Closed-Form Estimation of Finite-Order ARCH Models:

Asymptotic Theory and Finite-Sample Performance

Supplemental Appendix

Todd Prono

PRELIMINARIES. This Supplemental Appendix contains the statements and proofs of all Lemmas that support the paper's main Theorems, as well as the asymptotic properties for OLS estimation of the ARCH(1), GJR ARCH(1), and ARCH(p) with $1 < p \leq \infty$ models. Concerning notation, C denotes a constant that can assume different values in different places. For matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A} \geq \mathbf{B}$ means that every element in $\mathbf{A} \geq$ every corresponding element in $\mathbf{B}$. For a vector $\mathbf{y}$, $\delta_{\mathbf{y}}$ denotes the Dirac measure at $\mathbf{y}$. Finally, RV(κ_0) is shorthand for Regularly Varying with tail index κ_0.

LEMMA 1. *For ARCH processes that can be cast in terms of the SRE*

$$\sigma_t^2 = \omega_0 + \sigma_{t-1}^2 A_t, \tag{1}$$

let Assumptions A1 and A2 hold. Then Assumption A4 is sufficient for $E\left(\sigma_t^3\right) < \infty$.

Proof.

$$\begin{aligned} \sigma_t^3 &\leq \sigma_t^2 \times \left(\omega_0^{1/2} + \sigma_{t-1} A_t^{1/2}\right). \\ &\leq \left(\omega_0 + \sigma_{t-1}^2 A_t\right) \times \left(\omega_0^{1/2} + \sigma_{t-1} A_t^{1/2}\right) \\ &\leq \omega_0 \sigma_t + \omega_0^{1/2} \sigma_t^2 + \sigma_{t-1}^3 A_t^{3/2}, \end{aligned} \tag{2}$$

where the first inequality follows from the Triangle Inequality, and the third inequality uses $\sigma_t^2 - \omega_0 = \sigma_{t-1}^2 A_t$. Since $\left\{\sigma_t^2\right\}$ is strictly stationary (see; e.g., Mikosch, 1999, Corollary 1.4.38) with a well-defined second moment (see; e.g., Bollerslev, 1986, Theorem 1),

$$\begin{aligned} E\left(\sigma_t^3\right) &\leq C + E\left(\sigma_{t-1}^3\right) E\left(A_t^{3/2}\right) \\ &\leq C\left(1 + E\left(A^{3/2}\right) + E\left(A^{3/2}\right)^2 + \cdots + E\left(A^{3/2}\right)^{k-1}\right) + E\left(\sigma_{t-k}^3\right) E\left(A^{3/2}\right)^k. \end{aligned}$$

As a consequence, $\lim_{k\to\infty} E\left(\sigma_t^3\right) \leq \frac{C}{1-E\left(A^{3/2}\right)} < \infty$ if and only if $E\left(A^{3/2}\right) < 1$. ■

LEMMA 2. *For ARCH processes consistent with (1), let Assumptions A1–A2 and A4 hold. Consider the following lagged vectors for* $h \geq 0$*:*

$$\mathbf{Y}_h^{(i)} = \left(\begin{array}{ccc} |Y_0|^i, & \ldots, & |Y_h|^i \end{array} \right), \ i = 1, 2,$$

$$\mathbf{E}_h^{(2)} = \left(\begin{array}{ccccc} \epsilon_0^2, & A_1 \epsilon_1^2, & \prod_{j=1}^{2} A_j \epsilon_2^2, & \ldots, & \prod_{j=1}^{h} A_j \epsilon_h^2 \end{array} \right).$$

If σ *is* $RV(\kappa_0)$*, then* $\mathbf{Y}_h^{(2)}$ *is* $RV(\kappa_0/2)$*, and* $\mathbf{Y}_h^{(1)}$ *is* $RV(\kappa_0)$*.*

Proof. That σ is RV(κ_0); i.e.,

$$P(\sigma > x) \sim c_0 x^{-\kappa_0}, \quad n \to \infty, \tag{3}$$

where $c_0 = c_0 \left(\omega_0, \alpha_{1,0}, \alpha_{2,0}\right)$, and $\kappa_0 \in (3, \ p]$ is the unique solution to

$$E(A)^{\kappa_0/2} = 1$$

follows from Mikosch and Stărică (2000, Theorem 2.1).[1] Next,

$$\begin{aligned} \mathbf{Y}_h^{(2)} &= \left(\begin{array}{ccc} \sigma_0^2 \epsilon_0^2, & \ldots, & \sigma_h^2 \epsilon_h^2 \end{array} \right) \\ &= \left(\begin{array}{cccc} \sigma_0^2 \epsilon_0^2, & \sigma_0^2 A_1 \epsilon_1^2, & \ldots, & \sigma_{h-1}^2 A_h \epsilon_h^2 \end{array} \right) \\ &\quad + \omega_0 \times \left(\begin{array}{cccc} 0, & \epsilon_1^2, & \ldots, & \epsilon_h^2 \end{array} \right) \\ &= \mathbf{C}_h^{(2)} + \mathbf{R}_h^{(2)}. \end{aligned}$$

Since the tail of $\mathbf{R}_h^{(2)}$ is small relative to the tail of $\mathbf{C}_h^{(2)}$, the tail of $\mathbf{Y}_h^{(2)}$ is determined only by the tail of $\mathbf{C}_h^{(2)}$. By induction, then, the tail of $\mathbf{Y}_h^{(2)}$ is determined by the tail of $\sigma_0^2 \times \mathbf{E}_h^{(2)}$. Given (3) and Mikosch (1999, Proposition 1.5.9), $\sigma_0^2 \times \mathbf{E}_h^2$ is RV($\kappa_0/2$) by Mikosch (1999, Proposition 1.3.9(b)). Given $\mathbf{Y}_h^{(2)}$ is RV($\kappa_0/2$), $\mathbf{Y}_h^{(1)}$ is RV(κ_0) by Mikosch (1999, Proposition 1.5.9). ■

REMARK R1: Lemma 2 summarizes a collection of results for (G)ARCH processes proved elsewhere in the literature (see; e.g., Davis and Mikosch, 1998, and Mikosch and Stărică, 2000). Note that A3 is not influential in determining $\mathbf{Y}_h^{(i)}$ to be regularly varying.

LEMMA 3. *For the GJR ARCH*(1) *model, let Assumptions A1–A2 and A4 hold. Consider the following lagged vectors for* $h \geq 0$,

$$\mathbf{Y}_h^i = \left(\begin{array}{ccc} Y_0^i, & \ldots, & Y_h^i \end{array} \right), \quad i = 1, 3,$$

[1] The precise value of c_0 is given in Goldie (1991).

$$\mathbf{E}_h^{(1)} = \left(\begin{array}{ccccc} \epsilon_0, & |\epsilon_0|\,\epsilon_1, & |\epsilon_0|\,|\epsilon_1|\,\epsilon_2, & \ldots, & \prod_{i=0}^{h-1} |\epsilon_i|\,\epsilon_h \end{array} \right).$$

Then for all $\mathbf{y}_h^1 \in \mathbb{R}^{h+1} \setminus \{\mathbf{0}\}$, $\mathbf{Y}_h^1$ *is* $RV(\kappa_0)$, *and* $\mathbf{Y}_h^3$ *is* $RV(\kappa_0/3)$.

Proof. For the GJR ARCH(1) model,

$$\sigma_t^2\left(\omega_0, \boldsymbol{\alpha}_0\right) = \omega_0 + \alpha_{0,t-1} Y_{t-1}^2, \tag{4}$$

where $\boldsymbol{\alpha}_0 = \left(\alpha_{1,0}, \alpha_{2,0}\right)'$. Define

$$\underline{\alpha} = \min\left(\alpha_{1,0},\ \alpha_{2,0}\right) \leq \alpha_{0,t-1}, \quad \overline{\alpha} = \max\left(\alpha_{1,0},\ \alpha_{2,0}\right) \geq \alpha_{0,t-1} \quad \forall\ t. \tag{5}$$

Take a first-order Taylor Expansion of $\sigma_h\left(\omega_0, \boldsymbol{\alpha}_0\right)$ around $\underline{\omega}$ so that

$$\sigma_h\left(\omega_0, \boldsymbol{\alpha}_0\right) = \frac{\alpha_{0,h-1} Y_{h-1}^2}{\sigma_h\left(\underline{\omega}, \boldsymbol{\alpha}_0\right)} + \frac{\left(\omega_0 + \underline{\omega}\right)}{2\sigma_h\left(\underline{\omega}, \boldsymbol{\alpha}_0\right)}.$$

Then,

$$\begin{aligned}
\mathbf{Y}_h^1 &= \left(\begin{array}{ccccc} Y_0, & Y_1, & Y_2, & \ldots, & Y_h \end{array} \right) \\
&= \sigma_0 \times \left(\begin{array}{ccccc} \epsilon_0, & \sigma_0^{-1}\left(\frac{\alpha_{0,0} Y_0^2}{\sigma_1(\underline{\omega}, \boldsymbol{\alpha}_0)}\right)\epsilon_1, & \sigma_0^{-1}\left(\frac{\alpha_{0,1} Y_1^2}{\sigma_2(\underline{\omega}, \boldsymbol{\alpha}_0)}\right)\epsilon_2, & \ldots, & \sigma_0^{-1}\left(\frac{\alpha_{0,h-1} Y_{h-1}^2}{\sigma_h(\underline{\omega}, \boldsymbol{\alpha}_0)}\right)\epsilon_h \end{array} \right) \\
&\quad + \left(\omega_0 + \underline{\omega}\right) \times \left(\begin{array}{ccccc} 0, & \frac{\epsilon_1}{2\sigma_1(\underline{\omega}, \boldsymbol{\alpha}_0)}, & \frac{\epsilon_2}{2\sigma_2(\underline{\omega}, \boldsymbol{\alpha}_0)}, & \ldots, & \frac{\epsilon_h}{2\sigma_h(\underline{\omega}, \boldsymbol{\alpha}_0)} \end{array} \right) \\
&= \mathbf{C}_h^1 + \mathbf{R}_h^1
\end{aligned}$$

Since $\sigma_h^{-1}\left(\underline{\omega}, \boldsymbol{\alpha}_0\right)$ is bounded, the tail of $\mathbf{R}_h^1$ is light relative to the tail of $\mathbf{C}_h^1$. As a consequence, the tail of $\mathbf{C}_h^1$ determines the tail of $\mathbf{Y}_h^1$. Let $\mathbf{C}_h^1 = \sigma_0 \times \mathbf{E}_h^{(1)*}$. Since y_h^1 is bounded away from zero for all h,

$$\frac{\alpha_{0,h-1} Y_{h-1}^2}{\sigma_h\left(\underline{\omega}, \boldsymbol{\alpha}_0\right)} \leq \frac{\overline{\alpha} Y_{h-1}^2}{\sigma_h\left(\underline{\omega}, \boldsymbol{\alpha}_0\right)} \leq \frac{\overline{\alpha} Y_{h-1}^2}{\underline{\alpha}^{1/2}\left|Y_{h-1}\right|} = \frac{\overline{\alpha}}{\underline{\alpha}^{1/2}} \times \sigma_{h-1}\left|\epsilon_{h-1}\right|,$$

in which case,

$$\mathbf{E}_h^{(1)*} \leq \left(\begin{array}{ccccc} \epsilon_0, & \left(\frac{\overline{\alpha}}{\underline{\alpha}^{1/2}}\right) \times |\epsilon_0|\,\epsilon_1, & \left(\frac{\overline{\alpha}}{\underline{\alpha}^{1/2}}\right) \times \left(\frac{\sigma_1}{\sigma_0}\right) \times |\epsilon_1|\,\epsilon_2, & \ldots, & \left(\frac{\overline{\alpha}}{\underline{\alpha}^{1/2}}\right) \times \left(\frac{\sigma_{h-1}}{\sigma_0}\right) \times \left|\epsilon_{h-1}\right|\epsilon_h \end{array} \right).$$

Using the Triangle Inequality,

$$\frac{\sigma_1}{\sigma_0} \leq \frac{\omega_0^{1/2} + \alpha_{0,0}^{1/2}\,|Y_0|}{\sigma_0} \leq \frac{\omega_0^{1/2} + \overline{\alpha}^{1/2}\,|Y_0|}{\sigma_0} \leq C \times \frac{|Y_0|}{\sigma_0} = C \times |\epsilon_0|\,,$$

where the final inequality holds because y_h^1 is bounded away from zero for all h, and

$$\frac{\sigma_2}{\sigma_0} \leq C \times \frac{|Y_1|}{\sigma_0} = C \times \left(\frac{\sigma_1}{\sigma_0}\right) \times |\epsilon_1|.$$

Suppose that

$$\frac{\sigma_{h-2}}{\sigma_0} \leq C \times \prod_{i=0}^{h-3} |\epsilon_i|.$$

Then

$$\frac{\sigma_{h-1}}{\sigma_0} \leq C \times \left(\frac{\sigma_{h-2}}{\sigma_0}\right) \times \left|\epsilon_{h-2}\right| \leq C \times \prod_{i=0}^{h-2} |\epsilon_i|,$$

so that by induction,

$$\mathbf{E}_h^{(1)*} \leq C \times \mathbf{E}_h^{(1)}. \tag{6}$$

Since $E\left(\left|E_h^{(1)}\right|^{\kappa_0+\varepsilon}\right) < \infty$ for all h and some $\varepsilon > 0$, $\sigma_0 \times \mathbf{E}_h^{(1)}$ is RV(κ_0) by Lemma 2 and Basrak, Davis, and Mikosch (2002, Corollary A.2) for $d = 1$, meaning that the tail behavior of σ_0 determines the tail behavior of the product $\sigma_0 \times \mathbf{E}_h^{(1)}$. Since $\mathbf{C}_h^1 = \sigma_0 \times \mathbf{E}_h^{(1)*}$ is established to determine the tail behavior of $\mathbf{Y}_h^1$, given (6), σ_0 must also determine the tail behavior of $\mathbf{C}_h^1$. As a consequence, $\mathbf{Y}_h^1$ is RV(κ_0); in which case, $\mathbf{Y}_h^3$ is RV($\kappa_0/3$) along the same lines as Resnick (2007, proof of Proposition 7.6), since $\mathbf{Y}_h^{(2)}$ is RV($\kappa_0/2$) by Lemma 2. ■

REMARK R2: In the case of the GJR ARCH(1) model, Lemma 3 requires $\alpha_i > 0$ for $i = 1, 2$. Lemma 3 also applies to the special case where $\alpha_{1,0} = \alpha_{2,0} = \alpha_0$ (i.e., the ARCH(1) model). Under Lemma 3, regular variation of $\{Y_t\}$ follows minus any need for symmetry in the distribution of rescaled errors and so is consistent with A3 and complementary to Basrak, Davis, and Mikosch (2002, Corollary 3.5(B)). If the rescaled errors are, in fact, symmetrically distributed, then regular variation of $\{Y_t\}$ can also follow from regular variation of $\{|Y_t|\}$ as given by Lemma 2 and independence of $\{|Y_t|\}$ and $\{sign\,(\epsilon_t)\}$ so that Basrak, Davis, and Mikosch (2002, Corollary A.2) applies. Both Davis and Mikosch (1998, Lemma A.1) and Mikosch and Stărică (2000, Theorem 2.3) rely on this latter argument.

LEMMA 4. *Consider the GJR ARCH*(1) *model under the same Assumptions as Lemma 3. For the sequence of constants* $\{a_n\}$*, where*

$$nP\left(|\mathbf{Y}| > a_n\right) \longrightarrow 1, \qquad n \to \infty,$$

$|\mathbf{Y}| = \max_{m=0,\ldots,h} |Y_m|$, $a_n = n^{1/\kappa_0} L(n)$, *and* $L(\cdot)$ *is slowly-varying at* ∞,

$$N_n := \sum_{t=1}^{n} \delta_{a_n^{-1}\mathbf{Y}_t} \xrightarrow{d} N := \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \delta_{P_i \mathbf{Q}_{i,j}}, \tag{7}$$

where: (i) $\sum_{i=1}^{\infty}\delta_{P_i}$ *is a Poisson process on* $(0,\ \infty)$*; (ii) For* $\mathbf{Q}_{i,j} = \left(\ Q_{ij}^{(0)},\ \ \ldots,\ \ Q_{ij}^{(h)}\ \right)$, $\sum_{j=1}^{\infty}\delta_{\mathbf{Q}_{i,j}}$, $i \in \mathbb{N}$, *is an i.i.d. sequence of point processes on* $\mathbb{R}_+^{h+1} \setminus \{\mathbf{0}\}$ *with common distribution* Q*; (iii)* $\sum_{i=1}^{\infty}\delta_{P_i}$ *and* $\sum_{j=1}^{\infty}\delta_{\mathbf{Q}_{i,j}}$, $i \in \mathbb{N}$, *are mutually independent.*

Proof. The proof proceeds by verifying the following conditions from Davis and Mikosch (1998, Theorem 2.8):

CONDITION C1: (joint) regular variation of all finite-dimensional distributions of $\mathbf{Y}_t$

CONDITION C2: weak mixing for $\{\mathbf{Y}_t\}$

CONDITION C3: That

$$\lim_{k\to\infty}\lim_{n\to\infty} P\left(\bigvee_{k\le|t|\le r_n} |\mathbf{Y}_t| > a_n y \mid |\mathbf{Y}_0| > a_n y\right) = 0, \qquad y > 0,$$

where $\vee_i b_i = \max_i (b_i)$, and $r_n,\ m_n \to \infty$ are two integer sequences such that $n\phi_{m_n}/r_n \to 0$, $r_n m_n/n \to 0$, and ϕ_n is the mixing rate of $\{\mathbf{Y}_t\}$

Lemmas 2 and 3 establish (C1). $\{\mathbf{Y}_t\}$ is strongly mixing by Carrasco and Chen (2002, Corollary 6) when $\alpha_{1,0} = \alpha_{2,0}$ and by Carrasco and Chen (2002, Corollary 10) when $\alpha_{1,0} \neq \alpha_{2,0}$. Lastly, when $\alpha_{1,0} = \alpha_{2,0}$, (C3) follows immediately from Davis and Mikosch (1998, proof of Theorem 4.1). When $\alpha_{1,0} \neq \alpha_{2,0}$, note that

$$\begin{aligned} Y_t^2 &= \sigma_t^2\epsilon_t^2 \\ &= \alpha_{0,t-1}\epsilon_t^2 Y_{t-1}^2 + \omega_0\epsilon_t^2 \\ &= A_t^* Y_{t-1}^2 + B_t, \end{aligned}$$

where $A_t^* = \alpha_{1,0} \times I_{\{\epsilon_{t-1}\ge 0\}} + \alpha_{2,0} \times I_{\{\epsilon_{t-1}<0\}}$. Since $\left\{\left(\ A_t^*,\ \ B_t\ \right)\right\}$ is an i.i.d. sequence, $\{Y_t^2\}$ satisfies an SRE. In this case, (C3) follows along the same lines as Davis, Mikosch, and Basrak (1999, proof of Theorem 3.3). ■

REMARK R3: Lemma 4 is the nonstandard CLT upon which (weak) distributional convergence of the IV and OLS estimators discussed in the main paper and this Supplemental Appendix are based. A generalization of this Lemma applies to the ARCH(p) case (see Basrak, Davis, and Mikosch, 2002, Theorem 2.10). Specification of the distribution $\mathbf{Q}$ is found in Davis and Mikosch (1998, Theorem 2.8). Following from Lemma 4, for

$$\mathbf{Y}_t^{(l)} = \left(\ Y_t^l,\ \ \ldots,\ \ Y_{t+h}^l\ \right), \qquad l = 2, 3,$$

$$N_n := \sum_{t=1}^{n} \delta_{a_n^{-l} \mathbf{Y}_t^l} \xrightarrow{d} N := \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \delta_{P_i^l \mathbf{Q}_{i,j}^{(l)}}, \tag{8}$$

where $\mathbf{Q}_{i,j}^{(l)} = \left(\left(Q_{ij}^{(m)} \right)^l, \ m = 0, \ldots, h \right)$ by a continuous mapping argument.

LEMMA 5. *For the ARCH(1) model, let Assumptions A1–A2 and A4 hold. For $m = 0, \ldots, h$, define*

$$\widehat{\gamma}_{(Y,\ Y^2)}(m) = n^{-1} \sum_{t=1}^{n-m} Y_t Y_{t+m}^2, \qquad \gamma_{(Y,\ Y^2)}(m) = E\left(Y_0 Y_m^2\right).$$

Then for a $\kappa_0 \in (3,\ 6)$,

$$n a_n^{-3} \left(\widehat{\gamma}_{(Y,\ Y^2)}(m) - \gamma_{(Y,\ Y^2)}(m) \right) \xrightarrow{d} (V_m)_{m=0,\ldots,h}, \qquad h \geq 1, \tag{9}$$

where $V_0 := V_0^* - c_3^* \alpha_0^{3/2} \left(1 - c_3 \alpha_0^{3/2}\right)^{-1} V_0^{**}$, and $V_m := V_m^* + \alpha_0 V_{m-1}$.

Proof. For an $\varepsilon > 0$, consider

$$\begin{aligned} & a_n^{-3} \sum_t \left(Y_{t+1}^3 - E\left(Y_{t+1}^3\right) \right) \\ = \ & a_n^{-3} \sum_t \sigma_{t+1}^3 \left(\epsilon_{t+1}^3 - c_3^* \right) \times I_{\{|Y_t| \leq a_n \varepsilon\}} \\ & + a_n^{-3} \sum_t \sigma_{t+1}^3 \left(\epsilon_{t+1}^3 - c_3^* \right) \times I_{\{|Y_t| > a_n \varepsilon\}} \\ & + c_3^* a_n^{-3} \sum_t \left(\sigma_{t+1}^3 - E\left(\sigma_{t+1}^3\right) \right) \\ = \ & Ia + IIa + IIIa, \end{aligned} \tag{10}$$

where $\sigma_{t+1}^3 \equiv \sigma_{t+1}^3(\omega_0, \alpha_0)$. Let $\kappa \equiv \kappa_0/3$, and consider a $r \in (\kappa, 2)$. For a $\zeta > 0$,

$$\begin{aligned} P(|Ia| > \zeta) \quad & \leq \quad \left(\zeta^{-1} a_n^{-3}\right)^r E \left| \sum_t \sigma_{t+1}^3 \times I_{\{|Y_t| \leq a_n \varepsilon\}} \times \left(\epsilon_{t+1}^3 - c_3^* \right) \right|^r \\ & \leq \quad \left(\zeta^{-1} a_n^{-3}\right)^r 2nE \left(\sigma_{t+1}^{3r} \times I_{\{|Y_t| \leq a_n \varepsilon\}} \right) \times E \left| \left(\epsilon_{t+1}^3 - c_3^* \right) \right|^r \\ & \leq \quad \left(\zeta^{-1} a_n^{-3}\right)^r 2CnE \left(|Y_t|^{3r} \times I_{\{|Y_t| \leq a_n \varepsilon\}} \right) \times E \left| \left(\epsilon_{t+1}^3 - c_3^* \right) \right|^r \\ & \leq \quad \left(\zeta^{-1} a_n^{-3}\right)^r 2C \left(\frac{\kappa_0}{3r - \kappa_0} \right) (a_n \varepsilon)^{3r} nP(|Y| > a_n \varepsilon) \times E \left| \left(\epsilon_{t+1}^3 - c_3^* \right) \right|^r \\ & \longrightarrow \quad \zeta^{-r} 2C \left(\frac{\kappa_0}{3r - \kappa_0} \right) \varepsilon^{3r - \kappa_0} \times E \left| \left(\epsilon_{t+1}^3 - c_3^* \right) \right|^r, \qquad n \to \infty \\ & \longrightarrow \quad 0, \qquad \varepsilon \to 0. \end{aligned} \tag{11}$$

The first inequality follows from Markov's Inequality. Since for

$$M_n \equiv \sum_t \sigma_{t+1}^3 \times I_{\{|Y_t| \leq a_n \varepsilon\}} \times \left(\epsilon_{t+1}^3 - c_3^* \right),$$

$$E\left(M_{n+1} \mid M_n\right) = M_n + \sigma_{n+2}^3 E\left(I_{\left\{\left|Y_{n+1}\right| \leq a_{n+1}\varepsilon\right\}} \mid M_n\right) \times E\left(\left(\epsilon_{n+2}^3 - c_3^*\right) \mid M_n\right) = M_n \quad a.s.,$$

the second inequality follows from von Bahr and Esseen (1965, Theorem 2).[2] In the third inequality, the constant $C \in (0, \infty)$. The fourth inequality relies on

$$\begin{aligned} E\left(\left|Y_t\right|^{3r} \times I_{\{|Y_t| > a_n\varepsilon\}}\right) &= \int_0^{a_n\varepsilon} |y|^{3r} f(y)\, dy \\ &= -\kappa_0 \int_0^{a_n\varepsilon} |y|^{3r-\kappa_0-1} L(y)\, dy \\ &\sim \frac{\kappa_0}{(3r-\kappa_0)} |y|^{3r-\kappa_0} L(y)\ \Big|_0^{a_n\varepsilon} \\ &= \frac{\kappa_0}{(3r-\kappa_0)} (a_n\varepsilon)^{3r} P(|Y| > a_n\varepsilon), \end{aligned} \tag{12}$$

where the second equality follows from Mikosch (1999, Theorem 1.2.9), and the "$\sim$" is the result of Karamata's Theorem. Lastly, "$\longrightarrow$" as $n \to \infty$ follows from the properties of regular variation, while "$\longrightarrow$" as $\varepsilon \to 0$ follows given the defined support for r. Next, for IIa,

$$IIa = a_n^{-3} \sum_t Y_{t+1}^3 I_{\{|Y_t| > a_n\varepsilon\}} - c_3^* a_n^{-3} \sum_t \sigma_{t+1}^3 I_{\{|Y_t| > a_n\varepsilon\}}$$

A first-order Taylor Expansion of σ_{t+1}^3 around $\underline{\omega}$ is (with some simplification),

$$\sigma_{t+1}^3 = C\sigma_{t+1}(\underline{\omega}, \alpha_0) + \alpha_0 \sigma_{t+1}(\underline{\omega}, \alpha_0) Y_t^2, \tag{13}$$

so

$$a_n^{-3} \sum_t \sigma_{t+1}^3 I_{\{|Y_t| > a_n\varepsilon\}} = C a_n^{-3} \sum_t \sigma_{t+1}(\underline{\omega}, \alpha_0) I_{\{|Y_t| > a_n\varepsilon\}} + \alpha_0 a_n^{-3} \sum_t \sigma_{t+1}(\underline{\omega}, \alpha_0) Y_t^2 I_{\{|Y_t| > a_n\varepsilon\}}. \tag{14}$$

Next, let $\mathbf{x}_t = \left(\begin{array}{ccc} x_t^{(0)}, & \ldots, & x_t^{(h)} \end{array}\right) \in \mathbb{R}^{h+1} \setminus \{\mathbf{0}\}$, and define for $j \geq 1$,

$$\begin{aligned} T_{j,m,\varepsilon}\left(\sum_{i=1}^{\infty} n_i \delta_{\mathbf{x}_i}\right) &= \sum_{i=1}^{\infty} n_i \left(x_i^{(m)}\right)^j I_{\left\{\left|x_i^{(0)}\right| > \varepsilon\right\}}, \qquad m = 0, 1, \\ T_{j,m,\varepsilon}^{(a)}\left(\sum_{i=1}^{\infty} n_i \delta_{\mathbf{x}_i}\right) &= \sum_{i=1}^{\infty} n_i \left|x_i^{(m)}\right|^j I_{\left\{\left|x_i^{(0)}\right| > \varepsilon\right\}}, \qquad m = 0, 1, \\ T_{m,\varepsilon}^{(1)}\left(\sum_{i=1}^{\infty} n_i \delta_{\mathbf{x}_i}\right) &= \sum_{i=1}^{\infty} n_i x_i^{(0)} \left(x_i^{(m-1)}\right)^2 I_{\left\{\left|x_i^{(0)}\right| > \varepsilon\right\}}, \qquad m \geq 2, \end{aligned}$$

noting that the set $\left\{x \in \mathbb{R}^{h+1} \setminus \{\mathbf{0}\} : \left|x^{(m)}\right| > \varepsilon\right\}$ for any $m \geq 0$ is bounded away from the origin.

[2] The applicability of von Bahr and Esseen (1965, Theorem 2) in this general context is first noted by Vaynman and Beare (2014, proof of Lemma 1).

Then, for the first part of the decomposition in (14),

$$\begin{aligned} 0 &\leq a_n^{-3}\sum_t \sigma_{t+1}\left(\underline{\omega},\alpha_0\right) I_{\{|Y_t|>a_n\varepsilon\}} \\ &\leq \underline{\omega}^{1/2}a_n^{-3}\sum_t I_{\{|Y_t|>a_n\varepsilon\}} + \alpha_0 a_n^{-3}\sum_t |Y_t|\, I_{\{|Y_t|>a_n\varepsilon\}} \longrightarrow 0, \quad n\to\infty, \end{aligned}$$

since (for sufficiently large n),

$$n\left(n^{-1}\sum_t I_{\{|Y_t|>a_n\varepsilon\}}\right) \sim nP\left(|Y|>a_n\varepsilon\right) \longrightarrow \epsilon^{-\kappa_0}, \quad n\to\infty, \tag{15}$$

as in (11) and

$$a_n^{-1}\sum_t |Y_t|\, I_{\{|Y_t|>a_n\varepsilon\}} = T_{1,0,\varepsilon}^{(a)}\left(N_n\right) \xrightarrow{d} T_{1,0,\varepsilon}^{(a)}\left(N\right), \qquad n\to\infty, \tag{16}$$

by (7), Remark R3 and, given Vaynman and Beare (2014, Lemma A.2), and the continuous mapping theorem.[3] For the second part of the decomposition in (14),

$$\begin{aligned} \alpha_0^{3/2}a_n^{-3}\sum_t |Y_t|^3\, I_{\{|Y_t|>a_n\varepsilon\}} &\leq \alpha_0 a_n^{-3}\sum_t \sigma_{t+1}\left(\underline{\omega},\alpha_0\right) Y_t^2 I_{\{|Y_t|>a_n\varepsilon\}} \\ &\leq C a_n^{-3}\sum_t Y_t^2 I_{\{|Y_t|>a_n\varepsilon\}} + \alpha_0^{3/2}a_n^{-3}\sum_t |Y_t|^3\, I_{\{|Y_t|>a_n\varepsilon\}}, \end{aligned}$$

where the second inequality follows from the Triangle Inequality. Since

$$a_n^{-2}\sum_t Y_t^2 I_{\{|Y_t|>a_n\varepsilon\}} = T_{2,0,\varepsilon}\left(N_n\right) \xrightarrow{d} T_{2,0,\varepsilon}\left(N\right), \qquad n\to\infty \tag{17}$$

by the same argument that supports (16),

$$a_n^{-3}\sum_t \sigma_{t+1}^3 I_{\{|Y_t|>a_n\varepsilon\}} = \alpha_0^{3/2}a_n^{-3}\sum_t |Y_t|^3\, I_{\{|Y_t|>a_n\varepsilon\}} + o_P\left(1\right).$$

As a consequence,

$$IIa = T_{3,1,\varepsilon}\left(N_n\right) - c_3^*\alpha_0^{3/2}T_{3,0,\varepsilon}^{(a)}\left(N_n\right) + o_P\left(1\right).$$

Also, given the same argument that supports the simplification of III from Davis and Mikosch (1998, Section 4(B2), p. 2072),

$$\begin{aligned} IIIa &= c_3^*\alpha_0^{3/2}a_n^{-3}\sum_t \left(\omega_0+\alpha_0 Y_t^2\right)^{3/2} - E\left(\left(\omega_0+\alpha_0 Y_t^2\right)^{3/2}\right) \\ &= c_3^*\alpha_0^{3/2}a_n^{-3}\sum_t \left(|Y_t|^3 - E\,|Y_t|^3\right) + o_P\left(1\right). \end{aligned} \tag{18}$$

[3] Elsewhere in this Appendix, implicit in applications of the continuous mapping theorem to functions of N_n defined in Lemma 4 is Vaynman and Beare (2014, Lemma A.2).

Next, the same decomposition in (10) is also applicable to

$$a_n^{-3}\sum_t \left(\left|Y_{t+1}\right|^3 - E\left|Y_{t+1}\right|^3\right) = Ib + IIb + IIIb$$

where $\left|\epsilon_{t+1}\right|^3$ in Ib and IIb is centered around c_3. By the same argument that supports (11), for a $\zeta > 0$,

$$\lim_{n\to\infty}\lim_{\epsilon\to 0}\sup P\left(|Ib| > \zeta\right) = 0.$$

Reliance on (13), (16), and (17) produces

$$IIb = T_{3,1,\varepsilon}^{(a)}\left(N_n\right) - c_3\alpha_0^{3/2}T_{3,0,\varepsilon}^{(a)}\left(N_n\right) + o_P(1)$$

As a consequence,

$$a_n^{-3}\sum_t \left(\left|Y_{t+1}\right|^3 - E\left|Y_{t+1}\right|^3\right) = \left(1 - c_3\alpha_0^{3/2}\right)^{-1}\left(T_{3,1,\varepsilon}^{(a)}\left(N_n\right) - c_3\alpha_0^{3/2}T_{3,0,\varepsilon}^{(a)}\left(N_n\right)\right) + o_P(1),$$

noting that $IIIa = IIIb$. In addition,

$$\begin{aligned} a_n^{-3}\sum_t \left(Y_{t+1}^3 - E\left(Y_{t+1}^3\right)\right) &= T_{3,1,\varepsilon}\left(N_n\right) \\ &\quad -c_3^*\alpha_0^{3/2}\left(1 - c_3\alpha_0^{3/2}\right)^{-1}\left(T_{3,1,\varepsilon}^{(a)}\left(N_n\right) - c_3\alpha_0^{3/2}T_{3,0,\varepsilon}^{(a)}\left(N_n\right)\right) + o_P(1) \\ &\quad \xrightarrow{d} T_{3,1,\varepsilon}(N) - c_3^*\alpha_0^{3/2}\left(1 - c_3\alpha_0^{3/2}\right)^{-1}\left(T_{3,1,\varepsilon}^{(a)}(N) - c_3\alpha_0^{3/2}T_{3,0,\varepsilon}^{(a)}(N)\right) \\ &= S(\varepsilon,\infty) + c_3^*c_3\alpha_0^3\left(1 - c_3\alpha_0^{3/2}\right)^{-1}S^*(\varepsilon,\infty) \\ &\quad \xrightarrow{d} V_0^* + c_3^*c_3\alpha_0^3\left(1 - c_3\alpha_0^{3/2}\right)^{-1}V_0^{**}, \end{aligned} \tag{19}$$

where the first "$\xrightarrow{d}$" is as $n \to \infty$ and follows from (7), Remark R3, and the continuous mapping theorem, and the second "$\xrightarrow{d}$" is as $\varepsilon \to 0$ and follows from Davis and Hsing (1995, Proof of Theorem 3.1, pp. 897-898). As a consequence,

$$na_n^{-3}\left(\widehat{\gamma}_{(Y,\, Y^2)}(0) - \gamma_{(Y,\, Y^2)}(0)\right) \xrightarrow{d} V_0^* + c_3^*c_3\alpha_0^3\left(1 - c_3\alpha_0^{3/2}\right)^{-1}V_0^{**} =: V_0$$

Next consider

$$
\begin{aligned}
& a_n^{-3} \sum_t Y_t Y_{t+1}^2 - E\left(Y_t Y_{t+1}^2\right) \\
= \;& a_n^{-3} \sum_t Y_t \sigma_{t+1}^2 \left(\epsilon_{t+1}^2 - 1\right) \times I_{\{|Y_t| \le a_n \varepsilon\}} \\
& + a_n^{-3} \sum_t Y_t \sigma_{t+1}^2 \left(\epsilon_{t+1}^2 - 1\right) \times I_{\{|Y_t| > a_n \varepsilon\}} \\
& + a_n^{-3} \sum_t Y_t \sigma_{t+1}^2 - E\left(Y_t \sigma_{t+1}^2\right) \\
= \;& Ic + IIc + IIIc
\end{aligned} \tag{20}
$$

Again by the same arguments that establish Eq. (11), for a $\zeta > 0$,

$$\lim_{n\to\infty} \lim_{\epsilon \to 0} \sup P\left(|Ic| > \zeta\right) = 0.$$

Since

$$a_n^{-1} \sum_t Y_t I_{\{|Y_t| > a_n \varepsilon\}} = T_{1,0,\varepsilon}\left(N_n\right) \xrightarrow{d} T_{1,0,\varepsilon}\left(N\right), \qquad n \to \infty,$$

given the same arguments that support (16),

$$
\begin{aligned}
IIc &= a_n^{-3} \sum_t Y_t Y_{t+1}^2 I_{\{|Y_t| > a_n \varepsilon\}} + a_n^{-3} \sum_t Y_t^3 I_{\{|Y_t| > a_n \varepsilon\}} + o_P(1) \\
&= T_{2,\varepsilon}^{(1)}\left(N_n\right) - \alpha_0 T_{3,0,\varepsilon}\left(N_n\right) + o_P(1).
\end{aligned}
$$

Finally, since

$$a_n^{-3} \sum_t Y_t = n^{\frac{\kappa_0 - 6}{2\kappa_0}} \left(n^{-1/2} \sum_t Y_t\right) \longrightarrow 0, \quad n \to \infty,$$

by Ibragimov and Linnik (1971, Theorem 18.5.3), given that $\{Y_t\}$ is strongly mixing by Carrasco and Chen (2002, Corollary 6),

$$IIIc = \alpha_0 a_n^{-3} \sum_t Y_t^3 - E\left(Y_t^3\right) + o_P(1)$$

so that

$$
\begin{aligned}
a_n^{-3} \sum_t Y_t Y_{t+1}^2 - E\left(Y_t Y_{t+1}^2\right) &= T_{2,\varepsilon}^{(1)}\left(N_n\right) - \alpha_0 T_{3,0,\varepsilon}\left(N_n\right) + IIIc + o_P(1) \\
& \xrightarrow{d} V_1^* + \alpha_0 V_0,
\end{aligned}
$$

where, as is true elsewhere, "$\xrightarrow{d}$" is first as $n \to \infty$ and then as $\varepsilon \to 0$, given the same arguments that support (19). As a consequence,

$$n a_n^{-3} \left(\gamma_{n,\,(Y,\,Y^2)}(1) - \gamma_{(Y,\,Y^2)}(1)\right) \xrightarrow{d} V_1^* + \alpha_0 V_0 =: V_1. \tag{21}$$

Extending (21) to higher lags (i.e., $m > 1$) is a continuation of the arguments given above. ■

LEMMA 6. *For the GJR ARCH*(1) *model, let Assumptions A1–A2 and A4 hold. For* $m = 0, \ldots, h$, *define*

$$\widehat{\gamma}^{+}_{(Y,\ Y^2)}(m) = n^{-1}\sum_{t=1}^{n-m} Y^2_{t+m} Y_t \times I_{\{Y_t \geq 0\}}, \qquad \gamma^{+}_{(Y,\ Y^2)}(m) = E\left(Y^2_m Y_0 \times I_{\{Y_0 \geq 0\}}\right),$$

with $\widehat{\gamma}^{-}_{(Y,\ Y^2)}(m)$ *and* $\gamma^{-}_{(Y,\ Y^2)}(m)$ *defined analogously using* $I_{\{Y_t<0\}}$. *Then for a* $\kappa_0 \in (3,\ 6)$ *and* $h > 1$,

$$na_n^{-3}\left(\widehat{\gamma}^{+}_{(Y,\ Y^2)}(m) - \gamma^{+}_{(Y,\ Y^2)}(m)\right) \xrightarrow{d} \left(W^{+}_m\right)_{m=0,\ldots,h}, \tag{22}$$

and

$$na_n^{-3}\left(\widehat{\gamma}^{-}_{(Y,\ Y^2)}(m) - \gamma^{-}_{(Y,\ Y^2)}(m)\right) \xrightarrow{d} \left(W^{-}_m\right)_{m=0,\ldots,h}, \tag{23}$$

where

$$W^{+}_m = V^{+}_m + \alpha_{1,0} W^{+}_{m-1}, \qquad W^{-}_m = V^{-}_m + \alpha_{2,0} W^{-}_{m-1},$$

and both W^{+}_0 *and* W^{-}_0 *jointly depend on* V^{**}_0 *from the proof of Lemma 5.*

Proof. Let $I^{+}(m) \equiv I_{\{\epsilon_{t+m} \geq 0\}}$ and $I^{-}(m) \equiv I_{\{\epsilon_{t+m} < 0\}}$ for $m = 0, 1$, noting that $I^{+}(m) = I_{\{Y_{t+m} \geq 0\}}$ and $I^{-}(m) = I_{\{Y_{t+m} < 0\}}$. Then,

$$E\left(Y^3_{t+1} \times I^{+/-}(1)\right) = E\left(\sigma^3_{t+1}\right) c_3^{+/-},$$

where $c_3^{+/-} = E\left(\epsilon^3_{t+1} \times I^{+/-}(1)\right)$, and

$$\begin{aligned}
& a_n^{-3}\sum_t Y^3_{t+1} \times I^{+/-}(1) - E\left(Y^3_{t+1} \times I^{+/-}(1)\right) \\
= \;& a_n^{-3}\sum_t \sigma^3_{t+1}\left(\epsilon^3_{t+1} \times I^{+/-}(1) - c_3^{+/-}\right) \times I_{\{|Y_t| \leq a_n\varepsilon\}} \\
& + a_n^{-3}\sum_t \sigma^3_{t+1}\left(\epsilon^3_{t+1} \times I^{+/-}(1) - c_3^{+/-}\right) \times I_{\{|Y_t| > a_n\varepsilon\}} \\
& + c_3^{+/-} a_n^{-3}\sum_t \left(\sigma^3_{t+1} - E\left(\sigma^3_{t+1}\right)\right) \\
= \;& Ia^{+/-} + IIa^{+/-} + IIIa^{+/-}.
\end{aligned}$$

Given the same arguments that support (11), for a $\zeta > 0$, $\lim_{n\to\infty}\lim_{\varepsilon\to 0}\sup P\left(\left|Ia^{+/-}\right| > \zeta\right) = 0$. Consider next IIa^{+}. Given (4),

$$\sigma^3_{t+1}(\omega_0, \boldsymbol{\alpha}_0) = C\sigma_{t+1}(\underline{\omega}, \boldsymbol{\alpha}_0) + \alpha_{0,t}\sigma_{t+1}(\underline{\omega}, \boldsymbol{\alpha}_0) Y^2_t$$

by a first-order Taylor Expansion of σ_{t+1}^3 around $\underline{\omega}$. Then

$$\begin{aligned} a_n^{-3}\sum_t \sigma_{t+1}^3 \times I_{\{|Y_t|>a_n\varepsilon\}} &= Ca_n^{-3}\sum_t \sigma_{t+1}\left(\underline{\omega}, \boldsymbol{\alpha}_0\right) \times I_{\{|Y_t|>a_n\varepsilon\}} \\ &\quad + a_n^{-3}\sum_t \alpha_{0,t}\sigma_{t+1}\left(\underline{\omega}, \boldsymbol{\alpha}_0\right) Y_t^2 \times I_{\{|Y_t|>a_n\varepsilon\}}. \end{aligned}$$

Note that

$$\begin{aligned} 0 &\leq a_n^{-3}\sum_t \sigma_{t+1}\left(\underline{\omega}, \boldsymbol{\alpha}_0\right) \times I_{\{|Y_t|>a_n\varepsilon\}} \\ &\leq a_n^{-3}\sum_t \left(\underline{\omega}^{1/2} + \alpha_{0,t}^{1/2}\left|Y_t\right|\right) \times I_{\{|Y_t|>a_n\varepsilon\}} \\ &\leq \underline{\omega}^{1/2} a_n^{-3}\sum_t I_{\{|Y_t|>a_n\varepsilon\}} + \overline{\alpha}^{1/2} a_n^{-3}\sum_t \left|Y_t\right| \times I_{\{|Y_t|>a_n\varepsilon\}} \longrightarrow 0, \quad n \to \infty \end{aligned}$$

where the second inequality follows from the Triangle Inequality; the third inequality relies on (5), and "$\longrightarrow$" to zero follows from (15) and (16). Also note that, again based on (5),

$$\begin{aligned} a_n^{-3}\sum_t \alpha_{0,t}\sigma_{t+1}\left(\underline{\omega}, \boldsymbol{\alpha}_0\right) Y_t^2 \times I_{\{|Y_t|>a_n\varepsilon\}} &\geq \underline{\alpha} a_n^{-3}\sum_t \left(\underline{\omega} + \underline{\alpha} Y_t^2\right)^{1/2} Y_t^2 \times I_{\{|Y_t|>a_n\varepsilon\}} \\ &\geq \underline{\alpha}^{3/2} a_n^{-3}\sum_t \left|Y_t\right|^3 \times I_{\{|Y_t|>a_n\varepsilon\}}, \end{aligned}$$

and

$$\begin{aligned} a_n^{-3}\sum_t \alpha_{0,t}\sigma_{t+1}\left(\underline{\omega}, \boldsymbol{\alpha}_0\right) Y_t^2 \times I_{\{|Y_t|>a_n\varepsilon\}} &\leq \overline{\alpha} a_n^{-3}\sum_t \left(\underline{\omega} + \overline{\alpha} Y_t^2\right)^{1/2} Y_t^2 \times I_{\{|Y_t|>a_n\varepsilon\}} \\ &\leq \overline{\alpha}^{3/2} a_n^{-3}\sum_t \left|Y_t\right|^3 \times I_{\{|Y_t|>a_n\varepsilon\}} + \overline{\alpha}\underline{\omega}^{1/2} a_n^{-3}\sum_t Y_t^2 \times I_{\{|Y_t|>a_n\varepsilon\}} \\ &= \overline{\alpha}^{3/2} a_n^{-3}\sum_t \left|Y_t\right|^3 \times I_{\{|Y_t|>a_n\varepsilon\}} + o_P\left(1\right), \end{aligned}$$

where the equality follows from (17) so that there exists a constant C for which

$$\begin{aligned} IIa^+ &= a_n^{-3}\sum_t Y_t^3 \times I_{\left\{Y_{t+1}\geq 0\right\}} \times I_{\{|Y_t|>a_n\varepsilon\}} - c_3^+ a_n^{-3}\sum_t \sigma_{t+1}^3 \times I_{\{|Y_t|>a_n\varepsilon\}} \\ &= a_n^{-3}\sum_t Y_t^3 \times I_{\left\{Y_{t+1}\geq 0\right\}} \times I_{\{|Y_t|>a_n\varepsilon\}} - c_3^+ C a_n^{-3}\sum_t \left|Y_t\right|^3 \times I_{\{|Y_t|>a_n\varepsilon\}} + o_P\left(1\right). \end{aligned}$$

Based on $\mathbf{x}_t$ defined in the proof of Lemma 5 and for the same j and m, define

$$T_{j,m,\varepsilon}^+\left(\sum_{i=1}^{\infty} n_i \delta_{\mathbf{x}_i}\right) = \sum_{i=1}^{\infty} n_i \left(x_i^{(m)}\right)^j \times I_{\left\{x_i^{(m)}\geq 0\right\}} \times I_{\left\{\left|x_i^{(0)}\right|>\varepsilon\right\}},$$

and define $T_{j,m,\varepsilon}^-\left(\sum_{i=1}^{\infty} n_i \delta_{\mathbf{x}_i}\right)$ analogously, with $I_{\left\{x_i^{(m)}<0\right\}}$ replacing $I_{\left\{x_i^{(m)}\geq 0\right\}}$. Then

$$IIa^+ = T_{3,1,\varepsilon}^+\left(N_n\right) - c_3^+ C T_{3,0,\varepsilon}^{(a)}\left(N_n\right) + o_P\left(1\right).$$

Next, from

$$IIIa^{+}=c_3^{+}\left[a_n^{-3}\sum_t \sigma_{t+1}^3\times I_{\{Y_t\geq 0\}}-E\left(\sigma_{t+1}^3\times I_{\{Y_t\geq 0\}}\right)+a_n^{-3}\sum_t \sigma_{t+1}^3\times I_{\{Y_t<0\}}-E\left(\sigma_{t+1}^3\times I_{\{Y_t<0\}}\right)\right],$$

where

$$\begin{aligned}
& a_n^{-3}\sum_t \sigma_{t+1}^3\times I_{\{Y_t\geq 0\}}-E\left(\sigma_{t+1}^3\times I_{\{Y_t\geq 0\}}\right) \\
= \; & a_n^{-3}\sum_t \left(\omega_0+\alpha_{1,0}Y_t^2\right)^{3/2}\times I_{\{Y_t\geq 0\}}-E\left(\left(\omega_0+\alpha_{1,0}Y_t^2\right)^{3/2}\times I_{\{Y_t\geq 0\}}\right) \\
= \; & \alpha_{1,0}^{3/2}a_n^{-3}\sum_t \left|Y_t\right|^3\times I_{\{Y_t\geq 0\}}-E\left(\left|Y_t\right|^3\times I_{\{Y_t\geq 0\}}\right)+o_P(1),
\end{aligned}$$

with an analogous decomposition holding for $a_n^{-3}\sum_t \left(\sigma_{t+1}^3\times I_{\{Y_t<0\}}-E\left(\sigma_{t+1}^3\times I_{\{Y_t<0\}}\right)\right)$, follows that

$$\begin{aligned}
IIIa^{+} \;=\; & c_3^{+}\alpha_{1,0}^{3/2}a_n^{-3}\sum_t \left|Y_t\right|^3\times I_{\{Y_t\geq 0\}}-E\left(\left|Y_t\right|^3\times I_{\{Y_t\geq 0\}}\right) \\
& +c_3^{+}\alpha_{2,0}^{3/2}a_n^{-3}\sum_t \left|Y_t\right|^3\times I_{\{Y_t<0\}}-E\left(\left|Y_t\right|^3\times I_{\{Y_t<0\}}\right)+o_P(1).
\end{aligned}$$

As a consequence,

$$\begin{aligned}
& a_n^{-3}\sum_t \left(Y_{t+1}^3\times I_{\left\{Y_{t+1}\geq 0\right\}}-E\left(Y_{t+1}^3\times I_{\left\{Y_{t+1}\geq 0\right\}}\right)\right) \qquad (24)\\
= \; & \left(1-c_3^{+}\alpha_{1,0}^{3/2}\right)^{-1}\left[T_{3,1,\varepsilon}^{+}\left(N_n\right)-c_3^{+}CT_{3,0,\varepsilon}^{(a)}\left(N_n\right)\right] \\
& +c_3^{+}\alpha_{2,0}^{3/2}\left(1-c_3^{+}\alpha_{1,0}^{3/2}\right)^{-1}\left[a_n^{-3}\sum_t \left|Y_t\right|^3\times I_{\{Y_t<0\}}-E\left(\left|Y_t\right|^3\times I_{\{Y_t<0\}}\right)\right]+o_P(1).
\end{aligned}$$

The same arguments that establish (24) also establish

$$\begin{aligned}
& a_n^{-3}\sum_t \left(Y_{t+1}^3\times I_{\left\{Y_{t+1}<0\right\}}-E\left(Y_{t+1}^3\times I_{\left\{Y_{t+1}<0\right\}}\right)\right) \qquad (25)\\
= \; & \left(1-c_3^{-}\alpha_{2,0}^{3/2}\right)^{-1}\left[T_{3,1,\varepsilon}^{-}\left(N_n\right)-c_3^{-}CT_{3,0,\varepsilon}^{(a)}\left(N_n\right)\right] \\
& +c_3^{-}\alpha_{1,0}^{3/2}\left(1-c_3^{-}\alpha_{2,0}^{3/2}\right)^{-1}\left[a_n^{-3}\sum_t \left(Y_{t+1}^3\times I_{\{Y_t\geq 0\}}-E\left(Y_{t+1}^3\times I_{\{Y_t\geq 0\}}\right)\right)\right]+o_P(1).
\end{aligned}$$

From (24) and (25) then follows that

$$
\begin{aligned}
& a_n^{-3}\sum_t \left(Y_{t+1}^3 \times I_{\left\{Y_{t+1}\geq 0\right\}} - E\left(Y_{t+1}^3 \times I_{\left\{Y_{t+1}\geq 0\right\}}\right)\right) \\
& \xrightarrow{d} \left[1-\left(c_3^+\alpha_{1,0}^{3/2}+c_3^-\alpha_{2,0}^{3/2}\right)\right]^{-1} \times \left[\left(1-c_3^-\alpha_{2,0}^{3/2}\right)T_{3,1,\varepsilon}^+(N)+c_3^+\alpha_{2,0}^{3/2}T_{3,1,\varepsilon}^-(N)+c_3^+CT_{3,0,\varepsilon}^{(a)}(N)\right] \\
= & \left[1-\left(c_3^+\alpha_{1,0}^{3/2}+c_3^-\alpha_{2,0}^{3/2}\right)\right]^{-1} \times \left[S^+(\varepsilon,\infty)+c_3^+CS^*(\varepsilon,\infty)\right] \\
& \xrightarrow{d} \left[1-\left(c_3^+\alpha_{1,0}^{3/2}+c_3^-\alpha_{2,0}^{3/2}\right)\right]^{-1} \times \left[V_0^+ + c_3^+CV_0^{**}\right]
\end{aligned}
$$

where, as is true elsewhere, "$\xrightarrow{d}$" is first as $n\to\infty$ and then as $\varepsilon\to 0$, with each result following from the same, respective, arguments that support (19). As a consequence,

$$
na_n^{-3}\left(\widehat{\gamma}_{(Y,\,Y^2)}^+(0)-\gamma_{(Y,\,Y^2)}^+(0)\right) \xrightarrow{d} \left[1-\left(c_3^+\alpha_{1,0}^{3/2}+c_3^-\alpha_{2,0}^{3/2}\right)\right]^{-1} \times \left[V_0^+ + c_3^+CV_0^{**}\right] =: W_0^+.
$$

Moreover, since following parallel arguments,

$$
\begin{aligned}
& a_n^{-3}\sum_t \left(Y_{t+1}^3 \times I_{\left\{Y_{t+1}< 0\right\}} - E\left(Y_{t+1}^3 \times I_{\left\{Y_{t+1}< 0\right\}}\right)\right) \\
& \xrightarrow{d} \left[1-\left(c_3^+\alpha_{1,0}^{3/2}+c_3^-\alpha_{2,0}^{3/2}\right)\right]^{-1} \times \left[\left(1-c_3^+\alpha_{1,0}^{3/2}\right)T_{3,1,\varepsilon}^-(N)+c_3^-\alpha_{1,0}^{3/2}T_{3,1,\varepsilon}^+(N)+c_3^-CT_{3,0,\varepsilon}^{(a)}(N)\right] \\
= & \left[1-\left(c_3^+\alpha_{1,0}^{3/2}+c_3^-\alpha_{2,0}^{3/2}\right)\right]^{-1} \times \left[S^-(\varepsilon,\infty)+c_3^-CS^*(\varepsilon,\infty)\right] \\
& \xrightarrow{d} \left[1-\left(c_3^+\alpha_{1,0}^{3/2}+c_3^-\alpha_{2,0}^{3/2}\right)\right]^{-1} \times \left[V_0^- + c_3^-CV_0^{**}\right],
\end{aligned}
$$

$$
na_n^{-3}\left(\widehat{\gamma}_{(Y,\,Y^2)}^-(0)-\gamma_{(Y,\,Y^2)}^-(0)\right) \xrightarrow{d} \left[1-\left(c_3^+\alpha_{1,0}^{3/2}+c_3^-\alpha_{2,0}^{3/2}\right)\right]^{-1} \times \left[V_0^- + c_3^-CV_0^{**}\right] =: W_0^-.
$$

Next, define

$$
T_{m,\varepsilon}^+\left(\sum_{i=1}^{\infty}n_i\delta_{\mathbf{x}_i}\right)=\sum_{i=1}^{\infty}n_i x_i^{(0)}\left(x_i^{(m-1)}\right)^2 \times I_{\left\{x_i^{(0)}>\varepsilon\right\}}, \qquad m\geq 2,
$$

and consider

$$
\begin{aligned}
& a_n^{-3}\sum_t Y_{t+1}^2 Y_t \times I^{+/-}(0) - E\left(Y_{t+1}^2 Y_t \times I^{+/-}(0)\right) \\
= \; & a_n^{-3}\sum_t \sigma_{t+1}^2 Y_t \times I^{+/-}(0) \times \left(\epsilon_{t+1}^2-1\right) \times I_{\{|Y_t|\leq a_n\varepsilon\}} \\
& + a_n^{-3}\sum_t \sigma_{t+1}^2 Y_t \times I^{+/-}(0) \times \left(\epsilon_{t+1}^2-1\right) \times I_{\{Y_t> a_n\varepsilon\}} \\
& + a_n^{-3}\sum_t \sigma_{t+1}^2 Y_t \times I^{+/-}(0) - E\left(\sigma_{t+1}^2 Y_t \times I^{+/-}(0)\right) \\
= \; & Ib^{+/-} + IIb^{+/-} + IIIb^{+/-}.
\end{aligned}
$$

Again following the same arguments that support (11), $\lim_{n\to\infty}\lim_{\varepsilon\to 0}\sup P\left(\left|Ib^+\right|>\zeta\right)=0$ for a $\zeta>0$.

In addition,

$$\begin{aligned} IIb^{+} &= a_n^{-3}\sum_t Y_{t+1}^2 Y_t \times I_{\{Y_t > a_n\varepsilon\}} - Ca_n^{-3}\sum_t Y_t^3 \times I_{\{Y_t > a_n\varepsilon\}} + o_P(1) \\ &= T_{2,\varepsilon}^{+}(N_n) - CT_{3,0,\varepsilon}^{+}(N_n) + o_P(1), \end{aligned}$$

since

$$\begin{aligned} \underline{\alpha} a_n^{-3}\sum_t Y_t^3 \times I_{\{Y_t > a_n\varepsilon\}} + o_P(1) &\leq a_n^{-3}\sum_t \sigma_{t+1}^2 Y_t \times I_{\{Y_t > a_n\varepsilon\}} \\ &\leq \overline{\alpha} a_n^{-3}\sum_t Y_t^3 \times I_{\{Y_t > a_n\varepsilon\}} + o_P(1). \end{aligned}$$

As a consequence,

$$\begin{aligned} & a_n^{-3}\sum_t Y_{t+1}^2 Y_t \times I_{\{Y_t \geq 0\}} - E\left(Y_{t+1}^2 Y_t \times I_{\{Y_t \geq 0\}}\right) \qquad (26) \\ = \; & T_{2,\varepsilon}^{+}(N_n) - CT_{3,0,\varepsilon}^{+}(N_n) + \alpha_{1,0} a_n^{-3}\sum_t Y_t^3 \times I_{\{Y_t \geq 0\}} - E\left(Y_t^3 \times I_{\{Y_t \geq 0\}}\right) + o_P(1) \\ & \xrightarrow{d} V_1^{+} + \alpha_{1,0} W_0^{+}, \end{aligned}$$

where "$\xrightarrow{d}$" is first as $n \to \infty$ and then as $\varepsilon \to 0$ so that

$$n a_n^{-3}\left(\widehat{\gamma}_{(Y,\ Y^2)}^{+}(0) - \gamma_{(Y,\ Y^2)}^{+}(0)\right) \xrightarrow{d} V_1^{+} + \alpha_{1,0} W_0^{+} =: W_1^{+}.$$

Comparable arguments to those establishing (26) then also establish

$$\begin{aligned} & a_n^{-3}\sum_t Y_{t+1}^2 Y_t \times I_{\{Y_t < 0\}} - E\left(Y_{t+1}^2 Y_t \times I_{\{Y_t < 0\}}\right) \\ & \xrightarrow{d} V_1^{-} + \alpha_{2,0} W_0^{-} \end{aligned}$$

so that

$$n a_n^{-3}\left(\widehat{\gamma}_{(Y,\ Y^2)}^{-}(0) - \gamma_{(Y,\ Y^2)}^{-}(0)\right) \xrightarrow{d} V_1^{-} + \alpha_{2,0} W_0^{-} =: W_1^{-}. \qquad (27)$$

Extending (27) to higher lags (i.e., $m > 1$) is a continuation of the arguments given above. ■

LEMMA 7. *Let Assumptions A1*, A2 and A4* hold. For $m = 0, 1$ define*

$$\widehat{\gamma}_{Y^2}^{+}(m) = n^{-1}\sum_{t=1}^{n-m} Y_{t+m}^2 Y_t^2 \times I_{\{Y_t \geq 0\}}, \qquad \gamma_{Y^2}^{+}(m) = E\left(Y_m^2 Y_0^2 \times I_{\{Y_0 \geq 0\}}\right),$$

with $\widehat{\gamma}_{Y^2}^{-}(m)$ and $\gamma_{Y^2}^{-}(m)$ defined analogously using $I_{\{Y_t < 0\}}$. Then for a $\kappa_0 \in (4,\ 8)$,

$$n a_n^{-4}\left(\widehat{\gamma}_{Y^2}^{+}(m) - \gamma_{Y^2}^{+}(m)\right) \xrightarrow{d} \left(Q_m^{+}\right)_{m=0,1},$$

and
$$na_n^{-4}\left(\widehat{\gamma}^{-}_{Y^2}(m)-\gamma^{-}_{Y^2}(m)\right)\xrightarrow{d}\left(Q^{-}_m\right)_{m=0,1},$$

where
$$Q_1^{+}=U_1^{+}+\alpha_{1,0}Q_0^{+},\qquad Q_1^{-}=U_1^{-}+\alpha_{2,0}Q_0^{-},$$

jointly depend on U_1 from Proposition 1.

Proof. Following the notation introduced in the proof to Lemma 6, if $c_4^{+/-}=E\left(\epsilon_{t+1}^4\times I^{+/-}(1)\right)$, then

$$\begin{aligned}
& a_n^{-4}\sum_t Y_{t+1}^4\times I^{+/-}(1)-E\left(Y_{t+1}^4\times I^{+/-}(1)\right)\\
= \;& a_n^{-4}\sum_t\sigma_{t+1}^4\left(\epsilon_{t+1}^4\times I^{+/-}(1)-c_4^{+/-}\right)\times I_{\{|Y_t|\le a_n\varepsilon\}}\\
& +a_n^{-4}\sum_t\sigma_{t+1}^4\left(\epsilon_{t+1}^4\times I^{+/-}(1)-c_4^{+/-}\right)\times I_{\{|Y_t|>a_n\varepsilon\}}\\
& +c_4^{+/-}a_n^{-4}\sum_t\left(\sigma_{t+1}^4-E\left(\sigma_{t+1}^4\right)\right),
\end{aligned}$$

and

$$\begin{aligned}
& a_n^{-4}\sum_t Y_{t+1}^2Y_t^2\times I^{+/-}(0)-E\left(Y_{t+1}^2Y_t^2\times I^{+/-}(0)\right)\\
= \;& a_n^{-4}\sum_t\sigma_{t+1}^2Y_t^2\times I^{+/-}(0)\times\left(\epsilon_{t+1}^2-1\right)\times I_{\{|Y_t|\le a_n\varepsilon\}}\\
& +a_n^{-4}\sum_t\sigma_{t+1}^2Y_t^2\times I^{+/-}(0)\times\left(\epsilon_{t+1}^2-1\right)\times I_{\{Y_t>a_n\varepsilon\}}\\
& +a_n^{-4}\sum_t\sigma_{t+1}^2Y_t^2\times I^{+/-}(0)-E\left(\sigma_{t+1}^2Y_t^2\times I^{+/-}(0)\right).
\end{aligned}$$

Following the same, general, steps provided in the proof to Lemma 6 (while recognizing that σ_{t+1}^4 has an exact expression and, so, does not require a first-order Taylor approximation), it follows that
$$a_n^{-4}\sum_t Y_{t+1}^4\times I^{+}(1)-E\left(Y_{t+1}^4\times I^{+}(1)\right)\xrightarrow{d}\frac{U_0^{+}+c_4^{+}CU_0^{**}}{1-\left(c_4^{+}\alpha_{1,0}^2+c_4^{-}\alpha_{2,0}^2\right)}=:Q_0^{+},$$

where U_0^{**} is a component of U_1 in Proposition 1 and

$$a_n^{-4}\sum_t Y_{t+1}^2Y_t^2\times I^{+}(0)-E\left(Y_{t+1}^2Y_t^2\times I^{+}(0)\right)\xrightarrow{d}U_1^{+}+\alpha_{1,0}Q_0^{+}=:Q_1^{+}.$$

In addition, following from parallel arguments,

$$a_n^{-4}\sum_t Y_{t+1}^4\times I^{-}(1)-E\left(Y_{t+1}^4\times I^{-}(1)\right)\xrightarrow{d}\frac{U_0^{-}+c_4^{-}CU_0^{**}}{1-\left(c_4^{+}\alpha_{1,0}^2+c_4^{-}\alpha_{2,0}^2\right)}=:Q_0^{-}$$

and

$$a_n^{-4}\sum_t Y_{t+1}^2 Y_t^2 \times I^-(0) - E\left(Y_{t+1}^2 Y_t^2 \times I^-(0)\right) \xrightarrow{d} U_1^- + \alpha_{2,0} Q_0^- =: Q_1^-.$$

■

LEMMA 8. *For the ARCH(p) model, let Assumptions A1 and A2 hold. Then Assumption A7 is sufficient for* $E\left(\sigma_t^3\right) < \infty$.

Proof. The proof is by induction.

$$\begin{aligned}
\sigma_t^3 &\leq \sigma_t^2 \times \left(\omega_0^{1/2} + \sum_{i=1}^p \alpha_{i,0}^{1/2} \left|Y_{t-i}\right|\right) \\
&\leq \omega_0^{3/2} + \omega_0 \sum_{i=1}^p \alpha_{i,0}^{1/2} \left|Y_{t-i}\right| + \omega_0^{1/2} \sum_{i=1}^p \alpha_{i,0} Y_{t-i}^2 + \sum_{i=1}^p \sum_{j=1}^p \alpha_{i,0} \alpha_{j,0}^{1/2} Y_{t-i}^2 \left|Y_{t-j}\right|,
\end{aligned}$$

where the first inequality follows from the Triangle Inequality. Then, using Bollerslev (1986, Theorem 1),

$$\begin{aligned}
E\left(\sigma_t^3\right) &\leq C + \sum_{i=1}^p \sum_{j=1}^p \alpha_{i,0} \alpha_{j,0}^{1/2} E\left(Y_{t-i}^2 \left|Y_{t-j}\right|\right) \\
&\leq C + \sum_{i=1}^p \sum_{j=1}^p \alpha_{i,0} \alpha_{j,0}^{1/2} E\left|Y_{t-j}\right|^3 \\
&\leq C + c_3 \sum_{i=1}^p \sum_{j=1}^p \alpha_{i,0} \alpha_{j,0}^{1/2} E\left(\sigma_{t-j}^3\right)
\end{aligned}$$

From Lemma 1,

$$C + c_3 \alpha_{1,0}^{3/2} E\left(\sigma_{t-1}^3\right) \leq C + c_3 \alpha_{1,0}^{3/2} E\left(\sigma_t^3\right).$$

Suppose

$$C + c_3 \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} \alpha_{i,0} \alpha_{j,0}^{1/2} E\left(\sigma_{t-j}^3\right) \leq C + c_3 \left(\sum_{i=1}^{p-1} \sum_{j=1}^{p-1} \alpha_{i,0} \alpha_{j,0}^{1/2}\right) E\left(\sigma_t^3\right).$$

Then

$$\begin{aligned}
E\left(\sigma_t^3\right) &\leq C + c_3 \left(\sum_{i=1}^{p-1} \sum_{j=1}^{p-1} \alpha_{i,0} \alpha_{j,0}^{1/2}\right) E\left(\sigma_t^3\right) + c_3 \left(\sum_{i=1}^p \alpha_{i,0} \alpha_{p,0}^{1/2} + \sum_{j=1}^{p-1} \alpha_{i,0}^{1/2} \alpha_{p,0}\right) E\left(\sigma_{t-p}^3\right) \\
&\leq \overline{C} + D E\left(\sigma_{t-p}^3\right) \\
&\leq \overline{C}\left(1 + D + D^2 + \ldots\right) \\
&\leq \frac{\overline{C}}{1 - D} \\
&\leq \frac{C}{1 - c_3 \sum_{i=1}^p \sum_{j=1}^p \alpha_{i,0} \alpha_{j,0}^{1/2}}
\end{aligned}$$

■

LEMMA 9. *For the ARCH(p) model let Assumptions A1, A2 and A7 hold. Consider*

$$X_t = \mathbf{X}'_{t-1}\boldsymbol{\alpha}_0 + W_t \tag{28}$$

as it is defined in Section 2.3 of the main text and the set of instruments

$$\mathbf{Z}_{t-1} = \begin{pmatrix} Y_{t-1}, & \ldots, & Y_{t-h} \end{pmatrix}',$$

where, in this case, $h = p$. Given Assumption A3, $\mathbf{Z}_{t-1}$ identifies α_0.

Proof. The proof is by induction. When $p = 1$, $\mathbf{Z}_{t-1}$ identifies $\boldsymbol{\alpha}_0$ (see Section 2.1 in the main paper). From (28),

$$\begin{aligned} X_t &= \sum_{i=1}^{p-1} X_{t-i}\alpha_{i,0} + X_{t-p}\alpha_{p,0} + W_t \\ &= \widetilde{\mathbf{X}}'_{t-1}\widetilde{\boldsymbol{\alpha}}_0 + X_{t-p}\alpha_{p,0} + W_t. \end{aligned}$$

Let

$$\widetilde{\mathbf{Z}}_{t-1} = \begin{pmatrix} Y_{t-1}, & \ldots, & Y_{t-p+1} \end{pmatrix}',$$

and assume that $E\left(\widetilde{\mathbf{Z}}_{t-1}\widetilde{\mathbf{X}}'_{t-1}\right)$ is nonsingular. Then

$$\widetilde{\boldsymbol{\alpha}}_0 = E\left(\widetilde{\mathbf{Z}}_{t-1}\widetilde{\mathbf{X}}'_{t-1}\right)^{-1}\left[E\left(\widetilde{\mathbf{Z}}_{t-1}X_t\right) - E\left(\widetilde{\mathbf{Z}}_{t-1}X_{t-p}\right)\alpha_{p,0}\right]. \tag{29}$$

Further let

$$\begin{aligned} \mathbf{L}_0 &= E\left(Y_{t-p}\widetilde{\mathbf{X}}'_{t-1}\right)E\left(\widetilde{\mathbf{Z}}_{t-1}\widetilde{\mathbf{X}}'_{t-1}\right)^{-1}E\left(\widetilde{\mathbf{Z}}_{t-1}X_t\right), \\ \mathbf{M}_0 &= E\left(Y_{t-p}\widetilde{\mathbf{X}}'_{t-1}\right)E\left(\widetilde{\mathbf{Z}}_{t-1}\widetilde{\mathbf{X}}'_{t-1}\right)^{-1}E\left(\widetilde{\mathbf{Z}}_{t-1}X_{t-p}\right), \end{aligned}$$

noting that $\mathbf{M}_0$ is a scalar. Then given (29),

$$\alpha_{p,0} = \frac{E\left(Y_{t-p}X_t\right) - \mathbf{L}_0}{E\left(Y^3_{t-p}\right) - \mathbf{M}_0},$$

where $E\left(Y^3_{t-p}\right) - \mathbf{M}_0 \neq 0$ given A3 and Guo and Phillips (2001, Lemma 1). ■

LEMMA 10. *For the ARCH(p) model, let Assumptions A1, A2 and A7 hold. Then*

$$a_n^{-3}\sum_t \sigma_t^3 - E\left(\sigma_t^3\right) \xrightarrow{d} V_{0,\sigma}$$

when $\kappa_0 \in (3,\ 6)$, where $V_{0,\sigma}$ is $(\kappa_0/3)-$stable.

Proof.

$$
\begin{aligned}
& a_n^{-3} \sum_t \sigma_t^3 - E\left(\sigma_t^3\right) \\
= \; & a_n^{-3} \sum_t \left(\sigma_t^3 - E\left(\sigma_t^3\right)\right) \times I_{\{\sigma_t \leq a_n \varepsilon\}} \\
& + a_n^{-3} \sum_t \left(\sigma_t^3 - E\left(\sigma_t^3\right)\right) \times I_{\{\sigma_t > a_n \varepsilon\}} \\
= \; & Ia + IIa.
\end{aligned}
$$

Given Carrasco and Chen (2002, Proposition 12), $\{\sigma_t\}$ is strictly stationary. Then from Ia, given Lemma 8,

$$
\begin{aligned}
a_n^{-3} \sum_t E\left(\sigma_t^3\right) \times I_{\{\sigma_t \leq a_n \varepsilon\}} &= n^{\frac{\kappa_0 - 6}{2\kappa_0}} E\left(\sigma_t^3\right) n^{-1/2} \sum_t I_{\{\sigma_t \leq a_n \varepsilon\}} \\
& \longrightarrow 0,
\end{aligned}
$$

as $n \to \infty$ by the CLT in Ibragimov and Linnik (1971, Theorem 18.5.3), so that

$$
Ia = a_n^{-3} \sum_t \sigma_t^3 \times I_{\{\sigma_t \leq a_n \varepsilon\}} + o_p(1).
$$

Then, for a $\zeta > 0$,

$$
P\left(a_n^{-3} \sum_t \sigma_t^3 \times I_{\{\sigma_t \leq a_n \varepsilon\}} > \zeta\right) \leq \left(\zeta^{-1} a_n^{-3}\right) n E\left(\sigma^3 \times I_{\{\sigma \leq a_n \varepsilon\}}\right)
$$

by Markov's Inequality. Next, for the same r defined in the proof to Lemma 5, there exists a constant $C \in (0, \infty)$ such that

$$
\begin{aligned}
\left(\zeta^{-1} a_n^{-3}\right) n E\left(\sigma^3 \times I_{\{\sigma \leq a_n \varepsilon\}}\right) &\leq C \left(\zeta^{-1} a_n^{-3}\right)^r n E\left(\sigma^{3r} \times I_{\{\sigma \leq a_n \varepsilon\}}\right) \\
&\leq C \left(\frac{\kappa_0}{3r - \kappa_0}\right) \left(\zeta^{-1} a_n^{-3}\right)^r \left(a_n \varepsilon\right)^{3r} n P\left(\sigma > a_n \varepsilon\right),
\end{aligned}
$$

where the second inequality follows from the same arguments that support (12). As a consequence,

$$
\lim_{n \to \infty} \lim_{\varepsilon \to 0} \sup P\left(a_n^{-3} \sum_t \sigma_t^3 \times I_{\{\sigma_t \leq a_n \varepsilon\}} > \zeta\right) = 0,
$$

given the convergence results in (11). Next, since

$$
\begin{aligned}
a_n^{-3} \sum_t E\left(\sigma_t^3\right) \times I_{\{\sigma_t > a_n \varepsilon\}} &= n a_n^{-3} E\left(\sigma_t^3\right) n^{-1} \sum_t I_{\{\sigma_t > a_n \varepsilon\}} \\
&\sim a_n^{-3} E\left(\sigma_t^3\right) n P\left(\sigma_t > a_n \varepsilon\right) \\
& \longrightarrow 0,
\end{aligned}
$$

then

$$\begin{aligned} IIa &= a_n^{-3}\sum_t \sigma_t^3 \times I_{\{\sigma_t > a_n \varepsilon\}} + o_p(1) \\ &= T_{3,0,\epsilon}(N_n) + o_p(1) \end{aligned}$$

so that

$$a_n^{-3}\sum_t \sigma_t^3 - E\left(\sigma_t^3\right) \xrightarrow{d} T_{3,0,\epsilon}(N) \xrightarrow{d} V_{0,\sigma}$$

where the first "$\xrightarrow{d}$" is as $n \to \infty$ and the second as $\epsilon \to 0$. The first "$\xrightarrow{d}$" relies on Basrak, Davis, and Mikosch (2002, Corollary 3.5(B)) to establish $\left\{\left(\begin{array}{cc} Y_t, & \sigma_t \end{array}\right)\right\}$ as RV(κ_0) and Basrak, Davis, and Mikosch (2002, Theorem 2.10), which is a generalization of Lemma 4 to $\widetilde{\mathbf{Y}}_t$, since $\{\sigma_t\}$ is also strongly mixing given Carrasco and Chen (2002, Proposition 12). The second "$\xrightarrow{d}$", as is the case elsewhere in this Appendix, follows from Davis and Hsing (1995, Proof of Theorem 3.1, pp. 897-898). ■

LEMMA 11. *For the ARCH(p) model, let Assumptions A1, A2 and A7 hold. Then*

$$a_n^{-3}\sum_t Y_t^2 Y_{t+m} \xrightarrow{d} \left(R_{p,m}\right)_{m=1,\ldots,p},$$

when $\kappa_0 \in (3,\ 6)$.

Proof. To begin,

$$E\left(Y_t^2 Y_{t+m}\right) = E\left(Y_t^2 \sigma_{t+m} E\left(\epsilon_{t+m} \mid \digamma_{t-m+1}\right)\right) = 0.$$

Then,

$$\begin{aligned} a_n^{-3}\sum_t Y_t^2 Y_{t+m} &= a_n^{-3}\sum_t Y_t^2 Y_{t+m} \times I_{\{|Y_t| \le a_n \varepsilon\}} + a_n^{-3}\sum_t Y_t^2 Y_{t+m} I_{\{|Y_t| > a_n \varepsilon\}} \\ &= Ib + IIb \end{aligned}$$

For a $\zeta > 0$, using the same arguments that support the inequalities in (11),

$$\begin{aligned} P\left(|Ib| > \zeta\right) &\le \left(\zeta^{-1} a_n^{-3}\right)^r E\left|\sum_t Y_t^2 Y_{t+m} \times I_{\{|Y_t| \le a_n \varepsilon\}}\right|^r \qquad (30) \\ &\le \left(\zeta^{-1} a_n^{-3}\right)^r nE\left|Y_t^2 Y_{t+m} \times I_{\{|Y_t| \le a_n \varepsilon\}}\right|^r \\ &\le \left(\zeta^{-1} a_n^{-3}\right)^r nE\left(\left(\sigma_{t+m}^2\right)^{r/2} \times Y_t^{2r} \times I_{\{|Y_t| \le a_n \varepsilon\}} \times \left|\epsilon_{t+m}\right|^r\right) \\ &\le \left(\zeta^{-1} a_n^{-3}\right)^r nE\left(\left(\omega_0 + \sum_{i=1}^p \alpha_{i,0} Y_{t+m-i}^2\right)^{r/2} \times Y_t^{2r} \times I_{\{|Y_t| \le a_n \varepsilon\}}\right) \times E\left|\epsilon_{t+m}\right|^r \\ &\le C\left(\zeta^{-1} a_n^{-3}\right)^r nE\left(\left|Y_t\right|^{3r} \times I_{\{|Y_t| \le a_n \varepsilon\}}\right) \times E\left|\epsilon_{t+m}\right|^r, \end{aligned}$$

where in the final inequality, as is true elsewhere, the constant $C \in (0, \infty)$. Then,

$$\lim_{n\to\infty}\lim_{\varepsilon\to 0}\sup P\left(|Ib| > \zeta\right) = 0,$$

given (12) and the convergence results in (11). Next, building off of the definitions introduced in the proof of Lemma 5, consider

$$T_{m,\epsilon}^{(2)}\left(\sum_{i=1}^{\infty} n_i \delta_{\mathbf{x}_i}\right) = \sum_{i=1}^{\infty} n_i \left(x_i^{(0)}\right)^2 x_i^{(m-1)} I_{\left\{\left|x_i^{(0)}\right|>\varepsilon\right\}}, \qquad m \geq 2.$$

Then

$$\begin{aligned} a_n^{-3}\sum_t Y_t^2 Y_{t+m} &= Ib + T_{m,\epsilon}^{(2)}\left(N_n\right) \\ &\xrightarrow{d} T_{m,\epsilon}^{(2)}\left(N\right) \\ &\xrightarrow{d} R_{p,m}, \end{aligned}$$

where "$\xrightarrow{d}$" is as $n \to \infty$ first, and then as $\epsilon \to 0$. As for Lemma 10, the first "$\xrightarrow{d}$" relies on Basrak, Davis, and Mikosch (2002, Corollary 3.5(B) and Theorem 2.10) and the continuous mapping theorem. As is true elsewhere in this Appendix, the second "$\xrightarrow{d}$" follows from Davis and Hsing (1995, Proof of Theorem 3.1, pp. 897-898). ■

LEMMA 12. *For the ARCH(p) model, let Assumptions A1, A2 and A7 hold. Then, given the definitions of $\widehat{\gamma}_{(Y,\ Y^2)}(m)$ and $\gamma_{(Y,\ Y^2)}(m)$ in Lemma 5,*

$$na_n^{-3}\left(\widehat{\gamma}_{(Y,\ Y^2)}(m) - \gamma_{(Y,\ Y^2)}(m)\right) \xrightarrow{d} \left(V_{p,m}\right)_{m=0,\ldots,h} \tag{31}$$

for a $\kappa_0 \in (3,\ 6)$, where $V_{p,0} := V_{p,0}^* + c_3^* V_{0,\sigma}$, and $V_{p,m} := V_{p,m}^* - \alpha_{1,0} V_{p,m-1}$.

Proof. Begin by considering the following modification to (10)

$$\begin{aligned} & a_n^{-3}\sum_t \left(Y_{t+1}^3 - E\left(Y_{t+1}^3\right)\right) \\ = \; & a_n^{-3}\sum_t \sigma_{t+1}^3\left(\epsilon_{t+1}^3 - c_3^*\right) \times I_{\left\{\sigma_{t+1}\leq a_n\varepsilon\right\}} \\ & + a_n^{-3}\sum_t \sigma_{t+1}^3\left(\epsilon_{t+1}^3 - c_3^*\right) \times I_{\left\{\sigma_{t+1}> a_n\varepsilon\right\}} \\ & + c_3^* a_n^{-3}\sum_t \left(\sigma_{t+1}^3 - E\left(\sigma_{t+1}^3\right)\right) \\ = \; & Ia + IIa + IIIa \end{aligned}$$

introduced to deal with the complications posed by a multi-lag parameterization of σ_{t+1}^2. From this decomposition, for a $\zeta > 0$,

$$\lim_{n\to\infty}\lim_{\varepsilon\to 0}\sup P\left(|Ia| > \zeta\right) = 0,$$

given the arguments that support (11). Next,

$$\begin{aligned} IIa &= a_n^{-3}\sum_t Y_{t+1}^3 \times I_{\left\{|Y_{t+1}|>a_n\varepsilon\right\}} - c_3^* a_n^{-3}\sum_t \sigma_{t+1}^3 \times I_{\left\{\sigma_{t+1}>a_n\varepsilon\right\}} + o_P(1) \\ &= T_{3,0,\epsilon}(N_n) - c_3^* T_{3,0,\epsilon}^*(N_n) + o_P(1) \end{aligned}$$

where the first equality follows from Basrak, Davis and Mikosch (2002, proof of Theorem 3.6), and $T_{3,0,\epsilon}^*(N_n)$ denotes that N_n is defined in terms of σ_{t+m}, while $T_{3,0,\epsilon}(N_n)$ retains its definition from the proof of Lemma 5, where N_n is a function of Y_{t+m}. As a result,

$$\begin{aligned} a_n^{-3}\sum_t \left(Y_{t+1}^3 - E\left(Y_{t+1}^3\right)\right) &= T_{3,0,\epsilon}(N_n) - c_3^* T_{3,0,\epsilon}^*(N_n) + IIIa + o_P(1) \\ &\xrightarrow{d} V_{p,0}^* + c_3^* V_{0,\sigma}, \end{aligned} \tag{32}$$

where "$\xrightarrow{d}$" is as $n \to \infty$ first, and then as $\epsilon \to 0$. Here, "$\xrightarrow{d}$" follows from Basrak, Davis, and Mikosch (2002, Corollary 3.5(B) and Theorem 2.10), Lemma 10, and Davis and Hsing (1995, Theorem 3.1, pp. 897-898) and grants that

$$na_n^{-3}\left(\widehat{\gamma}_{(Y,\,Y^2)}(0) - \gamma_{(Y,\,Y^2)}(0)\right) \xrightarrow{d} V_{p,0} := V_{p,0}^* + c_3^* V_{0,\sigma}. \tag{33}$$

Consider next the decomposition in (20). From this decomposition,

$$\begin{aligned} P\left(|Ic| > \zeta\right) &\leq 2\left(\zeta^{-1}a_n^{-3}\right)^r nE\left|Y_t\sigma_{t+1}^2 \times I_{\{|Y_t|\leq a_n\varepsilon\}}\right|^r \times E\left|\epsilon_{t+1}^2 - 1\right|^r \\ &\leq 2\left(\zeta^{-1}a_n^{-3}\right)^r nE\left(|Y_t|^r\left(\omega_0 + \sum_{i=1}^p \alpha_{i,0}Y_{t+1-i}^2\right)^r \times I_{\{|Y_t|\leq a_n\varepsilon\}}\right) \times E\left|\epsilon_{t+1}^2 - 1\right|^r \\ &\leq 2C\left(\zeta^{-1}a_n^{-3}\right)^r nE\left(|Y_t|^{3r} \times I_{\{|Y_t|\leq a_n\varepsilon\}}\right) \times E\left|\epsilon_{t+1}^2 - 1\right|^r \end{aligned}$$

using similar arguments to those that support (30). As a consequence, as is true elsewhere,

$$\lim_{n\to\infty}\lim_{\varepsilon\to 0}\sup P\left(|Ic| > \zeta\right) = 0,$$

given (12) and the convergence results in (11). Next,

$$\begin{aligned} IIc &= a_n^{-3}\sum_t Y_t Y_{t+1}^2 \times I_{\{|Y_t|>a_n\varepsilon\}} - \alpha_{1,0}a_n^{-3}\sum_t Y_t^3 \times I_{\{|Y_t|>a_n\varepsilon\}} \\ &\quad - a_n^{-3}\sum_t\sum_{i=2}^p \alpha_{i,0}Y_tY_{t+1-i}^2 \times I_{\{|Y_t|>a_n\varepsilon\}} + o_P(1) \\ &= T_{2,\epsilon}^{(1)}(N_n) - \alpha_{1,0}T_{3,0,\epsilon}(N_n) - \sum_{i=2}^p \alpha_{i,0}T_{i,\epsilon}^{(2)}(N_n) + o_P(1). \end{aligned}$$

Finally,

$$IIIc = \alpha_{1,0} a_n^{-3} \sum_t Y_t^3 - E\left(Y_t^3\right) + a_n^{-3} \sum_t \sum_{i=2}^{p} \alpha_{i,0} Y_t Y_{t+1-i}^2 + o_P(1),$$

so that

$$\begin{aligned} a_n^{-3} \sum_t Y_t Y_{t+1}^2 - E\left(Y_t Y_{t+1}^2\right) &= Ic + T_{2,\epsilon}^{(1)}(N_n) - \alpha_{1,0} T_{3,0,\epsilon}(N_n) - \sum_{i=2}^{p} \alpha_{i,0} T_{i,\epsilon}^{(2)}(N_n) \\ &\quad + IIIc + o_P(1) \\ &\xrightarrow{d} V_{p,1}^* + \alpha_{1,0} V_{p,0}, \end{aligned}$$

where "$\xrightarrow{d}$" is with respect to $n \to \infty$ first (following from the same arguments that support convergence as $n \to \infty$ in (32) and Lemma 11) and $\epsilon \to 0$ second (as established elsewhere in this appendix) so that

$$n a_n^{-3} \left(\widehat{\gamma}_{(Y,\ Y^2)}(1) - \gamma_{(Y,\ Y^2)}(1) \right) \xrightarrow{d} V_{p,1} := V_{p,1}^* + \alpha_{1,0} V_{p,0}. \tag{34}$$

Extending (34) to higher lags (i.e., $m > 1$) is a continuation of the arguments given above. ■

OLS Estimation of the ARCH(1) Model

Recall that

$$Y_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 = \omega_0 + \alpha_0 Y_{t-1}^2$$

implies the second-order (centered) AR(1) model of

$$X_t = \alpha_0 X_{t-1} + W_t, \tag{35}$$

where $X_t \equiv Y_t^2 - \gamma_0$ and $\gamma_0 \equiv E\left(Y_t^2\right) = \frac{\omega_0}{1-\alpha_0}$.

ASSUMPTION A1*: Under A1(i), let $E\left|\epsilon_t\right|^j = c_j < \infty$ *for* $j > 4$.

A1* strengthens A1 from the main paper.

ASSUMPTION A4*: $E\left(A^l\right) < 1$ *for* $l \geq 2$.

A4* strengthens A4 from the main paper. Given A4* with $l = 2$,

$$E\left(X_t X_{t-m}\right) = \alpha_0^m E\left(X_t^2\right), \qquad m \geq 1, \tag{36}$$

so that OLS estimators for α_0 and ω_0 are

$$\widehat{\alpha}^{OLS} = \frac{\sum_t \widehat{X}_t \widehat{X}_{t-1}}{\sum_t \widehat{X}_{t-1}^2}, \tag{37}$$

$$\widehat{\omega}^{OLS} = \widehat{\gamma}\left(1 - \widehat{\alpha}^{OLS}\right). \tag{38}$$

Versions of (37) were first studied by Weiss (1986) and more recently by Guo and Phillips (2001).

PROPOSITION 1. *Consider the estimators in (37) and (38) for the model of (35). Let Assumptions A1*, A2, and A4* with $l = 2$ hold. Then*

$$\widehat{\alpha}^{OLS} \xrightarrow{a.s.} \alpha_0, \qquad \widehat{\omega}^{OLS} \xrightarrow{a.s.} \omega_0.$$

In addition,

$$na_n^{-4}\left(\widehat{\alpha}^{OLS} - \alpha_0\right) \xrightarrow{d} E\left(X_{t-1}^2\right)^{-1} U_1 \tag{39}$$

if $\kappa_0 \in (4,\ 8)$, *where* U_1 *is* $(\kappa_0/4)-$*stable, and*

$$na_n^{-4}\left(\widehat{\omega}^{OLS} - \omega_0\right) = -\gamma_0 na_n^{-4}\left(\widehat{\alpha}^{OLS} - \alpha_0\right) + o_p(1). \tag{40}$$

Alternatively, if Assumption A4 with* $l = 4$ *holds so that* $E\left(Y_t^8\right) < 8$ and $\kappa_0 \in (8,\ \infty)$, *then*

$$\sqrt{n}\left(\widehat{\alpha}^{OLS} - \alpha_0\right) \xrightarrow{d} N\left(0,\ E\left(X_{t-1}^2\right)^{-2} E\left(W_t^2 X_{t-1}^2\right)\right), \tag{41}$$

and

$$\sqrt{n}\left(\widehat{\omega}^{OLS} - \omega_0\right) \xrightarrow{d} N\left(0,\ \Sigma_{\omega_0}\right), \tag{42}$$

where

$$\Sigma_{\omega_0} = \Sigma_{\gamma_0} + E\left(X_{t-1}^2\right)^{-1}\left(\gamma_0^2 E\left(X_{t-1}^2\right)^{-1} E\left(W_t^2 X_{t-1}^2\right) - 2\sum_{s=1}^{\infty} E\left(W_t X_{t-1} Y_{t-s}^2\right)\right). \tag{43}$$

Proof. Recall that

$$\widehat{X}_t = X_t - \left(\widehat{\gamma} - \gamma_0\right), \tag{44}$$

and

$$\widehat{X}_t = \overline{c} + \alpha_0 \widehat{X}_{t-1} + W_t. \tag{45}$$

Given (44) and (45),

$$\widehat{\alpha}^{OLS} = \alpha_0 + \left(\sum_t \widehat{X}_{t-1}^2\right)^{-1}\left(\overline{c}\sum_t \widehat{X}_{t-1} - \left(\widehat{\gamma} - \gamma_0\right)\sum_t W_t + \sum_t W_t X_{t-1}\right). \tag{46}$$

Then $\widehat{\alpha}^{OLS} \overset{a.s.}{\to} \alpha_0$, and $\widehat{\omega}^{OLS} \overset{a.s.}{\to} \omega_0$ given the same arguments that establish consistency in the proof of Theorem 1 (see the main paper's Appendix). Next, given (44),

$$\begin{aligned} na_n^{-4}\left(\widehat{\alpha}^{OLS} - \alpha_0\right) &= E\left(X_{t-1}^2\right)^{-1}\left(a_n^{-4}\sum_t X_t X_{t-1} - E\left(X_t X_{t-1}\right)\right) + o_P(1) \qquad (47) \\ &\xrightarrow{d} E\left(X_{t-1}^2\right)^{-1} U_1, \end{aligned}$$

given Lemmas 2 and 3, Davis and Mikosch (1998), and von Bahr and Essen (1965, Theorem 2), where application of the latter permits $j \in (4,\ 8)$ in A1*.[4] Comparable to Theorem 1, this (weak) distributional convergence results relies on

$$a_n^{-4} \sum_t X_t X_{t-1} - E\left(X_t X_{t-1}\right) = a_n^{-4} \sum_t Y_t^2 Y_{t-1}^2 - E\left(Y_t^2 Y_{t-1}^2\right) + o_P(1)$$

since

$$a_n^{-4} \sum_t Y_t^2 - \gamma_0 = n^{\frac{\kappa_0 - 8}{2\kappa_0}} \left(n^{-1/2} \sum_t Y_t^2 - \gamma_0\right) \xrightarrow{d} 0 \tag{48}$$

by Ibragimov and Linnik (1971, Theorem 18.5.3). Also given (48),

$$n a_n^{-4} \left(\widehat{\omega}^{OLS} - \omega_0\right) = -\gamma_0 n a_n^{-4} \left(\widehat{\alpha}^{OLS} - \alpha_0\right) + o_P(1).$$

Finally, if $\kappa_0 \in (8,\ \infty)$, then given (46),

$$\begin{aligned}
\sqrt{n}\left(\widehat{\alpha}^{OLS} - \alpha_0\right) &= E\left(X_{t-1}^2\right)^{-1} \left(n^{-1/2} \sum_t W_t X_{t-1}\right) + o_P(1) \\
&\xrightarrow{d} N\left(0,\ E\left(X_{t-1}^2\right)^{-2} E\left(W_t^2 X_{t-1}^2\right)\right)
\end{aligned}$$

by Ibragimov and Linnik and the Slutsky Theorem, and

$$\begin{aligned}
\sqrt{n}\left(\widehat{\omega}^{OLS} - \omega_0\right) &= \sqrt{n}\left(\widehat{\gamma} - \gamma_0\right) - \gamma_0 \sqrt{n}\left(\widehat{\alpha}^{OLS} - \alpha_0\right) \\
&\xrightarrow{d} N\left(0,\ \Sigma_{\omega_0}\right)
\end{aligned}$$

where Σ_{ω_0} is defined in Theorem 1 of the main paper. ■

The OLS estimator in (37) depends on the first (sample) second-order autocovariance from (36). The resulting (weak) distributional limit in (39) follows immediately from Davis and Mikosch (1998) if $c_3^* = 0$, and $j = 8$ in A1. Under Proposition 1, in contrast, the asymptotic properties of $\widehat{\alpha}^{OLS}$ are unaffected by whether or not A3 holds. Moreover, given von Bahr and Esseen (1965, Theorem 2), $j \in (4,\ 8)$, instead, supports (39). The distribution of U_1 is similar to that of V_1 in Theorem 1 of the main paper but, nonetheless, is distinct because the former is based on fourth-order mixtures of Poisson and i.i.d. point processes (see Lemma 4 and Remark R3, as well as Davis and Hsing, 1995, Theorem 3.1), while the latter depends on third-order mixtures of these same processes. The general method of proof behind Proposition 1 and Theorem 1 in the main paper is analogous. Asymptotic normality under Proposition 1 mirrors Weiss (1986, Theorem 4.4). The heavy-tailed case of (39), where the rate of convergence is $n^{\frac{\kappa_0 - 4}{\kappa_0}}$, is closely related to Kristensen and Linton (2006, Theorem 2).

It is important to note that if $\kappa_0 \in (4,\ 8)$ and A3 holds, then $\widehat{\alpha}^{IV}$ in the main paper converges

[4] Application of von Bahr and Esseen (1965, Theorem 2) in this instance closely mirrors that in the proof of Lemma 5.

at a faster rate than does $\widehat{\alpha}^{OLS}$. Also, if $\kappa_0 \in (4,\ 8)$, then for

$$\widehat{\tau}_n^2 = n^{-1} \sum_t Y_t^8, \qquad n a_n^{-8} \widehat{\tau}_n^2 \xrightarrow{d} \widetilde{S}_0, \tag{49}$$

where $\widetilde{S}_0$ is $(\kappa_0/8)$ −stable (see Davis and Mikosch, 1998, Section 4B(1), for a closely-related result). As a consequence, normalizing the left-hand-side of (39) by $\widehat{\tau}_n$ enables inference on $\widehat{\alpha}^{OLS}$ to be conducted using the subsampling and bootstrapping methods discussed above in the context of Theorem 1 in the main paper. Lastly, the borderline case of $\kappa_0 = 8$ is not considered for the same reason that $\kappa_0 = 6$ is excluded from consideration in Theorem 1 in the main paper.

OLS Estimation of the GJR ARCH(1) Model

Recall that

$$Y_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 = \omega_0 + \alpha_{1,0} Y_{t-1}^2 \times I_{\{Y_{t-1} \geq 0\}} + \alpha_{2,0} Y_{t-1}^2 \times I_{\{Y_{t-1} < 0\}}$$

implies

$$\begin{aligned} X_t &= \alpha_{1,0} X_{1,t-1} + \alpha_{2,0} X_{2,t-1} + W_t \\ &= \mathbf{X}_{t-1}' \boldsymbol{\alpha}_0 + W_t, \end{aligned} \tag{50}$$

where $X_t \equiv Y_t^2 - \gamma_0$ and $\gamma_0 \equiv E\left(Y_t^2\right)$ as before, with

$$E\left(Y_t^2\right) = \frac{\omega_0 + \alpha_{1,0} Cov\left(Y_t^2, I_{\{Y_t \geq 0\}}\right) + \alpha_{2,0} Cov\left(Y_t^2, I_{\{Y_t < 0\}}\right)}{1 - \left(\alpha_{1,0} \times P\left(Y_t \geq 0\right) + \alpha_{2,0} \times P\left(Y_t < 0\right)\right)},$$

and

$$X_{1,t-1} = Y_{t-1}^2 \times I_{\{Y_{t-1} \geq 0\}} - E\left(Y_t^2 \times I_{\{Y_t \geq 0\}}\right), \qquad X_{2,t-1} = Y_{t-1}^2 \times I_{\{Y_{t-1} < 0\}} - E\left(Y_t^2 \times I_{\{Y_t < 0\}}\right).$$

ASSUMPTION A6*: $E\left(\mathbf{X}_{t-1}\mathbf{X}_{t-1}'\right)$ is nonsingular.

A6* is the analog to A6 in the main paper. It serves as the key identifying condition for the following OLS estimator:

$$\widehat{\boldsymbol{\alpha}}^{OLS} = \widehat{\mathbf{K}} \left(n^{-1} \sum_t \widehat{X}_t \widehat{\mathbf{X}}_{t-1} \right), \qquad \widehat{\mathbf{K}} = \left(n^{-1} \sum_t \widehat{\mathbf{X}}_{t-1} \widehat{\mathbf{X}}_{t-1}' \right)^{-1}. \tag{51}$$

PROPOSITION 2. *Consider the estimator in (51) for the model in (50), and let* $\mathbf{K}_0 = E\left(\mathbf{X}_{t-1}\mathbf{X}_{t-1}'\right)^{-1}$.
In addition, let Assumptions A1, A2, A4* with $l = 2$, and A6* hold. Then,*

$$\widehat{\boldsymbol{\alpha}}^{OLS} \xrightarrow{a.s.} \boldsymbol{\alpha}_0.$$

In addition,

$$na_n^{-4}\left(\widehat{\boldsymbol{\alpha}}^{OLS}-\boldsymbol{\alpha}_0\right)\xrightarrow{d}\mathbf{K}_0\mathbf{Q}_1^{(+,-)}\tag{52}$$

if $\kappa_0\in(4,\ 8)$, *where the vector* $Q_1^{(+,-)}=\left(\begin{array}{cc}Q_1^{+}, & Q_1^{-}\end{array}\right)'$ *is jointly* $(\kappa_0/4)-$*stable with components* Q_1^{+} *and* Q_1^{-} *defined in Lemma 7, if A4* with* $l=4$ *holds so that* $E\left(Y_t^8\right)<8$ *and* $\kappa_0\in(8,\ \infty)$, *then*

$$\sqrt{n}\left(\widehat{\boldsymbol{\alpha}}^{OLS}-\boldsymbol{\alpha}_0\right)\xrightarrow{d}N\left(0,\ \mathbf{K}_0E\left(W_t^2\mathbf{X}_{t-1}\mathbf{X}_{t-1}'\right)\mathbf{K}_0'\right).\tag{53}$$

Proof. From (51), using the expressions for $\widehat{\mathbf{X}}_{t-1}$ and $\widehat{X}_t$ as they relate to $\mathbf{X}_{t-1}$ and W_t, respectively (see the proof of Theorem 2 in the Appendix of the main paper),

$$\widehat{\boldsymbol{\alpha}}^{OLS}-\boldsymbol{\alpha}_0=\widehat{\mathbf{K}}\left[\bar{c}\left(n^{-1}\sum_t\mathbf{X}_{t-1}\right)+\left(\widehat{\mathbf{G}}-\mathbf{G}_0\right)\left(n^{-1}\sum_t W_t-1\right)+n^{-1}\sum_t\mathbf{X}_{t-1}W_t\right].\tag{54}$$

Then, given A6*, $\widehat{\boldsymbol{\alpha}}^{OLS}\xrightarrow{a.s.}\boldsymbol{\alpha}_0$ follows from the same arguments that establish (almost sure) consistency in the proof of Theorem 2. Next, let $\widehat{\mathbf{X}}_{t-1}=\mathbf{Z}_{t-1}^{(2)}-\mathbf{G}_0$. In the case where $\kappa_0\in(4,\ 8)$, consider

$$\begin{aligned}
na_n^{-4}\left(\widehat{\boldsymbol{\alpha}}^{OLS}-\boldsymbol{\alpha}_0\right) &= \mathbf{K}_0\left[a_n^{-4}\sum_t\mathbf{X}_{t-1}X_t-E\left(\mathbf{X}_{t-1}X_t\right)\right]+o_p(1)\\
&= \mathbf{K}_0\left[a_n^{-4}\sum_t\mathbf{Z}_{t-1}^{(2)}Y_t^2-E\left(\mathbf{Z}_{t-1}^{(2)}Y_t^2\right)\right]\\
&\quad -n^{\frac{\kappa_0-8}{2\kappa_0}}\left[\mathbf{G}_0n^{-1}\sum_t Y_t^2-E\left(Y_t^2\right)+\gamma_0n^{-1}\sum_t\mathbf{X}_{t-1}\right]+o_p(1)\\
&= \mathbf{K}_0\left[a_n^{-4}\sum_t\mathbf{Z}_{t-1}^{(2)}Y_t^2-E\left(\mathbf{Z}_{t-1}^{(2)}Y_t^2\right)\right]+o_p(1)\\
&\xrightarrow{d}\mathbf{K}_0\mathbf{Q}_1^{(+,-)},
\end{aligned}$$

where $\mathbf{Q}_1^{(+,-)}=\left(\begin{array}{cc}Q_1^{+}, & Q_1^{-}\end{array}\right)$; the third equality follows from the CLT of Ibragimov and Linnik (1971, Theorem 18.5.3), and (weak) convergence in distribution to a $(\kappa_0/4)-$stable limit follows from Lemma 7 and Samorodnitsky and Taqqu (1994, Theorem 2.1.5(c)). Finally, if $\kappa_0\in(8,\ \infty)$, then given (54), (53) follows along the same lines as given in the proof to Theorem 2. ■

Proposition 2 extends results from Davis and Mikosch (1998) to the GJR ARCH(1) model. Necessary for the proof of Proposition 2 is establishing the (weak) distributional limit of $n^{-1}\sum_t X_t\mathbf{X}_{t-1}$, (see Lemma 7). Given (49), normalizing the left-hand-side of (52) by $\widehat{\tau}_n$ produces

$$\sqrt{n}\left(\frac{\widehat{\boldsymbol{\alpha}}^{OLS}-\boldsymbol{\alpha}_0}{\widehat{\tau}_n}\right)\xrightarrow{d}\frac{\mathbf{K}_0\mathbf{Q}_1^{(+,-)}}{\widetilde{S}_0^{1/2}},$$

in which case, subsample and bootstrap confidence intervals for $\widehat{\boldsymbol{\alpha}}^{OLS}$ can also result as in the discussion that follows Proposition 1. Like Proposition 1, Proposition 2 does not require D in A1* to be symmetric. As a result, Proposition 2 can also apply to the same processes towards which Theorem 2 in the main paper is directed; provided (of course) that the requisite higher moments are well defined. In cases where $\kappa_0 \in (4,\ 6)$, however, $\widehat{\boldsymbol{\alpha}}^{IV}$ in Theorem 2 converges at a faster rate (although, to a different and stable distribution) than does $\widehat{\boldsymbol{\alpha}}^{OLS}$, and when $\kappa_0 \in [6,\ 8)$, $\widehat{\boldsymbol{\alpha}}^{IV}$ is $\sqrt{n}$ asymptotically normal. Moreover, and in contrast to the convergence rate differentials discovered between $\widehat{\alpha}^{IV}$ in Theorem 1 of the main paper and $\widehat{\alpha}^{OLS}$ in Proposition 1, improvements in the rate of convergence enjoyed by $\widehat{\boldsymbol{\alpha}}^{IV}$ over $\widehat{\boldsymbol{\alpha}}^{OLS}$ do not, necessarily, rely on skewness in the model's rescaled errors.

OLS Estimation of the ARCH(p) Model

Given

$$Y_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 = \omega_0 + \sum_{i=1}^{p} \alpha_{i,0} Y_{t-i}^2, \qquad 1 \le p < \infty,$$

the generalization of (35) is

$$X_t = \mathbf{X}_{t-1}' \boldsymbol{\alpha}_0 + W_t,$$

where $\boldsymbol{\alpha}_0 = \begin{pmatrix} \alpha_{1,0}, & \ldots, & \alpha_{p,0} \end{pmatrix}'$, and

$$\mathbf{X}_{t-1} = \begin{pmatrix} X_{t-1}, & \ldots, & X_{t-p} \end{pmatrix}'. \tag{55}$$

If A9 with $s = 2$ holds, then (51) with $\widehat{\mathbf{X}}_{t-1}$ defined as the feasible version of (55) is a (almost surely) consistent estimator of $\boldsymbol{\alpha}_0$ following the same method of proof for Proposition 2. Moreover, following the same method of proof for Lemmas 9–12, it can further be established that

$$n a_n^{-4} \left(\widehat{\boldsymbol{\alpha}}^{OLS} - \boldsymbol{\alpha}_0 \right) \xrightarrow{d} \mathbf{K}_0 \mathbf{U}_{p,p},$$

where the vector $\mathbf{U}_{p,p} = \begin{pmatrix} U_{p,1}, & \ldots, & U_{p,p} \end{pmatrix}'$ is jointly $(\kappa_0/4)$ −stable, reduces to U_1 from (39) in the special case where $p = 1$, but generally is not solely determined by functionals of the observable sequence $\{Y_t\}$. If A9 with $s = 4$ holds, then (53) is established following the same method of proof for Proposition 2 and echoes the result of Weiss (1986, Theorem 4.4). Confidence intervals for $\widehat{\boldsymbol{\alpha}}^{OLS}$ can be constructed from $\sqrt{n}\left(\frac{\widehat{\boldsymbol{\alpha}}^{OLS} - \boldsymbol{\alpha}_0}{\widehat{\tau}_n}\right)$ using (49), given either the subsample or bootstrap method discussed above in the context of Proposition 1.

References

[1] Basrak, B., R.A Davis & T. Mikosch (2002) Regular variation of GARCH processes. *Stochastic Processes and Their Applications* 99, 95-115.

[2] Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307-327.

[3] Carrasco, M. & X. Chen (2002) Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory* 18, 17-39.

[4] Davis, R.A. & T. Hsing (1995) Point process and partial sum convergence for weakly dependent random variables with infinite variance. *The Annals of Probability* 23, 879-917.

[5] Davis, R.A. & T. Mikosch (1998) The sample autocorrelations of heavy-tailed processes with applications to ARCH. *The Annals of Statistics* 26, 2049-2080.

[6] Davis, R.A., T. Mikosch and B. Basrak (1999) Sample ACF of multivariate stochastic recurrence equations with application to GARCH. Unpublished manuscript.

[7] Goldie, C.M. (1991) Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.* 1, 126-166.

[8] Guo, B. & P.C.B Phillips (2001) Efficient estimation of second moment parameters in ARCH models. Unpublished manuscript.

[9] Ibragimov, I.A. and Y.V. Linnik (1971) *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff: Groningen.

[10] Kristensen, D. & O. Linton (2006) A closed form estimator for the GARCH(1,1)-model. *Econometric Theory* 22, 323-327.

[11] Mikosch, T. (1999) Regular Variation, Subexponentiality and their applications in probability theory. Lecture notes for the workshop "Heavy Tails and Queques," EURANDOM, Eindhoven, Netherlands.

[12] Mikosch, T. & C. Stărică (2000) Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process. *The Annals of Statistics* 28, 1427-1451.

[13] Resnick, S.I. (2007) *Probabilistic and Statistical Modeling of Heavy Tailed Phenomena*. New York: Springer-Verlag.

[14] Vaynman, I. & B.K. Beare (2014) Stable limit theory for the variance targeting estimator, in Y. Chang, T.B. Fomby & J.Y. Park (eds), *Essays in Honor of Peter C.B. Phillips,* vol. 33 of *Advances in Econometrics:* Emerald Group Publishing Limited, chapter 24, 639-672.

[15] von Bahr, B., & C.G. Esseen (1965) Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Annals of Mathematical Statistics* 36, 299-303.

[16] Weiss, A.A. (1986) Asymptotic theory for ARCH models: estimation and testing. *Econometric Theory* 2, 107-131.