When Simplicity Offers a Benefit, Not a Cost: Closed-Form Estimation of the GARCH(1,1) Model that Enhances the Efficiency of Quasi-Maximum Likelihood

Supplemental Appendix
Todd Prono

A.1.1. Preliminaries

Contained in this Supplemental Appendix are both the statements and proofs of all Lemmas that support the paper’s main theorems. Concerning notation, $C$ denotes a generic constant that can assume different values in different places. For matrices $A$ and $B$, $A \geq B$ means that every element in $A$ $\geq$ every corresponding element in $B$. For a vector $y$, $\delta_y$ denotes the Dirac measure at $y$. For a random variable $X > 0$ with CDF $F_X(x)$, where $F_X(x) = 1 - F_X(x)$, if

$$\lim_{x \to \infty} \frac{F_X(tx)}{F_X(x)} = t^{-\kappa_0}, \quad \forall \ t > 0, \quad \kappa_0 \geq 0,$$

then $X$ is regularly varying with tail index $\kappa_0$. Finally, RV($\kappa_0$) is shorthand for regularly varying with tail index $\kappa_0$.

**Proposition 1** For a random variable $X > 0$, assume (1) holds. Then for a $p > 0$, $X^p$ is regularly varying with tail index $\kappa_0/p$.

**Proof.** Let $Y = X^p$. Then

$$\frac{F_Y(ty)}{F_Y(y)} = \frac{P(Y > ty)}{P(Y > y)} = \frac{P(Y^{1/p} > t^{1/p}y^{1/p})}{P(Y^{1/p} > y^{1/p})} = \frac{P(X > t^{1/p}x)}{P(X > x)} = \frac{F_X(t^{1/p}x)}{F_X(x)} = \frac{F_X(Bx)}{F_X(x)};$$

in which case, by (1),

$$\lim_{y \to \infty} \frac{F_Y(ty)}{F_Y(y)} = \lim_{x \to \infty} \frac{F_X(Bx)}{F_X(x)} = B^{-\kappa_0} = t^{-\kappa_0/p}.$$
A.1.2. Regular Variation

Consider the model
\[ Y_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim \text{i.i.d. } D(0, 1) \quad (2) \]
\[ \sigma_t^2 = \omega_0 + \alpha_0 Y_{t-1}^2 + \beta_0 \sigma_{t-1}^2, \quad (3) \]
where \( D \) is some unknown distribution. This model is the linear GARCH(1, 1) model of Bollerslev (1986). Also note that from (3),
\[ \sigma_t^2 = \omega_0 + \beta_0 A_t, \quad (4) \]

Lemma 2 For the model in (2) and (3), let Assumptions A1–A2 and A4 hold. Then
\[ E\left( A^{\kappa/2} \right) = 1 \quad (5) \]
has a unique and positive solution \( \kappa_0 \).

Proof. For a \( \kappa > 0 \), \( E\left( A^{\kappa} \right) \) is a continuous and convex (upwards) function of \( \kappa \). Since Assumption A4 is sufficient for \( E\left( A \right) < 1 \),

CONDITION C1: \( E\left( A^{\kappa} \right) < 1 \) for values of \( \kappa \) in some neighborhood of one.

Also, since \( P\left( A > 1 \right) > 0 \), and since there exists a value \( \bar{\kappa} \) of \( \kappa \) such that \( E\left( A^{\bar{\kappa}/2} \right) = \infty \),

CONDITION C2: \( E\left( A^{\kappa} \right) > 1 \) for sufficiently large \( \kappa \).

Conditions C1 and C2 together complete the proof. \( \blacksquare \)

Lemma 3 For the model in (2) and (3), under the same Assumptions as Lemma 2,
\[ P\left( \sigma > x \right) \sim C x^{-\kappa_0}, \quad (6) \]
and
\[ P\left( |Y| > x \right) = E\left( |\epsilon|^{\kappa_0} \right) \times P\left( \sigma > x \right), \quad x \to \infty. \quad (7) \]

Proof. Assumption A4 is sufficient to establish the sequence \( \{ \sigma_t^2 \} \) as strictly stationary (see; e.g., Mikosch, 1999, Corollary 1.4.39). Owing to the method used to establish (5) as having a unique and positive solution, there exists a small \( \eta > 0 \) such that
\[ E\left( A^{\kappa_0/2+\eta} \right) < \infty. \]
Consider then for (4)

2
CONDITION C3: \( \{A_t\} \) is an i.i.d. sequence.

CONDITION C4: \( \sigma_{t-1}^2 \) is independent of \( A_t \) for every \( t \).

C3 follows because \( A_t \) is only a function of \( \epsilon_{t-1} \). The validity of C4 depends on \( \sigma_{t-1}^2 \) being a function of \( \epsilon_{t-2}, \epsilon_{t-3}, \ldots, \epsilon_0 \). Given C3 and C4, (4) is a SRE (see also Mikosch and Stărică, 2000). At this point, all of the conditions in Goldie (1991, Theorem 4.1) are satisfied, in which case,

\[
P(\sigma^2 > x) \sim cx^{-\kappa_0/2}.
\]

The result in (6) then follows from Proposition 1. Lastly, summarizing a result originally from Breiman (1965), which is also stated as Mikosch (1999, Proposition 1.3.9(b)), consider two non-negative random variables \( X \) and \( Z \) that are also independent. If \( X \) is regularly varying with tail index \( \theta \), and \( E(Z^{\theta+\eta}) < \infty \) for an \( \eta \) as defined above, then

\[
P(XZ > x) \sim E(Z^{\kappa}) P(X > x).
\]

Since \( |Y| = \sigma |\epsilon| \), (7) immediately follows from (9).

Remark 4 Lemma 3 collects results available in the literature (see; e.g., Mikosch and Stărică, 2000, Theorem 2.1, and Basrak, Davis, and Mikosch, 2002, Theorem 3.1(B)).

Next, for \( 0 \leq h < \infty \), consider

\[
Y_h^{(i)} = \left( |Y_0|^i, \sigma_0^i, \ldots, |Y_h|^i, \sigma_h^i \right), \quad i = 1, 2,
\]

and

\[
\bar{Y}_m = \left( Y_0, \sigma_0, \ldots, Y_h, \sigma_h \right).
\]

Lemma 5 For the model in (2) and (3), under the same Assumptions as Lemma 2, \( Y_h^{(2)} \) is RV(\( \kappa_0/2 \)), while both \( Y_h^{(1)} \) and \( Y_h \) are RV(\( \kappa_0 \)).

Proof. Given (4) and (8), Mikosch and Stărică (2000, proof of Theorem 2.3(a)) establishes \( Y_h^{(2)} \) as RV(\( \kappa_0/2 \)), which, in turn, establishes \( Y_h^{(1)} \) as RV(\( \kappa_0 \)), given Mikosch (1999, Proposition 1.5.9), the multivariate extension of Proposition 1. Finally, \( Y_h \) is RV(\( \kappa_0 \)) by Basrak et al. (2002, proof of Corollary 3.5(B)).

Remark 6 Lemma 5 pieces together different results available in the literature. The SRE upon which this lemma depends is (4) (see the proof of Lemma 3 that establishes (4) as a valid SRE). In contrast, the SRE upon which Basrak et al. (2002, Corollary 3.5) is based is more closely related to

\[
X_t = \begin{pmatrix} Y_t^2 \\ \sigma_t^2 \end{pmatrix} = \begin{pmatrix} \alpha \epsilon_t^2 \\ \beta \epsilon_t^2 \end{pmatrix} X_{t-1} + \begin{pmatrix} \omega \epsilon_t^2 \\ \omega \end{pmatrix} = A_t X_{t-1} + B_t
\]

(10)
which is also a valid SRE, since it, too, satisfies Conditions C3 and C4.

A.1.3. Central Limit Theorem

Lemma 7 For the model in (2) and (3), under the same Assumptions as Lemma 2, let

\[ Y_t = \left( Y_t, \sigma_t, \ldots, Y_{t+h}, \sigma_{t+h} \right), \quad h < \infty, \]

and \{a_n\} be a sequence of constants satisfying

\[ nP (|Y| > a_n) \longrightarrow 1, \quad n \to \infty, \]

where \(|Y| = \max_{m=0,\ldots,h} |Y_m|; a_n = n^{1/\kappa_0} L(n), \) and \( L(\cdot) \) is slowly-varying at \( \infty. \) Then

\[ N_n := \sum_{t=1}^{n} \delta_{a_n Y_t}, \quad N := \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \delta_{Q_{ij}}, \]

where

1. \( \sum_{i=1}^{\infty} \delta_{P_i} \) is a Poisson process on \( \mathbb{R}_+ \)

2. \( \sum_{j=1}^{\infty} \delta_{Q_{ij}}, \quad i \in \mathbb{N}, \) is an i.i.d. sequence of point processes on \( \mathbb{R}_{+}^{h+1} \setminus \{0\} \) with common distribution \( Q \)

3. \( \sum_{i=1}^{\infty} \delta_{P_i} \) and \( \sum_{j=1}^{\infty} \delta_{Q_{ij}}, \quad i \in \mathbb{N}, \) are mutually independent

4. \( Q_{ij} = \left( Q_{ij}^{(m)} \right)_{m=0,\ldots,h} \)

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

CONDITION C8: (joint) regular variation of all finite-dimimensional distributions of \( \{Y_t\} \)

CONDITION C9: weak mixing for \( \{Y_t\} \)

CONDITION C10: that

\[ \lim_{k \to \infty} \lim_{n \to \infty} P \left( \bigvee_{k \leq |t| \leq r_n} |Y_t| > a_n y \mid |Y_0| > a_n y \right) = 0, \quad y > 0, \quad (11) \]

where \( \bigvee_{i} b_i = \max_{i} (b_i), \) 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 \( \phi_n \) is the mixing rate of \( \{Y_t\}. \)
Lemma 5 establishes Condition C8. \( \{Y_t\} \) is strongly mixing by Carrasco and Chen (2002, Corollary 6). Finally, by the definition of the sequence \( \{Y_t\} \) and as in Mikosch and Stărică (2000, proof of Theorem 3.1), it suffices to switch in (11) to the sequence \( \left\{ \left( Y_t^2, \sigma_t^2 \right) \right\} \) and to replace \( a_n y \) by \( a_n^2 y^2 \). Consequently, consider the SRE in (10). Recursive substitution establishes

\[
X_t = \prod_{i=1}^{t} A_i X_0 + \sum_{i=1}^{t} \prod_{j=i+1}^{t} A_j B_i \\
= I_{t,1} X_0 + I_{t,2} 
\]

Condition C10 is then established following Davis, Mikosch, and Basrak (1999, proof of Theorem 3.3).

**Remark 8** Lemma 7 is the (nonstandard) CLT upon which (weak) distributional convergence of the GARCH(1, 1) estimators in Sections 2.1 of the main paper are based and generalizes Mikosch and Stărică (2000, Theorem 3.1) by covering the case of an asymmetric \( D \). Given Remark 6, Lemma 7 complements Basrak et al. (2002, Theorem 2.10). Finally, given a continuous mapping argument, implied by Lemma 7 for \( Y_{(l)} \):

\[
N_n^{(l)} := \sum_{t=1}^{n} \delta_{a_n^{-l} Y_t} \overset{d}{\to} N^{(l)} := \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \delta P_i Q_{i,j}^{(l)},
\]

where

\[
Q_{i,j}^{(l)} = \left( \left( Q_{ij,Y}^{(m)} \right)^l, \left( Q_{ij,\sigma}^{(m)} \right)^l \right) \text{ for } m=0, ..., h \right) .
\]

**A.1.4. GARCH(1,1) Convergence Results**

From the model in (2) and (3) when \( \alpha_{1,0} = \alpha_{2,0} = \alpha_0 \),

\[
X_t = \phi_0 X_{t-1} + V_t, \quad V_t = W_t - \beta_0 W_{t-1},
\]

where \( X_t = Y_t^2 - \gamma_0 \), and \( \gamma_0 = E (Y_t^2) = \frac{\omega_0}{1-\phi_0} \).

**Lemma 9** For the model in (2) and (3), under the same Assumptions as Lemma 2,

\[
a_n^{-3} \sum_{t} \sigma_t^3 - E (\sigma^3) \overset{d}{\to} V_{0,\sigma},
\]

where \( \overset{d}{\to} \) is weak, and \( V_{0,\sigma} \) is \( (\kappa_0/3) \)-stable.
Proof. For an $\varepsilon > 0$,

$$a_n^{-3} \sum_t \sigma_t^3 - E (\sigma^3) = a_n^{-3} \sum_t (\sigma_t^3 - E (\sigma^3)) \times I_{\{\sigma_t > a_n \varepsilon\}} + a_n^{-3} \sum_t (\sigma_t^3 - E (\sigma^3)) \times I_{\{\sigma_t \leq a_n \varepsilon\}}$$

$$= Ia + Ila,$$

where $E (\sigma^3) < \infty$ by Prono (2018, Lemma 1). Then,

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

$$\sim a_n^{-3} E (\sigma^3) n P (\sigma_t > a_n \varepsilon)$$

$$\longrightarrow 0,$$

where "$\sim$" holds for sufficiently large $n$, and "$\longrightarrow"$ as $n \to \infty$ follows since

$$nP (\sigma_t > a_n \varepsilon) \longrightarrow \varepsilon^{-\kappa_0}, \quad n \to \infty,$$

so that

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

Next,

$$a_n^{-3} \sum_t E (\sigma^3) \times I_{\{\sigma_t \leq a_n \varepsilon\}} = n^{\frac{-6}{\kappa_0}} E (\sigma^3) n^{-1/2} \sum_t I_{\{\sigma_t \leq a_n \varepsilon\}}$$

$$\longrightarrow 0$$

as $n \to \infty$ by the CLT of Ibragimov and Linnik (1971, Theorem 18.5.3) applied to $n^{-1/2} \sum_t I_{\{\sigma_t \leq a_n \varepsilon\}}$ if $\kappa_0 < 6$, so that

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

Then, by Markov’s Inequality for a $\zeta > 0$,

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

(16)

In addition, for $\kappa \equiv \kappa_0 / 3$, and $r \in (\kappa, 2)$, there exists a constant $C \in (0, \infty)$ such that

$$n \left( \zeta^{-1} a_n^{-3} \right) E \left( \sigma^3 \times I_{\{\sigma \leq a_n \varepsilon\}} \right) \leq n C \left( \zeta^{-1} a_n^{-3} \right)^r E \left( \sigma^{3r} \times I_{\{\sigma \leq a_n \varepsilon\}} \right)$$

$$\leq n C \left( \zeta^{-1} a_n^{-3} \right)^r \int_0^{a_n \varepsilon} \sigma^{3r} f (\sigma) d\sigma$$

$$\leq n C \left( \zeta^{-1} a_n^{-3} \right)^r (-\kappa_0) \int_0^{a_n \varepsilon} \sigma^{3r-\kappa_0-1} L (\sigma) d\sigma,$$
where the last inequality follows from Mikosch (1999, Theorem 1.2.9). Since, by Karamata’s Theorem,

\[ \int_0^{a_n \varepsilon} \sigma^{3r-\kappa_0-1} L(\sigma) \, d\sigma \sim \frac{\sigma^{3r-\kappa_0}}{(3r-\kappa_0)} L(\sigma) \bigg|_{0}^{a_n \varepsilon}, \]

then

\[ n \left( \zeta^{-1} a_n^{-3} \right) \frac{E(\sigma^3 \times I_{\{\sigma \leq a_n \varepsilon\}})}{} \leq C \left( \zeta^{-1} a_n^{-3} \right)^{r} \left( \frac{\kappa_0}{3r-\kappa_0} \right) (a_n \varepsilon)^{3r} n P(\sigma > a_n \varepsilon) \]

\[ \to C \zeta^{-r} \left( \frac{\kappa_0}{3r-\kappa_0} \right) \varepsilon^{3r-\kappa_0}, \]

\[ \to 0, \]

where the first "\( \to \)" is as \( n \to \infty \) and follows from (15), while the second "\( \to \)" is as \( \varepsilon \to 0 \). As a consequence, \( \lim_{n \to \infty} \lim_{\varepsilon \to 0} \sup_{t} P(\left\{ a_n^{-3} \sum_{t} \sigma_t^3 \times I_{\{\sigma_t \leq a_n \varepsilon\}} > \zeta \right\}) = 0 \), and

\[ a_n^{-3} \sum_{t} \sigma_t^3 - E(\sigma^3) = a_n^{-3} \sum_{t} \sigma_t^3 \times I_{\{\sigma_t > a_n \varepsilon\}} + o_p(1). \]

Finally, let

\[ y_t = \left( y_t^{(0)}, y_t^{(0)}, \ldots, y_t^{(h)}, y_t^{(h)} \right) \in \mathbb{R}^{h+1} \setminus \{0\}, \]

and define

\[ T_{0,\varepsilon,\sigma} \left( \sum_{i=1}^{\infty} n_i \delta y_i \right) = \sum_{i=1}^{\infty} n_i \left( y_i^{(0)} \right)^{3} \times I_{\left\{ y_i^{(0)} > a_n \varepsilon \right\}}. \]

Since the set \( \{ y \in \mathbb{R}^{h+1} \setminus \{0\} : |y^{(m)}| > \varepsilon \} \) for any \( m \geq 0 \) is bounded away from the origin, and given Vaynman and Beare (2014, Lemma A.2), then

\[ a_n^{-3} \sum_{t} \sigma_t^3 - E(\sigma^3) = T_{0,\varepsilon,\sigma} (N_n) + o_p(1) \]

\[ \to T_{0,\varepsilon,\sigma} (N) \]

\[ \to V_{0,\sigma}, \]

where the first "\( \to \)" is as \( n \to \infty \) and follows from Lemma 7 and the continuous mapping theorem, while the second "\( \to \)" is as \( \varepsilon \to 0 \) and follows from Davis and Hsing (1995, Proof of Theorem 3.1, pp 897-898).

\[ \square \]

**Lemma 10** For the model in (2) and (3), under the same Assumptions as Lemma 2,

\[ a_n^{-3} \sum_{t} Y_t^2 \sigma_{t+m}^2 - E(\sigma_{t+m}^2) \to (V_{m,y})_{m=1,...,h}, \]

where "\( \to \)" continues to be weak, and \( V_{m,y} \) is \( (\kappa_0/3) \)-stable.
Proof. The (weak) convergence result in (20) is established for \( m = 1, 2 \). Generalizing to cases where \( m > 2 \) is an extension of the arguments given below. Given (4),

\[
a_n^{-3} \sum_t Y_t \sigma_{t+1}^2 - E (Y_t \sigma_{t+1}^2) = a_n^{-3} \sum_t \sigma_t^3 \left( \epsilon_t A_{t+1} - \alpha_0 c_3^s \right) \times I_{\{\sigma_t > \alpha_n \epsilon\}} + a_n^{-3} \sum_t \sigma_t^3 \left( \epsilon_t A_{t+1} - \alpha_0 c_3^s \right) \times I_{\{\sigma_t \leq \alpha_n \epsilon\}} + \alpha_0 c_3^s a_n^{-3} \sum_t \sigma_t^3 - E (\sigma_3^3) + o_p (1) \]

\[
= a_n^{-3} \sum_t \sigma_t^3 \epsilon_t A_{t+1} \times I_{\{\sigma_t > \alpha_n \epsilon\}} + a_n^{-3} \sum_t \sigma_t^3 \left( \epsilon_t A_{t+1} - \alpha_0 c_3^s \right) \times I_{\{\sigma_t \leq \alpha_n \epsilon\}} + o_p (1),
\]

where the first equality relies on

\[
a_n^{-1} \sum_t Y_t \xrightarrow{d} V_0, \tag{21}
\]

which follows given Lemma 7 and Davis and Hsing (1995, Theorem 3.1) and under which \( V_0 \) is \( \kappa_0 \)-stable, while the second equality follows from (19). Then for the same \( \tau \in (1, 2) \) in the proof of Lemma 9 and a \( \zeta > 0 \),

\[
P \left( \left| a_n^{-3} \sum_t \sigma_t^3 \times I_{\{\sigma_t \leq \alpha_n \epsilon\}} \times (\epsilon_t A_{t+1} - \alpha_0 c_3^s) \right| > \zeta \right) \leq \left( \zeta^{-1} a_n^{-3} \right)^{\tau} E \left| \sum_t \sigma_t^3 \times I_{\{\sigma_t \leq \alpha_n \epsilon\}} \times (\epsilon_t A_{t+1} - \alpha_0 c_3^s) \right|^{\tau} \leq 2 \left( \zeta^{-1} a_n^{-3} \right)^{\tau} n E \left( \sigma_t^3 \times I_{\{\sigma_t \leq \alpha_n \epsilon\}} \right) \times E \left[ \alpha_0 \left( \epsilon_t^3 - c_3^3 \right) + \beta_0 \epsilon_t \right],
\]

where the first inequality follows from Markov’s Inequality, and the second inequality follows from von Bahr and Esseen (1965, Theorem 2), since for

\[
M_n = \sum_t \sigma_t^3 \times I_{\{\sigma_t \leq \alpha_n \epsilon\}} \times (\epsilon_t A_{t+1} - \alpha_0 c_3^s),
\]

\[
E \left( M_{n+1} \mid M_n \right) = M_n \ a.s. \tag{22}
\]

Given (17),

\[
\lim_{n \to \infty} \limsup_{\epsilon \to 0} P \left( \left| a_n^{-3} \sum_t \sigma_t^3 \times I_{\{\sigma_t \leq \alpha_n \epsilon\}} \times (\epsilon_t A_{t+1} - \alpha_0 c_3^s) \right| > \zeta \right) = 0.
\]

Next, given (18), define

\[
T_{m, \epsilon, Y} \left( \sum_{i=1}^{\infty} n_i \delta_{y_i} \right) = \sum_{i=1}^{\infty} n_i \left( y_i^{(0)} \right) \left( y_i^{(m)} \right)^2 \times I_{\{y_i^{(0)} > a_n \epsilon\}}, \quad m \geq 1. \tag{22}
\]

\(^1\)The applicability of von Bahr and Esseen (1965, Theorem 2) in this general context is first recognized in Vaynman and Beare (2014, proof of Lemma A.1).
Then
\[
a_n^{-3} \sum_t Y_t \sigma_{t+2}^2 - E(Y_t \sigma_{t+2}^2) = T_{1, \varepsilon, Y}(N_n) + o_p(1) \tag{23}
\]
where the first "\( \overset{d}{\rightarrow} \)" is as \( n \to \infty \), the second "\( \overset{d}{\rightarrow} \)" as \( \varepsilon \to 0 \), and each convergence result follows from the same arguments that support (19). Next, consider
\[
a_n^{-3} \sum_t Y_t \sigma_{t+2}^2 - E(Y_t \sigma_{t+2}^2)
= a_n^{-3} \sum_t Y_t \sigma_{t+1}^2 (A_{t+2} - E(A)) \times I_{\{\sigma_t > a_n \varepsilon\}}
+ E(A) a_n^{-3} \sum_t Y_t \sigma_{t+1}^2 - E(Y_t \sigma_{t+1}^2) + o_p(1)
= a_n^{-3} \sum_t Y_t \sigma_{t+2}^2 \times I_{\{\sigma_t > a_n \varepsilon\}}
+ a_n^{-3} \sum_t Y_t \sigma_{t+1}^2 (A_{t+2} - E(A)) \times I_{\{\sigma_t \leq a_n \varepsilon\}} + o_p(1),
\]
where the first equality, again, relies on (4) and (21), while the second equality follows from (23). For
\[
IIb = \alpha_0 \omega_0 a_n^{-3} \sum_t Y_t \left( \varepsilon_{t+1}^2 - 1 \right) \times I_{\{\sigma_t \leq a_n \varepsilon\}} + \alpha_0 a_n^{-3} \sum_t \sigma_t^3 \varepsilon_t A_{t+1} \times \left( \varepsilon_{t+1}^2 - 1 \right) \times I_{\{\sigma_t \leq a_n \varepsilon\}}
= \alpha_0 a_n^{-3} \sum_t \sigma_t^3 \varepsilon_t A_{t+1} \times \left( \varepsilon_{t+1}^2 - 1 \right) \times I_{\{\sigma_t \leq a_n \varepsilon\}} + o_p(1),
\]
where the second equality relies on the CLT of Ibragimov and Linnik (1971, Theorem 18.5.3). Next, for a \( \zeta > 0 \),
\[
P \left( a_n^{-3} \left| \sum_t \sigma_t^3 \times I_{\{\sigma_t \leq a_n \varepsilon\}} \times (\alpha_0 \varepsilon_t^3 + \beta_0 \varepsilon_t) \times (\varepsilon_{t+1}^2 - 1) \right| > \zeta \right)
\leq 2 \left( \zeta^{-1} a_n^{-3} \right)^r n E \left( \sigma_t^3 \times I_{\{\sigma_t \leq a_n \varepsilon\}} \right)
\times E \left| \alpha_0 \varepsilon_t^3 + \beta_0 \varepsilon_t \right| \times E \left( \varepsilon_{t+1}^2 - 1 \right)^r,
\]
by Markov’s Inequality and von Bahr and Esseen (1965, Theorem 2), so that
\[
\lim_{n \to \infty} \limsup_{\varepsilon \to 0} P \left( a_n^{-3} \left| \sum_t \sigma_t^3 \times I_{\{\sigma_t \leq a_n \varepsilon\}} \times (\alpha_0 \varepsilon_t^3 + \beta_0 \varepsilon_t) \times (\varepsilon_{t+1}^2 - 1) \right| > \zeta \right) = 0;
\]
in which case,

\[ a_n^{-3} \sum_t Y_t \sigma_{t+2}^2 - E (Y_t \sigma_{t+2}^2) = a_n^{-3} \sum_t Y_t \sigma_{t+2}^2 \times I(\sigma_{t+2} > a_n \varepsilon) + o_p(1) \]

\[ = T_{2,\varepsilon,Y} (N_n) + o_p(1) \]

\[ \xrightarrow{d} T_{2,\varepsilon,Y} (N) \]

\[ \xrightarrow{d} V_{2,Y} , \]

where, as is true elsewhere, the first \( \xrightarrow{d} \) is as \( n \to \infty \), and the second \( \xrightarrow{d} \) is as \( \varepsilon \to 0 \). ■

**Lemma 11** For the model in (2) and (3), under the same Assumptions as Lemma 2,

\[ a_n^{-3} \sum_t Y_t^2 - m E (Y_t^2) \xrightarrow{d} \alpha_0^{-1} \left( V_{m+1,Y} - \beta_0 V_{m,Y} \right)_{m=1,\ldots,h}, \tag{24} \]

where, as is the case elsewhere, \( \xrightarrow{d} \) is weak, and the limits are \( (\kappa_0/3) \)-stable.

**Proof.** The (weak) convergence result in (24) is established for \( m = 1, 2 \). Generalizing to \( m > 2 \) is an extension of the results stated below. From (4),

\[ \epsilon_t^2 = \alpha_0^{-1} (A_{t+1} - \beta_0). \tag{25} \]

in which case,

\[ a_n^{-3} \sum_t Y_t \sigma_{t+1}^2 - E (Y_t \sigma_{t+1}^2) \]

\[ = \alpha_0^{-1} \left( a_n^{-3} \sum_t Y_t \sigma_{t+1}^2 A_{t+2} - E (Y_t \sigma_{t+1}^2 A_{t+2}) - \beta_0 a_n^{-3} \sum_t Y_t \sigma_{t+1}^2 E (Y_t \sigma_{t+1}^2) \right) \]

\[ = \alpha_0^{-1} \left( a_n^{-3} \sum_t Y_t \sigma_{t+2}^2 - E (Y_t \sigma_{t+2}^2) - \beta_0 a_n^{-3} \sum_t Y_t \sigma_{t+1}^2 E (Y_t \sigma_{t+1}) - \omega_0 a_n^{-3} \sum_t Y_t \sigma_{t+1}^2 \right) \]

\[ = \alpha_0^{-1} \left( a_n^{-3} \sum_t Y_t \sigma_{t+2}^2 - E (Y_t \sigma_{t+2}^2) - \beta_0 a_n^{-3} \sum_t Y_t \sigma_{t+1}^2 E (Y_t \sigma_{t+1}) + o_p(1) \right) \]

\[ \xrightarrow{d} \alpha_0^{-1} \left( V_{2,Y} - \beta_0 V_{1,Y} \right) , \]

where the second equality relies on (4), the third equality (21), and \( \xrightarrow{d} \) follows from Lemma 10. The
same arguments then support
\[
 a_n^{-3} \sum_t Y_t^2 - E(Y_t^2) + a_n^{-3} Y_{t+2} - E(Y_{t+2}^2)
 = a_0^{-1} \left( a_n^{-3} \sum_t Y_t^2 + a_n^{-3} Y_{t+2} - E(Y_t^2) \right)
 = a_0^{-1} \left( a_n^{-3} \sum_t Y_t^2 + a_n^{-3} Y_{t+2} - E(Y_t^2) + o_p(1) \right)
 \]
which completes the proof. 

**Lemma 12** For the model in (2) and (3), under the same Assumptions as Lemma 2,
\[
a_n^{-3} \sum_t Y_t^3 - E(Y_t^3) \xrightarrow{d} V_{0,Y},
\]
where "\( \xrightarrow{d} \)" is weak, and \( V_{0,Y} \) is \((\kappa_0/3)\)-stable.

**Proof.**
\[
a_n^{-3} \sum_t Y_t^3 - E(Y_t^3) = a_n^{-3} \sum_t \sigma_t^3 \left( \epsilon_t^3 - c_3^3 \right) \times I_{\{\sigma_t \leq a_n \epsilon\}}
+ a_n^{-3} \sum_t \sigma_t^3 \left( \epsilon_t^3 - c_3^3 \right) \times I_{\{\sigma_t > a_n \epsilon\}}
+ a_n^{-3} \sum_t \sigma_t^3 - E(\sigma^3)
= Ic + IIc + IIIc.
\]
As relied upon elsewhere, given Markov’s Inequality and von Bahr and Esseen (1965, Theorem 2), for a \( \zeta > 0 \) and a \( r \in (\kappa, 2) \) defined in the proof of Lemma 9,
\[
P(|Ic| > \zeta) \leq (\zeta^{-1} a_n^{-3})^r \left| \sum_t \sigma_t^3 \left( \epsilon_t^3 - c_3^3 \right) \times I_{\{\sigma_t \leq a_n \epsilon\}} \right|^r
\leq 2 (\zeta^{-1} a_n^{-3})^r n E \left( \sigma_t^{3r} \times I_{\{\sigma_t \leq a_n \epsilon\}} \right) \times E|\epsilon_t^3 - c_3^3|^r
\]
so that
\[
\lim_{n \to \infty} \limsup_{\zeta \to 0} P(|Ic| > \zeta) = 0
\]
by the arguments that support (17). Next, given (18), define
\[
T_{0,\epsilon,Y} \left( \sum_{i=1}^\infty n_i \tilde{Y}_i \right) = \sum_{i=1}^\infty n_i \left( Y_i^{(0)} \right)^3 \times I_{\{n_i > a_n \epsilon\}}.
\]
Then, given Lemma 7,
\[
a_n^{-3} \sum_t Y_t^3 - E(Y_t^3) = a_n^{-3} \sum_t Y_t^3 \times I\{\sigma_t > n\epsilon\} - c^3 a_n^{-3} \sum_t \sigma_t^3 \times I\{\sigma_t > n\epsilon\} + IIIc + o_p(1)
\]
\[
= T_{0,\epsilon,Y} (N_n) - c^3 T_{0,\epsilon,\sigma} (N_n) + IIIc
\]
\[
\overset{d}{\longrightarrow} T_{0,\epsilon,Y} (N)
\]
\[
\overset{d}{\longrightarrow} V_{0,Y},
\]
where \(T_{0,\epsilon,\sigma} (N_n)\) is defined in the proof of Lemma 9 and the sequential limiting results (first as \(n \to \infty\) and then as \(\epsilon \to 0\)) follow from the arguments given in that same proof. ■

Consider
\[
Z_{t-2} = \left( Y_{t-2}, \ldots, Y_{t-h} \right)'
\]
as a vector of (proper) instruments for \(X_{t-1}\) in (13). Then

\[
\hat{\phi}_{IV} = \hat{F} \left( n^{-1} \sum_t \hat{X}_t Z_{t-2} \right), \quad \hat{F} = \frac{\left( n^{-1} \sum_t \hat{X}_t Z_{t-2} \right)'}{\left( n^{-1} \sum_t \hat{X}_t Z_{t-2} \right)'} \Lambda \left( n^{-1} \sum_t \hat{X}_t Z_{t-2} \right).
\] (27)

**Theorem 13** Let
\[
F_0 = B_0^{-1} A_0',
\]
where
\[
A_0 = A_0 E(X_{t-1} Z_{t-2}), \quad B_0 = E(X_{t-1} Z_{t-2})' A_0.
\]
In addition, let Assumptions A1–A5 from the main paper hold. Then
\[
\hat{\phi}_{IV} \overset{a.s.}{\longrightarrow} \phi_0,
\]
and
\[
na_n^{-3} \left( \hat{\phi}_{IV} - \phi_0 \right) \overset{d}{\longrightarrow} \alpha_0^{-1} F_0 S,
\] (28)
where \(\kappa_0 \in (3, 6)\), "\(\overset{d}{\longrightarrow}\)" is weak,
\[
S = (V_{m+1.Y} - \beta_0 V_m.Y)_{m=2,\ldots,h+1},
\]
each \((V_m.Y)_{m=2,\ldots,h+1}\) is defined in Lemma 16, and \(S\) is jointly \((\kappa_0/3)\)-stable. If \(\kappa_0 \in (6, \infty)\) such that \(E(Y_t^6) < \infty\), then
\[
\sqrt{n} \left( \hat{\phi}_{IV} - \phi_0 \right) \overset{d}{\longrightarrow} N \left( 0, \frac{A_0' \Sigma \Sigma Z_{t-2} A_0}{B_0^2} \right),
\]
where
\[ \Sigma_{VZ_{-2}} = E \left( V_t^2 Z_{t-2} Z'_{t-2} \right) + 2E \left( V_t V_{t-1} Z_{t-2} Z'_{t-3} \right), \]
and \( V_t \) is defined in Theorem 1 of the main paper.

**Proof.** Since
\[ \hat{X}_t = X_t - (\hat{\gamma} - \gamma_0), \] (29)

\[ \hat{X}_t = c_0 + \phi_0 \hat{X}_{t-1} - \beta_0 W_{t-1} + W_t, \quad c_0 = (\hat{\gamma} - \gamma_0) \times (\phi_0 - 1) \] (30)
given (13). Also, since \( \{Y_t\} \) is strongly mixing (see the proof of Theorem 1 in the Appendix of the main paper), then given (29),
\[ n^{-1} \sum_t \hat{X}_{t-1} Z_{t-2} = n^{-1} \sum_t X_{t-1} Z_{t-2} - (\hat{\gamma} - \gamma_0) n^{-1} \sum_t Z_{t-2} \overset{a.s.}{\rightarrow} E \left( X_{t-1} Z_{t-2} \right) \]
by the Ergodic Theorem so that \( \hat{F} \overset{a.s.}{\rightarrow} F_0 \). Also, given (30),
\[ n^{-1} \sum_t \hat{X}_t Z_{t-2} = c_0 n^{-1} \sum_t Z_{t-2} + \phi_0 n^{-1} \sum_t \hat{X}_{t-1} Z_{t-2} - \beta_0 n^{-1} \sum_t W_{t-1} Z_{t-2} + n^{-1} \sum_t W_t Z_{t-2} \]
\[ \overset{a.s.}{\rightarrow} \phi_0 E \left( X_{t-1} Z_{t-2} \right). \]

Next,
\[ n a_n^{-3} \left( \hat{\phi}_{IV} - \phi_0 \right) = F_0 \left( a_n^{-3} \sum_t X_t Z_{t-2} - E \left( X_t Z_{t-2} \right) \right) + o_p(1) \]
\[ = F_0 \left( a_n^{-3} \sum_t Y_t^2 Z_{t-2} - E \left( Y_t^2 Z_{t-2} \right) \right) + o_p(1) \]
\[ \overset{d}{\rightarrow} \alpha_0^{-1} F_0 S \]
where the second equality follows from the arguments that support (XX) in the proof of Theorem 1 in the Appendix of the main paper, and \( S \) is jointly \((\kappa_0/3)\)—stable by Lemma 17 and Samorodnitsky and Taqqu
(1994, Theorem 2.1.5(c)). If $\kappa_0 \in (6, \infty)$ so that $E \left( Y_t^6 \right) < \infty$, then
\[
\sqrt{n} \left( \hat{\phi} - \phi_0 \right) = \sqrt{n} \left( \frac{\phi_0 \left( n^{-1} \sum_{t} X_t Z_{t-2} \right)'}{B} \widehat{A} - \phi_0 + \frac{\left( n^{-1} \sum_{t} V_t Z_{t-2} \right)'}{B} \widehat{A} + o_p(1) \right)
\]
\[
= \sqrt{n} \left( \frac{\left( n^{-1} \sum_{t} V_t Z_{t-2} \right)'}{B} \widehat{A} + o_p(1) \right)
\]
\[
\rightarrow N \left( 0, \frac{A_0' \Sigma_{V} Z_{t-2} A_0}{B_0^2} \right),
\]
where the limiting result uses the same CLT from the proof of Theorem 1.

Consistency of $\hat{\phi}_{IV}$ does not depend on consistency of $\hat{\gamma}$, and $\hat{\gamma}$ does not impact the limiting distribution of $\hat{\phi}_{IV}$. Necessary for $B_0 \neq 0$ is $E \left( Y_t^3 \right) \neq 0$, which illustrates the lack of identification that results if in A1, $D$ is a symmetric distribution. (28) depends on $j \in (3, 6)$ in A1, which is consistent with empirical findings. Lastly, the rate of convergence that applies in (28) is $n^{\frac{\kappa_0-3}{3}}$.

References


