Supplementary Online Appendix

to

Quasi-Maximum Likelihood Estimation and Bootstrap Inference in Fractional Time Series Models with Heteroskedasticity of Unknown Form

by

G. Cavaliere, M.Ø. Nielsen and A.M.R. Taylor

Date: March 5, 2017

S.1 Introduction

This supplement contains additional Monte Carlo results and proofs for our paper "Quasi-Maximum Likelihood Estimation and Bootstrap Inference in Fractional Time Series Models with Heteroskedasticity of Unknown Form". Equation references (S.n) for $n \geq 1$ refer to equations in this supplement and other equation references are to the main paper.

The supplement is organised as follows. Section S.2 presents additional Monte Carlo results and in Sections S.3 and S.4, respectively, we give proofs of the preliminary lemmas in Appendix A and of the variation bounds lemmas in Appendix B. Additional proofs for the QML estimator and the asymptotic tests are provided in Section S.5 and additional proofs for bootstrap inference are in Section S.6. All additional references are included at the end of the supplement.

For this entire supplement, all stated results and derivations shall be taken as conditional on $\sigma(\cdot)$. Due to the stochastic independence of $\{\sigma_t\}$ and $\{z_t\}$, see Assumption 1(b), and given the simple structure of conditional distributions on product spaces, this implies that $\{\sigma_t\}$ can be treated as fixed. In order to avoid repetition, this will not be repeated on every occasion. Where convergence obtains to a limit which does not depend on $\sigma(\cdot)$, it should be recalled that the stated convergence result also holds unconditionally.

S.2 Additional Monte Carlo Results

Tables S.1 and S.2 report results relating to tests on the autoregressive parameter a in (30). In particular, results are reported for the asymptotic LM_T, LR_T, W_T and RW_T tests of $H_{0,2} : a = 0$ against $H_{1,2} : a \neq 0$, along with their restricted wild bootstrap (Algorithm 1) and unrestricted wild bootstrap (Algorithm 2) counterparts. Finite sample size and power results are reported for $a = 0$ and $a = 1 + 5/\sqrt{T}$, respectively, in (30). Table S.1 relates to the case of a one-time shift in volatility, while Table S.2 reports results for the conditionally heteroskedastic Models A-I outlined in Section 5.3. Tables S.3 and S.4 report corresponding results for the joint tests of $H_{0,3} : d = 1 \cap a = 0$ against $H_{1,3} : d \neq 1 \cup a \neq 0$, with finite sample size and power results reported for $d = 1, a = 0$ and $\delta = (1, 5/3)'$, in (1) and (30). Finally, Figure S.1 reports finite sample power functions for the bootstrap tests of $H_{0,1} : d = 1$ against $H_{1,1} : d \neq 1$ for a range of values of δ in $d = 1 + \delta/\sqrt{T}$. The Monte Carlo DGP and set-up of these experiments are exactly as detailed in Section 5.1.

S.3 Proofs of Preliminary Lemmas

S.3.1 Proof of Lemma A.1

The result for condition (ii)(a) follows from Theorem 2.3 of McLeish (1974) and the comments in the two paragraphs following it. For condition (ii)(b) the result is Theorem 2.2 of Dvoretzky (1972).

S.3.2 Proof of Lemma A.2

The result for moments follows because $E(z_t z_{t-r_1} \cdots z_{t-r_{q-1}}) = E(E(z_t|\mathcal{F}_{t-1}) z_{t-r_1} \cdots z_{t-r_{q-1}}) = 0$ by the law of iterated expectations and the martingale difference property of z_t. To show the result for cumulants, we first have $\kappa_2(t, t-r) = E(z_t z_{t-r}) = 0$ because $r \geq 1$. When $q \geq 3$ we use the relation $E(z_t z_{t-r_1} \cdots z_{t-r_{q-1}}) = \sum_\pi \prod_{B \in \pi} \kappa(B)$, where π runs through the list of all partitions of $\{0, r_1, \ldots, r_{q-1}\}$ and B runs through the list of all blocks of the partition π. The required result then holds by induction on q because it has already been shown to hold for moments and for $q = 2$.

Table S.1: Tests of $H_{0,2}$: simulated size and power with one-time shift in unconditional volatility

τ	υ	T	λ	size				power			
				LM_T	LR_T	W_T	RW_T	LM_T	LR_T	W_T	RW_T
Panel A: asymptotic tests											
	1	100	1.00	6.32	7.93	8.93	10.34	44.73	59.89	45.36	53.25
	1	250	1.00	5.47	5.49	5.17	5.68	69.57	74.64	67.29	61.51
	1	∞	1.00	5.00	5.00	5.00	5.00	87.91	87.91	87.91	87.91
1/4	1/3	100	2.33	17.21	18.63	19.48	14.97	17.88	36.43	29.53	42.11
1/4	1/3	250	2.33	19.29	19.10	18.67	7.85	34.95	47.00	39.56	42.32
1/4	1/3	∞	2.33	19.95	19.95	19.95	5.00	53.57	53.57	53.57	53.57
1/4	3	100	1.24	9.18	11.25	12.45	12.00	36.67	52.71	29.50	41.33
1/4	3	250	1.24	8.27	8.27	7.82	5.72	59.46	65.89	58.03	54.23
1/4	3	∞	1.24	7.90	7.90	7.90	5.00	80.12	80.12	80.12	80.12
3/4	1/3	100	1.24	9.20	10.41	10.88	11.38	36.78	54.29	43.87	51.60
3/4	1/3	250	1.24	8.17	8.26	7.68	5.67	60.36	67.61	60.03	56.65
3/4	1/3	∞	1.24	7.90	7.90	7.90	5.00	80.12	80.12	80.12	80.12
3/4	3	100	2.33	19.31	21.29	22.73	16.61	14.07	32.55	20.12	33.64
3/4	3	250	2.33	19.56	19.66	19.15	8.31	31.65	44.31	35.54	38.14
3/4	3	∞	2.33	19.95	19.95	19.95	5.00	53.57	53.57	53.57	53.57
Panel B: wild bootstrap tests (Algorithm 1)											
	1	100	1.00	5.38	5.28	5.24	5.24	45.11	60.40	44.88	51.63
	1	250	1.00	5.06	5.03	5.14	5.26	69.58	74.63	67.68	62.30
	1	∞	1.00	5.00	5.00	5.00	5.00	87.91	87.91	87.91	87.91
1/4	1/3	100	2.33	6.31	6.43	6.08	5.39	20.86	39.55	31.03	42.21
1/4	1/3	250	2.33	5.88	5.97	5.87	5.12	36.70	48.28	40.22	42.24
1/4	1/3	∞	2.33	5.00	5.00	5.00	5.00	53.57	53.57	53.57	53.57
1/4	3	100	1.24	5.01	5.13	5.32	5.24	34.98	51.83	31.33	41.30
1/4	3	250	1.24	5.13	5.12	5.06	5.06	58.54	65.74	57.76	54.19
1/4	3	∞	1.24	5.00	5.00	5.00	5.00	80.12	80.12	80.12	80.12
3/4	1/3	100	1.24	5.85	5.59	5.06	5.03	38.86	55.98	42.84	50.54
3/4	1/3	250	1.24	5.00	5.09	5.17	4.93	59.96	67.61	59.65	55.69
3/4	1/3	∞	1.24	5.00	5.00	5.00	5.00	80.12	80.12	80.12	80.12
3/4	3	100	2.33	6.05	6.36	5.44	5.24	15.20	33.89	20.86	33.87
3/4	3	250	2.33	5.60	5.73	5.57	5.21	32.69	45.70	36.66	38.96
3/4	3	∞	2.33	5.00	5.00	5.00	5.00	53.57	53.57	53.57	53.57
Panel C: wild bootstrap tests (Algorithm 2)											
	1	100	1.00	6.07	8.72	7.05	6.70	47.83	72.84	31.10	39.56
	1	250	1.00	4.57	4.43	3.33	3.49	63.58	70.35	18.33	18.95
	1	∞	1.00	5.00	5.00	5.00	5.00	87.91	87.91	87.91	87.91
1/4	1/3	100	2.33	9.34	11.73	7.69	7.00	43.02	60.12	19.10	37.37
1/4	1/3	250	2.33	5.51	7.23	4.72	3.72	33.37	54.72	14.12	18.53
1/4	1/3	∞	2.33	5.00	5.00	5.00	5.00	53.57	53.57	53.57	53.57
1/4	3	100	1.24	6.50	9.64	8.21	7.96	45.47	68.69	29.03	40.55
1/4	3	250	1.24	4.42	4.86	3.47	3.52	50.12	62.99	14.78	15.56
1/4	3	∞	1.24	5.00	5.00	5.00	5.00	80.12	80.12	80.12	80.12
3/4	1/3	100	1.24	6.60	9.44	6.94	6.48	46.98	70.86	28.87	40.22
3/4	1/3	250	1.24	4.65	4.78	3.31	3.03	53.03	65.09	17.29	17.72
3/4	1/3	∞	1.24	5.00	5.00	5.00	5.00	80.12	80.12	80.12	80.12
3/4	3	100	2.33	10.42	10.84	8.70	8.78	41.91	52.38	17.80	37.52
3/4	3	250	2.33	5.19	7.18	4.88	4.19	29.05	52.17	13.66	18.27
3/4	3	∞	2.33	5.00	5.00	5.00	5.00	53.57	53.57	53.57	53.57

Notes: Entries for finite T are simulated rejection frequencies of the tests. Entries for $T = \infty$ are calculated as described in Remark 4.10. Power is measured at $\delta = 5$ and is size corrected for the asymptotic tests, but not for the bootstrap tests. All entries are based on $10,000$ Monte Carlo replications.

Table S.2: Tests of $H_{0,2}$: simulated size and power with conditionally heteroskedastic Models A–I

		size				power			
	T	LM_T	LR_T	W_T	RW_T	LM_T	LR_T	W_T	RW_T
Panel A: asymptotic tests									
Model A	100	12.88	13.78	14.42	12.18	24.12	43.50	31.57	47.63
	250	13.81	13.85	13.09	6.23	42.18	51.95	42.65	50.61
Model B	100	14.44	15.23	15.83	13.05	22.76	42.59	30.39	46.32
	250	18.04	18.08	17.16	7.52	28.33	39.96	30.25	46.19
Model C	100	11.66	12.95	13.65	12.21	27.20	46.13	35.38	46.52
	250	15.37	15.27	14.65	6.90	39.26	49.22	40.53	45.16
Model D	100	12.29	13.78	14.39	12.61	27.32	44.11	32.20	44.39
	250	19.30	19.31	18.42	8.72	27.75	40.50	31.22	41.66
Model E	100	16.39	17.28	17.76	13.78	16.42	35.97	29.51	45.83
	250	20.08	20.02	19.22	7.87	27.08	40.91	32.27	42.34
Model F	100	15.68	17.03	18.05	13.71	18.16	38.68	26.60	41.67
	250	23.28	23.18	22.82	9.19	22.06	34.90	27.77	38.90
Model G	100	14.68	16.48	17.15	13.91	19.60	37.53	26.52	39.62
	250	21.08	21.35	20.49	8.23	25.56	39.47	30.03	42.24
Model H	100	24.55	26.77	27.43	18.58	8.10	20.02	17.44	35.72
	250	35.58	36.05	35.79	12.85	8.68	17.51	15.66	29.81
Model I	100	23.48	25.82	26.82	18.33	8.43	22.57	15.96	32.56
	250	27.38	27.73	26.74	9.20	15.91	30.41	23.29	33.97
Panel B: wild bootstrap tests (Algorithm 1)									
Model A	100	6.04	6.35	5.88	5.20	30.99	48.76	36.41	47.84
	250	5.88	5.78	5.83	4.93	47.24	56.80	47.52	50.01
Model B	100	6.57	6.10	6.05	5.62	30.51	47.90	36.29	47.70
	250	6.52	6.57	6.67	5.77	43.74	53.95	44.87	47.84
Model C	100	5.72	5.72	5.29	4.98	30.78	48.83	34.86	44.83
	250	5.27	5.76	5.50	5.01	42.46	52.97	44.50	45.20
Model D	100	5.69	5.58	5.23	5.02	29.66	47.78	34.84	44.79
	250	6.68	6.59	6.66	5.83	39.31	49.98	41.99	44.10
Model E	100	6.45	6.32	5.57	4.91	23.94	42.24	32.12	43.95
	250	5.95	5.96	5.95	5.22	35.34	46.48	38.67	42.38
Model F	100	6.01	6.08	5.70	5.18	24.10	42.24	31.58	41.83
	250	6.25	6.43	6.39	5.54	31.81	43.06	35.80	39.76
Model G	100	5.91	6.34	5.97	5.64	24.95	43.32	32.75	43.66
	250	5.70	5.48	5.74	5.15	36.02	47.22	39.34	42.78
Model H	100	6.76	7.27	6.32	5.75	15.53	31.11	25.21	39.10
	250	6.98	7.11	7.21	6.02	18.40	29.85	24.85	32.86
Model I	100	6.63	7.26	6.10	5.49	13.75	29.71	19.98	33.43
	250	6.27	6.52	6.52	5.64	23.50	36.85	28.83	35.00
Panel C: wild bootstrap tests (Algorithm 2)									
Model A	100	8.70	10.53	7.70	7.10	47.96	65.68	23.12	38.38
	250	5.99	6.46	4.60	3.49	47.14	61.42	16.13	17.53
Model B	100	9.49	11.08	8.19	7.54	49.29	63.18	21.79	38.10
	250	6.94	7.62	5.38	4.25	45.13	58.62	16.63	18.64
Model C	100	7.58	10.21	7.38	7.09	44.98	65.80	24.77	38.66
	250	4.93	6.43	4.09	3.61	39.65	56.41	14.98	17.82
Model D	100	7.90	10.10	7.41	7.13	45.32	64.33	23.88	38.18
	250	6.56	7.76	5.06	4.48	38.69	55.01	14.40	18.47
Model E	100	9.52	11.18	7.80	6.82	45.44	61.11	21.03	36.92
	250	6.47	7.50	4.78	3.66	37.42	54.36	13.93	18.25
Model F	100	8.82	10.49	7.83	7.43	44.70	61.64	22.01	37.45
	250	6.75	8.00	5.23	4.37	33.54	51.87	13.12	18.57
Model G	100	8.81	11.11	8.04	7.73	44.96	62.65	22.88	37.77
	250	6.03	7.31	4.55	3.92	37.10	54.98	13.82	19.30
Model H	100	13.54	12.14	8.53	8.64	45.57	51.34	15.17	35.51
	250	9.53	10.31	6.33	5.41	31.01	42.05	9.83	19.29
Model I	100	13.42	12.40	8.95	9.51	43.72	47.81	14.78	36.28
	250	7.51	9.33	6.09	4.95	29.67	47.84	11.56	17.97

Notes: Entries are simulated rejection frequencies of the tests. Power is measured at $\delta = 5$ and is size corrected for the asymptotic tests, but not for the bootstrap tests. All entries are based on $10,000$ Monte Carlo replications.

Table S.3: Tests of $H_{0,3}$: simulated size and power with one-time shift in unconditional volatility

τ	υ	T	λ	size LM_T	size LR_T	size W_T	size RW_T	power LM_T	power LR_T	power W_T	power RW_T
Panel A: asymptotic tests											
	1	100	1	7.73	6.50	9.89	12.53	61.46	68.19	43.14	32.92
	1	250	1	5.91	5.39	5.77	6.28	65.53	69.32	64.89	63.80
	1	∞	1	5.00	5.00	5.00	5.00	70.33	70.33	70.33	70.33
1/4	1/3	100	2.33	25.46	23.32	27.03	20.06	29.92	42.28	21.10	12.96
1/4	1/3	250	2.33	27.27	26.06	26.56	11.08	32.89	39.43	31.17	21.50
1/4	1/3	∞	2.33	27.70	27.70	27.70	5.00	35.24	35.24	35.24	35.24
1/4	3	100	1.24	12.39	11.13	15.22	14.43	51.35	60.51	17.82	13.66
1/4	3	250	1.24	10.61	9.54	10.23	7.29	56.25	60.44	54.47	51.81
1/4	3	∞	1.24	9.01	9.01	9.01	5.00	60.13	60.13	60.13	60.13
3/4	1/3	100	1.24	12.30	10.40	13.94	14.25	49.18	59.45	33.59	23.29
3/4	1/3	250	1.24	9.87	9.16	9.39	7.23	56.98	61.84	55.31	51.39
3/4	1/3	∞	1.24	9.01	9.01	9.01	5.00	60.13	60.13	60.13	60.13
3/4	3	100	2.33	28.40	26.30	31.05	21.87	26.24	40.12	14.53	11.84
3/4	3	250	2.33	27.56	26.84	26.93	11.82	31.00	37.83	28.66	19.49
3/4	3	∞	2.33	27.70	27.70	27.70	5.00	35.24	35.24	35.24	35.24
Panel B: wild bootstrap tests (Algorithm 1)											
	1	100	1	5.45	5.06	5.20	5.12	62.06	68.10	40.31	33.03
	1	250	1	5.22	5.15	5.11	4.92	66.00	69.13	64.79	62.66
	1	∞	1	5.00	5.00	5.00	5.00	70.33	70.33	70.33	70.33
1/4	1/3	100	2.33	7.28	7.49	6.22	5.34	34.71	45.72	25.12	14.56
1/4	1/3	250	2.33	6.54	6.27	6.29	5.29	35.75	41.49	33.42	22.35
1/4	1/3	∞	2.33	5.00	5.00	5.00	5.00	35.24	35.24	35.24	35.24
1/4	3	100	1.24	5.51	5.47	5.54	5.32	52.29	60.57	22.38	17.57
1/4	3	250	1.24	5.26	5.28	5.17	5.06	56.80	60.56	54.72	51.22
1/4	3	∞	1.24	5.00	5.00	5.00	5.00	60.13	60.13	60.13	60.13
3/4	1/3	100	1.24	6.27	6.08	5.33	5.23	53.47	61.75	34.36	23.72
3/4	1/3	250	1.24	5.61	5.30	5.44	5.30	57.65	62.09	56.27	51.82
3/4	1/3	∞	1.24	5.00	5.00	5.00	5.00	60.13	60.13	60.13	60.13
3/4	3	100	2.33	7.16	7.79	5.68	5.19	30.04	43.15	17.38	13.04
3/4	3	250	2.33	6.24	6.52	6.48	5.70	34.38	40.31	32.02	21.84
3/4	3	∞	2.33	5.00	5.00	5.00	5.00	35.24	35.24	35.24	35.24
Panel C: wild bootstrap tests (Algorithm 2)											
	1	100	1	5.40	6.63	5.97	5.73	59.50	70.81	23.94	17.82
	1	250	1	4.82	4.94	3.86	3.76	64.73	67.84	53.59	48.69
	1	∞	1	5.00	5.00	5.00	5.00	70.33	70.33	70.33	70.33
1/4	1/3	100	2.33	8.38	8.66	7.63	6.63	37.11	48.74	21.61	14.43
1/4	1/3	250	2.33	6.20	6.81	5.05	3.76	34.06	42.98	21.54	9.97
1/4	1/3	∞	2.33	5.00	5.00	5.00	5.00	35.24	35.24	35.24	35.24
1/4	3	100	1.24	6.00	7.47	7.29	7.12	51.16	64.42	22.26	17.62
1/4	3	250	1.24	4.91	5.32	3.92	3.61	54.82	59.62	36.76	30.14
1/4	3	∞	1.24	5.00	5.00	5.00	5.00	60.13	60.13	60.13	60.13
3/4	1/3	100	1.24	6.47	7.67	6.02	5.56	51.91	64.94	23.15	15.84
3/4	1/3	250	1.24	5.25	5.28	4.05	3.70	56.55	61.22	43.37	34.56
3/4	1/3	∞	1.24	5.00	5.00	5.00	5.00	60.13	60.13	60.13	60.13
3/4	3	100	2.33	9.31	8.70	8.49	8.28	34.99	45.86	21.08	17.25
3/4	3	250	2.33	6.07	7.23	5.15	4.07	32.49	41.83	17.86	10.80
3/4	3	∞	2.33	5.00	5.00	5.00	5.00	35.24	35.24	35.24	35.24

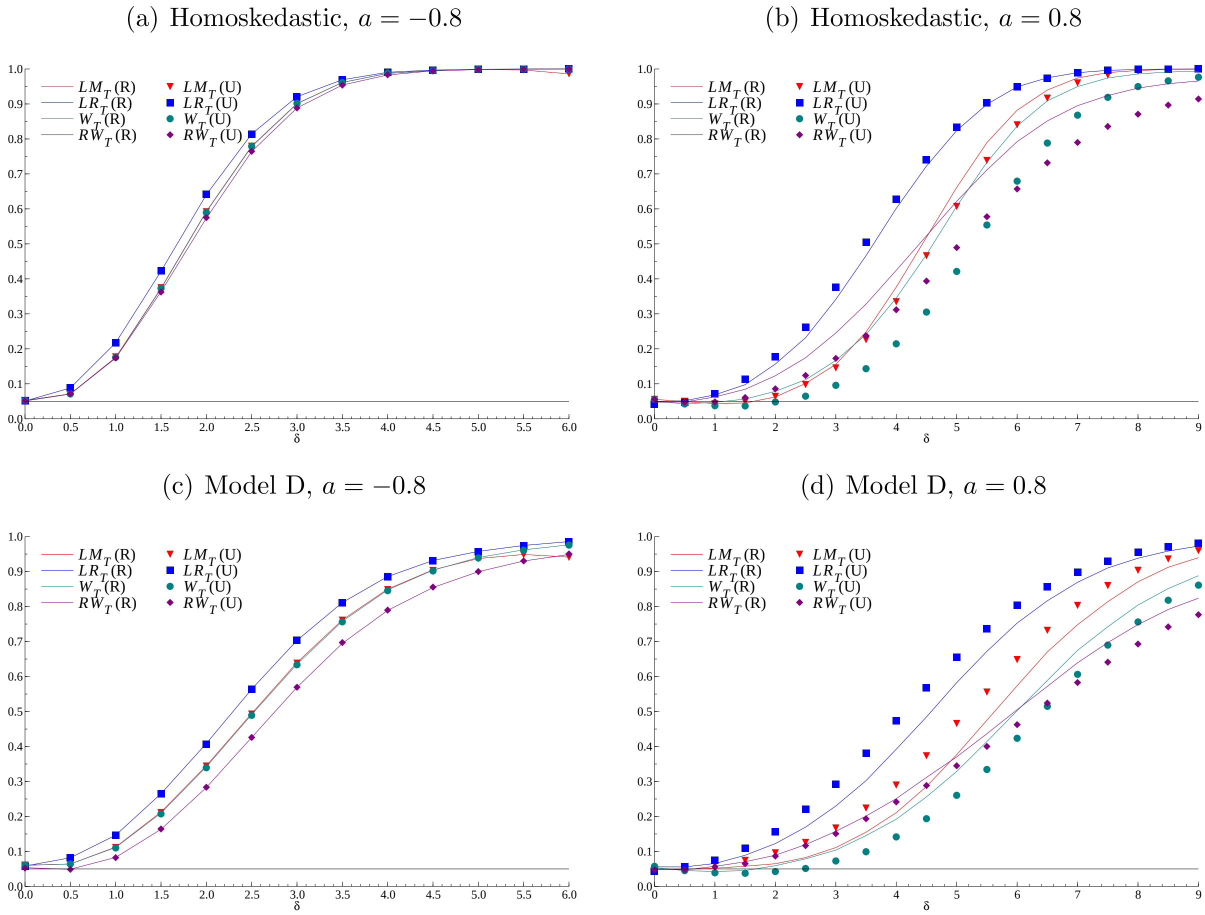
Notes: Entries for finite T are simulated rejection frequencies of the tests. Entries for $T = \infty$ are calculated as described in Remark 4.10. Power is measured at $\delta = [1, 5/3]'$ and is size corrected for the asymptotic tests, but not for the bootstrap tests. All entries are based on 10,000 Monte Carlo replications.

Table S.4: Tests of $H_{0,3}$: simulated size and power with conditionally heteroskedastic Models A–I

	T	size				power			
		LM_T	LR_T	W_T	RW_T	LM_T	LR_T	W_T	RW_T
Panel A: asymptotic tests									
Model A	100	19.87	18.05	20.98	15.47	33.65	42.68	21.37	17.13
	250	21.71	20.42	20.84	9.12	33.38	37.22	31.97	27.24
Model B	100	23.15	20.87	24.11	17.24	29.41	38.24	19.39	15.65
	250	28.15	26.92	27.41	11.17	21.07	25.05	19.82	22.55
Model C	100	16.90	14.75	18.51	15.92	40.67	51.81	25.13	17.37
	250	21.49	20.42	21.12	9.74	34.78	40.13	32.04	28.54
Model D	100	17.84	15.57	19.64	16.24	37.35	47.97	22.00	16.58
	250	26.11	24.98	25.59	11.24	25.06	31.29	23.08	23.05
Model E	100	23.92	21.63	24.80	17.83	28.73	39.72	20.13	15.40
	250	29.48	28.24	28.91	11.07	25.09	30.88	23.13	20.32
Model F	100	22.60	20.24	24.08	17.74	31.49	41.58	17.85	12.76
	250	32.62	31.61	32.07	12.35	21.36	26.53	19.58	16.63
Model G	100	21.58	19.52	23.06	18.21	29.23	40.78	17.60	12.57
	250	29.97	28.57	29.38	11.36	22.40	27.79	19.63	16.84
Model H	100	38.16	37.24	40.48	26.25	15.18	25.09	12.01	10.60
	250	53.11	52.29	53.24	20.04	10.11	13.86	9.17	8.62
Model I	100	36.30	35.21	39.21	24.39	16.12	24.88	13.07	11.24
	250	41.17	39.82	40.48	14.15	16.34	20.71	14.68	11.49
Panel B: wild bootstrap tests (Algorithm 1)									
Model A	100	7.13	7.31	6.24	5.15	41.76	50.46	29.35	18.82
	250	6.89	6.79	6.91	5.65	41.30	45.70	40.48	30.64
Model B	100	8.21	8.92	7.19	5.68	41.80	50.30	31.42	20.63
	250	8.75	9.09	8.94	6.20	38.46	42.33	38.10	28.06
Model C	100	6.39	6.50	5.32	4.90	45.76	55.31	29.20	19.90
	250	6.12	6.20	6.15	5.38	42.07	47.41	40.52	31.65
Model D	100	6.42	6.49	5.38	5.19	44.25	54.09	29.11	20.55
	250	7.22	7.24	6.93	6.09	38.96	44.23	37.21	28.55
Model E	100	7.07	7.79	5.91	4.92	37.63	48.20	27.25	17.24
	250	6.62	6.46	6.48	5.33	34.73	39.94	33.71	22.83
Model F	100	6.55	7.06	6.08	5.50	38.78	49.18	26.03	16.67
	250	6.98	6.84	6.84	5.41	31.90	37.15	30.39	20.35
Model G	100	6.68	6.68	5.85	5.60	37.58	47.85	25.36	17.51
	250	6.34	6.35	6.49	5.72	32.44	37.84	31.18	22.10
Model H	100	8.66	10.12	7.85	6.15	26.63	39.18	22.18	14.54
	250	9.13	10.00	9.12	6.28	23.11	29.36	22.35	13.67
Model I	100	8.43	9.68	6.98	5.49	24.54	36.25	18.66	13.16
	250	8.07	8.44	8.18	5.80	26.27	31.19	24.76	15.06
Panel C: wild bootstrap tests (Algorithm 2)									
Model A	100	7.58	8.80	7.71	6.22	41.75	53.33	23.74	14.47
	250	6.78	7.21	5.73	4.06	40.36	46.25	32.16	17.36
Model B	100	9.09	9.99	8.58	6.92	41.89	52.52	25.69	15.25
	250	8.72	9.39	7.53	4.66	37.71	43.16	31.31	15.86
Model C	100	7.10	7.81	7.01	6.34	45.74	58.42	23.29	16.21
	250	5.99	6.58	4.81	3.85	40.88	48.11	29.44	18.10
Model D	100	6.99	7.74	6.88	6.43	45.10	56.94	23.19	16.47
	250	7.09	7.56	5.81	4.64	37.78	45.11	26.83	16.47
Model E	100	8.12	8.97	7.59	6.02	39.02	50.87	23.25	15.21
	250	6.60	6.95	5.50	3.77	34.29	41.21	24.88	12.41
Model F	100	7.61	8.14	7.51	7.06	40.09	51.94	22.51	15.68
	250	7.13	7.29	5.85	4.30	31.23	38.47	21.43	11.02
Model G	100	7.79	7.89	7.60	7.13	38.87	50.89	21.80	15.79
	250	6.34	6.94	5.47	4.29	31.66	39.23	22.09	11.96
Model H	100	10.49	10.93	9.40	8.99	30.29	41.67	20.53	15.20
	250	9.89	10.41	8.23	5.69	24.18	30.99	17.25	9.01
Model I	100	10.34	10.48	10.03	9.34	28.66	38.31	21.86	17.18
	250	8.42	8.97	7.40	4.91	25.75	32.83	16.86	8.22

Notes: Entries are simulated rejection frequencies of the tests. Power is measured at $\delta = [1, 5/3]'$ and is size corrected for the asymptotic tests, but not for the bootstrap tests. All entries are based on 10,000 Monte Carlo replications.

Figure S.1: Finite sample power functions of bootstrap tests of $H_{0,1}$ with weakly dependent errors



Notes: Entries are simulated rejection frequencies of the tests measured at $\delta \in \{0.0, 0.5, 1.0, \ldots\}$. The notation (R) and (U) denotes the restricted and unrestricted bootstrap algorithms, respectively. All entries are based on 10,000 Monte Carlo replications.

S.3.3 Proofs of Lemmas A.3 and A.4

For the proof of Lemma A.3, see Lemma A.1 of Nielsen (2015) and Lemma B.3 of Johansen and Nielsen (2010), and for the proof of Lemma A.4, see Lemma B.4 of Johansen and Nielsen (2010).

S.3.4 Proof of Lemma A.5

First notice that

$$\left\| T^{-1} \sum_{t=1}^{T} \sigma_t^2 \sum_{n,m=1}^{t-1} \xi_n \xi_m' \sigma_{t-n} \sigma_{t-m} g_{t,n,m} - T^{-1} \sum_{t=1}^{T} \sigma_t^4 \sum_{n,m=1}^{t-1} \xi_n \xi_m' g_{t,n,m} \right\|$$

$$= \left\| T^{-1} \sum_{t=1}^{T} \sigma_t^2 \sum_{n,m=1}^{t-1} \xi_n \xi_m' (\sigma_{t-n} \sigma_{t-m} - \sigma_t^2) g_{t,n,m} \right\|$$

$$\leq K T^{-1} \sum_{t=1}^{T} \sum_{n=1}^{t-1} \sum_{m=n}^{t-1} ||\xi_n|| ||\xi_m|| |\sigma_{t-n} \sigma_{t-m} - \sigma_t^2| |g_{t,n,m}| = K(r_{1T} + r_{2T}),$$

where the inequality follows by Assumption 1(b)(i) and by symmetry in n and m, and where we defined

$$r_{1T} := \sum_{n=1}^{q_T}\sum_{m=n}^{q_T} ||\xi_n||||\xi_m|| \sup_t |g_{t,n,m}| T^{-1} \sum_{t=m+1}^{T} |\sigma_{t-n}\sigma_{t-m} - \sigma_t^2|,$$

$$r_{2T} := \sum_{n=1}^{T-1}\sum_{m=\max(n,q_T+1)}^{T-1} ||\xi_n||||\xi_m|| \sup_t |g_{t,n,m}| T^{-1} \sum_{t=m+1}^{T} |\sigma_{t-n}\sigma_{t-m} - \sigma_t^2|.$$

Let $q_T := \lfloor T^{\varkappa} \rfloor$ for $\varkappa \in (0,1)$ and $G := \sup_{t\in\mathbb{Z}} \sigma_t$, which is finite by Assumption 1(b)(i). Then

$$\left|\sigma_{t-n}\sigma_{t-m} - \sigma_t^2\right| \leq \sigma_t \left|\sigma_{t-n} - \sigma_t\right| + \sigma_{t-n}\left|\sigma_{t-m} - \sigma_t\right| \leq G\left(\left|\sigma_{t-n} - \sigma_t\right| + \left|\sigma_{t-m} - \sigma_t\right|\right)$$

such that, for $m \geq n \geq 1$,

$$\sum_{t=m+1}^{T} \left|\sigma_{t-n}\sigma_{t-m} - \sigma_t^2\right| \leq G \sum_{t=m+1}^{T} \left(\left|\sigma_{t-n} - \sigma_t\right| + \left|\sigma_{t-m} - \sigma_t\right|\right) \leq 2G \sum_{t=1}^{T-m} \left|\sigma_{t+m} - \sigma_t\right|.$$

Hence, using the fact that $\sigma_t = \sigma(t/T) \in \mathcal{D}([0,1])$ for $t = 1, \dots, T$, see Assumption 1(b)(ii),

$$\sup_{n,m=1,\dots,q_T} T^{-1} \sum_{t=m+1}^{T} |\sigma_{t-n}\sigma_{t-m} - \sigma_t^2| \leq 2G \sup_{m=1,\dots,q_T} T^{-1} \sum_{t=1}^{T-m} |\sigma_{t+m} - \sigma_t| \to 0 \text{ as } T \to \infty \tag{S.1}$$

by Lemma A.1 in Cavaliere and Taylor (2009). Now write

$$r_{1T} \leq \left(\sup_{n,m=1,\dots,q_T} T^{-1} \sum_{t=m+1}^{T} \left|\sigma_{t-n}\sigma_{t-m} - \sigma_t^2\right| \right) r_{11T}$$

with $r_{11T} := \sup_t \sum_{n=1}^{q_T} \sum_{m=n}^{q_T} ||\xi_n||||\xi_m|||g_{t,n,m}| < \infty$ by assumption. Because the first factor in r_{1T} converges to zero as $T \to \infty$ by (S.1), it follows that $r_{1T} \to 0$ as $T \to \infty$.

The term r_{2T} is bounded as, by another application of Assumption 1(b),

$$\begin{aligned} r_{2T} &\leq 4G^2 \sum_{n=1}^{T-1}\sum_{m=\max(n,q_T+1)}^{T-1} ||\xi_n||||\xi_m|| \sup_t |g_{t,n,m}| \\ &\leq 4G^2 \sum_{m=q_T+1}^{\infty}\sum_{n=1}^{\infty} ||\xi_n||||\xi_m|| \sup_t |g_{t,n,m}| \to 0 \end{aligned}$$

as $T \to \infty$ because it is a tail sum ($q_T \to \infty$) of the convergent sum $\sum_{n,m=1}^{\infty} ||\xi_n||||\xi_m|| \sup_t |g_{t,n,m}|$. This completes the proof.

S.4 Proofs of Variation Bounds Lemmas

We first present a lemma which contains uniform bounds on coefficient summations, which are used to prove the variation bounds in Lemmas B.1–B.3.

Lemma S.1. *Let* $\xi_T(u,v,k) := \max_{1\leq n,m\leq T} \sum_{t=\max(n,m)}^{T} |\zeta_{t-n}(-u,k)\zeta_{t-m}(-v,k)|$ *for coefficients* $\zeta_j(u,k)$ *satisfying* $\zeta_0(u,k) = 1$ *and* $\zeta_j(u,k) \leq c(\log j)^k j^{u-1}$ *for* $j \geq 1$*, where* $c > 0$ *does not depend on* u*,* k*, or* j*. Then:*

(i) *Uniformly for* $\min(u+1, v+1, u+v+1) \geq a$, *it holds that*

$$\xi_T(u,v,k) \leq \begin{cases} c(1+\log T)^{1+2k} T^{-a} & \text{if } a \leq 0, \\ c & \text{if } a > 0, \end{cases}$$

where the constant $c > 0$ *does not depend on* u, v, *or* T.

(ii) *For any* $u > 0, v > 0$ *it holds that*

$$\sum_{t=0}^{\infty} |\zeta_{|t-n|}(-u,k)\zeta_t(-v,k)| \leq c(\log|n|)^k |n|^{\max(-u-1,-v-1)},$$

where the constant $c > 0$ *does not depend on* u, v, *or* n.

S.4.1 Proof of Lemma S.1

Part (i) is Lemma A.7 of Johansen and Nielsen (2012). To show part (ii) when $n \geq 0$ we split the summation and find the bound

$$\sum_{t=0}^{\lfloor n/2 \rfloor} |\zeta_{|t-n|}(-u,k)\zeta_t(-v,k)| + \sum_{t=\lfloor n/2 \rfloor+1}^{n} |\zeta_{|t-n|}(-u,k)\zeta_t(-v,k)| + \sum_{t=n+1}^{\infty} |\zeta_{t-n}(-u,k)\zeta_t(-v,k)|$$

$$\leq c \sum_{t=0}^{\lfloor n/2 \rfloor} (n-t)^{-u-1}(\log(n-t))^k t^{-v-1}(\log t)^k + c \sum_{t=\lfloor n/2 \rfloor+1}^{n} (n-t)^{-u-1}(\log(n-t))^k t^{-v-1}(\log t)^k$$

$$+ \sum_{t=n+1}^{\infty} (t-n)^{-u-1}(\log(t-n))^k t^{-v-1}(\log t)^k$$

$$\leq c(n/2)^{-u-1}(\log(n/2))^k \sum_{t=0}^{\lfloor n/2 \rfloor} t^{-v-1}(\log t)^k + c(n/2)^{-v-1}(\log(n/2))^k \sum_{t=\lfloor n/2 \rfloor+1}^{n} (n-t)^{-u-1}(\log(n-t))^k$$

$$+ c(n+1)^{-v-1}(\log(n+1))^k \sum_{t=n+1}^{\infty} (t-n)^{-u-1}(\log(t-n))^k$$

$$\leq c(\log n)^k n^{\max(-u-1,-v-1)}.$$

When $n < 0$ we find the bound

$$c\sum_{t=0}^{\infty} (t-n)^{-u-1}(\log(t-n))^k t^{-v-1}(\log t)^k$$

$$\leq c(-n)^{-u-1}(\log(-n))^k \sum_{t=0}^{\infty} t^{-v-1}(\log t)^k \leq c|n|^{-u-1}(\log|n|)^k.$$

S.4.2 Proof of Lemma B.1

The proof is given in Lemma C.3 in Johansen and Nielsen (2010), which also applies under Assumption 1 on ε_t in place of their i.i.d. assumption.

S.4.3 Proof of Lemma B.2

We prove that, uniformly in $-1/2-\kappa \leq v \leq u \leq -1/2+\kappa$,

$$||M_{12NT}(u)||_2 \leq c(\log T)T^{-1/2+\kappa}N^{1/2+\kappa}, \tag{S.2}$$

$$||M_{12NT}(u) - M_{12NT}(v)||_2 \leq c|u-v|(\log T)^2 T^{-1/2+\kappa}N^{1/2+\kappa}, \tag{S.3}$$

$$||M_{11NT}(u)||_2 \leq c(\log T)T^{-1/2}N^{1/2+2\kappa}, \tag{S.4}$$

$$||M_{11NT}(u) - M_{11NT}(v)||_2 \leq c|u-v|(\log T)^2 T^{-1/2}N^{1/2+2\kappa}, \tag{S.5}$$

where the constant $c > 0$ does not depend on u, v, or T. Using the condition on α, the right-hand sides of (S.2)–(S.5) all converge to zero. Pointwise convergence in probability then follows from (S.2) and (S.4) and tightness on the interval $|u+1/2| \leq \kappa$ follows from (S.3) and (S.5) using the criterion (S.38). Together this implies uniform convergence in probability.

Proof of (S.2): First evaluate

$$EM_{12NT}(u)^2 = T^{-2} E \prod_{k=1}^{2} \sum_{t_k=N+1}^{T} \sum_{n_k=0}^{N-1} \sum_{m_k=N}^{t_k-1} \pi_{n_k}(-u)\pi_{m_k}(-u)\varepsilon_{t_k-n_k}\varepsilon_{t_k-m_k}.$$

The term $E(\prod_{k=1}^{2} \varepsilon_{t_k-n_k}\varepsilon_{t_k-m_k})$ is non-zero only if the two highest subscripts are equal, see Lemma A.2. However, $n_k < N \leq m_k$ such that $t_k - n_k > t_k - m_k$ for $k = 1, 2$. This leaves only one possibility, i.e., $t_1 - n_1 = t_2 - n_2$, in which case we eliminate $n_2 = t_2 - t_1 + n_1$ and note that $|t_1 - t_2| = |n_1 - n_2| \leq N$. In this case $EM_{12NT}(u)^2$ is

$$T^{-2} \sum_{\substack{t_1,t_2=N+1 \\ |t_1-t_2|\leq N}}^{T} \sum_{n_1=0}^{N-1} \sum_{m_1=N}^{t_1-1} \sum_{m_2=\max(N,t_2-t_1+n_1)}^{t_2-1} \pi_{n_1}(-u)\pi_{t_2-t_1+n_1}(-u)\pi_{m_1}(-u)\pi_{m_2}(-u)$$
$$\times \sigma_{t_1-n_1}^2 \sigma_{t_1-m_1}\sigma_{t_2-m_2} E(z_{t_1-n_1}^2 z_{t_1-m_1} z_{t_2-m_2}). \tag{S.6}$$

If, in this expression, $t_1 - m_1 = t_2 - m_2$ we eliminate $m_2 = t_2 - t_1 + m_1$ and the expectation is $\tau_{m_1-n_1,m_1-n_1}$. Then, with $\xi_T(u, v, k)$ defined in Lemma S.1, $\sum_{n_1=0}^{N-1} \pi_{n_1}(-u)\pi_{t_2-t_1+n_1}(-u) \leq \xi_N(u, u, 0)$ and $\sum_{m_1=N}^{t_1-1} \pi_{m_1}(-u)\pi_{t_2-t_1+m_1}(-u) \leq \xi_T(u, u, 0)$ by (A.1) of Lemma A.3, so the contribution to $EM_{12NT}(u)^2$ is bounded by

$$cT^{-2} \sum_{\substack{t_1,t_2=N+1 \\ |t_1-t_2|\leq N}}^{T} \xi_N(u, u, 0)\xi_T(u, u, 0).$$

The result when $t_1 - m_1 = t_2 - m_2$ now follows from Lemma S.1(i). If, on the other hand, $t_1 - m_1 \neq t_2 - m_2$ in (S.6), the expectation in (S.6) is $\kappa_4(t_1 - n_1, t_1 - n_1, t_1 - m_1, t_2 - m_2)$ and the contribution to $EM_{12NT}(u)^2$ is bounded by

$$cT^{-2} \sum_{\substack{t_1,t_2=N+1 \\ |t_1-t_2|\leq N}}^{T} \sum_{n_1=0}^{N-1} \pi_{n_1}(-u)\pi_{t_2-t_1+n_1}(-u)\pi_N(-u)^2$$
$$\times \sum_{m_1=N}^{t_1-1} \sum_{m_2=\max(N,t_2-t_1+n_1)}^{t_2-1} |\kappa_4(t_1 - n_1, t_1 - n_1, t_1 - m_1, t_2 - m_2)|$$
$$\leq cT^{-2} \sum_{\substack{t_1,t_2=N+1 \\ |t_1-t_2|\leq N}}^{T} \xi_N(u, u, 0)N^{-2u-2}$$

using Assumption 1(a)(iii),(b), and this proves the result.

Proof of (S.3): Next consider $||M_{12NT}(u) - M_{12NT}(v)||_2$ which is bounded by

$$||T^{-1} \sum_{t=N+1}^{T} (w_{1t}(u) - w_{1t}(v))w_{2t}(u)||_2 + ||T^{-1} \sum_{t=N+1}^{T} w_{1t}(v)(w_{2t}(u) - w_{2t}(v))||_2.$$

For the first term write $w_{1t}(u)-w_{1t}(v)=\sum_{n=0}^{N-1}(\pi_n(-u)-\pi_n(-v))\varepsilon_{t-n}=(u-v)\sum_{n=0}^{N-1}\zeta_n(-u,1)\varepsilon_{t-n}$, see (A.3) of Lemma A.3 and Lemma S.1. Now apply the same proof as for (S.2), noting that only a log-factor is added. The same proof can be used for the second term.

Proof of (S.4): Note that

$$\begin{aligned} E(T^{-1}\sum_{t=N+1}^{T} w_{1t}^2) &= T^{-1}\sum_{t=N+1}^{T}\sum_{n_1,n_2=0}^{N-1}\pi_{n_1}(-u)\pi_{n_2}(-u)E(\varepsilon_{t-n_1}\varepsilon_{t-n_2}) \\ &= T^{-1}\sum_{t=N+1}^{T}\sum_{n=0}^{N-1}\pi_n(-u)^2\sigma_{t-n}^2 \end{aligned}$$

such that the second moment of $M_{11NT}(u)$ is

$$EM_{11NT}(u)^2 = E(T^{-1}\sum_{t=N+1}^{T} w_{1t}^2)^2 - T^{-2}\sum_{t_1,t_2=N+1}^{T}\sum_{n,m=0}^{N-1}\pi_n(-u)^2\pi_m(-u)^2\sigma_{t_1-n}^2\sigma_{t_2-m}^2. \tag{S.7}$$

Now,

$$E(T^{-1}\sum_{t=N+1}^{T} w_{1t}^2)^2 = T^{-2}E\prod_{k=1}^{2}\sum_{t_k=N+1}^{T}\sum_{n_k=0}^{N-1}\sum_{m_k=0}^{N-1}\pi_{n_k}(-u)\pi_{m_k}(-u)\varepsilon_{t_k-n_k}\varepsilon_{t_k-m_k},$$

where again the two highest subscripts in $\prod_{k=1}^{2}\varepsilon_{t_k-n_k}\varepsilon_{t_k-m_k}$ have to be equal by Lemma A.2. By symmetry, there are three cases, which we now enumerate.

Case 1) Suppose first that $t_1-n_1=t_1-m_1$, i.e. $n_1=m_1$. If also $t_2-n_2=t_2-m_2$ the contribution is $T^{-2}\prod_{k=1}^{2}\sum_{t_k=N+1}^{T}\sum_{n_k=0}^{N-1}\pi_{n_k}(-u)^2\sigma_{t_k-n_k}^2$, which cancels with the second term of (S.7). If $t_2-n_2\neq t_2-m_2$, then both these terms have to be no greater than t_1-n_1 by Lemma A.2, so that $t_2\leq t_1-n_1+n_2$ and $m_2\geq t_2-t_1+n_1$. In this case the contribution is

$$\begin{aligned} &T^{-2}\sum_{t_1=N+1}^{T}\sum_{n_1,n_2=0}^{N-1}\sum_{t_2=N+1}^{\max(T,t_1-n_1+n_2)}\sum_{m_2=\max(0,t_2-t_1+n_1)}^{N-1}\pi_{n_1}(-u)^2\pi_{n_2}(-u)\pi_{m_2}(-u) \\ &\quad\times\sigma_{t_1-n_1}^2\sigma_{t_2-n_2}\sigma_{t_2-m_2}\kappa_4(t_1-n_1,t_1-n_1,t_2-n_2,t_2-m_2) \\ &\leq cT^{-2}\sum_{t_1=N+1}^{T}\sum_{n_1,n_2=0}^{N-1}\pi_{n_1}(-u)^2\pi_{n_2}(-u) \\ &\leq cT^{-1}\sum_{n_1,n_2=0}^{N-1}n_1^{-2u-2}n_2^{-u-1}\leq cT^{-1}N^{1/2+3\kappa}, \end{aligned}$$

where the first inequality is by Assumption 1(a)(iii),(b).

Case 2) If $t_1-n_1=t_2-n_2\geq t_k-m_k$ the restriction $|t_1-t_2|=|n_1-n_2|\leq N$ is implied such that the contribution is

$$\begin{aligned} &T^{-2}\sum_{\substack{t_1,t_2=N+1\\|t_1-t_2|\leq N}}^{T}\sum_{n_1=\max(0,t_1-t_2)}^{N-1}\sum_{m_1=n_1}^{N-1}\sum_{m_2=\max(0,t_2-t_1+n_1)}^{N-1}\pi_{n_1}(-u)\pi_{t_2-t_1+n_1}(-u)\pi_{m_1}(-u)\pi_{m_2}(-u) \\ &\quad\times\sigma_{t_1-n_1}^2\sigma_{t_1-m_1}\sigma_{t_2-m_2}E(z_{t_1-n_1}^2z_{t_1-m_1}z_{t_2-m_2}). \end{aligned}$$

If also $t_1 - m_1 = t_2 - m_2$, the expectation is $\tau_{m_1-n_1,m_1-n_1}$ and contribution is bounded by

$$cT^{-2} \sum_{\substack{t_1,t_2=N+1 \\ |t_1-t_2|\le N}}^{T} \left(\sum_{n=0}^{N-1} \pi_n(-u)\pi_{t_2-t_1+n}(-u) \right)^2$$

$$\le cT^{-2} \sum_{\substack{t_1,t_2=N+1 \\ |t_1-t_2|\le N}}^{T} \xi_N(u,u,0)^2 \le cT^{-1}N\xi_N(u,u,0)^2 \le c(\log T)^2 T^{-1} N^{1+4\kappa}$$

by Assumption 1(b) and Lemma S.1(i). If instead $t_1 - m_1 \neq t_2 - m_2$, the expectation is $\kappa_4(t_1-n_1, t_1-n_1, t_1-m_1, t_2-m_2)$ and the bound is

$$cT^{-2} \sum_{\substack{t_1,t_2=N+1 \\ |t_1-t_2|\le N}}^{T} \sum_{n_1=\max(0,t_1-t_2)}^{N-1} \pi_{n_1}(-u)^2 \pi_{t_2-t_1+n_1}(-u)^2$$

$$\times \sum_{m_1=n_1}^{N-1} \sum_{m_2=\max(0,t_2-t_1+n_1)}^{N-1} |\kappa_4(t_1-n_1, t_1-n_1, t_1-m_1, t_2-m_2)|$$

$$\le cT^{-2} \sum_{\substack{t_1,t_2=N+1 \\ |t_1-t_2|\le N}}^{T} \sum_{n_1=\max(0,t_1-t_2)}^{N-1} \pi_{n_1}(-u)^2 \pi_{t_2-t_1+n_1}(-u)^2 \le cT^{-1}N.$$

Case 3) If $t_1 - n_1 = t_2 - m_2$ and $t_1 - m_1 = t_2 - n_2$ the contribution is

$$T^{-2} \sum_{\substack{t_1,t_2=N+1 \\ |t_1-t_2|\le N}}^{T} \sum_{n_1=0}^{N-1} \sum_{m_1=0}^{N-1} \pi_{n_1}(-u)\pi_{t_2-t_1+m_1}(-u)\pi_{m_1}(-u)\pi_{t_2-t_1+n_1}(-u)\sigma_{t_1-n_1}^2 \sigma_{t_1-m_1}^2 \tau_{m_1-n_1,m_1-n_1}$$

$$\le c(\log T)^2 T^{-1} N^{1+4\kappa}$$

and if $t_1 - n_1 = t_2 - m_2$ and $t_1 - m_1 \neq t_2 - n_2$ (both no greater than $t_1 - n_1$ by Lemma A.2) the contribution is

$$T^{-2} \sum_{\substack{t_1,t_2=N+1 \\ |t_1-t_2|\le N}}^{T} \sum_{n_1=0}^{N-1} \sum_{m_1=n_1}^{N-1} \sum_{n_2=\max(0,t_2-t_1+n_1)}^{N-1} \pi_{n_1}(-u)\pi_{m_1}(-u)\pi_{n_2}(-u)\pi_{t_2-t_1+n_1}(-u)$$

$$\times \sigma_{t_1-n_1}^2 \sigma_{t_1-m_1} \sigma_{t_2-n_2} \kappa_4(t_1-n_1, t_1-n_1, t_1-m_1, t_2-n_2)$$

$$\le cT^{-2} \sum_{\substack{t_1,t_2=N+1 \\ |t_1-t_2|\le N}}^{T} \sum_{n_1=0}^{N-1} \pi_{n_1}(-u)^2 \pi_{t_2-t_1+n_1}(-u)^2 \le cT^{-1}N$$

in the same way as in case 2).

Proof of (S.5): Apply the same decomposition as in the proof of (S.3) and then use the same proof as for (S.4) with an extra log-factor.

S.4.4 Proof of Lemma B.3

The proof is given only for $k, l = 0$ since the derivatives just add a log-factor, see (A.1), which does not change the proof. Rearranging the summations the product moment

$M_T(u_1,u_2,\psi)$ is

$$T^{-1}\sum_{j,k=0}^{T-1}\pi_j(-u_1)\pi_k(-u_2)\sum_{n,m=0}^{\infty}\zeta_{1n}(\psi)\zeta_{2m}(\psi)\sum_{t=\max(j,k)+1}^{T}\varepsilon_{t-j-n}\varepsilon_{t-k-m}$$

$$=T^{-1}\sum_{j=0}^{T-1}\pi_j(-u_1)\sum_{n=0}^{\infty}\sum_{m=\max(0,j+n-T+1)}^{j+n}\zeta_{1n}(\psi)\zeta_{2m}(\psi)\pi_{j+n-m}(-u_2)\sum_{t=\max(j,j+n-m)+1}^{T}\varepsilon_{t-j-n}^2 \tag{S.8}$$

$$+2T^{-1}\sum_{j=0}^{T-1}\pi_j(-u_1)\sum_{n,m=0}^{\infty}\zeta_{1n}(\psi)\zeta_{2m}(\psi)\sum_{k=0}^{\min(T,j+n-m)-1}\pi_k(-u_2)\sum_{t=\max(j,k)+1}^{T}\varepsilon_{t-j-n}\varepsilon_{t-k-m}. \tag{S.9}$$

Since $T^{-1}\sum_{t=\max(j,j+n-m)+1}^{T}\varepsilon_{t-j-n}^2=O_p(1)$ uniformly in j,n,m it holds that $\sup_{(u_1,u_2,\psi)\in\tilde{\Theta}}|(S.8)|$ is

$$O_p\left(\sup_{(u_1,u_2,\psi)\in\tilde{\Theta}}\sum_{n=0}^{\infty}\sum_{m=\max(0,n-T+1)}^{T-1+n}|\zeta_{1n}(\psi)||\zeta_{2m}(\psi)|\sum_{j=\max(0,m-n)}^{\min(T-1,T-1+m-n)}|\pi_j(-u_1)||\pi_{j+n-m}(-u_2)|\right)$$

$$=O_p\left(\sup_{(u_1,u_2,\psi)\in\tilde{\Theta}}\sum_{n=0}^{\infty}\sum_{m=\max(0,n-T+1)}^{T-1+n}|\zeta_{1n}(\psi)||\zeta_{2m}(\psi)|\sum_{j=1+\max(0,m-n)}^{\min(T-1,T-1+m-n)}j^{-u_1-1}(j+n-m)^{-u_2-1}\right).$$

If $a>0$ the summation over j is bounded and then $\sup_{(u_1,u_2,\psi)\in\tilde{\Theta}}|(S.8)|=O_p(1)$ because $\sum_{n=0}^{\infty}|\zeta_{in}(\psi)|<\infty$ uniformly in $\psi\in\tilde{\Psi}, i=1,2$, showing (B.3) for (S.8). If $a\le 0$ the summation over j is $O_p((\log T)T^{-a})$ which is then also the bound for the supremum of (S.8), showing (B.4) for (S.8). In the case of (B.5), the summation over j is now

$$T^{-1}\sum_{j=1+\max(0,m-n)}^{\min(T-1,T-1+m-n)}(\frac{j}{T})^{-u_1-1}(\frac{j+n-m}{T})^{-u_2-1}$$
$$\le cT^{-1}\sum_{j=1+\max(0,m-n)}^{\min(T-1,T-1+m-n)}(\frac{j}{T})^{-1/2+\kappa_1}(\frac{j+n-m}{T})^{-1/2+\kappa_1}<\infty$$

because $\kappa_1>0$, as seen easily by integral approximation, which shows (B.5) for (S.8).

Next, we analyze (S.9). Summation by parts yields

$$\sum_{k=0}^{\min(T,j+n-m)-1}\pi_k(-u_2)\sum_{t=\max(j,k)+1}^{T}\varepsilon_{t-j-n}\varepsilon_{t-k-m}$$
$$=\pi_{\min(T,j+n-m)-1}(-u_2)\sum_{k=0}^{\min(T,j+n-m)-1}\sum_{t=\max(j,k)+1}^{T}\varepsilon_{t-j-n}\varepsilon_{t-k-m}$$
$$-\sum_{l=0}^{\min(T,j+n-m)-2}(\pi_{l+1}(-u_2)-\pi_l(-u_2))\sum_{k=0}^{l}\sum_{t=\max(j,k)+1}^{T}\varepsilon_{t-j-n}\varepsilon_{t-k-m}, \tag{S.10}$$

where

$$E \sup_{(u_1,u_2,\psi)\in\tilde{\Theta}} |(S.10)| \leq \sup_{(u_1,u_2,\psi)\in\tilde{\Theta}} |\pi_{\min(T,j+n-m)-1}(-u_2)| E \left| \sum_{k=0}^{\min(T,j+n-m)-1} \sum_{t=\max(j,k)+1}^{T} \varepsilon_{t-j-n}\varepsilon_{t-k-m} \right|$$

$$+ \sum_{l=0}^{\min(T,j+n-m)-2} \sup_{(u_1,u_2,\psi)\in\tilde{\Theta}} |\pi_{l+1}(-u_2) - \pi_l(-u_2)| E \left| \sum_{k=0}^{l} \sum_{t=\max(j,k)+1}^{T} \varepsilon_{t-j-n}\varepsilon_{t-k-m} \right|.$$

Note that $\sum_{k=0}^{l}\sum_{t=\max(j,k)+1}^{T} \varepsilon_{t-j-n}\varepsilon_{t-k-m} = \sum_{s=\max(1-m,1+j-l-m)}^{T-m} v_s$ with $v_s := \varepsilon_s \sum_{k=j+n-m-l}^{j+n-m} \varepsilon_{s-k}$ being an uncorrelated sequence that satisfies $E(v_s^2) \leq Kl$, such that

$$\left(E \left| \sum_{k=0}^{l} \sum_{t=\max(j,k)+1}^{T} \varepsilon_{t-j-n}\varepsilon_{t-k-m} \right| \right)^2 \leq E \left(\sum_{s=\max(1-m,1+j-l-m)}^{T-m} v_s \right)^2 \leq K(T+l-j)l.$$

It follows that $E \sup_{(u_1,u_2,\psi)\in\tilde{\Theta}} |(S.9)|$ is bounded by a constant times

$$\sup_{(u_1,u_2,\psi)\in\tilde{\Theta}} \sum_{j=0}^{T-1} |\pi_j(-u_1)| \sum_{n,m=0}^{\infty} |\zeta_{1n}(\psi)\zeta_{2m}(\psi)| |\pi_{\min(T,j+n-m)-1}(-u_2)| \tag{S.11}$$

$$+ \sup_{(u_1,u_2,\psi)\in\tilde{\Theta}} T^{-1} \sum_{j=0}^{T-1} |\pi_j(-u_1)| \sum_{n,m=0}^{\infty} |\zeta_{1n}(\psi)\zeta_{2m}(\psi)| \sum_{l=0}^{T-2} |\pi_{l+1}(-u_2) - \pi_l(-u_2)| (T+l-j)^{1/2} l^{1/2}. \tag{S.12}$$

The result for (S.11) follows as in the analysis of (S.8). To prove (B.3) and (B.4) for the term (S.12) we use (A.5) and that $\sum_{n=0}^{\infty} |\zeta_{in}(\psi)| < \infty$ uniformly in $\psi \in \tilde{\Psi}, i = 1, 2$, to obtain the bound

$$\sup_{(u_1,u_2,\psi)\in\tilde{\Theta}} T^{-1} \sum_{j=1}^{T-1} j^{-u_1-1} \sum_{l=1}^{T-2} l^{-u_2-3/2} (T+l-j)^{1/2}$$

$$\leq c \sup_{(u_1,u_2,\psi)\in\tilde{\Theta}} T^{-1} \sum_{l=1}^{T-2} l^{-u_2-3/2} \sum_{j=1}^{T+l-1} j^{-u_1-1} (T+l-j)^{1/2}$$

$$\leq c \sup_{(u_1,u_2,\psi)\in\tilde{\Theta}} c(\log T) T^{-1} \sum_{l=1}^{T-2} l^{-u_2-3/2} (T+l)^{\max(1/2,1/2-u_1)}$$

$$\leq c \sup_{(u_1,u_2,\psi)\in\tilde{\Theta}} c(\log T) T^{-1/2} \sum_{l=1}^{T-2} l^{-u_2-3/2+\max(0,-u_1)},$$

where the second inequality follows from Lemma A.4 and the third because $(T+l)^{\max(1/2,1/2-u_1)} = (T+l)^{1/2}(T+l)^{\max(0,-u_1)} \leq (2T)^{1/2} l^{\max(0,-u_1)}$. Since $-u_2-3/2+\max(0,-u_1) = -\min(u_2+1, u_1+u_2+1)-1/2 \leq -a-1/2$, the right-hand side is bounded by $c(\log T)^2 T^{-1/2} T^{\max(0,1/2-a)} = c(\log T)^2 T^{\max(-1/2,-a)}$ if $a > 0$ and $c(\log T) T^{-1/2} T^{1/2-a} = c(\log T) T^{-a}$ if $a \leq 0$. To prove (B.5) for the term (S.12) we note that $(T+l-j)^{1/2} \leq (2T)^{1/2}$ and find the simple bound

$$\sup_{(u_1,u_2,\psi)\in\bar{\Theta}} T^{-1} \sum_{j=1}^{T-1} (\frac{j}{T})^{-u_1-1} T^{-1} \sum_{l=1}^{T-2} (\frac{l}{T})^{-u_2-3/2} \leq c T^{-1} \sum_{j=1}^{T-1} (\frac{j}{T})^{-1/2+\kappa_1} T^{-1} \sum_{l=1}^{T-2} (\frac{l}{T})^{-1+\kappa_1} < \infty$$

because $\kappa_1 > 0$, as seen easily by integral approximation.

S.5 Proofs for QMLE and Asymptotic Tests

S.5.1 Proof of Theorem 1

The residual in (6) is given by $\varepsilon_t(\theta) = \sum_{n=0}^{t-1} b_n(\psi)\Delta_+^{d-d_{0,T}} u_{t-n}$, and clearly the convergence properties of $Q_T(\theta)$ in (8) depend on $\lim_{T\to\infty} d - d_{0,T} = d - d_0$. Define the untruncated processes

$$e_t(\psi) := c(L,\psi)\varepsilon_t = \sum_{n=0}^{\infty} c_n(\psi)\varepsilon_{t-n}, \tag{S.13}$$

$$\eta_t(\theta) := \Delta^{d-d_0} e_t(\psi) = \sum_{n=0}^{\infty} \varphi_n(\theta)\varepsilon_{t-n}, \tag{S.14}$$

where $\eta_t(\theta)$ is well-defined for $d - d_0 > -1/2$ and where we used

$$c(z,\psi) := b(z,\psi)a(z,\psi_{0,T}) = \frac{a(z,\psi_{0,T})}{a(z,\psi)} = \sum_{n=0}^{\infty} c_n(\psi)z^n, \tag{S.15}$$

$$\varphi_n(\theta) := \sum_{m=0}^{n} \pi_m(d_0 - d)c_{n-m}(\psi). \tag{S.16}$$

Again, we have suppressed the T subscript on the triangular arrays $e_t(\psi)$ and $\eta_t(\theta)$ and on the coefficients $c_n(\psi)$ and $\varphi_n(\theta)$.

From Assumption 3 and Lemma A.4, there exists a $T_0 \geq 1$ such that the coefficients $c_n(\psi)$ satisfy

$$|c_n(\psi)| = O(n^{-2-\zeta}) \text{ uniformly in } \psi \in \Psi \text{ and } T \geq T_0. \tag{S.17}$$

From Lemmas A.3 and A.4 the coefficients $\varphi_n(\theta)$ then satisfy

$$|\varphi_n(\theta)| = O(n^{\max(d_0-d-1,-2-\zeta)}) \text{ uniformly in } \psi \in \Psi \text{ and } T \geq T_0, \tag{S.18}$$

such that, in particular, when $d - d_0 > -1/2$, $\eta_t(\theta)$ is a linear process with square summable coefficients. Note that the uniformity in T in (S.17) and (S.18) obtains from the uniform bound on $a_n(\psi)$ in (5), when T is sufficiently large that $\psi_{0,T} \in \Psi$.

Let the deterministic function $Q(\theta)$ denote the pointwise probability limit of $Q_T(\theta)$, shown subsequently to be given by

$$Q(\theta) := \begin{cases} \int_0^1 \sigma(s)^2 \mathsf{d}s \sum_{n=0}^{\infty} \varphi_{0,n}(\theta)^2 & \text{if } d - d_0 > -1/2, \\ \infty & \text{if } d - d_0 \leq -1/2, \end{cases} \tag{S.19}$$

where $\varphi_{0,n}(\theta) := \sum_{m=0}^{n} \pi_m(d_0 - d)\sum_{k=0}^{n-m} b_k(\psi)a_{n-m-k}(\psi_0)$ is the same coefficient as in (S.16), but evaluated at ψ_0 instead of $\psi_{0,T}$. According to (S.19) the parameter space Θ is partitioned into three disjoint compact subsets, $\Theta_1 := \Theta_1(\kappa_1) = D_1 \times \Psi$, $\Theta_2 := \Theta_2(\kappa_1,\kappa_2) = D_2 \times \Psi$, and $\Theta_3 := \Theta_3(\kappa_2) = D_3 \times \Psi$, where $D_1 := D_1(\kappa_1) = D \cap \{d : d - d_0 \leq -1/2 - \kappa_1\}$, $D_2 := D_2(\kappa_1,\kappa_2) = D \cap \{d : -1/2 - \kappa_1 \leq d - d_0 \leq -1/2 + \kappa_2\}$, and $D_3 := D_3(\kappa_2) = D \cap \{d : d - d_0 \geq -1/2 + \kappa_2\}$, for some constants $0 < \kappa_2 < \kappa_1 < 1/2$ to be determined later. Here, special care is taken with respect to Θ_2, where the convergence of the objective function is non-uniform, as evident in (S.19). Clearly, $\theta_0 \in \Theta_3$ and if $d_1 > d_0 - 1/2$ then the choice $\kappa_2 = d_1 - d_0 + 1/2 > 0$ implies that Θ_1 and Θ_2 are empty in which case the proof is easily simplified accordingly.

The proof proceeds as follows. First, it is shown that for any $K > 0$ there exists a (fixed) $\bar{\kappa}_2 > 0$ such that

$$P(\inf_{\theta \in \Theta_1(\kappa_1) \cup \Theta_2(\kappa_1, \bar{\kappa}_2)} Q_T(\theta) > K) \to 1 \text{ as } T \to \infty. \quad \text{(S.20)}$$

This implies that $P(\hat{\theta} \in \Theta_3(\bar{\kappa}_2)) \to 1$ as $T \to \infty$, so that the relevant parameter space is reduced to $\Theta_3(\bar{\kappa}_2)$. From Theorem 5.7 of van der Vaart (1998) the desired result then follows if, for any fixed $\kappa_2 \in (0, 1/2)$,

$$\sup_{\theta \in \Theta_3(\kappa_2)} |Q_T(\theta) - Q(\theta)| \xrightarrow{p} 0 \text{ as } T \to \infty, \quad \text{(S.21)}$$

$$\inf_{\theta \in \Theta_3(\kappa_2) \cap \{\theta : |\theta - \theta_0| \geq \epsilon\}} Q(\theta) > Q(\theta_0) \text{ for all } \epsilon > 0. \quad \text{(S.22)}$$

Condition (S.21) entails uniform convergence of the objective function on Θ_3, and condition (S.22) ensures that the optimum of the limit function is uniquely attained at the true value. For the proofs of (S.20) and (S.21) we make repeated use of the following lemma, which is the non-bootstrap version of Lemma D.4 and shows that the problem can be simplified by considering the sum of squares of $\Delta_+^{d-d_0} e_t(\psi)$ rather than that of $\varepsilon_t(\theta)$ in the analysis of $Q_T(\theta)$. This serves two purposes: First, the truncation in the residual in the definition of $Q_T(\theta)$ can be dispensed with in the asymptotic analysis. Secondly, the fractional order of $e_t(\psi)$ is $d_0 - d$, which is fixed and corresponds to the definitions of the parameter sets D_i, while the fractional order of $\varepsilon_t(\theta)$ is $d_{0,T} - d$, which depends on T.

Lemma S.2. *Under the assumptions of Theorem 1 and* $0 < \kappa_1 < \min(1/2, \zeta/2 + 1/4)$ *it holds that*

$$\sup_{\theta \in \Theta_1} |T^{2(d-d_0)} \sum_{t=1}^{T} \varepsilon_t(\theta)^2 - T^{2(d-d_0)} \sum_{t=1}^{T} (\Delta_+^{d-d_0} e_t(\psi))^2| \xrightarrow{p} 0, \quad \text{(S.23)}$$

$$\sup_{\theta \in \Theta_2 \cup \Theta_3} |T^{-1} \sum_{t=1}^{T} \varepsilon_t(\theta)^2 - T^{-1} \sum_{t=1}^{T} (\Delta_+^{d-d_0} e_t(\psi))^2| \xrightarrow{p} 0. \quad \text{(S.24)}$$

S.5.1.1 Proof of Lemma S.2 First decompose

$$\sum_{t=1}^{T} \varepsilon_t(\theta)^2 - \sum_{t=1}^{T} (\Delta_+^{d-d_0} e_t(\psi))^2 = \sum_{t=1}^{T} \varepsilon_t(\theta)^2 - \sum_{t=1}^{T} (\Delta_+^{d-d_0} \sum_{n=0}^{t} b_n(\psi) u_{t-n})^2 \quad \text{(S.25)}$$

$$+ \sum_{t=1}^{T} (\Delta_+^{d-d_0} \sum_{n=0}^{t} b_n(\psi) u_{t-n})^2 - \sum_{t=1}^{T} (\Delta_+^{d-d_0} e_t(\psi))^2. \quad \text{(S.26)}$$

By the mean value theorem we find that

$$(S.25) = \frac{2\delta_\theta}{\sqrt{T}} \sum_{t=1}^{T} \left(\Delta_+^{\bar{v}} \sum_{n=0}^{t-1} b_n(\psi) u_{t-n} \right) \left(\frac{\partial}{\partial v} \Delta_+^{\bar{v}} \sum_{n=0}^{t-1} b_n(\psi) u_{t-n} \right)$$

for some intermediate value $\bar{v}$ between $d - d_{0,T}$ and $d - d_0$. We can apply Lemma B.3 directly to the right-hand side in both the non-stationary case (S.23) with normalization

$T^{2(d-d_0)}$ and in the nearly-stationary and stationary cases (S.24) with normalization T^{-1}. In either case, (S.25) is immediately shown to be uniformly negligible as required.

Next we write (S.26) as

$$(S.26) = \sum_{t=1}^{T} \Delta_+^{d-d_0} \sum_{n=0}^{t} b_n(\psi) u_{t-n} (\Delta_+^{d-d_0} \sum_{m=0}^{t} b_m(\psi) u_{t-m} - \Delta_+^{d-d_0} e_t(\psi)) \quad \text{(S.27)}$$

$$+ \sum_{t=1}^{T} \Delta_+^{d-d_0} e_t(\psi) (\Delta_+^{d-d_0} \sum_{n=0}^{t} b_n(\psi) u_{t-n} - \Delta_+^{d-d_0} e_t(\psi)) \quad \text{(S.28)}$$

and note that

$$\Delta_+^{d-d_0} \sum_{n=0}^{t} b_n(\psi) u_{t-n} - \Delta_+^{d-d_0} e_t(\psi) = - \sum_{j=0}^{t-1} \sum_{n=t-j}^{\infty} \pi_j(d_0 - d) b_n(\psi) u_{t-n-j} = \sum_{m=t}^{\infty} \phi_{tm} u_{t-m},$$

where $\phi_{tm} := - \sum_{j=0}^{t-1} \pi_j(d_0 - d) b_{m-j}(\psi)$ satisfies, see (5) and Lemmas A.3 and A.4,

$$\begin{aligned} \sup_{\psi \in \Psi} \sum_{m=t}^{\infty} |\phi_{tm}| &\leq c \sum_{m=t}^{\infty} \sum_{j=0}^{t-1} j^{d_0-d-1} (m-j)^{-2-\zeta} \\ &\leq c \sum_{j=0}^{t-1} j^{d_0-d-1} (t-j)^{-1-\zeta} \leq c(1 + \log t) t^{\max(d_0-d, -\zeta)-1}. \end{aligned} \quad \text{(S.29)}$$

Rewrite the term (S.28) as

$$(S.28) = \sum_{t=1}^{T} \sum_{j=0}^{t-1} \pi_j(d_0 - d) \sum_{n=0}^{\infty} b_n(\psi) \sum_{m=t}^{\infty} \phi_{tm} (u_{t-j-n} u_{t-m} - E(u_{t-j-n} u_{t-m})) \quad \text{(S.30)}$$

$$+ \sum_{t=1}^{T} \sum_{j=0}^{t-1} \pi_j(d_0 - d) \sum_{n=0}^{\infty} b_n(\psi) \sum_{m=t}^{\infty} \phi_{tm} E(u_{t-j-n} u_{t-m}). \quad \text{(S.31)}$$

The proof for (S.27) is identical to that for (S.28), except the summation over n in (S.27) is from t to ∞. For (S.31) we note that $\sup_t |E(u_t u_{t-n})| = \sup_t | \sum_{m=n}^{\infty} a_m(\psi_{0,T}) a_{m-n}(\psi_{0,T}) \sigma_{t-m}^2 | \leq c \sum_{m=n}^{\infty} m^{-2-\zeta} (m-n)^{-2-\zeta} \leq c n^{-2-\zeta}$ where $\zeta > 0$ is given in Assumption 3(ii), such that

$$\sum_{m=t}^{\infty} |E(u_{t-j-n} u_{t-m})| \leq c |t-j-n|^{-1-\zeta}.$$

Using also (S.29) it holds that

$$\begin{aligned} \sup_{\psi \in \Psi} \sum_{m=t}^{\infty} \phi_{tm} E(u_{t-j-n} u_{t-m}) &\leq (\sup_{\psi \in \Psi} \sum_{m=t}^{\infty} |\phi_{tm}|)(\sum_{m=t}^{\infty} |E(u_{t-j-n} u_{t-m})|) \\ &\leq c(1 + \log t) t^{\max(d_0-d, -\zeta)-1} |t-j-n|^{-1-\zeta}. \end{aligned}$$

It also holds that

$$\sup_{\psi \in \Psi} \sum_{n=0}^{\infty} |b_n(\psi)| |t-j-n|^{-1-\zeta} \leq c \sum_{n=0}^{\infty} n^{-2-\zeta} |t-j-n|^{-1-\zeta} \leq c(t-j)^{-1-\zeta}$$

by (5) and Lemma S.1(ii). Consequently,

$$\sup_{\psi\in\Psi}|(S.31)| \le c\sum_{t=1}^{T}(1+\log t)t^{\max(d_0-d,-\zeta)-1}\sum_{j=0}^{t-1}j^{d_0-d-1}(t-j)^{-1-\zeta}$$
$$\le c\sum_{t=1}^{T}(1+\log t)^2 t^{2\max(d_0-d,-\zeta)-2}$$

by Lemmas A.3 and A.4. Thus, $\sup_{\theta\in\Theta_1}T^{2(d-d_0)}|(S.31)| \le c(\log T)^3T^{-1} \to 0$ as $T\to\infty$ and $\sup_{\theta\in\Theta_2\cup\Theta_3}T^{-1}|(S.31)| \le c(\log T)^3T^{-1+2\kappa_1} \to 0$ as $T\to\infty$.

Changing the order of the summations, (S.30) is

$$-\sum_{j=0}^{T-1}\pi_j(d_0-d)\sum_{n=0}^{\infty}b_n(\psi)\sum_{m=j+1}^{\infty}\sum_{k=0}^{\min(m,T)-1}\pi_k(d_0-d)b_{m-k}(\psi)\sum_{t=\max(j,k)+1}^{\min(m,T)}v_t, \qquad \text{(S.32)}$$

where the summand $v_t := u_{t-j-n}u_{t-m} - E(u_{t-j-n}u_{t-m})$ is mean zero with autocovariances

$$Ev_tv_s = \sum_{k_1,k_2=0}^{\infty}\sum_{l_1,l_2=0}^{\infty}a_{k_1}(\psi_{0,T})a_{k_2}(\psi_{0,T})a_{l_1}(\psi_{0,T})a_{l_2}(\psi_{0,T})\sigma_{t-j-n-k_1}\sigma_{t-m-k_2}\sigma_{s-j-n-l_1}\sigma_{s-m-l_2}$$
$$\times[E(z_{t-j-n-k_1}z_{t-m-k_2}z_{s-j-n-l_1}z_{s-m-l_2}) - E(z_{t-j-n-k_1}z_{t-m-k_2})E(z_{s-j-n-l_1}z_{s-m-l_2})].$$

The expectations are non-zero only if the two highest subscripts are equal (Lemma A.2). Routine calculations using (5), Assumption 1, and Lemma S.1(ii) show that $|Ev_tv_s| \le c|s-t|^{-2-\zeta}$. Since the summation $\sum_{t=\max(j,k)+1}^{\min(m,T)}$ has at most m terms it follows that

$$E(\sum_{t=\max(j,k)+1}^{\min(m,T)}v_t)^2 = \sum_{t,s=\max(j,k)+1}^{\min(m,T)}E(v_tv_s) \le c\sum_{t,s=\max(j,k)+1}^{\min(m,T)}|t-s|^{-2-\zeta} \le cm$$

such that $E|\sum_{t=\max(j,k)+1}^{\min(m,T)}v_t| \le cm^{1/2}$. Using Lemma A.3 and $\sup_{\psi\in\Psi}\sum_{n=0}^{\infty}|b_n(\psi)| < \infty$, it now follows from (S.32) that (S.30) satisfies

$$E\sup_{\psi\in\Psi}|(S.30)| \le c\sum_{j=1}^{T-1}j^{d_0-d-1}\sup_{\psi\in\Psi}\sum_{m=j+1}^{T}\sum_{k=1}^{m-1}k^{d_0-d-1}|b_m(\psi)|m^{1/2} \qquad \text{(S.33)}$$
$$+c\sum_{j=1}^{T-1}j^{d_0-d-1}\sup_{\psi\in\Psi}\sum_{m=T+1}^{\infty}\sum_{k=1}^{T-1}k^{d_0-d-1}|b_m(\psi)|m^{1/2}. \qquad \text{(S.34)}$$

For (S.33) change the order of the summations,

$$\sup_{\psi\in\Psi}\sum_{m=j+1}^{T}\sum_{k=1}^{m-1}k^{d_0-d-1}|b_m(\psi)|m^{1/2} \le \sum_{k=1}^{T-1}k^{d_0-d-1}\sum_{m=\max(j,k)+1}^{T}m^{-3/2-\zeta}$$
$$\le c(\log T)T^{\max(d_0-d,0)}(j+1)^{-1/2-\zeta}.$$

Then the bounds for (S.33) are

$$\sup_{d\in D_1}T^{2(d-d_0)}(\log T)T^{\max(d_0-d,0)}\sum_{j=1}^{T-1}j^{d_0-d-3/2-\zeta} \le c(\log T)^2T^{-1/2+\max(-\kappa_1,-\zeta)},$$
$$\sup_{d\in D_2\cup D_3}T^{-1}(\log T)T^{\max(d_0-d,0)}\sum_{j=1}^{T-1}j^{d_0-d-3/2-\zeta} \le c(\log T)^2T^{-1/2+\kappa_1+\max(0,\kappa_1-\zeta)},$$

which shows the result for (S.33). Similarly, for (S.34),

$$\sup_{\psi\in\Psi}\sum_{m=T+1}^{\infty}\sum_{k=1}^{T-1}k^{d_0-d-1}|b_m(\psi)|m^{1/2}\le\sum_{k=1}^{T-1}k^{d_0-d-1}\sum_{m=T+1}^{\infty}m^{-3/2-\zeta}$$
$$\le c(\log T)T^{\max(0,d_0-d)-1/2-\zeta},$$

which gives the bounds

$$\sup_{d\in D_1}T^{2(d-d_0)}(\log T)T^{\max(0,d_0-d)-1/2-\zeta}\sum_{j=1}^{T-1}j^{d_0-d-1}\le c(\log T)^2T^{-1/2-\zeta},$$
$$\sup_{d\in D_2\cup D_3}T^{-1}(\log T)T^{\max(0,d_0-d)-1/2-\zeta}\sum_{j=1}^{T-1}j^{d_0-d-1}\le c(\log T)^2T^{-1/2-\zeta+2\kappa_1},$$

showing the result for (S.34) and hence concluding the proof.

S.5.1.2 Convergence on $\Theta_1(\kappa_1)$ First, if $\theta\in\Theta_1(\kappa_1)$ then $\varepsilon_t(\theta)$ should be normalized by $T^{d-d_0+1/2}$, and by Lemma S.2 the difference between $T^{2(d-d_0)+1}Q_T(\theta)$ and $T^{2(d-d_0)}\sum_{t=1}^{T}(\Delta_+^{d-d_0}e_t(\psi))^2$ is negligible in probability uniformly in $\theta\in\Theta_1$, so it suffices to consider the latter product moment. We apply the Beveridge-Nelson decomposition (C.4) and decompose the relevant product moment as

$$T^{2(d-d_0)}\sum_{t=1}^{T}(\Delta_+^{d-d_0}e_t(\psi))^2\ge\left(\sum_{n=0}^{\infty}c_n(\psi)\right)^2T^{2(d-d_0)}\sum_{t=1}^{T}(\Delta_+^{d-d_0}\varepsilon_t)^2 \tag{S.35}$$
$$+2\left(\sum_{n=0}^{\infty}c_n(\psi)\right)T^{2(d-d_0)}\sum_{t=1}^{T}\Delta_+^{d-d_0}\varepsilon_t\sum_{n=0}^{\infty}\tilde{c}_n(\psi)\Delta_+^{d-d_0+1}\varepsilon_{t-n}. \tag{S.36}$$

By the Cauchy-Schwarz inequality, (S.36) is bounded by

$$2\left(\sum_{n=0}^{\infty}c_n(\psi)\right)\left(T^{2(d-d_0)}\sum_{t=1}^{T}(\Delta_+^{d-d_0}\varepsilon_t)^2\right)^{1/2}\left(T^{2(d-d_0)}\sum_{t=1}^{T}(\sum_{n=0}^{\infty}\tilde{c}_n(\psi)\Delta_+^{d-d_0+1}\varepsilon_{t-n})^2\right)^{1/2}. \tag{S.37}$$

The term in the first parenthesis satisfies $0<|\sum_{n=0}^{\infty}c_n(\psi)|<\infty$ uniformly in $\psi\in\Psi$ for T sufficiently large by Assumption 3. For the term in the second parenthesis we define $M_T(d):=T^{2(d-d_0)}\sum_{t=1}^{T}(\Delta_+^{d-d_0}\varepsilon_t)^2$, which is $O_p(1)$ by Lemma B.1. To strengthen this to hold uniformly in $d\in D_1$ it is sufficient to show that $M_T(d)$ is tight as a function of the parameter. We prove tightness using the moment condition in Billingsley (1968, Theorem 12.3), which requires showing that $M_T(d)$ is tight for fixed $d\in D_1$ and that

$$||M_T(u_1)-M_T(u_2)||_2\le c|u_1-u_2| \tag{S.38}$$

for some constant $c>0$ that does not depend on T, u_1, or u_2. The tightness condition in (S.38) is satisfied by Lemma B.1, and hence the second term in (S.37) is $O_p(1)$ uniformly in $d\in D_1$.

The term inside the third parenthesis in (S.37) can be rewritten as

$$T^{2(d-d_0)}\sum_{t=1}^{T}\sum_{n,m=0}^{\infty}\tilde{c}_n(\psi)\tilde{c}_m(\psi)\sum_{j,k=0}^{t-1}\pi_j(d_0-d-1)\pi_k(d_0-d-1)\varepsilon_{t-j-n}\varepsilon_{t-k-m}$$
$$=T^{2(d-d_0)+1}\sum_{n,m=0}^{\infty}\tilde{c}_n(\psi)\tilde{c}_m(\psi)\sum_{j,k=0}^{T-1}\pi_j(d_0-d-1)\pi_k(d_0-d-1)T^{-1}\sum_{t=\max(j,k)+1}^{T}\varepsilon_{t-j-n}\varepsilon_{t-k-m},$$

where $E(T^{-1}\sum_{t=\max(j,k)+1}^{T}\varepsilon_{t-j-n}\varepsilon_{t-k-m})\leq c$ uniformly in $0\leq j,k\leq T-1$ and $\sum_{n=0}^{\infty}|\tilde{c}_n(\psi)|<\infty$ uniformly in $\psi\in\Psi$. Thus, the term inside the third parenthesis in (S.37) is a non-negative random variable with expectation

$$E\left(\sup_{\theta\in\Theta_1}T^{2(d-d_0)}\sum_{t=1}^{T}(\sum_{n=0}^{\infty}\tilde{c}_n(\psi)\Delta_+^{d-d_0+1}\varepsilon_{t-n})^2\right)\leq c\sup_{d\in D_1}T^{2(d-d_0)+1}\left(\sum_{j=0}^{T-1}|\pi_j(d_0-d-1)|\right)^2$$
$$\leq c\sup_{d\in D_1}T^{2(d-d_0)+1}(\sum_{j=0}^{T-1}j^{d_0-d-2})^2\leq c(\log T)^2T^{-2\kappa_1}$$

by application of Lemma A.3, thus showing that (S.36) converges to zero in probability uniformly in $\theta\in\Theta_1$.

Next, the term (S.35) is analyzed. By the Cauchy-Schwarz inequality,

$$T^{2(d-d_0)}\sum_{t=1}^{T}(\Delta_+^{d-d_0}\varepsilon_t)^2\geq T^{2(d-d_0)-1}(\sum_{t=1}^{T}\Delta_+^{d-d_0}\varepsilon_t)^2=(T^{d-d_0-1/2}\Delta_+^{d-d_0-1}\varepsilon_T)^2,$$

and we can write $T^{d-d_0-1/2}\Delta_+^{d-d_0-1}\varepsilon_T=T^{d-d_0-1/2}\sum_{j=0}^{T-1}\pi_j(d_0-d+1)\varepsilon_{T-j}=T^{d-d_0-1/2}\sum_{t=1}^{T}\pi_{T-t}(d_0-d+1)\varepsilon_t$ and apply Lemma A.1 with $U_{Tt}=T^{d-d_0-1/2}\pi_{T-t}(d_0-d+1)\varepsilon_t$, which is a martingale difference array by Assumption 1. Firstly, the Lindeberg condition (i) of Lemma A.1 is satisfied by Lyapunov's sufficient condition because $\sum_{t=1}^{T}EU_{Tt}^4=T^{4(d-d_0)-2}\sum_{t=1}^{T}\pi_{T-t}(d_0-d+1)^4\sigma_t^4Ez_t^4\leq cT^{-2}\sum_{t=1}^{T}\left(\frac{T-t}{T}\right)^{4(d_0-d)}\leq cT^{-1}\to 0$ as $T\to\infty$. Secondly, we verify condition (ii)(a) of Lemma A.1 by showing L_2-convergence. Thus,

$$E\left(\sum_{t=1}^{T}U_{Tt}^2-E\sum_{t=1}^{T}U_{Tt}^2\right)^2=\sum_{t,s=1}^{T}E(U_{Tt}^2U_{Ts}^2)-\sum_{t,s=1}^{T}E(U_{Tt}^2)E(U_{Ts}^2)$$
$$=T^{4(d-d_0)-2}\sum_{t,s=1}^{T}\pi_{T-t}(d_0-d+1)^2\pi_{T-s}(d_0-d+1)^2\sigma_t^2\sigma_s^2[E(z_t^2z_s^2)-E(z_t^2)E(z_s^2)]$$
$$=T^{4(d-d_0)-2}\sum_{t=1}^{T}\pi_{T-t}(d_0-d+1)^4\sigma_t^4[E(z_t^4)-E(z_t^2)^2] \quad \text{(S.39)}$$
$$+2T^{4(d-d_0)-2}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\pi_{T-t}(d_0-d+1)^2\pi_{T-s}(d_0-d+1)^2\sigma_t^2\sigma_s^2[E(z_t^2z_s^2)-E(z_t^2)E(z_s^2)]. \quad \text{(S.40)}$$

By Assumption 1(a)(ii),(b) and Lemma A.3, the term (S.39) is bounded by

$$cT^{4(d-d_0)-2}\sum_{t=1}^{T}(T-t)^{4(d_0-d)}\leq cT^{-1}\to 0.$$

The term (S.40) is

$$2T^{4(d-d_0)-2}\sum_{t=2}^{T}\sum_{r=1}^{t-1}\pi_{T-t}(d_0-d+1)^2\pi_{T-t+r}(d_0-d+1)^2\sigma_t^2\sigma_{t-r}^2\kappa_4(t,t,t-r,t-r)$$
$$\leq cT^{-2}\sum_{t=2}^{T}\left(\frac{T-t}{T}\right)^{2(d_0-d)}\left(\frac{T-1}{T}\right)^{2(d_0-d)}\sum_{r=1}^{t-1}|\kappa_4(t,t,t-r,t-r)|\leq cT^{-1}\to 0$$

using Assumption 1(a)(iii),(b) and Lemma A.3. Finally,

$$\begin{aligned}E\sum_{t=1}^{T}U_{Tt}^2 &= T^{2(d-d_0)-1}\sum_{t=1}^{T}\pi_{T-t}(d_0-d+1)^2\sigma_t^2\\ &= \frac{1}{\Gamma(d_0-d+1)^2}T^{-1}\sum_{t=1}^{T}(\frac{T-t}{T})^{2(d_0-d)}\sigma_t^2(1+o(1))\\ &\to \frac{1}{\Gamma(d_0-d+1)^2}\int_0^1(1-s)^{2(d_0-d)}\sigma(s)^2\mathsf{d}s =: V(d),\end{aligned}$$

and we conclude from Lemma A.1 and the above analysis that

$$G_T(d) := T^{2(d-d_0)-1}(\sum_{t=1}^{T}\Delta_+^{d-d_0}\varepsilon_t)^2 = (T^{d-d_0-1/2}\Delta_+^{d-d_0-1}\varepsilon_T)^2 \overset{w}{\to} V(d)\chi_1^2, \tag{S.41}$$

for any fixed $d\in D_1$, which shows the pointwise limit.

To strengthen the pointwise convergence in (S.41) to weak convergence in $\mathcal{C}(D_1)$, denoted $\Rightarrow$, it is sufficient to show that $G_T(d)$ is tight (stochastically equicontinuous) as a function of the parameter, which follows by the tightness condition (S.38) and Lemma B.1. Hence the convergence in (S.41) is strengthened to $G_T(d)\Rightarrow V(d)\chi_1^2$ in $\mathcal{C}(D_1)$. By the continuous mapping theorem applied to the $\inf_{d\in D_1}$ mapping, which is continuous because D_1 is compact, it then holds that $\inf_{d\in D_1}G_T(d)\overset{w}{\to}\inf_{d\in D_1}V(d)\chi_1^2$, which is positive almost surely. It follows that

$$\inf_{\theta\in\Theta_1}Q_T(\theta)\geq\inf_{\theta\in\Theta_1}(\sum\nolimits_{n=0}^{\infty}c_n(\psi))^2T^{2(d_0-d)-1}G_T(d)+o_p(1)$$

and, for any $K>0$,

$$P(\inf_{\theta\in\Theta_1}(\sum\nolimits_{n=0}^{\infty}c_n(\psi))^2T^{2(d_0-d)-1}G_T(d)>K)\to 1\text{ as }T\to\infty$$

because $\inf_{\psi\in\Psi}(\sum_{n=0}^{\infty}c_n(\psi))^2>0$ by Assumption 3 and $2(d_0-d)-1\geq 2\kappa_1>0$ for $d\in D_1$.

S.5.1.3 Convergence on $\Theta_2(\kappa_1,\kappa_2)$ First note that by (S.24) of Lemma S.2 it suffices to prove the result for $T^{-1}\sum_{t=1}^{T}(\Delta_+^{d-d_0}e_t(\psi))^2$. Letting $v:=d-d_0\in[-1/2-\kappa_1,-1/2+\kappa_2]$, $R_{1T}(v):=T^{-1}\sum_{t=1}^{T}(\Delta_+^{v}\varepsilon_t)^2$, and $R_{2T}(v,\psi):=T^{-1}\sum_{t=1}^{T}(\Delta_+^{v}\varepsilon_t)(\sum_{n=0}^{\infty}\tilde{c}_n(\psi)\Delta_+^{1+v}\varepsilon_{t-n})$, and applying the decomposition (C.4), the relevant product moment is

$$T^{-1}\sum_{t=1}^{T}(\Delta_+^{v}e_t(\psi))^2\geq\left(\sum_{n=0}^{\infty}c_n(\psi)\right)^2R_{1T}(v)+2\left(\sum_{n=0}^{\infty}c_n(\psi)\right)R_{2T}(v,\psi).$$

The second term, $R_{2T}(v,\psi)$, is $O_p(1)$ uniformly in $|v+1/2| \leq \kappa_1$ and $\psi \in \Psi$ by Lemma B.3 with $\tilde{\Psi} = \Psi, \zeta_{1n}(\psi) = 1_{\{n=0\}}, \zeta_{2n}(\psi) = \tilde{c}_n(\psi), u_1 = v \geq -1/2 - \kappa_1, u_2 = 1 + v \geq 1/2 - \kappa_1$ such that $a = \min(1/2 - \kappa_1, 1 - 2\kappa_2) > 0$.

To analyze $R_{1T}(v)$ decompose $\Delta_+^v \varepsilon_t$ as

$$\Delta_+^v \varepsilon_t = \sum_{n=0}^{N-1} \pi_n(-v)\varepsilon_{t-n} + \sum_{n=N}^{t-1} \pi_n(-v)\varepsilon_{t-n} = w_{1t} + w_{2t},\ t \geq N+1,$$

for some $N \geq 1$ to be determined. It then holds that

$$R_{1T}(v) \geq T^{-1} \sum_{t=N+1}^{T} (\Delta_+^v \varepsilon_t)^2 \geq T^{-1} \sum_{t=N+1}^{T} w_{1t}^2 + 2T^{-1} \sum_{t=N+1}^{T} w_{1t} w_{2t}. \tag{S.42}$$

Setting $N = N_T := \lfloor T^\alpha \rfloor$ with $0 < \alpha < \min(\frac{1/2-\kappa_1}{1/2+\kappa_1}, \frac{1/2}{1/2+2\kappa_1})$, noting that such an α exists because $0 < \kappa_1 < 1/2$, it follows from (B.2) of Lemma B.2 that the second term on the right-hand side of (S.42) converges in probability to zero uniformly in $|v+1/2| \leq \kappa_1$ and that

$$\sup_{|v+1/2|\leq\kappa_1} \left| T^{-1} \sum_{t=N+1}^{T} w_{1t}^2 - E\left(T^{-1} \sum_{t=N+1}^{T} w_{1t}^2\right)\right| \xrightarrow{p} 0 \text{ as } T \to \infty.$$

Thus, the right-hand side of (S.42) minus $E(T^{-1}\sum_{t=N+1}^{T} w_{1t}^2)$ converges uniformly in probability to zero as $T \to \infty$. It follows, see Assumption 1(b), that

$$\begin{aligned}
\left(\sum_{n=0}^{\infty} c_n(\psi)\right)^2 R_{1T}(v) &\geq \left(\sum_{n=0}^{\infty} c_n(\psi)\right)^2 E\left(T^{-1} \sum_{t=N+1}^{T} w_{1t}^2\right) + \mu_{1T}(\theta) \\
&= \left(\sum_{n=0}^{\infty} c_n(\psi)\right)^2 T^{-1} \sum_{t=N+1}^{T} \sum_{n=0}^{N-1} \pi_n(-v)^2 \sigma_{t-n}^2 + \mu_{1T}(\theta) \\
&\geq \left(\inf_{0\leq s\leq 1} \sigma(s)^2\right) \left(\sum_{n=0}^{\infty} c_n(\psi)\right)^2 T^{-1}(T-N) F_N(v) + \mu_{1T}(\theta),
\end{aligned}$$

where $F_N(v) = \sum_{n=0}^{N-1} \pi_n(-v)^2$ and $\mu_{1T}(\theta) \xrightarrow{p} 0$ as $T \to \infty$ uniformly in $|v+1/2| \leq \kappa_1$ and $\psi \in \Psi$.

S.5.1.4 Proof of Eqn. (S.20) Because of restrictions imposed on the κ_i in the analysis of the sets Θ_i, we need to be careful in the proof of (S.20). We need to show that, for any $K > 0, \eta > 0$, there exists a $\bar{\kappa}_2 > 0$ and a T_0 such that

$$P(\inf_{\theta\in\Theta_1(\kappa_1)\cup\Theta_2(\kappa_1,\bar{\kappa}_2)} Q_T(\theta) < K) \leq \eta$$

for all $T \geq T_0$. Since $\inf_{\theta\in\Theta_1\cup\Theta_2} Q_T(\theta) \leq \sum_{j=1}^{2} \inf_{\theta\in\Theta_j} Q_T(\theta)$, the two sets Θ_1 and Θ_2 can be considered separately.

First consider the set $\Theta_1(\kappa_1)$ with $\kappa_1 = \bar{\kappa}_1$ satisfying $0 < \bar{\kappa}_1 < \min(1/2, \zeta/2 + 1/4)$, and define $\bar{\Theta}_1 := \Theta_1(\bar{\kappa}_1)$. It holds from Section S.5.1.2 that $P(\inf_{\theta\in\bar{\Theta}_1} Q_T(\theta) > K) \to 1$ as $T \to \infty$, i.e., for any $K > 0, \eta > 0$, there exists a T_1 such that $P(\inf_{\theta\in\bar{\Theta}_1} Q_T(\theta) < K) \leq \eta/2$ for all $T \geq T_1$.

Second, having already fixed $\kappa_1 = \bar{\kappa}_1$, consider $\Theta_2(\bar{\kappa}_1, \kappa_2)$. From Section S.5.1.3 with $\kappa_1 = \bar{\kappa}_1$ and $\alpha = 1/6$,

$$Q_T(\theta) \geq \left(\inf_{0\leq s\leq 1} \sigma(s)^2\right)\left(\sum_{n=0}^{\infty} c_n(\psi)\right)^2 T^{-1}(T - T^{1/6})F_{T^{1/6}}(d-d_0) + \mu_T(\theta),$$

where $\mu_T(\theta) = O_p(1)$ as $T \to \infty$ uniformly in $d \in [d_0 - 1/2 - \bar{\kappa}_1, d_0 - 1/2 + \bar{\kappa}_1] \supset D_2$ and $\psi \in \Psi$. As in Section D.1.2, $F_{T^{1/6}}(d-d_0) \geq 1 + c(2\kappa_2)^{-1}(1-(T-1)^{-2\kappa_2/6})$ and $(2\kappa_2)^{-1}(1-(T-1)^{-2\kappa_2/6}) \to \infty$ as $(\kappa_2, T) \to (0, \infty)$. Because $(\sum_{n=0}^{\infty} c_n(\psi))^2 > 0$ uniformly in $\psi \in \Psi$ and $\inf_{0\leq s\leq 1}\sigma(s)^2 > 0$, it follows that for any $K > 0, \eta > 0$, there exists $\bar{\kappa}_2 > 0$ (small) and T_2 such that, with $\bar{\Theta}_2 := \Theta_2(\bar{\kappa}_1, \bar{\kappa}_2)$, $P(\inf_{\theta\in\bar{\Theta}_2} Q_T(\theta) < K) \leq \eta/2$ for all $T \geq T_2$.

Combining these results, for any $K > 0, \eta > 0$, there exists a $\bar{\kappa}_2 > 0$ such that

$$P(\inf_{\theta\in\bar{\Theta}_1\cup\bar{\Theta}_2} Q_T(\theta) < K) \leq \sum_{j=1}^{2} P(\inf_{\theta\in\bar{\Theta}_j} Q_T(\theta) < K) \leq \sum_{j=1}^{2} \eta/2 = \eta$$

for all $T \geq \max(T_1, T_2) = T_0$, which proves (S.20).

S.5.1.5 Convergence on $\Theta_3(\kappa_2)$ and Proof of Eqn. (S.21) First, by Lemma S.2, it suffices to demonstrate the result for $T^{-1}\sum_{t=1}^{T}(\Delta_+^{d-d_0} e_t(\psi))^2$. In this case, recall the untruncated process $\eta_t(\theta)$ defined in (S.14), and note that $\eta_t(\theta) - \Delta_+^{d-d_0} e_t(\psi) = \sum_{n=t}^{\infty} \pi_n(d_0 - d)e_{t-n}(\psi) = \sum_{n=t}^{\infty} \varphi_n(\theta)\varepsilon_{t-n}$, see (S.16), with

$$E(\eta_t(\theta) - \Delta_+^{d-d_0} e_t(\psi))^2 = \sum_{n=t}^{\infty} \varphi_n(\theta)^2 \sigma_{t-n}^2 \leq c\sum_{n=t}^{\infty} n^{2\max(d_0-d-1,-2-\zeta)} \leq ct^{-2\kappa_2} \to 0$$

for all $\theta \in \Theta_3$ (pointwise), using (S.18) and Assumption 1(b). It follows that

$$Q_T(\theta) = T^{-1}\sum_{t=1}^{T} \eta_t(\theta)^2 + o_p(1). \tag{S.43}$$

Next,

$$ET^{-1}\sum_{t=1}^{T}\eta_t(\theta)^2 = T^{-1}\sum_{t=1}^{T}\sum_{n=0}^{\infty}\varphi_n(\theta)^2\sigma_{t-n}^2 = T^{-1}\sum_{t=1}^{T}\sigma_t^2\sum_{n=0}^{\infty}\varphi_n(\theta)^2 + T^{-1}\sum_{t=1}^{T}\sum_{n=0}^{\infty}\varphi_n(\theta)^2(\sigma_{t-n}^2 - \sigma_t^2).$$

Let $q_T := \lfloor T^{\chi} \rfloor$ for some $\chi \in (0,1)$. Then the last term is bounded as

$$T^{-1}\sum_{t=1}^{T}\sum_{n=0}^{\infty}\varphi_n(\theta)^2(\sigma_{t-n}^2 - \sigma_t^2) \leq \sum_{n=0}^{q_T}\varphi_n(\theta)^2 T^{-1}\sum_{t=1}^{T}|\sigma_{t-n}^2 - \sigma_t^2| \tag{S.44}$$

$$+ \sum_{n=q_T+1}^{\infty}\varphi_n(\theta)^2 T^{-1}\sum_{t=1}^{T}|\sigma_{t-n}^2 - \sigma_t^2|. \tag{S.45}$$

Notice that $T^{-1}\sum_{t=1}^{T}|\sigma_{t-n}^2 - \sigma_t^2| = T^{-1}\sum_{t=1}^{n}|\sigma_{t-n}^2 - \sigma_t^2| + T^{-1}\sum_{t=n+1}^{T}|\sigma_{t-n}^2 - \sigma_t^2| \leq T^{-1}2n\sup_{t\in\mathbb{Z}}\sigma_t^2 + T^{-1}\sum_{t=n+1}^{T}|\sigma_{t-n}^2 - \sigma_t^2|$. Therefore, because $q_T = o(T)$ and $\sup_{t\in\mathbb{Z}}\sigma_t^2 \leq$

$M < \infty$ by Assumption 1(b)(i), we have

$$\sup_{n=1,\dots,q_T} T^{-1}\sum_{t=1}^{T}|\sigma_{t-n}^2-\sigma_t^2| \leq T^{-1}2q_T\sup_{t\in\mathbb{Z}}\sigma_t^2+\sup_{n=1,\dots,q_T}T^{-1}\sum_{t=1}^{T-n}|\sigma_{t+n}^2-\sigma_t^2|$$
$$=\sup_{n=1,\dots,q_T}T^{-1}\sum_{t=1}^{T-n}|\sigma_{t+n}^2-\sigma_t^2|+o(1)\to 0,$$

where the last convergence follows from Cavaliere and Taylor (2009, Lemma A.1). Because $\sum_{n=0}^{q_T}\varphi_n(\theta)^2 \leq \sum_{n=0}^{\infty}\varphi_n(\theta)^2 < \infty$ uniformly in $\theta \in \Theta_3$, see (S.18), it thus holds that $|(S.44)| \to 0$. Next, by Assumption 1(b)(i) and by (S.18) we have, respectively, $\sup_{t\in\mathbb{Z}}\sigma_t^2 \leq M < \infty$ such that $\sup_{t\in\mathbb{Z}}T^{-1}\sum_{t=1}^{T}|\sigma_{t-n}^2-\sigma_t^2| \leq 2M$ and $\sum_{n=q_T+1}^{\infty}\varphi_n(\theta)^2 \leq c\sum_{n=q_T+1}^{\infty}n^{2\max(d_0-d-1,-2-\zeta)} \leq cq_T^{-2\kappa_2} \to 0$ uniformly in $\theta \in \Theta_3$, and therefore $|(S.45)| \to 0$. Because $T^{-1}\sum_{t=1}^{T}\sigma_t^2 \to \int_0^1\sigma(s)^2\mathsf{d}s$ by Assumption 1(b)(ii) we thus have that $ET^{-1}\sum_{t=1}^{T}\eta_t(\theta)^2 = \int_0^1\sigma(s)^2\mathsf{d}s\sum_{n=0}^{\infty}\varphi_n(\theta)^2+o(1)$. To prove

$$T^{-1}\sum_{t=1}^{T}\eta_t(\theta)^2-\int_0^1\sigma(s)^2\mathsf{d}s\sum_{n=0}^{\infty}\varphi_n(\theta)^2 \xrightarrow{p} 0, \tag{S.46}$$

pointwise in $\theta \in \Theta_3$, it suffices to show L_2-convergence. In a similar way as in (S.39) and (S.40), we find that

$$E\left(T^{-1}\sum_{t=1}^{T}\eta_t(\theta)^2-ET^{-1}\sum_{t=1}^{T}\eta_t(\theta)^2\right)^2 = T^{-2}\sum_{t,s=1}^{T}E(\eta_t(\theta)^2\eta_s(\theta)^2)-T^{-2}\sum_{t,s=1}^{T}E(\eta_t(\theta)^2)E(\eta_s(\theta)^2)$$
$$= T^{-2}\sum_{t,s=1}^{T}\sum_{n_1,n_2=0}^{\infty}\sum_{m_1,m_2=0}^{\infty}\left(\prod_{i=1}^{2}\varphi_{n_i}(\theta)\varphi_{m_i}(\theta)\sigma_{t-n_i}\sigma_{s-m_i}\right)$$
$$\times [E(z_{t-n_1}z_{t-n_2}z_{s-m_1}z_{s-m_2})-E(z_{t-n_1}z_{t-n_2})E(z_{s-m_1}z_{s-m_2})],$$

where the expectations are zero unless the two highest subscripts are equal (Lemma A.2). By symmetry, we only need to consider three cases, which we now enumerate.

Case 1) $t-n_1 = t-n_2 = s-m_1 = s-m_2$, in which case the expectations and the σ_t's are uniformly bounded by Assumption 1 and we find the contribution

$$cT^{-2}\sum_{t=1}^{T}\left(\sum_{n=0}^{\infty}\varphi_n(\theta)^2\right)^2 \leq cT^{-1}\left(\sum_{n=0}^{\infty}n^{-1-2\kappa_2}\right)^2 \leq cT^{-1} \to 0$$

using (S.18).

Case 2) $t - n_1 = t - n_2 > s - m_1 \geq s - m_2$, where the contribution is

$$T^{-2} \sum_{t,s=1}^{T} \sum_{n=0}^{\infty} \sum_{m_1=\max(0,s-t+n+1)}^{\infty} \sum_{m_2=m_1}^{\infty} \varphi_n(\theta)^2 \varphi_{m_1}(\theta) \varphi_{m_2}(\theta)$$
$$\times \sigma_{t-n}^2 \sigma_{s-m_1} \sigma_{s-m_2} \kappa_4(t-n, t-n, s-m_1, s-m_2)$$
$$\leq cT^{-2} \sum_{t,s=1}^{T} \sum_{n=0}^{\infty} n^{-1-2\kappa_2} \max(0, s-t+n+1)^{-1-2\kappa_2}$$
$$\times \sum_{m_1=\max(0,s-t+n+1)}^{\infty} \sum_{m_2=m_1}^{\infty} |\kappa_4(t-n, t-n, s-m_1, s-m_2)|$$
$$\leq cT^{-2} \sum_{t,s=1}^{T} \sum_{n=0}^{\infty} n^{-1-2\kappa_2} \max(0, s-t+n+1)^{-1-2\kappa_2} \leq cT^{-2} \sum_{t,s=1}^{T} |t-s|^{-1-2\kappa_2} \leq cT^{-1} \to 0$$

using Assumption 1(a)(iii),(b) together with (S.18).

Case 3) $t - n_1 = s - m_1 > t - n_2 \geq s - m_2$, where we distinguish between the two subcases:

Case 3a) $t - n_2 = s - m_2$ with the contribution

$$T^{-2} \sum_{t,s=1}^{T} \sum_{n_1=\max(0,t-s)}^{\infty} \sum_{n_2=n_1+1}^{\infty} \varphi_{n_1}(\theta) \varphi_{n_2}(\theta) \varphi_{s-t+n_1}(\theta) \varphi_{s-t+n_2}(\theta) \sigma_{t-n_1}^2 \sigma_{t-n_2}^2 \tau_{n_2-n_1, n_2-n_1}$$
$$\leq cT^{-2} \sum_{t,s=1}^{T} \sum_{n_1=\max(0,t-s)}^{\infty} n_1^{-1/2-\kappa_2} (s-t+n_1)^{-1/2-\kappa_2} \sum_{n_2=n_1+1}^{\infty} n_2^{-1/2-\kappa_2} (s-t+n_2)^{-1/2-\kappa_2}$$
$$\leq cT^{-2} \sum_{t \geq s=1}^{T} \sum_{n_1=t-s}^{\infty} n_1^{-1/2-\kappa_2} (s-t+n_1)^{-1/2-\kappa_2} \sum_{n_2=n_1+1}^{\infty} n_2^{-1/2-\kappa_2} (s-t+n_2)^{-1/2-\kappa_2}$$
$$\leq cT^{-2} \sum_{t \geq s=1}^{T} \sum_{n_1=t-s}^{\infty} n_1^{-1/2-2\kappa_2} (s-t+n_1)^{-1/2-\kappa_2} \leq cT^{-2} \sum_{t \geq s=1}^{T} (t-s)^{-2\kappa_2} \leq cT^{-2\kappa_2} \to 0,$$

where we once again used (S.18) and Assumption 1(a)(ii),(b).

Case 3b) $t - n_2 > s - m_2$ with the contribution

$$T^{-2} \sum_{t,s=1}^{T} \sum_{n_1=\max(0,t-s)}^{\infty} \sum_{n_2=n_1+1}^{\infty} \sum_{m=s-t+n_2+1}^{\infty} \varphi_{n_1}(\theta) \varphi_{n_2}(\theta) \varphi_{s-t+n_1}(\theta) \varphi_m(\theta)$$
$$\times \sigma_{t-n_1}^2 \sigma_{t-n_2} \sigma_{s-m} \kappa_4(t-n_1, t-n_1, t-n_2, s-m)$$
$$\leq cT^{-2} \sum_{t \geq s=1}^{T} \sum_{n_1=t-s}^{\infty} n_1^{-1/2-2\kappa_2} (s-t+n_1)^{-1/2-\kappa_2} \leq cT^{-2\kappa_2} \to 0$$

as in Case 3a). This shows that (S.46) holds pointwise for all $\theta \in \Theta_3$.

Comparing the pointwise limit found in (S.46) with the definition of $Q(\theta)$ in (S.19), it remains only to show that

$$\sup_{\theta \in \Theta_3} \left| \sum_{n=0}^{\infty} \varphi_n(\theta)^2 - \sum_{n=0}^{\infty} \varphi_{0,n}(\theta)^2 \right| \to 0. \tag{S.47}$$

By the mean value theorem, the term inside the absolute value on the left-hand side is

$$\sum_{n=0}^{\infty}\left(\left(\sum_{m=0}^{n}\pi_m(d_0-d)c_{n-m}(\psi)\right)^2-\left(\sum_{l=0}^{n}\pi_l(d_0-d)\sum_{k=0}^{n-l}b_k(\psi)a_{n-l}(\psi_0)\right)^2\right)$$
$$=\frac{2\delta_\psi}{\sqrt{T}}\sum_{n=0}^{\infty}\sum_{m=0}^{n}\pi_m(d_0-d)\sum_{k=0}^{n-m}b_k(\psi)a_{n-m-k}(\bar{\psi})\sum_{l=0}^{n}\pi_l(d_0-d)\sum_{j=0}^{n-l}b_j(\psi)\dot{a}_{n-l-j}(\bar{\psi}),$$

where $\bar{\psi}$ is an intermediate value between ψ_0 and $\psi_{0,T}$. Taking the supremum of the absolute value we first find, using Assumption 3(iii) and Lemmas A.3 and A.4, that $\sup_{\psi\in\Psi}\sum_{j=0}^{n-l}|b_j(\psi)||\dot{a}_{n-l-j}(\bar{\psi})|\le c\sum_{j=0}^{n-l}j^{-2-\zeta}(n-l-j)^{-1-\zeta}\le c(n-l)^{-1-\zeta}$ and $\sup_{\psi\in\Psi}\sum_{k=0}^{n-m}|b_k(\psi)||a_{n-m-k}(\bar{\psi})|\le c\sum_{k=0}^{n-m}k^{-2-\zeta}(n-m-k)^{-2-\zeta}\le c(n-m)^{-2-\zeta}$. Thus, the left-hand side of (S.47) is bounded by

$$\frac{c}{\sqrt{T}}\sum_{n=0}^{\infty}\sum_{m=0}^{n}m^{-1/2-\kappa_2}(n-m)^{-2-\zeta}\sum_{l=0}^{n}l^{-1/2-\kappa_2}(n-l)^{-1-\zeta}\le\frac{c}{\sqrt{T}}\sum_{n=0}^{\infty}n^{-1/2-\kappa_2}n^{-1/2-\kappa_2}\le\frac{c}{\sqrt{T}}\to 0.$$

Combining (S.43), (S.46), and (S.47), we obtain the pointwise limit, i.e.

$$Q_T(\theta)\xrightarrow{p}Q(\theta). \tag{S.48}$$

The result (S.48) can be strengthened to uniform convergence in probability by showing that $T^{-1}\sum_{t=1}^{T}(\Delta_+^{d-d_0}e_t(\psi))^2$ is stochastically equicontinuous (or tight). From Newey (1991, Corollary 2.2) this holds if the derivative of $Q_T(\theta)$ is dominated uniformly in $\theta\in\Theta_3$ by a random variable $B_T=O_p(1)$. From Lemma B.3 with $u_1=u_2=d-d_0\ge -1/2+\kappa_2, a=2\kappa_2$, and $\tilde{\Psi}=\Psi$ it holds that $B_T=\sup_{\theta\in\Theta_3}|\frac{\partial}{\partial\theta}T^{-1}\sum_{t=1}^{T}(\Delta_+^{d-d_0}e_t(\psi))^2|=O_p(1)$ (noting that only summability of the linear coefficients is assumed in Lemma B.3 and this is satisfied uniformly on Θ by the derivatives of $c_n(\psi)$ by Assumption 3(iii)). This shows that $T^{-1}\sum_{t=1}^{T}(\Delta_+^{d-d_0}e_t(\psi))^2$ is stochastically equicontinuous on Θ_3 and hence that (S.48) holds uniformly in $\theta\in\Theta_3$ in view of Lemma S.2. Since the result holds for any κ_2 it proves (S.21).

S.5.1.6 Proof of Eqn. (S.22) Since $Q(\theta_0)=\int_0^1\sigma(s)^2\mathsf{d}s$ it is sufficient to prove that

$$\inf_{\theta\in\Theta_3\cap\{\theta:|\theta-\theta_0|\ge\epsilon\}}\sum_{n=0}^{\infty}\varphi_n(\theta)^2>1\text{ for all }\epsilon>0\text{ and all }\kappa_2\in(0,1/2).$$

Because $\varphi_0(\theta)=1$ for all $\theta\in\Theta_3$ by Assumption 3, it is clear that $\sum_{n=0}^{\infty}\varphi_n(\theta)^2=1+\sum_{n=1}^{\infty}\varphi_n(\theta)^2\ge 1$, and by Assumption 4 the inequality is strict for all $\theta\ne\theta_0$, which proves (S.22) by continuity of $\varphi_n(\cdot)$ and compactness of Θ_3.

S.5.2 Proof of Theorem 2

By consistency of $\hat{\theta}$, the asymptotic distribution theory for the QML estimator is obtained from the usual Taylor series expansion of the score function. That is,

$$0=T^{1/2}\frac{\partial Q_T(\hat{\theta})}{\partial\theta}=T^{1/2}\frac{\partial Q_T(\theta_{0,T})}{\partial\theta}+T^{1/2}\frac{\partial^2 Q_T(\bar{\theta})}{\partial\theta\partial\theta'}(\hat{\theta}-\theta_{0,T}), \tag{S.49}$$

where $\bar{\theta}$ is an intermediate value satisfying $|\bar{\theta}_i - \theta_{0,T,i}| \leq |\hat{\theta}_i - \theta_{0,T,i}|$ for $i = 1, \ldots, p+1$. Recalling the definition of $\xi_n(\theta_1, \theta_2)$ in (D.38), we note for this proof that, for example, $\xi_n(\theta_0, \theta) = \frac{\partial \varphi_{0,n}(\theta)}{\partial \theta}$, see (S.14) and (S.16). We also define $\xi_{0,n} := \xi_n(\theta_0, \theta_0) = [-n^{-1}, \gamma_n(\psi_0)']'$, which satisfies

$$\sum_{n=0}^{s} ||\xi_{0,n}|| = O(\log s) \text{ and } \sum_{n=0}^{s} (\xi_{0,n})_i^q = O(1) \text{ for any } q > 1, s \geq 2, i = 1, \ldots, p+1, \quad \text{(S.50)}$$

by Assumption 3(iii) and (5).

S.5.2.1 Convergence of the Score Function The normalized score function evaluated at the true value is

$$T^{1/2} \frac{\partial Q_T(\theta_{0,T})}{\partial \theta} = 2T^{-1/2} \sum_{t=1}^{T} \varepsilon_t(\theta_{0,T}) \hat{y}_{1,t-1} \text{ with } \hat{y}_{k,t-1} := \frac{\partial^k}{\partial \theta^{(k)}} \varepsilon_t(\theta_{0,T}).$$

Define also $S_T := 2T^{-1/2} \sum_{t=1}^{T} \varepsilon_t y_{1,t-1}$, where $y_{1,t-1} := \sum_{n=1}^{t-1} \xi_{0,n} \varepsilon_{t-n}$. That is, the first element of $y_{1,t-1}$ is $-\sum_{n=1}^{t-1} n^{-1} \varepsilon_{t-n}$ and the remaining p elements are given by $\sum_{n=1}^{t-1} \gamma_n(\psi_0) \varepsilon_{t-n}$. Similarly, the first element of $\hat{y}_{1,t-1}$ is $-\sum_{n=1}^{t-1} n^{-1} \varepsilon_{t-n}(\theta_{0,T})$ and the remaining elements are $\sum_{n=1}^{t-1} \dot{b}_n(\psi_{0,T}) u_{t-n}$.

We next show that

$$T^{1/2} \frac{\partial Q_T(\theta_{0,T})}{\partial \theta} - S_T = o_p(1). \quad \text{(S.51)}$$

The left-hand side of (S.51) is

$$T^{1/2} \frac{\partial Q_T(\theta_{0,T})}{\partial \theta} - S_T = 2T^{-1/2} \sum_{t=1}^{T} (\varepsilon_t(\theta_{0,T}) - \varepsilon_t) \hat{y}_{1,t-1} + 2T^{-1/2} \sum_{t=1}^{T} \varepsilon_t (\hat{y}_{1,t-1} - y_{1,t-1}),$$

where

$$\varepsilon_t(\theta_{0,T}) - \varepsilon_t = -\sum_{n=t}^{\infty} b_n(\psi_{0,T}) u_{t-n}$$

and

$$\hat{y}_{1,t-1} - y_{1,t-1} = \begin{bmatrix} -\sum_{n=1}^{t-1} n^{-1} \sum_{k=t-n}^{\infty} b_k(\psi_{0,T}) u_{t-n-k} \\ \sum_{n=1}^{t-1} \dot{b}_n(\psi_{0,T}) \sum_{k=t}^{\infty} a_k(\psi_{0,T}) \varepsilon_{t-k} \end{bmatrix}.$$

The first term on the right-hand side of (S.51) is then

$$-2T^{-1/2} \sum_{t=1}^{T} \sum_{n=t}^{\infty} b_n(\psi_{0,T}) u_{t-n} \hat{y}_{1,t-1},$$

which has second moment

$$4T^{-1} \sum_{t,s=1}^{T} \sum_{n=t}^{\infty} \sum_{m=s}^{\infty} b_n(\psi_{0,T}) b_m(\psi_{0,T}) E(u_{t-n} \hat{y}_{1,t-1} u_{s-m} \hat{y}_{1,s-1})$$
$$\leq KT^{-1} \sum_{t,s=1}^{T} \sum_{n=t}^{\infty} \sum_{m=s}^{\infty} |b_n(\psi_{0,T})||b_m(\psi_{0,T})| \leq KT^{-1} \sum_{t,s=1}^{T} t^{-1-\zeta} s^{-1-\zeta} \leq KT^{-1-2\zeta} \to 0,$$

see (5). The second term on the right-hand side of (S.51) is

$$\begin{bmatrix} -2T^{-1/2}\sum_{t=1}^{T}\varepsilon_t\sum_{n=1}^{t-1}n^{-1}\sum_{k=t-n}^{\infty}b_k(\psi_{0,T})u_{t-n-k} \\ 2T^{-1/2}\sum_{t=1}^{T}\varepsilon_t\sum_{n=1}^{t-1}\dot{b}_n(\psi_{0,T})\sum_{k=t}^{\infty}a_k(\psi_{0,T})\varepsilon_{t-k} \end{bmatrix}. \tag{S.52}$$

The first term in (S.52) has second moment

$$\begin{aligned}
&4T^{-1}\sum_{t,s=1}^{T}\sum_{n=1}^{t-1}\sum_{m=1}^{s-1}n^{-1}m^{-1}\sum_{k=t-n}^{\infty}\sum_{l=s-m}^{\infty}b_k(\psi_{0,T})b_l(\psi_{0,T})E(\varepsilon_t\varepsilon_s u_{t-n-k}u_{s-m-l})\\
&=4T^{-1}\sum_{t=1}^{T}\sum_{n,m=1}^{t-1}n^{-1}m^{-1}\sum_{k=t-n}^{\infty}\sum_{l=t-m}^{\infty}b_k(\psi_{0,T})b_l(\psi_{0,T})\sum_{r=0}^{\infty}\sum_{q=0}^{\infty}a_r(\psi_{0,T})a_q(\psi_{0,T})\sigma_t^2\sigma_{t-k-n-r}\sigma_{t-l-m-q}\\
&\quad\times(\kappa_4(t,t,t-k-n-r,t-l-m-q)+\kappa_2(t,t)\kappa_2(t-k-n-r,t-l-m-q))\\
&\le KT^{-1}\sum_{t=1}^{T}\sum_{n,m=1}^{t-1}n^{-1}m^{-1}\sum_{k=t-n}^{\infty}\sum_{l=t-m}^{\infty}|b_k(\psi_{0,T})||b_l(\psi_{0,T})|\sum_{r=0}^{\infty}\sum_{q=0}^{\infty}|a_r(\psi_{0,T})||a_q(\psi_{0,T})|\\
&\le KT^{-1}\sum_{t=1}^{T}\sum_{n,m=1}^{t-1}n^{-1}m^{-1}\sum_{k=t-n}^{\infty}\sum_{l=t-m}^{\infty}k^{-2-\zeta}l^{-2-\zeta}\\
&\le KT^{-1}\sum_{t=1}^{T}\sum_{n,m=1}^{t-1}n^{-1}m^{-1}(t-n)^{-1-\zeta}(t-m)^{-1-\zeta}\le KT^{-1}\sum_{t=1}^{T}t^{-2}\le KT^{-1}\to 0,
\end{aligned}$$

where the first two inequalities use Assumption 1(a)(iii),(b) and (5), and the fourth inequality uses Lemma A.4. The second term in (S.52) has second moment

$$\begin{aligned}
&4T^{-1}\sum_{t,s=1}^{T}\sum_{n=1}^{t-1}\sum_{m=1}^{s-1}\dot{b}_n(\psi_{0,T})\dot{b}_m(\psi_{0,T})\sum_{k=t}^{\infty}\sum_{l=s}^{\infty}a_k(\psi_{0,T})a_l(\psi_{0,T})E(\varepsilon_t\varepsilon_s\varepsilon_{t-k}\varepsilon_{s-l})\\
&\le KT^{-1}\sum_{t=1}^{T}(\sum_{k=t}^{\infty}|a_k(\psi_{0,T})|)^2\le KT^{-1}\sum_{t=1}^{T}(t^{-1-\zeta})^2\to 0
\end{aligned}$$

using Lemma A.2, Assumption 3(iii), and (5). Thus, each of the terms in (S.52), and hence those in (S.51), converge to zero in L_2-norm and therefore in probability.

Because $y_{1,t-1}$ is measurable with respect to the sigma-algebra $\mathcal{F}_{t-1} := \sigma(\{\varepsilon_s, s \le t-1\})$, it holds that $v_{Tt} := 2T^{-1/2}\varepsilon_t\sum_{n=1}^{t-1}\xi_{0,n}\varepsilon_{t-n} = 2T^{-1/2}\sigma_t z_t\sum_{n=1}^{t-1}\xi_{0,n}\sigma_{t-n}z_{t-n}$ is a MDS with respect to the filtration $\mathcal{F}_t$. To apply the central limit theorem for martingales, see Lemma A.1, we first verify the Lindeberg condition (i) via Lyapunov's sufficient condition that $\sum_{t=1}^{T}E||v_{Tt}||^{2+\epsilon}\to 0$ for some $\epsilon > 0$. Thus,

$$\begin{aligned}
T^{1+\epsilon/2}E||v_{Tt}||^{2+\epsilon}&\le KE\left(|z_t|^{2+\epsilon}(\sum_{n=1}^{t-1}||\xi_{0,n}|||z_{t-n}|)^{2+\epsilon}\right)\le K\left(\sum_{n=1}^{t-1}||\xi_{0,n}||\left(E(|z_t||z_{t-n}|)^{2+\epsilon}\right)^{1/(2+\epsilon)}\right)^{2+\epsilon}\\
&\le K\left(\sum_{n=1}^{t-1}||\xi_{0,n}||\right)^{2+\epsilon}\le K(\log T)^{2+\epsilon}
\end{aligned}$$

using Assumption 1(b), Minkowski's inequality, (S.50), and Assumption 5 with ϵ chosen such that $2\epsilon+4\le 8$. It follows that $\sum_{t=1}^{T}E||v_{Tt}||^{2+\epsilon}\le KT^{-\epsilon/2}(\log T)^{2+\epsilon}\to 0$.

Next, we verify condition (ii)(a) of Lemma A.1. The sum of squares of v_{Tt} is

$$4T^{-1}\sum_{t=1}^{T}\sigma_t^2 z_t^2 \sum_{n,m=1}^{t-1} \xi_{0,n}\xi_{0,m}'\sigma_{t-n}\sigma_{t-m}z_{t-n}z_{t-m}$$

$$= 4T^{-1}\sum_{t=1}^{T}\sigma_t^2 \sum_{n,m=1}^{t-1} \xi_{0,n}\xi_{0,m}'\sigma_{t-n}\sigma_{t-m}\tau_{n,m} \tag{S.53}$$

$$+ 4T^{-1}\sum_{t=1}^{T}\sigma_t^2 \sum_{n,m=1}^{t-1} \xi_{0,n}\xi_{0,m}'\sigma_{t-n}\sigma_{t-m}(z_t^2 z_{t-n}z_{t-m} - \tau_{n,m}). \tag{S.54}$$

The second moment of the (i,j)'th element of (S.54) is

$$16T^{-2}\sum_{t,s=1}^{T}\sigma_t^2\sigma_s^2 \sum_{n,m=1}^{s-1}\sum_{k,l=1}^{t-1}(\xi_{0,m})_i(\xi_{0,n})_j(\xi_{0,k})_i(\xi_{0,l})_j\sigma_{s-n}\sigma_{s-m}\sigma_{t-k}\sigma_{t-l}Cov(z_t^2 z_{t-k}z_{t-l}, z_s^2 z_{s-n}z_{s-m})$$

$$\leq KT^{-2}\sum_{t,s=1}^{T}\sum_{n,m=1}^{s-1}\sum_{k,l=1}^{t-1}||\xi_{0,m}||||\xi_{0,n}||||\xi_{0,k}||||\xi_{0,l}|||Cov(z_t^2 z_{t-k}z_{t-l}, z_s^2 z_{s-n}z_{s-m})|$$

$$= KT^{-2}\sum_{t=1}^{T}\sum_{n,m=1}^{t-1}\sum_{k,l=1}^{t-1}||\xi_{0,m}||||\xi_{0,n}||||\xi_{0,k}||||\xi_{0,l}|||Cov(z_t^2 z_{t-n}z_{t-m}, z_t^2 z_{t-k}z_{t-l})| \tag{S.55}$$

$$+ KT^{-2}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\sum_{n,m=1}^{s-1}\sum_{k,l=1}^{t-1}||\xi_{0,m}||||\xi_{0,n}||||\xi_{0,k}||||\xi_{0,l}|||Cov(z_t^2 z_{t-k}z_{t-l}, z_s^2 z_{s-n}z_{s-m})|. \tag{S.56}$$

For (S.55) we find the simple bound

$$KT^{-2}\sum_{t=1}^{T}\left(\sum_{k=1}^{t-1}||\xi_{0,k}||\right)^4 \leq KT^{-1}(\log T)^4 \to 0$$

using (S.50) and that z_t has finite eighth order moments by Assumption 5. The covariance in (S.56) is a combination of the cumulants of z_t up to order eight, where, apart from the eighth order term, each term is a product of two cumulants whose orders sum to eight. For the term with the eighth order cumulant we find the bound

$$T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\sum_{n,m=1}^{s-1}\sum_{k,l=1}^{t-1}|\kappa_8(t,t,t-k,t-l,s,s,s-n,s-m)| \leq KT^{-1} \to 0$$

by Assumption 5. There are no seventh order cumulants in (S.56) because they would be multiplied by a first order cumulant, which is zero. For the terms with products of sixth and second order cumulants we find, for example,

$$T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\sum_{n,m=1}^{s-1}\sum_{k,l=1}^{t-1}||\xi_{0,m}||||\xi_{0,n}||||\xi_{0,k}||||\xi_{0,l}|||\kappa_2(t-k,t-l)|\kappa_6(t,t,s,s,s-n,s-m)|$$

$$\leq KT^{-2}\sum_{t=2}^{T}\left(\sum_{s=1}^{t-1}\sum_{n,m=1}^{s-1}|\kappa_6(t,t,s,s,s-n,s-m)|\right)\left(\sum_{k=1}^{t-1}||\xi_{0,k}||^2\right) \leq KT^{-1}(\log T) \to 0$$

by (S.50) and Assumption 5. Another example is

$$\begin{aligned}
T^{-2}&\sum_{t=2}^{T}\sum_{s=1}^{t-1}\sum_{n,m=1}^{s-1}\sum_{k,l=1}^{t-1}||\xi_{0,m}||||\xi_{0,n}||||\xi_{0,k}||||\xi_{0,l}||\kappa_2(t,t)|\kappa_6(t-k,t-l,s,s,s-n,s-m)|\\
&\leq KT^{-2}\sum_{t=2}^{T}\sum_{k=1}^{t-1}\sum_{n,m=1}^{t-1}\sum_{s=\max(n,m)+1}^{t-1-k}||\xi_{0,m}||||\xi_{0,n}||||\xi_{0,k}||^2|\kappa_6(t-k,t-k,s,s,s-n,s-m)|\\
&\quad+KT^{-2}\sum_{t=2}^{T}\sum_{k,l=1}^{t-1}\sum_{s=t-\min(k,l)}^{t-1}\sum_{n,m=1}^{s-1}||\xi_{0,m}||||\xi_{0,n}||||\xi_{0,k}||||\xi_{0,l}|||\kappa_6(t-k,t-l,s,s,s-n,s-m)|\\
&\leq KT^{-2}\sum_{t=2}^{T}\sum_{k=1}^{t-1}||\xi_{0,k}||^2\sum_{n,m=1}^{t-1}\sum_{s=\max(n,m)+1}^{t-1-k}|\kappa_6(t-k,t-k,s,s,s-n,s-m)|\\
&\quad+KT^{-2}\sum_{s=1}^{T-1}\sum_{k=1}^{t-1}||\xi_{0,k}||\sum_{t=\max(k,s)+1}^{s+k-1}\sum_{n,m=1}^{s-1}\sum_{l=k}^{t-1}|\kappa_6(t-k,t-l,s,s,s-n,s-m)|
\end{aligned}$$

using Lemma A.2 and symmetry. Here, both terms are clearly $O(T^{-1}(\log T))$ by (S.50) and Assumption 5. The remaining products of sixth and second order cumulants, as well as products of lower order cumulants, are treated similarly, thus proving that (S.56) and hence (S.54) is $o_p(1)$.

By Lemma A.5 with $g_{t,n,m}=\tau_{n,m}$, (S.53) is, apart from a $o(1)$ term,

$$4T^{-1}\sum_{t=1}^{T}\sigma_t^4\sum_{n,m=1}^{t-1}\xi_{0,n}\xi_{0,m}'\tau_{n,m}=4T^{-1}\sum_{t=1}^{T}\sigma_t^4\sum_{n,m=1}^{\infty}\xi_{0,n}\xi_{0,m}'\tau_{n,m}-4T^{-1}\sum_{t=1}^{T}\sigma_t^4\sum_{n,m=t}^{\infty}\xi_{0,n}\xi_{0,m}'\tau_{n,m},$$

where the first term on the right-hand side is $4A_0T^{-1}\sum_{t=1}^{T}\sigma_t^4\to 4A_0\int_0^1\sigma^4(s)\mathsf{d}s$ and the second term on the right-hand side is bounded by $KT^{-1}\sum_{t=1}^{T}\sum_{n,m=t}^{\infty}||\xi_{0,n}||||\xi_{0,m}|||\tau_{n,m}|$, which converges to zero because it is the Cesàro mean of the sequence $\sum_{n,m=t}^{\infty}||\xi_{0,n}||||\xi_{0,m}|||\tau_{n,m}|$, which itself converges to zero as $t\to\infty$ since it is the tail of a convergent sum, see Assumption 1(a)(iii) and Remark 4.6.

It follows that the sum of squares of v_{Tt} satisfies

$$4T^{-1}\sum_{t=1}^{T}\sigma_t^2z_t^2\sum_{m,n=1}^{t-1}\xi_{0,m}\xi_{0,n}'\sigma_{t-m}\sigma_{t-n}z_{t-m}z_{t-n}\overset{p}{\to}4A_0\int_0^1\sigma^4(s)\mathsf{d}s. \tag{S.57}$$

Hence, by the central limit theorem for martingales, see Lemma A.1, we have $S_T\overset{w}{\to}N(0,4A_0\int_0^1\sigma^4(s)\mathsf{d}s)$ and therefore also $T^{1/2}\frac{\partial Q_T(\theta_{0,T})}{\partial\theta}\overset{w}{\to}N(0,4A_0\int_0^1\sigma^4(s)\mathsf{d}s)$ by (S.51).

S.5.2.2 Convergence of the Hessian The second derivative in (S.49) is tight (stochastically equicontinuous) by Newey (1991, Corollary 2.2) if its derivative is dominated uniformly in $d\in D_3,\psi\in\mathcal{N}_\delta(\psi_0)$ by a random variable $B_T=O_p(1)$. From Lemma B.3 with $u_1=u_2=d-d_{0,T}\geq-1/2+\kappa_2/2$ (for T sufficiently large) and $\tilde{\Psi}=\mathcal{N}_\delta(\psi_0)$ (noting that only summability of the linear coefficients is assumed in Lemma B.3 and this is satisfied uniformly on $\mathcal{N}_\delta(\psi_0)$ by the derivatives of $c_n(\psi)$ by Assumption 6) it holds that $B_T=\sup_{d\in D_3,\psi\in\mathcal{N}_\delta(\psi_0)}|\frac{\partial^3Q_T(\theta)}{\partial\theta^{(3)}}|=O_p(1)$, showing that the second derivative in (S.49) is tight. This result, together with $|\hat{\theta}-\theta_{0,T}|\overset{p}{\to}0$ (Theorem 1), implies by Lemma A.3 of

Johansen and Nielsen (2010) that the second derivative in (S.49) can be evaluated at the true value, $\theta_{0,T}$. Hence, we examine

$$\frac{\partial^2 Q_T(\theta_{0,T})}{\partial\theta\partial\theta'} = 2T^{-1}\sum_{t=1}^{T}\varepsilon_t(\theta_{0,T})\hat{y}_{2,t-1} + 2T^{-1}\sum_{t=1}^{T}\hat{y}_{1,t-1}\hat{y}_{1,t-1}',$$

and by the same argument as for the score, it is enough to consider $H_T := 2T^{-1}\sum_{t=1}^{T}\varepsilon_t\hat{y}_{2,t-1} + 2T^{-1}\sum_{t=1}^{T}y_{1,t-1}y_{1,t-1}'$. Because $\hat{y}_{2,t-1}$ is measurable with respect to $\mathcal{F}_t$, $\varepsilon_t\hat{y}_{2,t-1}$ is a MDS, and it has finite variance such that the first term of H_T is $o_p(1)$.

The second term of H_T is $2T^{-1}\sum_{t=1}^{T}\sum_{n,m=1}^{t-1}\xi_{0,n}\varepsilon_{t-n}\xi_{0,m}'\varepsilon_{t-m}$, which converges in L_2-norm, and hence in probability, to $2B_0\int_0^1\sigma^2(s)\mathsf{d}s$ exactly as in Section D.2.2 (just replacing $\xi_n^{\dagger}$ with $\xi_{0,n}$).

S.5.2.3 Proof of (17) To prove the result for $\hat{A}$ we write

$$\begin{aligned}\hat{A} = T^{-1}\sum_{t=1}^{T}\frac{\partial\ell_t(\hat{\theta},\hat{\sigma}^2)}{\partial\theta}\frac{\partial\ell_t(\hat{\theta},\hat{\sigma}^2)}{\partial\theta'} &= \frac{1}{\hat{\sigma}^4}T^{-1}\sum_{t=1}^{T}\varepsilon_t(\hat{\theta})^2\frac{\partial\varepsilon_t(\hat{\theta})}{\partial\theta}\frac{\partial\varepsilon_t(\hat{\theta})}{\partial\theta'} \\ &= \frac{1}{\hat{\sigma}^4}\left(T^{-1}\sum_{t=1}^{T}\varepsilon_t(\hat{\theta})^2\frac{\partial\varepsilon_t(\hat{\theta})}{\partial\theta}\frac{\partial\varepsilon_t(\hat{\theta})}{\partial\theta'} - T^{-1}\sum_{t=1}^{T}\varepsilon_t(\theta_{0,T})^2\hat{y}_{1,t-1}\hat{y}_{1,t-1}'\right) \quad (S.58) \\ &+ \frac{1}{\hat{\sigma}^4}\left(T^{-1}\sum_{t=1}^{T}\varepsilon_t(\theta_{0,T})^2\hat{y}_{1,t-1}\hat{y}_{1,t-1}' - T^{-1}\sum_{t=1}^{T}\varepsilon_t^2 y_{1,t-1}y_{1,t-1}'\right) \quad (S.59) \\ &+ \frac{1}{\hat{\sigma}^4}T^{-1}\sum_{t=1}^{T}\varepsilon_t^2 y_{1,t-1}y_{1,t-1}'. \quad (S.60)\end{aligned}$$

First of all, $\hat{\sigma}^2 = Q_T(\hat{\theta}) \overset{p}{\to} Q(\theta_0) = \int_0^1\sigma^2(s)\mathsf{d}s$ by the uniform convergence in (S.48), Theorem 1, and Johansen and Nielsen (2010, Lemma A.3).

Next, we decompose the (i,j)'th element of (S.58) and apply the Cauchy-Schwarz inequality,

$$\begin{aligned}&\frac{1}{\hat{\sigma}^4}T^{-1}\sum_{t=1}^{T}(\varepsilon_t(\hat{\theta})^2 - \varepsilon_t(\theta_{0,T})^2)\frac{\partial\varepsilon_t(\hat{\theta})}{\partial\theta_i}\frac{\partial\varepsilon_t(\hat{\theta})}{\partial\theta_j'} + \frac{1}{\hat{\sigma}^4}T^{-1}\sum_{t=1}^{T}\varepsilon_t(\theta_{0,T})^2\left(\frac{\partial\varepsilon_t(\hat{\theta})}{\partial\theta_i}\frac{\partial\varepsilon_t(\hat{\theta})}{\partial\theta_j'} - \hat{y}_{1,t-1,i}\hat{y}_{1,t-1,j}'\right) \\ &\leq \frac{1}{\hat{\sigma}^4}\left(T^{-1}\sum_{t=1}^{T}(\varepsilon_t(\hat{\theta})^2 - \varepsilon_t(\theta_{0,T})^2)^2\right)^{1/2}\left(T^{-1}\sum_{t=1}^{T}\left(\frac{\partial\varepsilon_t(\hat{\theta})}{\partial\theta_i}\frac{\partial\varepsilon_t(\hat{\theta})}{\partial\theta_j'}\right)^2\right)^{1/2} \quad (S.61) \\ &+ \frac{1}{\hat{\sigma}^4}\left(T^{-1}\sum_{t=1}^{T}\varepsilon_t(\theta_{0,T})^4\right)^{1/2}\left(T^{-1}\sum_{t=1}^{T}\left(\frac{\partial\varepsilon_t(\hat{\theta})}{\partial\theta}\frac{\partial\varepsilon_t(\hat{\theta})}{\partial\theta'} - \hat{y}_{1,t-1}\hat{y}_{1,t-1}'\right)^2\right)^{1/2}. \quad (S.62)\end{aligned}$$

The proofs for (S.61) and (S.62) are nearly identical, so we give only the former. The second large parenthesis in (S.61) is $O_p(1)$ by Lemma B.3. By the mean value theorem,

$$T^{-1}\sum_{t=1}^{T}(\varepsilon_t(\hat{\theta})^2 - \varepsilon_t(\theta_{0,T})^2)^2 = 4\sum_{i=1}^{p+1}(\hat{\theta}_i - \theta_{0,T,i})T^{-1}\sum_{t=1}^{T}(\varepsilon_t(\hat{\theta})^2 - \varepsilon_t(\theta_{0,T})^2)\frac{\partial\varepsilon_t(\bar{\theta})}{\partial\theta_i}$$

for an intermediate value, $\bar{\theta}$, between $\hat{\theta}$ and $\theta_{0,T}$. By another application of the Cauchy-Schwarz inequality,

$$T^{-1}\sum_{t=1}^{T}(\varepsilon_t(\hat{\theta})^2-\varepsilon_t(\theta_{0,T})^2)\frac{\partial\varepsilon_t(\bar{\theta})}{\partial\theta_i}\leq(T^{-1}\sum_{t=1}^{T}(\varepsilon_t(\hat{\theta})^2-\varepsilon_t(\theta_{0,T})^2)^2)^{1/2}(T^{-1}\sum_{t=1}^{T}(\frac{\partial\varepsilon_t(\bar{\theta})}{\partial\theta_i})^2)^{1/2},$$

which is also $O_p(1)$ by Lemma B.3. Because $(\hat{\theta}_i-\theta_{0,T,i})=O_p(T^{-1/2})$ by Theorem 2 and $\hat{\sigma}^4\xrightarrow{p}(\int_0^1\sigma^2(s)\mathsf{d}s)^2$, it follows that (S.61) is $o_p(1)$. Next, (S.59) is negligible by the exact same argument as in the proof of (S.51), and finally (S.60) is $\hat{\sigma}^{-4}\sum_{t=1}^{T}v_{Tt}v_{Tt}'/4\xrightarrow{p}(\int_0^1\sigma^2(s)\mathsf{d}s)^{-2}A_0\int_0^1\sigma^4(s)\mathsf{d}s=\lambda A_0$ by (S.57) and using $\hat{\sigma}^2\xrightarrow{p}\int_0^1\sigma^2(s)\mathsf{d}s$. It follows that $\hat{A}\xrightarrow{p}\lambda A_0$.

For the second result we find that

$$\hat{B}=-\frac{\partial^2L_T(\hat{\theta},\hat{\sigma}^2)}{\partial\theta\partial\theta'}=\frac{1}{2\hat{\sigma}^2}\frac{\partial^2Q_T(\hat{\theta},\hat{\sigma}^2)}{\partial\theta\partial\theta'}\xrightarrow{p}\frac{1}{2\int_0^1\sigma^2(s)\mathsf{d}s}2B_0\int_0^1\sigma^2(s)\mathsf{d}s=B_0$$

by the proof in Section S.5.2.2 and using $\hat{\sigma}^2\xrightarrow{p}\int_0^1\sigma^2(s)\mathsf{d}s$. Finally, it now follows straightforwardly, using Assumption 7 and Slutsky's Theorem, that $\hat{C}\xrightarrow{p}C_0$.

S.5.3 Proof of Theorem 3

Consider first the Wald statistic. From (16) of Theorem 2 we find, under $H_{1,T}$, that

$$\sqrt{T}(M'\hat{\theta}-m)\xrightarrow{w}N(\delta,M'C_0M).$$

It follows by (17) and the continuous mapping theorem that $W_T\xrightarrow{w}Y'F_0Y$. For the robust Wald statistic, the result follows in the same way by Theorem 2 and the continuous mapping theorem. Finally, the proofs for the LM and LR statistics apply standard mean-value or Taylor series expansions; for a textbook treatment, see for example Hayashi (2000, Section 7.4).

S.5.4 Proof of Theorem 4

Again, consider first the Wald statistic. Under the fixed alternative H_1 in (3) the true value is θ_0, i.e. $\delta_\theta=0$, and is such that $M'\theta_0=\bar{m}\neq m$. From Theorem 2 we then find

$$\sqrt{T}(M'\hat{\theta}-m)+\sqrt{T}(m-\bar{m})\xrightarrow{w}N(0,M'C_0M).$$

Since $\hat{B}\xrightarrow{p}B_0$ by (17), it follows that

$$\begin{aligned}W_T&=(\sqrt{T}(m-\bar{m})+O_p(1))'(M'B_0^{-1}M+o_p(1))^{-1}(\sqrt{T}(m-\bar{m})+O_p(1))\\&=T(m-\bar{m})'(M'B_0^{-1}M)^{-1}(m-\bar{m})+O_p(T^{1/2}).\end{aligned}$$

The proofs for the LM, LR, and robust Wald statistics again follow by standard expansions.

S.6 Additional Proofs for Bootstrap Estimator and Tests

S.6.1 Proof of Lemma D.1

Recall that $\hat{\varepsilon}_{c,t} = \varepsilon_t(\hat{\theta}) - T^{-1}\sum_{s=1}^{T}\varepsilon_s(\hat{\theta})$ and decompose as

$$\begin{aligned} T^{-1}\sum_{t=1}^{T}(\hat{\varepsilon}_{c,t}^2 - \varepsilon_t^2)^2 &= T^{-1}\sum_{t=1}^{T}(\varepsilon_t(\hat{\theta})^2 + (T^{-1}\sum_{s=1}^{T}\varepsilon_s(\hat{\theta}))^2 - 2\varepsilon_t(\hat{\theta})T^{-1}\sum_{s=1}^{T}\varepsilon_s(\hat{\theta}) - \varepsilon_t^2)^2 \\ &= T^{-1}\sum_{t=1}^{T}(\varepsilon_t(\hat{\theta})^2 - \varepsilon_t^2)^2 + (T^{-1}\sum_{s=1}^{T}\varepsilon_s(\hat{\theta}))^4 + 4T^{-1}\sum_{t=1}^{T}\varepsilon_t(\hat{\theta})^2(T^{-1}\sum_{s=1}^{T}\varepsilon_s(\hat{\theta}))^2 \\ &\quad + \text{cross product terms.} \end{aligned} \tag{S.63}$$

The cross product terms are asymptotically of the required order by the Cauchy-Schwarz inequality, after dealing with the first three terms on the right-hand side.

First we write $\varepsilon_t(\hat{\theta}) = \varepsilon_t(\hat{\theta}) - \varepsilon_t(\theta_{0,T}) + \varepsilon_t(\theta_{0,T}) - \varepsilon_t + \varepsilon_t$ and find that

$$T^{-1}\sum_{s=1}^{T}\varepsilon_s(\hat{\theta}) = T^{-1}\sum_{s=1}^{T}(\varepsilon_s(\hat{\theta}) - \varepsilon_s(\theta_{0,T})) + T^{-1}\sum_{s=1}^{T}(\varepsilon_s(\theta_{0,T}) - \varepsilon_s) + T^{-1}\sum_{s=1}^{T}\varepsilon_s, \tag{S.64}$$

where the last term is clearly $O_p(T^{-1/2})$ under Assumption 1. Using (6), the second term of (S.64) is $T^{-1}\sum_{s=1}^{T}\sum_{m=s}^{\infty} b_m(\psi_{0,T})u_{s-m}$, which has zero mean and variance bounded by

$$cT^{-2}\sum_{t,s=1}^{T}\sum_{m=s}^{\infty}\sum_{n=t}^{\infty} b_m(\psi_{0,T})b_n(\psi_{0,T}) \le cT^{-2}\sum_{t,s=1}^{T}\sum_{m=s}^{\infty}\sum_{n=t}^{\infty} m^{-2-\zeta}n^{-2-\zeta} \le cT^{-2}\sum_{t,s=1}^{T} s^{-1-\zeta}t^{-1-\zeta} \le cT^{-2},$$

see (5), so that the second term of (S.64) is $O_p(T^{-1})$ by L_2-convergence. For the first term of (S.64) we apply the mean value theorem,

$$T^{-1}\sum_{s=1}^{T}(\varepsilon_s(\hat{\theta}) - \varepsilon_s(\theta_{0,T})) = (\hat{\theta} - \theta_{0,T})'T^{-1}\sum_{t=1}^{T}\frac{\partial\varepsilon_t(\bar{\theta})}{\partial\theta},$$

where $\bar{\theta}$ is an intermediate value between $\hat{\theta}$ and $\theta_{0,T}$. By the Cauchy-Schwarz inequality and Lemma B.3, $T^{-1}\sum_{t=1}^{T}\frac{\partial\varepsilon_t(\bar{\theta})}{\partial\theta} \le (T^{-1}\sum_{t=1}^{T}(\frac{\partial\varepsilon_t(\bar{\theta})}{\partial\theta})^2)^{1/2} = O_p(1)$ when $\bar{d}$ is close to d_0. Since $(\hat{\theta} - \theta_{0,T}) = O_p(T^{-1/2})$ by Theorem 2, this shows that the first term of (S.64), and hence (S.64), is $O_p(T^{-1/2})$. Because $T^{-1}\sum_{t=1}^{T}\varepsilon_t(\hat{\theta})^2 = O_p(1)$ it follows that the second and third terms of (S.63) are both $O_p(T^{-1})$ such that we are left with the first term on the right-hand side of (S.63).

To deal with the first term of (S.63), we again write $\varepsilon_t(\hat{\theta}) = \varepsilon_t(\hat{\theta}) - \varepsilon_t(\theta_{0,T}) + \varepsilon_t(\theta_{0,T}) - \varepsilon_t + \varepsilon_t$ and find that

$$\begin{aligned} T^{-1}\sum_{t=1}^{T}(\varepsilon_t(\hat{\theta})^2 - \varepsilon_t^2)^2 &= T^{-1}\sum_{t=1}^{T}((\varepsilon_t(\hat{\theta}) - \varepsilon_t(\theta_{0,T}))^4 + T^{-1}\sum_{t=1}^{T}(\varepsilon_t(\theta_{0,T}) - \varepsilon_t)^4 \\ &\quad + \text{cross product terms.} \end{aligned} \tag{S.65}$$

Again, the cross product terms are asymptotically of the required order by the Cauchy-Schwarz inequality, if the first two terms on the right-hand side are dealt with. Using

(6), the second term on the right-hand side of (S.65) is $T^{-1}\sum_{t=1}^{T}(\sum_{m=t}^{\infty} b_m(\psi_{0,T})u_{t-m})^4$, which is a non-negative random variable with mean

$$T^{-1}\sum_{t=1}^{T} E(\sum_{m=t}^{\infty} b_m(\psi_{0,T})u_{t-m})^4 \leq cT^{-1}\sum_{t=1}^{T}(\sum_{m=t}^{\infty} b_m(\psi_{0,T}))^4 \leq cT^{-1}\sum_{t=1}^{T}(\sum_{m=t}^{\infty} m^{-2-\zeta})^4$$
$$\leq cT^{-1}\sum_{t=1}^{T} t^{-4-4\zeta} \leq cT^{-1},$$

see (5), which shows that the second term of (S.65) is $O_p(T^{-1})$ by L_1-convergence. For the first term of (S.65), we apply the mean value theorem followed by the Cauchy-Schwarz inequality,

$$\begin{aligned}
&T^{-1}\sum_{t=1}^{T}((\varepsilon_t(\hat{\theta}) - \varepsilon_t(\theta_{0,T}))^4 \\
&= 4\sum_{i=1}^{p+1}(\hat{\theta}_i - \theta_{0,T,i})T^{-1}\sum_{t=1}^{T}((\varepsilon_t(\hat{\theta}) - \varepsilon_t(\theta_{0,T}))^3 \frac{\partial \varepsilon_t(\bar{\theta})}{\partial \theta_i} \\
&\leq 4\sum_{i=1}^{p+1}(\hat{\theta}_i - \theta_{0,T,i})\left(T^{-1}\sum_{t=1}^{T}((\varepsilon_t(\hat{\theta}) - \varepsilon_t(\theta_{0,T}))^6\right)^{1/2}\left(T^{-1}\sum_{t=1}^{T}\left(\frac{\partial \varepsilon_t(\bar{\theta})}{\partial \theta_i}\right)^2\right)^{1/2},
\end{aligned}$$

where $\bar{\theta}$ is an intermediate value between $\hat{\theta}$ and $\theta_{0,T}$ and $T^{-1}\sum_{t=1}^{T}((\varepsilon_t(\hat{\theta}) - \varepsilon_t(\theta_{0,T}))^6$ is at most $O_p(1)$. Since $T^{-1}\sum_{t=1}^{T}(\frac{\partial \varepsilon_t(\bar{\theta})}{\partial \theta_i})^2 = O_p(1)$ by Lemma B.3 and $(\hat{\theta} - \theta_{0,T}) = O_p(T^{-1/2})$ by Theorem 2, this shows that the first term of (S.65) is $O_p(T^{-1/2})$, and hence completes the proof.

S.6.2 Proof of Lemma D.2

The proofs for the two cases $h = k+1$ and $h \leq m-1$ are identical, so we give only the former. First apply summation by parts,

$$\begin{aligned}
\sum_{j=m}^{k}\lambda_j(\theta)\sum_{t=k+2}^{T}\varepsilon^*_{t-j}\varepsilon^*_{t-k-1} &= \lambda_k(\theta)\sum_{j=m}^{k}\sum_{t=k+2}^{T}\varepsilon^*_{t-j}\varepsilon^*_{t-k-1} \\
&\quad - \sum_{q=m}^{k-1}(\lambda_{q+1}(\theta) - \lambda_q(\theta))\sum_{j=m}^{q}\sum_{t=k+2}^{T}\varepsilon^*_{t-j}\varepsilon^*_{t-k-1},
\end{aligned}$$

which implies that

$$\begin{aligned}
E^*\sup_{\theta}\left|\sum_{j=m}^{k}\lambda_j(\theta)\sum_{t=k+2}^{T}\varepsilon^*_{t-j}\varepsilon^*_{t-k-1}\right| &\leq \sup_{\theta}|\lambda_k(\theta)|E^*\left|\sum_{j=m}^{k}\sum_{t=k+2}^{T}\varepsilon^*_{t-j}\varepsilon^*_{t-k-1}\right| \\
&\quad + \sup_{\theta}\sum_{q=m}^{k-1}|\lambda_{q+1}(\theta) - \lambda_q(\theta)|E^*\left|\sum_{j=m}^{q}\sum_{t=k+2}^{T}\varepsilon^*_{t-j}\varepsilon^*_{t-k-1}\right|.
\end{aligned}$$

Next notice that, by Jensen's inequality,

$$\left(E^*\left|\sum_{j=m}^{k}\sum_{t=k+2}^{T}\varepsilon^*_{t-j}\varepsilon^*_{t-k-1}\right|\right)^2 \leq E^*\left|\sum_{j=m}^{k}\sum_{t=k+2}^{T}\varepsilon^*_{t-j}\varepsilon^*_{t-k-1}\right|^2$$
$$=\sum_{j=m}^{k}\sum_{j'=m}^{k}\sum_{t=k+2}^{T}\sum_{t'=k+2}^{T}E^*(\varepsilon^*_{t-j}\varepsilon^*_{t'-j'}\varepsilon^*_{t-k-1}\varepsilon^*_{t'-k-1})$$
$$=2\sum_{j=m}^{k}\sum_{t=k+2}^{T}E^*(\varepsilon^{*2}_{t-j}\varepsilon^{*2}_{t-k-1}) = 2\sum_{j=m}^{k}\sum_{t=k+2}^{T}\hat{\varepsilon}^2_{c,t-j}\hat{\varepsilon}^2_{c,t-k-1},$$

where, by the Cauchy-Schwarz inequality,

$$\sum_{j=m}^{k}\sum_{t=k+2}^{T}\hat{\varepsilon}^2_{c,t-j}\hat{\varepsilon}^2_{c,t-k-1} \leq \sum_{j=m}^{k}(\sum_{t=k+2}^{T}\hat{\varepsilon}^4_{c,t-j})^{1/2}(\sum_{t=k+2}^{T}\hat{\varepsilon}^4_{c,t-k-1})^{1/2} \leq \sum_{j=m}^{k}\sum_{t=1}^{T}\hat{\varepsilon}^4_{c,t} = (k-m+1)O_p(T).$$

Therefore,

$$E^*\sup_\theta\left|\sum_{j=m}^{k}\lambda_j(\theta)\sum_{t=k+2}^{T}\varepsilon^*_{t-j}\varepsilon^*_{t-k-1}\right| \leq \sup_\theta|\lambda_k(\theta)|k^{1/2}O_p(T^{1/2}) + \sup_\theta\sum_{q=m}^{k-1}|\lambda_{q+1}(\theta)-\lambda_q(\theta)|q^{1/2}O_p(T^{1/2})$$
$$\leq O_p(T^{1/2}k^{g+1/2}) + O_p(T^{1/2})\sum_{q=m}^{k}q^{g-1/2},$$

which proves the result.

S.6.3 Proof of Lemma D.3

Reversing the summations, we find $M^*_{12NT}(u) = T^{-1}\sum_{n=0}^{N-1}\pi_n(-u)\sum_{m=N}^{T-1}\pi_m(-u)\sum_{t=m+1}^{T}\varepsilon^*_{t-n}\varepsilon^*_{t-m}$, and we apply Lemma D.2 with $g=-1/2+\kappa$,

$$E^*\sup_{|u+1/2|\leq\kappa}\left|\sum_{m=N}^{T-1}\pi_m(-u)\sum_{t=m+1}^{T}\varepsilon^*_{t-n}\varepsilon^*_{t-m}\right| = O_p(T^{1/2+\kappa}),$$

which implies that

$$E^*\sup_{|u+1/2|\leq\kappa}|M^*_{12NT}(u)| \leq \sup_{|u+1/2|\leq\kappa}T^{-1}\sum_{n=0}^{N-1}|\pi_n(-u)|O_p(T^{1/2+\kappa}) = O_p(N^{\kappa+1/2}T^{\kappa-1/2}),$$

so that $\sup_{|u+1/2|\leq\kappa}|T^{-1}\sum_{t=N+1}^{T}w^*_{1t}w^*_{2t}| = o^*_p(1)$, in probability, by setting $N=T^\alpha$ with $\alpha<(1/2-\kappa)/(1/2+\kappa)$.

Next, we decompose $M^*_{11NT}(u)$ as

$$M^*_{11NT}(u) = T^{-1}\sum_{t=N+1}^{T}\sum_{n=0}^{N-1}\pi_n(-u)^2(\varepsilon^{*2}_{t-n}-\sigma^2_{t-n}) \quad (S.66)$$
$$+T^{-1}\sum_{t=N+1}^{T}\sum_{n\neq m=0}^{N-1}\pi_n(-u)\pi_m(-u)\varepsilon^*_{t-n}\varepsilon^*_{t-m}, \quad (S.67)$$

where

$$E^* \sup_{|u+1/2|\leq\kappa} |(S.67)| = \sup_{|u+1/2|\leq\kappa} \sum_{n\neq m=0}^{N-1} |\pi_n(-u)||\pi_m(-u)| E^* \left| T^{-1} \sum_{t=N+1}^{T} \varepsilon^*_{t-n}\varepsilon^*_{t-m} \right|$$
$$\leq c \sum_{n\neq m=0}^{N-1} n^{\kappa-1/2} m^{\kappa-1/2} O_p(T^{-1/2}) = O_p(N^{2\kappa+1}T^{-1/2}),$$

with the first inequality following from (D.57). Thus, $E^* \sup_{|u+1/2|\leq\kappa} |(S.67)| = o_p(1)$ when $N = T^\alpha$ with $\alpha < 1/(4\kappa+2)$. We decompose (S.66) as

$$(S.66) = T^{-1} \sum_{t=N+1}^{T} \sum_{n=0}^{N-1} \pi_n(-u)^2 (\varepsilon^{*2}_{t-n} - \hat{\varepsilon}^2_{c,t-n}) \tag{S.68}$$
$$+ T^{-1} \sum_{t=N+1}^{T} \sum_{n=0}^{N-1} \pi_n(-u)^2 (\hat{\varepsilon}^2_{c,t-n} - \varepsilon^2_{t-n}) \tag{S.69}$$
$$+ T^{-1} \sum_{t=N+1}^{T} \sum_{n=0}^{N-1} \pi_n(-u)^2 (\varepsilon^2_{t-n} - \sigma^2_{t-n}), \tag{S.70}$$

and show that each of these terms are asymptotically negligible (in the sense of $\overset{p^*}{\to}_p 0$). First,

$$E^* \sup_{|u+1/2|\leq\kappa} |(S.68)| \leq \sup_{|u+1/2|\leq\kappa} \sum_{n=0}^{N-1} \pi_n(-u)^2 E^* \left| T^{-1} \sum_{t=N+1}^{T} (\varepsilon^{*2}_{t-n} - \hat{\varepsilon}^2_{c,t-n}) \right|$$
$$\leq c \sum_{n=0}^{N-1} n^{2\kappa-1} O_p(T^{-1/2}) = O_p(N^{2\kappa}T^{-1/2}),$$

where the second inequality follows by (D.32). Thus, $E^* \sup_{|u+1/2|\leq\kappa} |(S.68)| = o_p(1)$ for $N = T^\alpha$ with $\alpha < 1/(4\kappa)$.

Next, using $\varepsilon_t = \sigma_t z_t$,

$$\left(E \left| T^{-1} \sum_{t=N+1}^{T} \sigma^2_{t-n}(z^2_{t-n} - 1) \right| \right)^2 \leq E \left(T^{-1} \sum_{t=N+1}^{T} \sigma^2_{t-n}(z^2_{t-n} - 1) \right)^2$$
$$= T^{-2} \sum_{t,s=N+1}^{T} \sigma^2_{t-n}\sigma^2_{s-n}\kappa_4(t-n, t-n, s-n, s-n)$$
$$\leq cT^{-2} \sum_{t,s=N+1}^{T} |\kappa_4(t-n, t-n, s-n, s-n)| \leq cT^{-1}$$

by Assumption 1(a)(iii),(b). Thus,

$$E \sup_{|u+1/2|\leq\kappa} |(S.70)| \leq c \sup_{|u+1/2|\leq\kappa} \sum_{n=0}^{N-1} \pi_n(-u)^2 T^{-1/2} = O(N^{2\kappa}T^{-1/2}),$$

such that $\sup_{|u+1/2|\leq\kappa} |(S.70)| = o_p(1)$ for $N = T^\alpha$ with $\alpha < 1/(4\kappa)$.

For the term (S.69), we apply the Cauchy-Schwarz inequality,

$$|(S.69)| \leq \sum_{n=0}^{N-1} \pi_n(-u)^2 \left(T^{-1} \sum_{t=N+1}^{T} (\hat{\varepsilon}_{c,t-n}^2 - \varepsilon_{t-n}^2)^2 \right)^{1/2},$$

where the last term is $O_p(T^{-1/2})$ by Lemma D.1, uniformly in $n = 0, \ldots, N-1$ and $N = 1, \ldots, T-1$, and the first term satisfies $\sup_{|u+1/2| \leq \kappa} \sum_{n=0}^{N-1} \pi_n(-u)^2 \leq cN^{2\kappa}$. Thus, $\sup_{|u+1/2| \leq \kappa} |(S.69)| = O_p(N^{2\kappa} T^{-1/2})$ which is $o_p(1)$ when $N = T^\alpha$ with $\alpha < 1/(4\kappa)$.

S.6.4 Proof of Lemma D.4

The bootstrap residual is

$$\begin{aligned}
\varepsilon_t^*(\theta) &= \sum_{n=0}^{t-1} b_n(\psi) \Delta_+^{d-\check{d}} \sum_{m=0}^{t-n-1} a_m(\check{\psi}) \varepsilon_{t-n-m}^* \\
&= \sum_{n=0}^{\infty} b_n(\psi) \Delta_+^{d-\check{d}} \sum_{m=0}^{\infty} a_m(\check{\psi}) \varepsilon_{t-n-m}^* = \Delta_+^{d-\check{d}} e_t^*(\psi),
\end{aligned}$$

where the first equality is the definition in (25), the second is because $\varepsilon_t^* = 0$ for $t \leq 0$ in step (iii) of Algorithms 1 and 2, and the final equality is by definition of $e_t^*(\psi)$ and $\check{c}(L, \psi)$, see (D.6) and (D.7). The results (D.13) and (D.14) are trivial consequences of $\varepsilon_t^*(\theta) = \Delta_+^{d-\check{d}} e_t^*(\psi)$.

S.6.5 Proof for remainder in Eqn. (D.17)

With $\check{M}_T^*(d) := T^{2(d-\check{d})} \sum_{t=1}^{T} (\Delta_+^{d-\check{d}} \varepsilon_t^*)^2$ we find from (D.16) that

$$\begin{aligned}
T^{2(d-\check{d})} \sum_{t=1}^{T} (\Delta_+^{d-\check{d}} e_t^*(\psi))^2 &= \left(\sum_{n=0}^{\infty} \check{c}_n(\psi) \right)^2 \check{M}_T^*(d) + T^{2(d-\check{d})} \sum_{t=1}^{T} \left(\sum_{n=0}^{\infty} \bar{c}_n(\psi) \Delta_+^{d-\check{d}+1} \varepsilon_{t-n}^* \right)^2 \\
&\quad + 2 \left(\sum_{n=0}^{\infty} \check{c}_n(\psi) \right) T^{2(d-\check{d})} \sum_{t=1}^{T} \Delta_+^{d-\check{d}} \varepsilon_t^* \sum_{n=0}^{\infty} \bar{c}_n(\psi) \Delta_+^{d-\check{d}+1} \varepsilon_{t-n}^*
\end{aligned}$$

such that (because the second term on the right-hand side is non-negative)

$$\begin{aligned}
q_{1,T}^*(\theta) &= 2 \left(\sum_{n=0}^{\infty} \check{c}_n(\psi) \right) T^{2(d-\check{d})} \sum_{t=1}^{T} \Delta_+^{d-\check{d}} \varepsilon_t^* \sum_{n=0}^{\infty} \bar{c}_n(\psi) \Delta_+^{d-\check{d}+1} \varepsilon_{t-n}^* \\
&\leq 2 \left(\sum_{n=0}^{\infty} \check{c}_n(\psi) \right) \check{M}_T^*(d)^{1/2} \left(T^{2(d-\check{d})} \sum_{t=1}^{T} (\sum_{n=0}^{\infty} \bar{c}_n(\psi) \Delta_+^{d-\check{d}+1} \varepsilon_{t-n}^*)^2 \right)^{1/2} \quad \text{(S.71)}
\end{aligned}$$

using the Cauchy-Schwarz inequality. The term in the first parenthesis satisfies $0 < |\sum_{n=0}^{\infty} \check{c}_n(\psi)| < \infty$ almost surely uniformly in $\psi \in \Psi$.

Next, we show that $\check{M}_T^*(d) = O_p^*(1)$, in probability, uniformly in $d \in \check{D}_1$. For the pointwise argument, first note that

$$\begin{aligned}
E^* \check{M}_T^*(d) &= T^{2(d-\check{d})} \sum_{t=1}^{T} \sum_{j,k=0}^{t-1} \pi_j(\check{d} - d) \pi_k(\check{d} - d) E^*(\varepsilon_{t-j}^* \varepsilon_{t-k}^*) \\
&= T^{2(d-\check{d})} \sum_{t=1}^{T} \sum_{j=0}^{t-1} \pi_j(\check{d} - d)^2 \hat{\varepsilon}_{c,t-j}^2 = T^{2(d-\check{d})+1} \sum_{j=0}^{T-1} \pi_j(\check{d} - d)^2 T^{-1} \sum_{t=j+1}^{T} \hat{\varepsilon}_{c,t-j}^2,
\end{aligned}$$

where the second equality follows by uncorrelatedness of ε_t^*, conditional on the original data, the third equality by reversing the order of the summations, and where $T^{-1}\sum_{t=j+1}^{T}\hat{\varepsilon}_{c,t-j}^2 = O_p(1)$ uniformly in $j = 0,\ldots,T-1$. Thus, by Lemma A.3,

$$E^*\check{M}_T^*(d) = O_p(1)T^{-1}\sum_{j=0}^{T-1}(j/T)^{2(\check{d}-d-1)} \leq O_p(1)T^{-1}\sum_{j=0}^{T-1}(j/T)^{-1+2\kappa_1},$$

where the inequality applies the definition of $\check{D}_1$ and $T^{-1}\sum_{j=0}^{T-1}(j/T)^{-1+2\kappa_1} \to \int_0^1 u^{-1+2\kappa_1}\mathsf{d}u < \infty$ because $-1+2\kappa_1 > -1$. Thus, $\check{M}_T^*(d) = O_p^*(1)$, in probability, pointwise for any $d \in \check{D}_1$. To strengthen this to hold uniformly in $d \in D_1^\dagger$ it is sufficient to show that $\check{M}_T^*(d)$ is tight (in probability) as a stochastic process on the space of continuous functions indexed by the parameter d. Using the mean value theorem, the tightness condition in (D.19) is satisfied by the same proof as the pointwise proof that $\check{M}_T^*(d) = O_p^*(1)$, in probability, except the derivative means we apply (A.2) from Lemma A.3 and find $T^{-1}\sum_{j=0}^{T-1}(j/T)^{-1+2\kappa_1}(1+\log|j/T|) \to \int_0^1 u^{-1+2\kappa_1}(1+\log|u|)\mathsf{d}u < \infty$ because $-1+2\kappa_1 > -1$. It follows that the second term on the right-hand side of (S.71), i.e. $\check{M}_T^*(d)$, is $O_p^*(1)$ in probability, uniformly in $d \in \check{D}_1$.

The term inside the second large parenthesis in (S.71) can be rewritten as

$$T^{2(d-\check{d})}\sum_{t=1}^{T}\sum_{n,m=0}^{\infty}\bar{c}_n(\psi)\bar{c}_m(\psi)\sum_{j,k=0}^{t-1}\pi_j(\check{d}-d-1)\pi_k(\check{d}-d-1)\varepsilon_{t-j-n}^*\varepsilon_{t-k-m}^*$$
$$= T^{2(d-\check{d})+1}\sum_{n,m=0}^{\infty}\bar{c}_n(\psi)\bar{c}_m(\psi)\sum_{j,k=0}^{T-1}\pi_j(\check{d}-d-1)\pi_k(\check{d}-d-1)T^{-1}\sum_{t=\max(j,k)+1}^{T}\varepsilon_{t-j-n}^*\varepsilon_{t-k-m}^*.$$

Taking the supremum we find the bound

$$\sup_{\theta\in\check{\Theta}_1} T^{2(d-\check{d})+1}\sum_{n,m=0}^{\infty}|\bar{c}_n(\psi)\bar{c}_m(\psi)|\sum_{j,k=0}^{T-1}|\pi_j(\check{d}-d-1)\pi_k(\check{d}-d-1)|\left|T^{-1}\sum_{t=\max(j,k)+1}^{T}\varepsilon_{t-j-n}^*\varepsilon_{t-k-m}^*\right|, \tag{S.72}$$

which is $o_p^*(1)$, in probability, thereby implying that $\sup_{\theta\in\check{\Theta}_1}|(S.71)| = o_p^*(1)$, in probability. To see that (S.72) is $o_p^*(1)$, in probability, note that

$$E^*\left|T^{-1}\sum_{t=\max(j,k)+1}^{T}\varepsilon_{t-j-n}^*\varepsilon_{t-k-m}^*\right| \leq T^{-1}\sum_{t=\max(j,k)+1}^{T}|\hat{\varepsilon}_{c,t-j-n}||\hat{\varepsilon}_{c,t-k-m}|$$
$$\leq \left(T^{-1}\sum_{t=\max(j,k)+1}^{T}\hat{\varepsilon}_{c,t-j-n}^2\right)^{1/2}\left(T^{-1}\sum_{t=\max(j,k)+1}^{T}\hat{\varepsilon}_{c,t-k-m}^2\right)^{1/2} \leq T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_{c,t}^2.$$

Then (S.72) is a non-negative random variable with (conditional) expectation

$$\sup_{\theta\in\check{\Theta}_1} T^{2(d-\check{d})+1}\sum_{n,m=0}^{\infty}|\bar{c}_n(\psi)\bar{c}_m(\psi)|\sum_{j,k=0}^{T-1}|\pi_j(\check{d}-d-1)\pi_k(\check{d}-d-1)|T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_{c,t}^2,$$

where $\sum_{n=0}^{\infty} |\bar{c}_n(\psi)| < \infty$ almost surely uniformly in $\psi \in \Psi$. This leaves the bound

$$E^* \sup_{\theta \in \check{\Theta}_1} T^{2(d-\check{d})} \sum_{t=1}^{T} (\sum_{n=0}^{\infty} \bar{c}_n(\psi) \Delta_+^{d-\check{d}+1} \varepsilon^*_{t-n})^2 \leq c \sup_{d \in \check{D}_1} T^{2(d-\check{d})+1} \left(\sum_{j=0}^{T-1} |\pi_j(\check{d}-d-1)| \right)^2 T^{-1} \sum_{t=1}^{T} \hat{\varepsilon}_{c,t}^2$$

$$= O_p(\sup_{d \in \check{D}_1} T^{2(d-\check{d})+1} (\sum_{j=0}^{T-1} j^{\check{d}-d-2})^2) = O_p((\log T)^2 T^{-2\kappa_1})$$

by application of Lemma A.3.

S.6.6 Proof for remainder in Eqn. (D.18)

By independence (conditional on the original data) of ε^*_t we find $\sum_{t=1}^{T} E^*(U_{Tt}^{*2}|\mathcal{F}^*_{t-1}) = T^{2(d-\check{d}-1/2)} \sum_{t=1}^{T} \pi_{T-t}(\check{d}-d+1)^2 \hat{\varepsilon}_{c,t}^2$, such that

$$q_{2,T}(d) = T^{2(d-\check{d}-1/2)} \sum_{t=1}^{T} \pi_{T-t}(\check{d}-d+1)^2 (\hat{\varepsilon}_{c,t}^2 - \varepsilon_t^2) \tag{S.73}$$

$$+ \sum_{t=1}^{T} (T^{2(d-\check{d}-1/2)} \pi_{T-t}(\check{d}-d+1)^2 - T^{2(d-d^\dagger-1/2)} \pi_{T-t}(d^\dagger - d+1)^2) \varepsilon_t^2 \tag{S.74}$$

$$+ T^{2(d-d^\dagger-1/2)} \sum_{t=1}^{T} \pi_{T-t}(d^\dagger - d+1)^2 \varepsilon_t^2 - V^\dagger(d). \tag{S.75}$$

Applying the Cauchy-Schwarz inequality, (S.73) is bounded as

$$|(S.73)| \leq \left(T^{4(d-\check{d})-1} \sum_{t=1}^{T} \pi_{T-t}(\check{d}-d+1)^4 \right)^{1/2} \left(T^{-1} \sum_{t=1}^{T} (\hat{\varepsilon}_{c,t}^2 - \varepsilon_t^2)^2 \right)^{1/2},$$

where the term in the second parenthesis is $o_p(1)$ by Lemma D.1 and the term in the first parenthesis is bounded (uniformly for $d \in \check{D}_1$) using Lemma A.3 as $cT^{-1} \sum_{t=1}^{T} (\frac{T-t}{T})^{4(\check{d}-d)} \leq cT^{-1} \sum_{t=1}^{T} (\frac{T-t}{T})^{2+4\kappa_1} \leq c$.

To analyze (S.74), we apply the mean value theorem and note that the derivative of $f(\check{d}) := T^{2(d-\check{d}-1/2)} \pi_{T-t}(\check{d}-d+1)^2$ is bounded as

$$\left| \frac{\partial f(\check{d})}{\partial \check{d}} \right| \leq c(1 + \log |\tfrac{T-t}{T}|) T^{2(d-\check{d}-1/2)} (T-t)^{2(\check{d}-d)}$$

using (A.2) of Lemma A.3. Then, $(S.74) = (\check{d} - d^\dagger) \sum_{t=1}^{T} \frac{\partial f(\bar{d})}{\partial \check{d}} \varepsilon_t^2$, where $\bar{d}$ is an intermediate value between $\check{d}$ and $d^\dagger$. For any $\epsilon > 0$, $|\bar{d} - \check{d}| \leq \epsilon$ with probability converging to one, so that

$$\sup_{d \in \check{D}_1} |(S.74)| \leq |\check{d} - d^\dagger| T^{-1} \sum_{t=1}^{T} (1 + \log |\tfrac{T-t}{T}|) (\tfrac{T-t}{T})^{1-2\kappa_1-2\epsilon} \varepsilon_t^2$$

$$\leq |\check{d} - d^\dagger| T^{-1} \sum_{t=1}^{T} \varepsilon_t^2 = O_p(|\check{d} - d^\dagger|) = o_p(1).$$

Finally, (S.75) is $o_p(1)$ by the same argument as the corresponding term in Section S.5.1.2.

S.6.7 Proof of Bound for $R^*_{2T}(\check{v},\psi)$ in Eqn. (D.21)

To bound $R^*_{2T}(\check{v},\psi)$ we note that the summation over n can be truncated at $n = t-1$ because $\varepsilon^*_t = 0$ for $t \le 0$, and we decompose as

$$
\begin{aligned}
R^*_{2T}(\check{v},\psi) &= T^{-1}\sum_{t=1}^{T}\sum_{j=0}^{t-1}\pi_j(-\check{v})\varepsilon^*_{t-j}\sum_{n=0}^{t-1}\bar{c}_n(\psi)\sum_{k=0}^{t-n-1}\pi_k(-\check{v}-1)\varepsilon^*_{t-n-k} \\
&= T^{-1}\sum_{t=1}^{T}\sum_{j=0}^{t-1}\pi_j(-\check{v})\sum_{n=0}^{j-1}\bar{c}_n(\psi)\pi_{j-n}(-\check{v}-1)\varepsilon^{*2}_{t-j} &\text{(S.76)}\\
&\quad + T^{-1}\sum_{t=1}^{T}\sum_{j=0}^{t-1}\pi_j(-\check{v})\sum_{n=0}^{t-1}\bar{c}_n(\psi)\sum_{k=j-n+1}^{t-n-1}\pi_k(-\check{v}-1)\varepsilon^*_{t-j}\varepsilon^*_{t-n-k} &\text{(S.77)}\\
&\quad + T^{-1}\sum_{t=1}^{T}\sum_{j=0}^{t-1}\pi_j(-\check{v})\varepsilon^*_{t-j}\sum_{n=0}^{t-1}\bar{c}_n(\psi)\sum_{k=0}^{j-n-1}\pi_k(-\check{v}-1)\varepsilon^*_{t-n-k}. &\text{(S.78)}
\end{aligned}
$$

We give the proofs for (S.76) and (S.77) only, since the proof for (S.78) is the same as that for (S.77). Reversing the summations,

$$
(S.76) = \sum_{j=0}^{T-1}\pi_j(-\check{v})\sum_{n=0}^{j-1}\bar{c}_n(\psi)\pi_{j-n}(-\check{v}-1)T^{-1}\sum_{t=j+1}^{T}\varepsilon^{*2}_{t-j},
$$

where $T^{-1}\sum_{t=j+1}^{T}\varepsilon^{*2}_{t-j} = O^*_p(1)$, in probability, uniformly in $j = 0,\ldots,T-1$, which leaves the bound

$$
\begin{aligned}
\sup_{\theta\in\check{\Theta}_2}|(S.76)| &\le \sup_{\theta\in\check{\Theta}_2} c\sum_{j=0}^{T-1} j^{-\check{v}-1}\sum_{n=0}^{j-1}|\bar{c}_n(\psi)|(j-n)^{-\check{v}-2}O^*_p(1) \\
&= \sup_{\theta\in\check{\Theta}_2} c\sum_{n=0}^{T-1}|\bar{c}_n(\psi)|\sum_{j=n+1}^{T-1} j^{-\check{v}-1}(j-n)^{-\check{v}-2}O^*_p(1) = O^*_p(1),
\end{aligned}
$$

in probability, using Lemma A.3 and that $\sum_{n=0}^{T-1}|\bar{c}_n(\psi)| < \infty$ almost surely uniformly in $\psi \in \Psi$.

Next write $(S.77) = T^{-1}\sum_{n=0}^{T-1}\bar{c}_n(\psi)\sum_{k=n+1}^{T-1}\pi_k(-\check{v}-1)\sum_{j=0}^{k-1+n}\pi_j(-\check{v})\sum_{t=k+n+1}^{T}\varepsilon^*_{t-j}\varepsilon^*_{t-n-k}$ and apply Lemma D.2 with $g = -1/2+\kappa_1$,

$$
E^*\sup_{d\in\check{D}_2}\left|\sum_{j=0}^{k-1+n}\pi_j(-\check{v})\sum_{t=k+n+1}^{T}\varepsilon^*_{t-j}\varepsilon^*_{t-n-k}\right| = O_p(T^{1/2}(k+n)^{\kappa_1}),
$$

and hence

$$
E^*\sup_{\theta\in\check{\Theta}_2}|(S.77)| \le \sup_{\theta\in\check{\Theta}_2} T^{-1/2}\left|\sum_{n=0}^{T-1}\bar{c}_n(\psi)\right|\sum_{k=n+1}^{T-1}k^{-\check{v}-2}(k+n)^{\kappa_1}O_p(1) \le O_p((\log T)T^{\max(2\kappa_1-1,-1/2)}),
$$

because $\sum_{n=0}^{T-1}|\bar{c}_n(\psi)| < \infty$ almost surely uniformly in $\psi\in\Psi$. Thus, $\sup_{\theta\in\check{\Theta}_2}|(S.77)| = o^*_p(1)$, in probability.

S.6.8 Proof of variance of (D.63)

The variance of the (i, j)'th element of (D.63) is, apart from an asymptotically negligible term due to (D.64),

$$4T^{-2}\sum_{t,s=1}^{T}\sum_{m,n=1}^{s-1}\sum_{k,l=1}^{t-1}(\xi_m^\dagger)_i(\xi_n^\dagger)_j(\xi_k^\dagger)_i(\xi_l^\dagger)_j\sigma_{s-m}\sigma_{s-n}\sigma_{t-k}\sigma_{t-l}$$
$$\times\left[E(z_{s-m}z_{s-n}z_{t-k}z_{t-l})-E(z_{s-m}z_{s-n})E(z_{t-k}z_{t-l})\right]$$
$$\leq KT^{-2}\sum_{t,s=1}^{T}\sum_{m,n=1}^{s-1}\sum_{k,l=1}^{t-1}||\xi_m^\dagger||||\xi_n^\dagger||||\xi_k^\dagger||||\xi_l^\dagger|||E(z_{s-m}z_{s-n}z_{t-k}z_{t-l})-E(z_{s-m}z_{s-n})E(z_{t-k}z_{t-l})|,$$

using Assumption 1(b) to bound the σ_t's. Here, the expectations are zero unless the two highest subscripts are equal (Lemma A.2). By symmetry, we only need to consider three cases as follows.

Case 1) $s-m=s-n=t-k=t-l$, in which case the expectations are uniformly bounded by Assumption 1 and we find the contribution $cT^{-2}\sum_{t=1}^{T}(\sum_{n=0}^{\infty}||\xi_n^\dagger||^2)^2 \leq cT^{-1}\to 0$ using (D.39).

Case 2) $s-m=s-n>t-k\geq t-l$, where the contribution is

$$cT^{-2}\sum_{t,s=1}^{T}\sum_{n=1}^{s-1}\sum_{k,l=1}^{t-1}||\xi_n^\dagger||^2||\xi_k^\dagger||||\xi_l^\dagger|||\kappa_4(s-n,s-n,t-k,t-l)|$$
$$\leq cT^{-2}\sum_{s=1}^{T}\sum_{n=1}^{s-1}\sum_{k=1}^{T-1}||\xi_n^\dagger||^2||\xi_k^\dagger||\sum_{t=k+1}^{T}\sum_{l=1}^{t-1}|\kappa_4(s-n,s-n,t-k,t-l)|\leq cT^{-1}(\log T)\to 0$$

using Assumption 1(a)(iii) and (D.39).

Case 3) $s-m=t-k>s-n\geq t-l$, where we distinguish between the two subcases:

Case 3a) $s-n=t-l$ with the contribution

$$cT^{-2}\sum_{t,s=1}^{T}\sum_{m,n=\max(0,s-t)}^{s-1}||\xi_m^\dagger||||\xi_n^\dagger||||\xi_{t-s+m}^\dagger||||\xi_{t-s+n}^\dagger||\tau_{n-m,n-m}$$
$$\leq cT^{-2}\sum_{t,s=1}^{T}\sum_{m,n=\max(0,s-t)}^{s-1}||\xi_m^\dagger||||\xi_n^\dagger||||\xi_{t-s+m}^\dagger||||\xi_{t-s+n}^\dagger||\leq cT^{-1}(\log T)^3\to 0,$$

where we once again used Assumption 1(a)(ii) and (D.39).

Case 3b) $s-n>t-l$ with the contribution

$$cT^{-2}\sum_{t,s=1}^{T}\sum_{m,n=1}^{s-1}\sum_{l=1}^{t-1}||\xi_m^\dagger||||\xi_n^\dagger||||\xi_k^\dagger||||\xi_l^\dagger|||\kappa_4(s-m,s-m,s-n,t-l)|\leq cT^{-2}(\log T)^2\to 0$$

as in Case 2).

S.6.9 Proofs of Theorems 7 and 8

These results follow from Theorem 6 in the same way that Theorems 3 and 4 follow from Theorem 2.

Additional References

Dvoretzky, A. (1972), Asymptotic normality for sums of dependent random variables, in L.M. LeCam, J. Neyman, and E.L. Scott (eds.), *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability Vol. 2: Probability Theory*, University of California Press, Berkeley, 513–535.

McLeish, D.L. (1974), Dependent central limit theorems and invariance principles, *Annals of Statistics* 2, 620–628.

Newey, W.K. (1991), Uniform convergence in probability and stochastic equicontinuity, *Econometrica* 59, 1161–1167.

van der Vaart, A.W. (1998), *Asymptotic Statistics*, Cambridge University Press, Cambridge.