HYBRID-GARCH: A Generic Class of Models for Volatility Predictions using High Frequency Data — Technical Appendix

Xilong Chen

SAS Institute Inc.

Eric Ghysels

University of North Carolina at Chapel Hill

Fangfang Wang

University of Illinois at Chicago

S1 Notation

The purpose of this section is to collect the notation used in this supplementary file.

Let $\|X\|_p = (E|X|^p)^{1/p}$ for $X \in L^p(\Omega, \mathcal{F}, P)$ and $p < \infty$. $\|A\| = \sqrt{tr(A^T A)}$ for $A \in \mathbb{R}^{n\times n}$ or $A \in \mathbb{R}^{n\times 1}$ and $n \geq 1$. We write $A > 0$ if A is a positive definite matrix, and $A \geq 0$ if A is positive semi-definite. If A is finite element by element, then we write $A < \infty$.

We use ∇ to denote the vector differential operator (w.r.t θ) so that ∇f is the gradient (column vector) of scaler function f, and $Hess(f)$ the Hessian matrix of f, i.e., $ent_{i,j} Hess(f) = \partial_i \partial_j f$ where ∂_k denotes the partial derivative w.r.t. the k^{th} parameter in $\theta = (\alpha, \beta, \gamma, \phi)$. For a vector ϕ, ∂_ϕ represents the partial derivative w.r.t. a component of ϕ (say ϕ_i), and ∂^2_ϕ is treated as $\partial_{\phi_i}\partial_{\phi_j}$, and ∇_ϕ is a vector differential operator w.r.t. ϕ.

S2 More details on Assumption 2.4

Assumption 2.4 essentially guarantees that the HYBRID process is non-negative and measurable, and satisfies identifiability if it is parameterized. Conditions (1) and (2) are very standard. Here we give more explanations on condition (3) which also pertains to the choice of Φ.

Example 1. Consider the HYBRID process driven by MIDAS component with an exponential Almon lag polynomial:

$$H_t(\phi) = \sum_{j=0}^{m-1} (\tilde{\gamma} + b_j(\eta)) r^2_{t-j/m}, \qquad \text{(S2.1)}$$

and

$$b_j(\eta) = \frac{\exp\{\eta_1(j/m) + \eta_2(j/m)^2\}}{\sum_{k=0}^{m-1} \exp\{\eta_1(k/m) + \eta_2(k/m)^2\}}, \quad \tilde{\gamma} > 0, \eta_1, \eta_2 \in \mathbb{R}, \phi = (\tilde{\gamma}, \eta_1, \eta_2)^T.$$

For easy discussion we let $m = 5$.

1. When $\tilde{\gamma} > 0$ and $\eta_1, \eta_2 \neq 0$. Note that $\partial H_t/\partial\tilde{\gamma} = \sum_{j=0}^{m-1} r^2_{t-j/m}$, $\partial H_t/\partial\eta_1 = \sum_{j=0}^{m-1}(\partial b_j/\partial\eta_1) r^2_{t-j/m}$, and $\partial H_t/\partial\eta_2 = \sum_{j=0}^{m-1}(\partial b_j/\partial\eta_2) r^2_{t-j/m}$. For $c = (c_1, c_2, c_3, c_4, c_5)^\top \in \mathbb{R}^5$, suppose that $c_1 + c_2 H_t + c_3 \partial H_t/\partial\tilde{\gamma} + c_4 \partial H_t/\partial\eta_1 + c_5 \partial H_t/\partial\eta_2 = 0$, which is equivalent to $c_1 + \sum_{j=0}^{m-1}[c_2(\tilde{\gamma} + b_j) + c_3 + c_4(\partial b_j/\partial\eta_1) + c_5(\partial b_j/\partial\eta_2)] r^2_{t-j/m} = 0$. Hence $c_1 = 0$, and $c_2(\tilde{\gamma}+b_j)+c_3+c_4(\partial b_j/\partial\eta_1)+c_5(\partial b_j/\partial\eta_2) = 0$ for $j = 0, 1, 2, 3, 4$.

   Note that $\sum_{j=0}^{m-1} b_j(\eta) = 1$. We have $c_2(m\tilde{\gamma} + 1) + mc_3 = 0$, and $c_2(b_j - 1/m) + c_4(\partial b_j/\partial\eta_1) + c_5(\partial b_j/\partial\eta_2) = 0$, $\forall\ j$, or equivalently

$$\begin{pmatrix} b_0 - 1/5 & \partial b_0/\partial\eta_1 & \partial b_0/\partial\eta_2 \\ b_1 - 1/5 & \partial b_1/\partial\eta_1 & \partial b_1/\partial\eta_2 \\ b_2 - 1/5 & \partial b_2/\partial\eta_1 & \partial b_2/\partial\eta_2 \\ b_3 - 1/5 & \partial b_3/\partial\eta_1 & \partial b_3/\partial\eta_2 \\ b_4 - 1/5 & \partial b_4/\partial\eta_1 & \partial b_4/\partial\eta_2 \end{pmatrix} \begin{pmatrix} c_2 \\ c_4 \\ c_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \qquad \text{(S2.2)}$$

   Since

$$\frac{\partial b_j(\eta)}{\partial\eta_1} = b_j(\eta)\left(\frac{j}{m} - \sum_{k=0}^{m-1}\frac{k}{m} b_k(\eta)\right), \quad \frac{\partial b_j(\eta)}{\partial\eta_2} = b_j(\eta)\left(\frac{j^2}{m^2} - \sum_{k=0}^{m-1}\frac{k^2}{m^2} b_k(\eta)\right).$$

   the rank of the coefficient matrix in (S2.2) is 3. We have $c_2 = c_4 = c_5 = 0$, and hence $c_3 = 0$ as well. It follows that 1, H_t, and each component of $\partial_\phi(H_t)$ are linearly independent, when $\tilde{\gamma} > 0$ and $\eta_1, \eta_2 \neq 0$.

2. When $\tilde{\gamma} > 0$, and either $\eta_1 \neq 0, \eta_2 = 0$ or $\eta_1 = 0, \eta_2 \neq 0$, 1, H_t, and each component of $\partial_\phi(H_t)$ are linearly independent. The proof is similar to (1).

3. When $\tilde{\gamma} > 0$ and $\eta_1 = \eta_2 = 0$, $H_t(\phi) = \sum_{j=0}^{m-1}(\tilde{\gamma} + 1/m) r^2_{t-j/m}$. For $c = (c_1, c_2, c_3) \in R^3$, $c_1 + c_2 H_t + c_3 \partial H_t/\partial\tilde{\gamma} = 0$ is equivalent to $c_1 + \sum_{j=0}^{m-1}[c_2(\tilde{\gamma} + 1/m) + c_3] r^2_{t-j/m} = 0$, which implies $c_1 = 0$, and $c_2(\tilde{\gamma} + 1/m) + c_3 = 0$. Because c_2 and c_3 may not be zero at the same time, 1, H_t, and each component of $\partial_\phi(H_t)$ are linearly *dependent*.

The above discussion shows that if Φ is a connected subset of $\{(\tilde{\gamma}, \eta_1, \eta_2) : \tilde{\gamma} > 0, \eta_1^2 + \eta_2^2 \neq 0\}$, $H_t(\phi)$ satisfies condition (3). ■

Example 2. Consider the HYBRID process in equation (18), i.e.,

$$H_t(\phi) = \sum_{j=0}^{m-1} \Psi_j(\phi_1) NIC(\phi_2, r_{t-j/m}), \quad \sum_{j=0}^{m-1} \Psi_j(\phi_1) = 1, \tag{S2.3}$$

where $\phi = (\phi_1, \phi_2)$, and the weights $(\Psi_0(\phi_1), \Psi_1(\phi_1), \ldots, \Psi_{m-1}(\phi_1))^\top$ are determined by a low-dimensional functional specification. In this example, we will discuss how to choose weights and the parameter space Φ in order to meet condition (3). Two NIC specifications are considered:

$$NIC(\phi_2, r) = br^2 1_{r \geq 0} + \delta r^2 \mathbf{1}_{r<0}, \tag{S2.4}$$

$$NIC(\phi_2, r) = b(r - \delta)^2. \tag{S2.5}$$

Hence $\phi_2 = (b, \delta)$. The degenerate case that $\phi_1 = 0$ and/or $\phi_2 = 0$ is excluded from the discussion.

(1) Consider first $NIC(\phi_2, r) = br^2 1_{r\geq 0} + \delta r^2 1_{r<0}$ where $b \neq 0$, $\delta \neq 0$. For $c = (c_1, c_2, c_3, c_4, c_5)$, $c_1 + c_2 H_t + c_3 \partial H_t/\partial b + c_4 \partial H_t/\partial \delta + c_5^T \nabla_{\phi_1} H_t = 0$ is equivalent to

$$\begin{aligned} c_1 &+ \sum_{j=0}^{m-1} \left[c_2 \Psi_j b + c_3 \Psi_j + b c_5^T \nabla_{\phi_1} \Psi_j \right] 1_{r_{t-j/m} \geq 0} r_{t-j/m}^2 \\ &+ \sum_{j=0}^{m-1} \left[c_2 \Psi_j \delta + c_4 \Psi_j + \delta c_5^T \nabla_{\phi_1} \Psi_j \right] 1_{r_{t-j/m} < 0} r_{t-j/m}^2 = 0. \end{aligned}$$

Because $1_{r_{t-j/m} \geq 0}$ and $1_{r_{t-j/m}<0}$ are linearly independent, we have $c_1 = 0$, $c_2 \Psi_j b + c_3 \Psi_j + b c_5^T \nabla_{\phi_1} \Psi_j = 0$, $c_2 \Psi_j \delta + c_4 \Psi_j + \delta c_5^T \nabla_{\phi_1} \Psi_j = 0$ for $j = 0, 1, \ldots, m-1$. Note that $\sum_{j=0}^{m-1} \Psi_j = 1$. It follows that $c_2 b + c_3 = 0$, $c_2 \delta + c_4 = 0$, and $c_5^T \nabla_{\phi_1} \Psi_j = 0$ ($\forall\ j$). Moreover, c_5 is 0 if the weights satisfy Assumption S2.1 below.

Assumption S2.1. *The rank of the matrix* $(\nabla_{\phi_1} \Psi_0, \nabla_{\phi_1} \Psi_1, ..., \nabla_{\phi_1} \Psi_{m-1})$ *is same as the dimension of* ϕ_1.

But c_2,c_3,c_4 may not be zeros. Therefore, 1, H_t, and each component of $\partial_\phi(H_t)$ are linearly *dependent*.

In order to have $H_t(\phi)$ meet condition (3), one should consider $NIC(\phi_2, r) = r^2 1_{r\geq 0} + \delta r^2 1_{r<0}$ or $NIC(\phi_2, r) = br^2 1_{r\geq 0} + r^2 1_{r<0}$ and the weights satisfy Assumption S2.1.

(2) The HYBIRD process $H_t(\phi)$ with $NIC(\phi_2, r) = b(r - \delta)^2$ $(b > 0, \delta \neq 0)$ does not meet condition (3). The proof is similar. However, $H_t(\phi)$ with $NIC(\phi_2, r) = (r - c)^2$ and weights satisfying Assumption S2.1 will satisfy condition (3). ■

S3 Proofs

We first present some useful results. The following lemmas are stated under Assumptions 2.1 and 2.4.

Lemma S3.1. *Under Assumptions 3.1 and 3.3(1), $\partial_i V_{t|t-1}(\theta)$, $\partial_i\partial_j V_{t|t-1}(\theta)$ are strictly stationary ergodic for $\theta \in \mathcal{C}$ and $i, j \in \{1, 2, \dots, d{+}3\}$. Moreover under the additional Assumption 3.3(2), $E(\sup_{\theta\in\mathcal{C}} V_{t|t-1}(\theta))^2$, $E(\sup_{\theta\in\mathcal{C}} |\partial_i V_{t|t-1}(\theta)|)^2$, and $E(\sup_{\theta\in\mathcal{C}} |\partial_i\partial_j V_{t|t-1}(\theta)|)^2$ are bounded.*

Proof: Note that H_t, $\partial_i H_t$, $\partial_i\partial_j H_t$ are strictly stationary ergodic. $V_{t|t-1}(\theta) = \frac{\alpha}{1-\beta} + \gamma \sum_{k=0}^{\infty} \beta^k H_{t-1-k}(\phi)$ a.s for $\theta \in \mathcal{C}$. It is easy to check that $\sum_{k=0}^{\infty} \partial_i(\gamma\beta^k)$ and $\sum_{k=0}^{\infty} \partial_i\partial_j(\gamma\beta^k)$ are absolutely summable uniformly on $\mathcal{C}$, which implies that $\partial_i V_{t|t-1} = \partial_i(\alpha/(1-\beta)) + \sum_{k=0}^{\infty} \partial_i(\gamma\beta^k H_{t-1-k}(\phi))$ a.s. and $\partial_i\partial_j V_{t|t-1} = \partial_i\partial_j(\alpha/(1-\beta)) + \sum_{k=0}^{\infty} \partial_i\partial_j(\gamma\beta^k H_{t-1-k}(\phi))$ a.s., and hence they are strictly stationary ergodic.

Since $\mathcal{C}$ is bounded, one can always find constants (say) $c_1 > 0$, $c_2 > 0$ and $0 < c_3 < 1$ such that $V_{t|t-1}(\theta) \leq c_1 + c_2 \sum_{k=0}^{\infty} c_3^k \sup_{\phi\in\overline{\Phi^0}} H_{t-1-k}(\phi)$. Note that $\sum_{k=0}^{\infty} c_3^k (\sup_{\phi\in\overline{\Phi^0}} H_{t-1-k}(\phi))^2 < \infty$ a.s. due to Assumption 3.3(2). We have $V_{t|t-1}(\theta)^2 \leq 2c_1^2 + \frac{2c_2^2}{1-c_3} \sum_{k=0}^{\infty} c_3^k (\sup_{\phi\in\overline{\Phi^0}} H_{t-1-k}(\phi))^2$ a.s. due to the Cauchy-Schwarz inequality and hence $E(\sup_{\theta\in\mathcal{C}} V_{t|t-1}(\theta))^2$ is O(1). Similarly $E(\sup_{\theta\in\mathcal{C}} |\partial_i V_{t|t-1}(\theta)|)^2$ and $E(\sup_{\theta\in\mathcal{C}} |\partial_i\partial_j V_{t|t-1}(\theta)|)^2$ are O(1). ■

Lemma S3.2. *Fix $\theta \in \mathcal{C}$. If $p^T \nabla V_{t|t-1}(\theta) = 0$ a.s. for any $t \in \mathbb{Z}$, then $p \equiv 0$.*

Proof: Let $p = (p_1, p_2, p_3, p_4) \in \mathbb{R}^{d+3}$, where p_4 is of the same dimension as ϕ. Note that $\nabla V_{t+1|t}(\theta) = \nabla\alpha + (\nabla\beta)V_{t|t-1}(\theta) + \beta(\nabla V_{t|t-1}(\theta)) + \nabla(\gamma H_t(\phi))$. $p^T\nabla V_{t|t-1}(\theta) = 0$ a.s. implies $p_1 + p_2 V_{t|t-1}(\theta) + p_3 H_t(\phi) + \gamma p_4^T \nabla_\phi H_t(\phi) = 0$ a.s. Since $p_3 H_t(\phi) + \gamma p_4^T \nabla_\phi H_t(\phi) \in \mathcal{I}_t$, p_2=0 and hence $p_1 + p_3 H_t(\phi) + \gamma p_4^T \nabla_\phi H_t(\phi) = 0$ a.s. Assumption 2.4 implies $p_1 = p_3 = p_4 = 0$ (since $\gamma > 0$). ■

Lemma S3.3. *For $\theta \in \mathcal{C}$, $V_{t|t-1}(\theta) = V_{t|t-1}(\theta_0)$ a.s. $\forall\ t \in \mathbb{Z}$ if and only if $\theta = \theta_0$.*

Proof: Sufficiency is apparent. We need to check the necessity. If $V_{t|t-1}(\theta) = V_{t|t-1}(\theta_0)$ a.s. for $t \in \mathbb{Z}$, then $\alpha - \alpha_0 + (\beta - \beta_0)V_{t|t-1}(\theta_0) + (\gamma H(\phi, \vec{r}_t) - \gamma_0 H(\phi_0, \vec{r}_t)) = 0$ a.s. Since $V_{t|t-1}(\theta_0) \in \mathcal{I}_{t-1}$ and $\gamma H(\phi, \vec{r}_t) - \gamma_0 H(\phi_0, \vec{r}_t) \in \mathcal{I}_t$, we have $\beta = \beta_0$ and hence $(\alpha - \alpha_0) + (\gamma H(\phi, \vec{r}_t) - \gamma_0 H(\phi_0, \vec{r}_t)) = 0$ a.s. Note that $\gamma H(\phi, \vec{r}_t) - \gamma_0 H(\phi_0, \vec{r}_t) = H(\bar{\phi}, \vec{r}_t)(\gamma - \gamma_0) + \bar{\gamma}(\phi - \phi_0)^T \nabla_\phi H(\bar{\phi}, \vec{r}_t)$ where $(\bar{\gamma}, \bar{\phi})$ is between (γ, ϕ) and (γ_0, ϕ_0) and it may *depend on* t. Assumption 2.4 indicates that $\alpha = \alpha_0$, $\gamma = \gamma_0$ and $\phi = \phi_0$. In other words $\theta = \theta_0$. ■

Lemma S3.4. *Suppose that $E(\sup_{\phi\in\overline{\Phi^0}} H(\phi, \vec{r}_t))^\delta < \infty$ for some $\delta > 0$, and inequality (8) holds. Then we have*

$$E(\sup_{\theta\in\mathcal{C}} |\partial_i V_{t|t-1}/V_{t|t-1}|)^v < \infty, \quad E(\sup_{\theta\in\mathcal{C}} |\partial_i\partial_j V_{t|t-1}/V_{t|t-1}|)^v < \infty \quad \forall v > 0. \tag{S3.1}$$

Proof: $|\partial_\phi H(\phi,\vec{x})/H(\phi,\vec{x})|$ and $|\partial^2_\phi H(\phi,\vec{x})/H(\phi,\vec{x})|$ are bounded on $\mathcal{C}$. Suppose that the upper bound is $M_1 > 0$. Note that $|\partial_i(\alpha/(1-\beta))| \le (1/\alpha + 1/(1-\beta))\alpha/(1-\beta)$ and $|\partial_i(\gamma\beta^k)| \le (1/\gamma + k/\beta)\gamma\beta^k$. (8) implies that $|\partial_i V_{t|t-1}| \le |\partial_i(\alpha/(1-\beta))| + \sum_{k=0}^\infty C(k)\gamma\beta^k H_{t-1-k}(\phi)$ where $C(k) = M_1 + 1/\gamma + k/\beta$. Therefore, on $\mathcal{C}$, $|\partial_i V_{t|t-1}(\theta)/V_{t|t-1}(\theta)| \le (1/\alpha + 1/(1-\beta)) + C(N) + (1-\beta)/\alpha \sum_{k>N} C(k)\gamma\beta^k H_{t-1-k}(\phi)$, for $N \in \mathbb{N}$. Because one can always find constants $M_2 > 0$ and $0 < \rho_* < 1$ such that $(1-\beta)/\alpha C(k)\gamma\beta^k \le M_2\rho_*^k$, $1/\alpha + 1/(1-\beta) < M_2$ and $C(N) \le M_2 N$ on $\mathcal{C}$, we have for $\theta \in \mathcal{C}$, $|\partial_i V_{t|t-1}(\theta)/V_{t|t-1}(\theta)| \le M_2 + M_2 N + M_2 \sum_{k>N} \rho_*^k H_{t-1-k}(\phi)$. The rest of discussion is similar to the proof of Lemma 5.2 of Berkes et al. (2003), and hence we have $E\sup_{\theta\in\mathcal{C}} |\partial_i V_{t|t-1}/V_{t|t-1}|^v < \infty$ for any $v > 0$.

The second inequality in (S3.1) follows from a similar argument. ∎

Lemma S3.5. *Let $\varepsilon_t(\theta) = RV_t - V_{t|t-1}(\theta)$, and $\mathcal{F}_{t-m}^{t+m} = \sigma(r_s, t-m-1+1/m \le s \le t+m)$. Suppose that $Er_s^8 < \infty$, r_s is strictly stationary, and Assumption 3.3(3) is true. For $k \in \{1,\ldots,d+3\}$ and $\theta \in \mathcal{C}$, $\|\varepsilon_t\partial_k\varepsilon_t\|_2 < \infty$ and $\sup_t \|\varepsilon_t\partial_k\varepsilon_t - E(\varepsilon_t\partial_k\varepsilon_t|\mathcal{F}_{t-m}^{t+m})\|_2 \le C\rho^m$ for some constants $C > 0$ and $0 < \rho < 1$. Therefore $\{\varepsilon_t\partial_k\varepsilon_t, t \in \mathbb{Z}\}$ is near epoch dependent on $\{\vec{r}_t\}$. This is also true when RV_t is replaced with R_t^2.*

Proof: Let $Z_t = \varepsilon_t\partial_k\varepsilon_t$. Note that $E\sup_{\theta\in\mathcal{C}} V_{t|t-1}^4(\theta) < \infty$ and $E\sup_{\theta\in\mathcal{C}}(\partial_k V_{t|t-1}(\theta))^4 < \infty$, which follows from an argument similar to the proof of Lemma S3.1. We have $\|Z_t\|_2 \le \|\varepsilon_t\|_4\|\partial_k\varepsilon_t\|_4 < \infty$.

Since $\varepsilon_t(\theta) = RV_t - \frac{\alpha}{1-\beta} - \gamma\sum_{j=0}^\infty \beta^j H_{t-1-j}(\phi)$, it can be written as $\varepsilon_t(\theta) = \sum_{j=0}^\infty c_j(\theta)\tilde{H}_{t-j}(\theta)$ where $c_0(\theta) = 1$, $\tilde{H}_t(\theta) = RV_t - \alpha/(1-\beta)$, and $c_j(\theta) = -\gamma\beta^{j-1}$, $\tilde{H}_{t-j}(\theta) = H_{t-j}(\phi)$ for $j \ge 1$. Hence

$$Z_t = \left(\sum_{0\le i,j\le m} + \sum_{0\le i\le m, j>m} + \sum_{i>m, j\ge 0}\right) c_i\tilde{H}_{t-i}\partial_k(c_j\tilde{H}_{t-j}) \doteq Z_t^{(m)} + \xi_t^{(m)} + \eta_t^{(m)}. \tag{S3.2}$$

Note that $\|\xi_t^{(m)}\|_2 \le \sum_{0\le i\le m, j>m} |c_i\partial_k(c_j)|\|\tilde{H}_{t-i}\tilde{H}_{t-j}\|_2 + |c_ic_j|\|\tilde{H}_{t-i}\partial_k(\tilde{H}_{t-j})\|_2$. Since there exist $0 < \rho < 1$ and $M > 0$ such that $|c_i| < M\rho^i$ and $|\partial_k c_i| < M\rho^i$ for $i \ge 0$, $\|\xi_t^{(m)}\|_2 \le 2M^2B_1/(1-\rho)\rho^{m+1}$. Similarly, $\|\eta_t^{(m)}\|_2 \le 2M^2B_1/(1-\rho)\rho^{m+1}$. Note that $\|Z_t - E(Z_t|\mathcal{F}_{t-m}^{t+m})\|_2 \le \|Z_t - Z_t^{(m)}\|_2$. Therefore $\sup_t \|Z_t - E(Z_t|\mathcal{F}_{t-m}^{t+m})\|_2 \le C\rho^m$ for some constants $C > 0$ and $0 < \rho < 1$. ∎

Lemma S3.6. *Let $l_t(\theta) = \log V_{t|t-1} + RV_t/V_{t|t-1}(\theta)$. Suppose that r_s is strictly stationary. Then*

(1) $E\sup_{\theta\in\mathcal{C}} |l_t(\theta)| < \infty$ if $E\sup_{\phi\in\overline{\Phi^0}} H(\phi,\vec{r}_t) < \infty$.

(2) Suppose that $E\sup_{\phi\in\overline{\Phi^0}} H(\phi,\vec{r}_t) < \infty$ and inequality (8) holds. Then $E\sup_{\theta\in\mathcal{C}} |\partial_i l_t(\theta)| < \infty$ and $E\sup_{\theta\in\mathcal{C}} |\partial_i\partial_j l_t(\theta)| < \infty$. If additionally assume that $Er^{4+v} < \infty$ for some $v > 0$, then $E(\sup_{\theta\in\mathcal{C}} |\partial_i l_t(\theta)|)^2 < \infty$.

This is also true when RV_t is replaced with R_t^2.

Proof: (1) Note that $\log\alpha \le l_t(\theta) \le \log V_{t|t-1}(\theta) + RV_t/\alpha$. Hence $|l_t(\theta)| \le \max(|\log\alpha|, V_{t|t-1}(\theta) + RV_t/\alpha)$. Since $E\sup_{\theta\in\mathcal{C}} V_{t|t-1}(\theta) < \infty$ which follows from an argument similar to the proof of Lemma S3.1, we have $E\sup_{\theta\in\mathcal{C}} |l_t(\theta)| < \infty$.

(2) Note that $\partial_i l_t = (1 - RV_t/V_{t|t-1})\partial_i V_{t|t-1}/V_{t|t-1}$ and $\partial_i\partial_j l_t = \big(1 - RV_t/V_{t|t-1}\big)(\partial_i\partial_j V_{t|t-1}/V_{t|t-1}) + \big(2RV_t/V_{t|t-1} - 1\big)(\partial_i V_{t|t-1}/V_{t|t-1})(\partial_j V_{t|t-1}/V_{t|t-1})$. We have, due to due to Lemma S3.4,

$$\begin{aligned}
E\sup_{\theta\in\mathcal{C}}|\partial_i l_t(\theta)| \le& E(\sup_{\theta\in\mathcal{C}}(1+RV_t/\alpha))^2 E(\sup_{\theta\in\mathcal{C}}\partial_i V_{t|t-1}/V_{t|t-1})^2 < \infty,\\
E(\sup_{\theta\in\mathcal{C}}|\partial_i l_t(\theta)|)^2 \le& E(\sup_{\theta\in\mathcal{C}}(1+RV_t/\alpha))^{2+v/2} E(\sup_{\theta\in\mathcal{C}}\partial_i V_{t|t-1}/V_{t|t-1})^{(4+v)/(2+v)} < \infty,\\
E\sup_{\theta\in\mathcal{C}}|\partial_i\partial_j l_t(\theta)| \le& E(\sup_{\theta\in\mathcal{C}}(1+RV_t/\alpha))^2 E(\sup_{\theta\in\mathcal{C}}(\partial_i\partial_j V_{t|t-1}/V_{t|t-1}))^2\\
&+ E(\sup_{\theta\in\mathcal{C}}(2RV_t/\alpha+1))^2 E(\sup_{\theta\in\mathcal{C}}(\partial_i V_{t|t-1}/V_{t|t-1})(\partial_j V_{t|t-1}/V_{t|t-1}))^2 < \infty.
\end{aligned}$$

■

S3.1 Proofs of Propositions 3.1, and 3.3

Proof of Proposition 3.1: Note that $\|RV_t - V_{t|t-1}(\theta)\|_2^2 = \|RV_t - \sigma^2_{t|t-1}\|_2^2 + \|V_{t|t-1}(\theta) - \sigma^2_{t|t-1}\|_2^2$ for all θ's. Hence $\min_{\theta\in\mathcal{C}}\|RV_t - V_{t|t-1}(\theta)\|_2 = \|RV_t - V_{t|t-1}(\theta_0)\|_2$. Suppose there exists $\theta_1 \in \mathcal{C}$ such that $\|RV_t - V_{t|t-1}(\theta_1)\|_2 = \min_{\theta\in\mathcal{C}}\|RV_t - V_{t|t-1}(\theta)\|_2$. It implies $\|V_{t|t-1}(\theta_1) - \sigma^2_{t|t-1}\|_2 = 0$, or $V_{t|t-1}(\theta_1) = V_{t|t-1}(\theta_0)$ a.s. Therefore $\theta_1 = \theta_0$, which follows from Lemma S3.3. ■

Proof of Proposition 3.3: It suffices to justify the first equality. Define $l_t(\theta) = \log V_{t|t-1}(\theta) + R_t^2/V_{t|t-1}(\theta)$. Due to Lemmas S3.6 and S3.3, $E\sup_{\theta\in\mathcal{C}}|l_t(\theta)| < \infty$ and $E(l_t(\theta) - l_t(\theta_0)) = E\left(\frac{V_{t|t-1}(\theta_0)}{V_{t|t-1}(\theta)} - 1 - \log\frac{V_{t|t-1}(\theta_0)}{V_{t|t-1}(\theta)}\right) > 0$ if $\theta \ne \theta_0$. Therefore $El_t(\theta)$ is uniquely minimized at θ_0. ■

S3.2 Proof of Theorem 3.1

Let $\varepsilon_t(\theta) \doteq RV_t - V_{t|t-1}(\theta)$, $\tilde\varepsilon_t(\theta) \doteq RV_t - \tilde V_t(\theta)$, $O_T(\theta) \doteq 1/T\sum_{t=1}^T \varepsilon_t^2(\theta)$, $\tilde O_T(\theta) \doteq 1/T\sum_{t=1}^T \tilde\varepsilon_t^2(\theta)$. The proof is started with $\hat\theta_T^{mdrv} \doteq \arg\min_{\theta\in\mathcal{C}} O_T(\theta)$.

Lemma S3.7. *Under Assumptions 2.1, 2.2, 2.4, and 3.1, $\hat\theta_T^{mdrv}$ is identifiably unique and converges to θ_0 a.s.*

Proof: Note that $\varepsilon_t(\theta)$ is strictly stationary ergodic and $E\sup_{\theta\in\mathcal{C}}(\varepsilon_t(\theta))^2 < \infty$ (see Lemma S3.1). $O_T(\theta) - E(\varepsilon_t^2(\theta))$ converges to 0 a.s. uniformly on $\mathcal{C}$ due to uniform SLLN. Moreover θ_0 is identifiable unique. The results follow from Lemma A.1 of Goncalves and White (2004) and Theorem 3.3 of Gallant and White (1988). ■

Lemma S3.8. *Under Assumptions 2.1, 2.2, 2.4, and 3.1,* $\lim_{T\to\infty}\sup_{\theta\in\mathcal{C}}|O_T(\theta)-\tilde{O}_T(\theta)|\overset{a.s.}{=}0$.

Proof: Note that there exists $\kappa>1$ such that $\lim_{t\to\infty}\kappa^t\sup_{\theta\in\mathcal{C}}|V_{t|t-1}(\theta)-\tilde{V}_t(\theta)|\overset{a.s.}{=}0$ according to Theorem 3.1 of Bougerol (1993) or Theorem 2.8 of Straumann and Mikosch (2006). In other words, $\forall\delta>0$, $\exists\ T_0>0$ such that $\kappa^t\sup_{\theta\in\mathcal{C}}|\varepsilon_t(\theta)-\tilde{\varepsilon}_t(\theta)|<\delta$ for $t>T_0$. Hence $\sup_{\theta\in\mathcal{C}}|\tilde{\varepsilon}_t^2(\theta)-\varepsilon_t^2(\theta)|\le 2\delta\kappa^{-t}\sup_{\theta\in\mathcal{C}}|\varepsilon_t(\theta)|+\delta^2\kappa^{-2t}$ when $t>T_0$. Since under Assumption 3.1 $E\sup_{\theta\in\mathcal{C}}|\varepsilon_t(\theta)|$ is bounded away from 0, $E\log\sup_{\theta\in\mathcal{C}}|\varepsilon_t(\theta)|$ is finite as well. Considering Lemma 2.1 of Straumann and Mikosch (2006), we have $\overline{\lim}_{t\to\infty}\sup_{\theta\in\mathcal{C}}|\tilde{\varepsilon}_t^2(\theta)-\varepsilon_t^2(\theta)|=0$ a.s., and hence $\overline{\lim}_{T\to\infty}\sup_{\theta\in\mathcal{C}}|O_T(\theta)-\tilde{O}_T(\theta)|\le\overline{\lim}_{T\to\infty}\frac{1}{T}\sum_{t=1}^T\sup_{\theta\in\mathcal{C}}|\tilde{\varepsilon}_t^2(\theta)-\varepsilon_t^2(\theta)|=0$ a.s. ■

Proof of Theorem 3.1:

(1) Due to Lemmas S3.7 and S3.8. (2) Let $Z_t=\varepsilon_t(\theta_0)\partial_k\varepsilon_t(\theta_0)$. $EZ_t=0$. Lemma S3.5 implies that $\{Z_t\}$ is near epoch dependent on $\{\vec{r}_t\}$ and $\sup_t\|Z_t-E(Z_t|\mathcal{F}_{t-m}^{t+m})\|_2\le C\rho^m$ for some constants $C>0$ and $0<\rho<1$. Let $\Omega_T=var(\frac{1}{\sqrt{T}}\sum_{t=1}^T Z_t)$. Note that $\Omega_T=\gamma(0)+2\sum_{k=1}^{T-1}(1-k/T)\gamma(k)$ where $\gamma(k)=cov(Z_k,Z_0)$. For $k>0$

$$|\gamma(2k)|=|E(Z_tZ_{t-2k})|\le C\rho^k\|Z_t\|_2+12\|Z_t\|_{2+v_2}^2\alpha(k)^{v_2/(2+v_2)}. \tag{S3.3}$$

Therefore $\sum_{k=0}^\infty|\gamma(k)|<\infty$ under assumption 3.2 and thus $\lim_{T\to\infty}\Omega_T$ exists and is finite. ■

The proof of Theorem 3.1(3) needs the following lemmas.

Lemma S3.9. *Under Assumptions 2.1, 2.2, 2.4, and 3.1,*

$$\limsup_{N\to\infty}\lim_{T\to\infty}\sup_{\theta\in B(\theta_0,1/N)\cap\mathcal{C}}\|Hess(O_T)(\theta)-2\Sigma^{md}\|=0 \quad a.s. \tag{S3.4}$$

where $B(\theta_0,1/N)=\{\theta\in\mathbb{R}^{d+3}:\|\theta-\theta_0\|<1/N\}$ *and* $0<\Sigma^{md}=E\nabla V_{t|t-1}(\theta_0)(\nabla V_{t|t-1}(\theta_0))'<\infty$.

Proof: Since $\theta_0\in\mathcal{C}^0$, $B(\theta_0,1/N)\cap\mathcal{C}$ is not empty for sufficiently large N. H in Scenario 1 meets Assumption 2.4 automatically. Note that $Hess(O_T)=\frac{1}{T}\sum_{t=1}^T Hess(\varepsilon_t^2)$, and $\partial_i\partial_j\varepsilon_t^2=2\varepsilon_t\partial_i\partial_j\varepsilon_t+2\partial_i\varepsilon_t\partial_j\varepsilon_t$. $E\partial_i\partial_j\varepsilon_t^2(\theta_0)=2E\partial_i\varepsilon_t(\theta_0)\partial_j\varepsilon_t(\theta_0)$ due to $\partial_i\partial_j\varepsilon_t\in I_{t-1}$. Hence $EHess(\varepsilon_t^2)(\theta_0)=2\Sigma^{md}$. Clearly, $\Sigma^{md}\ge 0$ and O(1). Suppose that there exists $p\in\mathbb{R}^d$ such that $p'E\nabla\varepsilon_t(\theta_0)(\nabla\varepsilon_t(\theta_0))'p=0$, which is equivalent to $p'\nabla V_{t|t-1}(\theta_0)=0$ a.s. for all t. Lemma S3.2 implies $p\equiv 0$ and hence $\Sigma^{md}>0$.

Note that $\sup_{\theta\in B(\theta_0,1/N)\cap\mathcal{C}}\|Hess(O_T)(\theta)-2\Sigma^{md}\|\le\sup_{\theta\in\mathcal{C}}\|Hess(O_T)(\theta)-EHess(\varepsilon_1^2)(\theta)\|$ $+\ E\sup_{\theta\in B(\theta_0,1/N)\cap\mathcal{C}}\|Hess(\varepsilon_1^2)(\theta)-Hess(\varepsilon_1^2)(\theta_0)\|$, and $E\sup_{\theta\in\mathcal{C}}\|Hess(\varepsilon_t^2)(\theta)\|$ is O(1) uniformly in t due to Lemma S3.1. (S3.4) follows from the dominated convergence theorem and uniform SLLN. ■

Lemma S3.10. *Under Assumptions 2.1, 2.2, 2.4, 3.1, 3.2 and* $\Omega^{mdrv} > 0$, $\sqrt{T}(\hat{\theta}_T^{mdrv} - \theta_0) \Rightarrow N(0, (\Sigma^{md})^{-1}\Omega^{mdrv}(\Sigma^{md})^{-1})$, *where* $\Sigma^{md} = E\nabla V_{t|t-1}(\theta_0)(\nabla V_{t|t-1}(\theta_0))'$.

Proof: Note that $-\nabla O_T(\theta_0) = Hess(O_T)(\bar{\theta}_T)(\hat{\theta}_T^{mdrv} - \theta_0)$ where $\bar{\theta}_T$ is between θ_0 and $\hat{\theta}_T^{mdrv}$. Since $\bar{\theta}_T$ converges to θ_0 a.s., Lemma S3.9 implies that $Hess(O_T)(\bar{\theta}_T)$ converges to $2\Sigma^{md}$ a.s. Note that $2\Sigma^{md}$ is invertible – see the proof of Lemma S3.9. We have $\sqrt{T}(\hat{\theta}_T^{mdrv} - \theta_0) = -(2\Sigma^{md})^{-1}(1 + o_p(1))\sqrt{T}\nabla O_T(\theta_0)$. The asymptotic normality follows if $\sqrt{T}\nabla O_T(\theta_0)$ converges to $N(0, 4\Omega^{mdrv})$ in distribution. Therefore we just need to show that $\sqrt{T}p^T\nabla O_T(\theta_0)$ converges to $N(0, 4p^T\Omega^{mdrv}p)$ in distribution for any $p \in \mathbb{R}^{d+3}$ due to the Cramér-Wold device.

Note that $\sqrt{T}p^T\nabla O_T(\theta_0) = \frac{2}{\sqrt{T}}\sum_{t=1}^T Z_t$ where $Z_t = \sum_{k=1}^{d+3} p_k Y_{k,t}$ and $Y_{k,t} = \varepsilon_t(\theta_0)\partial_k\varepsilon_t\theta_0$. Let $\Omega_T = var(\frac{1}{\sqrt{T}}\sum_{t=1}^T Z_t)$. The random matrix Ω_T is O(1) and is uniformly positive definite, hence Ω_T^{-1} is O(1). Consider $X_{Tt} \doteq Z_t/\sqrt{T\Omega_T}$. $E(X_{Tt}) = 0$ and $Var(\sum_{t=1}^T X_{Tt}) = 1$. $\{X_{Tt}\}$ is near epoch dependent on $\{\vec{r}_t\}$ of size 1 due to Lemma S3.5 and $\{\vec{r}_t\}$ is $\alpha-$mixing of size $-(2+v_2)/v_2$. Note also that $\|Z_t\|_{2+v_2} < \infty$, and $T(1/\sqrt{T\Omega_T})^2$ is O(1). An application of Theorem 3.6 of Davidson (1992) yields that $\sum_{t=1}^T X_{Tt}$ converges to N(0,1) in distribution and hence $\sqrt{T}p^T\nabla O_T(\theta_0)$ converges to $N(0, 4p^T\Omega^{mdrv}p)$ in distribution. ■

Lemma S3.11. *Under Assumptions 2.1, 2.2, 2.4, and 3.1,*

$$\lim_{T\to\infty}\sup_{\theta\in\mathcal{C}}\sqrt{T}\|\nabla O_T(\theta) - \nabla\tilde{O}_T(\theta)\| \overset{a.s.}{=} 0.$$

Proof: Note that for $t \geq 1$, $\tilde{V}_t(\theta) = \alpha\frac{1-\beta^t}{1-\beta} + \beta^t v + \sum_{k=0}^{t-1}\gamma\beta^k H_{t-k}(\phi)$, and $\partial_i\tilde{V}_t(\theta) = \partial_i\left(\alpha\frac{1-\beta^t}{1-\beta}\right) + \partial_i(\beta^t)v + \sum_{k=0}^{t-1}\partial_i(\gamma\beta^k H_{t-k}(\phi))$. Note also that $\partial_i V_{t|t-1}(\theta) = \partial_i(\alpha/(1-\beta)) + \sum_{k=0}^{\infty}\partial_i(\gamma\beta^k H_{t-1-k}(\phi))$ (see Lemma S3.1). It is easy to check that both $\partial_i V_{t+1|t}(\theta)$ and $\partial_i\tilde{V}_t(\theta)$ satisfy

$$\partial_i X_t = \partial_i\alpha + (\partial_i\beta)X_{t-1} + \beta(\partial_i X_{t-1}) + \partial_i(\gamma H_t(\phi)), \quad t \in \mathbb{Z}^+, \tag{S3.5}$$

for each i. Since under Assumption 3.1 the conditions of Proposition 6.1 of Straumann and Mikosch (2006) are met, then $\partial_i V_{t|t-1}(\theta)$ is the unique stationary ergodic solution to (S3.5) and $\lim_{t\to\infty}\kappa_1^t\sup_{\theta\in\mathcal{C}}|\partial_i V_{t|t-1}(\theta) - \partial_i\tilde{V}_t(\theta)| \overset{a.s.}{=} 0$ for some $\kappa_1 > 1$, which implies $\lim_{t\to\infty}\kappa_1^t\sup_{\theta\in\mathcal{C}}|\partial_i\varepsilon_t(\theta) - \partial_i\tilde{\varepsilon}_t(\theta)| \overset{a.s.}{=} 0$. In other words, $\forall\delta > 0$, $\exists T_0 > 0$ such that $\sup_{\theta\in\mathcal{C}}|\partial_i\varepsilon_t(\theta) - \partial_i\tilde{\varepsilon}_t(\theta)| < \kappa_1^{-t}\delta$ and $\sup_{\theta\in\mathcal{C}}|\varepsilon_t(\theta) - \tilde{\varepsilon}_t(\theta)| < \kappa^{-t}\delta$ for $t > T_0$ (the second inequality is from (1)). Consequently,

$$\sqrt{t}\sup_{\theta\in\mathcal{C}}|\varepsilon_t(\theta)\partial_i\varepsilon_t(\theta) - \tilde{\varepsilon}_t(\theta)\partial_i\tilde{\varepsilon}_t(\theta)| \leq \delta\left[\sqrt{t}\kappa^{-2t}\delta + \sqrt{t}\kappa^{-t}\sup_{\theta\in\mathcal{C}}|\varepsilon_t(\theta)| + \sqrt{t}\kappa_1^{-t}\sup_{\theta\in\mathcal{C}}|\partial_i\varepsilon_t(\theta)|\right]$$

for $t > T_0$. Note that $E\sup_{\theta\in\mathcal{C}}|\partial_i\varepsilon_t(\theta)|$ and $E\sup_{\theta\in\mathcal{C}}|\varepsilon_t(\theta)|$ are bounded away from 0. Same as the discussion in (1), we have $\lim_{t\to\infty}\sqrt{t}\sup_{\theta\in\mathcal{C}}|\varepsilon_t(\theta)\partial_i\varepsilon_t(\theta) - \tilde{\varepsilon}_t(\theta)\partial_i\tilde{\varepsilon}_t(\theta)| = 0$ a.s. Therefore,

$$\lim_{T\to\infty}\sup_{\theta\in\Theta}\sqrt{T}|\partial_i O_T(\theta) - \partial_i\tilde{O}_T(\theta)| \leq \lim_{T\to\infty}\frac{1}{\sqrt{T}}\sum_{t=1}^T 2\sup_{\theta\in\Theta}|\varepsilon_t(\theta)\partial_i\varepsilon_t(\theta) - \tilde{\varepsilon}_t(\theta)\partial_i\tilde{\varepsilon}_t(\theta)| \overset{a.s.}{=} 0.$$

■

Proof of Theorem 3.1(3): It suffices to that $\sqrt{T}(\tilde{\theta}_T^{mdrv} - \hat{\theta}_T^{mdrv}) = o_p(1)$. Note that

$$\nabla O_T(\tilde{\theta}_T^{mdrv}) - \nabla O_T(\hat{\theta}_T^{mdrv}) = Hess(O_T)(\bar{\theta}_T)(\tilde{\theta}_T^{mdrv} - \hat{\theta}_T^{mdrv}),$$

where $\bar{\theta}_T \in \mathcal{C}$ is between $\tilde{\theta}_T^{mdrv}$ and $\hat{\theta}_T^{mdrv}$. Since $\bar{\theta}_T$ converges to θ_0 a.s., $Hess(O_T)(\bar{\theta}_T)$ converges to $2\Sigma^{md}$ a.s. Note that $2\Sigma^{md}$ is invertible – see the proof of Lemma S3.9. We have $\sqrt{T}(\tilde{\theta}_T^{mdrv} - \hat{\theta}_T^{mdrv}) = -(2\Sigma^{md})^{-1}(1+o_p(1))\sqrt{T}(\nabla O_T(\tilde{\theta}_T^{mdrv}) - \nabla O_T(\hat{\theta}_T^{mdrv}))$. Note also that $\sqrt{T}(\nabla O_T(\tilde{\theta}_T^{mdrv}) - \nabla O_T(\hat{\theta}_T^{mdrv})) = \sqrt{T}(\nabla O_T(\tilde{\theta}_T^{mdrv}) - \nabla \tilde{O}_T(\tilde{\theta}_T^{mdrv}))$ converges to 0 a.s. due to Lemma S3.11. Therefore $\sqrt{T}(\tilde{\theta}_T^{mdrv} - \hat{\theta}_T^{mdrv})$ converges to 0 in probability. ■

S3.3 Proof of Theorem 3.2

It suffices to show the proof of $\tilde{\theta}_T^{mdrv}$. Use the notation introduced in section S3.2.

(1) Similar to the proof of Theorem 3.1(1).

(2) Need to show $\sqrt{T}(\hat{\theta}_T^{mdrv} - \theta_0) \Longrightarrow N\left(0, (\Sigma^{md})^{-1}\Omega^{mdrv}(\Sigma^{md})^{-1}\right)$:

Lemma S3.9 still holds under Scenario 2. The first paragraph in the proof of Lemma S3.10 is true under Scenario 2, and we only need to revise the proof in the second paragraph, i.e., the asymptotic normality of $\sqrt{T}p^T\nabla O_T(\theta_0)$. Note that $Y_{k,t}(\theta_0) \equiv \varepsilon_t(\theta_0)\partial_k\varepsilon_t(\theta_0)$ is strictly stationary ergodic with finite second moment due to $Er^8 < \infty$ and Assumption 3.3(3). $Y_{i,t}(\theta_0)$ is a *martingale difference sequence*. Then $\sqrt{T}p^T\nabla O_T(\theta_0)$ $= \frac{1}{\sqrt{T}}\sum_{t=1}^T 2p^T Y_t(\theta_0)$ converges to $N(0, 4p^T\Omega^{mdrv}p)$ in distribution due to martingale central limit theorem where

$$\Omega^{mdrv} = EY_t(\theta_0)Y_t(\theta_0)^T = E[(RV_t - V_{t|t-1}(\theta_0))^2\nabla V_{t|t-1}(\theta_0)\nabla V_{t|t-1}(\theta_0)^T].$$

Note that $p^T\Omega^{mdrv}p > 0$ if and only if $p^T\nabla V_{t|t-1}(\theta_0) \neq 0$ a.s. Hence Ω^{mdrv} is positive definite.

Note that $\lim_{T\to\infty}\sqrt{T}(\tilde{\theta}_T^{mdrv} - \hat{\theta}_T^{mdrv}) = 0$ in probability, which follows from an argument similar to the proof of Theorem 3.1 (since Lemmas S3.8 and S3.9 are true under Scenario 2). $\sqrt{T}(\tilde{\theta}_T^{mdrv} - \theta_0)$ converges to $N(0, (\Sigma^{md})^{-1}\Omega^{mdrv}(\Sigma^{md})^{-1})$ in distribution.

S3.4 Proof of Theorem 3.3

It suffices to show the proof regarding $\tilde{\theta}_T^{lhrv}$. Define $l_t(\theta) = \log V_{t|t-1} + RV_t/V_{t|t-1}(\theta)$ and $\tilde{l}_t(\theta) = \log\tilde{V}_t + RV_t/\tilde{V}_t(\theta)$. Let $L_T(\theta) \equiv \frac{1}{T}\sum_{t=1}^T l_t(\theta)$ and $\tilde{L}_T(\theta) \equiv \frac{1}{T}\sum_{t=1}^T \tilde{l}_t(\theta)$. Suppose that $\hat{\theta}_T^{lhrv}$ is the solution to $\min_{\theta\in\mathcal{C}} L_T(\theta)$.

Lemma S3.12. *Under Assumptions 2.3, 3.1 and $E\sup_{\phi\in\overline{\Phi^0}} H(\phi, \vec{r}_t) < \infty$, $\hat{\theta}_T^{lhrv}$ is identifiably unique and it converges to θ_0 a.s.*

Proof: l_t is strictly stationary ergodic, and $E\sup_{\theta\in\mathcal{C}}|l_t(\theta)| < \infty$ (see Lemma S3.6). $L_T(\theta)$ converges to $EL_T(\theta) = El_1(\theta)$ a.s. uniformly on $\mathcal{C}$ due to the uniform SLLN. Moreover θ_0 is the unique minimizer of $L(\theta)$. The results follow from Lemma A.1 of Goncalves and White (2004) and Theorem 3.3 of Gallant and White (1988). ■

Lemma S3.13. *Suppose inequality (8) holds. Let* $B(\theta_0, 1/N) = \{\theta \in \mathbb{R}^{d+3} : \|\theta-\theta_0\| < 1/N\}$. *Under Assumptions 2.3, 3.1 and* $E\sup_{\phi\in\overline{\Phi^0}} H(\phi, \vec{r}_t) < \infty$,

$$\limsup_{N\to\infty} \lim_{T\to\infty} \sup_{\theta\in B(\theta_0,1/N)\cap\mathcal{C}} \|Hess(L_T)(\theta) - \Sigma^{lh}\| = 0 \quad a.s. \tag{S3.6}$$

where $0 < \Sigma^{lh} = E\left(V_{t|t-1}^{-2}(\theta_0)\nabla V_{t|t-1}(\theta_0)\nabla V_{t|t-1}(\theta_0)'\right) < \infty$.

Proof: Since $\theta_0 \in \mathcal{C}^0$, $B(\theta_0, 1/N)\cap\mathcal{C}$ is not empty for sufficiently large N. Note that

$$\partial_i\partial_j L_T = \frac{1}{T}\sum_{t=1}^{T}\partial_i\partial_j l_t = \frac{1}{T}\sum_{t=1}^{T}\left(1-\frac{RV_t}{V_{t|t-1}}\right)\frac{\partial_i\partial_j V_{t|t-1}}{V_{t|t-1}} + \left(\frac{2RV_t}{V_{t|t-1}}-1\right)\frac{\partial_i V_{t|t-1}\partial_j V_{t|t-1}}{V_{t|t-1}^2}.$$

$\partial_i\partial_j l_t$ is strictly stationary ergodic. And $E\sup_{\theta\in\mathcal{C}}|\partial_i\partial_j l_t(\theta)| < \infty$ by Lemma S3.6. $\Sigma^{lh} > 0$ because $p^{'}\nabla V_{t|t-1}(\theta_0) \neq 0$ a.s. for non-zero $p \in \mathbb{R}^{d+3}$ (see Lemma S3.2).

Note that $\sup_{\theta\in B(\theta_0,1/N)\cap\mathcal{C}}\|Hess(L_T)(\theta) - \Sigma^{lh}\| \le \sup_{\theta\in\mathcal{C}}\|Hess(L_T)(\theta) - EHess(l_1)(\theta)\| + E\sup_{\theta\in B(\theta_0,1/N)\cap\mathcal{C}}\|Hess(l_1)(\theta) - Hess(l_1)(\theta_0)\|$, and $E\sup_{\theta\in\mathcal{C}}\|Hess(l_t)(\theta)\|$ is O(1) uniformly in t. Thus (S3.6) follows from the dominated convergence theorem and uniform SLLN. ■

Lemma S3.14. *Suppose that* $Er^{4+v} < \infty$ *for* $v > 0$ *and inequality (8) holds. Under Assumptions 2.3 , 3.1, and* $E\sup_{\phi\in\overline{\Phi^0}} H(\phi, \vec{r}_t) < \infty$, $\sqrt{T}\nabla L_T(\theta_0) \Longrightarrow N(0, \Omega^{lhrv})$ *where* $\Omega^{lhrv} = E\left(V_{t|t-1}^{-4}(\theta_0)(RV_t - V_{t|t-1}(\theta_0))^2\nabla V_{t|t-1}(\theta_0)\nabla V_{t|t-1}(\theta_0)'\right) > 0$.

Proof: Note that $\nabla L_T = \frac{1}{T}\sum_{t=1}^{T}\nabla l_t = \frac{1}{T}\sum_{t=1}^{T}\left(1 - RV_t/V_{t|t-1}\right)\nabla V_{t|t-1}/V_{t|t-1}$. ∇l_t is strictly stationary ergodic. $E(\partial_i l_t(\theta_0))^2 < \infty$ due to Lemma S3.6. And $E(RV_t|\mathcal{F}_{t-1}) = V_{t|t-1}(\theta_0)$. Hence $\{\partial_i l_t(\theta_0), t \in \mathbb{Z}\}$ is a martingale difference sequence. Note also that Ω^{lhrv} is positive definite because $p'\nabla V_{t|t-1}(\theta_0) \neq 0$ a.s. for $p \neq 0$. The asymptotic normality follows from the martingale central limit theorem and the Cramer-Wold device. ■

Lemma S3.15. *Under assumptions 2.3 and 3.1 and* $E\sup_{\phi\in\overline{\Phi^0}} H(\phi, \vec{r}_t) < \infty$,

$$\lim_{T\to\infty}\sup_{\theta\in\mathcal{C}}|L_T(\theta) - \tilde{L}_T(\theta)| \overset{a.s.}{=} 0. \tag{S3.7}$$

Proof: Note that $L_T(\theta) - \tilde{L}_T(\theta) = \frac{1}{T}\sum_{t=1}^{T}(l_t(\theta) - \tilde{l}_t(\theta))$. It suffices to show $\lim_{t\to\infty}\sup_{\theta\in\mathcal{C}}|l_t(\theta) - \tilde{l}_t(\theta)| = 0$ a.s. Since $|l_t(\theta) - \tilde{l}_t(\theta)| \le |\log V_{t|t-1}(\theta) - \log\tilde{V}_t(\theta)| + \left|\frac{1}{V_{t|t-1}(\theta)} - \frac{1}{\tilde{V}_t(\theta)}\right| \le (1/\alpha + RV_t/\alpha^2)|V_{t|t-1}(\theta) - \tilde{V}_t(\theta)|$, we have $\sup_{\theta\in\mathcal{C}}|l_t(\theta) - \tilde{l}_t(\theta)| \le (1/\alpha_u + RV_t/\alpha_u^2)$

$\sup_{\theta\in\mathcal{C}} |V_{t|t-1}(\theta) - \tilde{V}_t(\theta)|$ for some $\alpha_u > 0$. Note also that $E\log^+ RV_t < \infty$ and $\lim_{t\to\infty} \kappa^t \sup_{\theta\in\mathcal{C}} |V_{t|t-1}(\theta) - \tilde{V}_t(\theta)| \overset{a.s.}{=} 0$ for some $\kappa > 1$ due to Lemma S3.8. $\lim_{t\to\infty}\sup_{\theta\in\mathcal{C}} |l_t(\theta) - \tilde{l}_t(\theta)| = 0$ a.s. by Lemma 2.1 of Straumann and Mikosch (2006). ■

Lemma S3.16. *Suppose that inequality (8) holds. Under assumptions 2.3 and 3.1, and* $E\sup_{\phi\in\overline{\Phi^0}} H(\phi, \vec{r}_t) < \infty$,

$$\lim_{T\to\infty} \sqrt{T} \sup_{\theta\in\mathcal{C}} \|\nabla L_T(\theta) - \nabla\tilde{L}_T(\theta)\| \overset{a.s.}{=} 0. \qquad \text{(S3.8)}$$

Proof: Since $\sqrt{T}\sup_{\theta\in\mathcal{C}} |\partial_i L_T(\theta) - \partial_i \tilde{L}_T(\theta)| \le \frac{1}{\sqrt{T}} \sum_{t=1}^T \sup_{\theta\in\mathcal{C}} |\partial_i l_t(\theta) - \partial_i \tilde{l}_t(\theta)|$ for each i, it suffices to show $\lim_{t\to\infty} \sqrt{t}\sup_{\theta\in\mathcal{C}} |\partial_i l_t(\theta) - \partial_i\tilde{l}_t(\theta)| = 0$ a.s. Note that

$$\partial_i l_t - \partial_i \tilde{l}_t = \left(1 - \frac{RV_t}{V_{t|t-1}}\right)\frac{\partial_i V_{t|t-1}}{V_{t|t-1}} - \left(1 - \frac{RV_t}{\tilde{V}_t}\right)\frac{\partial_i \tilde{V}_t}{\tilde{V}_t}.$$

Applying the mean value theorem to $\partial_i l_t - \partial_i \tilde{l}_t$, we have

$$|\partial_i l_t - \partial_i \tilde{l}_t| \le \frac{|\partial_i \tilde{V}_t - \partial_i V_{t|t-1}| + |\partial_i V_{t|t-1}|}{\alpha^2}\left(\frac{2RV_t}{\alpha} + 1\right)|V_{t|t-1} - \tilde{V}_t| + \left(1 + \frac{RV_t}{\alpha}\right)\frac{1}{\alpha}|\partial_i V_{t|t-1} - \partial_i \tilde{V}_t|.$$

Note that

$$\begin{aligned}&\sqrt{t}\frac{|\partial_i \tilde{V}_t - \partial_i V_{t|t-1}|}{\alpha^2}\left(\frac{2RV_t}{\alpha} + 1\right)|V_{t|t-1} - \tilde{V}_t| \\ =&\kappa_1^t |\partial_i \tilde{V}_t - \partial_i V_{t|t-1}| \kappa^t |V_{t|t-1} - \tilde{V}_t| \sqrt{t}\kappa^{-t}\kappa_1^{-t}\alpha^{-2}\left(\frac{2RV_t}{\alpha} + 1\right),\end{aligned}$$

$$\sqrt{t}\frac{|\partial_i V_{t|t-1}|}{\alpha^2}\left(\frac{2RV_t}{\alpha} + 1\right)|V_{t|t-1} - \tilde{V}_t| = \kappa^t |V_{t|t-1} - \tilde{V}_t|\sqrt{t}\kappa^{-t}\frac{|\partial_i V_{t|t-1}|}{\alpha^2}\left(\frac{2RV_t}{\alpha} + 1\right),$$

$$\sqrt{t}\left(1 + \frac{RV_t}{\alpha}\right)\frac{1}{\alpha}|\partial_i V_{t|t-1} - \partial_i \tilde{V}_t| = \kappa_1^t |\partial_i V_{t|t-1} - \partial_i \tilde{V}_t|\sqrt{t}\kappa_1^{-t}\left(1 + \frac{RV_t}{\alpha}\right)\frac{1}{\alpha}.$$

Since $E\log^+ RV_t < \infty$, and

$$E\log^+(\sup_{\theta\in\mathcal{C}} |\partial_i V_{t|t-1}(\theta)| RV_t) \le E\log^+(\sup_{\theta\in\mathcal{C}} |\partial_i V_{t|t-1}(\theta)|) + E\log^+ RV_t < \infty,$$

and $\lim_{t\to\infty} \kappa_1^t \sup_{\theta\in\mathcal{C}} |\partial_i V_{t|t-1} - \partial_i \tilde{V}_t| \overset{a.s.}{=} 0$, $\lim_{t\to\infty} \kappa^t \sup_{\theta\in\mathcal{C}} |V_{t|t-1} - \tilde{V}_t| \overset{a.s.}{=} 0$ (κ and κ_1 are defined in the proof of Lemma S3.8), we have $\lim_{t\to\infty}\sqrt{t}\sup_{\theta\in\mathcal{C}} |\partial_i l_t(\theta) - \partial_i \tilde{l}_t(\theta)|=0$ a.s. ■

Proof of Theorem 3.3: The results follow from an argument similar to the proof of Theorem 3.1. ■

S3.5 Proofs of Proposition 4.1 and Corollary 4.1

Proof of Proposition 4.1: As shown in Drost and Werker (1996), (a, b, c) relates to $(\theta, \omega, \lambda, v_L^*)$ in the following way: letting $h = 1/m$, $a = \omega(1 - e^{-\theta h})h$, $c = e^{-\theta h} - b$ and $|b| < 1$ is the solution to $\frac{b}{1+b^2} = \frac{\rho e^{-\theta h} - 1}{\rho(1+e^{-2\theta h})-2}$, where $\rho = \frac{4(e^{-\theta h}-1+\theta h)+2\theta h(1+(v/2+\theta h)(1-\lambda)/\lambda)}{1-e^{-2\theta h}}$, and $v = (\theta v_L^*)/(1-\lambda)$.

Note that $\rho = 1 + h\theta(1+1/\lambda) + \theta^2 h^2/\lambda + \tilde{v}(1 + h\theta + \theta^2 h^2/3) + o(h^2)$ where $\tilde{v} = (v/2)(1-\lambda)/\lambda = \theta v_L^*/(2\lambda)$. Therefore when $v_L^* > \mathbf{0}$, $b = 1 - h\theta(1+\phi) + o(h)$ and $c = e^{-\theta h} - b = h\theta\phi + o(h)$ where $\phi = \sqrt{1 + 1/\tilde{v}} - 1 = \sqrt{1 + 2\lambda/(\theta v_L^*)} - 1$. It implies that, as m goes to ∞, $\beta_m = b^m$ goes to $e^{-\theta(1+\phi)}$, $\frac{c}{1-b} = \frac{e^{-\theta h} - b}{1-b}$ tends to $\frac{\phi}{1+\phi}$, $\frac{d_m}{m} = \frac{1-(b+c)^m}{m(1-b-c)}$ tends to $\theta^{-1}(1 - e^{-\theta})$, $\alpha_m = \frac{ma(1-b^m)}{1-(b+c)}\left(1 - \frac{cd_m}{m(1-b)}\right)$ tends to $\omega\left(1 - e^{-\theta(1+\phi)}\right)\left(1 - \frac{\phi}{1+\phi}\theta^{-1}(1-e^{-\theta})\right)$, and $\gamma_m = cd_m$ tends to $(1 - e^{-\theta})\phi$.

Note that $\lim_{m\to\infty} \sum_{i=1}^{m} e^{-\theta(1+\phi)(t-t_{i-1})} r_{t_i}^2 = \int_{(t-1,t]} e^{-\theta(1+\phi)(t-s)} d[p,p]_s$ in probability where $t_i = t - 1 + i/m$ (see Protter (2004)). For any $\epsilon > 0$,

$$P\left(\left|\sum_{j=0}^{m-1} \beta_m^{j/m} r_{t-j/m}^2 - \sum_{i=1}^{m} e^{-\theta(1+\phi)(t-t_{i-1})} r_{t_i}^2\right| > \epsilon\right) \leq \frac{\omega}{\epsilon}(|\log(\beta_m) + \theta(1+\phi)|/2 + \theta(1+\phi)/m).$$

Therefore $\limsup_m P\left(\left|\sum_{j=0}^{m-1} \beta_m^{j/m} r_{t-j/m}^2 - \sum_{i=1}^{m} e^{-\theta(1+\phi)(t-t_{i-1})} r_{t_i}^2\right| > \epsilon\right) = 0$ and (14) is proved.

When $v_L^* = \mathbf{0}$, we have $b = 1 - \sqrt{h\theta\lambda} + o(h^{1/2})$ and $c = \sqrt{h\theta\lambda} + o(h^{1/2})$. Therefore, as m goes to ∞, $\beta_m = b^m$ tends to 0, $\frac{c}{1-b}$ tends to 1, $\frac{d_m}{m}$ tends to $\theta^{-1}(1 - e^{-\theta})$, $\alpha_m = \frac{ma(1-b^m)}{1-(b+c)}(1 - \frac{cd_m}{m(1-b)})$ tends to $\omega\left(1 - \theta^{-1}(1 - e^{-\theta})\right)$, and $\frac{\gamma_m}{\sqrt{m}} = \sqrt{\lambda/\theta}(1 - e^{-\theta})$.

We next show that $\sqrt{m}\sum_{j=0}^{m-1} \beta_m^{j/m} r_{t-j/m}^2$ converges to $(\theta\lambda)^{-1/2}\sigma_t^2$ in L^2, which is equivalent to show that $\lim_{m\to\infty} mc\sum_{j=0}^{m-1} b^j r_{t-j/m}^2 = \sigma_t^2$ in L^2. Let $\widetilde{RV}_t = \sum_{j=0}^{m-1} b^j r_{t-j/m}^2$. Note that

$$E(\widetilde{RV}_t^2) = \sum_{j=0}^{m-1} b^{2j}(kh^2\omega^2) + 2\sum_{j<i} b^{i+j}\left[h^2\omega^2 + \frac{\omega^2\lambda}{1-\lambda}\frac{e^{h\theta}(1-e^{-h\theta})^2}{\theta^2}e^{-(i-j)\theta/m}\right],$$

$$E(\widetilde{RV}_t\sigma_t^2) = \sum_{j=0}^{m-1} b^j E\left(\int_{t-j/m-1/m}^{t-j/m} \sigma_t\sigma_u dL_u\right)^2 = \frac{\omega^2}{m}\left[\frac{1-b^m}{1-b} + \frac{m\lambda}{1-\lambda}\theta^{-1}(1-e^{-\theta/m})\frac{1-b^m e^{-\theta}}{1-be^{-\theta/m}}\right].$$

Therefore,

$$E\left[mc\sum_{j=0}^{m-1} b^j r_{t-j/m}^2 - \sigma_t^2\right]^2$$
$$= \frac{\omega^2}{1-\lambda} + \underbrace{k\omega^2c^2\frac{1-b^{2m}}{1-b^2}}_{T_1} + \underbrace{2\omega^2c^2\sum_{j<i} b^{i+j}}_{T_2} + \underbrace{2m^2c^2\frac{\omega^2\lambda}{1-\lambda}\frac{e^{h\theta}(1-e^{-h\theta})^2}{\theta^2}\sum_{j<i} b^{i+j}e^{-(i-j)\theta/m}}_{T_3}$$

$$- \underbrace{2c\omega^2 \left[\frac{1-b^m}{1-b} + \frac{m\lambda}{1-\lambda}\theta^{-1}(1-e^{-\theta/m})\frac{1-b^m e^{-\theta}}{1-be^{-\theta/m}} \right]}_{T_4}.$$

Note that, as $m \Rightarrow \infty$, $T_1 \Rightarrow 0$, $T_2 \Rightarrow \omega^2$, $T_3 \Rightarrow \frac{\omega^2\lambda}{1-\lambda}$, and $T_4 \Rightarrow \frac{2\omega^2}{1-\lambda}$. Therefore $mc\sum_{j=0}^{m-1} b^{(m)j} r_{t-j/m}^2$ converges to σ_t^2 in L^2. ■

Proof of Corollary 4.1: Sufficiency follows from the fact that for $s > 0$, $\lim_{m\to\infty} P(\sup_{0\le t\le s} |V_{t+1|t}^{(m)} - E_t([p,p]_{t+1} - [p,p]_t)| \ge \varepsilon) \le \sum_{t=0}^{s} \lim_{m\to\infty} P(|V_{t+1|t}^{(m)} - E_t([p,p]_{t+1} - [p,p]_t)| \ge \varepsilon)=0$. To prove necessity, suppose $\{V_{t+1|t}^{(m)}, t\}_{m\ge1}$ converges to $\{E_t([p,p]_{t+1} - [p,p]_t), t\}$ uniformly on compacts in probability when jumps are present. It follows that $V_{t+1|t}^{(m)}$ converges to $E_t([p,p]_{t+1} - [p,p]_t)$ in probability for each t, and hence $H_t^{(m)}$ will converge to $\left(\sigma_t^2 - e^{-\theta(1+\phi)}\sigma_{t-1}^2 - \omega(1-e^{-\theta(1+\phi)})/(1+\phi)\right)/(\theta\phi)$ in probability, which however contradicts Proportion 4.1. ■

Table 1: Small sample property of various estimators, GARCH Diffusion

The table displays estimation of α_m, β_m, γ_m (and g for the MEM estimation procedure) of a GARCH diffusion process appearing in equation (9) ($\eta = 0$) with sample size 500 (Panel I:III) and sample size 1000 (Panel IV:VI), where the true values of α_m, β_m, γ_m are shown in the first line of each panel. The estimators considered are: *mdrv*, defined in (4), and the companion estimator *mdr*2, replacing RV by R^2, as well as (quasi-)likelihood-based estimators *lhr*2, defined in (6), and *lhrv*, defined in (7). The table also includes the *mem* method described in subsection 3.2.2. The numbers in the parenthesis are MSE for *lhr*2, relative MSE (with respect to *lhr*2) for *lhrv*, *mdr*2, *mdrv*, *mem*. For g, we only report sample variance.

	α_m	β_m	γ_m	g
		Panel I: m = 24, T = 500		
True Value	0.021560	0.606483	0.452303	
lhr2	0.028485 (0.000239)	0.574957 (0.021370)	0.519297 (0.078285)	
lhrv	0.027866 (**0.299575**)	0.592717 (**0.085556**)	0.463234 (**0.055312**)	
mdr2	0.047512 (14.672334)	0.554378 (2.350239)	0.624201 (8.267090)	
mdrv	0.029201 (0.731973)	0.603333 (0.153910)	0.444092 (0.097035)	
mem	0.003728 (1.825360)	0.639632 (0.163212)	0.439080 (0.295075)	7.793987 (0.312895)
		Panel II: m = 144, T = 500		
True Value	0.020402	0.294540	1.161865	
lhr2	0.045201 (0.003203)	0.283460 (0.043350)	1.658002 (2.818349)	
lhrv	0.026274 (0.098454)	0.285434 (0.044476)	1.183518 (0.006573)	
mdr2	0.068484 (2.404729)	0.277976 (1.392734)	2.723028 (8.974830)	
mdrv	0.029991 (**0.060709**)	0.289406 (0.070331)	1.166936 (0.012388)	
mem	0.002005 (0.150026)	0.308290 (**0.038025**)	1.159722 (**0.004771**)	26.320110 (5.386544)
		Panel III: m = 288, T = 500		
True Value	0.019472	0.177589	1.659011	
lhr2	0.043392 (0.003154)	0.192863 (0.040788)	3.269031 (22.558997)	
lhrv	0.023372 (**0.026421**)	0.172963 (**0.020106**)	1.680080 (**0.001080**)	
mdr2	0.072641 (2.886375)	0.195808 (1.321687)	6.074646 (7.459430)	
mdrv	0.028927 (0.060610)	0.175222 (0.047458)	1.667509 (0.002766)	
mem	0.020868 (0.738234)	0.175481 (0.062443)	1.935688 (0.176230)	35.562272 (100.568095)
		Panel IV: m = 24, T = 1000		
True Value	0.021560	0.606483	0.452303	
lhr2	0.027721 (0.000070)	0.582393 (0.011110)	0.493834 (0.037197)	
lhrv	0.027869 (**0.808425**)	0.590124 (**0.101996**)	0.465888 (**0.064436**)	
mdr2	0.038240 (23.955470)	0.568033 (3.269306)	0.579447 (12.799811)	
mdrv	0.026646 (1.013888)	0.603776 (0.184239)	0.447793 (0.134316)	
mem	0.002638 (6.866083)	0.640964 (0.235021)	0.434746 (0.068354)	7.730200 (0.187919)
		Panel V: m = 144, T = 1000		
True Value	0.020402	0.294540	1.161865	
lhr2	0.031422 (0.001167)	0.299256 (0.031683)	1.353009 (0.707128)	
lhrv	0.023179 (0.112965)	0.290395 (**0.029173**)	1.171383 (0.012677)	
mdr2	0.053584 (3.304682)	0.288688 (1.653277)	2.071579 (17.816284)	
mdrv	0.026522 (**0.078156**)	0.291540 (0.056826)	1.162984 (0.028757)	
mem	0.000591 (0.355572)	0.310927 (0.029184)	1.150915 (**0.008941**)	26.387767 (2.853065)
		Panel VI: m = 288, T = 1000		
True Value	0.019472	0.177589	1.659011	
lhr2	0.034068 (0.001155)	0.197662 (0.028785)	2.132745 (3.979507)	
lhrv	0.021192 (**0.029212**)	0.175443 (**0.016039**)	1.669002 (**0.003373**)	
mdr2	0.058739 (4.439753)	0.200768 (1.623452)	4.015785 (20.044682)	
mdrv	0.025429 (0.075973)	0.175460 (0.044330)	1.668098 (0.010440)	
mem	0.014977 (0.189630)	0.177831 (0.064847)	1.960999 (1.213615)	35.716633 (92.418742)

Bibliography

Berkes, I., L. Horvath, and P. Kokoszka (2003). GARCH processes: structure and estimation. *Bernoulli*, 201–227.

Bougerol, P. (1993). Kalman filtering with random coefficients and contractions. *SIAM Journal on Control and Optimization 31*, 942.

Davidson, J. (1992). A central limit theorem for globally nonstationary near-epoch dependent functions of mixing processes. *Econometric theory 8*(3), 313–329.

Drost, F. C. and B. J. Werker (1996). Closing the gap: Continuous time garch modeling. *Journal of Econometrics 74*, 31–57.

Gallant, A. and H. White (1988). *A unified theory of estimation and inference for nonlinear dynamic models*. Basil Blackwell Oxford.

Goncalves, S. and H. White (2004). Maximum likelihood and the bootstrap for nonlinear dynamic models. *Journal of Econometrics 119*(1), 199–219.

Protter, P. (2004). *Stochastic integration and differential equations*. Springer Verlag.

Straumann, D. and T. Mikosch (2006). Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equations approach. *Annals of Statistics 34*(5), 2449.