UNIVERSITY OF CAMBRIDGE
Faculty of Economics

Institute for New Economic Thinking

CAMBRIDGE WORKING PAPERS IN ECONOMICS
CAMBRIDGE-INET WORKING PAPERS

A Coupled Component GARCH Model for Intraday and Overnight Volatility

Oliver Linton
University of Cambridge

Jianbin Wu
Nanjing University

Abstract

We propose a semi-parametric coupled component GARCH model for intraday and overnight volatility that allows the two return series to have different properties. We adopt a dynamic conditional score model with t-distributed innovations that captures the very heavy tails of overnight returns. We propose a several-step estimation procedure that captures the nonparametric slowly moving components by kernel estimation and the dynamic parameters by estimated maximum likelihood. We establish the consistency, asymptotic normality, and semiparametric efficiency of our semiparametric estimation procedures. We extend the modelling to the multivariate case where we allow time varying correlation between stocks. We apply our model to the study of Dow Jones industrial average component stocks, CRSP size-based portfolios, and size-based portfolios in four large international markets over the period 1993-2017. We show that the ratio of overnight to intraday volatility has actually increased in importance for Dow Jones stocks during the last two decades. This ratio has also increased for large stocks in the CRSP database, but decreased for small stocks in CRSP. Notably, the slope increases monotonically from the smallest-cap decile to the largest-cap decile. This pattern also exists in other international markets. The multivariate model shows that overnight and intraday correlations have both increased, but overnight correlations have increased more substantially during recent crises than intraday correlations.

Reference Details

1879 Cambridge Working Papers in Economics
2018/25 Cambridge-INET Working Paper Series

Published 14 September 2018

Key Words DCS, GAS, GARCH, size-based portfolios, Testing
JEL Codes C12 C13

Websites www.econ.cam.ac.uk/cwpe
www.inet.econ.cam.ac.uk/working-papers

A Coupled Component GARCH Model for Intraday and Overnight Volatility*

Oliver Linton† University of Cambridge

Jianbin Wu‡ Nanjing University

September 14, 2018

Abstract

We propose a semi-parametric coupled component GARCH model for intraday and overnight volatility that allows the two return series to have different properties. We adopt a dynamic conditional score model with t-distributed innovations that captures the very heavy tails of overnight returns. We propose a several-step estimation procedure that captures the non-parametric slowly moving components by kernel estimation and the dynamic parameters by estimated maximum likelihood. We establish the consistency, asymptotic normality, and semi-parametric efficiency of our semiparametric estimation procedures. We extend the modelling to the multivariate case where we allow time varying correlation between stocks. We apply our model to the study of Dow Jones industrial average component stocks, CRSP size-based portfolios, and size-based portfolios in four large international markets over the period 1993-2017. We show that the ratio of overnight to intraday volatility has actually increased in importance for Dow Jones stocks during the last two decades. This ratio has also increased for large stocks in the CRSP database, but decreased for small stocks in CRSP. Notably, the slope increases monotonically from the smallest-cap decile to the largest-cap decile. This pattern also exists in other international markets. The multivariate model shows that overnight and intraday correlations have both increased, but overnight correlations have increased more substantially during recent crises than intraday correlations.

KEYWORDS: DCS, GAS, GARCH, size-based portfolios, Testing

JEL CLASSIFICATION: C12, C13

*We would like to thank Greg Connor, Jinyong Hahn, Andrew Harvey, Hashem Pesaran, Piet Sercu, Haihan 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 Economics, Nanjing, China. Email:wujianbin@nju.edu.cn

1 Introduction

The balance between intraday and overnight returns is of considerable interest as it potentially 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 versus continuous trading, the effect of high frequency trading on market quality, and the globalization and connectedness of international markets. We propose a bivariate time series model for intraday and overnight returns that respects their temporal ordering and permits the two processes to have different marginal properties, and to feedback into each other, and allows for both short run and long components. In particular, our volatility model for each return series 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, where those deviations depend on previous intraday and overnight shocks. We adopt a dynamic conditional score (DCS) 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 intraday and overnight). 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 the tail thickness of the conditional distributions. The short run dynamic process allows for leverage effects and separates the overnight shock from the intraday shock. We also introduce a multivariate model that allows for time varying correlations between stocks and between overnight and intraday returns.

We apply our model to the study of 26 Dow Jones industrial average component stocks over the period 1993-2017, a period that 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 (HFT) strategies in the period post 2005 has lead to an increase in volatility, see Boehmer, Fong, and Wu (2015) and Linton, O'Hara, and Zigrand (2013). To address this, 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 HFT was becoming more prevalent. There are a number of studies that have investigated this question with natural experiments methodology (Brogaard, 2011), but the conclusions one can draw from such work are event specific. We model the volatility process with a view to addressing this hypothesis in a more general way. One implication of this hypothesis is that ceteris paribus 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 volatility from the intraday segment with the daily return volatility 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 (unconditional) variances are not very reliable. We use our dynamic two component model, which allows for both fast and slow dynamic components to volatility, as is now common practice (Engle and Lee, 1999; Engle and Rangel, 2008; Hafner and Linton, 2010; Rangel

and Engle, 2012; and Han and Kristensen, 2015). 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, but the parameter estimates we employ are robust to this phenomenon. Our methodology 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 the long run component of overnight volatility has actually increased in importance during this period relative to the long run component of intraday volatility (although intraday volatility is still generally higher than overnight volatility). We provide a formal test statistic that confirms quantitively the strength of this effect; our test can be interpreted as carrying out a difference in difference analysis but in ratio form, Imbens and Wooldridge (2007). This finding 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 that overnight return shocks 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 a 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 have slowly increased more or less from the beginning of the period.

We also apply our model to size-sorted portfolios of CRSP stocks over the period 1993-2017. We find that the ratio of overnight to intraday volatility has indeed increased for large stocks, but has decreased for small stocks especially in the 1990s. Notably, the slope increases monotonically from the smallest-cap to the largest-cap decile, and the ratio of overnight to intraday volatility is typically high during recent crises. From the multivariate model, we find that small stocks had rather weak co-movement with the market in the early 1990s. However the co-movement has increased considerably during the recent period, although it still remains smaller than that of large stocks. In general, the overnight correlation increases more substantially than the intraday correlation during recent crises. We further apply our model to four large international markets (U.S., U.K., Germany and Japan). The results are similar.

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 open price to forecast intraday volatility or to update the close to close volatility forecast and also to take into account 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.

We work only with the return series, although for some stocks intraday transaction and quote records are available for the duration of our study, which would permit the computation of realized volatility measures, which are for some the preferred measure of intraday volatility. This however would pose some additional questions in terms of the joint modelling of discrete time returns and real-

ized volatility, and puts an imbalance between the measurement of intraday and overnight volatility.[1] Furthermore, it would be problematic to implement some of those techniques on the small CRSP stocks in the early part of the sample period, so it is not a silver bullet. Instead we do make use of alternative measures of market (SP500) volatility - the VIX (which includes overnight volatility) and the Rogers and Satchell (1991) intraday volatility measure - to conduct a robustness check. We find that these measures confirm the finding regarding the rise of overnight volatility relative to intraday after 2004.

Related Literature. Overnight returns have recently attracted much attention in empirical finance. Many find overnight and intraday returns behave entirely differently, and overnight returns tend to outperform intraday returns. Specifically, Cooper, Cliff, and Gulen (2008) suggest that the U.S. equity premium over the period 1993-2006 is solely due to overnight returns. Kelly and Clark (2011) find the overnight returns are on average larger than the intraday returns. Berkman, Koch, Tuttle, and Zhang (2012) find a significant positive mean overnight return and a significant negative mean intraday return. They suggest stocks that have recently attracted the attention of retail investors tend to have higher net retail buying at the open, leading to high overnight returns that followed by intraday reversals. Aboody, Even-Tov, Lehavy, and Trueman (2018) suggest overnight returns can serve as a measure of firm-specific investor sentiment, and find short-term persistence in overnight returns. Polk, Lou, and Skouras (2018) link investor heterogeneity to the strong persistence of the overnight and intraday returns. They find an overnight versus intraday tug of war in strategy risk premium, and the risk premium is earned entirely overnight for the largest stocks. Besides the difference in expected returns, overnight returns are found less volatile (French and Roll, 1986; Lockwood and Linn, 1990; Aretz and Bartram, 2015), but more leptokurtic than intraday returns in the U.S. market (Ng and Masulis, 1995; Blanc, Chicheportiche, and Bouchaud, 2014).

Tsiakas (2008) proposed a stochastic volatility model for daytime returns with feedback from night to day and leverage effects built in. He assumed Gaussian innovations; he did not model the overnight returns. In the literature on realized volatility, many authors have considered how to incorporate overnight returns into daily variance modeling and forecasting, by scaling the intraday measure (e.g., Martens, 2002 and Fleming, Kirby, and Ostdiek, 2003), or by combining daytime realized volatility and the squared overnight return with optimally chosen weight parameters (e.g., Hansen and Lunde, 2005). However, these authors also did not model the overnight returns either. Andersen, Bollerslev, and Huang (2011) decomposed the total daily return variability into the continuous sample path variance, the discontinuous intraday jumps, and the overnight variance. For this overnight variance, they used an augmented GARCH-t type structure with the immediately preceding daytime realized volatility as an additional explanatory variable. Blanc, Chicheportiche, and Bouchaud (2014) employ a quadratic ARCH model with flexible dynamics for both intraday and overnight returns; they also allow for feedback from overnight to intraday returns and leverage effects. They use a t distributed shock to drive each process and to define an estimation algorithm. They impose a pooling assumption on the model parameters across 280 S&P500 stocks that are continually in the index over 2000-2009,

[1]Our main findings involve averages of the daily volatility series and so the efficiency gain of realized volatility may not be so large in this context.

and assume stationarity over the period in question.

Our paper is closely related to the Generalized Autoregressive Score models or the Beta-t-(E)GARCH model. Creal, Koopman, and Lucas (2013) introduced a general class of time series models called Generalized Autoregressive Score models (GAS). Simultaneously, Harvey, Chakravarty et al. (2008) developed a score driven model specifically for volatilities, called the Beta-t-(E)GARCH model, built on exactly the same philosophy. Harvey (2013) settles on the dynamic conditional score model terminology, and we follow that nomenclature. This paper is also related to the work of Engle and Rangel (2008) and Hafner and Linton (2010) about incorporating long run volatilities.Engle and Rangel (2008) introduced nonparametric slowly varying trends into GARCH models; Hafner and Linton (2010) propose a multivariate extension and develop the distribution theory for inference.

2 The Model and its 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 open price on day t and P_t^C denotes the close price on day t. Daily close to close returns satisfy $r_t = r_t^D + r_t^N$. The timeline is illustrated below

$$\cdots \longrightarrow P_{t-1}^C \quad \Big| \underbrace{\overbrace{\longrightarrow}^{Night\ \mathrm{t}}}_{r_t^N} \; P_t^O \; \underbrace{\overbrace{\longrightarrow}^{Day\ \mathrm{t}}}_{r_t^D} \; P_t^C \Big| \overbrace{\longrightarrow}^{Night\ \mathrm{t+1}} \quad P_{t+1}^O \longrightarrow \cdots$$

We do not distinguish between weekend, holiday weekends, and ordinary midweek over night periods, although we comment on this issue in the concluding section below.

Our model allows intraday returns to depend on overnight returns with the same t, but overnight returns just depend on lagged variables. We 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 what could be misspricing effects or short run effects such as might arise from the market microstructure, Scholes and Williams (1977).

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

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

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 $j = D, N$:

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

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 $\sigma^D(\cdot)$ and $\sigma^N(\cdot)$ integrate to zero to achieve identification, which is a different approach from Hafner and Linton (2010). In this case we do not have to restrict the parameters of the short run dynamic processes for identification. In the following, j is always used to denote D, N without further mentioning.

Regarding the short run dynamic part of (2), we adopt a dynamic conditional score approach, Creal, Koopman, and Lucas (2011) and Harvey and Luati (2014). The conditional (scale) score function associated with the t-distributed shocks is $(1 - x^2(\nu+1)/(\nu\sigma^2 + x^2))/\sigma^2$, and so we take as innovation processes

$$m_t^j = \frac{(1+v_j)(e_t^j)^2}{v_j \exp(2\lambda_t^j) + (e_t^j)^2} - 1, \quad v_j > 0 \tag{4}$$

where $e_t^j = \exp(-\sigma^j(t/T))u_t^j$ for $j = D, N$. Note that $\text{var}(e_t^j|\mathcal{F}_{t-1}) = \exp(2\lambda_t^j)\text{var}(\varepsilon_t^j) = \exp(2\lambda_t^j)v_j/(v_j - 2)$ and m_t^j is a bounded function of e_t^j with $E(m_t^j|\mathcal{F}_{t-1}) = 0$. We suppose that λ_t^D and λ_t^N are linear combinations of past values of the shocks determined by m_t^j, $j = D, N$:

$$\begin{aligned} \lambda_t^D &= \omega_D(1-\beta_D) + \beta_D\lambda_{t-1}^D + \gamma_D m_{t-1}^D + \rho_D m_t^N \\ &\quad + \gamma_D^*(m_{t-1}^D + 1)\text{sign}(e_{t-1}^D) + \rho_D^*(m_t^N + 1)\text{sign}(e_t^N) \end{aligned} \tag{5}$$

$$\begin{aligned} \lambda_t^N &= \omega_N(1-\beta_N) + \beta_N\lambda_{t-1}^N + \gamma_N m_{t-1}^N + \rho_N m_{t-1}^D \\ &\quad + \rho_N^*(m_{t-1}^D + 1)\text{sign}(e_{t-1}^D) + \gamma_N^*(m_{t-1}^N + 1)\text{sign}(e_{t-1}^N). \end{aligned} \tag{6}$$

This gives two dynamic processes for the short run scale of the overnight and intraday return. The parameters ρ_D, ρ_D^* capture the effect of overnight shocks on intraday volatility, while ρ_N, ρ_N^* capture the effects of intraday shocks on overnight volatility; we call $\rho_D, \rho_D^*, \rho_N, \rho_N^*$ feedback parameters that couple together the processes λ_t^D, λ_t^N, whereas $\gamma_D, \gamma_D^*, \gamma_N, \gamma_N^*$ are capturing the effect of shocks from previous same type of period on same type of period volatility. We allow for leverage effects through the parameters $\gamma_D^*, \rho_D^*, \rho_N^*$, and γ_N^*.[2] The parameters β_D, β_N measure the persistence of the volatility processes. We set the intercepts this way so that ω_D is the unconditional mean of λ_t^D and ω_N is the unconditional mean of λ_t^N; we may consider $\exp(\omega_D - \omega_N)$ to measure the relative mean volatility contribution of the daily process and the overnight process. Let

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

be the finite dimensional parameters of interest. The two unknown functions $\sigma^D(\cdot)$ and $\sigma^N(\cdot)$ complete the semiparametric model for the process $\{u_t\}$.

[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, $beta\,(1/2, v_j/2)$.

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 or not obvious why 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 argues that 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 down weighted, and in such a way driven by the shape of the error distribution.[3] The DCS model has the incidental advantage that there are analytic expressions for moments, autocorrelation functions, multi-step forecasts, and the mean squared forecast errors.

Before introducing our estimation procedure we comment on some properties of our model that are useful in applications. In Supplementary Material we prove that if $|\beta_j| < 1$, $j = D, N$, then e_t^j and λ_t^j are strongly stationary and β-mixing with exponential decay. We note that although the conditional distribution of returns is symmetric about the mean, the unconditional distribution implied by our model may be asymmetric because of the conditional mean process and the asymmetric news impact curve that we allow for, He, Silvennoinen, and Teräsvirta (2008).

One use of our model is to improve risk management for day trading by updating volatility and value at risk estimates based on the opening price. In particular, the dynamic intraday value at risk conditional on overnight returns and past information is as follows

$$VaR_t^D(\alpha) = \mu_t^D + s_t^D t_\alpha(v_D), \tag{7}$$

$$\mu_t^D = E(r_t^D|\mathcal{F}_{t-1}, r_t^N) = \mu_D - \delta r_t^N - \Pi_{11} r_{t-1}^D - \Pi_{12} r_{t-1}^N$$

$$s_t^D = \mathrm{sd}(r_t^D|\mathcal{F}_{t-1}, r_t^N) = \sqrt{\frac{v_D}{v_D - 2}} \exp(\lambda_t^D) \exp(\sigma^D(t/T)),$$

where $t_\alpha(v)$ is the α quantile of the t-distribution with degrees of freedom v. Here, $\mathcal{F}_{t-1}$ is the 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 the sigma field generated by $\{r_t^N, r_{t-1}^D, r_{t-1}^N, r_{t-2}^D, r_{t-2}^N, \ldots\}$. This process depends on all components of the model, the dynamic mean, the dynamic variance, the trend variance, and the tail thickness of the shock distribution. To obtain the value at risk given only past intraday returns say, requires some further arguments and this is presented in Supplementary Material.

3 Estimation

Suppose that we know δ, μ, Π and hence u_t^j, $j = D, N$. In practice these can be replaced by root-T consistent estimators, and we shall not detail the properties of the mean estimators in the sequel as these are well known, and we shall drop them from the notation for convenience in the theoretical analysis. We next describe how we estimate the unknown quantities ϕ and $\sigma^j(.)$. For any $\alpha > 0$, we

[3]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.

have for $j = N, D$,

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

where c^j is a constant that depends in a complicated way on the parameter vector ϕ and on α. Therefore, we can estimate $\sigma^N(s), \sigma^D(s)$ as follows with kernel technology. Let $K(u)$ be a kernel with support $[-1,1]$ and h a bandwidth, and let $K_h(.) = K(./h)/h$. Then let

$$\widetilde{\sigma}^j(s) = \frac{1}{\alpha}\log\left(\frac{1}{T}\sum_{t=1}^{T} K_h(s - t/T)\left|u_t^j\right|^\alpha\right) \tag{8}$$

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 $[0,1]$, Härdle and Linton (1994). One may apply more sophisticated adjustments such as Jones, Linton, and Nielsen (1995) that preserves positivity but reduces the bias in the boundary region. For identification, we recenter $\widetilde{\sigma}^j(t/T)$ as

$$\widetilde{\sigma}^j(t/T) = \widetilde{\sigma}^j(t/T) - \frac{1}{T}\sum_{t=1}^{T}\widetilde{\sigma}^j(t/T). \tag{9}$$

Note that $\widetilde{\sigma}^j(u)$ can be written as $\widetilde{\sigma}^j(u) = \sum_{i=1}^{\infty}\widetilde{\theta}_i^j\psi_i^j(s)$ for some coefficients $\widetilde{\theta}_i^j$ (that satisfy $\sum_{i=1}^{\infty}|\widetilde{\theta}_i^j| < \infty$) determined uniquely by the estimator $\widetilde{\sigma}^j(u)$, that is, we can represent the kernel estimator as a sieve estimator with a potentially infinite number of coefficients, see Supplementary Material. We will use this representation for notational convenience, that is, we will represent $\widetilde{\sigma}^j(.)$ in terms of $\{\widetilde{\theta}_i^j\}_{i=1}^{\infty}$ or just $\widetilde{\theta}$ for shorthand. In practice, the bandwidth may be chosen by some rule of thumb method.

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)

$$\begin{aligned} l_T(\phi;\widetilde{\theta}) &= \frac{1}{T}\sum_{t=1}^{T} l_t(\phi;\widetilde{\theta}) = \frac{1}{T}\sum_{t=1}^{T}\left(l_t^N(\phi;\widetilde{\theta}) + l_t^D(\phi;\widetilde{\theta})\right), \\ l_t^j(\phi;\widetilde{\theta}) &= -\lambda_t^j(\phi;\widetilde{\theta}) - \frac{v_j+1}{2}\ln\left(1 + \frac{(\widetilde{e}_t^j)^2}{v_j\exp(2\lambda_t^j(\phi;\widetilde{\theta}))}\right) + \ln\Gamma\left(\frac{v_j+1}{2}\right) - \frac{1}{2}\ln v_j - \ln\Gamma\left(\frac{v_j}{2}\right), \end{aligned} \tag{10}$$

where Γ is the gamma function and $\lambda_t^j(\phi;\widetilde{\theta})$ are defined in (5) and (6). 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 $\phi \in \Phi$. 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 $\sigma^D(\cdot), \sigma^N(\cdot)$ with the local likelihood function in Severini and Wong (1992) given $\widetilde{\eta}_t^j$ and $\widetilde{v}_j$, i.e., we maximize the objective function

$$\widetilde{L}_T^j(\gamma; \widetilde{\lambda}^j, s) = -\frac{1}{T}\sum_{t=1}^{T} K_h(s - t/T)\left[\gamma + \frac{\widetilde{v}_j + 1}{2}\ln\left(1 + \frac{(\widetilde{\eta}_t^j \exp(-\gamma))^2}{\widetilde{v}_j}\right)\right] \tag{11}$$

with respect to $\gamma \in \mathbb{R}$, for $j = D, N$ separately, where $\widetilde{\lambda}^j = (\widetilde{\lambda}_1^j, \ldots, \widetilde{\lambda}_T^j)^\intercal$. Here, we also use a boundary kernel for $s \in [0, h] \cup [1 - h, 1]$. In practice we use Newton-Raphson iterations making use of the analytic derivatives of the objective functions, which are given in (11) in Supplementary Material. Harvey (2013) gives some discussion about computational issues. 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 (8) and (9)

STEP 2. Estimate ϕ by optimizing $l_T(\phi; \widetilde{\theta})$with respect to $\phi \in \Phi$ (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 $\sigma^j(s)$. Rescale $\widehat{\sigma}^j(t/T) = \widehat{\sigma}^j(t/T) - \frac{1}{T}\sum_{t=1}^{T}\widehat{\sigma}^j(t/T) = \sum_{i=1}^{\infty}\widehat{\theta}_i^j\psi_i^j(s)$. Update ϕ by optimizing $l_T(\phi; \widehat{\theta})$with respect to $\phi \in \Phi$ to give $\widehat{\phi}$.

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. The proofs of all results are given in Supplementary Material. Let $h_t^j = \lambda_t^j + \sigma^j(t/T)$, and let:

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

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

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

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

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

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. $0 \le |\beta_j| < 1$.

3. h_t^j *starts from the 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 functions* σ^j *are twice continuously differentiable on* $[0,\ 1], j = D, N$.

5. $E|u_t^j|^{(2+\delta)\alpha} < \infty$ *for some* $\delta > 0$, $j = D, N$.

6. *The function* $l(\phi) = E(l_T(\phi; \theta_0))$ *is uniquely maximized at* $\phi = \phi_0$.

7. *The kernel function* K *is bounded, symmetric about zero with compact support, that is* $K(s) = 0$ for all $|s| > C_1$ with some $C_1 < \infty$. *Moreover, it is Lipschitz, that is* $|K(s) - K(s')| \le L|s - s'|$ *for some* $L < \infty$ *and all* $s, s' \in \mathbb{R}$. *Denote* $||K||_2^2 = \int K(s)^2 ds$.

8. $h(T) \to 0$,*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 (2012). But we only require that $E|u_t^j|^{\alpha(2+\delta)} < \infty$, since we use $\widetilde{\sigma}^j(s) = \log(T^{-1}\sum_{t=1}^{T} K_h(s - t/T)|u_t^j|^\alpha)/\alpha$. This is in line with the fact that the fourth-order moment of overnight returns may not exist for some datasets. 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.

Lemma 1 in Supplementary Material gives the uniform convergence rate of the initial estimator $\widetilde{\sigma}^j(s)$, which is close to $T^{-2/5}$ when $h = O(T^{-1/5})$. The proof mainly follows Theorem 3 in Vogt and Linton (2014). We note that our initial estimator is robust to the specification of the short run dynamic process in the sense that Lemma 1 continues to hold under the weak dependence assumptions for whatever stationary mixing process is assumed for λ_t^j.

We next present an important orthogonality condition that allows us to establish a simple theory for the parametric component.

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*

$$\frac{1}{T}\sum_{t=1}^{T} E\left[\frac{\partial l_t(\phi_0;\theta_0)}{\partial \theta_k}\frac{\partial l_t(\phi_0;\theta_0)}{\partial \phi_i}\right] = o(T^{-1/2}).$$

The proof of Theorem 1 is provided in Supplementary Material. Theorem 1 implies that the score functions with respect to θ and ϕ are asymptotically orthogonal. The intuition behind this is that σ^j is a function of a deterministic variable, t/T, while $\lambda_t^j$is a stationary process independent of time t. The cross product of their score functions can be somehow separated, see Linton (1993) for a similar result. The asymptotic orthogonality implies that the particular asymptotic property of $\tilde{\phi}$ and $\hat{\phi}$ in Theorem 2 follows.

Define the asymptotic information matrix

$$\mathcal{I}(\phi_0) = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} E\left[\frac{\partial l_t(\phi_0;\theta_0)}{\partial \phi}\frac{\partial l_t(\phi_0;\theta_0)}{\partial \phi^{\intercal}}\right].$$

Theorem 2 *Suppose that Assumptions A1-A8 hold. Then*

$$\sqrt{T}\left(\tilde{\phi} - \phi_0\right) = \sqrt{T}\left(\widehat{\phi} - \phi_0\right) + o_P(1) \Longrightarrow N\left(0, \mathcal{I}(\phi_0)^{-1}\right).$$

Theorem 3 *Suppose that Assumptions A1-A8 hold. Then for* $s \in (0,1)$

$$\sqrt{Th}\begin{pmatrix} \widehat{\sigma}^D(s) \\ \widehat{\sigma}^N(s) \end{pmatrix} - \begin{pmatrix} \sigma_0^D(s) \\ \sigma_0^N(s) \end{pmatrix} \Longrightarrow N\left(0, ||K||_2^2 \begin{pmatrix} \frac{(v_D+3)}{2v_D} & 0 \\ 0 & \frac{(v_N+3)}{2v_N} \end{pmatrix}\right). \tag{12}$$

Theorem 2 shows that $\tilde{\phi}$ and $\hat{\phi}$ have the same asymptotic property and are efficient. The form of the limiting variance in (12) is consistent with the known Fisher information for the estimation of scale parameters 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 Supplementary Material. The information matrix, $\mathcal{I}(\phi_0)$, can be computed explicitly. We can conduct inference with Theorem 2 and Theorem 3 using plug-in estimates of the unknown quantities. In the application we present various Wald statistics for testing hypotheses about ϕ such as: the absence of leverage effects, the absence of feedback effects, and the equality of intraday and overnight parameters.

Test of constancy of the ratio of long run components. We next provide a test of the constancy of the ratio of long run overnight to intraday volatility. We consider the null hypothesis to be

$$H_0 : \exp\left(\sigma_0^N(s)\right) = \rho \exp\left(\sigma_0^D(s)\right) \text{ for some } \rho \in \mathbb{R}_+, \quad \text{for all } s \in (0,1), \tag{13}$$

versus the general alternative that the ratio $\exp\left(\sigma_0^N(s)\right) / \exp\left(\sigma_0^D(s)\right)$ is time varying. 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. Therefore, consider the t-ratio

$$\widehat{t}(s) = \frac{\sqrt{Th}\left(\widehat{\rho}(s) - \widehat{\rho}\right)}{\sqrt{\widehat{\omega}(s)}}, \tag{14}$$

$$\widehat{\rho}(s) = \frac{\exp(\widehat{\sigma}^N(s))}{\exp(\widehat{\sigma}^D(s))}, \qquad \widehat{\rho} = \int_0^1 \frac{\exp(\widehat{\sigma}^N(s))}{\exp(\widehat{\sigma}^D(s))} ds$$

$$\widehat{\omega}(s) = \widehat{\rho}^2 ||K||_2^2 \left(\frac{\widehat{v}_N + 3}{2\widehat{v}_N} + \frac{\widehat{v}_D + 3}{2\widehat{v}_D} \right).$$

Large values of $|\widehat{t}(s)|$ are inconsistent with the null hypothesis. It follows from Theorem 3 that for $s \in (0,1)$, $\widehat{t}(s) \Longrightarrow N(0,1)$ under the null hypothesis. We may carry out the pointwise test statistic or confidence interval based on this. We also consider an integrated version of this, specifically let

$$\tau = \frac{\int \widehat{t}(s)^2 dW_T(s) - a_T}{b_T}, \tag{15}$$

where $W_T(.)$ is some weighting function, for example Lebesgue measure on $[0,1]$, and a_T, b_T are constants. This test statistic is similar to for example Fan and Li (1996). Under the null hypothesis, $E(\widehat{t}(s)^2) \simeq 1$ so we take $a_T = 1$. Under the null hypothesis

$$\text{var}\left(\int \widehat{t}(s)^2 dW_T(s) \right) = E\left(\int \int \widehat{t}(s)^2 \widehat{t}(r)^2 dW_T(s) dW_T(r) \right) - 1.$$

In the special case that W_T is the measure that puts equal mass on the points $s_1, \ldots, s_M$ with $M = O(Th)$ so that $\widehat{t}(s_l)$ and $\widehat{t}(s_k)$ are asymptotically independent for $l \neq k$, we may take $b_T = \sqrt{2}$, because $E(\widehat{t}(s)^4) \simeq 3$. Under the null hypothesis, $\tau \Longrightarrow N(0,1)$, while under the alternative hypothesis $\tau \to \infty$ with probability one. This testing strategy is well suited to detect general alternatives to the null hypothesis of constancy of the volatility ratio.

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.

We consider two approaches to modelling the conditional mean. 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. This dynamic model is similar to that considered in the univariate section. In the application we also consider an alternative modelling approach when

we have also market returns. In this case, we specify r_{it}^j using a microstructure-adjusted Market Model

$$r_{it}^D = a_i^D + \beta_i^{DD} r_{mt}^D + \beta_i^{DN} r_{mt}^N + u_{it}^D \tag{16}$$
$$r_{it}^N = a_i^N + \beta_i^{NN} r_{mt}^N + \beta_i^{ND} r_{mt-1}^D + u_{it}^N, \tag{17}$$

where r_{mt}^D and r_{mt}^N are the market intraday and overnight returns and r_{it}^D and r_{it}^N are the returns of stock i. Including a lagged return in the market model to account for microstructure goes back to Scholes and Williams (1977).

We now consider the specification of the variance equation for the errors u_t. We suppose that

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

where: ε_{it}^j are i.i.d. shocks (mutually independent across i, j, and t, identically distributed over t) 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. They allow slowly evolving correlation between stocks in the day or night, and for those correlations to vary by stock and over time.

We can write the covariance matrices in terms of the correlation matrices and the variances as follows

$$\Sigma^j(s) = \text{diag}\left(\exp(\sigma^j(s))\right) R^j(s) \text{diag}\left(\exp(\sigma^j(s))\right), \quad j = D, N, \tag{18}$$

with $\text{diag}\left(\exp(\sigma^j(s))\right)$ being the volatility matrix and $R^j(s)$ being the correlation matrix with unit diagonal elements and off-diagonal elements $R_{il}^j(s)$ with $-1 \leq R_{il}^j(s) \leq 1$. For identification, we still assume $\int_0^1 \sigma_i^j(s)ds = 0$, for $i \in \{1, \ldots, n\}$ and $j = D, N$.

As with the univariate model, define $e_t^j = \text{diag}(\exp(\lambda_t^j))\varepsilon_t^j \in \mathbb{R}^n$, 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 $\phi_i = (\omega_{iD}, \beta_{iD}, \gamma_{iD}, \gamma_{iD}^*, \rho_{iD}, \rho_{iD}^*, v_{iD}, \omega_{iN}, \beta_{iN}, \gamma_{iN}, \gamma_{iN}^*, \rho_{iN}, \rho_{iN}^*, v_{iN})^\intercal \in \Phi \subset \mathbb{R}^{14}$, and let $\phi = (\phi_1^\intercal, \ldots, \phi_n^\intercal)^\intercal$ denote all the dynamic parameters.

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

$$l_T(\phi, \Sigma(\cdot)) = \frac{1}{T}\sum_{t=1}^{T} l_t^N(\phi, \Sigma(\cdot)) + l_t^D(\phi, \Sigma(\cdot))$$

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

We first define an initial estimator for $\Sigma^j(t/T)$ 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 $\Sigma^j(t/T)$ that is robust to heavy tails, we estimate the volatility parameter

$$\widetilde{\sigma}_i^j(s) = \frac{1}{\alpha}\log\left(\frac{1}{T}\sum_{t=1}^T K_h(s-t/T)\left|u_{it}^j\right|^\alpha\right), \tag{19}$$

and then rescale $\widetilde{\sigma}^j(t/T)$ as

$$\widetilde{\sigma}_i^j(t/T) = \widetilde{\sigma}_i^j(t/T) - \frac{1}{T}\sum_{t=1}^T \widetilde{\sigma}_i^j(t/T). \tag{20}$$

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

$$\widetilde{R}_{ik}^j(s) = \frac{\sum_{t=1}^T K_h(s-t/T)u_{ik}^j u_{ik}^j}{\sqrt{\sum_{t=1}^T K_h(s-\frac{t}{T})u_{it}^j u_{it}^j \sum_{t=1}^T K_h(s-\frac{t}{T})u_{kt}^j u_{kt}^j}} \tag{21}$$

for $s \in (0,1)$, and boundary modification as previously detailed. Alternatively, we can use a robust correlation estimator. Omitting the superscript $j = D, N$ here, we may compute the pairwise Kendall tau

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

Then 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, $\hat{\rho}_{k,l}(s) = \sin(\frac{\pi}{2}\hat{\tau}_{k,l}(s))$. In some cases, the matrix of pairwise correlations must be adjusted to ensure that the resulting matrix is positive definite.

We have

$$\widetilde{\Sigma}^j(s) = \mathrm{diag}\left(\exp(\widetilde{\sigma}^j(s))\right)\widetilde{R}^j(s)\mathrm{diag}\left(\exp(\widetilde{\sigma}^j(s))\right),\ j = D, N. \tag{22}$$

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 (10) for each $i = 1, \ldots, n$. 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 (13) and (14) in Supplementary Material.

To summarize, the estimation algorithm is as follows.

Algorithm

STEP 1. Estimate $\mathbb{D}, \mu, \Pi$ (or the CAPM structure)by least squares, $\widetilde{\sigma}^j(u)$, $u \in [0,1]$, $j = N, D$ from (19) and (20), and $\widetilde{R}^j_{ik}$ from (21).

STEP 2. Let $\widetilde{e}^j_t = \widetilde{\Sigma}^j(\frac{t}{T})^{-1/2} u^j_t$. We obtain $\widetilde{\phi}_i$ by maximizing the univariate log-likelihood function of $\widetilde{e}^j_{it}$ in (10) for each $i = 1, \ldots, n$.

STEP 3. We replace λ^j_t with $\widetilde{\lambda}^j_t = \lambda^j_t(\widetilde{\phi}; \widetilde{\theta})$. For each $\Sigma_j(t/T)$, denote $\Theta^j = (\Sigma^j)^{-1/2}$. Then obtain $\widehat{\Theta}$ with the local likelihood function given $\widetilde{\lambda}^j_t$ and $\widetilde{v}_j$, with the Newton-Raphson iterations making use of the derivatives of the objective functions.

STEP 4. Repeat Steps 2-3 to update $\widehat{\phi}$ and $\widehat{\Sigma}^j(t/T)$ until convergence.

Our multivariate model can be considered as a diagonal DCS EGARCH model with a slowly moving correlation matrix. Assuming diagonality on the short run component λ^j_t enables us to estimate the model easily and rapidly. In particular, 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. We do not provide the distribution theory here for space reasons but it follows by similar arguments to given for the univariate case.

Blanc, Chicheportiche, and Bouchaud (2014) impose a pooling assumption in their modelling, which translates here to the restriction that $\phi_i = \phi_1$ for all $i = 1, \ldots, n$. This improves efficiency when the restriction is true. We can test the restriction by a standard Wald procedure or Likelihood ratio statistic. In the application we find these pooling restrictions are strongly rejected by the data.

6 Empirical application

In this section, we apply our coupled-component GARCH model to empirical data. We first apply the model to the Dow Jones stocks, and we report detailed estimates and examine the out-of-sample forecast performance. Then, we investigate the overnight to intraday volatility and correlation in the U.S. size-based portfolios. Finally, we extend our study to the four largest international stock markets (U.S., U.K., Japan, and Germany).

6.1 Application to Dow Jones stocks

6.1.1 Data and preliminary analysis

We investigate 26 components of the Dow Jones industrial average index during the period of 4 January 1993 to 29 December 2017. The 26 stocks are AAPL, MSFT,, XOM, JNJ, INTC, WMT,

CVX, UNH, CSCO, HD, PFE, BA, VZ, PG, KO, MRK, DIS, IBM, GE, MCD, MMM, NKE, UTX, CAT, AXP, and TRV. [4] The data are obtained from Datastream, and the prices have been adjusted for corporate actions. We define overnight returns as the log price change between the close of one trading day and the opening of the next trading day. We do not incorporate weekend and holiday effects into our model as they are not the focus of this paper. In addition, although the weekend effect is documented by studies such as 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, including Mehdian and Perry (2001) and Steeley (2001). In addition, Sullivan, Timmermann, and White (2001) claim that many calendar effects arise from data-snooping.

Many studies, such as Cooper, Cliff, and Gulen (2008) and Berkman, Koch, Tuttle, and Zhang (2012), find significantly higher overnight returns. Cooper, Cliff, and Gulen (2008) even suggest that the U.S. equity premium during their research period is solely due to overnight returns. Figure A.1 in Supplementary Material plots the cumulative returns for these 26 stocks. There is no clear dominance of positive overnight returns from these Dow Jones stocks.

Berkman, Koch, Tuttle, and Zhang (2012) find significant positive mean overnight returns of +10 basis points (bp) per day, along with -7 bp for intraday returns from the 3000 largest U.S. stocks. Following Berkman, Koch, Tuttle, and Zhang (2012), we first compute the cross-sectional mean returns for each day, then compute the time-series mean and standard deviation of these values. The mean intraday return is 2.05 bp with a standard error of 1.12 bp, while the mean overnight return is 1.68 bp with a standard error of 0.71 bp. The difference between overnight and intraday means is not statistically significant.

Table A.1 in Supplementary Material provides summary statistics for intraday and overnight returns. Compared with intraday returns, overnight returns exhibit more negative skewness and leptokurtosis. Specifically, 9 of these 26 stocks exhibit negative intraday skewness, while 25 of these 26 stocks have negative overnight skewness. The largest sample kurtosis for overnight returns is 935.78, which suggests the non-existence of the population kurtosis. We find that the per hour variance of intraday returns is roughly 12 times the per hour variance of overnight returns, which is somewhat less than the range of 13-100 times found by French and Roll (1986).[5]

[4] These stocks are constituents of the Dow Jones index according to the constituent list in May 2018. The V, GS, and DWDP.K are excluded because they do not have prices available in 1993. The JPM is excluded because its open price from 2 September 1993 to 4 January 1995 is missing in Datastream.

[5] Suppose that hourly stock returns satisfy

$$r_{ht} \sim \mu_h, \sigma_h^2, \kappa_{3h}, \kappa_{4h},$$

which is consistent with French and Roll (1986). Daily (based on a 6-hour trading day) and weekend (66 hours from Friday close to Monday open) returns should then satisfy

$$r_{Dt} \sim 6\mu_h, 6\sigma_h^2, \frac{\kappa_{3h}}{\sqrt{6}}, \frac{\kappa_{4h}}{6} \quad ; \quad r_{Wt} \sim 66\mu_h, 66\sigma_h^2, \frac{\kappa_{3h}}{\sqrt{(66)}}, \frac{\kappa_{4h}}{66}.$$

In fact, overnight returns including weekend returns are very leptokurtic.

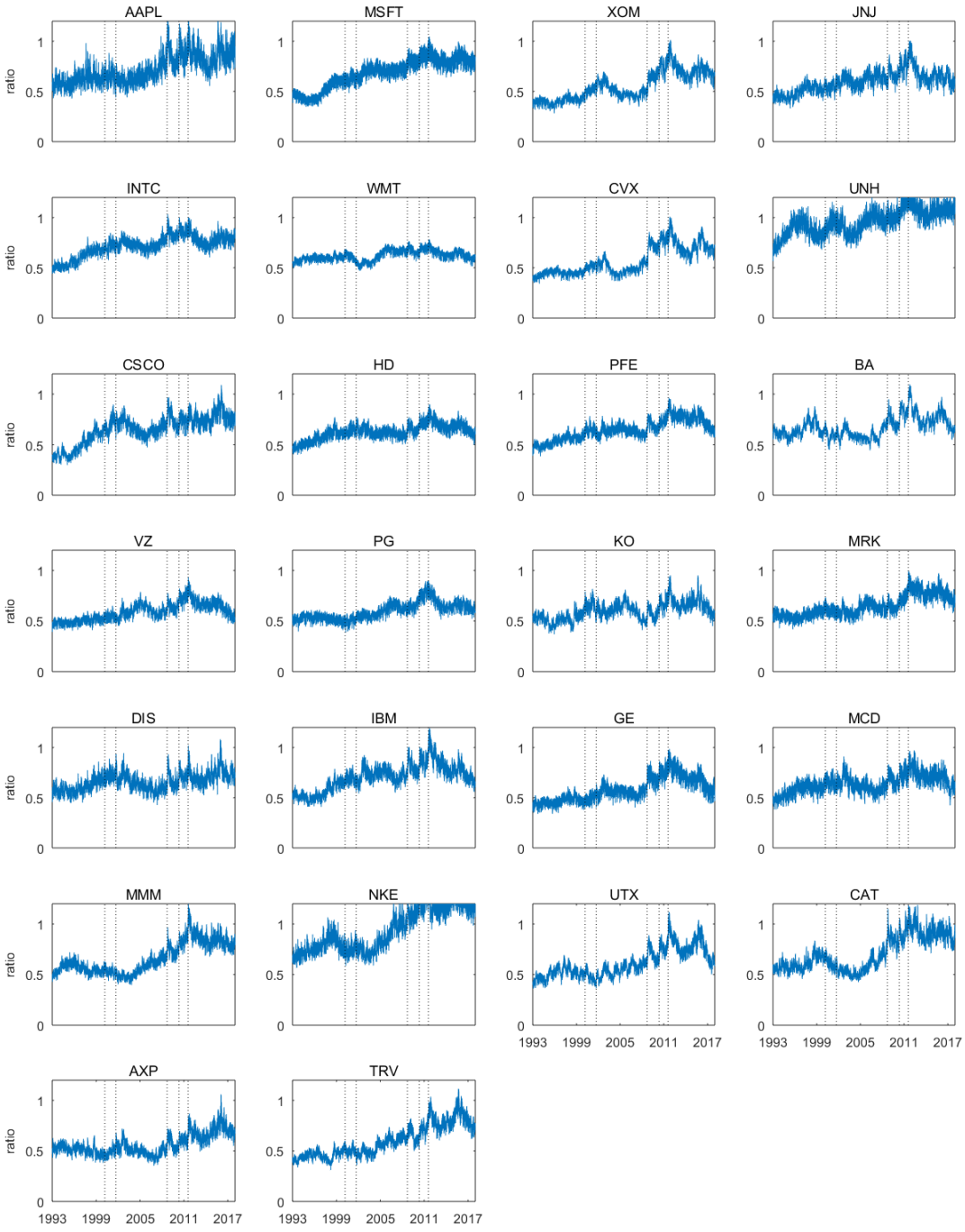
6.1.2 Results of the univariate model

We estimate the univariate coupled-component GARCH model for each Dow Jones stocks. The estimates and their robust standard errors in the mean equations are reported in Table A.2 in Supplementary Material. We multiply returns by 100 to give more readable coefficients. 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 26stocks have significant Π_{11} values that are all negative, and 7 of 28 stocks have significant δ values which are all positive. This outcome suggests that both overnight and intraday returns tend to have a negative effect on the subsequent intraday return. 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 A.3 in Supplementary Material gives the estimates of dynamic parameters. Parameters β_D and β_N are significantly different from 1, and ρ_D, γ_D, ρ_N and γ_N are positive and significant. In addition, we find significant leverage effects, with negative and significant ρ_D^*, γ_D^*, ρ_N^*, and γ_N^*, which suggest higher volatility after negative returns.

We are also concerned about the difference between overnight and intraday parameters. Table A.4 in Supplementary Material reports Wald tests with the null hypothesis that the intraday and overnight parameters are equal within each stock. The parameter ω_D, which determines the unconditional short-run scale, is significantly larger than ω_N. The overnight degree-of-freedom parameter is around 3, which is significantly smaller than the intraday counterpart at approximately 8. Both are in line with the descriptive statistics in Table A.1 as well as previous studies suggesting 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 weakly affect future intraday volatilities, except for the very next one, but have a substantial impact on future overnight volatilities. This inconsistency is probably because the dynamic conditional score model shrinks the impact of extreme overnight observations. After this shrinkage, the effect of overnight innovations on parameter estimation becomes closer to the intraday innovations.

Many papers have argued that the introduction of high-frequency trading in the period post 2005 has led to an increase in volatility. Figure A.2 in Supplementary Material plots the intraday and overnight volatilities, $\sqrt{\frac{\nu_j}{\nu_j-2}\exp(2\lambda_t^j + 2\sigma^j(\frac{t}{T}))}$, for $j = D, N$. The five dashed 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), 6 May 2010 (flash crash), and 1 August 2011 (August 2011 stock markets fall). The intraday volatility (red lines) significantly dominates the overnight volatility (black lines) in the first half of the study period, but this domination gradually disappears, especially after the 2008 financial crisis. In addition, the intraday volatilities after 2005 are in general smaller than those before 2005, except for the financial crisis period. This finding is contrary to the typical argument that high-frequency trading increases volatilities. To further investigate this point, we plot the ratios of overnight to intraday volatility in Figure 1. All stocks exhibit upward trends



This figure shows the dynamic ratio of overnight to intraday volatility, based on the univariate coupled-component model, with one subplot for each stock. The five dashed 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), 6 May 2010 (flash crash), and 1 August 2011 (August 2011 stock markets fall), respectively. Intraday and overnight volatiles are defined as $\sqrt{\frac{\nu_j}{\nu_j-2}exp(2\lambda_t^j + 2\sigma^j(\frac{t}{T}))}$, for $j = D, N$.

Figure 1: Ratios of overnight to intraday volatility: univariate model

over the 25-year period considered here, and many of them experience peaks around August 2011, corresponding to the August 2011 stock markets fall.

6.1.3 Constancy of the ratio of overnight to intraday volatility

The long-run intraday and overnight components, $\sigma^D(t/T)$ and $\sigma^N(t/T)$, and their 95% point-wise confidence intervals are depicted in Figure A.3 in Supplementary Material. Most stocks arrive at their first peaks around 10 March 2000, corresponding to the dot-com bubble event, while some arrive at around September 2011, right after 9-11. The intraday components reach their second peaks during the financial crisis in September 2008, while overnight components continue to rise until around 2011. Roughly speaking, the intraday components are larger than the overnight ones before the first peaks, but smaller after the financial crisis of September 2008. However, it is imperative to remember that the long-run components are constructed with rescaling $\int_0^1 \sigma(s)ds = 0$. In general, the intraday volatility is still larger.

We test the constancy of the ratio of long run overnight to intraday volatility. Figure A.4 in Supplementary Material displays the test statistics $\widehat{t}(s)$ 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).

Cumulatively, this evidence indicates that the overnight volatility has increased in importance during the 25-year period considered here, relative to the intraday volatility for the Dow Jones stocks.

6.1.4 Volatility forecast comparison

We also compare our coupled-component GARCH model with its one-component version for the open-to-close returns to assess 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-sided weight function $K_h(s - t/T) = 0$, for $t >= \tau$, and then rescale $K_h(s - t/T)$ to obtain a sum of 1. Table A.5 in Supplementary Material 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.

6.1.5 Diagnostic tests

Ljung-Box tests on the absolute and squared standardized residuals are used to verify whether the coupled-component GARCH model is adequate to capture the heteroskedasticity, shown in Table A.6

in Supplementary Material. With the absolute form, strong heteroskedasticity exists in both intraday and overnight returns but disappears in the standardized residuals, implying that our model captures the heteroskedasticity well. However, we are sometimes unable to detect the heteroskedasticity in overnight returns with squared values. In general, the use of the absolute form is more robust when the distribution is heavy tailed.

Figure A.5 in Supplementary Material displays the quantile-quantile (Q-Q) plots of the intraday innovations, comparing these with the student t distribution with $\hat{\nu}_D$ degrees of freedom. The points in the Q-Q plots approximately lie on a line, showing that the intraday innovations closely approximate the t distribution. Figure A.6 in in Supplementary Material 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.

6.1.6 Results of the multivariate model

Figure A.7 in Supplementary Material presents the long-run correlations between intraday or overnight returns. Each subplot presents the averaged correlations between that individual stock and the remaining stocks. The correlations exhibit an obvious upward trend during the sample period of 1998-2016. In the 1990s, the overnight correlations and intraday correlations are both around 0.2, albeit with fluctuations. In the period 2000 to 2007, intraday correlations start to increase and are larger than the overnight correlations. However, during the period 2008 to 2016, overnight correlations increase substantially to around 0.7 in 2011 and remain higher than 0.5, while intraday correlations peak in around 2008 but the correlations are seldom larger than 0.5. Both correlations start to decrease in 2017.

Figure A.8 in Supplementary Material plots the eigenvalues of the dynamic covariance matrices, as well as their the scaled eigenvalues (the eigenvalues divided by the sum of eigenvalues). The dynamic of eigenvalues reinforces the previous remark that the stock markets experienced high intraday risk in the 9-11 attacks in 2001 and in the 2008 financial crisis, while stock markets experienced high overnight risk in around 2011. The largest eigenvalue represents a strong common component, illustrating that a large proportion of the market financial risk can be explained by a single factor. The largest eigenvalue increases substantially during our research period. The second and third largest eigenvalues still account for a considerable proportion of risk in the volatile period from 2000 to 2002, but become rather insignificant in the volatile period from 2008 to 2011. The largest intraday eigenvalue proportion reaches its peak in 2008, while the largest overnight eigenvalue proportion remains consistently high until 2011. Remarkably, the largest eigenvalue explains nearly 50% of intraday risk in the 2008 financial crisis and 70% of overnight risk in the August 2011 stock markets fall. The overnight eigenvalue proportion is much higher than its intraday counterpart in the period 2008 to 2016. Generally speaking, the market risk in the crisis period from 2008 to 2011 can be largely explained by a single-factor structure, in particular, the overnight risk. This is in line with the finding of Li, Viktor, and George (2017) that stocks returns tend to obey an exact one-factor structure at times of market-wide jump events.

One concern is that our initial correlation estimator is based on the Pearson product moment

correlation. This Pearson estimator may perform poorly because of the heavy tails of overnight innovations. Therefore, we also try the robust correlation estimator in the initial step, yet the results remain unchanged, as shown in Figure A.9 in Supplementary Material. This figure plots the largest scaled eigenvalue of the estimated covariance matrix to assess the difference between using robust (in black) and non-robust (in red) correlation estimators in the initial step. We use dashed lines for the initial estimators and solid lines for the updated estimators. Despite the large difference of initial estimators, particularly for overnight returns, the updated estimators are roughly similar. Like the eigenvalues, the updated covariances themselves are also robust to a different initial estimator.

6.2 Application to CRSP size-based portfolios

6.2.1 Data and preliminary analysis

In this section, we investigate 10 size-based portfolios with CRSP stocks from January 1993 to December 2017. The prices are adjusted for stock splits and dividends with the cumulative factor in CRSP. Stocks with a close price less than 1 dollar at the end of the previous year are excluded. Stocks with a non-active trade status are excluded for that day, as well as stocks with overnight or intraday returns larger than 50% in absolute value.[6] CRSP sorts all stocks into 10 deciles based on their market capitalization values and provides the portfolio assignments for each stock in each year. We construct three versions of value-weighted intraday and overnight returns for these 10 size-based deciles, according to the assignment of three market types: NYSE/AMEX/NASDAQ, NASDAQ, and NYSE.

Table A.8 in Supplementary Material reports descriptive statistics for overnight and intraday returns of the deciles constructed with NYSE/AMEX/NASDAQ stocks (Panel a). The standard errors are estimated based on the standard deviations of the value-weighted returns across time. Overnight returns tend to be larger than intraday returns for large and middle stocks, while overnight returns tend to be smaller than intraday returns for small stocks. In particular, with the NYSE/AMEX/NASDAQ stocks, the mean overnight returns are significantly positive in decile 5-10 and significantly negative in decile 1-3. In contrast, the intraday returns are significantly negative in decile 3-7, while significantly positive in decile 1. If only with the NYSE size-based portfolios, most overnight and intraday returns are insignificant, except that the overnight returns are still significantly positive in the large stocks (decile 8-9) and almost significantly positive in the largest decile. If only with the NASDAQ size-based portfolios, most overnight and intraday returns are significant. Namely, overnight returns are significantly positive in large stocks (decile 6-10) and significantly negative in small stocks (decile 1-4), while intraday returns are significantly negative in large and middle stocks (decile 4-10). Most median values are larger than the mean values, due to the negative skewness. The median values of intraday returns are positive in all size-based portfolios, while the median values of overnight returns are positive in middle-cap and large-cap portfolios but negative in small-cap portfolios. We next compare large and small stocks. Small stocks have significantly higher intra-

[6] In CRSP, if the close price is not available on any given trading day, the number in the price field has a negative sign to indicate that it is a bid/ask average and not an actual close price. We exclude these negative close prices.

day returns than large stocks but significantly lower overnight returns than large stocks with the NYSE/AMEX/NASDAQ portfolios and with the NASDAQ portfolios. This pattern is consistent with Polk, Lou, and Skouras (2018) in that they suggest that the well-known small-stock effect is entirely an intraday phenomenon.

Figure A.10 in Supplementary Material plots the sample autocorrelation function of intraday and overnight returns for deciles with stocks on NYSE/AMEX/NASDAQ. Both intraday and overnight returns exhibit large and significant positive autocorrelations in small-cap deciles. Notice that the autocorrelations of overnight returns decay at a very slow rate, which suggests the existence of long memory. This finding is in line with Aboody, Even-Tov, Lehavy, and Trueman (2018), who document that overnight returns are persistent for periods extending several weeks and who argue that short-term persistence is stronger for harder-to-value firms.

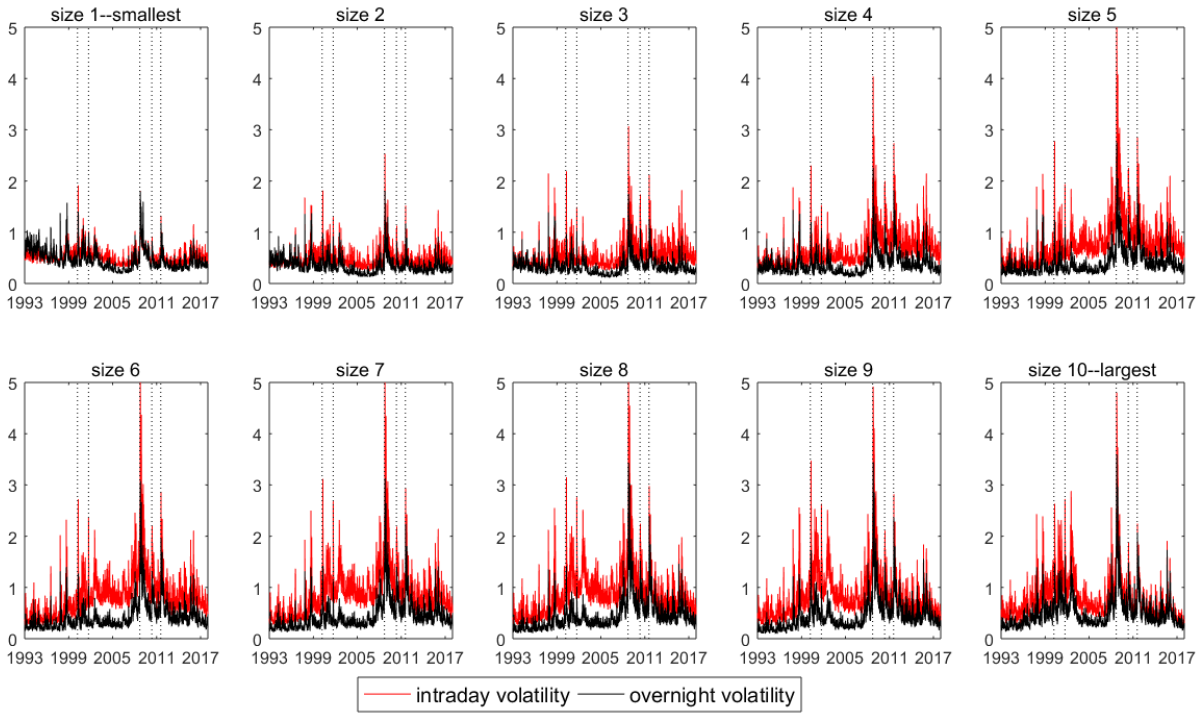
The overnight returns in large-cap deciles (deciles 5-10) show significant negative first-order autocorrelations, with a magnitude around -0.1. For intraday deciles, only the largest-cap decile exhibits significantly negative first-order autocorrelation, around -0.07. In addition to the autocorrelations, Figure A.11 in Supplementary Material shows the cross correlations between overnight returns and intraday returns. The correlation between r_t^D and r_t^N on the same day is around 0.1 and significant in most cap deciles, but it is insignificant in the largest decile. The cross correlation between r_{t-1}^D and r_t^N is significant for all cap deciles, with a magnitude that is even larger than the correlation between r_t^D and r_t^N in some deciles. These findings indicate strong positive effects from intraday (overnight) returns to the subsequent overnight (intraday) returns.

6.2.2 Overnight and intraday volatilities of CRSP size-based portfolios

We estimate the univariate coupled-component GARCH model with the intraday and overnight returns in each size-based portfolio. Table A.9 in Supplementary Material reports the estimates and their robust standard errors in the mean equations. Most portfolios exhibit negative and significant δ, indicating that overnight returns have positive effects on the subsequent intraday returns. The smallest decile has a positive but insignificant δ, and the largest decile has a positive and insignificant δ. These positive effects seem inconsistent with the strong reversal effects reported by Berkman, Koch, Tuttle, and Zhang (2012). But note that their reversal effects describe the cross-sectional difference in returns, while our positive effects describe the time-series properties in returns.

Table A.10 in Supplementary Material reports the estimates of dynamic parameters for the size-sorted portfolios of all NYSE, AMEX, and NASDAQ stocks. Parameters β_D and β_N are significantly different from 1, and ρ_D, γ_D, ρ_N and γ_N are significantly positive. The leverage effects are also significant, suggesting higher volatility after negative returns. The overnight degrees of freedom are larger than 4 and less heavy tailed than that of individual Dow Jones stocks.

The main advantage of the coupled-component model, relative to the traditional one-component version, is that it allows investigating both the overnight and intraday volatilities. Figure 2 displays the intraday and overnight volatilities for the size-sorted portfolios of all NYSE, AMEX, and NASDAQ stocks, $\sqrt{\frac{\nu_j}{\nu_j-2}\exp(2\lambda_t^j + 2\sigma^j(\frac{t}{T}))}$, for $j = D, N$. The five dashed vertical lines from left to right indicate the dates 10 March 2000 (dot-com bubble), 11 September 2001 (the September 11



This figure shows the estimated intraday (in red) and overnight (in black) volatilities, $\sqrt{\frac{\nu_j}{\nu_j-2}exp(2\lambda_t^j + 2\sigma^j(\frac{t}{T}))}$, for size-sorted deciles of NYSE/AMEX/NASDAQ stocks. The five dashed 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), 6 May 2010 (flash crash), and 1 August 2011 (August 2011 stock markets fall), respectively.

Figure 2: Intraday and overnight volatilities: NYSE/AMEX/NASDAQ

attacks), 16 September 2008 (financial crisis), 6 May 2010 (flash crash), and 1 August 2011 (August 2011 stock markets fall). The overnight volatilities (black lines) are typically high during the 2008 financial crisis period for all deciles. Since Asian and European markets are active during the U.S. overnight, we may expect overnight volatilities are driven more by oversea shocks. We indeed find overnight volatilities remain high throughout the period 2009 to 2012, the period with European sovereign debt crisis. After that, overnight volatilities become smaller, but still higher than the pre-crisis level. Figure 3 presents the ratio of overnight to intraday volatility of those deciles. The ratio exhibits a downward trend in small-cap portfolios (in particular before 2000) and an upward trend in large-cap portfolios. Notably, the trend changes monotonically from the smallest-cap portfolio to the largest-cap portfolio. The ratios are typically high during the 2011 stock markets fall probably due to the fear of the European sovereign debt crisis. This finding is consistent with the fact that, during the 2011 crisis, many large shocks stemmed from European markets and occurred overnight in the U.S. market.

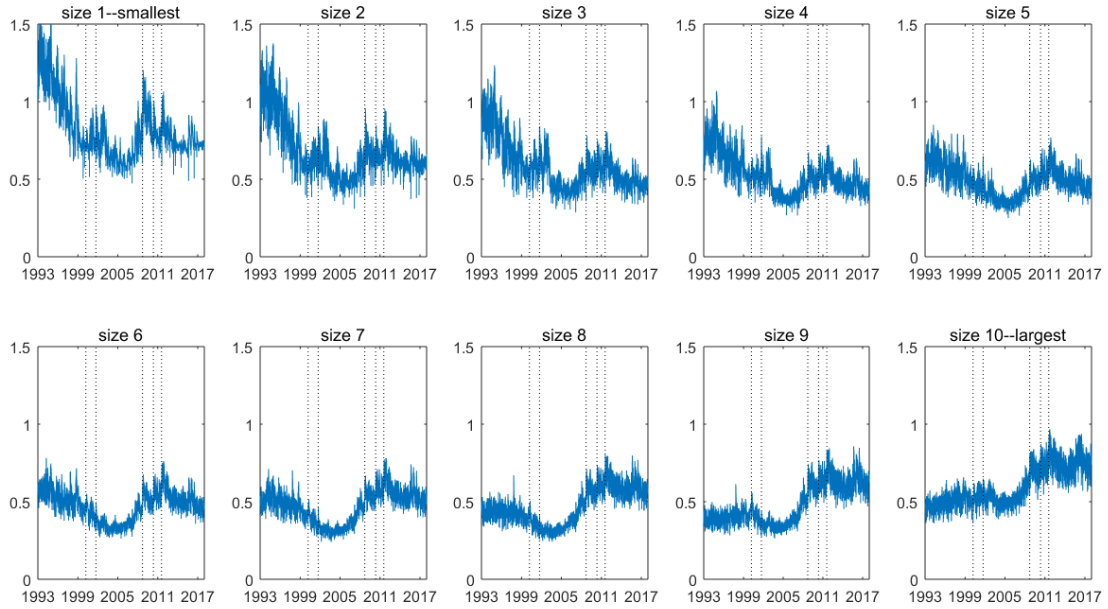
Results remain unchanged when we consider NASDAQ stocks or NYSE stocks alone (Figure A.12 in Supplementary Material). However, the slopes are generally flatter for NYSE stocks. It is probably because the small stocks in NYSE are in general much larger than the small stocks in NASDAQ.

With the same approach, we also conduct 10 beta-sorted portfolios and 10 standard deviation-sorted portfolios. No large difference is observed in the overnight to intraday volatility ratio across these portfolios (Figure A.13 in Supplementary Material). Nearly all beta-sorted and standard deviation-sorted deciles exhibit increasing overnight to intraday volatility ratios, with their highest values during the European sovereign debt crisis.

6.2.3 Explain the dynamics of the overnight to intraday volatility ratio

The empirical results show that the ratio of overnight to intraday volatility for large stocks has increased during the last 25 years when accounting for both slowly changing and rapidly changing components. This finding conflicts with what is often argued with regard to the change in market structure and the predatory practices of certain traders. Portfolios of small stocks on the other hand seem to exhibit a different trend. A number of possible explanations exist for this phenomenon, and we focus on two aspects: the change in trading mechanisms and the variation in international linkage.

First, the trading mechanism has evolved considerably from the early part of the sample. For instance, in the early 1990s, the minimum tick size was largely 12.5 cents, and NASDAQ dealers were found colluding implicitly to avoid odd-eighth quotes in order to maintain spreads of at least 25 cents (Christie, Harris, and Schultz (1994)). In June 1997, the minimum tick sizes decreased to 6.25 cents with the implementation of new SEC rules, and further became a penny after the decimalization in 2001, as a result, spreads for most stocks declined (Hasbrouck (2007)). We indeed observe that many stocks (even stocks with low prices) often had have 25 cents jumps in their open prices in the 1990s. For stocks with prices less than 5 dollars, 25 cents overnight jumps means larger than 5% overnight returns, which may lead to substantial overnight volatilities. With a cross-tabulation analysis of the stock price and stock size, we can easily find that small stocks tend to have low prices. This finding



This figure plots the ratio of overnight to intraday volatility for portfolios formed on size. Decile 1 is the portfolio with the smallest market capitalizations and decile 10 is the portfolio with the largest market capitalizations. The intraday and overnight volatilities are $\sqrt{\frac{\nu_j}{\nu_j-2}exp(2\lambda_t^j + 2\sigma^j(\frac{t}{T}))}$ for $j = D, N$, respectively. The five dashed 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), 6 May 2010 (flash crash), and 1 August 2011 (August 2011 stock markets fall), respectively.

Figure 3: Ratio of overnight to intraday volatility of size-based portfolios: NYSE/AMEX/NASDAQ

could explain the decrease of the overnight to intraday volatility ratio in small-cap portfolios during the 1990s.[7] In the following parts, we focus on the post-decimalization period, after February 2000, when most stocks have fairly small spreads.

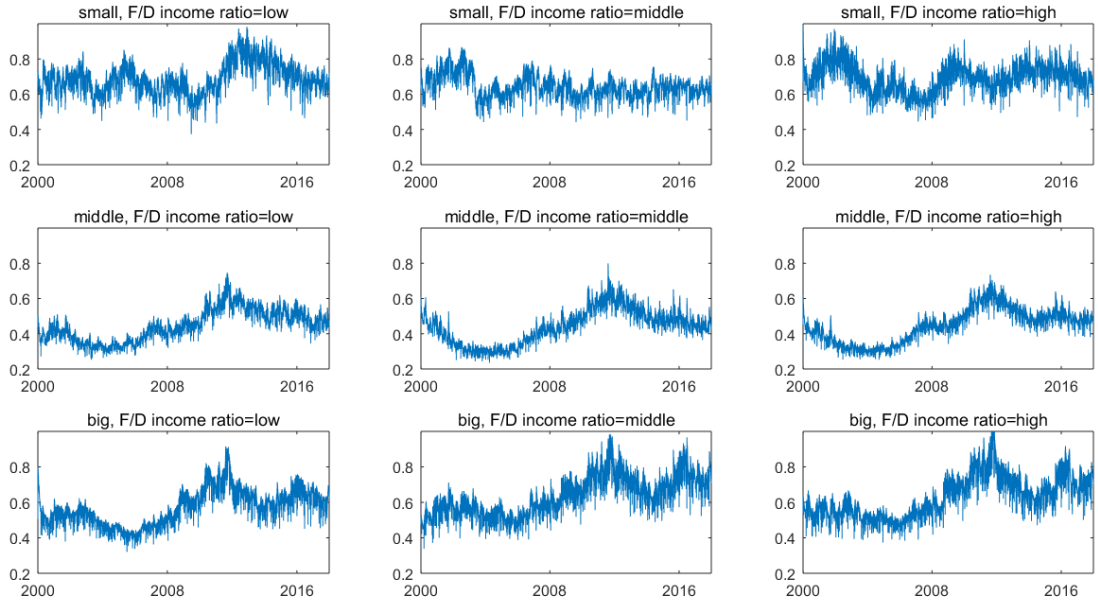
During 2004, NASDAQ launched opening cross and closing cross to give a single open and close price. To investigate whether these crosses changed the dynamics of the overnight to intraday volatility, we use NYSE as the control group, and we estimate the coupled-component GARCH model for the value-weighted NASDAQ index and the value-weighted NYSE index. Figure A.15 in Supplementary Material plots the overnight volatility, intraday volatility, and overnight to intraday volatility ratio for both NASDAQ and NYSE. The dashed vertical line indicates the last trading day in October 2004, when NASDAQ had introduced the opening and closing crosses. Both overnight and intraday volatilities are rather small for both NASDAQ and NYSE in 2004, and the launch of opening and closing crosses does not seem to change the overnight to intraday volatility ratio.

Next, we examine whether the variation in international linkage can explain the dynamics of the overnight to intraday volatility ratio. Intuitively, stocks with stronger international linkage should be more sensitive to news about international events that is released while the U.S. market is closed.

First, we use the ratio of foreign to domestic income as a proxy of international linkage. Companies with a larger proportion of income from foreign countries are presumed to have more exposure to global shocks. We collect the yearly data of pretax foreign income and pretax domestic income from COMPU.S.TAT and link them with stock prices obtained from CRSP by index 'CRU.S.IP'. About 10 percent of NYSE, AMEX, and NASDAQ stocks obtain their income data. At the end of each year, these stocks (excluding the stocks with share prices less than 1 dollar) are allocated to three size groups (small, middle, and large), with the assignment provided in CRSP. Specifically, stocks with size assignment 1, 2, or 3 are allocated to the small group; those with assignment 4, 5, 6, or 7 are allocated to the middle group; and those with assignment 8, 9, or 10 are allocated to the large group. Independently, we sort NYSE, AMEX, and NASDAQ stocks by the ratio of pretax foreign income to pretax domestic income (F/D income) and split them into three groups, using the 30th and 70th percentiles as breakpoints. The intersections of these two sorts produce nine portfolios. We then estimate the coupled-component GARCH model with the value-weighted returns of these nine portfolios. Figure 4 shows that the overnight to intraday ratios remain nearly unchanged across portfolios with different F/D income.

Second, we investigate the overnight to intraday volatility ratio across industries since some industries are supposed to have higher international linkage. We use the Fama-French 12 industry classification and form 12×3 portfolios by size and industry. The coupled-component GARCH model is then estimated for the value-weighted returns of these 36 portfolios. In general, the main variation in overnight to intraday volatility ratio still comes from the size dimension (Figure 5). These results are consistent with our findings in the previous section about Dow Jones stocks. Within the Dow Jones there is some variation across firms in terms of the percentage revenue that is earned overseas, from 85% in Intel to 0% in United Health Care, and yet both of these stocks experience a similar

[7]We construct portfolios on size and price, and we find low-price groups had a clearer downward overnight to intraday volatility ratio during 1990s. After controlling the price, the variation across size becomes weaker.



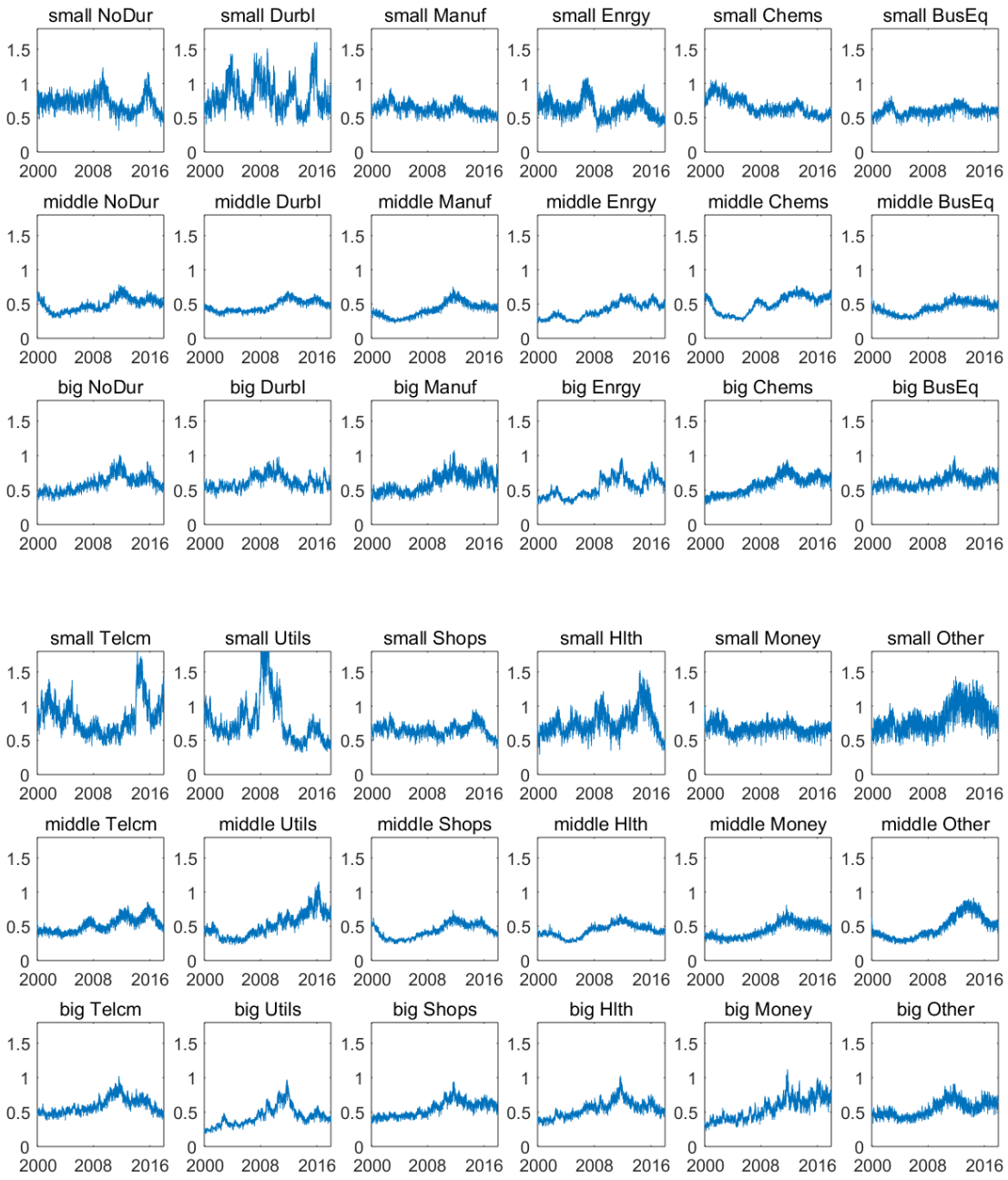
This figure plots the overnight to intraday volatility ratio for the nine portfolios formed according to size and F/D income sorts, where F/D income is the ratio of pretax foreign to domestic income.

Figure 4: Overnight to intraday volatility ratio: size and F/D income

upward trend in the ratio of overnight to intraday volatility.

Third, we use the correlation of daily returns to measure the international linkage. Specifically, for each stock, we compute the yearly sample correlation between its daily returns and the global excluding U.S. daily returns. The global ex-U.S.. index is obtained from the Kenneth French data library. We sort all stocks by their correlations in the previous year and split them into 10 groups. Independently, we split stocks into three size groups. The intersections of the sorts on correlation and size produce 30 portfolios. Figure 6 plots the overnight to intraday volatility ratio of these 30 portfolios. The overall slope of the overnight to intraday ratio changes noticeably with international correlations. The volatility ratios increase considerably in the deciles with high international correlations, indicating the financial integration in recent decades. In contrast, the ratios of portfolios with smaller international correlations remain much more stable. We also find that high international correlations stocks are likely in the large-cap portfolios, consistent with Eun, Huang, and Lai (2008). This finding can explain the phenomenon of large stocks exhibiting an increasing ratio of overnight to intraday volatility in recent decades.

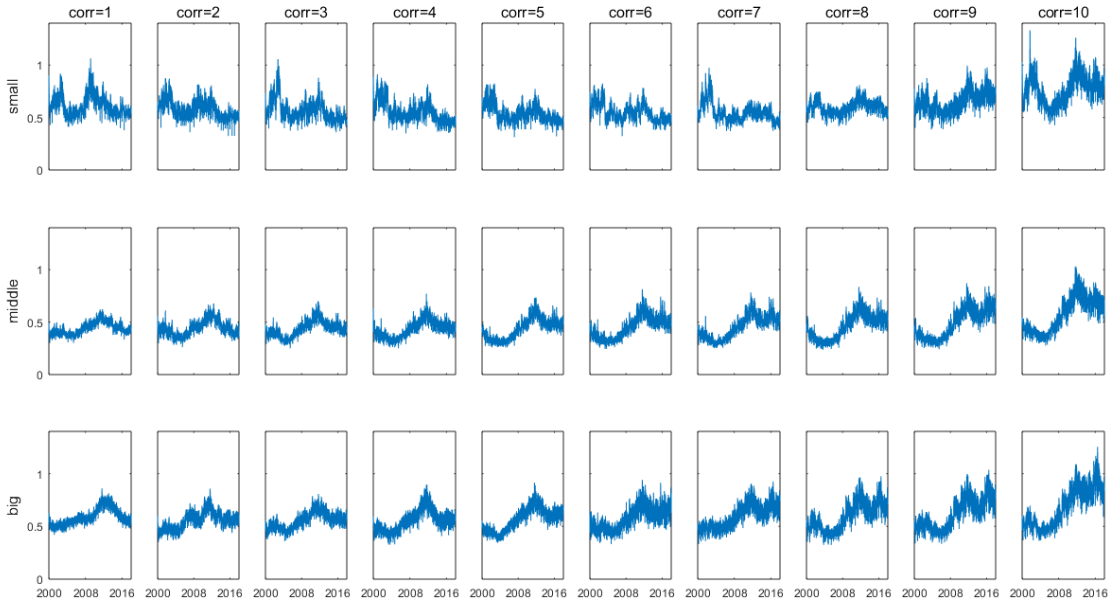
In sum, the change in the minimal tick size, which results in the improvement of market liquidity, explains the high overnight to intraday ratio in small stocks during the 1990s. The variation in international linkage, measured by the yearly correlation, also explains the size pattern in the overnight and intraday volatility ratio. But the foreign to domestic income ratio and industry category can not explain these size pattern.



This figure plots the overnight to intraday volatility ratio for portfolios formed according to size and industry. We use the Fama-French 12 industry classification.

Figure 5: Overnight to intraday volatility ratio: size and industry

Some other factors may also explain this phenomenon. For instance, one argument is that the introduction of the EMini contract and its overnight trade ability from 9 September 1997 has allowed investors to hedge their overnight positions, which should reduce the riskiness of overnight positions. The liquidity of this contract has slowly grown during our study period and so this may make a contribution to the changing overnight volatility, except that it should reduce the overnight volatility. In addition, changes in the market structure and technology for trading large firms shares have led to improvements in market quality when the market is open; that is, the electronic trading mechanism delivers less volatility and the market absorbs information faster than it used to. This is partly consistent with the findings of Boehmer, Fong, and Wu (2015) that algorithmic trading has "systematically negative effects on the liquidity of small or low-priced stocks, and AT increases volatility more in those stocks". We leave these for further study.



This figure plots the overnight to intraday volatility ratio for portfolios formed according to size and international correlations. For each stock, we compute its yearly sample correlation with the global excluding U.S. returns based on daily data. We then sort all stocks by their correlations in the previous year and split them into 10 groups.

Figure 6: Overnight to intraday volatility ratio: size and international correlation

6.2.4 Alternative volatility measures

For a robustness check of the observed rising trend of most portfolios, we investigate the ratio of VIX to the Rogers and Satchell (1991) volatility (RS volatility). The idea is that the VIX measures one-month-ahead volatility and total volatility including presumably intraday and overnight, whereas the aggregated RS volatility only includes intraday volatility. Therefore, the ratio reflects the intraday

to overnight variability to some extent, although it is quite noisy. The published open price of the S&P500 index may not reflect all the accumulated overnight information (Lin, Engle, and Ito (1994) and many others) since some component stocks may open a few minutes later. An alternative proxy for the S&P500 open price is the special opening quote (SOQ), suggested by Ahoniemi and Lanne (2013). This index is calculated using the opening values of each of the 500 component stocks, and it is also used as the typical settlement price for S&P 500 index futures. The data of VIX and SOQ are obtained from Datastream. Figure A.14 in Supplementary Material presents: (1) the RS volatility on the daily S&P500 returns, $V_{rs,t}$; (2) one-month-ahead RS volatility, $\sqrt{\sum_{i=1}^{22} V_{rs,t+1}^2}$; (3) VIX; and (4) the ratio of VIX to the one-month-ahead RS volatility. The ratio of VIX to the one-month-ahead RS volatility decreases before 2004 and increases afterwards. This outcome is consistent with our previous findings.

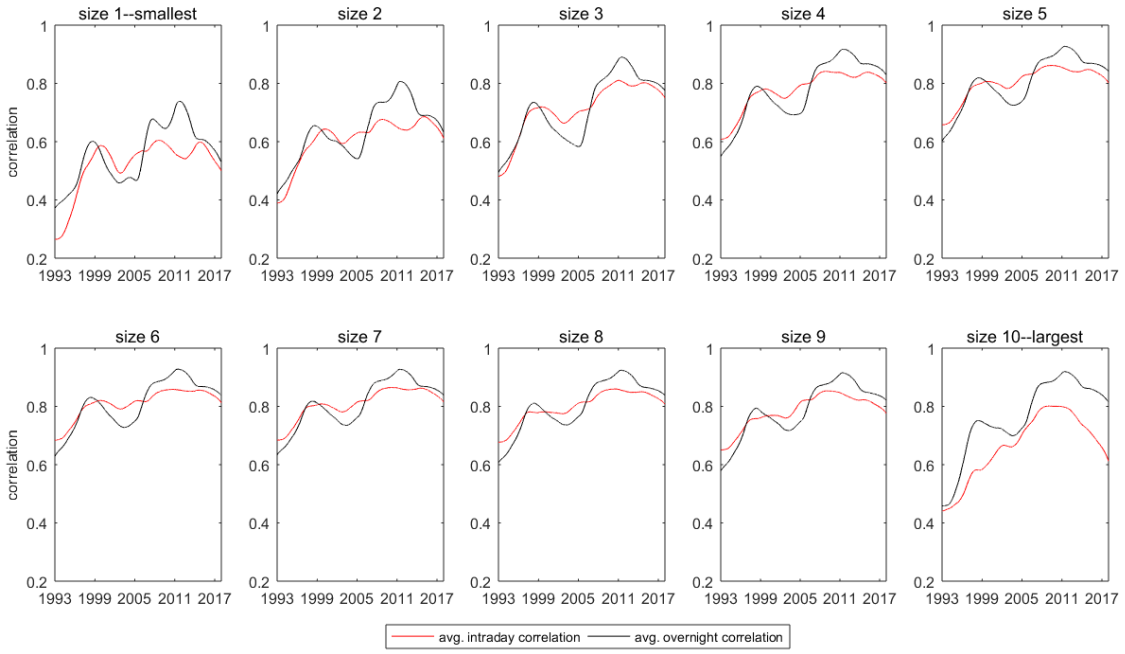
6.2.5 Overnight and intraday long-run correlations of size-based portfolios

We next investigate the correlations between size-based portfolios during overnight and intraday periods with the multivariate model. Each subplot in Figure 7 presents the average value of the long-run intraday (in red) and overnight (in black) correlations between that decile and the remaining deciles. All correlations increase gradually. Compared with large stocks, small stocks still co-move less with other stocks. Both overnight and intraday correlations increase during the period 2008-2016, especially the overnight correlations, which become larger than the intraday correlations. This outcome is in line with our findings from the Dow Jones stocks that overnight correlations tend to increase more dramatically during crisis periods.

6.3 Application to four large international stock markets

Given the remarkable size pattern in the U.S. market, we extend our empirical study to international markets. We choose the four largest international stock markets (U.S., U.K., Germany, and Japan), still with the research period from 3 4 January 1993 to 29 December 2017. The U.S. stock market and the Japanese stock market have almost opposite intraday and overnight periods, which enables us to better investigate the role of overseas information in the overnight and intraday volatility dynamics. We first construct five size-based portfolios for each market and then estimate the coupled-component GARCH model for each portfolio, similar to the procedure in the last subsection with the U.S. market.

We are also concerned that our results are affected by how portfolio construction is done. For instance, with the yearly portfolio re-construction, the number of stocks in the portfolios jumps at the beginning of each year because some stocks in the portfolios die throughout the year, whereas new stocks are added only at the beginning of the next year. The jumps in the stock numbers would lead to jumps in portfolio volatilities. We are not able to maintain the continuity in returns and in volatilities. In addition, the volatilities of portfolio returns are determined not only by the volatilities of component stocks but also by their correlations. The increase in portfolio volatility could be attributed to the increase in correlations. Hence, as an alternative approach, we estimate the GARCH model for each individual stock instead of estimating the GARCH model for each portfolio.



Each panel presents the average long-run intraday(overnight) correlations between that size decile and the remaining size deciles, as implied by the multivariate coupled-component model.

Figure 7: Long-run correlations between size-based portfolios

The individual stocks we investigate are the constituents of well-known market indices. We choose the S&P 500, S&P 400, and S&P 600 in the U.S. market; FTSE 100, FTSE 250, and FTSE small in the U.K. market; DAX 30, MDAX 50, and SDAX 50 in the German market; and Topix 100, Topix 400, and Topix small in the Japanese market. These indices cover the large-cap, middle-cap, and small-cap stocks in each market.

It is worth mentioning that using the opening and close prices of market indices directly is inappropriate. As we stated before, the published open prices of the market indices may not reflect all the accumulated overnight information because some component stocks may open a few minutes later.

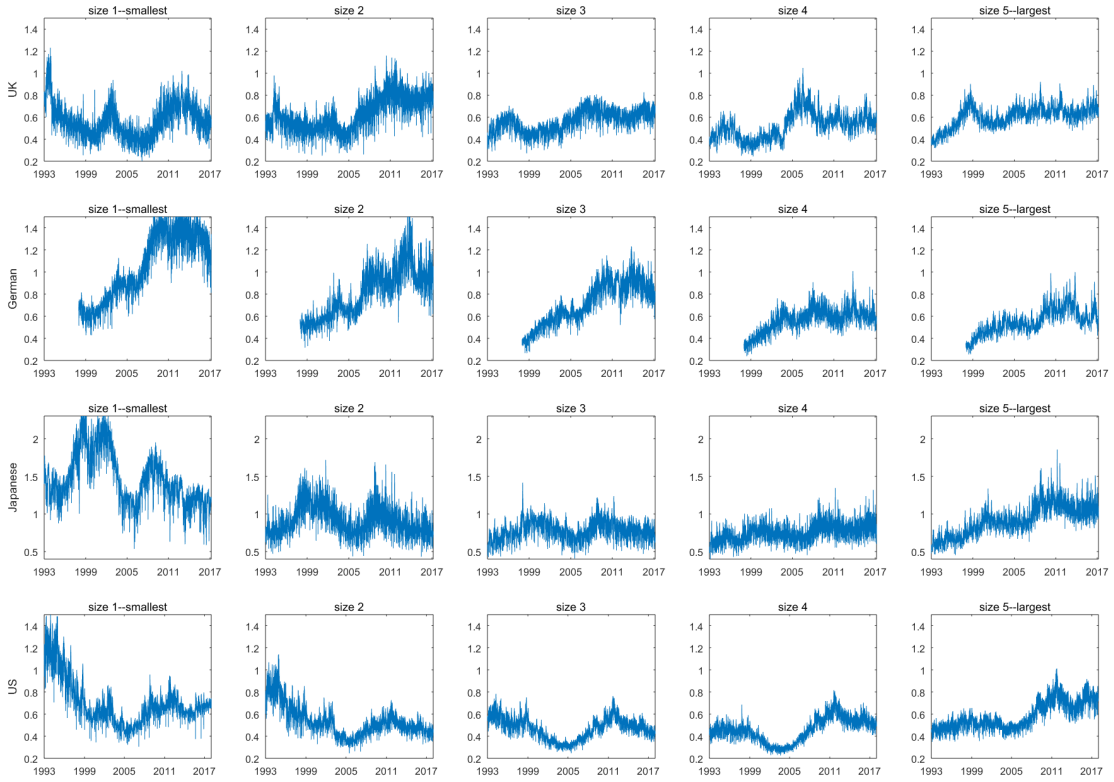
6.3.1 Results with size-based portfolios in international markets

In this part, we investigate the overnight and intraday volatility of size-based portfolios in the U.S., U.K., German, and Japanese stock markets. The data for the U.K., German, and Japanese markets are obtained from Datastream, while the data for the U.S. market are still those from CRSP, as in the last subsection. We exclude stocks with prices below one dollar in the U.S. market (below one pound in the U.K. market, below one euro in the German market, and below 100 yen in the Japanese market). [8] For the U.K. stocks, we only choose those traded in the London stock exchange. For many German firms, there are two prices in Datastream, one for XETRA and the other for Deutsche

[8]For stocks that stop trading on an exchange, Datastream keeps the value from the last trading day and displays it as the current one. This practice yields zero daily returns but repeated non-zero overnight and intraday returns. We discard those repeated values.

Boerse AG. These two prices often differed in the early 1990s. We choose the one from Xetra.

For each market, we sort stocks based on their market capitalizations and form five size-based portfolios. The size breakpoints are independent for each market, a different approach than in some research with unified global size breakpoints. Figure 8 plots the overnight to intraday ratios for the size-based portfolios in the four stock markets. The largest-cap portfolio has an upward overnight to intraday volatility ratio in all four of these markets, and the ratios are typically high after 2007, with the largest peaks during 2011 in U.S., Japan, and Germany.



This figure plots the overnight to intraday volatility ratio of the five size-based portfolios in the four international markets (U.K., Germany, Japan, and U.S.).

Figure 8: Overnight to intraday volatility ratio: size-based portfolios in four international markets

6.3.2 Results with the constituents of major international stock indices

For each of the four markets, we choose the well-known large-cap, middle-cap, and small-cap indices. We further divide the constituents of each index into two subgroups, large and small, according to their recent market capitalizations, with the exception of the German market indices (DAX 30, MDAX 50, and SDAX 50), which contain a relatively small number of stocks.[9] The group assignments

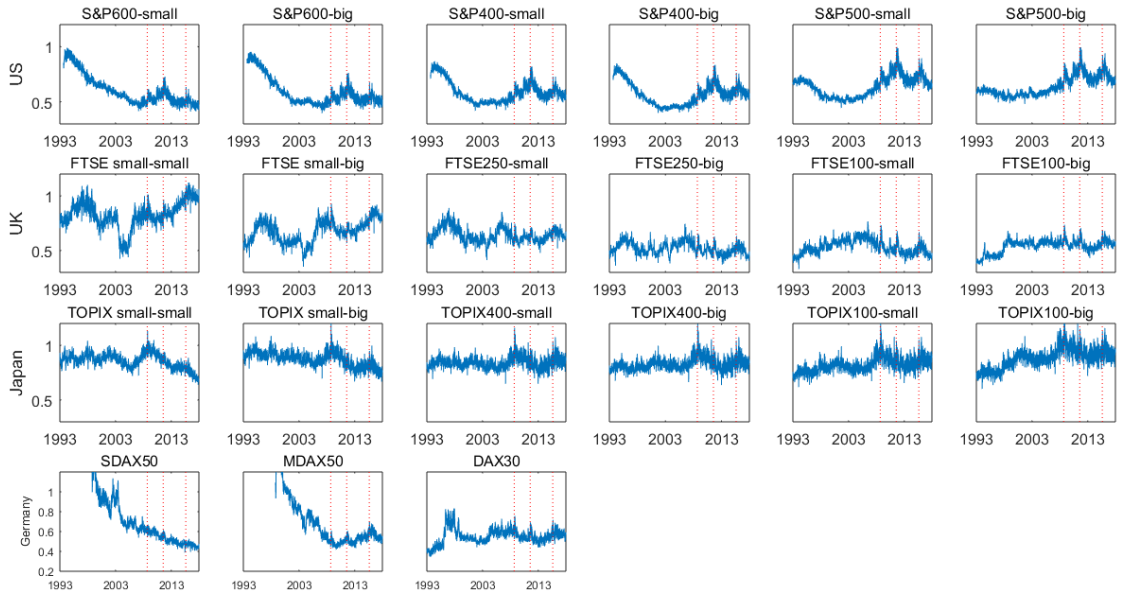
[9]Our constituents are based on the lists in May 2018, as well as the market values.

remain unchanged during the whole research period. The data for these constituents are obtained from Datastream. All opening and close prices have been adjusted for capital actions by Datastream.

In the U.S. stock market, we investigate the S&P 500, S&P 400, and S&P 600, which are indexes of the large-cap, middle-cap, and small-cap stocks, respectively. These three indices cover 1500 stocks. We exclude stocks with less than 1000 trading days and also exclude the stocks that ever have overnight to intraday volatility ratios larger than 5 or an estimated overnight degree of freedom, ν_N, less than 2. footnoteOur model requires the degree of freedom, ν_N and ν_D, larger than 2. It holds for most stocks, but it is sometimes violated, especially for small stocks with wild overnight returns. In total, there are 478, 355, and 525 stocks left in the S&P 500, S&P 400, and S&P 600, respectively. Figure 9 presents the average volatility ratios for each subgroup of these three indices. We plot the ratios of the S&P 400 and S&P 600 stocks from August 1993 because most of those stocks had zero overnight returns before then. A remarkable size pattern remains in the overnight to intraday volatility. Specifically, the overnight to intraday ratios decrease substantially in the 1990s for all subgroups in the U.S. market, especially the S&P 400 and S&P 600 groups, with the ratio decreasing from slightly less than 1 to around 0.6. The magnitude of the decrease is smaller for large-cap stocks, and we hardly observe the decrease in the S&P-big stocks. After 2015, all these ratios in the U.S. market start to increase, especially the large-cap and middle-cap stocks. The overall slope of the volatility ratios still changes monotonically from the small-cap subgroup (S&P 600-small) to the big-cap group (S&P 500-big). Remarkably, all the U.S. groups peak during stock market crashes, namely in the 2008 financial crisis, in the 2011 August stock markets fall, in the 2015 August black Monday, and in the 2016 January global fall.

In the U.K. stock market, we use the FTSE 100, FTSE 250, and FTSE SMALL, covering around 620 stocks in total. After excluding the stocks with less than 1000 trading days and stocks with overnight to intraday ratios larger than 5, there are 98, 175, and 176 stocks left in the FTSE 100, FTSE 250, and FTSE SMALL, respectively. Similarly, the overnight to intraday volatility ratio still shows an upward trend in the biggest-cap subgroup (FTSE 100-big) in the 1990s. The ratio is around 0.4 in 1993 but increases gradually to around 0.6 in 2000, and then starts to fluctuate around 0.6. However, the ratio of the small-cap stocks (FTSE SMALL) does not show a decreasing trend during the research period. The FEST SMALL has an evident low overnight to intraday volatility ratio during 2004 and 2005, because many zeros overnight returns occur in this period.

In the German stock market, we investigate the DAX 30, MDAX 50, and SDAX 50, covering around 130 stocks in total. After excluding the stocks with less than 1000 trading days and stocks with overnight to intraday ratios larger than 5, there are 27, 31, and 32 stocks left in the DAX 30, MDAX 50, and SDAX 50, respectively. Since each index has a relatively small number of stocks, we do not further divide them into subgroups. The ratios before July 1997 are not presented in the figure for MDAX and SDAX because most of their current constituents were not traded during that period. Again, the average overnight to intraday volatility ratio of the DAX 30 increases from 0.4 in 1993 to 0.6 in 2000, very similar to that of the FTSE 100-big stocks. In contrast, the ratio of the SDAX shows a dramatic downward trend during the 20-year period considered here, with a value around 1.3 in 1998 to around 0.4 in 2017. The MDAX also has a dramatic decrease in the early



This figure plots the overnight to intraday volatility ratio for the major international indices. The components of each index are further divided into two subgroups according to their current market value, except for the German indices. The three dashed vertical lines from left to right indicate the dates 16 September 2008 (financial crisis), 1 August 2011 (August 2011 stock markets fall), and 24 August 2015 (the 2015-2016 market fall), respectively.

Figure 9: Overnight to intraday volatility ratio: major market indices

years, from 1.2 in 1998 to around 0.5 in 2002, but it fluctuates around 0.5 afterward.

In the Japanese stock markets, we use the Topix 100, Topix 400, and Topix SMALL, covering around 2000 stocks. After excluding stocks with less than 1000 trading days or with overnight to intraday ratios larger than 5, there are 98, 373, and 1112 stocks left in the DAX 30, MDAX 50, and SDAX 50, respectively. The overnight to intraday volatility ratio shows an upward trend in the biggest-cap group (TOPIX 100-big) and a downward trend in the TOPIX SAMLL stocks. All ratios reach their peaks in the 2008 financial crisis, and have their second largest peaks at the beginning of 2016.

We are also concerned that the dynamic of the overnight to intraday volatility ratios is subject to new stocks bias. Specifically, some current constituents were not traded at the beginning of our research period, and the averaged ratio of the overnight to intraday volatility may increase if newly added stocks have higher ratios. Hence, we exclude the stocks that were not traded in 1993, but the results remain nearly unchanged.

In general, the results based on individual stocks are in some extent similar to those from the size-based portfolios, but the increase of the overnight to intraday volatility ratio for large stocks is less evident. Specifically, the ratio has still increased substantially during the recent 25 years for large stocks but has decreased for small stocks in the U.S. market. The Japanese market shows similar patterns, although they are less evident. The ratio for the large stocks in the U.K. and

German markets also increased in the 1990s, but not afterward. This result is possibly because the U.S. market opens during the trading time of the U.K. and the Germany markets, so many overnight announcements in the U.S. market are reflected in the intraday prices of the U.K. and German markets. The less evident increase in individual stocks suggests that the increase of stock correlations is also an important factor driving up the overnight to intraday ratio in the large-cap portfolios, especially in U.K. and Germany.

We also find that the volatility ratio of most stocks reaches their peak during a stock market crisis, for instance, in the 2008 financial crisis, the 2011 markets fall, and the 2015-2016 market fall. Together with the previous finding that overnight correlations also increase more substantially than intraday correlations during a crisis. These findings may explain the high overnight returns found in recent empirical finance study.

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 fluctuation. The large sample properties of the estimators are provided. For the univariate model we show that one can adaptively estimate the parameters of the dynamic process in the presence of the unknown slowly varying trend.

The empirical results show that the ratio of overnight to intraday volatility for especially large stocks has increased during the last 25 years when accounting for both slowly changing and rapidly changing components. This is contrary to what is often argued with regard to the change in market structure and the effects of high frequency trading. Portfolios of small stocks on the other hand seem to exhibit a different trend. The main explanation for this phenomenon is perhaps the variation in international linkage. Stocks with higher international correlations show considerably higher overnight to intraday volatility ratio, and likewise with larger market capitalization. The changes of minimal tick size make a contribution to the downward trend of small stocks during 1990s.

We found various other results. First, we found in the multivariate model that (slowly moving) correlations between assets have increased during our sample period. In addition, overnight correlations increase more substantially than intraday correlations during recent crises. We also found that the information in overnight returns is valuable for updating the forecast of the close to close volatility.

In our modelling we have not separated midweek overnight components from weekend components. We may extend the model to allow multiple different components reflecting weekend different from intraweek overnight, but at the cost of estimating many more parameters. We are also considering how to extend the model to allow stocks traded in different time zones, Lin, Engle, and Ito (1994).

References

Aboody, D., O. Even-Tov, R. Lehavy, and B. Trueman (2018), "Overnight returns and firm-specific investor sentiment." *Journal of Financial and Quantitative Analysis*, 53, 485–505. Cited By 0.

Ahoniemi, Katja and Markku Lanne (2013), "Overnight stock returns and realized volatility." *International Journal of Forecasting*, 29, 592–604.

Andersen, Torben G, Tim Bollerslev, and Xin Huang (2011), "A reduced form framework for modeling volatility of speculative prices based on realized variation measures." *Journal of Econometrics*, 160, 176–189.

Aretz, Kevin and Söhnke M Bartram (2015), "Making money while you sleep? anomalies in international day and night returns." Working paper, available at: https://ssrn.com/abstract=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.

Boehmer, Ekkehart, Kingsley YL Fong, and Juan Julie Wu (2015), "International evidence on algorithmic trading."

Brogaard, Jonathan (2011), "High frequency trading and volatility." Working paper.

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.

Christie, William G., Jeffrey H. Harris, and Paul H. Schultz (1994), "Why did nasdaq market makers stop avoiding oddeighth quotes?" *Journal of Finance*, 49, 1841–1860.

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://ssrn.com/abstract=1004081.

Creal, Drew, Siem Jan Koopman, and André Lucas (2011), "A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations." *Journal of Business & Economic Statistics*, 29, 552–563.

Creal, Drew, Siem Jan Koopman, and André Lucas (2013), "Generalized autoregressive score models with applications." *Journal of Applied Econometrics*, 28, 777–795.

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

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.

Eun, Cheol S., Wei Huang, and Sandy Lai (2008), "International diversification with large- and small-cap stocks." *Journal of Financial & Quantitative Analysis*, 43, 489–523.

Fan, Yanqin and Qi Li (1996), "Consistent model specification tests: omitted variables and semiparametric functional forms." *Econometrica*, 64, 865–890.

Fleming, Jeff, Chris Kirby, and Barbara Ostdiek (2003), "The economic value of volatility timing using realized volatility." *Journal of Financial Economics*, 67, 473–509.

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.

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.

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.

Härdle, Wolfgang and Oliver Linton (1994), "Applied nonparametric methods." *Handbook of Econometrics*, 4, 2295–2339.

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.

Harvey, Andrew C, Tirthankar Chakravarty, et al. (2008), *Beta-t-(e) garch*. University of Cambridge, Faculty of Economics.

Hasbrouck, Joel (2007), *Empirical market microstructure: The institutions, economics, and econometrics of securities trading*. Oxford University Press.

He, Changli, Annastiina Silvennoinen, and Timo Teräsvirta (2008), "Parameterizing unconditional skewness in models for financial time series." *Journal of Financial Econometrics*, 6, 208–230.

Imbens, G.W. and J.M. Wooldridge (2007), "Difference-in-differences estimation." Nber lecture notes in econometrics, lecture notes 10.

Jones, M.C., O. Linton, and J.P. Nielsen (1995), "A simple bias reduction method for density estimation." *Biometrika*, 82, 327–338. Cited By 93.

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.

Li, Jia, Todorov Viktor, and Tauchen George (2017), "Jump factor models in large cross-sections." Technical report.

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* (Bol Georg, Nakhaeizadeh Gholamreza, T. Rachev Svetlozarm, Ridder Thomas, and Vollmer Karl-Heinz, eds.), 149–156, Springer.

Linton, Oliver (1993), "Adaptive estimation in arch models." *Econometric Theory*, 9, 539–569.

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.

Martens, Martin (2002), "Measuring and forecasting S&P 500 index-futures volatility using high-frequency data." *Journal of Futures Markets*, 22, 497–518.

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.

Polk, Christopher, Dong Lou, and Spyros Skouras (2018), "A tug of war: overnight versus intraday expected returns." *Journal of Financial Economics*, In Press.

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*, 1, 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*, 20, 1768–1802.

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.

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

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

Supplementary Material to "A Coupled Component GARCH Model for Intraday and Overnight Volatility"

Oliver Linton
University of Cambridge

Jianbin Wu
Nanjing University

September 14, 2018

1 Proof of the large sample properties

1.1 Lemmas

Lemma 1 *Suppose that Assumptions A1-A8 hold. Then,*

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

$$\int_0^1\left(\widetilde{\sigma}^{j}(u)-\sigma_0^{j}(u)\right)^2 du = O_p\left(h^2+\sqrt{\frac{1}{Th}}\right).$$

Furthermore $\left\|\tilde{\theta}-\theta\right\|^2 = O_p\left(h^2+\sqrt{\frac{1}{Th}}\right)$.

Proof of Lemma 1 Denote $H^j(s)=\exp(\sigma^j(s))$. We drop the superscript j in what follows 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 the 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 (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

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

where $\bar{H}(s)$ is 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 $\widetilde{\sigma}(s)^*$ has vanished due to 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

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

Therefore, the uniform convergence rate is

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

The L_2 rate follows by similar arguments.

Recall that $\sigma(s) = \sum_{j=1}^{\infty}\theta_j\psi_j(s)$ for the orthogonal basis ψ_j. By construction $\tilde{\sigma}(s)$ is a member of the same normed space as $\sigma(s)$, in which case we can write $\tilde{\sigma}(s) = \sum_{j=1}^{\infty}\tilde{\theta}_j\psi_j(s)$ for coefficients $\tilde{\theta}_j$, $j = 1, 2, \ldots$. that satisfy $\sum_{j=1}^{\infty}|\tilde{\theta}_j| < \infty$. In particular, let

$$Q(\theta) = \int_0^1\left(\widetilde{\sigma}(s) - \int_0^1\widetilde{\sigma}(u)du - \sum_{k=1}^{\infty}\theta_k\psi_k(s)\right)^2 ds.$$

We have for $k = 1, 2, \ldots$

$$\frac{\partial Q}{\partial\theta_k}(\theta) = \int_0^1\left(\widetilde{\sigma}(s) - \int_0^1\widetilde{\sigma}(u)du - \sum_{k=1}^{\infty}\theta_k\psi_k(s)\right)\psi_k(s)ds$$

and so

$$\widetilde{\theta}_k = \int_0^1\left(\widetilde{\sigma}(s) - \int_0^1\widetilde{\sigma}(u)du\right)\psi_k(s)ds = \int_0^1\widetilde{\sigma}(s)\psi_k(s)ds,$$

since $\int_0^1\psi_k(s)ds = 0$. We have $Q(\widetilde{\theta}) = 0$. The coefficients satisfy $\widetilde{\theta}_k - \theta_k = \int_0^1\left(\widetilde{\sigma}(s) - \sigma(s)\right)\psi_k(s)ds$.

We have

$$\begin{aligned}\int\left(\tilde{\sigma}(s) - \sigma(s)\right)^2 ds &= \int\left(\sum_{j=1}^{\infty}\left(\tilde{\theta}_j - \theta_j\right)\psi_j(s)\right)^2 ds \\ &= \sum_{j=1}^{\infty}\left(\tilde{\theta}_j - \theta_j\right)^2\int\psi_j^2(s)ds \\ &= \sum_{j=1}^{\infty}\left(\tilde{\theta}_j - \theta_j\right)^2 = \left\|\tilde{\theta} - \theta\right\|^2\end{aligned}$$

under the assumption that ψ_j are orthonormal. So given the L_2 rate of convergence of $\tilde{\sigma}$ we have the same convergence rate for the implied coefficients. ■

Lemma 2 *If $|\beta_j| < 1$, $j = D, N$, then e_t^j and λ_t^j are strictly stationary and β-mixing with exponential decay.*

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

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

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 $E|X_t|^s < \infty$. 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 stationary 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 $X_t = \begin{pmatrix} \lambda_t^D & \lambda_t^N & m_t^D & m_t^N & \text{sign}\left(e_t^D\right) & \text{sign}\left(e_t^N\right) \end{pmatrix}^\intercal$. ■

Lemma 3 *Suppose that Assumptions A1-A4 hold. Then,*

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

Proof. By (3) and (4), 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

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

Since $\left(B_{t-1}\otimes B_{t-1}\right)\left(A_{t-1}\otimes A_{t-1}\right)$ and B_tA_t are i.i.d, and $EB_tA_t = EB_tEA_t$, we obtain

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

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

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

with the first term

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

and the second term

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

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

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

Therefore,

$$\left\|E\mathrm{vec}\left(\left(\prod_{i=j}^{k-1} B_{t-i}A_{t-i}\right)\Lambda_{t-k}\begin{pmatrix}\lambda_{t-j}^D-\omega_D & 0\end{pmatrix}\right)\right\|_\infty \le c_T\left\|EB_tEA_t\right\|^{k-j-1} \tag{2}$$

with

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

Substituting (2) into (1) gives

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

provided that $\|EB_tEA_t\|_\infty < 1$ and $\|E\left(B_{t-1}A_{t-1} \otimes B_{t-1}A_{t-1}\right)\|_\infty < \|EB_tEA_t\|_\infty$ .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*

$$\begin{aligned} \begin{pmatrix} \frac{\partial}{\partial \beta_D} h_t^D \\ \frac{\partial}{\partial \beta_D} h_t^N \end{pmatrix} &= A_t \begin{pmatrix} \lambda_{t-1}^D - \omega_D \\ 0 \end{pmatrix} + A_t B_{t-1} \begin{pmatrix} \frac{\partial}{\partial \beta_D} h_{t-1}^D \\ \frac{\partial}{\partial \beta_D} h_{t-1}^N \end{pmatrix} \\ &= \sum_{j=1}^{\infty} A_t \prod_{i=1}^{j-1} B_{t-i}A_{t-i} \begin{pmatrix} \lambda_{t-j}^D - \omega_D \\ 0 \end{pmatrix}. \end{aligned} \tag{3}$$

$$\begin{aligned} \begin{pmatrix} \frac{\partial h_t^D}{\partial \sigma^D(t-k/T)} \\ \frac{\partial h_t^N}{\partial \sigma^D(t-k/T)} \end{pmatrix} &= A_t B_{t-1} \begin{pmatrix} \frac{\partial h_{t-1}^D}{\partial \sigma^D(t-k/T)} \\ \frac{\partial h_{t-1}^N}{\partial \sigma^D(t-k/T)} \end{pmatrix} \\ &= A_t \left(\prod_{i=1}^{k-1} B_{t-i}A_{t-i} \right) \Lambda_{t-k}, k > 1 \\ \begin{pmatrix} \frac{\partial h_t^D}{\partial \sigma^D(t/T)} \\ \frac{\partial h_t^N}{\partial \sigma^D(t/T)} \end{pmatrix} &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \text{ and } \begin{pmatrix} \frac{\partial h_t^D}{\partial \sigma^D(t-1/T)} \\ \frac{\partial h_t^N}{\partial \sigma^D(t-1/T)} \end{pmatrix} = A_t \begin{pmatrix} a_{t-1}^{DD} \\ a_{t-1}^{ND} \end{pmatrix}, \end{aligned} \tag{4}$$

with $\Lambda_t = \begin{pmatrix} a_t^{DD} \\ a_t^{ND} \end{pmatrix}$ .If the top-Lyapunov exponent of the sequence of A_tB_{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

$$\begin{aligned} h_t^D &= \sigma^D(t/T) - \beta_D \sigma^D(\frac{t-1}{T}) + \omega_D(1-\beta_D) + \beta_D h_{t-1}^D + \gamma_D m_{t-1}^D \\ &\quad + \rho_D m_t^N + \gamma_D^*(m_{t-1}^D + 1)\text{sign}(u_{t-1}^D) + \rho_D^*(m_t^N + 1)\text{sign}(u_t^N) \end{aligned} \tag{5}$$

$$\begin{aligned} h_t^N &= \sigma^N(t/T) - \beta_N \sigma^N(\frac{t-1}{T}) + \omega_N(1-\beta_N) + \beta_N h_{t-1}^N + \gamma_N m_{t-1}^N \\ &\quad + \rho_N m_{t-1}^D + \rho_N^*(m_{t-1}^D + 1)\text{sign}(u_{t-1}^D) + \gamma_N^*(m_{t-1}^N + 1)\text{sign}(u_{t-1}^N). \end{aligned} \tag{6}$$

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 (5) and (6) with respect to β_D gives

$$\begin{aligned}\frac{\partial h_t^D}{\partial \beta_D} &= -\sigma^D(\frac{t-1}{T}) - \omega_D + h_{t-1}^D + \beta_D \frac{\partial}{\partial \beta_D} h_{t-1}^D + \frac{\partial}{\partial \beta_D}\gamma_D m_{t-1}^D + \frac{\partial}{\partial \beta_D}\rho_D m_t^N \\ &\quad + \frac{\partial}{\partial \beta_D}\gamma_D^*(m_{t-1}^D+1)\text{sign}(u_{t-1}^D) + \frac{\partial}{\partial \beta_D}\rho_D^*(m_t^N+1)\text{sign}(u_t^N) \qquad (7) \\ \frac{\partial h_t^N}{\partial \beta_D} &= \beta_N \frac{\partial}{\partial \beta_D} h_{t-1}^N + \frac{\partial}{\partial \beta_D}\gamma_N m_{t-1}^N + \frac{\partial}{\partial \beta_D}\rho_N m_{t-1}^D \\ &\quad + \frac{\partial}{\partial \beta_D}\rho_N^*(m_{t-1}^D+1)\text{sign}(u_{t-1}^D) + \frac{\partial}{\partial \beta_D}\gamma_N^*(m_{t-1}^N+1)\text{sign}(u_{t-1}^N) \qquad (8)\end{aligned}$$

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

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

Substituting them back into (7) and (8) gives

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

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, Moulines, and Stoffer (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

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

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

and (4) 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 *Suppose that Assumptions A1-A4 hold. Then, we have*

$$\frac{1}{T} \sum_{t=1}^{T} E\left(\frac{\partial h_t^D}{\partial \theta} \right) = 0.$$

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

1.2 Proof of Main Results

1.2.1 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 the following explanation). 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 $E[(m_t^j)^2] < \infty$. Therefore, we can write

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

To prove the Theorem, it then suffices to show that

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

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

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

when the limit exists. Thus we obtain,

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

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 $\sum_{t=1}^{T} \psi_i^D\left(\frac{t}{T}\right)$ gives

$$\begin{aligned}
\frac{1}{T}\sum_t \psi_i^D\left(\frac{t-k}{T}\right) &= \frac{1}{T}\sum_t \psi_i^D\left(\frac{t}{T}\right) - \frac{1}{T}\frac{k}{T}\sum_t \psi_i^{D'}\left(\frac{t}{T}\right) + O\left(\frac{k}{T}\right)^2 \\
&= O\left(\frac{1}{T}\right) + O\left(\frac{k}{T}\right) + O\left(\frac{k}{T}\right)^2 \\
&= O\left(\frac{k}{T}\right).
\end{aligned}$$

Hence, it suffices to show

$$\sum_{k=0}^{T} \left\| kE\left(\begin{pmatrix} \frac{\partial h_t^D}{\partial \sigma^D(\frac{t-k}{T})} \\ \frac{\partial h_t^N}{\partial \sigma^D(\frac{t-k}{T})} \end{pmatrix} \begin{pmatrix} \frac{\partial h_t^D}{\partial \beta_D} & \frac{\partial h_t^N}{\partial \beta_D} \end{pmatrix} \right) \right\|_\infty < \infty,$$

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

$$\begin{aligned}\frac{\partial l_t^D}{\partial v_D} &= -\frac{1}{2}\ln\left(1+\frac{(u_t^D)^2}{v_D\exp(2h_t^D)}\right)+\frac{\partial}{\partial v_D}\left(\ln\Gamma\left(\frac{v_D+1}{2}\right)-\ln\Gamma\left(\frac{v_D}{2}\right)\right)-\frac{1}{2v_D}\\ &\quad+\frac{v_D+1}{2\left(1+\frac{(u_t^D)^2}{v_D\exp(2h_t^D)}\right)}\frac{(u_t^D)^2}{v_D^2\exp(2h_t^D)}\left(1+2v_D\frac{\partial h_t^D}{\partial v_D}\right)+\frac{\partial h_t^D}{\partial v_D} \qquad (9)\\ \frac{\partial l_t^N}{\partial v_D} &= \frac{v_N+1}{2\left(1+\frac{(u_t^N)^2}{v_N\exp(2h_t^N)}\right)}\frac{(u_t^N)^2}{v_N^2\exp(2h_t^N)}\left(1+2v_N\frac{\partial h_t^N}{\partial v_D}\right)+\frac{\partial h_t^N}{\partial v_D}.\end{aligned}$$

Then we have

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

The first term vanishes by Lemma 5. Then we can use the same procedure above to obtain $\sum_{t=1}^T E\left(\frac{\partial h_t^D}{\partial v_D}\frac{\partial h_t^D}{\partial\theta}\right)=O(1)$, and to finish the proof for v_D. ■

1.2.2 Proof of Theorem 2

By the triangle inequality,

$$\sup_{\phi\in\Phi}\left|l_T(\phi;\widetilde{\theta})-l(\phi)\right|\le\sup_{\phi\in\Phi}\left|l_T(\phi;\widetilde{\theta})-l_T(\phi;\theta_0)\right|+\sup_{\phi\in\Phi}\left|l_T(\phi;\theta_0)-l(\phi)\right|,$$

where $l(\phi)=E(l_T(\phi;\theta_0))$. By the identification condition $l(\phi)$ is uniquely maximized at $\phi=\phi_0$ and standard arguments (Harvey (2013)) show that the second term is $o_p(1)$. The first term is also $o_p(1)$ by the uniform consistency of $\tilde{\sigma}(s)$ in Lemma 1 and the smoothness of the objective function in $\tilde{\sigma}(t/T)$ and equivalently θ_k.

We next turn to asymptotic normality. The general strategy is that we first show the estimators obtained by maximizing $l_T(\phi;\widetilde{\theta})$ and $l_T(\phi;\theta_0)$ have the same asymptotic distribution, provided $||\tilde{\theta}-$

$\theta_0||$ converges to 0. As a result, the asymptotic property of $\tilde{\phi}$ follows as in the parametric model with known $\sigma(t/T)$.

Following Severini and Wong (1992), the expansion of $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{\partial l_t(\phi_0,\tilde{\theta})}{\partial\phi}$ at θ_0 gives

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{\partial l_t(\phi_0;\tilde{\theta})}{\partial\phi}=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{\partial l_t(\phi_0;\theta_0)}{\partial\phi}+\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\sum_{k}\frac{\partial^2 l_t(\phi_0;\theta_0)}{\partial\phi\partial\theta_k}(\tilde{\theta}_k-\theta_{k,0})\right)+o_p(1). \tag{10}$$

According to Theorem 1, we have $\sum_{t=1}^{T}E(\frac{\partial^2 l_t(\phi;\theta)}{\partial\phi_i\partial\theta_k})=O(1)$, for each k and i, where $k\in\{1,\ldots,\infty\}$ and $i\in\{1,\ldots,14\}$. It follows that

$$\sum_{t=1}^{T}\frac{\partial^2 l_t(\phi_0;\theta_0)}{\partial\phi_i\partial\theta_k}=O_p(\sqrt{T}).$$

Given that the dimension of the sieve space grows slowly and $\tilde{\theta}$ converges to θ, the second term in (10) is of order $o_p(1)$.

Therefore, the asymptotic property of $\tilde{\phi}$ can be obtained with a similar procedure to Harvey (2013). He gives the consistency and asymptotic normality of the estimator for the parametric beta-t-egarch model. The basic idea is that 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.

Obviously, $\hat{\phi}$ has the same limiting distribution as $\tilde{\phi}$, since $\sum_{t=1}^{T}\frac{\partial l_t(\phi_0,\tilde{\theta})}{\partial\phi}$ and $\sum_{t=1}^{T}\frac{\partial l_t(\phi_0,\theta)}{\partial\phi}$ have the same asymptotic property. ■

1.2.3 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:

$$\begin{aligned}\frac{\partial L_T^j(\sigma^j;s)}{\partial\sigma^j}&=\frac{1}{T}\sum_{t=1}^{T}K_h(s-t/T)\left[-(v_j+1)b_t^j(\sigma^j)+1\right]\\ \frac{\partial^2 L_T^j(\sigma^j;s)}{\partial\sigma^{j2}}&=2(v_j+1)\frac{1}{T}\sum_{t=1}^{T}K_h(s-t/T)\left[b_t^j(\sigma^j)\left(1-b_t^j(\sigma^j)\right)\right],\end{aligned} \tag{11}$$

where

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

We have

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

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

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

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 $\eta_t^j=\exp(\sigma^j(t/T))\varepsilon_t^j$. It follows that

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

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. ■

2 Derivatives in the multivariate model

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

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

Then

$$
\begin{aligned}
\mathrm{d}l_t &= \mathrm{d}\log|\Theta| - \sum_{i=1}^{n}\frac{(v_i+1)\left(\iota_i^{\intercal}\mathrm{diag}\left(\exp(-\lambda_t)\right)\Theta u_t\right)\exp(-\lambda_{it})}{v_i+\left(\iota_i^{\intercal}\mathrm{diag}\left(\exp(-\lambda_t)\right)\Theta u_t\right)^2}\mathrm{tr}u_t\iota_i^{\intercal}\mathrm{d}\Theta \\
&= \mathrm{tr}\left(\Theta^{-1}\mathrm{d}\Theta\right) - \mathrm{tr}\left(\sum_{i=1}^{n}\frac{(v_i+1)\exp(-2\lambda_{it})\iota_i^{\intercal}\Theta u_t}{v_i+\exp(-2\lambda_{it})\left(\iota_i^{\intercal}\Theta u_t\right)^2}u_t\iota_i^{\intercal}\mathrm{d}\Theta\right) \\
&= \mathrm{tr}\left[\left(\Theta^{-1} - \sum_{i=1}^{n}\frac{(v_i+1)\exp(-2\lambda_{it})\iota_i^{\intercal}\Theta u_t}{v_i+\exp(-2\lambda_{it})\left(\iota_i^{\intercal}\Theta u_t\right)^2}u_t\iota_i^{\intercal}\right)\mathrm{d}\Theta\right] \\
&= \left[\mathrm{vec}\left(\Theta^{-1} - \sum_{i=1}^{n}\iota_i u_t^{\intercal}\frac{(v_i+1)\exp(-2\lambda_{it})\iota_i^{\intercal}\Theta u_t}{v_i+\exp(-2\lambda_{it})\left(\iota_i^{\intercal}\Theta u_t\right)^2}\right)\right]^{\intercal}\mathrm{dvec}\Theta \\
&= \left[\mathrm{vec}\left(\Theta^{-1} - \sum_{i=1}^{n}\iota_i u_t^{\intercal}\frac{(v_i+1)\exp(-2\lambda_{it})\iota_i^{\intercal}\Theta u_t}{v_i+\exp(-2\lambda_{it})\left(\iota_i^{\intercal}\Theta u_t\right)^2}\right)\right]^{\intercal}D_n\mathrm{dvech}\Theta, \qquad (12)
\end{aligned}
$$

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

$$
\begin{aligned}
\frac{\partial L_T(\Theta;\lambda_t,s)}{\partial\mathrm{vech}\Theta} = &-\frac{1}{T}D_n^{\intercal}\mathrm{vec}\sum_{i=1}^{n}\left(\iota_i\sum_{t=1}^{T}\left(K_h(s-t/T)u_t^{\intercal}\frac{(v_i+1)\exp(-2\lambda_{it})\iota_i^{\intercal}\Theta u_t}{v_i+\exp(-2\lambda_{it})\left(\iota_i^{\intercal}\Theta u_t\right)^2}\right)\right) \\
&+ D_n^{\intercal}\mathrm{vec}\left(\Theta^{-1}\right). \qquad (13)
\end{aligned}
$$

To compute the Hessian matrix, we evaluate the differential of the Jacobian matrix in (12)

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

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

$$
\begin{aligned}
&\frac{\partial^2 L_T(\Theta;\lambda_t,s)}{\partial\mathrm{vech}\Theta\partial\left(\mathrm{vech}\Theta\right)^{\intercal}} \\
&= -D_n^{\intercal}\left(\sum_{i=1}^{n}\left(\sum_{t=1}^{T}\frac{K_h(s-t/T)\left(v_i-\exp(-2\lambda_{it})\left(\iota_i^{\intercal}\Theta u_t\right)^2\right)(v_i+1)}{T\left(v_i+\exp(-2\lambda_{it})\left(\iota_i^{\intercal}\Theta u_t\right)^2\right)^2\exp(2\lambda_{it})}u_t u_t^{\intercal}\right)\otimes\left(\iota_i\iota_i^{\intercal}\right)\right)D_n \\
&- D_n^{\intercal}\left(\Theta^{-1}\otimes\Theta^{-1}\right)D_n \qquad (14)
\end{aligned}
$$

■

3 Variances of r_t^N and r_t^D conditional on past information

In this section, we first give the conditional and unconditional moments of λ_t. Based on that, we can obtain the variance of r_t^N and r_t^D conditional on past information.

3.1 Conditional and unconditional moments of λ_t

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

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

with $\Lambda_t = \exp(2\omega_D(1-\beta_D) + 2\beta_D\lambda_{t-1}^D + 2\gamma_D m_{t-1}^D + 2\gamma_D^*(m_{t-1}^D+1)\text{sign}(e_{t-1}^D)$. 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,

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

with ${}_1F_1$ the Kummer's function, we have

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

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

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

When λ_t^j starts from infinite past,

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

Therefore, we obtain

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

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

3.2 Variances of r_t^N and r_t^D conditional on past information

The variances of r_t^N and r_t^D conditional on $\mathcal{F}_{t-1}$ are

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

In last subsection we show that

$$
\begin{aligned}
E\left[\exp(2\lambda_t^D)|\mathcal{F}_{t-1}\right]&=\frac{1}{2}\exp(-2\rho_D)\Lambda_t\ \times_1 F_1(1/2,1/2+v_N/2,(2\rho_D+2\rho_D^*)(v_N+1))\\
&\quad+\frac{1}{2}\exp(-2\rho_D)\Lambda_t\ \times_1 F_1(1/2,1/2+v_N/2,(2\rho_D-2\rho_D^*)(v_N+1)),
\end{aligned}
\tag{16}
$$

and

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

From this we obtain an explicit formula for $\operatorname{var}(r_t^D|\mathcal{F}_{t-1})$ and $\operatorname{var}\left(r_t^D+r_t^N|\mathcal{F}_{t-1}\right)$, which can be used for forecasting.

4 Tables and figures

	intraday				overnight			
	mean	std.dev.	skew	kurt	mean	std.dev.	skew	kurt
AAPL	-0.0003	0.0234	0.4624	9.9250	0.0009	0.0173	-8.2036	318.6199
MSFT	0.0003	0.0162	0.1571	6.2404	0.0002	0.0106	-0.6479	33.7075
XOM	0.0004	0.0126	0.1274	12.0193	-0.0001	0.0069	-0.9669	15.2169
JNJ	0.0003	0.0114	0.1154	7.1950	0.0001	0.0069	-3.1942	87.6901
INTC	0.0000	0.0194	0.1992	8.0585	0.0004	0.0137	-2.7660	54.5180
WMT	-0.0001	0.0142	0.1265	8.4767	0.0003	0.0082	-0.5287	16.8276
CVX	0.0001	0.0133	0.0781	11.2908	0.0002	0.0072	-0.9472	13.6858
UNH	0.0003	0.0192	-0.1528	15.9404	0.0004	0.0112	-3.0667	69.9750
CSCO	-0.0001	0.0223	0.8577	25.9271	0.0007	0.0147	-3.8152	108.3076
HD	0.0001	0.0165	0.3053	7.5083	0.0004	0.0106	-3.9397	117.0914
PFE	-0.0001	0.0146	0.0103	6.5375	0.0003	0.0095	-1.9437	44.6513
BA	0.0003	0.0153	-0.0112	6.7304	0.0001	0.0101	-2.5133	62.5035
VZ	0.0001	0.0138	0.4653	8.1210	0.0000	0.0077	-0.5386	16.4611
PG	0.0008	0.0119	-0.0339	9.9414	-0.0005	0.0083	-19.0614	935.7817
KO	0.0005	0.0119	0.0663	9.3238	-0.0003	0.0070	-0.6990	17.2768
MRK	0.0003	0.0143	-0.0408	8.3037	-0.0002	0.0095	-6.0575	173.5046
DIS	0.0003	0.0154	0.1754	7.7958	0.0000	0.0106	-1.2228	51.4031
IBM	0.0004	0.0144	0.0549	7.9465	-0.0000	0.0096	-0.6096	43.7477
GE	-0.0002	0.0153	-0.0095	11.9089	0.0003	0.0099	0.1355	31.4138
MCD	0.0005	0.0132	0.1176	10.6598	-0.0001	0.0081	-0.9658	23.4707
MMM	0.0003	0.0123	-0.0108	7.6709	0.0001	0.0071	-0.4690	19.6496
NKE	0.0006	0.0170	0.1205	10.6060	-0.0001	0.0102	-2.0427	51.4836
UTX	0.0001	0.0140	-0.3085	10.2302	0.0004	0.0081	-1.7679	42.1281
CAT	-0.0001	0.0168	0.0265	6.1293	0.0006	0.0108	-0.8354	19.9062
AXP	0.0003	0.0185	-0.0528	10.6319	0.0002	0.0104	-1.1468	31.7965
TRV	0.0002	0.0158	-0.1092	17.2694	0.0001	0.0086	-1.7539	66.5588

This table gives the summary statistics of the intraday and overnight returns for Dow Jones stocks.

Table A.1: Summary statistics of intraday and overnight returns for Dow Jones stocks

	δ	μ_D	μ_N	Π_{11}	Π_{12}	Π_{21}	Π_{22}
AAPL	0.0485	-0.0354	0.0935	-0.0609	0.0775	-0.0053	0.0017
	(0.0444)	(0.0292)	(0.0216)	(0.0172)	(0.0258)	(0.0200)	(0.0172)
MSFT	-0.0188	0.0290	0.0250	-0.0594	0.0472	0.0059	-0.0216
	(0.0317)	(0.0201)	(0.0131)	(0.0184)	(0.0260)	(0.0109)	(0.0151)
XOM	-0.0410	0.0434	-0.0132	-0.0873	-0.0197	-0.0298	-0.0201
	(0.0458)	(0.0157)	(0.0086)	(0.0222)	(0.0440)	(0.0152)	(0.0230)
JNJ	0.0764	0.0326	0.0046	-0.0315	-0.0059	0.0201	0.0290
	(0.0319)	(0.0142)	(0.0086)	(0.0186)	(0.0332)	(0.0124)	(0.0194)
INTC	0.0334	0.0005	0.0426	-0.0582	0.0638	0.0074	-0.0478
	(0.0276)	(0.0241)	(0.0169)	(0.0182)	(0.0309)	(0.0123)	(0.0193)
WMT	0.0647	-0.0069	0.0332	-0.0466	0.0289	-0.0000	0.0178
	(0.0352)	(0.0177)	(0.0102)	(0.0184)	(0.0320)	(0.0097)	(0.0174)
CVX	-0.0823	0.0142	0.0181	-0.0536	-0.0393	-0.0077	-0.0313
	(0.0430)	(0.0166)	(0.0090)	(0.0210)	(0.0443)	(0.0137)	(0.0221)
UNH	-0.0332	0.0257	0.0387	0.0245	-0.0399	0.0035	-0.0110
	(0.0600)	(0.0240)	(0.0140)	(0.0211)	(0.0403)	(0.0151)	(0.0211)
CSCO	0.1418	-0.0102	0.0665	-0.0662	-0.0096	0.0232	-0.0149
	(0.1021)	(0.0278)	(0.0182)	(0.0210)	(0.0309)	(0.0109)	(0.0178)
HD	0.0456	0.0103	0.0350	0.0231	-0.0503	0.0244	-0.0071
	(0.0363)	(0.0205)	(0.0132)	(0.0191)	(0.0280)	(0.0122)	(0.0190)
PFE	0.1436	-0.0053	0.0318	0.0039	-0.0094	-0.0053	0.0326
	(0.0286)	(0.0182)	(0.0118)	(0.0169)	(0.0278)	(0.0115)	(0.0192)
BA	-0.0334	0.0292	0.0121	-0.0040	0.0305	-0.0096	0.0165
	(0.0310)	(0.0190)	(0.0125)	(0.0193)	(0.0394)	(0.0140)	(0.0184)
VZ	0.0534	0.0120	0.0013	-0.0409	-0.0275	-0.0090	0.0004
	(0.0425)	(0.0171)	(0.0096)	(0.0197)	(0.0368)	(0.0116)	(0.0184)
PG	0.0788	0.0867	-0.0474	-0.0627	0.0782	-0.0267	-0.0211
	(0.0262)	(0.0148)	(0.0106)	(0.0218)	(0.0278)	(0.0135)	(0.0144)
KO	0.0549	0.0536	-0.0265	-0.0200	0.0466	-0.0171	0.0463
	(0.0352)	(0.0148)	(0.0087)	(0.0195)	(0.0347)	(0.0126)	(0.0168)
MRK	0.0014	0.0310	-0.0162	-0.0010	-0.0090	0.0114	0.0018
	(0.0284)	(0.0177)	(0.0118)	(0.0197)	(0.0258)	(0.0116)	(0.0187)
DIS	0.0344	0.0302	0.0024	-0.0350	-0.0003	-0.0117	-0.0107
	(0.0291)	(0.0191)	(0.0132)	(0.0197)	(0.0367)	(0.0131)	(0.0218)
IBM	-0.0115	0.0414	-0.0017	-0.0405	0.0466	0.0029	-0.0584
	(0.0298)	(0.0179)	(0.0119)	(0.0175)	(0.0250)	(0.0109)	(0.0177)
GE	0.1175	-0.0200	0.0317	-0.0293	0.0243	0.0012	0.0221
	(0.0489)	(0.0190)	(0.0124)	(0.0281)	(0.0419)	(0.0171)	(0.0281)
MCD	0.1784	0.0515	-0.0083	-0.0297	0.0307	-0.0184	0.0171
	(0.0472)	(0.0164)	(0.0100)	(0.0195)	(0.0291)	(0.0112)	(0.0175)
MMM	-0.0267	0.0275	0.0088	-0.0148	-0.0040	-0.0294	-0.0176
	(0.0326)	(0.0152)	(0.0088)	(0.0177)	(0.0295)	(0.0108)	(0.0199)
NKE	0.0334	0.0551	-0.0069	0.0115	-0.0116	-0.0059	-0.0147
	(0.0351)	(0.0212)	(0.0127)	(0.0202)	(0.0276)	(0.0117)	(0.0152)
UTX	-0.0574	0.0079	0.0409	-0.0275	-0.0160	0.0011	-0.0316
	(0.0711)	(0.0174)	(0.0101)	(0.0200)	(0.0313)	(0.0104)	(0.0164)
CAT	0.0238	-0.0058	0.0562	-0.0038	-0.0257	0.0170	-0.0121
	(0.0351)	(0.0208)	(0.0133)	(0.0175)	(0.0249)	(0.0114)	(0.0174)
AXP	-0.0642	0.0265	0.0176	-0.0573	0.0235	-0.0061	-0.0499
	(0.0449)	(0.0230)	(0.0129)	(0.0214)	(0.0436)	(0.0138)	(0.0261)
TRV	0.1572	0.0242	0.0078	-0.0423	-0.0557	-0.0207	-0.0282
	(0.0673)	(0.0196)	(0.0106)	(0.0299)	(0.0448)	(0.0191)	(0.0214)
average	0.0353	0.0211	0.0172	-0.0311	0.0067	-0.0033	-0.0082
pool est.	0.0404	0.0209	0.0170	-0.0321	0.0132	-0.0013	-0.0098
pool s.e.	(0.0038)	(0.0022)	(0.0014)	(0.0018)	(0.0032)	(0.0012)	(0.0019)

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

Table A.2: Estimates of the mean equations with Dow Jones stocks

	β_D	γ_D	ρ_D	γ_D^*	ρ_D^*	ν_D	ω_D
AAPL	0.8939	0.0466	0.0479	-0.0187	-0.0093	7.3824	0.5180
	(0.0243)	(0.0066)	(0.0088)	(0.0037)	(0.0038)	(0.6270)	(0.0205)
MSFT	0.9311	0.0469	0.0620	-0.0120	-0.0080	9.1070	0.1991
	(0.0118)	(0.0052)	(0.0072)	(0.0032)	(0.0035)	(0.9149)	(0.0251)
XOM	0.9582	0.0399	0.0309	-0.0114	-0.0180	10.4484	-0.0396
	(0.0074)	(0.0046)	(0.0045)	(0.0027)	(0.0029)	(1.1757)	(0.0281)
JNJ	0.9408	0.0433	0.0468	-0.0216	-0.0114	7.6657	-0.1929
	(0.0095)	(0.0049)	(0.0059)	(0.0032)	(0.0033)	(0.6855)	(0.0257)
INTC	0.9630	0.0254	0.0292	-0.0123	-0.0125	10.6042	0.3897
	(0.0076)	(0.0035)	(0.0050)	(0.0023)	(0.0027)	(1.2275)	(0.0247)
WMT	0.9625	0.0294	0.0333	-0.0107	-0.0086	6.9672	-0.0052
	(0.0101)	(0.0045)	(0.0060)	(0.0027)	(0.0028)	(0.5684)	(0.0266)
CVX	0.9660	0.0302	0.0280	-0.0181	-0.0158	12.1562	0.0465
	(0.0061)	(0.0038)	(0.0042)	(0.0023)	(0.0027)	(1.6333)	(0.0276)
UNH	0.9590	0.0372	0.0353	-0.0216	-0.0121	6.3959	0.2983
	(0.0081)	(0.0046)	(0.0059)	(0.0031)	(0.0030)	(0.4884)	(0.0275)
CSCO	0.9484	0.0328	0.0387	-0.0203	-0.0197	9.1977	0.3973
	(0.0084)	(0.0044)	(0.0053)	(0.0030)	(0.0029)	(0.8801)	(0.0234)
HD	0.9603	0.0327	0.0346	-0.0235	-0.0167	7.4076	0.1840
	(0.0068)	(0.0041)	(0.0053)	(0.0029)	(0.0029)	(0.6287)	(0.0267)
PFE	0.9558	0.0323	0.0411	-0.0163	-0.0042	8.5713	0.1039
	(0.0092)	(0.0045)	(0.0060)	(0.0028)	(0.0028)	(0.8350)	(0.0261)
BA	0.9589	0.0336	0.0289	-0.0100	-0.0122	7.2430	0.1563
	(0.0093)	(0.0047)	(0.0056)	(0.0026)	(0.0027)	(0.6080)	(0.0259)
VZ	0.9669	0.0281	0.0294	-0.0075	-0.0158	9.0911	0.0432
	(0.0067)	(0.0039)	(0.0047)	(0.0023)	(0.0026)	(0.9090)	(0.0273)
PG	0.9504	0.0340	0.0381	-0.0169	-0.0142	7.5308	-0.1665
	(0.0093)	(0.0045)	(0.0054)	(0.0030)	(0.0032)	(0.6385)	(0.0246)
KO	0.9688	0.0267	0.0265	-0.0149	-0.0161	7.6841	-0.1440
	(0.0068)	(0.0040)	(0.0050)	(0.0025)	(0.0026)	(0.7088)	(0.0274)
MRK	0.9168	0.0469	0.0529	-0.0175	-0.0105	6.5537	0.0572
	(0.0188)	(0.0064)	(0.0086)	(0.0037)	(0.0036)	(0.4884)	(0.0224)
DIS	0.9522	0.0348	0.0418	-0.0153	-0.0114	8.3087	0.1375
	(0.0089)	(0.0046)	(0.0059)	(0.0028)	(0.0030)	(0.7628)	(0.0255)
IBM	0.9421	0.0402	0.0437	-0.0205	-0.0083	7.0256	0.0311
	(0.0123)	(0.0057)	(0.0067)	(0.0032)	(0.0032)	(0.5721)	(0.0247)
GE	0.9552	0.0419	0.0452	-0.0201	-0.0177	8.7128	0.0551
	(0.0070)	(0.0047)	(0.0055)	(0.0028)	(0.0029)	(0.8412)	(0.0294)
MCD	0.9119	0.0422	0.0560	-0.0143	-0.0098	6.9078	-0.0575
	(0.0225)	(0.0069)	(0.0086)	(0.0037)	(0.0037)	(0.5493)	(0.0216)
MMM	0.9726	0.0216	0.0305	-0.0190	-0.0110	6.2805	-0.1271
	(0.0051)	(0.0035)	(0.0044)	(0.0024)	(0.0026)	(0.4491)	(0.0270)
NKE	0.9683	0.0308	0.0319	-0.0154	-0.0082	6.2075	0.1867
	(0.0072)	(0.0046)	(0.0057)	(0.0027)	(0.0025)	(0.4620)	(0.0291)
UTX	0.9603	0.0310	0.0349	-0.0211	-0.0159	7.5010	0.0104
	(0.0072)	(0.0046)	(0.0051)	(0.0027)	(0.0028)	(0.6710)	(0.0264)
CAT	0.9747	0.0238	0.0284	-0.0162	-0.0111	7.0967	0.2441
	(0.0051)	(0.0040)	(0.0046)	(0.0021)	(0.0022)	(0.5959)	(0.0294)
AXP	0.9735	0.0344	0.0322	-0.0161	-0.0137	8.2837	0.2124
	(0.0047)	(0.0040)	(0.0043)	(0.0024)	(0.0026)	(0.7934)	(0.0347)
TRV	0.9618	0.0439	0.0421	-0.0144	-0.0109	6.8122	0.0290
	(0.0074)	(0.0051)	(0.0059)	(0.0029)	(0.0029)	(0.5562)	(0.0317)
average	0.9528	0.0350	0.0381	-0.0164	-0.0124	7.9670	0.0987
pool est.	0.9844	0.0324	0.0349	-0.0065	-0.0026	6.8668	-1.0370
pool s.e.	(0.0005)	(0.0010)	(0.0007)	(0.0003)	(0.0003)	(0.1054)	(0.0376)

Continued on the next page.

Table A.3: Estimates of the dynamic parameters in the univariate coupled component models with Dow Jones stocks

	β_N	γ_N	ρ_N	γ_N^*	ρ_N^*	ν_N	ω_N
AAPL	0.9087	0.0674	0.0685	-0.0132	-0.0186	2.5831	-0.4023
	(0.0143)	(0.0087)	(0.0069)	(0.0046)	(0.0044)	(0.0913)	(0.0259)
MSFT	0.9411	0.0621	0.0463	-0.0086	-0.0083	2.7236	-0.7218
	(0.0103)	(0.0080)	(0.0054)	(0.0039)	(0.0036)	(0.1001)	(0.0294)
XOM	0.9720	0.0306	0.0306	-0.0148	-0.0126	4.6413	-0.8132
	(0.0053)	(0.0046)	(0.0037)	(0.0027)	(0.0025)	(0.2715)	(0.0323)
JNJ	0.9633	0.0309	0.0330	-0.0152	-0.0240	3.6899	-0.9392
	(0.0057)	(0.0049)	(0.0041)	(0.0028)	(0.0028)	(0.1838)	(0.0281)
INTC	0.9616	0.0377	0.0325	-0.0139	-0.0144	2.7884	-0.4623
	(0.0101)	(0.0074)	(0.0053)	(0.0035)	(0.0029)	(0.1051)	(0.0292)
WMT	0.9684	0.0300	0.0303	-0.0105	-0.0125	3.0230	-0.8379
	(0.0064)	(0.0052)	(0.0041)	(0.0029)	(0.0028)	(0.1266)	(0.0295)
CVX	0.9800	0.0228	0.0251	-0.0115	-0.0139	3.8001	-0.8119
	(0.0037)	(0.0041)	(0.0030)	(0.0024)	(0.0022)	(0.2027)	(0.0342)
UNH	0.9488	0.0502	0.0492	-0.0108	-0.0219	2.1709	-0.8229
	(0.0087)	(0.0074)	(0.0059)	(0.0041)	(0.0040)	(0.0789)	(0.0310)
CSCO	0.9580	0.0444	0.0394	-0.0155	-0.0206	3.1447	-0.4090
	(0.0073)	(0.0065)	(0.0048)	(0.0032)	(0.0032)	(0.1264)	(0.0302)
HD	0.9605	0.0435	0.0359	-0.0204	-0.0230	2.9114	-0.6871
	(0.0068)	(0.0061)	(0.0046)	(0.0036)	(0.0034)	(0.1228)	(0.0311)
PFE	0.9620	0.0451	0.0357	-0.0086	-0.0149	3.0342	-0.7291
	(0.0073)	(0.0061)	(0.0046)	(0.0031)	(0.0031)	(0.1295)	(0.0321)
BA	0.9756	0.0234	0.0297	-0.0118	-0.0157	2.7995	-0.6943
	(0.0054)	(0.0054)	(0.0041)	(0.0026)	(0.0024)	(0.1152)	(0.0320)
VZ	0.9681	0.0347	0.0310	-0.0213	-0.0114	3.1998	-0.8313
	(0.0055)	(0.0054)	(0.0040)	(0.0032)	(0.0028)	(0.1453)	(0.0312)
PG	0.9514	0.0335	0.0324	-0.0210	-0.0221	3.1983	-1.0012
	(0.0086)	(0.0057)	(0.0049)	(0.0034)	(0.0032)	(0.1402)	(0.0252)
KO	0.9739	0.0383	0.0230	-0.0212	-0.0190	3.3124	-0.9450
	(0.0044)	(0.0051)	(0.0038)	(0.0029)	(0.0027)	(0.1467)	(0.0342)
MRK	0.9347	0.0452	0.0488	-0.0136	-0.0209	2.8925	-0.7636
	(0.0141)	(0.0073)	(0.0071)	(0.0039)	(0.0038)	(0.1203)	(0.0259)
DIS	0.9521	0.0394	0.0482	-0.0185	-0.0231	3.0567	-0.6560
	(0.0086)	(0.0067)	(0.0055)	(0.0035)	(0.0033)	(0.1302)	(0.0287)
IBM	0.9576	0.0463	0.0398	-0.0156	-0.0222	2.6542	-0.8477
	(0.0069)	(0.0064)	(0.0049)	(0.0034)	(0.0031)	(0.1004)	(0.0309)
GE	0.9606	0.0488	0.0462	-0.0223	-0.0217	4.0438	-0.6823
	(0.0060)	(0.0057)	(0.0047)	(0.0031)	(0.0031)	(0.2042)	(0.0345)
MCD	0.9431	0.0398	0.0393	-0.0159	-0.0212	3.0874	-0.8367
	(0.0103)	(0.0064)	(0.0053)	(0.0035)	(0.0034)	(0.1321)	(0.0254)
MMM	0.9701	0.0378	0.0308	-0.0139	-0.0247	2.7494	-1.0112
	(0.0059)	(0.0057)	(0.0049)	(0.0033)	(0.0032)	(0.1160)	(0.0334)
NKE	0.9545	0.0418	0.0448	-0.0175	-0.0186	2.1900	-0.9093
	(0.0101)	(0.0072)	(0.0062)	(0.0038)	(0.0039)	(0.0756)	(0.0298)
UTX	0.9742	0.0314	0.0301	-0.0217	-0.0192	3.0986	-0.8425
	(0.0041)	(0.0048)	(0.0040)	(0.0029)	(0.0027)	(0.1389)	(0.0340)
CAT	0.9715	0.0331	0.0396	-0.0115	-0.0202	2.7066	-0.6204
	(0.0055)	(0.0059)	(0.0048)	(0.0029)	(0.0029)	(0.1092)	(0.0351)
AXP	0.9748	0.0392	0.0379	-0.0213	-0.0227	3.5189	-0.6356
	(0.0041)	(0.0055)	(0.0040)	(0.0031)	(0.0028)	(0.1703)	(0.0401)
TRV	0.9725	0.0445	0.0375	-0.0135	-0.0186	2.7308	-0.9772
	(0.0051)	(0.0066)	(0.0046)	(0.0031)	(0.0030)	(0.1105)	(0.0390)
average	0.9600	0.0401	0.0379	-0.0155	-0.0187	3.0673	-0.7650
pool est.	0.9819	0.0448	0.0306	-0.0087	-0.0119	2.9195	-1.6942
pool s.e.	(0.0006)	(0.0010)	(0.0009)	(0.0005)	(0.0004)	(0.0235)	(0.0365)

This table presents the estimates of the dynamic parameters 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 A.3: Estimates of the dynamic parameters in the univariate coupled component models with Dow Jones stocks (cont.)

	$\beta_D = \beta_N$	$\gamma_D = \gamma_N$	$\rho_D = \rho_N$	$\gamma_D^* = \gamma_N^*$	$\rho_D^* = \rho_N^*$	$\nu_D = \nu_N$	$\omega_D = \omega_N$	$\gamma_N = \rho_D$	$(\beta_D, \gamma_D, \rho_D, \gamma_D^*, \rho_D^*) = (\beta_N, \gamma_N, \rho_N, \gamma_N^*, \rho_N^*)$
AAPL	0.4151	0.0377	0.0435	0.3566	0.1238	0.0000	0.0000	0.0204	0.0000
MSFT	0.2755	0.1077	0.0632	0.4994	0.9542	0.0000	0.0000	0.9905	0.1502
XOM	0.0086	0.1299	0.9602	0.3611	0.1673	0.0000	0.0000	0.9444	0.0569
JNJ	0.0015	0.0714	0.0445	0.1289	0.0036	0.0000	0.0000	0.0035	0.0013
INTC	0.8272	0.1144	0.6035	0.6919	0.6275	0.0000	0.0000	0.1661	0.2030
WMT	0.3816	0.9349	0.6629	0.9595	0.3274	0.0000	0.0000	0.5270	0.7146
CVX	0.0007	0.1598	0.5616	0.0378	0.5835	0.0000	0.0000	0.1917	0.0147
UNH	0.1352	0.1227	0.0734	0.0351	0.0541	0.0000	0.0000	0.0197	0.0098
CSCO	0.1193	0.1329	0.9078	0.2625	0.8379	0.0000	0.0000	0.3181	0.0077
HD	0.9682	0.1249	0.8408	0.4931	0.1554	0.0000	0.0000	0.1032	0.1870
PFE	0.3310	0.0806	0.4280	0.0572	0.0100	0.0000	0.0000	0.4842	0.0055
BA	0.0101	0.1194	0.8855	0.6147	0.3219	0.0000	0.0000	0.2627	0.0157
VZ	0.8093	0.2987	0.7780	0.0004	0.2581	0.0000	0.0000	0.2800	0.0046
PG	0.8966	0.9388	0.4021	0.3522	0.0850	0.0000	0.0000	0.4153	0.2592
KO	0.3146	0.0533	0.5554	0.0861	0.4268	0.0000	0.0000	0.0132	0.0071
MRK	0.1612	0.8459	0.6448	0.4568	0.0481	0.0000	0.0000	0.3101	0.1931
DIS	0.9901	0.5488	0.3388	0.4718	0.0097	0.0000	0.0000	0.6606	0.0056
IBM	0.0944	0.4629	0.6043	0.2844	0.0022	0.0000	0.0000	0.6905	0.0040
GE	0.2330	0.3307	0.8871	0.6107	0.3594	0.0000	0.0000	0.4652	0.0176
MCD	0.0733	0.7976	0.0614	0.7449	0.0236	0.0000	0.0000	0.0467	0.0283
MMM	0.5296	0.0133	0.9672	0.2115	0.0008	0.0000	0.0000	0.1225	0.0007
NKE	0.0387	0.1604	0.0719	0.6335	0.0222	0.0000	0.0000	0.0880	0.0759
UTX	0.0078	0.9547	0.4004	0.8796	0.3925	0.0000	0.0000	0.4476	0.0062
CAT	0.3085	0.1572	0.0438	0.1859	0.0106	0.0000	0.0000	0.2758	0.0023
AXP	0.6457	0.4545	0.2875	0.1735	0.0171	0.0000	0.0000	0.1199	0.0003
TRV	0.0358	0.9315	0.4881	0.8417	0.0730	0.0000	0.0000	0.6667	0.0173

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 A.4: Wald tests based on the univariate model with Dow Jones stocks

	student t log-likelihood				quasi Gaussian log-likelihood			
	$l_{cogarch}$	l_{garch}	GW stat.	p-val.	$l_{cogarch}$	l_{garch}	GW stat.	p-val.
AAPL	1.2657	1.2753	1.7634	0.1842	1.3065	1.3141	0.4312	0.5114
MSFT	1.1802	1.1861	0.7936	0.3730	1.1981	1.2087	1.1950	0.2743
XOM	1.0570	1.0625	0.4534	0.5007	1.0742	1.0852	1.1503	0.2835
JNJ	0.9272	0.9357	2.3027	0.1292	0.9506	0.9603	1.9578	0.1617
INTC	1.2798	1.2862	0.9449	0.3310	1.3008	1.3071	0.6065	0.4361
WMT	1.1544	1.1669	4.0642	0.0438	1.1931	1.2100	4.0114	0.0452
CVX	1.2181	1.2235	0.6455	0.4217	1.2346	1.2440	1.1811	0.2771
UNH	1.2165	1.2221	1.9359	0.1641	1.2370	1.2493	6.0665	0.0138
CSCO	1.1261	1.1399	2.3208	0.1277	1.1313	1.1481	2.2340	0.1350
HD	1.1719	1.1767	0.5672	0.4514	1.1959	1.2052	0.9658	0.3257
PFE	1.0962	1.1017	0.3038	0.5815	1.1377	1.1362	0.0091	0.9239
BA	1.3022	1.3165	8.4277	0.0037	1.3546	1.3723	6.2982	0.0121
VZ	1.2136	1.2101	0.2579	0.6116	1.2317	1.2231	0.7008	0.4025
PG	0.9218	0.9340	2.8485	0.0915	0.9513	0.9751	4.0699	0.0437
KO	0.9105	0.9222	3.5267	0.0604	0.9440	0.9650	5.2306	0.0222
MRK	1.1660	1.1728	1.6389	0.2005	1.1972	1.2112	2.6679	0.1024
DIS	1.0950	1.0977	0.1423	0.7060	1.1430	1.1348	0.2034	0.6520
IBM	1.1104	1.1241	4.8065	0.0284	1.1229	1.1414	5.5228	0.0188
GE	1.2193	1.2466	6.0812	0.0137	1.2767	1.3229	3.6455	0.0562
MCD	1.0193	1.0189	0.0020	0.9642	1.0659	1.0791	0.4837	0.4868
MMM	0.9235	0.9297	0.5145	0.4732	0.9563	0.9589	0.0261	0.8717
NKE	1.4594	1.4616	0.2770	0.5987	1.4681	1.4681	0.0000	0.9990
UTX	1.0734	1.0773	0.1882	0.6644	1.1276	1.1273	0.0003	0.9852
CAT	1.4229	1.4310	1.8841	0.1699	1.4499	1.4574	0.5936	0.4410
AXP	1.1714	1.1779	1.4314	0.2315	1.1903	1.2024	2.5820	0.1081
TRV	1.0654	1.0949	10.9938	0.0009	1.0947	1.1206	4.3900	0.0362

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 A.5: GW tests based on the univariate model for Dow Jones stocks

	$\lvert\epsilon_d\rvert$	$\lvert\epsilon_n\rvert$	$\lvert r_d\rvert$	$\lvert r_n\rvert$	ϵ_d^2	ϵ_n^2	r_d^2	r_n^2
AAPL	0.8161	0.2252	0.0000	0.0000	0.1106	0.9982	0.0000	0.9346
MSFT	0.1657	0.5292	0.0000	0.0000	0.1868	0.9999	0.0000	0.0000
XOM	0.2944	0.4334	0.0000	0.0000	0.6872	0.8370	0.0000	0.0000
JNJ	0.1181	0.6729	0.0000	0.0000	0.1503	0.9677	0.0000	0.0186
INTC	0.3359	0.2107	0.0000	0.0000	0.6822	0.9965	0.0000	0.0002
WMT	0.3929	0.1679	0.0000	0.0000	0.9992	0.9768	0.0000	0.0000
CVX	0.3463	0.9379	0.0000	0.0000	0.6090	0.6604	0.0000	0.0000
UNH	0.0109	0.6807	0.0000	0.0000	0.4687	0.9980	0.0000	0.0000
CSCO	0.1087	0.7825	0.0000	0.0000	0.6531	1.0000	0.0000	0.9200
HD	0.0680	0.8480	0.0000	0.0000	0.1221	0.9993	0.0000	0.2121
PFE	0.4283	0.2030	0.0000	0.0000	0.1371	1.0000	0.0000	0.0034
BA	0.0170	0.5694	0.0000	0.0000	0.0844	0.9983	0.0000	0.0000
VZ	0.0154	0.1840	0.0000	0.0000	0.2066	0.0005	0.0000	0.0000
PG	0.4321	0.9061	0.0000	0.0000	0.5520	1.0000	0.0000	1.0000
KO	0.0429	0.1192	0.0000	0.0000	0.0580	0.8637	0.0000	0.0000
MRK	0.2432	0.8689	0.0000	0.0000	0.0000	1.0000	0.0000	1.0000
DIS	0.8809	0.8123	0.0000	0.0000	0.8944	1.0000	0.0000	0.0000
IBM	0.9042	0.6381	0.0000	0.0000	0.7996	0.9997	0.0000	0.0000
GE	0.7434	0.2261	0.0000	0.0000	0.0001	0.9989	0.0000	0.0000
MCD	0.3076	0.6982	0.0000	0.0000	0.0179	0.8558	0.0000	0.0000
MMM	0.0993	0.1647	0.0000	0.0000	0.6664	0.9683	0.0000	0.0000
NKE	0.0137	0.3498	0.0000	0.0000	0.0307	0.9993	0.0000	0.9998
UTX	0.2131	0.4790	0.0000	0.0000	0.0008	1.0000	0.0000	0.0005
CAT	0.8231	0.0107	0.0000	0.0000	0.9362	0.9968	0.0000	0.0000
AXP	0.3672	0.0598	0.0000	0.0000	0.3445	0.9923	0.0000	0.0000
TRV	0.7044	0.2810	0.0000	0.0000	0.0094	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 A.6: Diagnostic checking for GARCH effects in Dow Jones stocks

	β_D	γ_D	ρ_D	γ_D^*	ρ_D^*	ν_D	ω_D
AAPL	0.8209	0.0666	0.0723	-0.0225	0.0049	5.0414	0.4866
	(0.0498)	(0.0086)	(0.0118)	(0.0057)	(0.0056)	(0.3013)	(0.0194)
MSFT	0.8550	0.0624	0.0794	0.0018	0.0041	7.4535	0.1480
	(0.0265)	(0.0069)	(0.0093)	(0.0041)	(0.0045)	(0.6494)	(0.0209)
XOM	0.7401	0.0759	0.0766	0.0004	-0.0040	8.6252	-0.0985
	(0.2115)	(0.0228)	(0.0256)	(0.0034)	(0.0037)	(1.0690)	(0.0231)
JNJ	0.6549	0.0742	0.0799	-0.0055	0.0176	6.3548	-0.2999
	(0.0755)	(0.0085)	(0.0097)	(0.0058)	(0.0069)	(0.4752)	(0.0174)
INTC	0.8928	0.0356	0.0382	-0.0051	-0.0038	7.0961	0.3706
	(0.0374)	(0.0070)	(0.0096)	(0.0033)	(0.0036)	(0.5904)	(0.0190)
WMT	0.7801	0.0545	0.0800	-0.0056	0.0061	6.7666	-0.0424
	(0.0534)	(0.0082)	(0.0126)	(0.0049)	(0.0055)	(0.5396)	(0.0181)
CVX	0.9413	0.0382	0.0371	-0.0125	-0.0053	8.0180	0.0012
	(0.0161)	(0.0059)	(0.0069)	(0.0030)	(0.0032)	(0.7854)	(0.0241)
UNH	0.8913	0.0500	0.0756	-0.0279	-0.0085	5.5086	0.2686
	(0.0246)	(0.0059)	(0.0090)	(0.0039)	(0.0038)	(0.3877)	(0.0231)
CSCO	0.8615	0.0577	0.0644	-0.0209	-0.0139	6.6629	0.4197
	(0.0338)	(0.0081)	(0.0107)	(0.0044)	(0.0041)	(0.5053)	(0.0201)
HD	0.7042	0.0792	0.0800	-0.0180	-0.0025	7.1077	0.1901
	(0.0608)	(0.0086)	(0.0111)	(0.0052)	(0.0057)	(0.6185)	(0.0184)
PFE	0.8939	0.0418	0.0597	-0.0051	0.0068	6.8400	0.0690
	(0.0380)	(0.0079)	(0.0136)	(0.0040)	(0.0043)	(0.5364)	(0.0208)
BA	0.6283	0.0652	0.0524	-0.0019	0.0205	7.4862	0.1528
	(0.1236)	(0.0079)	(0.0097)	(0.0059)	(0.0068)	(-0.6935)	(0.0169)
VZ	0.8421	0.0451	0.0775	0.0011	-0.0067	7.0443	-0.0159
	(0.0348)	(0.0072)	(0.0101)	(0.0042)	(0.0044)	(0.5798)	(0.0189)
PG	0.9030	0.0337	0.0728	-0.0134	-0.0089	6.6076	-0.2771
	(0.0264)	(0.0041)	(0.0096)	(0.0037)	(0.0030)	(0.4268)	(0.0584)
KO	0.7442	0.0628	0.0798	-0.0127	0.0046	7.6692	-0.1914
	(0.0583)	(0.0086)	(0.0107)	(0.0048)	(0.0058)	(0.7301)	(0.0184)
MRK	0.7085	0.0503	0.0800	-0.0115	-0.0027	6.4271	0.0341
	(0.0600)	(0.0072)	(0.0103)	(0.0053)	(0.0055)	(0.5181)	(0.0176)
DIS	0.8399	0.0360	0.0799	-0.0053	0.0001	7.0943	0.1094
	(0.0495)	(0.0061)	(0.0149)	(0.0042)	(0.0043)	(0.5746)	(0.0181)
IBM	0.8581	0.0539	0.0800	-0.0093	0.0031	5.6215	-0.0583
	(0.0289)	(0.0070)	(0.0110)	(0.0042)	(0.0046)	(0.3833)	(0.0206)
GE	0.8943	0.0570	0.0799	-0.0132	0.0013	6.2792	-0.0056
	(0.0242)	(0.0080)	(0.0107)	(0.0040)	(0.0043)	(0.4449)	(0.0229)
MCD	0.7019	0.0700	0.0800	-0.0119	0.0035	6.5152	-0.1040
	(0.0512)	(0.0090)	(0.0107)	(0.0058)	(0.0056)	(0.5211)	(0.0181)
MMM	0.9557	0.0147	0.0456	-0.0134	-0.0015	5.2657	-0.3001
	(0.0071)	(0.0041)	(0.0063)	(0.0029)	(0.0026)	(0.2927)	(0.0532)
NKE	0.7149	0.0569	0.0795	-0.0089	0.0041	5.5639	0.1775
	(0.1116)	(0.0087)	(0.0184)	(0.0056)	(0.0057)	(0.3727)	(0.0173)
UTX	0.8262	0.0618	0.0763	-0.0062	-0.0067	6.6471	-0.0375
	(0.0342)	(0.0071)	(0.0105)	(0.0046)	(0.0048)	(0.5314)	(0.0197)
CAT	0.9237	0.0337	0.0407	-0.0107	0.0019	6.3817	0.2577
	(0.0251)	(0.0060)	(0.0113)	(0.0033)	(0.0036)	(0.4747)	(0.0200)
AXP	0.9766	0.0307	0.0291	-0.0101	-0.0084	7.4893	0.2164
	(0.0066)	(0.0031)	(0.0031)	(0.0021)	(0.0022)	(0.6929)	(0.0408)
TRV	0.9098	0.0560	0.0688	-0.0095	0.0019	6.2834	0.0014
	(0.0160)	(0.0060)	(0.0083)	(0.0037)	(0.0038)	(0.4759)	(0.0239)

Continued on the next page.

Table A.7: Estimates of the dynamic parameters in the multivariate coupled component model with Dow Jones stocks

	β_N	γ_N	ρ_N	γ_N^*	ρ_N^*	ν_N	ω_N
AAPL	0.9186	0.0799	0.0779	-0.0024	-0.0029	2.1504	-0.5227
	(0.0159)	(0.0101)	(0.0094)	(0.0045)	(0.0044)	(0.0692)	(0.0297)
MSFT	0.8875	0.0682	0.0523	0.0020	0.0049	2.3927	-0.9326
	(0.0206)	(0.0087)	(0.0073)	(0.0045)	(0.0044)	(0.0802)	(0.0234)
XOM	0.8984	0.0484	0.0181	-0.0038	0.0001	3.7163	-0.9379
	(0.0547)	(0.0118)	(0.0051)	(0.0029)	(0.0029)	(0.1854)	(0.0264)
JNJ	0.9537	0.0213	0.0257	0.0015	-0.0095	3.2702	-1.1164
	(0.0197)	(0.0089)	(0.0092)	(0.0030)	(0.0029)	(0.1490)	(0.0229)
INTC	0.8111	0.0799	0.0670	0.0001	-0.0041	2.2523	-0.5951
	(0.0325)	(0.0109)	(0.0090)	(0.0060)	(0.0054)	(0.0750)	(0.0214)
WMT	0.8222	0.0257	0.0476	0.0055	-0.0101	2.7032	-0.9731
	(0.0574)	(0.0083)	(0.0085)	(0.0057)	(0.0051)	(0.1009)	(0.0180)
CVX	0.9684	0.0168	0.0244	-0.0016	-0.0055	3.1493	-0.9066
	(0.0076)	(0.0044)	(0.0038)	(0.0025)	(0.0024)	(0.1426)	(0.0249)
UNH	0.8183	0.0766	0.0696	0.0149	-0.0225	2.0431	-0.9057
	(0.0549)	(0.0105)	(0.0091)	(0.0047)	(0.0041)	(0.0695)	(0.0344)
CSCO	0.8905	0.0718	0.0619	0.0021	-0.0192	2.5697	-0.5082
	(0.0235)	(0.0101)	(0.0083)	(0.0048)	(0.0048)	(0.0905)	(0.0242)
HD	0.8926	0.0242	0.0583	-0.0004	-0.0153	2.6526	-0.7484
	(0.0226)	(0.0061)	(0.0099)	(0.0040)	(0.0044)	(0.0991)	(0.0202)
PFE	0.9301	0.0488	0.0422	0.0058	0.0003	2.7134	-0.8251
	(0.0184)	(0.0069)	(0.0081)	(0.0043)	(0.0039)	(0.1037)	(0.0243)
BA	0.6984	0.0340	0.0725	0.0014	0.0002	2.4611	-0.7913
	(0.1696)	(0.0101)	(0.0221)	(0.0047)	(0.0054)	(0.0905)	(0.0192)
VZ	0.9240	0.0447	0.0357	-0.0087	-0.0000	2.8176	-0.9588
	(0.0145)	(0.0075)	(0.0057)	(0.0036)	(0.0037)	(0.1086)	(0.0231)
PG	0.9067	0.0401	0.0326	-0.0080	-0.0080	2.8072	-1.2496
	(0.0340)	(0.0054)	(0.0048)	(0.0034)	(0.0039)	(0.1301)	(0.0487)
KO	0.8519	0.0435	0.0427	-0.0081	-0.0150	2.9273	-1.1082
	(0.0361)	(0.0082)	(0.0082)	(0.0047)	(0.0050)	(0.1176)	(0.0194)
MRK	0.9149	0.0477	0.0234	0.0027	-0.0045	2.6650	-0.8864
	(0.0456)	(0.0196)	(0.0074)	(0.0042)	(0.0040)	(0.1046)	(0.0232)
DIS	0.8194	0.0583	0.0784	-0.0089	-0.0003	2.6393	-0.7454
	(0.0302)	(0.0085)	(0.0087)	(0.0054)	(0.0053)	(0.0990)	(0.0206)
IBM	0.8641	0.0635	0.0624	-0.0095	-0.0154	2.2896	-1.0126
	(0.0247)	(0.0092)	(0.0080)	(0.0052)	(0.0048)	(0.0750)	(0.0225)
GE	0.9075	0.0599	0.0579	0.0063	-0.0174	3.1200	-0.8152
	(0.0148)	(0.0070)	(0.0063)	(0.0041)	(0.0041)	(0.1322)	(0.0246)
MCD	0.6463	0.0318	0.0679	0.0077	-0.0216	2.8182	-0.9581
	(0.1440)	(0.0090)	(0.0075)	(0.0063)	(0.0062)	(0.1115)	(0.0174)
MMM	0.9639	0.0275	0.0216	-0.0005	-0.0103	2.6088	-1.1856
	(0.0076)	(0.0053)	(0.0056)	(0.0028)	(0.0030)	(0.0983)	(0.0465)
NKE	0.6992	0.0323	0.0754	-0.0100	-0.0105	2.1239	-1.0059
	(0.0805)	(0.0107)	(0.0116)	(0.0070)	(0.0071)	(0.0691)	(0.0184)
UTX	0.9229	0.0386	0.0429	-0.0075	-0.0091	2.6970	-0.9454
	(0.0189)	(0.0076)	(0.0072)	(0.0040)	(0.0040)	(0.1047)	(0.0231)
CAT	0.9313	0.0521	0.0505	0.0079	-0.0122	2.3309	-0.6774
	(0.0152)	(0.0087)	(0.0072)	(0.0045)	(0.0041)	(0.0825)	(0.0254)
AXP	0.9814	0.0181	0.0280	-0.0104	-0.0138	3.0156	-0.7092
	(0.0042)	(0.0029)	(0.0025)	(0.0017)	(0.0018)	(0.1365)	(0.0416)
TRV	0.9368	0.0631	0.0453	-0.0041	-0.0062	2.5516	-1.0782
	(0.0117)	(0.0080)	(0.0061)	(0.0039)	(0.0039)	(0.0924)	(0.0285)

This table gives the estimates of the dynamic parameters in the multivariate coupled component model, and their asymptotic standard errors in parenthesis.

Table A.7: Estimates of the dynamic parameters in the multivariate coupled component model with Dow Jones stocks(cont.)

Panel a: size-based portfolios of NYSE/AMEX/NASDAQ stocks

	mean		s.e.		median			
decile	day(%)	night(%)	day(%)	night(%)	day(%)	night(%)	T stat.	avg. no.
1-smallest	0.0235	-0.2145	0.0076	0.0068	0.0503	-0.1403	23.3770	356.2961
2	0.0048	-0.0934	0.0071	0.0050	0.0573	-0.0540	11.3449	513.3339
3	-0.0238	-0.0312	0.0081	0.0049	0.0458	-0.0110	0.7832	577.7025
4	-0.0345	0.0002	0.0100	0.0055	0.0400	0.0048	-3.0362	606.8326
5	-0.0393	0.0147	0.0121	0.0061	0.0344	0.0239	-3.9901	617.5727
6	-0.0401	0.0202	0.0130	0.0065	0.0334	0.0275	-4.1401	631.0932
7	-0.0359	0.0236	0.0136	0.0068	0.0411	0.0302	-3.9237	641.4180
8	-0.0240	0.0187	0.0138	0.0069	0.0427	0.0319	-2.7744	652.2880
9	-0.0191	0.0223	0.0132	0.0069	0.0599	0.0369	-2.7725	665.5434
10-largest	-0.0078	0.0202	0.0122	0.0080	0.0308	0.0380	-1.9151	687.3414

Panel b: size-based portfolios of NYSE stocks

	mean		s.e.		median			
size	day(%)	night(%)	day(%)	night(%)	day(%)	night(%)	T stat.	avg. no.
1-smallest	-0.0035	-0.0073	0.0078	0.0046	0.0379	-0.0007	0.4216	212.7596
2	-0.0157	0.0048	0.0094	0.0049	0.0289	0.0040	-1.9407	229.7850
3	-0.0088	0.0040	0.0106	0.0054	0.0400	0.0070	-1.0732	232.3495
4	-0.0150	0.0100	0.0121	0.0060	0.0364	0.0077	-1.8538	232.4569
5	-0.0014	0.0088	0.0130	0.0063	0.0500	0.0085	-0.7002	234.6683
6	-0.0008	0.0095	0.0133	0.0066	0.0488	0.0140	-0.6909	233.0124
7	0.0002	0.0103	0.0126	0.0065	0.0530	0.0193	-0.7098	237.1872
8	0.0021	0.0135	0.0123	0.0066	0.0516	0.0289	-0.8183	237.0842
9	-0.0010	0.0192	0.0124	0.0071	0.0462	0.0312	-1.4140	240.4760
10-largest	0.0016	0.0121	0.0117	0.0079	0.0321	0.0299	-0.7499	241.3931

Panel c: size-based portfolios of NASDAQ stocks

	mean		s.e.		median			
decile	day(%)	night(%)	day(%)	night(%)	day(%)	night(%)	T stat.	avg. no.
1-smallest	0.0237	-0.3032	0.0095	0.0094	0.0466	-0.1860	24.4218	185.3430
2	0.0001	-0.1418	0.0081	0.0064	0.0442	-0.0877	13.7931	265.5334
3	-0.0050	-0.0789	0.0082	0.0055	0.0477	-0.0433	7.4848	296.0316
4	-0.0364	-0.0306	0.0092	0.0056	0.0330	-0.0059	-0.5414	318.8368
5	-0.0468	0.0040	0.0118	0.0063	0.0388	0.0168	-3.7825	329.8977
6	-0.0544	0.0243	0.0144	0.0071	0.0294	0.0322	-4.9052	334.0138
7	-0.0547	0.0249	0.0153	0.0076	0.0187	0.0386	-4.6738	341.5321
8	-0.0570	0.0360	0.0156	0.0077	0.0280	0.0520	-5.3511	340.0634
9	-0.0546	0.0345	0.0162	0.0081	0.0323	0.0475	-4.9159	351.3039
10-largest	-0.0450	0.0529	0.0175	0.0104	0.0352	0.0762	-4.8000	367.1565

This table gives the summary statistics for the intraday and overnight returns of 10 size-based portfolios: mean, median, standard error of the mean, T statistics, and the average number of stocks. The T statistics are used to test the null hypothesis that the mean value of overnight returns equals the mean value of intraday returns.

Table A.8: Descriptive statistics for intraday and overnight returns with the size-sorted portfolios

Panel A: size-based portfolios of NYSE/AMEX/NASDAQ stocks							
decile	δ	μ_D	μ_N	Π_{11}	Π_{12}	Π_{21}	Π_{22}
1-smallest	0.0108	0.0079	-0.1462	0.2909	-0.0403	0.0008	0.3194
	(0.0223)	(0.0079)	(0.0067)	(0.0189)	(0.0186)	(0.0166)	(0.0193)
2	-0.0740	-0.0009	-0.0757	0.2880	-0.0457	0.0484	0.1929
	(0.0341)	(0.0070)	(0.0051)	(0.0214)	(0.0296)	(0.0168)	(0.0267)
3	-0.0697	-0.0202	-0.0270	0.2453	-0.0699	0.0595	0.0901
	(0.0446)	(0.0078)	(0.0047)	(0.0241)	(0.0395)	(0.0154)	(0.0315)
4	-0.1068	-0.0303	0.0014	0.1220	-0.0304	0.0341	0.0017
	(0.0473)	(0.0097)	(0.0054)	(0.0258)	(0.0503)	(0.0159)	(0.0329)
5	-0.1111	-0.0381	0.0167	0.0280	-0.0065	0.0267	-0.0653
	(0.0573)	(0.0119)	(0.0060)	(0.0304)	(0.0596)	(0.0160)	(0.0307)
6	-0.0996	-0.0395	0.0224	0.0099	-0.0110	0.0177	-0.0723
	(0.0575)	(0.0130)	(0.0065)	(0.0289)	(0.0601)	(0.0153)	(0.0300)
7	-0.0833	-0.0351	0.0265	0.0181	-0.0063	0.0224	-0.0905
	(0.0543)	(0.0136)	(0.0068)	(0.0263)	(0.0565)	(0.0141)	(0.0289)
8	-0.0965	-0.0236	0.0213	0.0161	0.0009	0.0253	-0.1034
	(0.0572)	(0.0138)	(0.0068)	(0.0273)	(0.0573)	(0.0144)	(0.0294)
9	-0.0994	-0.0180	0.0258	0.0451	-0.0087	0.0415	-0.1216
	(0.0566)	(0.0132)	(0.0069)	(0.0265)	(0.0546)	(0.0142)	(0.0289)
10-largest	-0.0212	-0.0085	0.0227	-0.0644	0.0159	0.0511	-0.1034
	(0.0455)	(0.0123)	(0.0080)	(0.0254)	(0.0451)	(0.0180)	(0.0300)
Panel b: size-based portfolios of NYSE stocks							
decile	δ	μ_D	μ_N	Π_{11}	Π_{12}	Π_{21}	Π_{22}
1-smallest	-0.2438	-0.0033	-0.0075	0.1786	-0.0444	0.0882	-0.0560
	(0.0885)	(0.0076)	(0.0045)	(0.0316)	(0.0909)	(0.0290)	(0.0529)
2	-0.1716	-0.0148	0.0063	0.0611	-0.0008	0.0703	-0.0828
	(0.0834)	(0.0093)	(0.0048)	(0.0370)	(0.0917)	(0.0266)	(0.0545)
3	-0.1313	-0.0088	0.0048	0.0108	0.0109	0.0548	-0.0858
	(0.0804)	(0.0106)	(0.0053)	(0.0355)	(0.0869)	(0.0248)	(0.0529)
4	-0.1096	-0.0151	0.0114	0.0145	0.0189	0.0312	-0.0963
	(0.0704)	(0.0121)	(0.0059)	(0.0322)	(0.0742)	(0.0196)	(0.0421)
5	-0.0804	-0.0014	0.0097	0.0043	0.0066	0.0293	-0.1097
	(0.0705)	(0.0131)	(0.0063)	(0.0330)	(0.0738)	(0.0179)	(0.0390)
6	-0.0991	-0.0011	0.0105	0.0079	0.0313	0.0215	-0.1076
	(0.0669)	(0.0133)	(0.0066)	(0.0324)	(0.0704)	(0.0175)	(0.0353)
7	-0.1285	0.0000	0.0115	0.0249	0.0209	0.0275	-0.1146
	(0.0626)	(0.0127)	(0.0064)	(0.0296)	(0.0637)	(0.0159)	(0.0328)
8	-0.1307	0.0019	0.0150	0.0275	0.0079	0.0328	-0.1180
	(0.0636)	(0.0124)	(0.0066)	(0.0300)	(0.0610)	(0.0168)	(0.0328)
9	-0.0899	-0.0007	0.0213	0.0117	-0.0148	0.0393	-0.1044
	(0.0593)	(0.0125)	(0.0071)	(0.0278)	(0.0576)	(0.0176)	(0.0355)
10-largest	-0.0284	0.0016	0.0132	-0.0716	0.0099	0.0479	-0.0901
	(0.0464)	(0.0117)	(0.0078)	(0.0263)	(0.0468)	(0.0206)	(0.0322)
Panel c: size-based portfolios of NASDAQ stocks							
decile	δ	μ_D	μ_N	Π_{11}	Π_{12}	Π_{21}	Π_{22}
1-smallest	0.1363	0.0101	-0.2104	0.2371	-0.0268	-0.0405	0.3029
	(0.0182)	(0.0101)	(0.0093)	(0.0167)	(0.0162)	(0.0169)	(0.0171)
2	0.0298	-0.0037	-0.1089	0.2701	-0.0266	-0.0029	0.2315
	(0.0252)	(0.0080)	(0.0063)	(0.0194)	(0.0213)	(0.0150)	(0.0194)
3	0.0169	-0.0061	-0.0663	0.2841	-0.0303	0.0262	0.1572
	(0.0327)	(0.0080)	(0.0054)	(0.0220)	(0.0296)	(0.0148)	(0.0261)
4	0.0207	-0.0294	-0.0258	0.2432	-0.0571	0.0384	0.1115
	(0.0393)	(0.0087)	(0.0054)	(0.0232)	(0.0351)	(0.0144)	(0.0272)
5	-0.0421	-0.0412	0.0045	0.1158	-0.0428	0.0122	0.0130
	(0.0431)	(0.0115)	(0.0063)	(0.0255)	(0.0450)	(0.0139)	(0.0261)
6	-0.0456	-0.0528	0.0257	0.0197	-0.0236	0.0054	-0.0470
	(0.0494)	(0.0142)	(0.0071)	(0.0283)	(0.0501)	(0.0133)	(0.0239)
7	-0.0453	-0.0539	0.0265	0.0033	-0.0244	0.0010	-0.0574
	(0.0504)	(0.0152)	(0.0076)	(0.0262)	(0.0498)	(0.0125)	(0.0231)
8	-0.0508	-0.0549	0.0393	0.0239	-0.0226	0.0093	-0.0752
	(0.0467)	(0.0156)	(0.0077)	(0.0237)	(0.0471)	(0.0117)	(0.0231)
9	0.0281	-0.0505	0.0390	0.0467	-0.0438	0.0245	-0.0896
	(0.0557)	(0.0161)	(0.0081)	(0.0237)	(0.0471)	(0.0121)	(0.0226)
10-largest	0.0523	-0.0498	0.0610	-0.0444	0.0513	0.0500	-0.1114
	(0.0416)	(0.0175)	(0.0105)	(0.0251)	(0.0385)	(0.0129)	(0.0236)

Estimates of mean equation parameters for 10 size-based portfolios, with standard errors in parentheses. The mean equation is specified as $\begin{pmatrix} 1 & \delta \\ 0 & 1 \end{pmatrix} \begin{pmatrix} r_t^D \\ r_t^N \end{pmatrix} = \begin{pmatrix} \mu_D \\ \mu_N \end{pmatrix} + \begin{pmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{pmatrix} \begin{pmatrix} r_{t-1}^D \\ r_{t-1}^N \end{pmatrix} + \begin{pmatrix} u_t^D \\ u_t^N \end{pmatrix}$.

Table A.9: Estimates of the mean equations with the size-sorted portfolios

	size 1-small	size 2	size 3	size 4	size 5	size 6	size 7	size 8	size 9	size 10-large
β_D	0.9180	0.8998	0.9096	0.9305	0.9265	0.9323	0.9449	0.9460	0.9523	0.9584
	(0.0128)	(0.0122)	(0.0103)	(0.0086)	(0.0088)	(0.0084)	(0.0065)	(0.0062)	(0.0054)	(0.0047)
γ_D	0.0396	0.0483	0.0558	0.0472	0.0493	0.0419	0.0388	0.0378	0.0336	0.0261
	(0.0049)	(0.0049)	(0.0051)	(0.0047)	(0.0046)	(0.0041)	(0.0038)	(0.0039)	(0.0038)	(0.0034)
ρ_D	0.0241	0.0460	0.0454	0.0429	0.0391	0.0350	0.0288	0.0307	0.0317	0.0344
	(0.0042)	(0.0051)	(0.0052)	(0.0048)	(0.0050)	(0.0048)	(0.0044)	(0.0045)	(0.0043)	(0.0041)
γ_D^*	-0.0158	-0.0225	-0.0252	-0.0238	-0.0251	-0.0255	-0.0259	-0.0271	-0.0299	-0.0344
	(0.0026)	(0.0030)	(0.0031)	(0.0028)	(0.0029)	(0.0029)	(0.0026)	(0.0026)	(0.0026)	(0.0028)
ρ_D^*	-0.0173	-0.0231	-0.0209	-0.0185	-0.0203	-0.0209	-0.0215	-0.0258	-0.0274	-0.0279
	(0.0029)	(0.0032)	(0.0029)	(0.0026)	(0.0028)	(0.0027)	(0.0026)	(0.0028)	(0.0028)	(0.0029)
ν_D	13.7745	11.4980	10.2268	9.2325	10.5935	12.8001	12.8073	12.6206	11.3425	9.0860
	(1.9787)	(1.3791)	(1.0784)	(0.8894)	(1.1528)	(1.6057)	(1.5659)	(1.4768)	(1.2167)	(0.8980)
ω_D	-0.7088	-0.8519	-0.7678	-0.6147	-0.4630	-0.3592	-0.3110	-0.3126	-0.3897	-0.4744
	(0.0205)	(0.0208)	(0.0223)	(0.0236)	(0.0231)	(0.0222)	(0.0233)	(0.0233)	(0.0238)	(0.0242)
β_N	0.9404	0.9221	0.9318	0.9403	0.9353	0.9451	0.9476	0.9448	0.9519	0.9603
	(0.0067)	(0.0083)	(0.0069)	(0.0065)	(0.0070)	(0.0062)	(0.0061)	(0.0063)	(0.0055)	(0.0046)
γ_N	0.0324	0.0478	0.0474	0.0531	0.0529	0.0427	0.0410	0.0419	0.0425	0.0349
	(0.0037)	(0.0045)	(0.0045)	(0.0049)	(0.0055)	(0.0051)	(0.0052)	(0.0055)	(0.0053)	(0.0046)
ρ_N	0.0367	0.0438	0.0417	0.0453	0.0464	0.0443	0.0423	0.0480	0.0439	0.0324
	(0.0041)	(0.0046)	(0.0045)	(0.0048)	(0.0047)	(0.0045)	(0.0044)	(0.0047)	(0.0045)	(0.0039)
γ_N^*	-0.0230	-0.0313	-0.0274	-0.0254	-0.0264	-0.0245	-0.0262	-0.0296	-0.0291	-0.0350
	(0.0025)	(0.0032)	(0.0028)	(0.0027)	(0.0030)	(0.0028)	(0.0029)	(0.0033)	(0.0031)	(0.0032)
ρ_N^*	-0.0140	-0.0166	-0.0224	-0.0228	-0.0244	-0.0254	-0.0292	-0.0284	-0.0309	-0.0328
	(0.0022)	(0.0027)	(0.0027)	(0.0027)	(0.0029)	(0.0028)	(0.0029)	(0.0030)	(0.0031)	(0.0029)
ν_N	12.2527	8.1456	6.8208	5.5303	5.0590	4.8222	4.5947	4.2923	4.3768	5.1477
	(1.4976)	(0.6646)	(0.4786)	(0.3448)	(0.2903)	(0.2725)	(0.2512)	(0.2270)	(0.2277)	(0.3029)
ω_N	-0.8920	-1.2634	-1.3574	-1.3540	-1.3175	-1.2754	-1.2702	-1.3012	-1.3293	-1.0885
	(0.0231)	(0.0230)	(0.0240)	(0.0274)	(0.0264)	(0.0267)	(0.0268)	(0.0271)	(0.0286)	(0.0275)

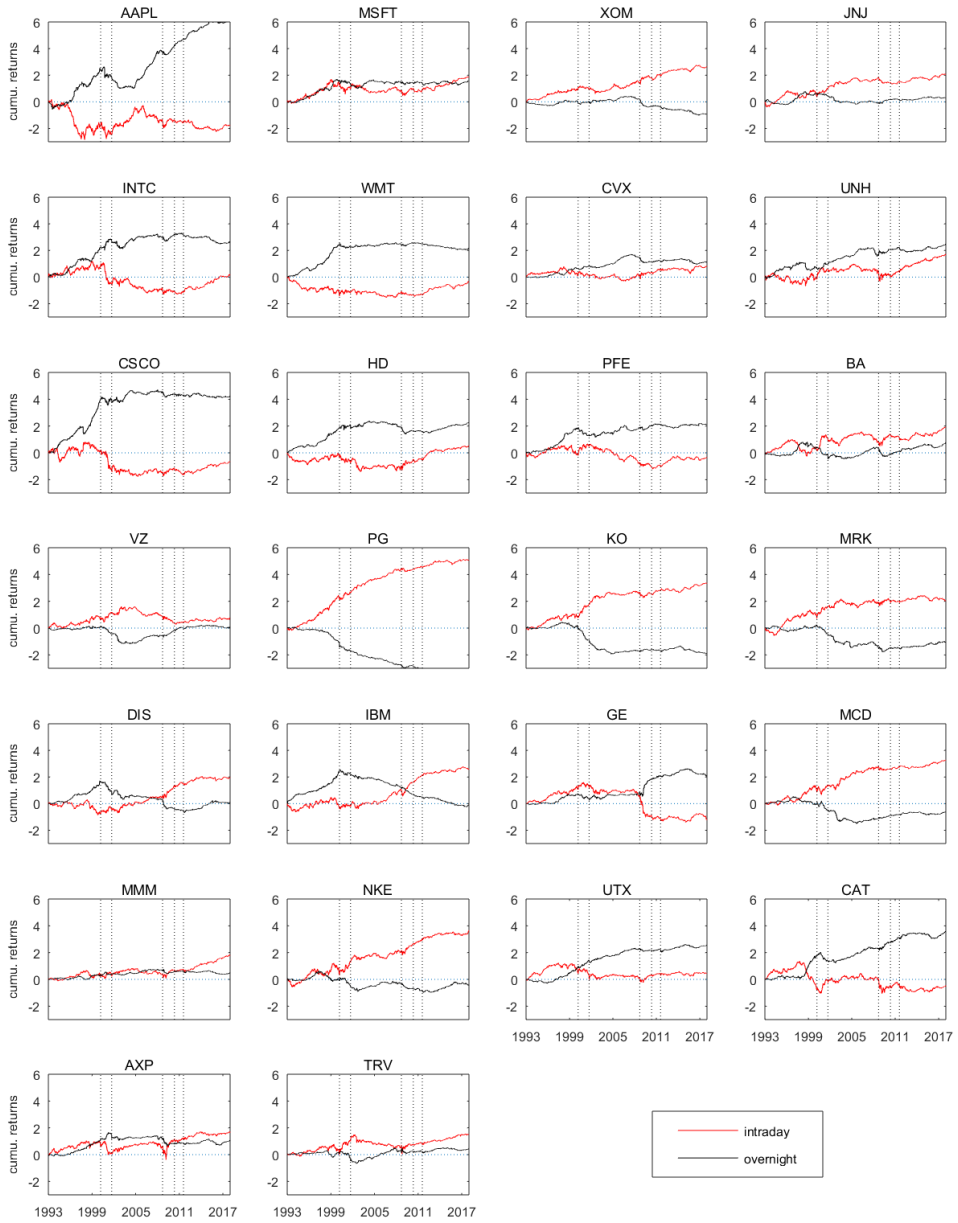
Estimates of the dynamic parameters for size-based portfolios with standard errors in parentheses with the univariate model.

Table A.10: Estimates of the dynamic parameters in the univariate model with the size-sorted portfolios

	size 1-small	size 2	size 3	size 4	size 5	size 6	size 7	size 8	size 9	size 10-large
β_D	0.9215	0.9022	0.9114	0.9316	0.9270	0.9328	0.9455	0.9469	0.9537	0.9609
	(0.0122)	(0.0119)	(0.0101)	(0.0085)	(0.0087)	(0.0084)	(0.0064)	(0.0061)	(0.0052)	(0.0044)
γ_D	0.0394	0.0483	0.0557	0.0471	0.0493	0.0419	0.0388	0.0377	0.0336	0.0258
	(0.0049)	(0.0049)	(0.0051)	(0.0047)	(0.0046)	(0.0041)	(0.0038)	(0.0039)	(0.0038)	(0.0034)
ρ_D	0.0240	0.0461	0.0455	0.0430	0.0391	0.0350	0.0288	0.0307	0.0318	0.0341
	(0.0042)	(0.0051)	(0.0052)	(0.0048)	(0.0050)	(0.0048)	(0.0044)	(0.0045)	(0.0043)	(0.0040)
γ_D^*	-0.0154	-0.0221	-0.0252	-0.0239	-0.0252	-0.0256	-0.0259	-0.0271	-0.0299	-0.0343
	(0.0026)	(0.0030)	(0.0031)	(0.0028)	(0.0029)	(0.0029)	(0.0026)	(0.0026)	(0.0025)	(0.0027)
ρ_D^*	-0.0182	-0.0235	-0.0206	-0.0182	-0.0201	-0.0207	-0.0214	-0.0257	-0.0272	-0.0278
	(0.0030)	(0.0032)	(0.0029)	(0.0026)	(0.0028)	(0.0027)	(0.0026)	(0.0028)	(0.0028)	(0.0029)
ν_D	13.7437	11.5287	10.2093	9.2259	10.5870	12.7914	12.7853	12.6009	11.3025	9.0657
	(1.9742)	(1.3868)	(1.0744)	(0.8878)	(1.1508)	(1.6034)	(1.5602)	(1.4722)	(1.2080)	(0.8907)
ω_D	-0.7076	-0.8518	-0.7684	-0.6149	-0.4631	-0.3594	-0.3115	-0.3132	-0.3913	-0.4784
	(0.0208)	(0.0210)	(0.0225)	(0.0238)	(0.0232)	(0.0222)	(0.0234)	(0.0235)	(0.0242)	(0.0248)
β_N	0.9436	0.9239	0.9334	0.9413	0.9360	0.9456	0.9485	0.9459	0.9536	0.9626
	(0.0063)	(0.0080)	(0.0067)	(0.0064)	(0.0070)	(0.0061)	(0.0059)	(0.0062)	(0.0054)	(0.0043)
γ_N	0.0333	0.0483	0.0476	0.0533	0.0530	0.0428	0.0410	0.0420	0.0426	0.0347
	(0.0037)	(0.0045)	(0.0045)	(0.0049)	(0.0055)	(0.0051)	(0.0052)	(0.0055)	(0.0053)	(0.0046)
ρ_N	0.0362	0.0439	0.0419	0.0453	0.0464	0.0443	0.0423	0.0479	0.0438	0.0320
	(0.0040)	(0.0045)	(0.0045)	(0.0047)	(0.0047)	(0.0045)	(0.0044)	(0.0047)	(0.0044)	(0.0038)
γ_N^*	-0.0249	-0.0332	-0.0280	-0.0255	-0.0263	-0.0245	-0.0261	-0.0295	-0.0288	-0.0347
	(0.0026)	(0.0033)	(0.0028)	(0.0027)	(0.0030)	(0.0028)	(0.0029)	(0.0033)	(0.0031)	(0.0031)
ρ_N^*	-0.0126	-0.0148	-0.0215	-0.0226	-0.0243	-0.0254	-0.0292	-0.0283	-0.0309	-0.0331
	(0.0022)	(0.0027)	(0.0027)	(0.0027)	(0.0029)	(0.0028)	(0.0029)	(0.0030)	(0.0031)	(0.0029)
ν_N	11.9461	8.0501	6.7869	5.5198	5.0562	4.8200	4.5925	4.2899	4.3730	5.1354
	(1.4158)	(0.6481)	(0.4737)	(0.3433)	(0.2900)	(0.2723)	(0.2510)	(0.2268)	(0.2275)	(0.3018)
ω_N	-0.8904	-1.2640	-1.3589	-1.3549	-1.3178	-1.2759	-1.2709	-1.3019	-1.3309	-1.0932
	(0.0238)	(0.0233)	(0.0243)	(0.0277)	(0.0265)	(0.0269)	(0.0270)	(0.0274)	(0.0292)	(0.0284)

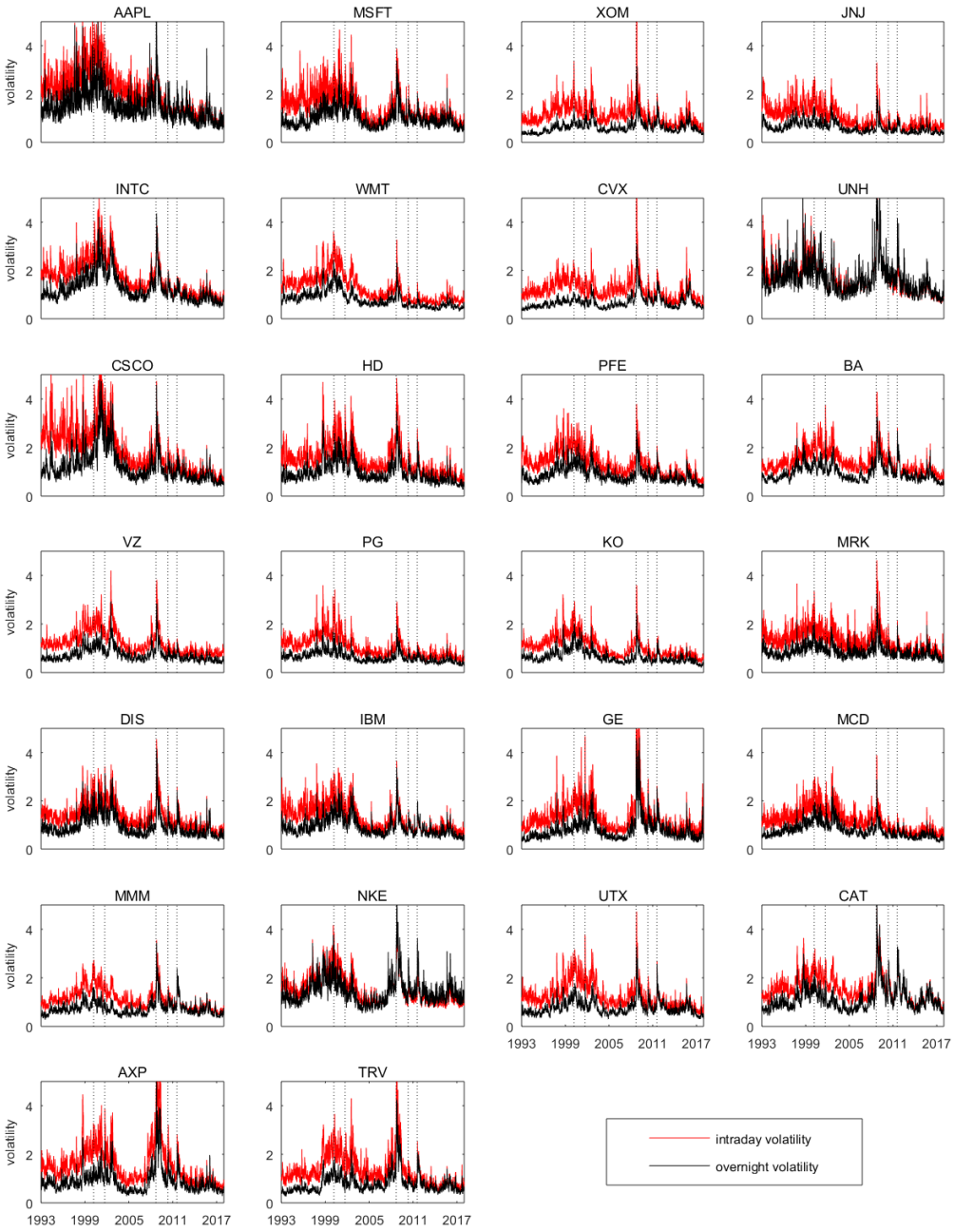
Estimates of the dynamic parameters for size portfolios with standard errors in parentheses with the multivariate model.

Table A.11: Estimates of the dynamic parameters in the multivariate model with the size-sorted portfolios



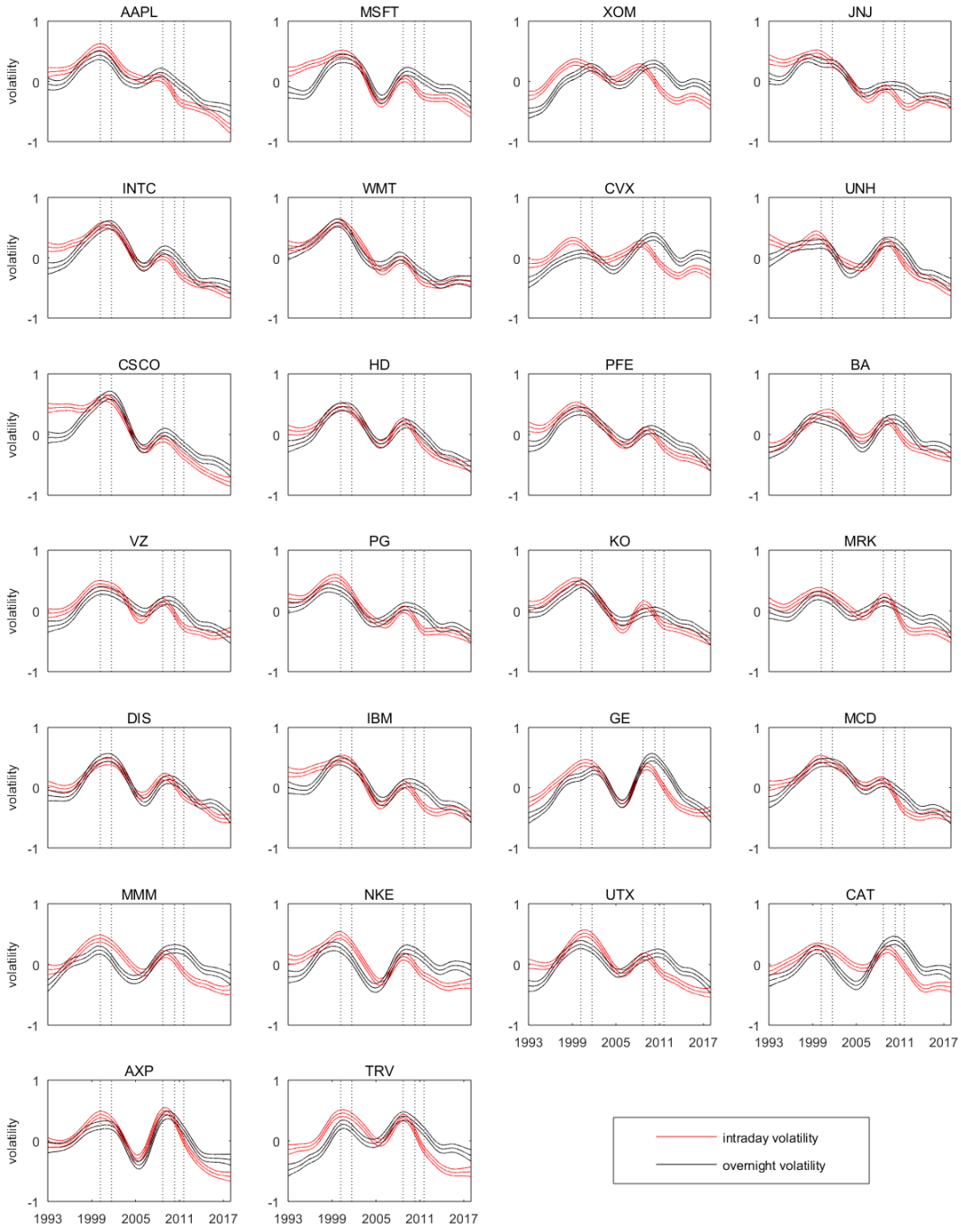
This figure shows the cumulative intraday (in red) and cumulative overnight (in black) returns with one subplot for each stock.

Figure A.1: Cumulative intraday and overnight returns of Dow Jones stocks



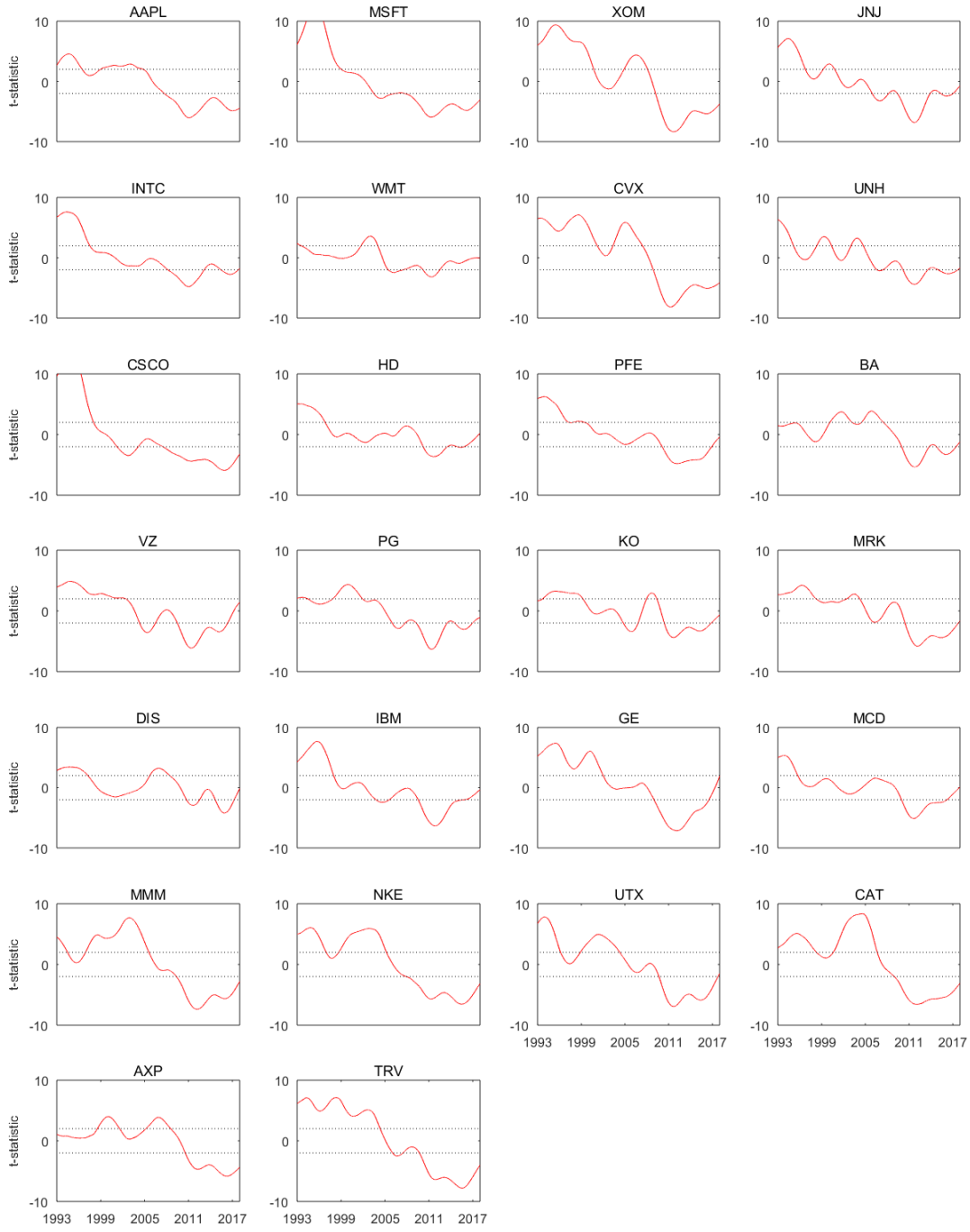
This figure shows the estimated intraday (in red) and overnight (in black) volatilities, $\sqrt{\frac{\nu_j}{\nu_j-2}exp(2\lambda_t^j + 2\sigma^j(\frac{t}{T}))}$, based on the univariate coupled component model, with one subplot for each stock. The five dashed vertical lines from left to right represent the dates: 10 March 2000 (dot-com bubble), 17 September 2001 (the September 11 attacks), 16 September 2008 (financial crisis), 6 May 2010 (flash crash) and 1 August 2011 (August 2011 stock markets fall), respectively.

Figure A.2: Intraday and overnight volatilities of Dow Jones stocks



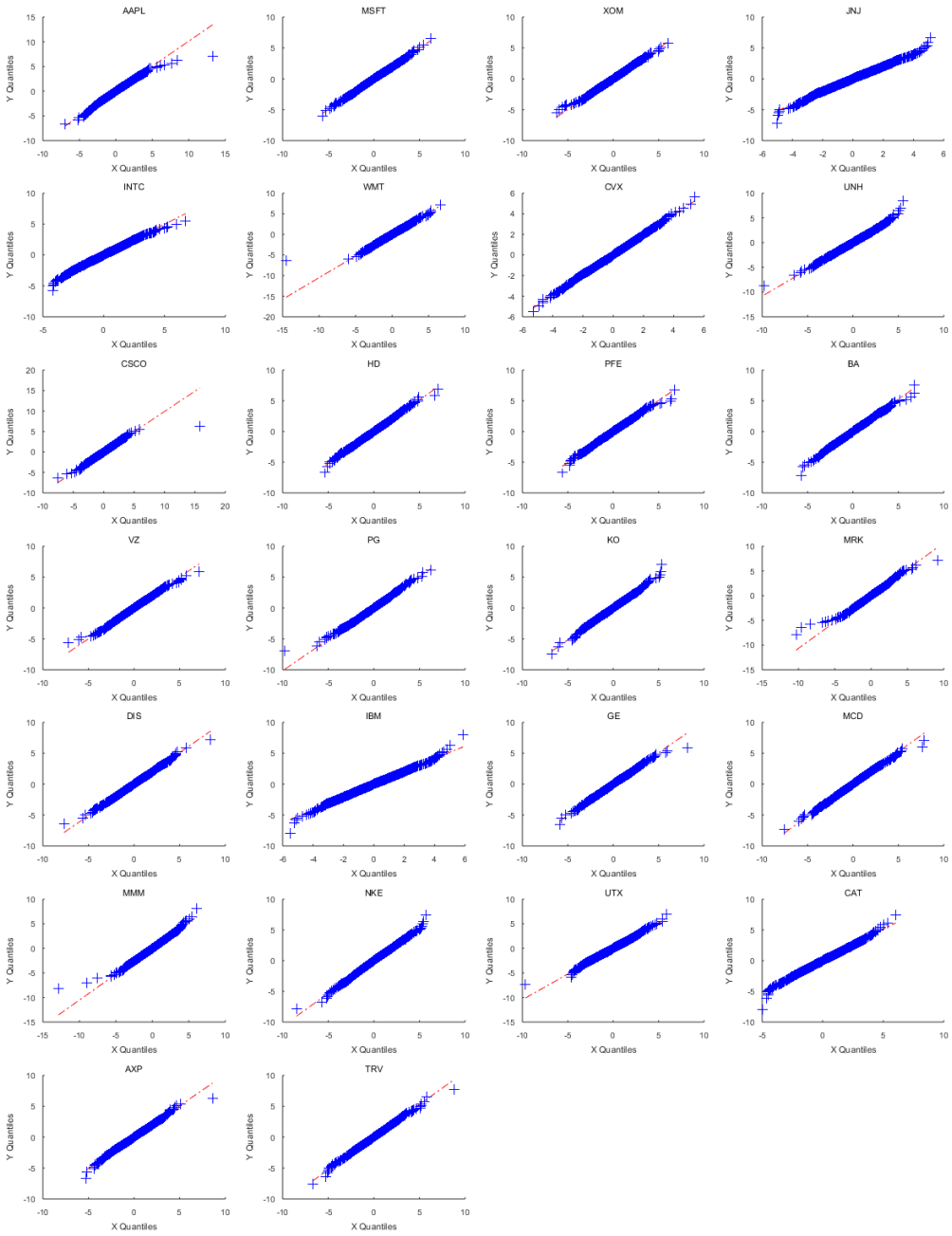
This figure shows the estimated intraday (in red) and overnight (in black) long run components,$\sigma^D(t/T)$ and $\sigma^N(t/T)$, based on the univariate coupled component model, with one subplot for each stock. The five dashed vertical lines from left to right represent the dates: 10 March 2000(dot-com bubble), 17 September 2011(after the September 11 attacks), 16 September 2008(financial crisis), 6 May 2010 (flash crash) and 1 August 2011 (August 2011 stock markets fall), respectively.

Figure A.3: Long run component $\sigma^j(\frac{t}{T})$ in the univariate model with Dow Jones stocks



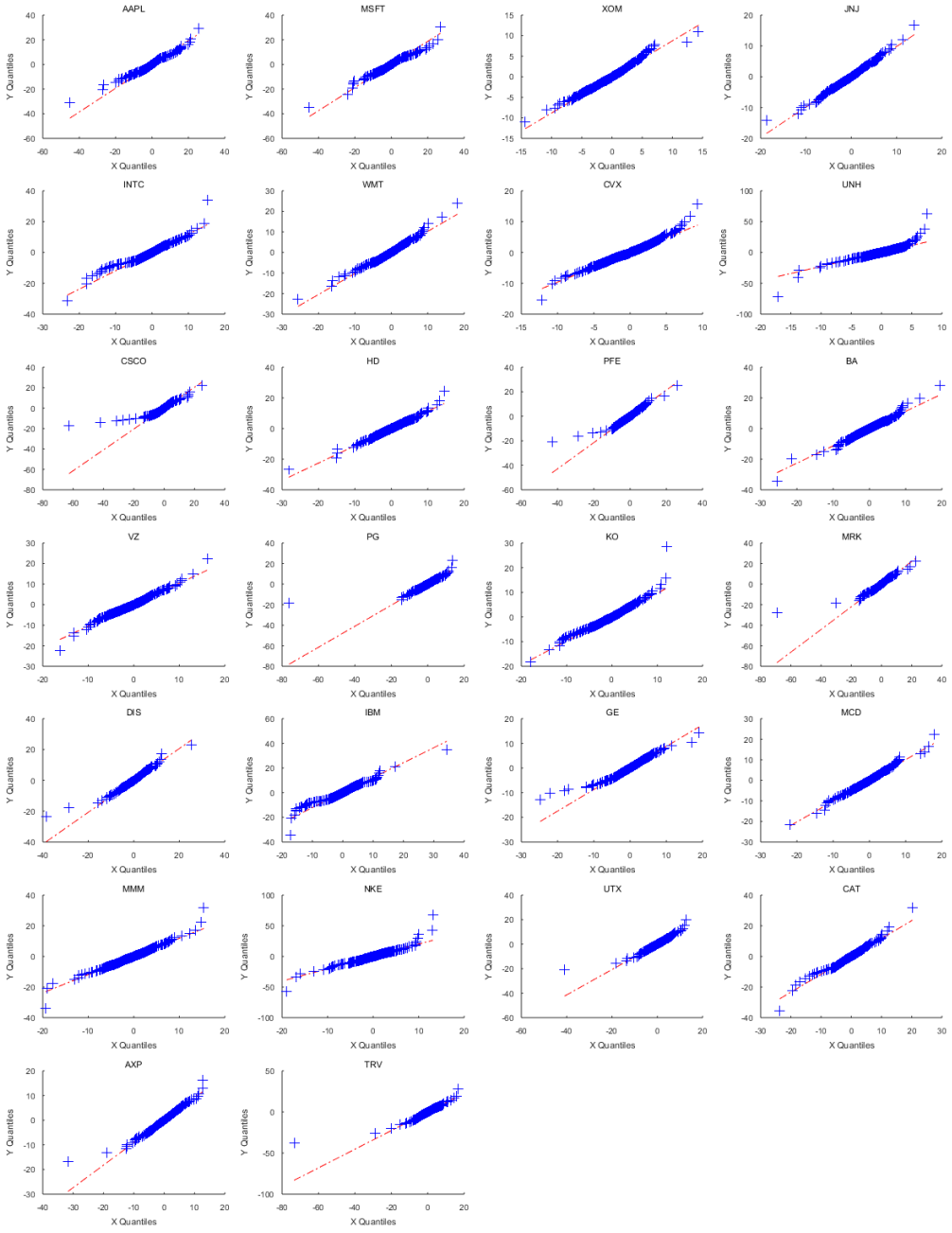
The red lines represent the statistics of the ratio tests, with the null hypothesis $H_0 : \exp\left(\sigma_0^N(t/T)\right) = \rho \exp\left(\sigma_0^D(t/T)\right)$. The black lines indicate the 95% confidence intervals of the statistics under the null. The five dashed vertical lines from left to right represent the dates: 10 March 2000(dot-com bubble), 17 September 2001 (after the September 11 attacks), 16 September 2008(financial crisis), 6 May 2010 (flash crash) and 1 August 2011 (August 2011 stock markets fall), respectively.

Figure A.4: Statistics of ratio tests in the univariate model with Dow Jones stocks



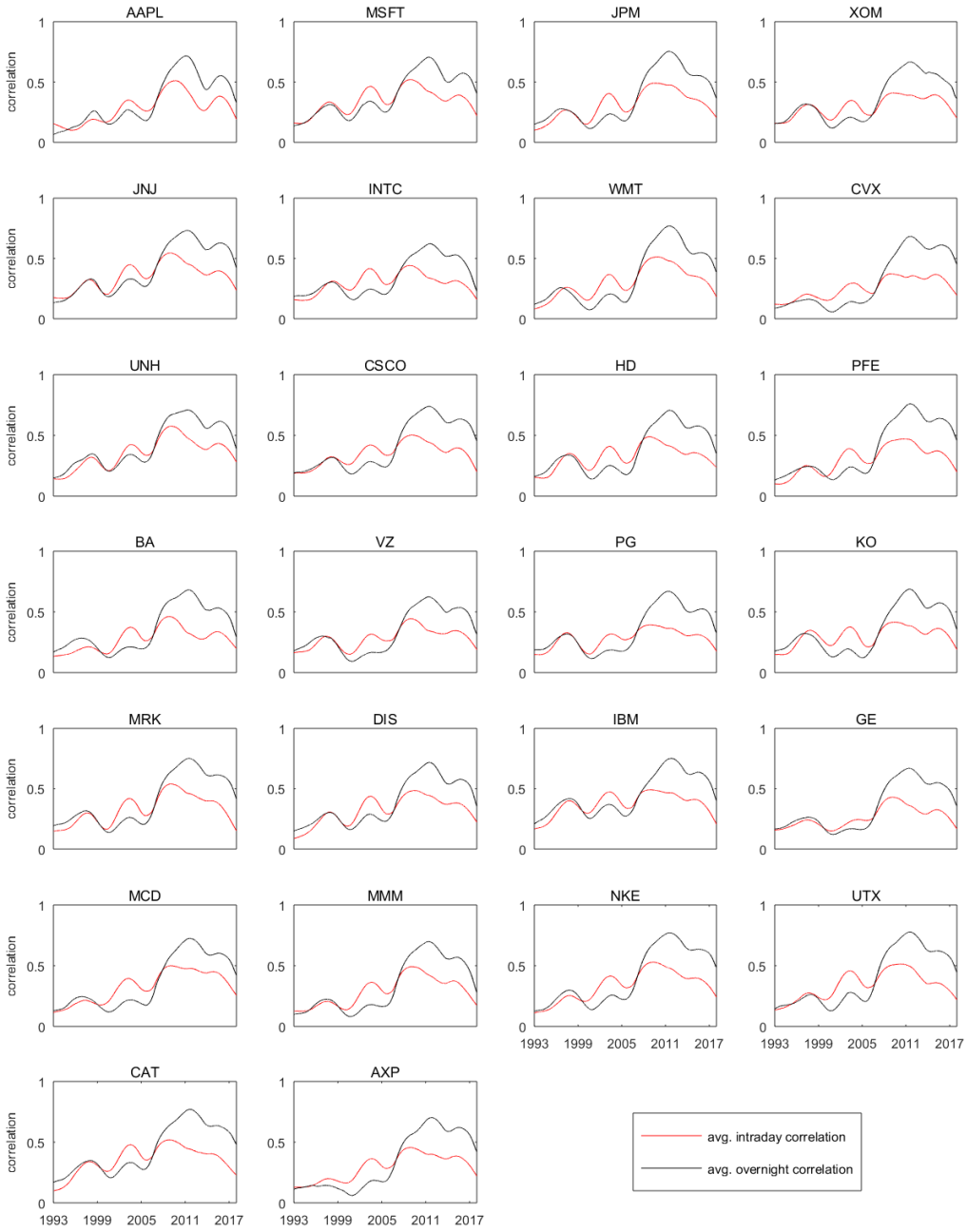
This figure displays Q-Q plots of the quantiles of the intraday innovations (X axis), versus the theoretical quantiles of the student t distribution with the $\hat{\nu}_D$ degrees of freedom (Y axis), with one panel for each stock.

Figure A.5: QQ plot of the intraday innovations in the univariate model with Dow Jones stocks



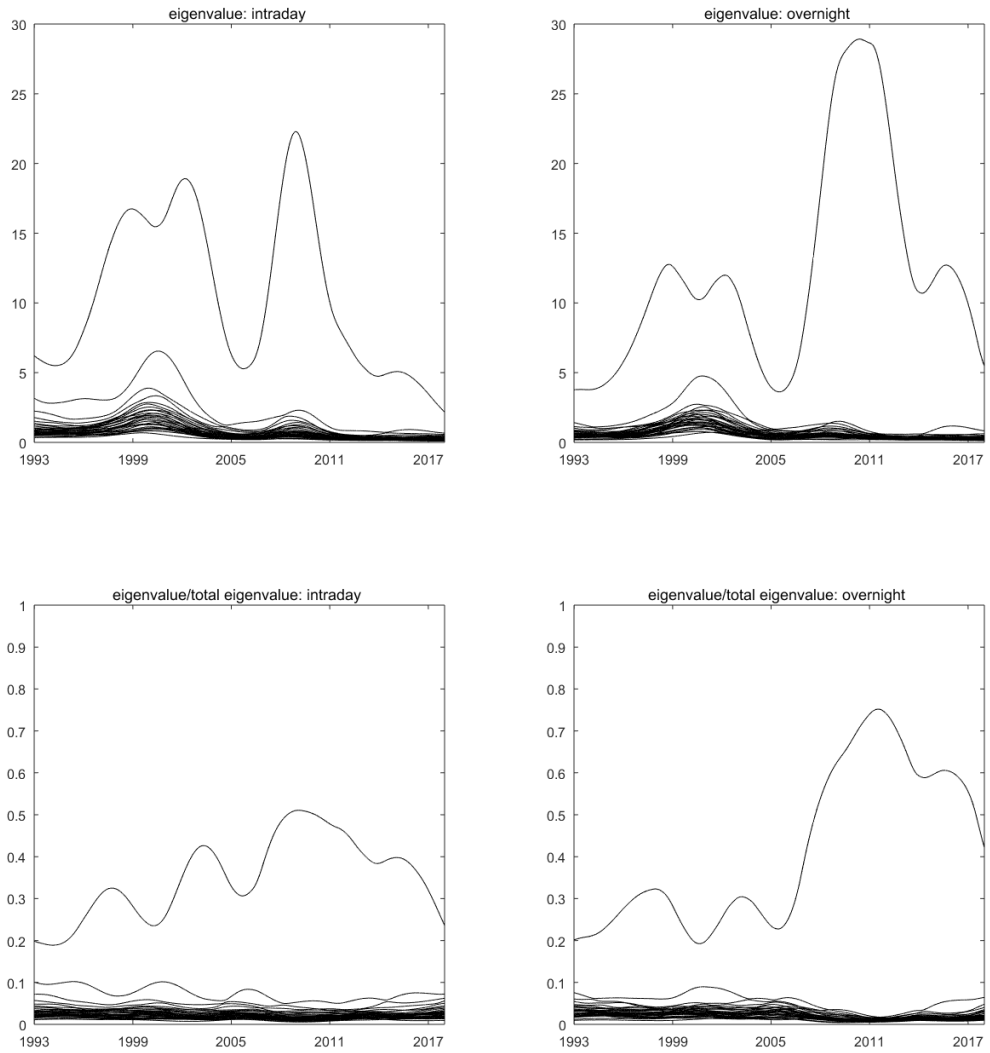
This figure displays Q-Q plots of the quantiles of the overnight innovations (X axis), versus the theoretical quantiles of the student t distribution with the $\hat{\nu}_D$ degrees of freedom (Y axis), with one panel for each stock.

Figure A.6: QQ plot of the overnight innovations in the univariate model with Dow Jones stocks



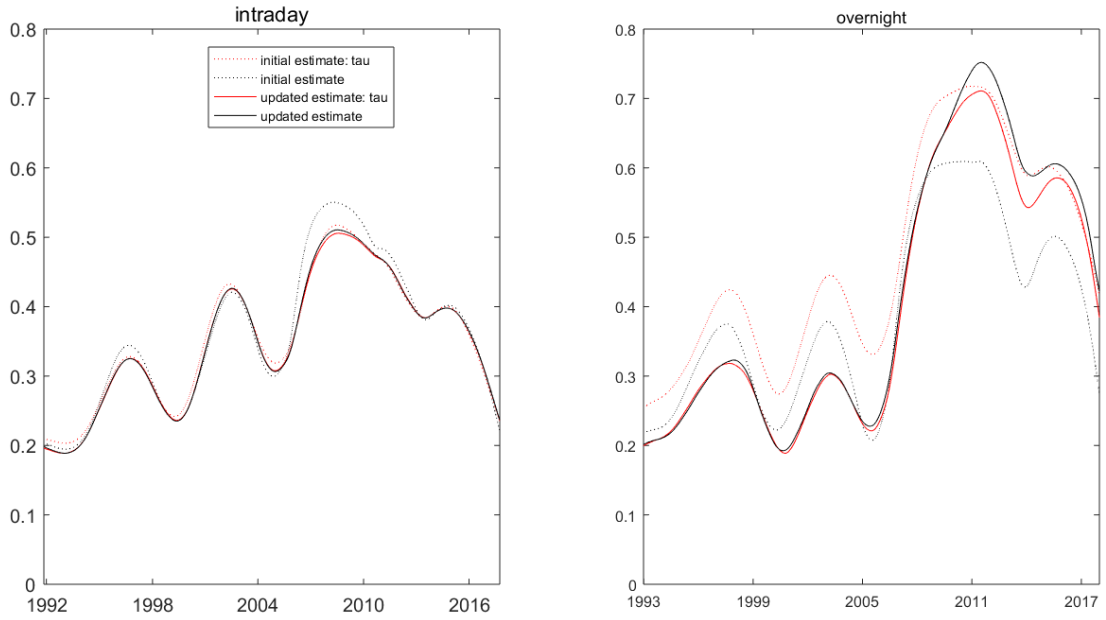
Each panel presents the average long run intraday(overnight) correlations between that individual stock and the remaining stocks, as implied by the multivariate coupled component model.

Figure A.7: Long run correlations between Dow Jones stocks



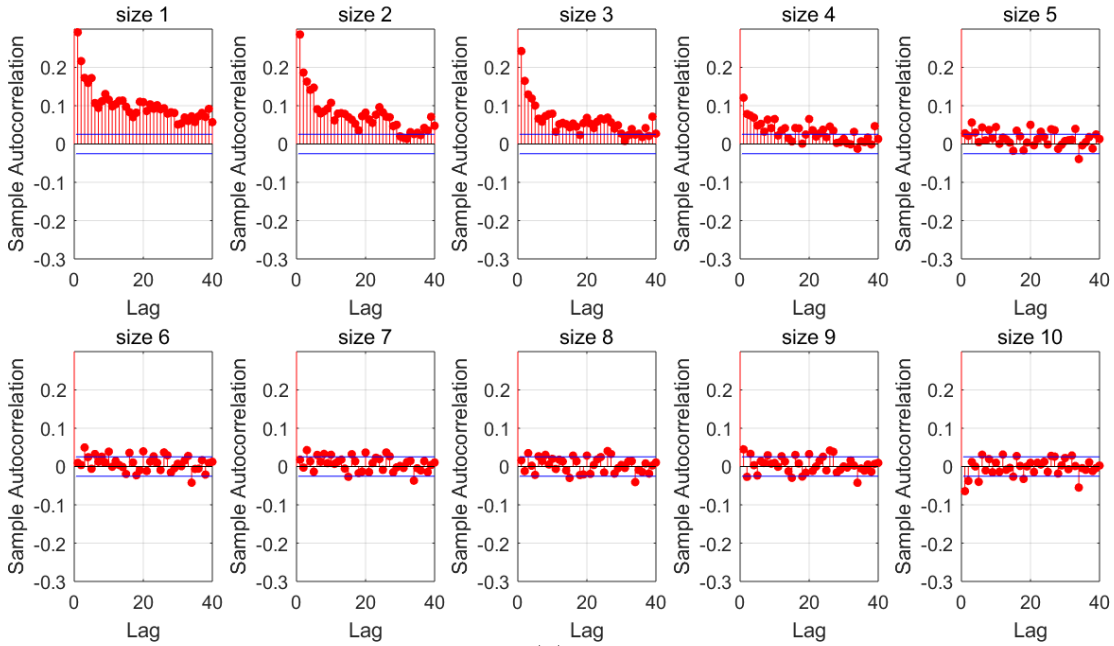
The figure plots the eigenvalues and the scaled eigenvalues (the eigenvalue divided by the sum of all eigenvalues) of the covariance matrix.

Figure A.8: Eigenvalues of correlation matrices for Dow Jones stocks

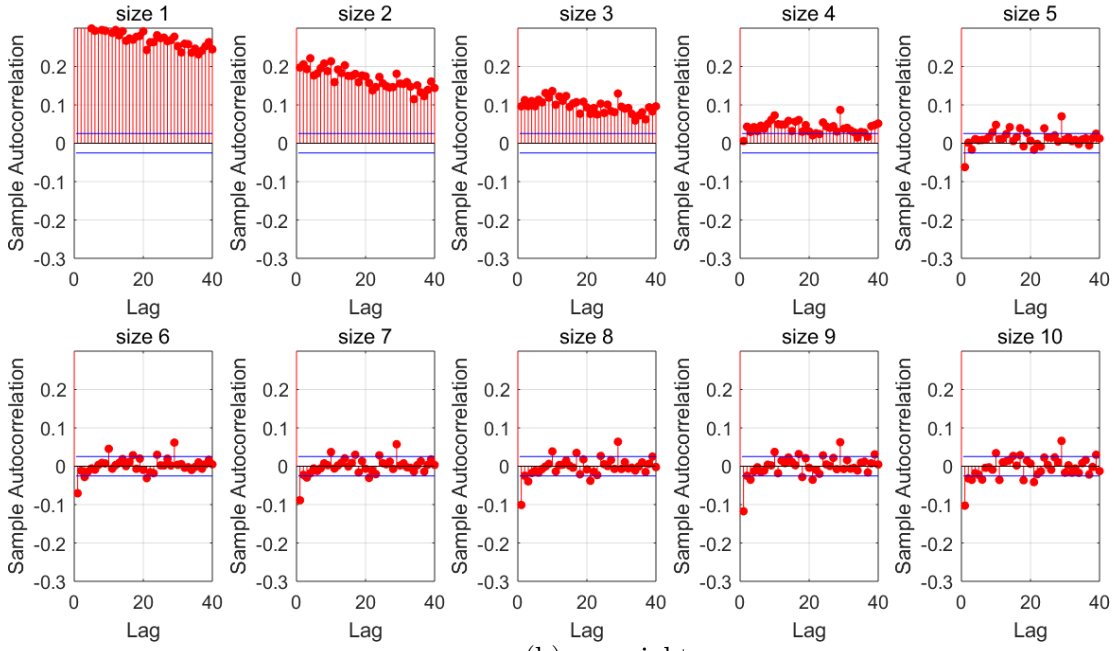


The figure plots the largest scaled eigenvalue (the eigenvalue divided by the sum of all eigenvalues) of the estimated intraday and overnight covariance matrix. Red (black) lines are used to indicate the use of the non-robust (robust) correlation matrix in the initial step. Solid (dashed) lines are further used to indicate the updated (initial) estimators,

Figure A.9: Comparison of robust and non-robust initial correlation estimator for Dow Jones stocks



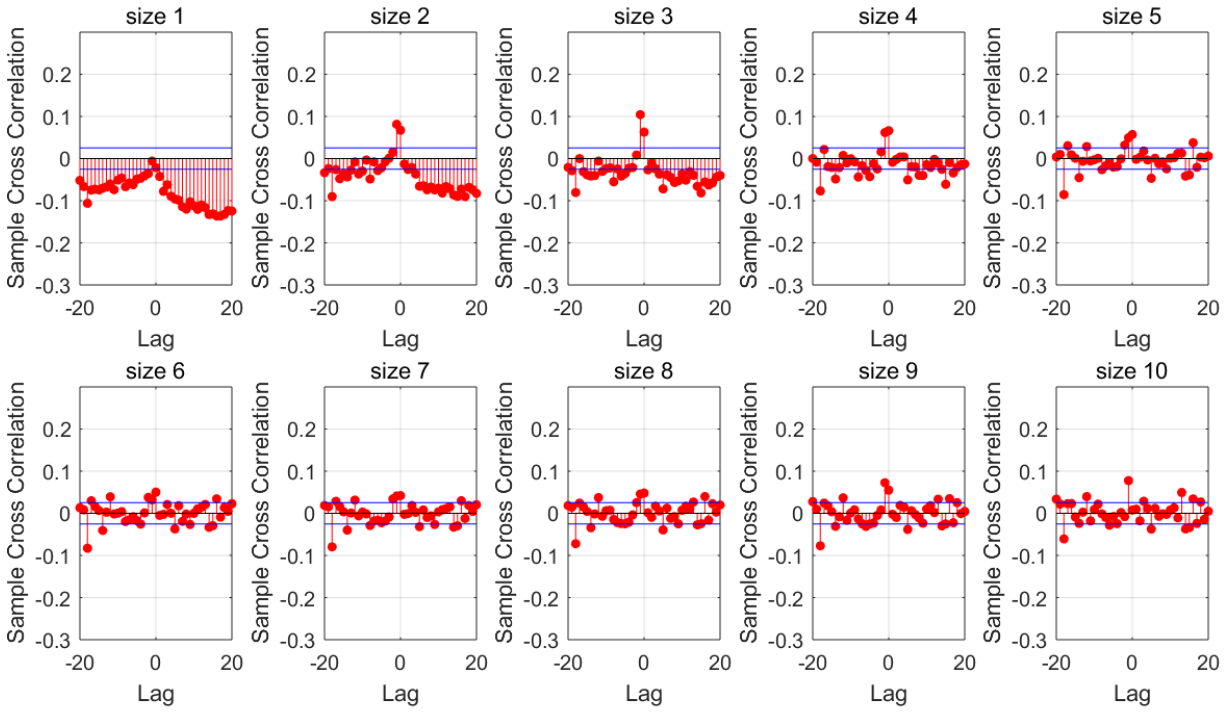
(a) intraday



(b) overnight

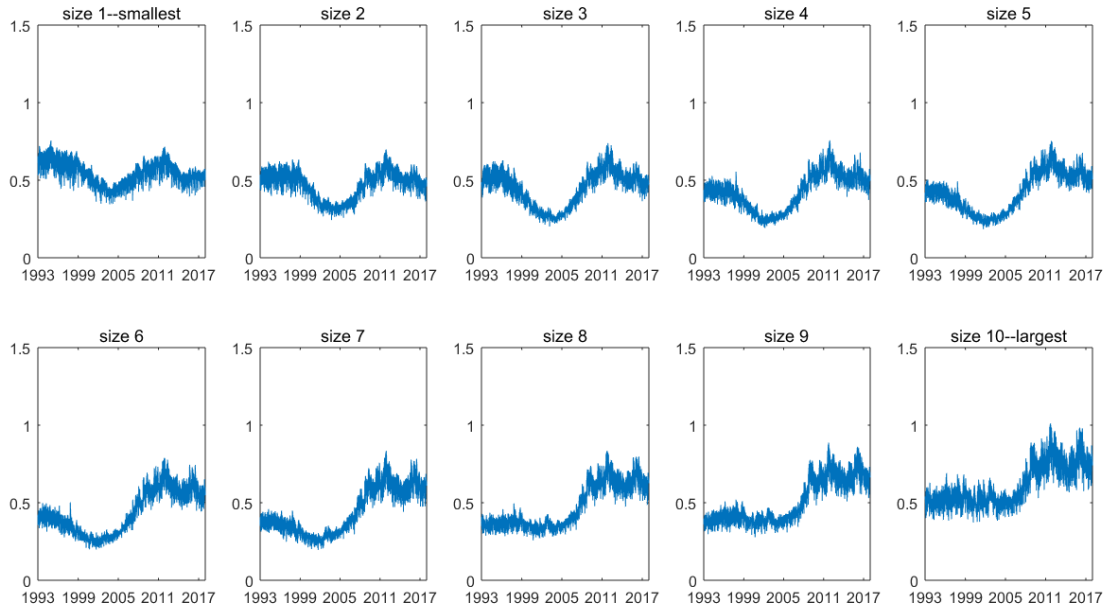
This figure plots the autocorrelations of intraday (overnight) returns for the size-based portfolios on NYSE/AMEX/NASDAQ stocks in Panel a (Panel b).

Figure A.10: Autocorrelations of intraday and overnight returns for the size-based portfolios: NYSE/AMEX/NASDAQ

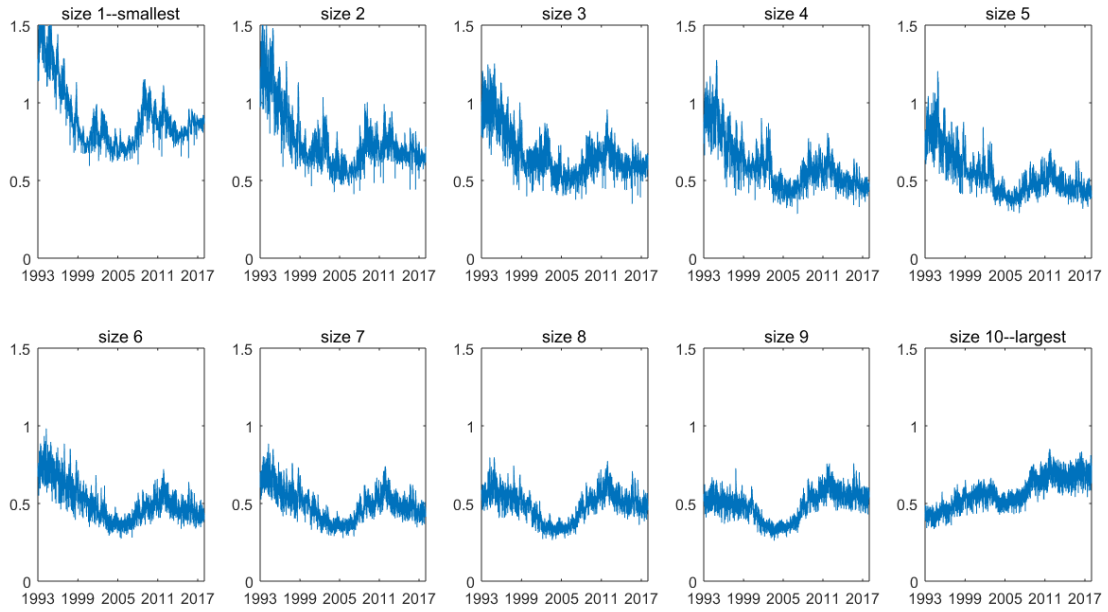


Cross correlations between overnight and intraday returns, r_t^N and r_{t+Lag}^D, for the size-based portfolios in NYSE/AMEX/NASDAQ

Figure A.11: Cross correlations between overnight and intraday returns for the size-based portfolios: NYSE/AMEX/NASDAQ



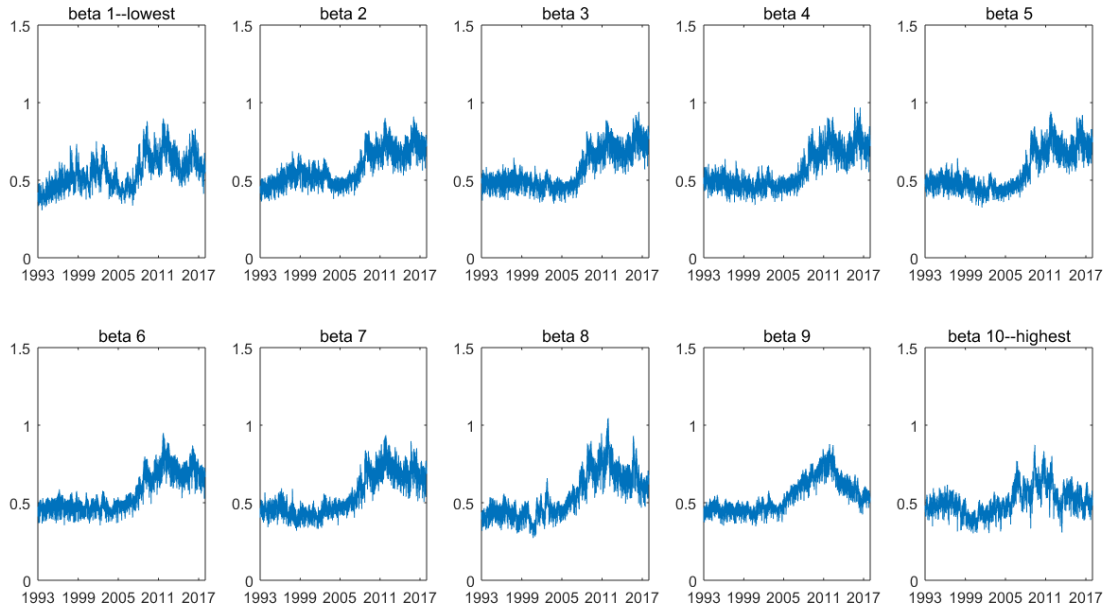
(a) NYSE



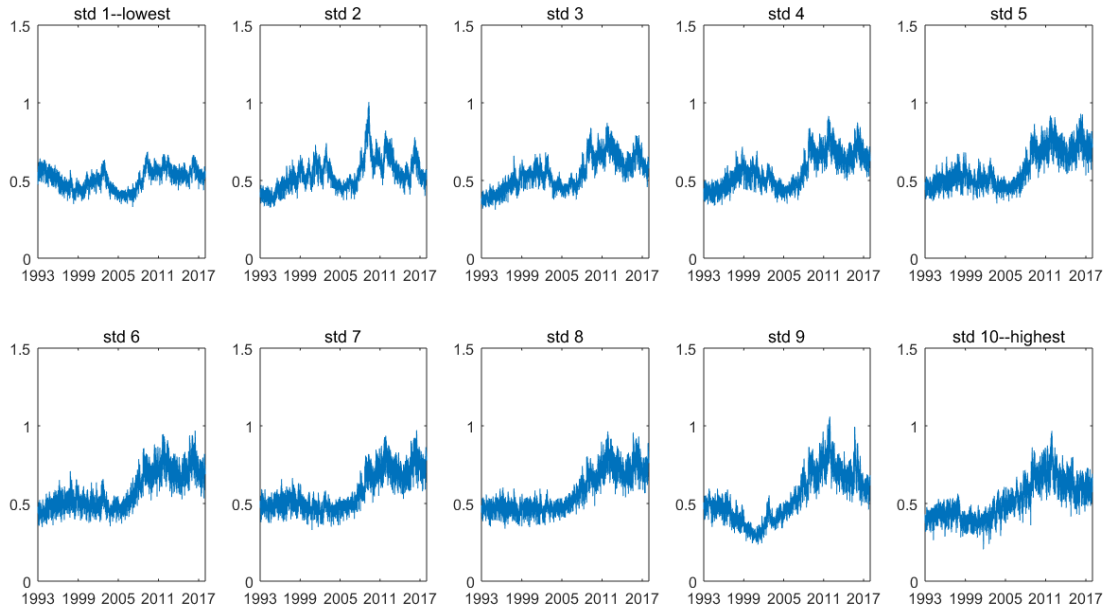
(b) NASDAQ

Panel a (Panel b) plots the ratio of overnight to intraday volatility for size-based portfolios for NYSE (NASDAQ) stocks: decile 1 with the smallest market capitalizations and decile 10 with the largest market capitalizations. Intraday and overnight volatilities are defined as $\sqrt{\frac{\nu_j}{\nu_j-2}exp(2\lambda_t^j + 2\sigma^j(\frac{t}{T}))}$, for $j = D, N$, respectively.

Figure A.12: Ratio of overnight to intraday volatility for the size-based portfolios: NYSE and NASDAQ



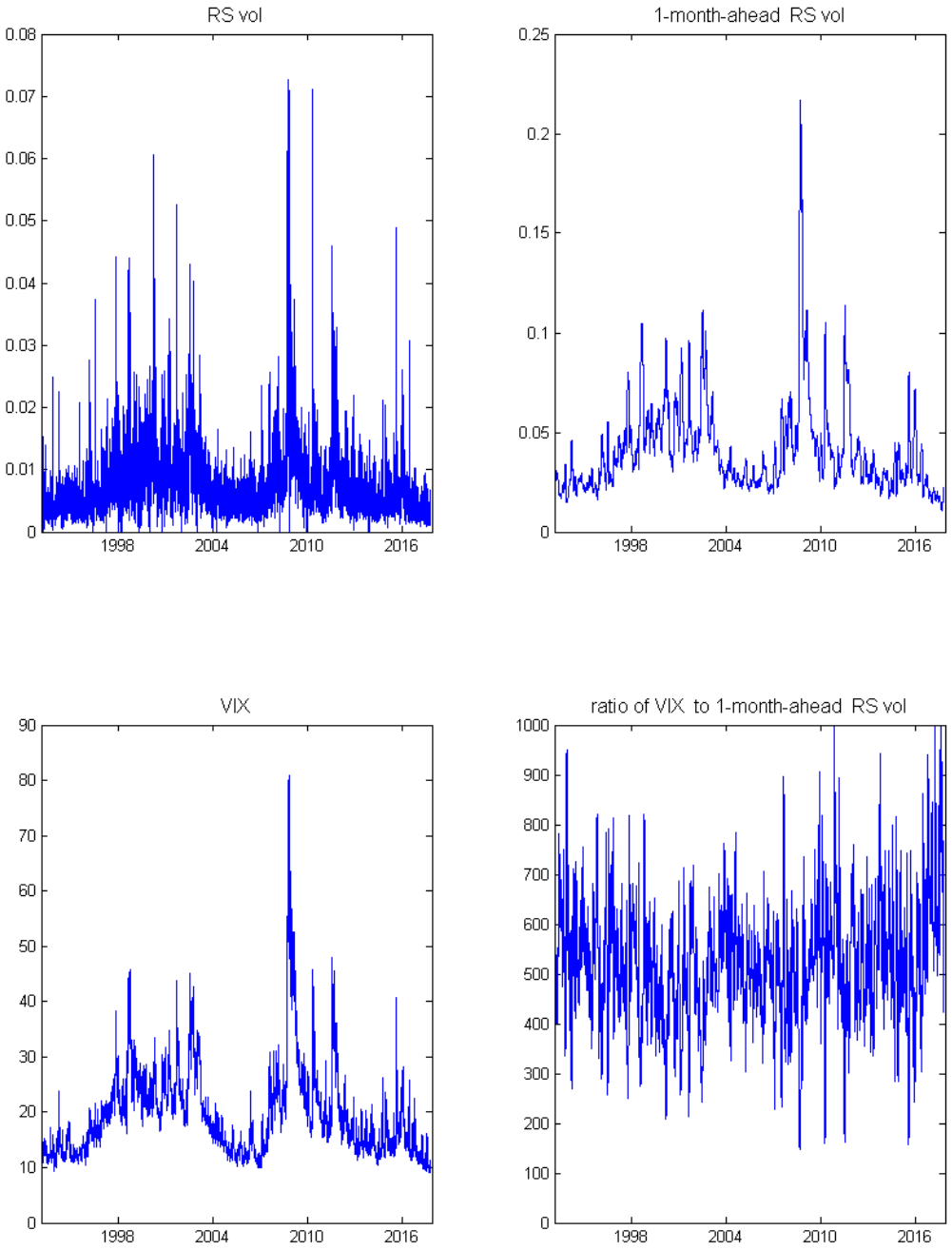
(a) beta based deciles



(b) standard deviation based deciles

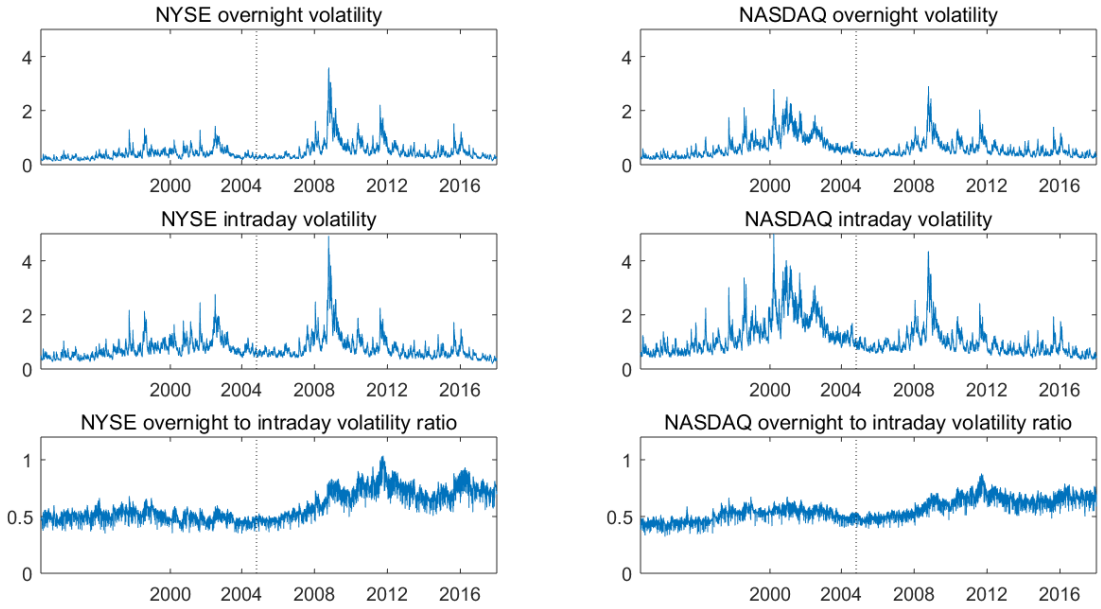
Panel a plots the ratio of overnight to intraday volatility for portfolios formed on beta. Decile 1 has the largest beta, around 1.53 on average, while decile 10 has the smallest beta, around 0.25 on average. Panel b plots the ratio for standard deviation sorted deciles. Decile 1 has the largest standard deviation, around 0.06 on average, and decile 10 has the smallest standard deviation, around 0.009 on average. Intraday and overnight volatilities are defined as $\sqrt{\frac{\nu_j}{\nu_j-2}exp(2\lambda_t^j + 2\sigma^j(\frac{t}{T}))}$, for $j = D, N$, respectively.

Figure A.13: Ratio of overnight to intraday volatility for portfolios formed on beta or standard deviation: NYSE/AMEX/ NASDAQ



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.

Figure A.14: RS, VIX and their ratio



This figure plots the overnight volatility, intraday volatility and the overnight to intraday volatility ratio for value-weighted NASDAQ index and value-weighted NYSE index. The dashed vertical line indicates the last trading day in October 2004, when NASDAQ had introduced the opening and closing crosses.

Figure A.15: Overnight and intraday volatilities: NASDAQ and NYSE

References

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

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

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

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

Severini, Thomas A and Wing Hung Wong (1992), "Profile likelihood and conditionally parametric models." *The Annals of Statistics*, 20, 1768–1802.

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

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

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