Empirical likelihood and self-weighting approach for hypothesis testing of infinite variance processes and its applications

Fumiya Akashi

Research Associate
Department of Applied Mathematics
Waseda University

August 17, 2016

Boston University/Keio University Workshop 2016

This work was supported by Grant-in-Aid for Young Scientists (B) (16K16022, Fumiya Akashi)

Outline

Outline

Introduction

Electrical engineering, Hydrology,
Finance, physical systems

$\Downarrow$

Heavy-tailed data $y_1, \cdots, y_n$

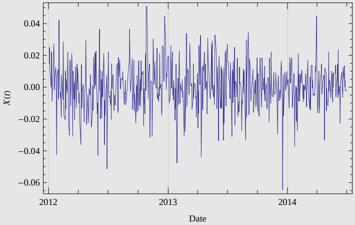
- $P(y_t > x) \sim cx^{-\alpha} \quad (x \to \infty)$

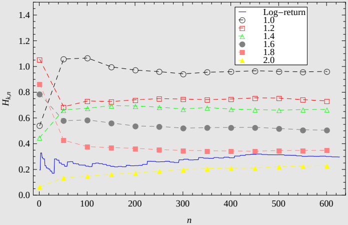
α: tail-index

$$0 < \alpha < 2 \Rightarrow \mathbb{E}[y_t^2] = \infty.$$

Introduction (cont.)

Figure: Log-stock return process of Ford and Hill-plot for y_t





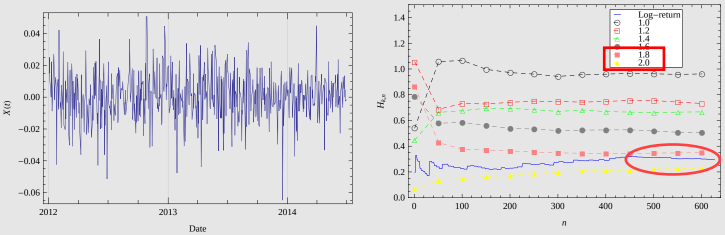
$\Rightarrow \{y_t : t \in \mathbb{Z}\}$: Infinite variance process

Example: Stable AR-model

$$y_t = \beta^\top X_{t-1} + e_t, \quad X_{t-1} = (y_{t-1}, \ldots, y_{t-p})^\top$$
$$\Rightarrow \mathbb{E}[||X_{t-1}||^2] = \infty$$

Introduction (cont.)

Figure: Log-stock return process of Ford and Hill-plot for y_t



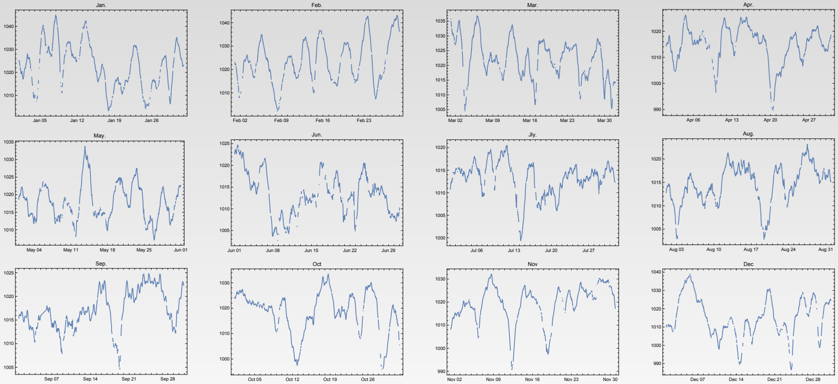
$\Rightarrow \{y_t : t \in \mathbb{Z}\}$: Infinite variance process

Example: Stable AR-model

$$y_t = \beta^\top X_{t-1} + e_t, \quad X_{t-1} = (y_{t-1}, \ldots, y_{t-p})^\top$$
$$\Rightarrow \mathbb{E}[||X_{t-1}||^2] = \infty$$

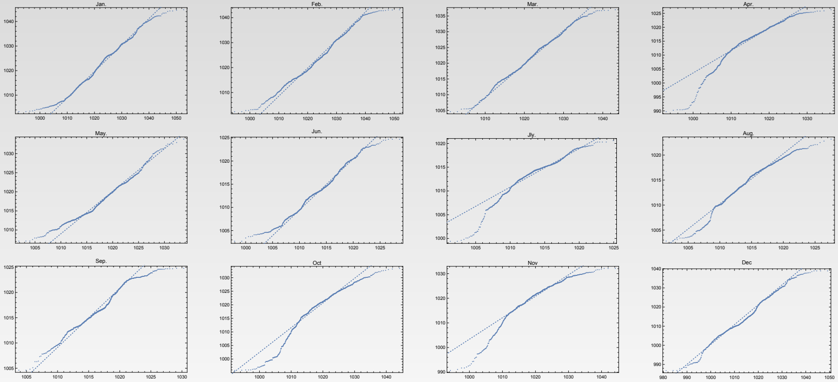
Introduction (cont.)

Figure: Atmospheric pressure data in Chicago, 2015 (mbar)



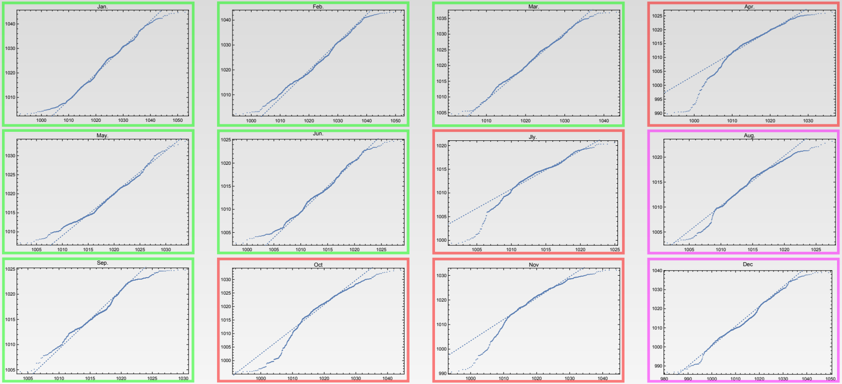
Introduction (cont.)

Figure: QQ-plot for atmospheric pressure data in Chicago, 2015 (mbar)



Introduction (cont.)

Figure: QQ-plot for atmospheric pressure data in Chicago, 2015 (mbar)



Gaussian tail or Heavy tail ?? (or Another tail ??)

Introduction (cont.)

EL approach for...

- Akashi, Liu & Taniguchi (2015) $\cdots$ Stable process + FD-SN-EL $\tilde{r}_n(\theta_0)$

Problem:

- Limit distributions contain α & σ (unknown).

$\Rightarrow$ Self-Weighting (SW) approach (Ling (2005), Pan et al. (2007))
$\Rightarrow$ Least Absolution Deviation (LAD) based statistic (Chen et al. (2008))

Main aim

- L_1-SW-EL based statistic for IV-process:

$$\rho_n^* := \inf_{R\beta=c}\{-2\log r_n^*(\beta)\} - \inf_{\beta\in\mathcal{B}}\{-2\log r_n^*(\beta)\}$$

(r_n^*: L_1-SW-EL function)

$$\rho_n^* \xrightarrow{\mathcal{L}} \chi^2 \quad \text{(pivotal limit distribution)}$$

Outline

Fundamental settings

Model: ARMA(p,q) process

$$y_t = \sum_{j=1}^{p} b_j y_{t-j} + e_t + \sum_{j=1}^{q} a_j e_{t-j},$$

where

- $\beta = (b_1, \ldots, b_p, a_1, \ldots, a_q)^\top \in \mathrm{Int}(\mathcal{B})$ ($\mathcal{B}$: compact parameter space),
- $\{e_t : t \in \mathbb{Z}\}$: i.i.d. r.v.s. with $\mathrm{med}(e_1) = 0$.

Remark 1

$\{e_t : t \in \mathbb{Z}\}$ can be infinite variance r.v.s.

Fundamental settings (cont.)

Nested linear hypothesis:

$$H : R\beta = c$$

where

$$\begin{cases} R : r \times (p+q) \\ c : r \times 1 \end{cases} \qquad (r < p+q).$$

Examples

(i) $R = (O_{q\times p}, I_{q\times q})$ & $c = 0_q$
$\Rightarrow H : a_j = 0$ for $j = 1, ..., q$ (test for serial correlation)

(ii) $R = (I_{p_1\times p_1},\ O_{p_1\times(p+q-p_1)})$ & $c = 0_{p_1}$
$\Rightarrow H : b_j = 0$ for $j = 1, \ldots, p_1$ (variable selection)

Fundamental settings (cont.)

$$e_t = y_t - \sum_{j=1}^{p} b_j y_{t-j} - \sum_{j=1}^{q} a_j e_{t-j} \quad \& \quad \mathrm{med}(e_t) = 0$$

Self-weighted LAD estimator (Pan et al. (2007))

$$\hat{\beta}_n = \arg\min_{\beta \in \mathcal{B}} Q(\beta),$$

where

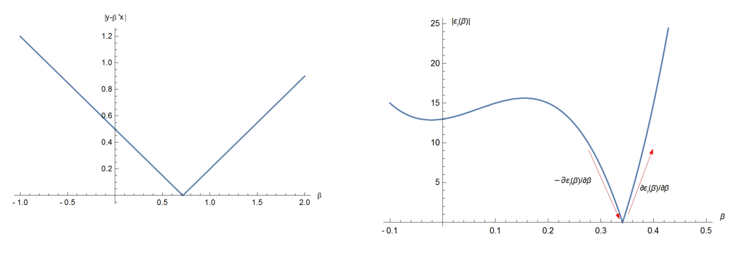
- $$Q(\beta) = \sum_{t=u+1}^{n} w_t \left| \varepsilon_t(\beta) \right|,$$
- $u \geq \max\{p, q\} + 1$: some starting point,
- $$\varepsilon_t(\beta) = \begin{cases} 0 & (t \leq 0) \\ y_t - \sum_{j=1}^{p} b_j y_{t-j} - \sum_{j=1}^{q} a_j \varepsilon_{t-j}(\beta) & (1 \leq t \leq n) \end{cases},$$
- w_t : self-weights of the form $w_t = {}^{\exists} w(y_{t-1}, y_{t-2}, \ldots)$

Fundamental settings (cont.)

Remark 2

Because of the truncation, $\varepsilon_t(\beta_0) \neq e_t$ but $\varepsilon_t(\beta_0) \approx e_t$.

Figure: Example of $\varepsilon_t(\beta)$ for AR(p) case (left) & general ARMA case



Fundamental settings (cont.)

Pan et al. (2007) showed that $\hat{\beta}_n$ is asymptotically normal & SW-Wald test:

$$W_n := n\left(R\hat{\beta}_n - c\right)^\top \hat{H}\left(R\hat{\beta}_n - c\right) \xrightarrow{\mathcal{L}} \chi_r^2.$$

Problem:

$\hat{H}$ contains unknown quantity $f(0)$.

Main aim

- Remove unknown quantities from statistic/limit distribution.

Fundamental settings (cont.)

$Y_1, \ldots, Y_n \sim$ i.i.d., $F(x)$
Hypothesis: $H : \mathbb{E}_F[Y_i] = 0$
Nonparametric likelihood:

$$L(F) = \prod_{i=1}^{n} \{F(Y_i) - F(Y_i-)\}$$

$\rightarrow$ Empirical distribution function

$$F_n(y) := \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(Y_i \leq y)$$

maximize $L(F)$ ($L(F_n) = n^{-n}$).
$\rightarrow$ Profile nonparametric likelihood ratio under H:

$$r_n = \frac{\sup\{L(F) : F \in \mathcal{F}_H\}}{L(F_n)},$$

where

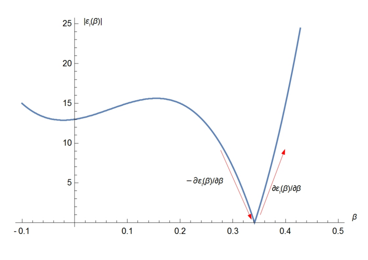
$$\mathcal{F}_H = \left\{F(y) = \sum_{i=1}^{n} v_i \mathbb{I}(Y_i \leq y) : \sum_{i=1}^{n} v_i Y_i = 0, \sum_{i=1}^{n} v_i = 1, 0 \leq v_i \leq 1\right\}$$

Fundamental settings (cont.)

For EL, we need “moment condition”

$$\mathbb{E}[{}^{\exists}g_t(\beta_0)] = 0_{p+q}.$$

$Q(\beta)$: NOT differentiable w.r.t β.



- $Q(\beta) = \sum_{t=u+1}^{n} \varepsilon_t(\beta)^2$

  $\Rightarrow$ Minimizer of $Q(\beta)$: Solution to $\displaystyle\sum_{t=u+1}^{n} \varepsilon_t(\beta) \frac{\partial \varepsilon_t(\beta)}{\partial \beta} = 0_{p+q}$

Fundamental settings (cont.)

- $Q(\beta) = \sum_{t=u+1}^{n} w_t \left|\varepsilon_t(\beta)\right|$

  $\Rightarrow$ Minimizer of $Q(\beta)$: Solution to $\displaystyle\sum_{t=u+1}^{n} w_t \mathrm{sign}\left\{\varepsilon_t(\beta)\right\} \frac{\partial \varepsilon_t(\beta)}{\partial \beta} = 0_{p+q}$

Definition 1 (L_1-based Self-weighted moment function)

$$g_t^*(\beta) := w_t \mathrm{sign}\left\{\varepsilon_t(\beta)\right\} A_t(\beta) \quad (t = u+1, \ldots, n),$$

where

- $u \geq \max\{p, q\} + 1$ (starting point),
- $A_t(\beta) = \dfrac{\partial \varepsilon_t(\beta)}{\partial \beta}$.

Fundamental settings (cont.)

Example of self-weights (Pan et al. (2007))

$$w_t = \left(1 + \sum_{k=1}^{t-1} k^{-\gamma} |y_{t-k}|\right)^{-2} \qquad (\gamma > 2)$$

Remark 3

Numerical results are not sensitive w.r.t. choice of γ or w_t's.

Fundamental settings (cont.)

Definition 2 (L_1-SW-EL statistic)

$$r_n^*(\beta) = \sup\left\{\prod_{t=u+1}^{n} nv_t : \sum_{t=u+1}^{n} v_t g_t^*(\beta) = 0_{p+q}, \sum_{t=u+1}^{n} v_t = 1, 0 \le v_t \le 1\right\}$$

Remark 4

$r_n^*(\beta)$: Nonparametric likelihood ratio function

Outline

Main results

Let $Q_t = (V_{t-1}, \ldots, V_{t-p}, W_{t-1}, \ldots, W_{t-q})^\top$,

$$V_t - b_{01} V_{t-1} - \cdots - b_{0p} V_{t-p} = -e_t,$$
$$W_t + a_{01} W_{t-1} + \cdots + a_{0q} W_{t-q} = -e_t.$$

Assumption 1

(i) β_0: the unique solution to $\mathbb{E}[g_t^*(\beta_0)] = 0_{p+q}$.

(ii) $\beta_0 \in \mathrm{Int}(\mathcal{B})$ & $\mathcal{B}$ is compact in $\mathbb{R}^{p+q}$.

(ii) $b(z) = 1 + \beta_1 z + \cdots + \beta_p z^p \neq 0$ & $a(z) = 1 + a_1 z + \cdots + a_q z^q \neq 0$ in $\{z : |z| \leq 1\}$.

(iii) $b(z)$ & $a(z)$ have no common zeros.

(iv) $^\exists \delta > 0$ s.t. $\mathbb{E}[|e_t|^\delta] < \infty$.

(iv) $\mathbb{E}[(w_t + w_t^2)(\|Q_t(\beta_0)\|^2 + \|Q_t(\beta_0)\|^3)] < \infty$.

(v) $\Omega = \mathbb{E}[w_t^2 Q_t Q_t^\top]$ is nonsingular.

Main results (cont.)

Theorem 1

Suppose that Assumption 1 holds. Then, under $H : R\beta = c$,

$$\rho_n^* \xrightarrow{\mathcal{L}} \chi_r^2 \text{ as } n \to \infty.$$

Main results (cont.)

Remark 5

- Limit distribution $\cdots$ pivotal $\Leftrightarrow$ does not contain any unknown quantity.
- Test for nested linear hypothesis
-

$$\text{SN: } \hat{g}(\beta_0) = \frac{1}{n-u} \sum_{t=u+1}^{n} \frac{1}{\{\sum_{s=1}^{n} y_s^2\}^{1/2}} g_t(\beta_0)$$

$\rightarrow$ Limit distribution contains α (tail-index of e_t)

$\rightarrow$ Rate of convergence contains α & $\neq \sqrt{n}$

$$\text{SW: } \sqrt{n}\hat{g}^*(\beta_0) = \frac{\sqrt{n}}{n-u} \sum_{t=u+1}^{n} w_t \text{sign}\{\varepsilon_t(\beta_0)\} A_t(\beta_0) \xrightarrow{\mathcal{L}} N(0_{p+q}, \Omega)$$

Outline

Example 1: Test of serial correlation

- Model: ARMA(1,1) process

$$y_t = by_{t-1} + e_t + ae_{t-1},$$

where

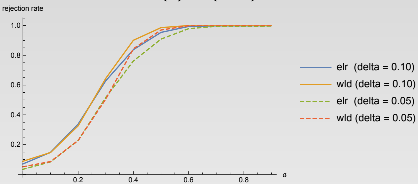
  - $|b| < 1$, $|a| < 1$ & $b + a \neq 0$
  - $e_t \sim \begin{cases} \text{(a)} & N(0,1) \\ \text{(b)} & (5/3)^{-1/2} t_5 \\ \text{(c)} & (3)^{-1/2} t_3 \\ \text{(d)} & 2^{-1/2} L(0,1) \end{cases}$
- Testing problem: $H : a = 0$
- b: a nuisance parameter

- Sample size: $n = 200, 400$
- Number of iteration: 1000
- Nominal level: $\delta = 0.10, 0.05$
- SW-Wald-test (Pan et al. (2007))

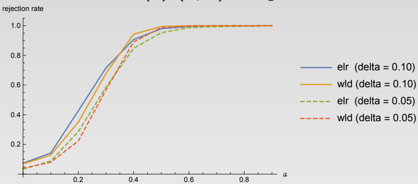
Example 1: Test of serial correlation (cont.)

Figure: Rejection rate of the tests ($b = 0.5$, $n = 200$)

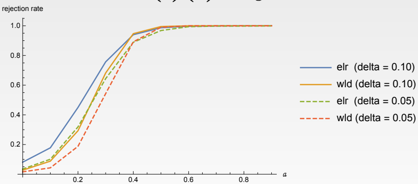
(a) $N(0,1)$



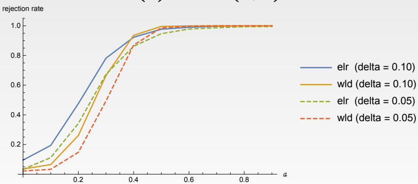
(b) $(5/2)^{-1/2}t_5$



(c) $(3)^{-1/2}t_3$



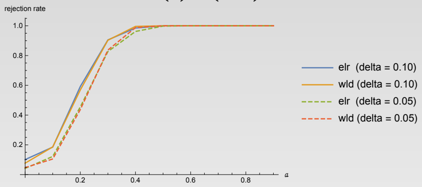
(d) $2^{-1/2}L(0,1)$



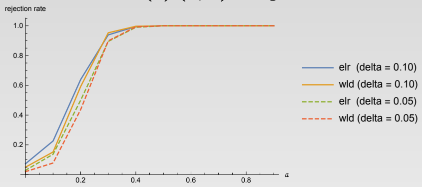
Example 1: Test of serial correlation (cont.)

Figure: Rejection rate of the tests ($b = 0.5$, $n = 400$)

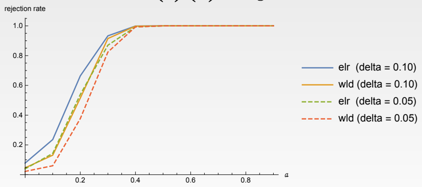
(a) $N(0,1)$



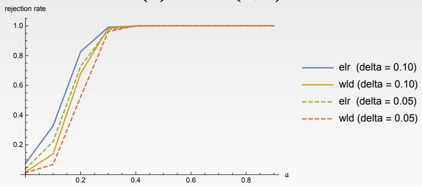
(b) $(5/2)^{-1/2}t_5$



(c) $(3)^{-1/2}t_3$



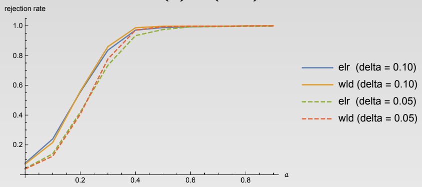
(d) $2^{-1/2}L(0,1)$



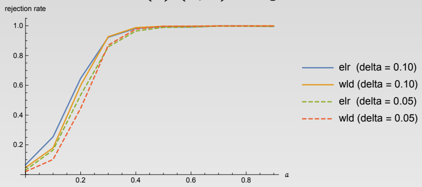
Example 1: Test of serial correlation (cont.)

Figure: Rejection rate of the tests ($b = 0.9$, $n = 200$)

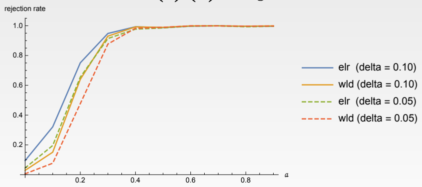
(a) $N(0,1)$



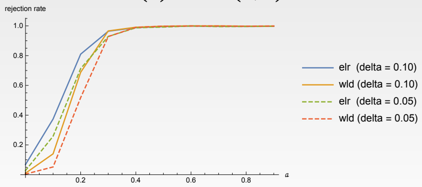
(b) $(5/2)^{-1/2}t_5$



(c) $(3)^{-1/2}t_3$

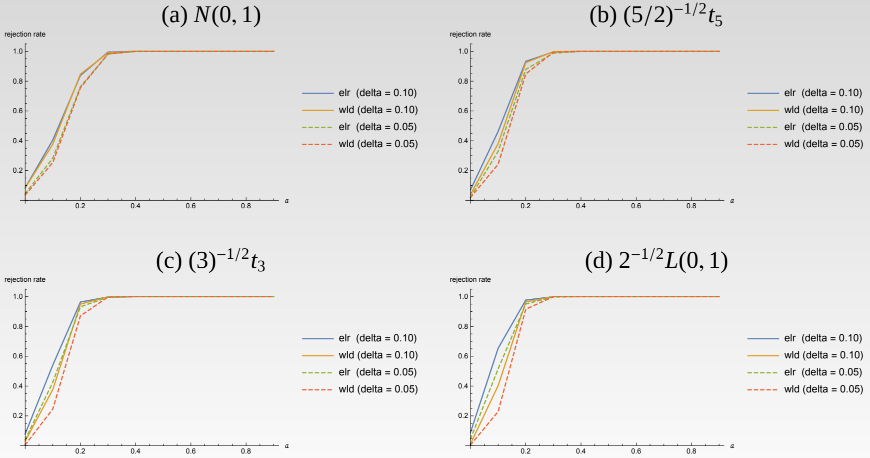


(d) $2^{-1/2}L(0,1)$



Example 1: Test of serial correlation (cont.)

Figure: Rejection rate of the tests ($b = 0.9$, $n = 400$)



Outline

Concluding remarks

- Pivotal limit distribution
  $\Leftrightarrow$ We do not need to estimate unknown parameters (e.g., α & σ of stable distribution)
- L_1-SW-EL-test improves power of test in $\left\{ \begin{array}{ll} \text{(i)} & \text{heavy-tailed noise case} \\ \text{(ii)} & \text{near unit-root} \end{array} \right\}$.

Extension 1: SW-GEL test statistics (EL, ET, CU)

Extension 2: FD-SW-GEL test statistics $\rightarrow$ Nonparametric model & quantities

Thank you for listening.

Reference

[1] Akashi, F.
(Submitted for publication).
Self-weighted empirical likelihood method for testing problem of infinite variance processes.

[2] Akashi, F., Liu, Y., and Taniguchi, M.
(2015).
An empirical likelihood approach for symmetric α-stable processes.
Bernoulli, 21(4):2093–2119.

[3] Chen, K., Ying, Z., Zhang, H., and Zhao, L.
(2008).
Analysis of least absolute deviation.
Biometrika, 95(1):107–122.

[4] Hall, P. and Heyde, C.
(1980).
Martingale limit theory and applications.
Academic, New York.

Reference (cont.)

[5] Ling, S.
(2005).
Self-weighted least absolute deviation estimation for infinite variance autoregressive models.
Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67(3):381–393.

[6] Owen, A. B.
(1988).
Empirical likelihood ratio confidence intervals for a single functional.
Biometrika, 75(2):237–249.

[7] Pan, J., Wang, H., and Yao, Q.
(2007).
Weighted least absolute deviations estimation for arma models with infinite variance.
Econometric Theory, 23(05):852–879.

[8] Parente, P. M. and Smith, R. J.
(2011).
Gel methods for nonsmooth moment indicators.
Econometric Theory, 27(01):74–113.

Appendix: SW-Wald-test (Pan et al. (2007))

$$W_n = n(R\hat{\beta}_n - c)^\top \left\{ \frac{1}{4\hat{f}(0)^2} R\hat{\Sigma}^{-1}\hat{\Omega}\hat{\Sigma}^{-1}R^\top \right\}^{-1} (R\hat{\beta}_n - c) = \frac{n\hat{a}_n^2}{\hat{w}},$$

where

$$\hat{\beta}_n = (\hat{b}_n, \hat{a}_n)^\top = \arg\min_{\beta \in \mathcal{B}} \sum_{t=u+1}^{n} w_t |\varepsilon_t(\beta)|,$$

$$\hat{f}(0) = \frac{1}{1.06 n^{-1/5} \sum_{t=u+1}^{n} w_t} \sum_{t=u+1}^{n} w_t K\left(\frac{\varepsilon_t(\hat{\beta}_n)}{1.06 n^{-1/5}} \right),$$

$$K(x) = \frac{\exp(-x)}{\{1 + \exp(-x)\}^2}, \quad \hat{w} = \frac{1}{4\hat{f}(0)^2} R\hat{\Sigma}^{-1}\hat{\Omega}\hat{\Sigma}^{-1}R^\top,$$

$$\hat{\Sigma} = \frac{1}{n-u} \sum_{t=u+1}^{n} w_t A_t(\hat{\beta}_n) A_t(\hat{\beta}_n)^\top, \quad \hat{\Omega} = \frac{1}{n-u} \sum_{t=u+1}^{n} w_t^2 A_t(\hat{\beta}_n) A_t(\hat{\beta}_n)^\top.$$

$\Rightarrow W_n \xrightarrow{\mathcal{L}} \chi_r^2$