begin{pmatrix} \lambda_{t-2}^D - \omega_D & 0 \end{pmatrix} A_{t-1}^T B_{t-1}^T A_t^T \\
&+ \ldots \\
&+ \operatorname{vec} A_t \left(\prod_{i=1}^{k-1} B_{t-i} A_{t-i} \right) \Lambda_{t-k} \begin{pmatrix} \lambda_{t-k+1}^D - \omega_D & 0 \end{pmatrix} A_{t-k+2}^T B_{t-k+2}^T \ldots A_{t-1}^T B_{t-1}^T A_t^T \\
&= \sum_{j=1}^{k-1} (A_t \otimes A_t) \left(\prod_{i=1}^{j-1} (B_{t-i} \otimes B_{t-i}) (A_{t-i} \otimes A_{t-i}) \right) \operatorname{vec} \left(\left(\prod_{i=j}^{k-1} B_{t-i} A_{t-i} \right) \Lambda_{t-k} \begin{pmatrix} \lambda_{t-j}^D - \omega_D & 0 \end{pmatrix} \right).
\end{aligned}
$$

Since $(B_{t-1} \otimes B_{t-1})(A_{t-1} \otimes A_{t-1})$ and $B_t A_t$ are i.i.d, and $E B_t A_t = E B_t E A_t$, we obtain

$$
\begin{aligned}
&E \operatorname{vec} \begin{pmatrix} \frac{\partial h_t^D}{\partial \sigma^D(t-k/T)} \\ \frac{\partial h_t^N}{\partial \sigma^D(t-k/T)} \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial \beta_D} h_t^D & \frac{\partial}{\partial \beta_D} h_t^N \end{pmatrix} \\
&= \sum_{j=1}^{k-1} E(A_t \otimes A_t) E(B_{t-i} \otimes B_{t-i}) (A_{t-i} \otimes A_{t-i})^{j-1} E \operatorname{vec} \left(\left(\prod_{i=j}^{k-1} B_{t-i} A_{t-i} \right) \Lambda_{t-k} \begin{pmatrix} \lambda_{t-j}^D - \omega_D & 0 \end{pmatrix} \right).
\end{aligned} \tag{20}
$$

By (15), we can express λ_{t-1}^D as a function of $\left\{ \left(m_{t-i}^D, m_{t-i+1}^N \right), i > 1 \right\}$. Note that B_t, A_t,and Λ_t are independent of $\left\{ \left(m_s^D, m_s^N \right), s \neq t \right\}$. Therefore, we have

$$
\begin{aligned}
&E \left(\left(\prod_{i=j}^{k-1} B_{t-i} A_{t-i} \right) \Lambda_{t-k} \left(\lambda_{t-j}^D - \omega_D \right) \right) \\
&= \gamma_D E \left(\prod_{i=j}^{k-1} B_{t-i} A_{t-i} \right) \Lambda_{t-k} \sum\nolimits_{i=j+1}^{k} \beta_D^{i-1} \left(m_{t-i}^D + \left(m_{t-i}^D + 1 \right) \operatorname{sign}(e_{t-i}^D) \right) \\
&+ \rho_D E \left(\prod_{i=j}^{k-1} B_{t-i} A_{t-i} \right) \Lambda_{t-k} \sum\nolimits_{i=j+1}^{k} \beta_D^{i-1} \left(m_{t-i+1}^N + (m_{t-i+1}^N + 1) \operatorname{sign}(e_{t-i+1}^N) \right),
\end{aligned}
$$

with the first term

$$\begin{aligned}
&\left\| E\left(\prod_{i=j}^{k-1} B_{t-i}A_{t-i}\right)\Lambda_{t-k}\sum\nolimits_{i=j+1}^{k}\beta_D^{i-1}\left(m_{t-i}^D+\left(m_{t-i}^D+1\right)\operatorname{sign}(e_{t-i}^D)\right)\right\|_\infty \\
&\leq \left(\sum\nolimits_{i=j+1}^{k-1}\beta_D^{i-1}\right)\left\|E\left(B_t\left(m_t^D+(m_t^D+1)\operatorname{sign}(e_t^D)\right)A_t\right)\right\|_\infty \|EB_tEA_t\|_\infty^{k-j-1}\|E\Lambda_t\|_\infty \\
&+\beta_D^{k-j}\left\|E\Lambda_{t-k}\left(m_{t-k}^D+(m_{t-k}^D+1)\operatorname{sign}(e_{t-k}^D)\right)\right\|_\infty \|EB_tEA_t\|_\infty^{k-j} \\
&\leq \frac{\beta_D}{1-\beta_D}\left\|E\left(B_t\left(m_t^D+(m_t^D+1)\operatorname{sign}(e_t^D)\right)A_t\right)\right\|_\infty \|E\Lambda_t\|_\infty \|EB_tEA_t\|_\infty^{k-j-1} \\
&+\beta_D^{k-j}\left\|E\Lambda_{t-k}\left(m_{t-k}^D+(m_{t-k}^D+1)\operatorname{sign}(e_{t-k}^D)\right)\right\|_\infty \|EB_tEA_t\|_\infty^{k-j}
\end{aligned}$$

and the second term

$$\begin{aligned}
&\left\| E\left(\prod_{i=j}^{k-1} B_{t-i}A_{t-i}\right)\Lambda_{t-k}\sum\nolimits_{i=j+1}^{k}\beta_D^{i-1}\left(m_{t-i+1}^N+(m_{t-i+1}^N+1)\operatorname{sign}(e_{t-i+1}^N)\right)\right\|_\infty \\
&\leq \frac{\beta_D}{1-\beta_D}\left\|E\left(B_t\left(m_t^N+(m_t^N+1)\operatorname{sign}(e_t^N)\right)A_t\right)\right\|_\infty \|E\Lambda_t\|_\infty \|EB_tEA_t\|_\infty^{k-j-1}.
\end{aligned}$$

According to the definition of $\|\|_\infty$,

$$\left\|E\operatorname{vec}\left(\left(\prod_{i=j}^{k-1} B_{t-i}A_{t-i}\right)\Lambda_{t-k}\begin{pmatrix}\lambda_{t-j}^D-\omega_D & 0\end{pmatrix}\right)\right\|_\infty \leq \left\|E\left(\left(\prod_{i=j}^{k-1} B_{t-i}A_{t-i}\right)\Lambda_{t-k}\begin{pmatrix}\lambda_{t-j}^D-\omega_D & 0\end{pmatrix}\right)\right\|_\infty$$

Therefore,

$$\left\|E\operatorname{vec}\left(\left(\prod_{i=j}^{k-1} B_{t-i}A_{t-i}\right)\Lambda_{t-k}\begin{pmatrix}\lambda_{t-j}^D-\omega_D & 0\end{pmatrix}\right)\right\|_\infty \leq c_T\left\|EB_tEA_t\right\|^{k-j-1} \tag{21}$$

with

$$\begin{aligned}
c_T = &\frac{\beta_D}{1-\beta_D}|\gamma_D|\left\|E\left(B_t\left(m_t^D+(m_t^D+1)\operatorname{sign}(e_t^D)\right)A_t\right)\right\|_\infty \|E\Lambda_t\|_\infty \\
&+\|EB_tEA_t\|_\infty\left\|E\Lambda_{t-k}\left(m_{t-k}^D+(m_{t-k}^D+1)\operatorname{sign}(e_{t-k}^D)\right)\right\|_\infty \\
&+\frac{\beta_D}{1-\beta_D}|\rho_D|\left\|E\left(B_t\left(m_t^N+(m_t^N+1)\operatorname{sign}(e_t^N)\right)A_t\right)\right\|_\infty \|E\Lambda_t\|_\infty.
\end{aligned}$$

Substituting (21) into (20) gives

$$\begin{aligned}
&\left\| E\text{vec}\begin{pmatrix}\frac{\partial h_t^D}{\partial\sigma^D(t-k/T)}\\ \frac{\partial h_t^N}{\partial\sigma^D(t-k/T)}\end{pmatrix}\begin{pmatrix}\frac{\partial}{\partial\beta_D}h_t^D & \frac{\partial}{\partial\beta_D}h_t^N\end{pmatrix}\right\|_\infty \\
&\leq \sum_{j=1}^{k-1}\left\|E\left(A_t\otimes A_t\right)\right\|_\infty\left\|E\left(B_{t-i}\otimes B_{t-i}\right)\left(A_{t-i}\otimes A_{t-i}\right)\right\|_\infty^{j-1}c_T\left\|EB_tEA_t\right\|_\infty^{k-j-1} \\
&\leq c_T\left\|E\left(A_t\otimes A_t\right)\right\|_\infty\sum_{j=1}^{k-1}\left\|E\left(B_{t-i}\otimes B_{t-i}\right)\left(A_{t-i}\otimes A_{t-i}\right)\right\|_\infty^{j-1}\left\|EB_tEA_t\right\|_\infty^{k-j-1} \\
&\leq c_T\left\|E\left(A_t\otimes A_t\right)\right\|_\infty\frac{\left\|EB_tEA_t\right\|_\infty^{k-2}}{1-\frac{\|E(B_{t-i}\otimes B_{t-i})(A_{t-i}\otimes A_{t-i})\|_\infty}{\|EB_tEA_t\|_\infty}},
\end{aligned}$$

provided that $\left\|EB_tEA_t\right\|_\infty<1$ and $\left\|E\left(B_{t-1}A_{t-1}\otimes B_{t-1}A_{t-1}\right)\right\|_\infty<\left\|EB_tEA_t\right\|_\infty$ .It is then straightforward to show

$$\sum_k k\left\|E\text{vec}\begin{pmatrix}\frac{\partial h_t^D}{\partial\sigma^D(t-k/T)}\\ \frac{\partial h_t^N}{\partial\sigma^D(t-k/T)}\end{pmatrix}\begin{pmatrix}\frac{\partial}{\partial\beta_D}h_t^D & \frac{\partial}{\partial\beta_D}h_t^N\end{pmatrix}\right\|_\infty<\infty$$

and thereby

$$\sum_k k\left\|E\begin{pmatrix}\frac{\partial h_t^D}{\partial\sigma^D(t-k/T)}\\ \frac{\partial h_t^N}{\partial\sigma^D(t-k/T)}\end{pmatrix}\begin{pmatrix}\frac{\partial}{\partial\beta_D}h_t^D & \frac{\partial}{\partial\beta_D}h_t^N\end{pmatrix}\right\|_\infty<\infty.$$

■

Lemma 4 *The score functions of h_t^j with respect to β_D, v_D and $\sigma^j(t/T)$ are*

$$\begin{aligned}
\begin{pmatrix}\frac{\partial}{\partial\beta_D}h_t^D\\ \frac{\partial}{\partial\beta_D}h_t^N\end{pmatrix} &= A_t\begin{pmatrix}\lambda_{t-1}^D-\omega_D\\ 0\end{pmatrix}+A_tB_{t-1}\begin{pmatrix}\frac{\partial}{\partial\beta_D}h_{t-1}^D\\ \frac{\partial}{\partial\beta_D}h_{t-1}^N\end{pmatrix} \\
&= \sum_{j=1}^{\infty}A_t\prod_{i=1}^{j-1}B_{t-i}A_{t-i}\begin{pmatrix}\lambda_{t-j}^D-\omega_D\\ 0\end{pmatrix}.
\end{aligned} \tag{22}$$

$$\begin{aligned}
\begin{pmatrix}\frac{\partial h_t^D}{\partial\sigma^D(t-k/T)}\\ \frac{\partial h_t^N}{\partial\sigma^D(t-k/T)}\end{pmatrix} &= A_tB_{t-1}\begin{pmatrix}\frac{\partial h_{t-1}^D}{\partial\sigma^D(t-k/T)}\\ \frac{\partial h_{t-1}^N}{\partial\sigma^D(t-k/T)}\end{pmatrix} \\
&= A_t\left(\prod_{i=1}^{k-1}B_{t-i}A_{t-i}\right)\Lambda_{t-k}, k>1
\end{aligned} \tag{23}$$

$$\begin{pmatrix}\frac{\partial h_t^D}{\partial\sigma^D(t/T)}\\ \frac{\partial h_t^N}{\partial\sigma^D(t/T)}\end{pmatrix}=\begin{pmatrix}1\\ 0\end{pmatrix};\ \textit{and}\ \begin{pmatrix}\frac{\partial h_t^D}{\partial\sigma^D(t-1/T)}\\ \frac{\partial h_t^N}{\partial\sigma^D(t-1/T)}\end{pmatrix}=A_t\begin{pmatrix}a_{t-1}^{DD}\\ a_{t-1}^{ND}\end{pmatrix},$$

with $\Lambda_t = \begin{pmatrix} a_t^{DD} \\ a_t^{ND} \end{pmatrix}$ *.If the top-Lyapunov exponent of the sequence of* $A_t B_{t-1}$ *is strictly negative,* $\begin{pmatrix} \frac{\partial}{\partial \beta_D} h_t^D \\ \frac{\partial}{\partial \beta_D} h_t^N \end{pmatrix}$, $\begin{pmatrix} \frac{\partial h_t^D}{\partial \sigma^D (t-k/T)} \\ \frac{\partial h_t^N}{\partial \sigma^D (t-k/T)} \end{pmatrix}$ *and* $\begin{pmatrix} \frac{\partial h_t^D}{\partial \sigma^D (t-k/T)} \\ \frac{\partial h_t^N}{\partial \sigma^D (t-k/T)} \end{pmatrix} \begin{pmatrix} \frac{\partial h_t^D}{\partial \beta_D} & \frac{\partial h_t^N}{\partial \beta_D} \end{pmatrix}$ *are strictly stationary.*

Proof. Since $h_t^j = \lambda_t^j + \sigma^j(t/T)$, we can write h_t^j in a recursive formula as

$$\begin{aligned} h_t^D &= \sigma^D(t/T) - \beta_D \sigma^D(\frac{t-1}{T}) + \omega_D(1-\beta_D) + \beta_D h_{t-1}^D + \gamma_D m_{t-1}^D \\ &\quad + \rho_D m_t^N + \gamma_D^*(m_{t-1}^D + 1)\text{sign}(u_{t-1}^D) + \rho_D^*(m_t^N + 1)\text{sign}(u_t^N) \qquad (24) \\ h_t^N &= \sigma^N(t/T) - \beta_N \sigma^N(\frac{t-1}{T}) + \omega_N(1-\beta_N) + \beta_N h_{t-1}^N + \gamma_N m_{t-1}^N \\ &\quad + \rho_N m_{t-1}^D + \rho_N^*(m_{t-1}^D + 1)\text{sign}(u_{t-1}^D) + \gamma_N^*(m_{t-1}^N + 1)\text{sign}(u_{t-1}^N). \qquad (25) \end{aligned}$$

and m_t^D and m_t^N can be expressed as

$$\begin{aligned} m_t^D &= \frac{(1+v_D)(u_t^D)^2 \exp(-2h_t^D)}{v_D + (u_t^D)^2 \exp(-2h_t^D)} - 1, \quad v_D > 0 \\ m_t^N &= \frac{(1+v_N)(u_t^N)^2 \exp(-2h_t^N)}{v_N + (u_t^N)^2 \exp(-2h_t^N)} - 1, \quad v_N > 0. \end{aligned}$$

Taking the first order derivative of equation (24) and (25) with respect to β_D gives

$$\begin{aligned} \frac{\partial h_t^D}{\partial \beta_D} &= -\sigma^D(\frac{t-1}{T}) - \omega_D + h_{t-1}^D + \beta_D \frac{\partial}{\partial \beta_D} h_{t-1}^D + \frac{\partial}{\partial \beta_D} \gamma_D m_{t-1}^D + \frac{\partial}{\partial \beta_D} \rho_D m_t^N \\ &\quad + \frac{\partial}{\partial \beta_D} \gamma_D^*(m_{t-1}^D + 1)\text{sign}(u_{t-1}^D) + \frac{\partial}{\partial \beta_D} \rho_D^*(m_t^N + 1)\text{sign}(u_t^N) \qquad (26) \\ \frac{\partial h_t^N}{\partial \beta_D} &= \beta_N \frac{\partial}{\partial \beta_D} h_{t-1}^N + \frac{\partial}{\partial \beta_D} \gamma_N m_{t-1}^N + \frac{\partial}{\partial \beta_D} \rho_N m_{t-1}^D \\ &\quad + \frac{\partial}{\partial \beta_D} \rho_N^*(m_{t-1}^D + 1)\text{sign}(u_{t-1}^D) + \frac{\partial}{\partial \beta_D} \gamma_N^*(m_{t-1}^N + 1)\text{sign}(u_{t-1}^N) \qquad (27) \end{aligned}$$

and the derivatives of m_{t-1}^D and m_{t-1}^N are

$$\begin{aligned} \frac{\partial}{\partial \beta_D} m_{t-1}^D &= \frac{\partial m_{t-1}^D}{\partial h_{t-1}^D} \frac{\partial}{\partial \beta_D} h_{t-1}^D = -2\,(v_D + 1)\, b_{t-1}^D \left(1 - b_{t-1}^D\right) \frac{\partial}{\partial \beta_D} h_{t-1}^D \\ \frac{\partial}{\partial \beta_D} m_{t-1}^N &= \frac{\partial m_{t-1}^N}{\partial h_{t-1}^N} \frac{\partial}{\partial \beta_D} h_{t-1}^N = -2\,(v_N + 1)\, b_{t-1}^N \left(1 - b_{t-1}^N\right) \frac{\partial}{\partial \beta_D} h_{t-1}^N. \end{aligned}$$

Substituting them back into (26) and (27) gives

$$\frac{\partial h_t^D}{\partial \beta_D} = \lambda_{t-1}^D - \omega_D + \left(\beta_D + a_{t-1}^{DD}\right) \frac{\partial}{\partial \phi} h_{t-1}^D + a_t^{DN} \frac{\partial}{\partial \phi} h_t^N$$
$$\frac{\partial h_t^N}{\partial \beta_D} = 0 + \left(\beta_N + a_{t-1}^{NN}\right) \frac{\partial}{\partial \phi} h_{t-1}^N + a_{t-1}^{ND} \frac{\partial}{\partial \phi} h_{t-1}^D$$

with the matrix form

$$\begin{pmatrix} \frac{\partial}{\partial \beta_D} h_t^D \\ \frac{\partial}{\partial \beta_D} h_t^N \end{pmatrix} = A_t \begin{pmatrix} \lambda_{t-1}^D - \omega_D \\ 0 \end{pmatrix} + A_t B_{t-1} \begin{pmatrix} \frac{\partial}{\partial \beta_D} h_{t-1}^D \\ \frac{\partial}{\partial \beta_D} h_{t-1}^N \end{pmatrix}.$$

Note that $A_t B_{t-1}$ and $A_t \begin{pmatrix} \lambda_{t-1}^D - \omega_D \\ 0 \end{pmatrix}$ are strictly stationary and ergodic, by Theorem 4.27 in Douc et al. (2014), when the top-Lyapunov exponent of the sequence of $A_t B_{t-1}$ is strictly negative, $\begin{pmatrix} \frac{\partial}{\partial \beta_D} h_t^D \\ \frac{\partial}{\partial \beta_D} h_t^N \end{pmatrix}$ converges and is strictly stationary.

Likewise, taking the first order derivative of h_t^j with respect to $\sigma^D\left(\frac{t-k}{T}\right)$ yields

$$\frac{\partial h_t^D}{\partial \sigma^D((t-k)/T)} = \left(\beta_D + a_{t-1}^{DD}\right) \frac{\partial h_{t-1}^D}{\partial \sigma^D((t-k)/T)} + a_t^{DN} \frac{\partial h_t^N}{\partial \sigma^D((t-k)/T)}, \ k > 1$$
$$\frac{\partial h_t^D}{\partial \sigma^D(t/T)} = 1, \ \frac{\partial h_t^D}{\partial \sigma^D((t-1)/T)} = a_{t-1}^{DD} + a_t^{DN} a_{t-1}^{ND}$$

$$\frac{\partial h_t^N}{\partial \sigma^D((t-k)/T)} = \left(\beta_N + a_{t-1}^{NN}\right) \frac{\partial h_{t-1}^N}{\partial \sigma^D((t-k)/T)} + a_{t-1}^{ND} \frac{\partial h_{t-1}^D}{\partial \sigma^D((t-k)/T)}, \ k > 1$$
$$\frac{\partial h_t^N}{\partial \sigma^D(t/T)} = 0, \ \frac{\partial h_t^N}{\partial \sigma^D((t-1)/T)} = a_{t-1}^{ND},$$

and (23) follows. Similarly, $\begin{pmatrix} \frac{\partial h_t^D}{\partial \sigma^D((t-k)/T)} \\ \frac{\partial h_t^N}{\partial \sigma^D((t-k)/T)} \end{pmatrix}$ is strictly stationary across time t.

Finally, we can write

$$\begin{pmatrix} \frac{\partial h_t^D}{\partial \sigma^D((t-k)/T)} \\ \frac{\partial h_t^N}{\partial \sigma^D((t-k)T)} \\ \frac{\partial h_t^D}{\partial \beta_D} \\ \frac{\partial h_t^N}{\partial \beta_D} \end{pmatrix} = \begin{pmatrix} A_t B_{t-1} & 0 \\ 0 & A_t B_{t-1} \end{pmatrix} \begin{pmatrix} \frac{\partial h_{t-1}^D}{\partial \sigma^D((t-k)/T)} \\ \frac{\partial h_{t-1}^N}{\partial \sigma^D((t-k)/T)} \\ \frac{\partial h_{t-1}^D}{\partial \beta_D} \\ \frac{\partial h_{t-1}^N}{\partial \beta_D} \end{pmatrix} + \begin{pmatrix} A_t \begin{pmatrix} \lambda_{t-1}^D - \omega_D \\ 0 \end{pmatrix} \\ 0 \end{pmatrix}.$$

Both $\begin{pmatrix} \frac{\partial h_t^D}{\partial \sigma^D((t-k)/T)} \\ \frac{\partial h_t^N}{\partial \sigma^D((t-k)/T)} \\ \frac{\partial h_t^D}{\partial \beta_D} \\ \frac{\partial h_t^N}{\partial \beta_D} \end{pmatrix}$ and $\begin{pmatrix} \frac{\partial h_t^D}{\partial \sigma^D((t-k)/T)} \\ \frac{\partial h_t^N}{\partial \sigma^D((t-k)/T)} \end{pmatrix} \begin{pmatrix} \frac{\partial h_t^D}{\partial \beta_D} & \frac{\partial h_t^N}{\partial \beta_D} \end{pmatrix}$ are strictly stationary, since the top-Lyapunov exponent of the sequence $\begin{pmatrix} A_t B_{t-1} & 0 \\ 0 & A_t B_{t-1} \end{pmatrix}$, same as that of $A_t B_{t-1}$, is strictly negative by assumption. ■

Lemma 5 *When Assumption A1-A4 holds, we have* $\frac{1}{T}\sum_{t=1}^{T} E\frac{\partial h_t^D}{\partial \theta} = 0$.

Proof. Similar to the proof of Theorem 1, we only need to show $\sum_{t=1}^{T} k \left\| E \begin{pmatrix} \frac{\partial h_t^D}{\partial \sigma^N((t-k)/T)} \\ \frac{\partial h_t^N}{\partial \sigma^N((t-k)T)} \end{pmatrix} \right\|_\infty < \infty$. Note that $E \begin{pmatrix} \frac{\partial h_t^D}{\partial \sigma^N((t-k)/T)} \\ \frac{\partial h_t^N}{\partial \sigma^N((t-k)/T)} \end{pmatrix} = EA_t B_{t-1} A_{t-1} B_{t-2} .. A_{t-k+2} B_{t-k+1} A_{t-k+1} \Lambda_{t-k} = A\left(BA\right)^{k-1} \Lambda$, when $k > 1$. Obviously, $\sum_{t=1}^{T} k \left\| E \begin{pmatrix} \frac{\partial h_t^D}{\partial \sigma^N((t-k)/T)} \\ \frac{\partial h_t^N}{\partial \sigma^N((t-k)/T)} \end{pmatrix} \right\|_\infty < \infty$. ■

8.3 Appendix 3: Derivatives in the multivariate model

We now give the first and second order derivatives of the global log-likelihood function in the multivariate model, given λ_t and v. Without subscripts j, the log-likelihood function, ignoring some unnecessary parts, is

$$l_t = \log|\Theta| - \sum_{i=1}^{n} \left(\frac{v_i + 1}{2} \ln \left(1 + \frac{(\iota_i^\intercal \mathrm{diag}\left(\exp(-\lambda_t)\right) \Theta u_t)^2}{v_i} \right) \right).$$

Then

$$
\begin{aligned}
\mathrm{d}l_t &= \mathrm{d}\log|\Theta| - \sum_{i=1}^{n} \frac{(v_i+1)\left(\iota_i^\intercal \mathrm{diag}\left(\exp(-\lambda_t)\right)\Theta u_t\right)\exp(-\lambda_{it})}{v_i + \left(\iota_i^\intercal \mathrm{diag}\left(\exp(-\lambda_t)\right)\Theta u_t\right)^2} \mathrm{tr} u_t \iota_i^\intercal \mathrm{d}\Theta \\
&= \mathrm{tr}\left(\Theta^{-1}\mathrm{d}\Theta\right) - \mathrm{tr}\left(\sum_{i=1}^{n} \frac{(v_i+1)\exp(-2\lambda_{it})\iota_i^\intercal \Theta u_t}{v_i + \exp(-2\lambda_{it})\left(\iota_i^\intercal \Theta u_t\right)^2} u_t \iota_i^\intercal \mathrm{d}\Theta\right) \\
&= \mathrm{tr}\left[\left(\Theta^{-1} - \sum_{i=1}^{n} \frac{(v_i+1)\exp(-2\lambda_{it})\iota_i^\intercal \Theta u_t}{v_i + \exp(-2\lambda_{it})\left(\iota_i^\intercal \Theta u_t\right)^2} u_t \iota_i^\intercal\right)\mathrm{d}\Theta\right] \\
&= \left[\mathrm{vec}\left(\Theta^{-1} - \sum_{i=1}^{n} \iota_i u_t^\intercal \frac{(v_i+1)\exp(-2\lambda_{it})\iota_i^\intercal \Theta u_t}{v_i + \exp(-2\lambda_{it})\left(\iota_i^\intercal \Theta u_t\right)^2}\right)\right]^\intercal \mathrm{dvec}\Theta \\
&= \left[\mathrm{vec}\left(\Theta^{-1} - \sum_{i=1}^{n} \iota_i u_t^\intercal \frac{(v_i+1)\exp(-2\lambda_{it})\iota_i^\intercal \Theta u_t}{v_i + \exp(-2\lambda_{it})\left(\iota_i^\intercal \Theta u_t\right)^2}\right)\right]^\intercal D_n \mathrm{dvech}\Theta, \qquad (28)
\end{aligned}
$$

where D_n is the duplication matrix so that $\mathrm{vec}\Theta = D_n \mathrm{vech}\Theta$. Therefore, the first order derivative of the global log-likelihood function is

$$
\begin{aligned}
\frac{\partial L_T(\Theta;\lambda_t,s)}{\partial \mathrm{vech}\Theta} &= -\frac{1}{T} D_n^\intercal \mathrm{vec} \sum_{i=1}^{n}\left(\iota_i \sum_{t=1}^{T}\left(K_h(s-t/T) u_t^\intercal \frac{(v_i+1)\exp(-2\lambda_{it})\iota_i^\intercal \Theta u_t}{v_i + \exp(-2\lambda_{it})\left(\iota_i^\intercal \Theta u_t\right)^2}\right)\right) \\
&\quad + D_n^\intercal \mathrm{vec}\left(\Theta^{-1}\right). \qquad (29)
\end{aligned}
$$

To compute the Hessian matrix, we evaluate the differential of the Jacobian matrix in (28)

$$
\begin{aligned}
&\mathrm{dvec} D_n^\intercal \left(\Theta^{-1} - \sum_{i=1}^{n} \iota_i u_t^\intercal \frac{(v_i+1)\exp(-2\lambda_{it})\iota_i^\intercal \Theta u_t}{v_i + \left(\exp(-2\lambda_{it})\iota_i^\intercal \Theta u_t\right)^2}\right) D_n \\
&= D_n^\intercal \mathrm{dvec}\Theta^{-1} - D_n^\intercal \mathrm{vec} \sum_{i=1}^{n}\left(\mathrm{d}\frac{(v_i+1)\exp(-2\lambda_{it})\iota_i^\intercal \Theta u_t}{v_i + \exp(-2\lambda_{it})\left(\iota_i^\intercal \Theta u_t\right)^2}\right)\iota_i u_t^\intercal \\
&= D_n^\intercal \mathrm{dvec}\Theta^{-1} - D_n^\intercal \sum_{i=1}^{n} \frac{\left(v_i - \exp(-2\lambda_{it})\left(\iota_i^\intercal \Theta u_t\right)^2\right)(v_i+1)\exp(-2\lambda_{it})}{\left(v_i + \exp(-2\lambda_{it})\left(\iota_i^\intercal \Theta u_t\right)^2\right)^2} \mathrm{vec}\left(\iota_i \iota_i^\intercal \mathrm{d}\Theta u_t u_t^\intercal\right) \\
&= -D_n^\intercal \left(\Theta^{-1} \otimes \Theta^{-1}\right) D_n \mathrm{dvech}\Theta \\
&- D_n^\intercal \sum_{i=1}^{n} \frac{\left(v_i - \exp(-2\lambda_{it})\left(\iota_i^\intercal \Theta u_t\right)^2\right)(v_i+1)\exp(-2\lambda_{it})}{\left(v_i + \exp(-2\lambda_{it})\left(\iota_i^\intercal \Theta u_t\right)^2\right)^2} \left(u_t u_t^\intercal\right) \otimes \left(\iota_i \iota_i^\intercal\right) D_n \mathrm{dvech}\Theta.
\end{aligned}
$$

The Hessian matrix of the global log-likelihood function is thus

$$\frac{\partial^2 L_T(\Theta;\lambda_t,s)}{\partial \mathrm{vech}\Theta \partial\left(\mathrm{vech}\Theta\right)^{\intercal}}$$
$$= -D_n^{\intercal}\left(\sum_{i=1}^{n}\left(\sum_{t=1}^{T}\frac{K_h(s-t/T)\left(v_i-\exp(-2\lambda_{it})\left(\iota_i^{\intercal}\Theta u_t\right)^2\right)(v_i+1)}{T\left(v_i+\exp(-2\lambda_{it})\left(\iota_i^{\intercal}\Theta u_t\right)^2\right)^2\exp(2\lambda_{it})}u_t u_t^{\intercal}\right)\otimes\left(\iota_i\iota_i^{\intercal}\right)\right)D_n$$
$$- D_n^{\intercal}\left(\Theta^{-1}\otimes\Theta^{-1}\right)D_n \tag{30}$$

■