Supplementary Material to

Efficient Estimation of the Parameter Path in Unstable Time Series Models

by Ulrich K. Müller and Philippe-Emmanuel Petalas

1 Additional Risk Comparisons

1.1 Linear Trend Parameter Path

Consider the model of Section 2.3 with a linear trend specification for the parameter path

$$\theta_t = \theta_0 + \beta \frac{t}{T}. \qquad (*)$$

Condition 2 allows for the special case where $G(s) = Zs$, $Z \sim \mathcal{N}(0, c^2/H)$, so that by Theorem 1, large sample weighted average risk minimizing decisions relative to the weighting function $\beta \sim \mathcal{N}(0, c^2/HT)$ are obtained by replacing the original likelihood by the approximations (7) and (8), or (23).

We compare the following modes of inference: (i) MLE estimation of θ_0 and β from the log-likelihood $\sum l_t(\theta_0 + \beta t/T)$ with sandwich covariance matrix (trend MLE); (ii) Linear trend model estimated using approximation (7) and (8) with c known (kn c, trnd.LL) (that is, inference from the posterior $\mathcal{N}(\hat{\theta}\mathbf{e} + \Sigma\hat{\mathbf{s}}, \Sigma)$, where Σ_δ in Σ is generated by $G(s) = Zs$, $Z \sim \mathcal{N}(0, c^2/H)$); (iii) Linear trend model estimated using approximation (23) as in Table 1 with c known (kn c, trnd.Kal); (iv) Equal probability mixture of linear trend model estimated using approximation (7) and (8) with $c \in C = \{0, 5, \ldots, 50\}$ (un c, trnd.LL); (v) Equal probability mixture of linear trend model estimated using approximation (23) as in Table 1 with $c \in C = \{0, 5, \ldots, 50\}$ (un c, trnd.Kal); (vi) the two path estimators considered in Table 1 of the paper ("un c, LL" and "un c, Kal" abbreviated for "unknown c, Local Level" and "unknown c, Kalman", respectively). [For the trnd.XX methods, the monikers "LL" and "Kal" are misnomers; they are merely supposed to indicate application of the pseudo models (7) and (8), and (23), respectively, as in Table 1.]

To estimate weighted average risk, we draw data from model (10) and (*) with $\theta_0 \sim \mathcal{N}(0, 100)$ and $\beta \sim \mathcal{N}(0, c^2/HT)$. All Tables in this supplement are based on 8,000 replications. Table 2 reports weighted average risk relative to "trend MLE" inference. Theorem 1 implies that "kn c, trnd.LL" and "kn c, trnd.Kal" are large sample weighted average risk minimizing, so these entries should be smallest. As can be seen from Table 2, this approximation holds up reasonably well as long as c is not too big (and the shrinking relative to the MLE implied by these methods is especially helpful for moderate amounts of time variation). As expected, knowledge of the linear trend form of the parameter path leads to lower risks compared to those obtained from "un c, LL" and "un c, Kal", especially for c large.

Table 3 reports relative weighted average risk of the same inference methods under the weighting function considered in Table 1, $\theta_0 \sim \mathcal{N}(0, 100)$ and $\theta_t - \theta_{t-1} \sim i.i.d.\mathcal{N}(0, c^2/HT^2)$.

Table 2: Weighted Average Risk relative to trend MLE, Gaussian Linear Trend Weight Function

	$c=0$			$c=4$			$c=8$			$c=12$		
df	∞	12	6	∞	12	6	∞	12	6	∞	12	6
						Average Square Loss, $T=160$						
kn c, trnd.LL	0.49	0.51	0.54	0.81	0.78	0.74	1.07	1.02	0.95	1.65	1.47	1.25
kn c, trnd.Kal	0.49	0.51	0.54	0.80	0.77	0.74	0.95	0.93	0.89	1.03	1.01	0.98
un c, trnd.LL	0.70	0.72	0.74	0.86	0.84	0.81	1.10	1.05	0.97	1.67	1.49	1.27
un c, trnd.Kal	0.71	0.72	0.74	0.86	0.84	0.82	0.99	0.97	0.93	1.07	1.06	1.02
un c, LL	0.82	0.83	0.84	1.07	1.03	0.97	1.57	1.44	1.27	2.54	2.20	1.80
un c, Kal	0.83	0.84	0.82	1.05	1.00	0.94	1.32	1.25	1.13	1.55	1.45	1.30
						Average Square Loss, $T=480$						
kn c, trnd.LL	0.50	0.50	0.52	0.81	0.77	0.72	0.98	0.94	0.89	1.19	1.12	1.03
kn c, trnd.Kal	0.50	0.50	0.52	0.81	0.77	0.72	0.94	0.92	0.87	0.99	0.98	0.94
un c, trnd.LL	0.71	0.73	0.73	0.87	0.84	0.81	1.01	0.98	0.92	1.22	1.15	1.05
un c, trnd.Kal	0.72	0.73	0.73	0.87	0.85	0.81	0.99	0.97	0.92	1.04	1.04	0.98
un c, LL	0.86	0.87	0.86	1.09	1.06	0.98	1.45	1.36	1.21	1.94	1.78	1.51
un c, Kal	0.86	0.87	0.84	1.09	1.04	0.95	1.36	1.29	1.14	1.56	1.48	1.30
						Endpoint Interval Estimation Loss, $T=160$						
kn c, trnd.LL	0.49	0.51	0.53	0.83	0.79	0.73	0.96	0.93	0.87	1.14	1.10	0.99
kn c, trnd.Kal	0.49	0.51	0.53	0.83	0.80	0.74	0.95	0.93	0.88	1.01	1.00	0.95
un c, trnd.LL	0.72	0.72	0.70	0.87	0.85	0.80	1.01	1.00	0.93	1.18	1.14	1.05
un c, trnd.Kal	0.73	0.72	0.70	0.88	0.86	0.81	1.02	1.01	0.95	1.09	1.08	1.03
un c, LL	0.92	0.91	0.86	1.05	1.03	0.95	1.28	1.25	1.12	1.60	1.54	1.33
un c, Kal	0.94	0.93	0.88	1.06	1.03	0.95	1.26	1.23	1.12	1.44	1.40	1.26
						Endpoint Interval Estimation Loss, $T=480$						
kn c, trnd.LL	0.50	0.50	0.52	0.85	0.80	0.74	0.95	0.92	0.87	1.01	0.98	0.93
kn c, trnd.Kal	0.50	0.50	0.52	0.85	0.80	0.74	0.95	0.93	0.88	0.99	0.97	0.93
un c, trnd.LL	0.74	0.74	0.73	0.90	0.85	0.81	1.02	0.98	0.92	1.07	1.04	0.99
un c, trnd.Kal	0.75	0.75	0.73	0.91	0.86	0.81	1.03	0.99	0.93	1.06	1.04	1.00
un c, LL	0.96	0.94	0.92	1.09	1.03	0.98	1.29	1.21	1.13	1.46	1.37	1.28
un c, Kal	0.97	0.95	0.93	1.09	1.04	0.99	1.28	1.20	1.13	1.42	1.34	1.26

Table 3: Weighted Average Risk relative to trend MLE, Gaussian Random Walk Weight Function

	$c = 0$			$c = 4$			$c = 8$			$c = 12$		
df	∞	12	6	∞	12	6	∞	12	6	∞	12	6
Average Square Loss, $T = 160$												
kn c, trnd.LL	0.50	0.49	0.53	0.89	0.85	0.80	1.07	1.05	1.01	1.23	1.22	1.18
kn c, trnd.Kal	0.50	0.49	0.53	0.88	0.84	0.80	1.00	0.99	0.96	1.04	1.04	1.03
un c, trnd.LL	0.71	0.70	0.73	0.93	0.89	0.86	1.09	1.07	1.02	1.24	1.23	1.19
un c, trnd.Kal	0.72	0.71	0.74	0.93	0.89	0.86	1.02	1.01	0.98	1.06	1.06	1.05
un c, LL	0.82	0.82	0.83	0.94	0.94	0.91	0.93	0.96	0.97	0.97	1.01	1.03
un c, Kal	0.83	0.82	0.82	0.92	0.91	0.89	0.80	0.84	0.87	0.69	0.74	0.80
Average Square Loss, $T = 480$												
kn c, trnd.LL	0.49	0.50	0.52	0.89	0.85	0.78	1.01	1.00	0.96	1.08	1.08	1.05
kn c, trnd.Kal	0.49	0.50	0.52	0.88	0.84	0.78	0.99	0.97	0.95	1.01	1.00	0.99
un c, trnd.LL	0.71	0.72	0.73	0.93	0.90	0.86	1.03	1.01	0.98	1.09	1.08	1.06
un c, trnd.Kal	0.72	0.73	0.73	0.93	0.90	0.86	1.00	1.00	0.97	1.02	1.02	1.01
un c, LL	0.87	0.87	0.86	0.97	0.97	0.95	0.86	0.91	0.96	0.76	0.83	0.91
un c, Kal	0.88	0.86	0.84	0.96	0.95	0.93	0.80	0.86	0.90	0.63	0.70	0.79
Endpoint Interval Estimation Loss, $T = 160$												
kn c, trnd.LL	0.50	0.50	0.50	0.80	0.75	0.70	0.88	0.85	0.80	0.92	0.90	0.86
kn c, trnd.Kal	0.50	0.50	0.50	0.80	0.75	0.70	0.87	0.84	0.80	0.89	0.86	0.83
un c, trnd.LL	0.71	0.71	0.68	0.87	0.79	0.75	0.93	0.89	0.83	0.94	0.92	0.88
un c, trnd.Kal	0.72	0.71	0.68	0.88	0.80	0.76	0.93	0.89	0.85	0.92	0.89	0.87
un c, LL	0.90	0.90	0.84	0.85	0.83	0.82	0.62	0.64	0.68	0.50	0.53	0.57
un c, Kal	0.92	0.92	0.85	0.84	0.83	0.82	0.58	0.61	0.66	0.40	0.44	0.50
Endpoint Interval Estimation Loss, $T = 480$												
kn c, trnd.LL	0.49	0.49	0.52	0.84	0.80	0.74	0.93	0.90	0.84	0.95	0.93	0.88
kn c, trnd.Kal	0.49	0.49	0.52	0.84	0.80	0.74	0.93	0.90	0.85	0.94	0.92	0.88
un c, trnd.LL	0.74	0.74	0.74	0.92	0.86	0.80	0.99	0.96	0.90	0.98	0.96	0.92
un c, trnd.Kal	0.74	0.74	0.74	0.92	0.87	0.81	1.00	0.97	0.91	0.98	0.97	0.93
un c, LL	0.96	0.94	0.94	0.88	0.89	0.90	0.59	0.67	0.74	0.41	0.48	0.55
un c, Kal	0.98	0.95	0.94	0.88	0.89	0.89	0.57	0.65	0.72	0.37	0.44	0.52

1.2 Single Break Parameter Path

To compare the suggested path inference with the approach in the "single break" literature, we conducted further risk calculations in the model of Section 2.3, but with a focus on parameter paths with a single break at fraction $\rho \in [0,1]$ of the sample. Specifically, we simulated data from (10) in the paper with parameter path given by

$$\theta_t = \theta_0 + \frac{c}{\sqrt{HT}}\mathbf{1}[t < \rho T] \qquad (**)$$

and $\theta_0 = 0$. We compared our suggested methods with methods that estimate the pre and post break parameter via MLE (with sandwich covariance matrix) in two subsamples, where the subsamples are determined by (i) ρ is known; (ii) ρ is estimated by least squares from the model $y_t^2 = \mu_1\mathbf{1}[t < \hat{\rho}T] + \mu_2\mathbf{1}[t > \hat{\rho}T] + e_t$ with $\hat{\rho}$ constrained to $[0.15; 0.85]$ (LS.sqr); (iii) ρ is estimated by least squares from the model $|y_t| = \mu_1\mathbf{1}[t < \hat{\rho}T] + \mu_2\mathbf{1}[t > \hat{\rho}T] + e_t$ with $\hat{\rho}$ constrained to $[0.15; 0.85]$ (LS.abs). Methods of this type have been used to date the great moderation (see Stock and Watson (2002), McConnell and Perez-Quiros (2000)). In addition, we also include the two versions of path inference considered in Table 1 ("unkn c, LL" and "unkn c, Kal" abbreviated for "unknown c, Local Level" and "unknown c, Kalman", respectively), as well as inference based on the assumption that there is no break, i.e. full sample MLE with sandwich covariance matrix (FS MLE).

It becomes apparent from Tables 4-6 that knowledge of the break date is very helpful, but "unkn c, LL" and "unkn c, Kal" compare quite favorably to least-squares break date based inference, at least as long as c is small to moderate.

Table 7 reports relative weighted average risk of the same inference methods under the weighting function considered in Table 1, $\theta_0 \sim \mathcal{N}(0, 100)$ and $\theta_t - \theta_{t-1} \sim i.i.d.\mathcal{N}(0, c^2/HT^2)$.

Table 4: Risk in Model (**) relative to MLE with ρ known, $\rho = 0.5$

	$c = 0$			$c = 4$			$c = 8$			$c = 12$		
df	∞	12	6	∞	12	6	∞	12	6	∞	12	6
Average Square Loss, $T = 160$												
LS.sqr	2.30	2.20	1.97	2.37	2.24	2.00	3.04	2.83	2.40	4.06	3.79	3.14
LS.abs	2.20	2.09	1.89	2.40	2.21	1.91	2.89	2.56	2.10	3.20	2.92	2.37
unkn c, LL	0.82	0.83	0.82	1.67	1.50	1.25	3.20	2.79	2.20	6.15	5.08	3.75
unkn c, Kal	0.84	0.84	0.80	1.62	1.45	1.20	2.73	2.43	1.95	4.29	3.68	2.86
FS MLE	0.49	0.51	0.52	2.52	2.01	1.46	8.84	6.70	4.39	20.1	15.1	9.60
Average Square Loss, $T = 480$												
LS.sqr	2.38	2.32	2.12	2.43	2.32	2.13	2.83	2.68	2.37	3.24	3.13	2.77
LS.abs	2.23	2.15	1.95	2.44	2.24	1.93	2.84	2.51	2.04	3.03	2.66	2.15
unkn c, LL	0.86	0.87	0.85	1.69	1.52	1.26	2.85	2.54	2.04	4.42	3.78	2.92
unkn c, Kal	0.86	0.87	0.83	1.68	1.50	1.23	2.69	2.41	1.94	3.73	3.25	2.57
FS MLE	0.50	0.50	0.51	2.50	1.96	1.38	8.60	6.42	4.03	19.1	14.1	8.57
Endpoint Interval Estimation Loss, $T = 160$												
LS.sqr	2.02	2.12	1.97	1.45	1.52	1.51	1.21	1.25	1.20	1.61	1.73	1.56
LS.abs	2.01	2.14	2.01	1.65	1.72	1.68	1.24	1.27	1.25	1.22	1.23	1.19
unkn c, LL	1.26	1.27	1.16	1.54	1.48	1.29	1.96	1.87	1.60	2.26	2.16	1.88
unkn c, Kal	1.30	1.30	1.18	1.64	1.57	1.38	2.15	2.05	1.77	2.60	2.49	2.16
FS MLE	0.68	0.71	0.71	3.48	2.51	1.47	15.81	12.4	7.70	31.7	26.3	18.4
Endpoint Interval Estimation Loss, $T = 480$												
LS.sqr	2.09	2.06	2.07	1.52	1.59	1.68	1.19	1.20	1.26	1.24	1.26	1.31
LS.abs	1.99	1.96	2.00	1.64	1.64	1.73	1.26	1.25	1.31	1.18	1.14	1.14
unkn c, LL	1.35	1.31	1.26	1.64	1.52	1.39	2.06	1.89	1.69	2.39	2.17	1.95
unkn c, Kal	1.37	1.32	1.28	1.71	1.58	1.45	2.20	2.02	1.81	2.63	2.39	2.16
FS MLE	0.70	0.70	0.71	3.35	2.29	1.37	15.1	11.0	6.58	30.0	23.2	15.8

Table 5: Risk in Model (**) relative to MLE with ρ known, $\rho = 0.25$

	$c = 0$			$c = 4$			$c = 8$			$c = 12$		
df	∞	12	6	∞	12	6	∞	12	6	∞	12	6
Average Square Loss, $T = 160$												
LS.sqr	2.26	2.19	1.99	2.06	2.00	1.86	2.23	2.13	1.91	2.94	2.69	2.24
LS.abs	2.16	2.08	1.91	2.14	2.02	1.80	2.31	2.12	1.78	2.67	2.39	1.94
unkn c, LL	0.81	0.83	0.82	1.58	1.43	1.23	2.96	2.61	2.09	5.22	4.39	3.29
unkn c, Kal	0.82	0.83	0.81	1.55	1.40	1.19	2.64	2.35	1.90	3.86	3.36	2.65
FS MLE	0.49	0.50	0.53	1.97	1.63	1.24	6.67	5.17	3.48	15.3	11.7	7.59
Average Square Loss, $T = 480$												
LS.sqr	2.37	2.30	2.14	2.22	2.18	2.05	2.30	2.20	2.04	2.73	2.50	2.22
LS.abs	2.23	2.14	1.96	2.27	2.12	1.87	2.45	2.18	1.86	2.70	2.37	1.93
unkn c, LL	0.86	0.87	0.85	1.63	1.46	1.24	2.84	2.52	2.02	4.20	3.64	2.86
unkn c, Kal	0.86	0.86	0.84	1.62	1.44	1.20	2.72	2.42	1.93	3.67	3.24	2.60
FS MLE	0.50	0.50	0.52	2.00	1.59	1.18	6.56	4.92	3.18	14.4	10.6	6.63
Endpoint Interval Estimation Loss, $T = 160$												
LS.sqr	2.54	2.58	2.42	1.73	1.84	1.88	1.07	1.15	1.24	1.03	1.02	1.01
LS.abs	2.53	2.61	2.46	1.87	1.99	2.03	1.19	1.28	1.34	1.05	1.04	1.04
unkn c, LL	1.59	1.55	1.42	1.88	1.77	1.56	2.55	2.32	1.96	3.24	2.93	2.45
unkn c, Kal	1.63	1.58	1.44	1.98	1.87	1.65	2.71	2.48	2.12	3.52	3.21	2.71
FS MLE	0.85	0.86	0.87	1.51	1.22	0.93	6.09	4.29	2.48	16.7	12.5	7.71
Endpoint Interval Estimation Loss, $T = 480$												
LS.sqr	2.57	2.56	2.54	1.80	1.95	2.06	1.14	1.26	1.41	1.02	1.04	1.07
LS.abs	2.44	2.44	2.46	1.87	2.00	2.05	1.25	1.30	1.42	1.05	1.08	1.09
unkn c, LL	1.66	1.63	1.55	1.94	1.83	1.68	2.55	2.32	2.03	3.12	2.84	2.46
unkn c, Kal	1.68	1.65	1.57	2.00	1.89	1.73	2.62	2.40	2.12	3.22	2.95	2.58
FS MLE	0.86	0.87	0.87	1.44	1.18	0.93	5.16	3.54	2.04	13.9	9.94	5.66

Table 6: Risk in Model (**) relative to MLE with ρ known, $\rho = 0.75$

	$c = 0$			$c = 4$			$c = 8$			$c = 12$		
df	∞	12	6	∞	12	6	∞	12	6	∞	12	6
					Average Square Loss, $T = 160$							
LS.sqr	2.30	2.16	1.99	2.65	2.44	2.23	3.29	3.00	2.64	4.32	3.90	3.39
LS.abs	2.20	2.05	1.91	2.57	2.32	2.07	3.03	2.67	2.25	3.44	2.99	2.49
unkn c, LL	0.82	0.82	0.82	1.60	1.39	1.20	3.50	2.91	2.27	6.74	5.47	4.09
unkn c, Kal	0.84	0.82	0.81	1.56	1.35	1.15	3.17	2.66	2.10	5.63	4.68	3.60
FS MLE	0.49	0.50	0.53	2.00	1.60	1.24	6.62	4.96	3.40	14.6	10.8	7.10
					Average Square Loss, $T = 480$							
LS.sqr	2.36	2.31	2.14	2.56	2.50	2.29	2.92	2.86	2.54	3.25	3.27	2.98
LS.abs	2.22	2.14	1.97	2.50	2.33	2.04	2.80	2.59	2.14	3.00	2.74	2.26
unkn c, LL	0.85	0.87	0.86	1.60	1.43	1.20	3.04	2.65	2.06	4.89	4.21	3.21
unkn c, Kal	0.86	0.87	0.84	1.58	1.41	1.17	2.83	2.49	1.94	4.07	3.60	2.84
FS MLE	0.50	0.50	0.52	1.99	1.59	1.18	6.50	4.90	3.16	14.1	10.5	6.52
					Endpoint Interval Estimation Loss, $T = 160$							
LS.sqr	1.40	1.40	1.36	1.79	1.67	1.45	1.99	2.02	1.83	2.69	2.83	2.81
LS.abs	1.39	1.41	1.38	1.93	1.80	1.57	1.84	1.79	1.52	1.74	1.71	1.55
unkn c, LL	0.88	0.84	0.80	1.12	0.99	0.88	1.56	1.43	1.24	1.90	1.81	1.66
unkn c, Kal	0.90	0.85	0.81	1.15	1.02	0.91	1.57	1.44	1.26	1.71	1.67	1.59
FS MLE	0.47	0.47	0.49	5.18	3.62	2.13	18.16	14.25	9.95	32.0	26.2	19.5
					Endpoint Interval Estimation Loss, $T = 480$							
LS.sqr	1.46	1.46	1.42	1.80	1.76	1.55	1.69	1.76	1.75	1.62	1.88	2.08
LS.abs	1.39	1.40	1.37	1.92	1.78	1.58	1.70	1.66	1.51	1.47	1.46	1.37
unkn c, LL	0.94	0.93	0.86	1.21	1.10	0.95	1.61	1.53	1.28	1.68	1.67	1.54
unkn c, Kal	0.96	0.94	0.87	1.23	1.13	0.98	1.66	1.57	1.32	1.78	1.74	1.59
FS MLE	0.49	0.50	0.49	5.23	3.64	1.96	18.3	14.3	9.19	32.2	26.0	18.0

Table 7: Weighted Average Risk relative to MLE with $\rho = 0.5$, Gaussian Random Walk Weight Function

	$c = 0$			$c = 4$			$c = 8$			$c = 12$		
df	∞	12	6	∞	12	6	∞	12	6	∞	12	6
				Average Square Loss, $T = 160$								
LS.sqr	2.27	2.19	1.98	1.56	1.61	1.60	1.09	1.15	1.21	0.95	0.99	1.03
LS.abs	2.17	2.09	1.90	1.54	1.57	1.54	1.07	1.11	1.14	0.89	0.92	0.94
unkn c, LL	0.83	0.82	0.81	0.88	0.87	0.86	0.80	0.83	0.86	0.80	0.84	0.87
unkn c, Kal	0.84	0.82	0.80	0.86	0.85	0.84	0.69	0.73	0.77	0.57	0.61	0.68
FS MLE	0.50	0.49	0.52	1.09	0.99	0.87	1.62	1.53	1.37	1.86	1.81	1.69
				Average Square Loss, $T = 480$								
LS.sqr	2.39	2.29	2.16	1.61	1.68	1.75	1.10	1.18	1.29	0.93	0.97	1.06
LS.abs	2.24	2.12	1.98	1.57	1.60	1.60	1.08	1.13	1.19	0.89	0.93	0.98
unkn c, LL	0.88	0.86	0.86	0.90	0.90	0.90	0.74	0.79	0.85	0.63	0.69	0.77
unkn c, Kal	0.89	0.86	0.84	0.89	0.89	0.88	0.69	0.74	0.80	0.52	0.58	0.67
FS MLE	0.50	0.50	0.51	1.09	0.98	0.84	1.59	1.48	1.31	1.80	1.73	1.60
				Endpoint Interval Estimation Loss, $T = 160$								
LS.sqr	2.09	2.03	1.92	1.41	1.47	1.48	0.96	0.98	1.03	0.89	0.91	0.91
LS.abs	2.05	2.05	1.97	1.42	1.46	1.50	0.93	0.96	1.00	0.83	0.85	0.85
unkn c, LL	1.27	1.23	1.14	0.80	0.83	0.87	0.43	0.46	0.52	0.34	0.36	0.39
unkn c, Kal	1.30	1.25	1.16	0.80	0.83	0.87	0.40	0.44	0.50	0.28	0.30	0.34
FS MLE	0.70	0.68	0.68	1.99	1.78	1.50	2.12	2.07	2.00	1.94	1.94	1.93
				Endpoint Interval Estimation Loss, $T = 480$								
LS.sqr	2.10	2.08	2.05	1.39	1.52	1.58	0.89	0.96	1.05	0.80	0.83	0.88
LS.abs	2.02	1.99	1.97	1.38	1.45	1.53	0.88	0.93	0.99	0.78	0.79	0.83
unkn c, LL	1.36	1.33	1.27	0.81	0.89	0.97	0.39	0.48	0.55	0.27	0.32	0.37
unkn c, Kal	1.39	1.34	1.28	0.81	0.89	0.96	0.38	0.46	0.54	0.24	0.29	0.35
FS MLE	0.70	0.70	0.70	2.02	1.80	1.53	2.11	2.14	2.07	1.92	1.99	2.00

2 Proofs of Additional Lemmas

Lemma 3 *Under Condition 1:*

(i) $T^{-1/2}\sum_{t=1}^{[\cdot T]} s_t(\theta_0) \Rightarrow \int_0^{\cdot} \Gamma^{1/2}(l)dW(l)$, *where* W *is a* $k\times 1$ *standard Wiener process*

(ii) $\sup_{t\leq T,\{v_t\}_{t=1}^T\in\mathcal{B}_T^T,\{\tilde{v}_t\}_{t=1}^T\in\mathcal{B}_T^T} T^{-1}||\sum_{s=1}^t(2\int_0^1 \lambda h_s(\theta_0+v_s+\lambda\tilde{v}_s)d\lambda-\Gamma_s)|| \xrightarrow{p} 0$ *and* $\sup_{t\leq T,\{v_t\}_{t=1}^T\in\mathcal{B}_T^T,\{\tilde{v}_t\}_{t=1}^T\in\mathcal{B}_T^T} T^{-1}||\sum_{s=1}^t(\int_0^1 h_s(\theta_0+\lambda(v_s-\tilde{v}_s))d\lambda-\Gamma_s)|| \xrightarrow{p} 0$, *where* $B_T=\{\theta: ||\theta-\theta_0||<b_T\}$ *with* $b_T\to 0$, *and* $B_T^T = B_T\times\cdots\times B_T$

(iii) $\hat{u}=T^{1/2}(\hat{\theta}-\theta_0)=O_p(1)$

(iv) $T^{-1/2}\sum_{t=1}^{[\cdot T]} s_t(\hat{\theta}) \Rightarrow \int_0^{\cdot}\Gamma(l)^{1/2}dW(l)-\int_0^{\cdot}\Gamma(l)dl(\int_0^1\Gamma(l)dl)^{-1}\int_0^1\Gamma(l)^{1/2}dW(l)$

(v) $\sup_{\lambda\in[0,1]}||T^{-1}\sum_{t=1}^{[\lambda T]} s_t(\hat{\theta})s_t(\hat{\theta})'-\int_0^{\lambda}\Gamma(l)dl|| \xrightarrow{p} 0$ *and* $T^{-1}\sum s_t(\theta_0)s_t(\theta_0)'=O_p(1)$

(vi) $\sup_{\lambda\in[0,1]}||T^{-1}\sum_{t=1}^{[\lambda T]} h_t(\hat{\theta})-\int_0^{\lambda}\Gamma(l)dl|| \xrightarrow{p} 0.$

Proof. (i) Fix any $k\times 1$ vector v with $v'v=1$, and let $\eta_t=v's_t(\theta_0)$. Then $\{\eta_t,\mathfrak{F}_t\}$ is a martingale difference array and $T^{-1}\sum_{t=1}^T E[|\eta_t|^{2+\varepsilon}|\mathfrak{F}_{t-1}] \leq T^{-1}\sum_{t=1}^T E[||s_t(\theta_0)||^{2+\varepsilon}|\mathfrak{F}_{t-1}] = O_p(1)$ by Condition 1 (MDA). Let $\omega_\eta^2=\int_0^1 v'\Gamma(l)vdl$ and $g(\lambda)=\int_0^{\lambda} v'\Gamma(l)vdl/\omega_\eta^2$, which is a continuous and strictly increasing function on the unit interval, so that it has an inverse g^{-1}. By Corollary 3.8 of McLeish (1974), $T^{-1/2}\sum_{t=1}^{[g^{-1}(\cdot)T]}\eta_t \Rightarrow \omega_\eta W_\eta(\cdot)$, where W_η is a standard scalar Wiener process and the convergence is on the space of cadlag functions on the unit interval, equipped with the Skorohod norm. By the continuous mapping theorem, we hence obtain $T^{-1/2}\sum_{t=1}^{[\cdot T]}\eta_t \Rightarrow \omega_\eta W_\eta(g(\cdot)) \sim v'\int_0^{\cdot}\Gamma(l)^{1/2}dW(l)$ and the result follows from the Functional Cramer-Wold device (see, for instance, Theorem 29.16 of Davidson (1994)).

(ii) We have

$$T^{-1}||\sum_{s=1}^t(2\int_0^1 \lambda h_s(\theta_0+v_s+\lambda\tilde{v}_s)d\lambda-\Gamma_s)||$$
$$\leq T^{-1}||\sum_{s=1}^t(2\int_0^1 \lambda h_s(\theta_0+v_s+\lambda\tilde{v}_s)d\lambda-h_s(\theta_0))||+T^{-1}||\sum_{s=1}^t(\Gamma_s-h_s(\theta_0))||.$$

Now $\sup_{t\leq T}T^{-1}||\sum_{s=1}^t(\Gamma_s-h_s(\theta_0))|| \xrightarrow{p} 0$ by Condition 1 (LLLN) and $\sup_{\lambda\in[0,1]}||T^{-1}\sum_{s=1}^{[\lambda T]}\Gamma_s-\int_0^{\lambda}\Gamma(s)ds||\to 0$, and

$$\sup_{t\leq T,\{v_t\}_{t=1}^T\in\mathcal{B}_T^T,\{\tilde{v}_t\}_{t=1}^T\in\mathcal{B}_T^T} T^{-1}||\sum_{s=1}^t 2\int_0^1 \lambda(h_s(\theta_0+v_s+\lambda\tilde{v}_s)-h_s(\theta_0))d\lambda||$$
$$\leq 2T^{-1}\sum_{t=1}^T \sup_{v\in\mathcal{B}_T}||h_t(\theta_0+2v)-h_t(\theta_0))|| \xrightarrow{p} 0$$

by Condition 1 (LLLN). The second claim follows similarly.

(iii) For any $\varepsilon > 0$,

$$\begin{aligned} P(||\hat{\theta} - \theta_0|| &\geq \varepsilon) \leq P(\sup_{||\theta-\theta_0||\geq\varepsilon} T^{-1}\sum[l_t(\theta) - l_t(\theta_0)] \geq -K(\varepsilon)) \\ &\leq 1 - P(\sup_{||\theta-\theta_0||\geq\varepsilon} T^{-1}\sum \sup_{||v||<T^{-1/2+\eta}} [l_t(\theta+v) - l_t(\theta_0)] < -K(\varepsilon)) \to 0 \end{aligned}$$

by Condition 1 (ID) and so $\hat{\theta} \xrightarrow{p} \theta_0$.

Further, as $\hat{\theta} \xrightarrow{p} \theta_0$, there exists a sequence of decreasing balls $\mathcal{T}_T$ around θ_0 such that $P(\hat{\theta} \in \mathcal{T}_T) \to 1$. For $v \in \Theta_0$, we have by the fundamental theorem of calculus applied row by row that $s_t(\theta_0 + v) - s_t(\theta_0) = \left(-\int_0^1 h_t(\theta_0 + \lambda v)d\lambda\right) v$ almost surely for $t = 1, \cdots, T$. Let T be large enough so that $\mathcal{T}_T \subset \Theta_0$, and define $h_t^S = \int_0^1 h_t(\theta_0 + \lambda(\hat{\theta} - \theta_0))d\lambda$ if $\hat{\theta} \in \mathcal{T}_T$, and $h_t^S = \tilde{h}_t$ otherwise, so that from the first order condition $\mathbf{1}[\hat{\theta} \in \mathcal{T}_T]\sum s_t(\hat{\theta}) = 0$, we obtain

$$\mathbf{1}[\hat{\theta} \in \mathcal{T}_T]\left(T^{-1/2}\sum s_t(\theta_0) - \left(T^{-1}\sum h_t^S\right)T^{1/2}(\hat{\theta} - \theta_0)\right) = 0 \tag{1}$$

almost surely for $t = 1, \cdots, T$. From part (i), $T^{-1/2}\sum s_t(\theta_0) = O_p(1)$. Applying the result of part (ii), we obtain $T^{-1}\sum h_t^S - T^{-1}\sum \Gamma_t \xrightarrow{p} 0$. But $T^{-1}\sum \Gamma_t \to \int_0^1 \Gamma(l)dl$, which is positive definite, so the result follows from (1) and $P(\hat{\theta} \in \mathcal{T}_T) \to 1$.

(iv) Proceed as in the proof of part (iii) to obtain

$$\mathbf{1}[\hat{\theta} \in \mathcal{T}_T]\left(T^{-1/2}\sum_{s=1}^{t} s_s(\hat{\theta}) - T^{-1/2}\sum_{s=1}^{t} s_s(\theta_0) + \left(T^{-1}\sum_{s=1}^{t} h_s^S\right)T^{1/2}(\hat{\theta} - \theta_0)\right) = 0$$

almost surely, so that

$$\sup_{t\leq T}||T^{-1/2}\sum_{s=t}^{T} s_s(\hat{\theta})|| \leq \sup_{t\leq T}||T^{-1/2}\sum_{s=t}^{T} s_s(\theta_0)|| + T^{1/2}\sup_{t\leq T}||T^{-1}\sum_{s=t}^{T} h_s^S|| \cdot ||\hat{\theta} - \theta_0|| + o_p(1)$$

and the result follows from parts (i), (ii) and (iii) of this Lemma and the CMT.

(v) From the proof of part (iii), $\mathbf{1}[\hat{\theta} \in \mathcal{T}_T](s_t(\hat{\theta}) - s_t(\theta_0) + h_s^S(\hat{\theta} - \theta_0)) = 0$, almost surely for $t = 1, \cdots, T$, so that

$$\begin{aligned} &\sup_{\lambda\in[0,1]}||T^{-1}\sum_{t=1}^{[\lambda T]} s_t(\hat{\theta})s_t(\hat{\theta})' - T^{-1}\sum_{t=1}^{[\lambda T]} s_t(\theta_0)s_t(\theta_0)'|| \\ &\qquad \leq 2||\hat{u}||T^{-1}\sum||h_t^S||T^{-1/2}\sup_{t\leq T}||s_t(\theta_0)|| + T^{-1}||\hat{u}||^2\sup_{t\leq T}||h_t^S||T^{-1}\sum||h_t^S|| \end{aligned}$$

with probability $P(\hat{\theta} \in \mathcal{T}_T) \to 1$. Now $||\hat{u}|| = O_p(1)$ by part (iii), and $T^{-1}\sum||h_t^S|| = O_p(1)$ by a calculation similar to the proof of part (ii) and Condition 1 (LLLN), and $T^{-1/2}\sum_{t=1}^{[\cdot T]} s_t(\theta_0) \Rightarrow \int_0^{\cdot}\Gamma^{1/2}(l)dW(l)$ implies $T^{-1/2}\sup_{t\leq T}||s_t(\theta_0)|| \xrightarrow{p} 0$, and also

$T^{-1}\sup_{t\leq T}||h_t^S|| \leq T^{-1}\sup_{t\leq T}||h_t(\theta_0)|| + T^{-1}\sup_{t\leq T,\theta\in\mathcal{T}_T}||h_t(\theta_0) - h_t(\theta)|| \xrightarrow{p} 0$ by Condition 1 (LLLN) (i) and (iii), so that $\sup_{\lambda\in[0,1]}||T^{-1}\sum_{t=1}^{[\lambda T]} s_t(\hat\theta)s_t(\hat\theta)' - T^{-1}\sum_{t=1}^{[\lambda T]} s_t(\theta_0)s_t(\theta_0)'|| \xrightarrow{p} 0$.

Let $v \in \mathbb{R}^k$, and define $\eta_t = v's_t(\theta_0)$. Then from Condition 1 (MDA), $\{\eta_t, \mathfrak{F}_t\}$ is a martingale difference array with conditional variance process $V_{\eta,t}^2 = v'E[s_t(\theta_0)s_t(\theta_0)'|\mathfrak{F}_{t-1}]v$, and $\sup_{\lambda\in[0,1]}|T^{-1}\sum_{t=1}^{[\lambda T]} V_{\eta,t}^2 - v'(\int_0^\lambda \Gamma(l)dl)v| \xrightarrow{p} 0$. Note that $T^{-1}\sum_{t=1}^{T} V_{\eta,t}^2 \leq ||v||^2(T^{-1}\sum_{t=1}^{T} E[||s_t(\theta_0)||^{2+\epsilon}|\mathfrak{F}_{t-1}])^{2/(2+\epsilon)} = O_p(1)$ implies $T^{-1}\sum E[\eta_t^2\mathbf{1}[|\eta_t| > T^{1/2}c]|\mathfrak{F}_{t-1}] \xrightarrow{p} 0$ for all $0 < c < \infty$, so that from Theorem 2.23 of Hall and Heyde (1980), $\sup_{t\leq T}|T^{-1}\sum_{s=1}^{t}(\eta_s^2 - V_{\eta,s}^2)| = \sup_{\lambda\in[0,1]}|T^{-1}v'\sum_{t=1}^{[\lambda T]}(s_t(\theta_0)s_t(\theta_0)' - E[s_t(\theta_0)s_t(\theta_0)'|\mathfrak{F}_{t-1}])v| \xrightarrow{p} 0$, so that also $\sup_{\lambda\in[0,1]}|T^{-1}v'\sum_{t=1}^{[\lambda T]}(s_t(\theta_0)s_t(\theta_0)' - \int_0^\lambda \Gamma(l)dl)v| \xrightarrow{p} 0$. This holds for arbitrary $v \in \mathbb{R}^k$, so in particular jointly for all vectors v_j, $j = 1, \cdots, 2^k$ with elements that are either zero or one. It is easy to see that if $v_j'A_0v_j = v_j'A_1v_j$ for all such v_j, $j = 1, \cdots, 2^k$ for two symmetric matrices A_0 and A_1, then $A_0 = A_1$, and both results follow.

(vi) Follows from parts (ii) and (iii). ■

Lemma 4 *Under Conditions 1 and 2, there exists a sequence of real numbers α_T with $\alpha_T \to \infty$ and $T^{-1/2}\alpha_T \to 0$ such that*

(i) $\int w(\theta_0 + T^{-1/2}u)E_\delta(1 - \mathcal{A}_T(u)\mathcal{S}_T(\boldsymbol{\delta}))LR_T(u, \boldsymbol{\delta})du \xrightarrow{p} 0$

(ii) $E_\delta(1 - \mathcal{S}_T(\boldsymbol{\delta}))LR_T(0, \boldsymbol{\delta} - \mathbf{e}\bar\delta) \xrightarrow{p} 0$

Proof. (i) For any choice of α_T, we have

$$\begin{aligned} & |\int w(\theta_0 + T^{-1/2}u)E_\delta(1 - \mathcal{A}_T(u)\mathcal{S}_T(\boldsymbol{\delta}))LR_T(u, \boldsymbol{\delta})du| \\ \leq & |\int w(\theta_0 + T^{-1/2}u)E_\delta(1 - \mathcal{S}_T(\boldsymbol{\delta}))LR_T(u, \boldsymbol{\delta})du| \\ & \qquad + |\int w(\theta_0 + T^{-1/2}u)E_\delta\mathcal{S}_T(\boldsymbol{\delta})(1 - \mathcal{A}_T(u))LR_T(u, \boldsymbol{\delta})du| \\ = & \rho_1 + \rho_2. \end{aligned}$$

For $\rho_1 = \int w(\theta_0 + T^{-1/2}u)E_\delta(1 - \mathcal{S}_T(\boldsymbol{\delta}))LR_T(u, \boldsymbol{\delta})du$, note that by Markov's inequality, for any $\epsilon > 0$

$$\begin{aligned} P(\rho_1 &> \epsilon) \leq \epsilon^{-1}E\rho_1 \\ &= \epsilon^{-1}\int\int w(\theta_0 + T^{-1/2}u)E_\delta(1 - \mathcal{S}_T(\boldsymbol{\delta}))LR_T(u, \boldsymbol{\delta})du f_T(\theta_0, 0)d\mu_T \\ &= \epsilon^{-1}\int w(\theta_0 + T^{-1/2}u)E_\delta(1 - \mathcal{S}_T(\boldsymbol{\delta}))\int f_T(\theta_0 + T^{-1/2}u, \boldsymbol{\delta})d\mu_T du \end{aligned}$$

where the interchange of the order of integration in the second equality follows from Fubini's Theorem. For any fixed u and $\boldsymbol{\delta}$, if for all $t \leq T$, $\theta_0 + T^{-1/2}u + \delta_t \in \Theta$, then $f_T(\theta_0 + T^{-1/2}u, \boldsymbol{\delta})$ is a probability density with respect to μ_T, and $\int f_T(\theta_0 + T^{-1/2}u, \boldsymbol{\delta})d\mu_T = 1$. If for some t,

$\theta_0 + T^{-1/2}u + \delta_t \notin \Theta$, then $f_T(\theta_0 + T^{-1/2}u, \boldsymbol{\delta}) = 0$, and also $\int f_T(\theta_0 + T^{-1/2}u, \boldsymbol{\delta}) d\mu_T = 0$. Therefore, $\sup_{u\in\mathbb{R}^k, \boldsymbol{\delta}\in\mathbb{R}^{Tk}} \int f_T(\theta_0 + T^{-1/2}u, \boldsymbol{\delta}) d\mu_T \leq 1$, and we obtain

$$\begin{aligned} P(\rho_1 &> \epsilon) \leq \epsilon^{-1} \int w(\theta_0 + T^{-1/2}u) E_\delta(1 - \mathcal{S}_T(\boldsymbol{\delta}))du \\ &= \epsilon^{-1} \int w(\theta_0 + T^{-1/2}u)du \; E_\delta \mathbf{1}[T^{1/2} \sup_{t\leq T} ||\delta_t|| > T^\eta]. \end{aligned}$$

Now by a change of variable $\int w(\theta_0 + T^{-1/2}u)du = T^{k/2} \int w(\theta)d\theta = T^{k/2}$. Furthermore, let $\bar{G} = \sup_{t\leq T, T, i\leq k} |G_{(i)}(t/T)|$, where $G_{(i)}(s)$ is the ith element of $G(s)$. Then $\bar{G} \leq \sup_{i,s\in[0,1]} |G_{(i)}(s)|$, which is bounded with probability one. By Borell's inequality (see, for instance, Pollard (2002), p. 279), this implies that the tail probability of $\bar{G}$ decays exponentially. Therefore, with $\eta > 0$, $T^{k/2} E_\delta \mathbf{1}[T^{1/2} \sup_{t\leq T} ||\delta_t|| > T^\eta] \to 0$. Since ϵ is arbitrary, this implies $\rho_1 \xrightarrow{p} 0$.

For ρ_2, note that for any fixed n, by Condition 1 (ID), there exists $T^*(n)$ such that for all $T > T^*(n)$,

$$P(\sup_{||\theta-\theta_0||\geq n^{-1}} T^{-1} \sum \sup_{||v||<T^{-1/2+\eta}, \theta+v\in\Theta} (l_t(\theta + v) - l_t(\theta_0)) < -K(n^{-1})) \geq 1 - n^{-1}.$$

For any T, let n_T be the largest n such that simultaneously, $T > \sup_{n'\leq n} T^*(n')$, $T^{1/2}K(n^{-1}) > 1$ and $T^{-1/4}n < 1$. Note that $n_T \to \infty$, since for any fixed n, $T^*(n+1)$ and $n+1$ are finite and $K((n+1)^{-1}) > 0$. By construction,

$$P(\sup_{||\theta-\theta_0||\geq n_T^{-1}} T^{-1} \sum \sup_{||v||<T^{-1/2+\eta}, \theta+v\in\Theta} (l_t(\theta + v) - l_t(\theta_0)) < -K(n_T^{-1})) \geq 1 - n_T^{-1}. \quad (2)$$

Now set $\alpha_T = T^{1/2} n_T^{-1} = o(T^{1/2})$. Note that

$$\begin{aligned} \mathcal{S}_T(\boldsymbol{\delta})(1 - \mathcal{A}_T(u))LR_T(u, \boldsymbol{\delta}) &= \mathcal{S}_T(\boldsymbol{\delta})(1 - \mathcal{A}_T(u)) \exp[\sum (l_t(\theta_0 + T^{-1/2}u + \delta_t) - l_t(\theta_0))] \\ &\leq (1 - \mathcal{A}_T(u)) \exp[\sum \sup_{||v||<T^{-1/2+\eta}} (l_t(\theta_0 + T^{-1/2}u + v) - l_t(\theta_0))] \\ &\leq \exp[\sup_{||\theta-\theta_0||\geq n_T^{-1}} \sum \sup_{||v||<T^{-1/2+\eta}} (l_t(\theta + v) - l_t(\theta_0))]. \end{aligned}$$

Hence, with probability of at least $1 - n_T^{-1} \to 1$,

$$\begin{aligned} \rho_2 &\leq \int w(\theta_0 + T^{-1/2}u)du \cdot \exp\left[\sup_{||\theta-\theta_0||\geq n_T^{-1}} \sum \sup_{||v||<T^{-1/2+\eta}} (l_t(\theta + v) - l_t(\theta_0)) \right] \\ &\leq T^{k/2} \exp\left[-TK(n_T^{-1})\right] \leq T^{k/2} \exp\left[-T^{1/2}\right] \to 0 \end{aligned}$$

where the last inequality holds since $T^{1/2}K(n_T^{-1}) > 1$ by construction of n_T.

(ii) Similarly to the reasoning concerning ρ_1 in the proof of part (i), for any $\epsilon > 0$, by Markov's inequality

$$\begin{aligned} P(E_\delta(1-\mathcal{S}_T(\boldsymbol{\delta}))LR_T(0,\boldsymbol{\delta}-\mathbf{e}\bar{\delta}) &> \epsilon) \leq \epsilon^{-1}EE_\delta(1-\mathcal{S}_T(\boldsymbol{\delta}))LR_T(0,\boldsymbol{\delta}-\mathbf{e}\bar{\delta}) \\ &= E_\delta(1-\mathcal{S}_T(\boldsymbol{\delta}))\int f_T(\theta_0,\boldsymbol{\delta}-\mathbf{e}\bar{\delta})d\mu_T \\ &\leq E_\delta(1-\mathcal{S}_T(\boldsymbol{\delta}))\to 0. \end{aligned}$$

■

Lemma 5 *Let $\Sigma_\Xi(u)$ be a $Tk\times Tk$ matrix consisting of $k\times k$ blocks $\Xi_{i,j}(u)$, $i,j=1,\cdots,T$, possibly dependent on u and define $c_T^U=\sup_{i,j\leq T,u\in\mathbb{R}^k}||\Xi_{i,j}(u)||$. Under Condition 2, there exists a constant c_G independent of u and T such that*
(i) $|\operatorname{tr}((F^{-1}\Sigma_\delta F'^{-1})\Sigma_\Xi(u))|\leq c_T^U c_G$
(ii) $|\operatorname{tr}((F^{-1}\Sigma_\delta F'^{-1})\Sigma_\Xi(u)(F^{-1}\Sigma_\delta F'^{-1})\Sigma_\Xi(u))|\leq (c_T^U)^2c_G^2$.

Proof. Note that for $1<i\leq j$, the i,jth $k\times k$ block of $F^{-1}\Sigma_\delta F'^{-1}$ is given by

$$\kappa_G(i/T,j/T)-\kappa_G((i-1)/T,j/T)+\kappa_G((i-1)/T,(j-1)/T)-\kappa_G(i/T,(j-1)/T)$$

by $\kappa_G(1/T,j/T)-\kappa_G(1/T,(j-1)/T)$ for $i=1<j$, and by $\kappa_G(1/T,1/T)$ for $i=j=1$.

If $i=j$ and $((i-1)/T,i/T]\cap\tau=\varnothing$, due to the symmetry of κ_G and by the Fundamental Theorem of Calculus

$$\begin{aligned} ||\kappa_G(i/T,i/T)-\kappa_G((i-1)/T,i/T)|| &\leq T^{-1}\sup_{r,s\in[0,1]\setminus\tau}||\nabla_1^-\kappa_G(r,s)|| \\ ||\kappa_G(i/T,(i-1)/T)-\kappa_G((i-1)/T,(i-1)/T)|| &\leq T^{-1}\sup_{r,s\in[0,1]\setminus\tau}||\nabla_1^+\kappa_G(r,s)|| \end{aligned}$$

where $\nabla_1^-\kappa_G(r,s)$ and $\nabla_1^+\kappa_G(r,s)$ are the left and right partial derivatives of κ_G with respect to the first argument, so that in this case, the i,ith block has a norm that is bounded by $T^{-1}c_D=T^{-1}(\sup_{r,s\in[0,1]\setminus\tau}||\nabla_1^-\kappa_G(r,s)||+\sup_{r,s\in[0,1]\setminus\tau}||\nabla_1^+\kappa_G(r,s)||)$.

If $((j-1)/T,j/T]\cap\tau\neq\varnothing$ and $((i-1)/T,i/T]\cap\tau\neq\varnothing$, then

$$\begin{aligned} &||\kappa_G(i/T,j/T)-\kappa_G((i-1)/T,j/T)+\kappa_G((i-1)/T,(j-1)/T)-\kappa_G(i/T,(j-1)/T)|| \\ &\qquad\leq 4\sup_{r,s\in[0,1]}||\kappa_G(r,s)||=c_J \end{aligned}$$

which is also a valid bound for $||\kappa_G(1/T,1/T)||$.

If $1<i<j$ and $((j-1)/T,j/T]\cap\tau=((i-1)/T,i/T]\cap\tau=\varnothing$, then by the Fundamental Theorem of Calculus

$$\begin{aligned} &||\kappa_G(i/T,j/T)-\kappa_G((i-1)/T,j/T)+\kappa_G((i-1)/T,(j-1)/T)-\kappa_G(i/T,(j-1)/T)|| \\ &\qquad\leq T^{-2}\sup_{r\neq s,r,s\in[0,1]\setminus\tau}||\frac{\partial^2\kappa_G(r,s)}{\partial r\partial s}||=T^{-2}c_O. \end{aligned}$$

Also, if $1 < i < j$ and $((j-1)/T, j/T] \cap \tau \neq \varnothing$, and $((i-1)/T, i/T] \cap \tau = \varnothing$, then by the Fundamental Theorem of Calculus

$$\begin{aligned} ||\kappa_G(i/T, j/T) - \kappa_G((i-1)/T, j/T)|| &\leq T^{-1} \sup_{r\in[0,1]\backslash\tau} ||\frac{\partial \kappa_G(r,s)}{\partial r}|| \\ ||\kappa_G(i/T, (j-1)/T) - \kappa_G((i-1)/T, (j-1)/T)|| &\leq T^{-1} \sup_{r\in[0,1]\backslash\tau} ||\frac{\partial \kappa_G(r,s)}{\partial r}|| \end{aligned}$$

so that the norm of the i,jth block is bounded by $T^{-1}c_C = 2T^{-1}\sup_{r\in[0,1]\backslash\tau}||\frac{\partial\kappa_G(r,s)}{\partial r}||$, which is also a valid bound for $||\kappa_G(1/T, j/T) - \kappa_G(1/T, (j-1)/T)||$.

We can hence decompose

$$TF^{-1}\Sigma_\delta F'^{-1} = \Sigma_D + \Sigma_O + \Sigma_C + \Sigma_J$$

where Σ_D is a block diagonal matrix whose i,ith $k \times k$ block has a norm that is bounded by c_D ("the variance of the increments of the continuous part of $\boldsymbol{\delta}$"), Σ_O is a $Tk \times Tk$ matrix whose i,jth block has a norm that is bounded by $T^{-1}c_O$ ("the covariance of the increments of the continuous part of $\boldsymbol{\delta}$"), $\Sigma_C = \sum_{l=1}^{q} \Sigma_{C,l}$ with $\Sigma_{C,i}$ $Tk \times Tk$ matrices whose only nonzero $k \times k$ blocks are in one (block) row and column and correspond to the jump at time τ_i, and these nonzero blocks have a norm that is bounded by c_C ("the covariance between the jumps and the increments of $\boldsymbol{\delta}$") and Σ_J with q^2 nonzero $k \times k$ blocks whose norm is bounded by $c_J T$ ("the variance of the jumps"), and all these bounds are uniform in i,j and T.

Let A and B be $Tk \times Tk$ matrices with i,jth $k \times k$ block $[A]_{i,j}$ and $[B]_{i,j}$, respectively. Note that the i,jth $k \times k$ block of AB, $[AB]_{i,j}$ satisfies

$$||[AB]_{i,j}|| = ||\sum_{l=1}^{T} A_{i,l}B_{l,j}|| \leq \sum_{l=1}^{T} ||A_{i,l}|| \cdot ||B_{l,j}|| \leq [\bar{A}\bar{B}]_{i,j}$$

where for any $Tk \times Tk$ matrix C with i,jth $k \times k$ block $[C]_{i,j}$, $\bar{C}$ denotes a $T \times T$ matrix whose i,jth element $[\bar{C}]_{i,j}$ is at least as large as $\mathbf{1}[||[C]_{i,j}|| > 0] \sup_{s\leq T, t\leq T} ||[C]_{s,t}||$. Also

$$\begin{aligned} ||[ABCB]_{i,j}|| &= ||\sum_{l=1}^{T}[AB]_{i,l}[CB]_{l,j}|| \leq \sum_{l=1}^{T} ||[AB]_{i,l}|| \cdot ||[CB]_{l,j}|| \\ &\leq \sum_{l=1}^{T}[\bar{A}\bar{B}]_{i,l}[\bar{C}\bar{B}]_{l,j} = [(\bar{A}\bar{B})(\bar{C}\bar{B})]_{i,j}. \end{aligned}$$

Hence, using $|\operatorname{tr}[AB]_{i,i}| \leq k||[AB]_{i,i}||$, we obtain

$$|\operatorname{tr} AB| \leq k \operatorname{tr} \bar{A}\bar{B} \quad \text{and} \quad |\operatorname{tr} ABCB| \leq k \operatorname{tr} \bar{A}\bar{B}\bar{C}\bar{B}.$$

Note that we can choose $\bar{\Sigma}_O = T^{-1}c_O\mathbf{e}_0\mathbf{e}_0'$, $\bar{\Sigma}_D = c_D I_T$, $\bar{\Sigma}_J = Tc_J\boldsymbol{\iota}_\tau\boldsymbol{\iota}_\tau'$, $\bar{\Sigma}_C = c_C(\boldsymbol{\iota}_\tau\mathbf{e}_0' + \mathbf{e}_0\boldsymbol{\iota}_\tau')$ and $\bar{\Sigma}_\Xi(u) = c_U\mathbf{e}_0\mathbf{e}_0'$ where $\boldsymbol{\iota}_\tau$ is a $T\times 1$ vector with elements $[\boldsymbol{\iota}_\tau]_j = \mathbf{1}[((j-1)/T, j/T]\cap\tau \neq \varnothing]$ and $\mathbf{e}_0$ is a $T\times 1$ vector of ones.

(i) We compute

$$\begin{aligned}
|\operatorname{tr}(F^{-1}\Sigma_\delta F'^{-1})\Sigma_\Xi(u)| &= T^{-1}|\operatorname{tr}(\Sigma_D+\Sigma_O+\Sigma_C+\Sigma_J)\Sigma_\Xi(u)| \\
&\leq kT^{-1}\operatorname{tr}(\bar{\Sigma}_D+\bar{\Sigma}_O+\bar{\Sigma}_C+\bar{\Sigma}_J)\bar{\Sigma}_\Xi(u) \\
&= kc_T^U T^{-1}\operatorname{tr}(c_D I_T + T^{-1}c_O\mathbf{e}_0\mathbf{e}_0' + Tc_J\boldsymbol{\iota}_\tau\boldsymbol{\iota}_\tau' + c_C(\boldsymbol{\iota}_\tau\mathbf{e}_0' + \mathbf{e}_0'\boldsymbol{\iota}_\tau))\mathbf{e}_0\mathbf{e}_0' \\
&= kc_T^U(c_D + c_O + c_J q^2 + 2c_C q).
\end{aligned}$$

(ii) We compute

$$\begin{aligned}
& |\operatorname{tr}(F^{-1}\Sigma_\delta F'^{-1})\Sigma_\Xi(u)(F^{-1}\Sigma_\delta F'^{-1})\Sigma_\Xi(u)| \\
= {} & T^{-2}\operatorname{tr}(\Sigma_D+\Sigma_O+\Sigma_C+\Sigma_J)\Sigma_\Xi(u)(\Sigma_D+\Sigma_O+\Sigma_C+\Sigma_J)\Sigma_\Xi(u) \\
\leq {} & T^{-2}k\operatorname{tr}(\bar{\Sigma}_D+\bar{\Sigma}_O+\bar{\Sigma}_C+\bar{\Sigma}_J)\bar{\Sigma}_\Xi(u)(\bar{\Sigma}_D+\bar{\Sigma}_O+\bar{\Sigma}_C+\bar{\Sigma}_J)\bar{\Sigma}_\Xi(u) \\
= {} & T^{-2}k(c_T^U\mathbf{e}_0'[c_D I_T + T^{-1}c_O\mathbf{e}_0\mathbf{e}_0' + Tc_J\boldsymbol{\iota}_\tau\boldsymbol{\iota}_\tau' + c_C(\boldsymbol{\iota}_\tau\mathbf{e}_0' + \mathbf{e}_0'\boldsymbol{\iota}_\tau)]\mathbf{e}_0)^2.
\end{aligned}$$

■

Lemma 6 *Under Condition 1:*

(i) There exists a sequence of random variables $\tilde{C}_T = O_p(1)$ satisfying $\tilde{C}_T^{-1} = O_p(1)$ such that

$$\sup_{v\in\mathbb{R}^k, T}(E_\delta\exp[-2(\boldsymbol{\delta} - T^{-1/2}\mathbf{e}v)'D_{\tilde{h}}(\boldsymbol{\delta} - T^{-1/2}\mathbf{e}v)] - \exp[-\tfrac{1}{2}\tilde{C}_T||v||^2]) \leq 0.$$

(ii) Suppose the $k\times 1$ vectors ξ_t satisfy $\sup_{t\leq T}||T^{-1/2}\sum_{s=1}^t \xi_s|| \xrightarrow{p} 0$, the $k\times k$ matrix functions $\zeta_t : \mathbb{R}^k \mapsto \mathbb{R}^{k\times k}$ satisfy $\sup_{t\leq T, u\in\mathbb{R}^k}||T^{-1}\sum_{s=1}^t \zeta_s(u)|| \xrightarrow{p} 0$, the $k\times k$ matrices Ξ_{it} satisfy $\sup_{t\leq T}||T^{-1}\sum_{s=1}^t \Xi_{is}|| \xrightarrow{p} 0$, $i = 1,2,3$. Then, with $\boldsymbol{\xi} = (\xi_1', \cdots, \xi_T')'$, $D_\zeta(u) = \operatorname{diag}(\zeta_1(u), \cdots, \zeta_T(u))$, $\Xi_i = (\Xi_{i1}', \cdots, \Xi_{iT}')'$, $i = 1,2$ and $D_\Xi = \operatorname{diag}(\Xi_{31}, \cdots, \Xi_{3T})$

$$\underline{\kappa}_T\exp[\underline{\Delta}_T||v||^2] \leq E_\delta\exp[\boldsymbol{\xi}'\boldsymbol{\delta} + T^{-1/2}v'\mathbf{e}'D_\zeta(u)\boldsymbol{\delta} - \tfrac{1}{2}\boldsymbol{\delta}'(T^{-1}\Xi_1\Xi_2' + D_\Xi)\boldsymbol{\delta}] \leq \overline{\kappa}_T\exp[\bar{\Delta}_T||v||^2]$$

uniformly in v and T, where the scalar random variables $\underline{\kappa}_T$, $\underline{\Delta}_T$, $\overline{\kappa}_T$ and $\bar{\Delta}_T$ do not depend on u or v and $\underline{\kappa}_T \xrightarrow{p} 1$, $\overline{\kappa}_T \xrightarrow{p} 1$, $\underline{\Delta}_T \xrightarrow{p} 0$ and $\bar{\Delta}_T \xrightarrow{p} 0$.

(iii) If $T^{-1/2}\sum_{t=1}^{[\cdot T]}\hat{s}_t \Rightarrow S_L(\cdot)$, then $E_\delta\exp[4\hat{\mathbf{s}}'\boldsymbol{\delta}] = O_p(1)$.

(iv) If $J_T \in \mathcal{D}$ is a nonstochastic sequence converging to $J \in \mathcal{D}$, where $\mathcal{D}$ is the set of cadlag functions on the unit interval, then

$$\sup_T E_\delta\exp[T^{1/2}J_T(1)'(\delta_T - \bar{\delta}) - T^{1/2}\sum J_T((t-1)/T)'(\delta_t - \delta_{t-1})] < \infty$$

with $\delta_0 = 0$.

Proof. (i) A direct calculation yields

$$\frac{E_\delta \exp[-2(\boldsymbol{\delta} - T^{-1/2}\mathbf{e}v)' D_{\tilde{h}}(\boldsymbol{\delta} - T^{-1/2}\mathbf{e}v)]}{E_\delta \exp[-2\boldsymbol{\delta}' D_{\tilde{h}} \boldsymbol{\delta}]} = \exp[-8T^{-1}v'\mathbf{e}'(D_{\tilde{h}}^{-1} + 4\Sigma_\delta)^{-1}\mathbf{e}v].$$

We have $\mathbf{e}'(D_{\tilde{h}}^{-1} + 4\Sigma_\delta)^{-1}\mathbf{e} = \mathbf{e}'D_{\tilde{h}}^{1/2}(I_{Tk} + 4D_{\tilde{h}}^{1/2}\Sigma_\delta D_{\tilde{h}}^{1/2})^{-1}D_{\tilde{h}}^{1/2}\mathbf{e}$, so that

$$\begin{aligned} T^{-1}||\mathbf{e}'(D_{\tilde{h}}^{-1} + 4\Sigma_\delta)^{-1}\mathbf{e}|| &\leq T^{-1}||\mathbf{e}'D_{\tilde{h}}\mathbf{e}|| \\ T^{-1}||\mathbf{e}'(D_{\tilde{h}}^{-1} + 4\Sigma_\delta)^{-1}\mathbf{e}|| &\geq ||(T^{-1}\mathbf{e}'D_{\tilde{h}}\mathbf{e})^{-1}||^{-1}(1 + 4||D_{\tilde{h}}^{1/2}\Sigma_\delta D_{\tilde{h}}^{1/2}||)^{-1} \end{aligned}$$

where $T^{-1}\mathbf{e}'D_{\tilde{h}}\mathbf{e} \xrightarrow{p} \int_0^1 \Gamma(l)dl$ and

$$\begin{aligned} ||D_{\tilde{h}}^{1/2}\Sigma_\delta D_{\tilde{h}}^{1/2}|| &\leq \operatorname{tr} D_{\tilde{h}}\Sigma_\delta \\ &\leq k \sup_{s\in[0,1]} ||\kappa_G(s,s)||T^{-1} \operatorname{tr} D_{\tilde{h}} \xrightarrow{p} k \sup_{s\in[0,1]} ||\kappa_G(s,s)|| \operatorname{tr} \textstyle\int_0^1 \Gamma(l)dl \end{aligned}$$

and finally $E_\delta \exp[-2\boldsymbol{\delta}' D_{\tilde{h}}\boldsymbol{\delta}] \leq 1$ a.s.

(ii) Let $\mathbf{U} = T^{-1/2}F'D_\zeta(u)'\mathbf{e}v$ and $\Sigma_\varepsilon = T^{-1}\Xi_1\Xi_2' + D_\Xi$. We first show the result for the upper bound. By the Cauchy-Schwarz inequality

$$\begin{aligned} E_\delta \exp[\boldsymbol{\xi}'\boldsymbol{\delta} + \boldsymbol{\delta}'F'^{-1}\mathbf{U} - \tfrac{1}{2}\boldsymbol{\delta}'\Sigma_\varepsilon\boldsymbol{\delta}] &\leq (E_\delta \exp[4\boldsymbol{\delta}'\boldsymbol{\xi}])^{1/4}(E_\delta \exp[4\boldsymbol{\delta}'F'^{-1}\mathbf{U}])^{1/4}(E_\delta \exp[-\boldsymbol{\delta}'\Sigma_\varepsilon\boldsymbol{\delta}])^{1/2} \\ &= \exp[2\boldsymbol{\xi}'\Sigma_\delta\boldsymbol{\xi} + 2\mathbf{U}'F^{-1}\Sigma_\delta F'^{-1}\mathbf{U}](E_\delta \exp[-\boldsymbol{\delta}'\Sigma_\varepsilon\boldsymbol{\delta}])^{1/2}. \end{aligned}$$

Now

$$\begin{aligned} \mathbf{U}'F^{-1}\Sigma_\delta F'^{-1}\mathbf{U} &= T^{-1}v'\mathbf{e}'D_\zeta(u)FF^{-1}\Sigma_\delta F'^{-1}F'D_\zeta(u)'\mathbf{e}v \\ &\leq ||v||^2 \operatorname{tr} T^{-1}F'D_\zeta(u)'\mathbf{e}\mathbf{e}'D_\zeta(u)FF^{-1}\Sigma_\delta F'^{-1}. \end{aligned}$$

But the norm of the i,jth $k \times k$ block of $T^{-1}F'D_\zeta(u)'\mathbf{e}\mathbf{e}'D_\zeta(u)F$ is bounded by $(\sup_{t\leq T, u\in\mathbb{R}^k} ||T^{-1}\sum_{s=t}^T \zeta_s(u)||)^2 \xrightarrow{p} 0$. Hence, by Lemma 5 (i), $\bar{\Delta}_T = \sup_{u\in\mathbb{R}^k} \operatorname{tr}(T^{-1}F'D_\zeta(u)'\mathbf{e}\mathbf{e}'D_\zeta(u)F)F^{-1}\Sigma_\delta F'^{-1} \xrightarrow{p} 0$, and $\mathbf{U}'F^{-1}\Sigma_\delta F'^{-1}\mathbf{U} \leq \bar{\Delta}_T||v||^2$. Similarly, also $\boldsymbol{\xi}'\Sigma_\delta\boldsymbol{\xi} = \operatorname{tr} F'\boldsymbol{\xi}\boldsymbol{\xi}'FF^{-1}\Sigma_\delta F'^{-1} \xrightarrow{p} 0$.

For each T, let A_G be a $Tk \times \ell$ full column rank matrix that spans the column null space of $F^{-1}\Sigma_\delta F'^{-1}$, and B_A the $Tk \times (Tk - \ell)$ matrix such that $B_A'B_A = I_{Tk-\ell}$ and $B_AB_A' = M_A = I_{kT} - A_G(A_G'A_G)^{-1}A_G'$ (if $F^{-1}\Sigma_\delta F'^{-1}$ is full rank, define $B_A = I_{Tk}$). Then

$$E_\delta \exp[-\boldsymbol{\delta}'\Sigma_\varepsilon\boldsymbol{\delta}] = E_\delta \exp[-\boldsymbol{\delta}'F'^{-1}B_AB_A'F'\Sigma_\varepsilon FB_AB_A'F^{-1}\boldsymbol{\delta}].$$

Note that the covariance matrix of $B_A'F^{-1}\boldsymbol{\delta}$, $B_A'F^{-1}\Sigma_\delta F^{-1\prime}B_A$ is positive definite, and $M_AF^{-1}\Sigma_\delta F^{-1\prime}M_A = F^{-1}\Sigma_\delta F^{-1\prime}$. Let λ_i, $i = 1, \cdots, kT - \ell$ be the eigenvalues of the symmetric matrix

$$\Sigma_S = (B_A'F^{-1}\Sigma_\delta F^{-1\prime}B_A)^{1/2}B_A'F'\Sigma_\varepsilon FB_A(B_A'F^{-1}\Sigma_\delta F^{-1\prime}B_A)^{1/2}.$$

Now the i,jth block of $F'\Sigma_\varepsilon F$, $i \leq j$ equals $(T^{-1}\sum_{s=i}^T \Xi_{1s})(T^{-1}\sum_{s=j}^T \Xi_{1s})' + T^{-1}\sum_{s=j}^T \Xi_{2s}$, whose norm is $o_p(1)$ uniformly in i,j by assumption, so that by Lemma 5,

$$\sum_{i=1}^{Tk-\ell} \lambda_i = \operatorname{tr} B_A' F^{-1}\Sigma_\delta F^{-1\prime} B_A B_A' F'\Sigma_\varepsilon F B_A = \operatorname{tr} F^{-1}\Sigma_\delta F^{-1\prime} F'\Sigma_\varepsilon F \xrightarrow{p} 0 \tag{3}$$

and also

$$\begin{aligned}\sum_{i=1}^{Tk-\ell} \lambda_i^2 &= \operatorname{tr} B_A' F^{-1}\Sigma_\delta F'^{-1} B_A B_A' F'\Sigma_\varepsilon F B_A B_A' F^{-1}\Sigma_\delta F'^{-1} B_A B_A' F'\Sigma_\varepsilon F B_A \qquad (4)\\ &= \operatorname{tr} F^{-1}\Sigma_\delta F'^{-1} F'\Sigma_\varepsilon F F^{-1}\Sigma_\delta F'^{-1} F'\Sigma_\varepsilon F \xrightarrow{p} 0.\end{aligned}$$

Let $\mathcal{L}_T = \mathbf{1}[\sup_{i \leq kT-\ell} |\lambda_i| \leq 1/4]$, and define $\tilde{\Sigma}_\varepsilon = \mathcal{L}_T \Sigma_\varepsilon$, $\tilde{\Sigma}_S = \mathcal{L}_T \Sigma_S$ and $\tilde{\lambda}_i = \mathcal{L}_T \lambda_i$, $i = 1, \cdots, Tk-\ell$. Note that $E(1-\mathcal{L}_T) \leq P((\sum_{i=1}^{Tk-\ell} \lambda_i^2)^{1/2} > 1/4) \to 0$ by (4), so that it suffices to show $\mathcal{L}_T E_\delta \exp[-\boldsymbol{\delta}'\Sigma_\varepsilon \boldsymbol{\delta}] \xrightarrow{p} 1$. We compute

$$\begin{aligned}\mathcal{L}_T E_\delta \exp[-\boldsymbol{\delta}'\Sigma_\varepsilon \boldsymbol{\delta}] &= \mathcal{L}_T E_\delta \exp[-\boldsymbol{\delta}' F'^{-1} B_A B_A' F' \tilde{\Sigma}_\varepsilon F B_A B_A' F^{-1} \boldsymbol{\delta}]\\ &= \mathcal{L}_T |B_A' F^{-1}\Sigma_\delta F'^{-1} B_A|^{-1/2} |2 B_A' F' \tilde{\Sigma}_\varepsilon F B_A + (B_A' F^{-1}\Sigma_\delta F'^{-1} B_A)^{-1}|^{-1/2}\\ &= \mathcal{L}_T |I_{Tk-\ell} + 2\tilde{\Sigma}_S|^{-1/2}.\end{aligned}$$

Since for $x \in [-1/2, 1/2]$, $x - x^2 \leq \ln(1+x) \leq x$, we find

$$\mathcal{L}_T 2 \sum_{i=1}^{Tk-\ell} (\tilde{\lambda}_i - 2\tilde{\lambda}_i^2) \leq \mathcal{L}_T \sum_{i=1}^{Tk-\ell} \ln(1 + 2\tilde{\lambda}_i) = \mathcal{L}_T \ln |I_{Tk-\ell} + 2\tilde{\Sigma}_S| \leq \mathcal{L}_T 2 \sum_{i=1}^{Tk-\ell} \tilde{\lambda}_i$$

and the result follows from (3) and (4).

For the lower bound, note that by Jensen's inequality,

$$E_\delta \exp[\boldsymbol{\xi}'\boldsymbol{\delta} + T^{-1/2} v' \mathbf{e}' D_\zeta(u) \boldsymbol{\delta} - \tfrac{1}{2}\boldsymbol{\delta}'\Sigma_\varepsilon \boldsymbol{\delta}] \geq (E_\delta \exp[-\boldsymbol{\xi}'\boldsymbol{\delta} - T^{-1/2} v' \mathbf{e}' D_\zeta(u) \boldsymbol{\delta} + \tfrac{1}{2}\boldsymbol{\delta}'\Sigma_\varepsilon \boldsymbol{\delta}])^{-1}$$

and proceeding as for the upper bound yields the result.

(iii) Using the formula for the moment generating function of a multivariate normal, we find

$$E_\delta \exp[4 \textstyle\sum s_t(\hat{\theta})'\delta_t] = \exp[8 \operatorname{tr} F' \hat{\mathbf{s}} \hat{\mathbf{s}}' F (F^{-1}\Sigma_\delta F^{-1\prime})].$$

The i,jth $k \times k$ block of $F'\hat{\mathbf{s}}\hat{\mathbf{s}}'F$ is given by $(T^{-1/2}\sum_{t=i}^T s_t(\hat{\theta}))(T^{-1/2}\sum_{t=j}^T s_t(\hat{\theta}))'$, whose norm is $O_p(1)$ uniformly in i,j by assumption. Hence applying Lemma 5 yields $E_\delta \exp[4\sum s_t(\hat{\theta})'\delta_t] = O_p(1)$.

(iv) Let $\bar{\mathbf{J}}_T = (J_T(0)', J_T(1/T)', \cdots, J_T((T-1)/T)')'$ and $\hat{\mathbf{J}} = \mathbf{e} J_T(1) - \bar{\mathbf{J}}_T - T^{-1/2} F' D_\Gamma \mathbf{e} \hat{\Gamma}^{-1} J_T(1)$. We compute

$$\begin{aligned}& E_\delta \exp[T^{1/2} J_T(1)'(\delta_T - \bar{\delta}) - T^{1/2} \textstyle\sum J_T((t-1)/T)'(\delta_t - \delta_{t-1})]\\ =\; & E_\delta \exp[(\mathbf{e} J_T(1) - \bar{\mathbf{J}}_T - T^{-1/2} F' D_\Gamma \mathbf{e} \hat{\Gamma}^{-1} J_T(1))' F^{-1} \boldsymbol{\delta}]\\ =\; & \exp[\tfrac{1}{2} \operatorname{tr} \hat{\mathbf{J}} \hat{\mathbf{J}}' (F^{-1}\Sigma_\delta F^{-1\prime})].\end{aligned}$$

But the i,jth $k \times k$ block of $\hat{\mathbf{J}}\hat{\mathbf{J}}'$ is given by $[J_T(1)(I_k - (T^{-1}\sum_{s=i}^T \Gamma_s)\hat{\Gamma}^{-1}) - J_T((i-1)/T)][J_T(1)(I_k - (T^{-1}\sum_{s=j}^T \Gamma_s)\hat{\Gamma}^{-1}) - J_T((j-1)/T)]'$, whose norm is $O(1)$ uniformly in i, j, T by assumption, so that the result follows from Lemma 5. ■

In the following lemma, we write $\int G^{*\prime}\Gamma^{1/2}dW$ for $\int_0^1 G^*(s)\Gamma(s)^{1/2}dW(s)$, $\int G^{*\prime}\Gamma G^*$ for $\int_0^1 G^*(s)'\Gamma(s)G^*(s)ds$ and so forth.

Lemma 7 *Under Conditions 1 and 2,*

$$E_\delta \overline{LR}_T(\boldsymbol{\delta}) \Rightarrow E_G \exp[\int G^{*\prime}\Gamma^{1/2}dW - \tfrac{1}{2}\int G^{*\prime}\Gamma G^*]$$

where $G^*(s) = G(s) - (\int \Gamma(\lambda)d\lambda)^{-1}\int \Gamma(\lambda)G(\lambda)d\lambda$.

Proof. By Lemma 2 (ii),

$$E_\delta \overline{LR}_T(\boldsymbol{\delta}) - E_\delta \exp[\boldsymbol{\delta}'\hat{\mathbf{s}} - \tfrac{1}{2}\boldsymbol{\delta}' D_\Gamma(\boldsymbol{\delta} - \mathbf{e}\bar{\delta})] \xrightarrow{p} 0.$$

Note that, by the summation by parts formula, with $\delta_0 = 0$,

$$\sum_{t=1}^T \hat{s}_t'\delta_t = \delta_T' \sum_{t=1}^T \hat{s}_t - \sum_{t=1}^T (\sum_{s=1}^{t-1} \hat{s}_s)'(\delta_t - \delta_{t-1}).$$

Now by Lemma 3 (iii), $S_T(\cdot) = T^{-1/2}\sum_{t=1}^{[\cdot T]} \hat{s}_t(\theta_0) \Rightarrow S(\cdot)$, where the convergence is on the space $\mathcal{D}_{[0,1]}$ of cadlag functions on the unit interval in the Skorohod metric, and $S(1) = 0$. By 11.7.2 of Dudley (2002) there exists a sequence of stochastic processes $\tilde{S}_T \in \mathcal{D}_{[0,1]}$ defined on some probability space $(\tilde{\mathcal{F}}, \tilde{\mathfrak{F}}, \tilde{P})$ and event $\tilde{A} \in \tilde{\mathfrak{F}}$ with $\tilde{P}(\tilde{A}) = 1$, such that $\tilde{S}_T$ has the same distribution as S_T, $\tilde{S}$ has the same distribution as S (and is continuous with $\tilde{S}(0) = \tilde{S}(1) = 0$) and $\tilde{S}_T(\cdot, \tilde{\omega}) \to \tilde{S}(\cdot, \tilde{\omega})$ for all $\tilde{\omega} \in \tilde{A}$. Denote by $(\tilde{\mathcal{F}}_p, \tilde{\mathfrak{F}}_p, \tilde{P}_p)$ the probability space obtained as the product space of $(\tilde{\mathcal{F}}, \tilde{\mathfrak{F}}, \tilde{P})$ and $(\mathcal{F}_G, \mathfrak{F}_G, P_G)$, where G of Condition 2 is a stochastic process defined on $(\mathcal{F}_G, \mathfrak{F}_G, P_G)$ (so that E_δ denotes integration with respect to a measure induced by P_G). By this construction,

$$\begin{aligned}\widetilde{LR}_T(\boldsymbol{\delta}, S_T) &= \exp[T^{1/2}\tilde{S}_T(1)'\delta_T - T^{1/2}\sum \tilde{S}_T((t-1)/T)'(\delta_t - \delta_{t-1}) - T^{1/2}\tilde{S}_T(1)'\bar{\delta} \\ &\quad -\tfrac{1}{2}\sum \delta_t'\Gamma_t\delta_t + \tfrac{1}{2}(T^{-1/2}\sum \Gamma_t\delta_t)'\hat{\Gamma}^{-1}(T^{-1/2}\sum \Gamma_t\delta_t)]\end{aligned}$$

is a random variable defined on $(\tilde{\mathcal{F}}_p, \tilde{\mathfrak{F}}_p, \tilde{P}_p)$, and $E_\delta \widetilde{LR}_T(\boldsymbol{\delta})$ defined on $(\tilde{\mathcal{F}}, \tilde{\mathfrak{F}}, \tilde{P})$ and $E_\delta \exp[(\boldsymbol{\delta} - \mathbf{e}\bar{\delta})' s_0 - \frac{1}{2}\boldsymbol{\delta}' D_\Gamma(\boldsymbol{\delta} - \mathbf{e}\bar{\delta})]$ defined on $(\mathcal{F}, \mathfrak{F}, P)$ have the same distribution for all T (since they are functions of $\tilde{S}_T$ and S_T, respectively). It therefore suffices to find the limiting distribution of $E_\delta \widetilde{LR}_T(\boldsymbol{\delta})$.

With $\bar{\mathbf{S}}_T(\tilde{\omega}) = (\tilde{S}_T(0,\tilde{\omega})', \tilde{S}_T(1/T,\tilde{\omega})', \cdots, \tilde{S}_T((T-1)/T,\tilde{\omega})')'$ and $\bar{\mathbf{S}}$ defined analogously, $T^{1/2}\sum \tilde{S}_T((t-1)/T,\tilde{\omega})'(\delta_t - \delta_{t-1}) = \bar{\mathbf{S}}_T(\tilde{\omega})'F^{-1}\boldsymbol{\delta}$, so that for any $\tilde{\omega} \in \tilde{A}$,

$$E_\delta[(\bar{\mathbf{S}}_T(\tilde{\omega}) - \bar{\mathbf{S}}(\tilde{\omega}))'F^{-1}\boldsymbol{\delta}\boldsymbol{\delta}'F^{-1\prime}(\bar{\mathbf{S}}_T(\tilde{\omega}) - \bar{\mathbf{S}}(\tilde{\omega}))] = \operatorname{tr} F^{-1}\Sigma_\delta F^{-1\prime}(\bar{\mathbf{S}}_T(\tilde{\omega}) - \bar{\mathbf{S}}(\tilde{\omega}))(\bar{\mathbf{S}}_T(\tilde{\omega}) - \bar{\mathbf{S}}(\tilde{\omega}))'.$$

But the i,jth $k \times k$ block of $(\bar{\mathbf{S}}_T(\tilde{\omega}) - \bar{\mathbf{S}}(\tilde{\omega}))(\bar{\mathbf{S}}_T(\tilde{\omega}) - \bar{\mathbf{S}}(\tilde{\omega}))'$ is equal to

$$(\tilde{S}_T((i-1)/T,\tilde{\omega}) - \tilde{S}((i-1)/T,\tilde{\omega}))(\tilde{S}_T((j-1)/T,\tilde{\omega}) - \tilde{S}((j-1)/T,\tilde{\omega}))'$$

whose norm converges to zero uniformly in i and j. Therefore, by Lemma 5, $(\bar{\mathbf{S}}_T(\tilde{\omega}) - \bar{\mathbf{S}}(\tilde{\omega}))'F^{-1}\boldsymbol{\delta} \xrightarrow{p} 0$ in P_G, and hence

$$\exp[T^{1/2}\tilde{S}_T(1,\tilde{\omega})'(\delta_T - \bar{\delta}) - \bar{\mathbf{S}}_T(\tilde{\omega})'F^{-1}\boldsymbol{\delta} - \tfrac{1}{2}\boldsymbol{\delta}'D_\Gamma(\boldsymbol{\delta} - \mathbf{e}\bar{\delta})] - \exp[-\bar{\mathbf{S}}(\tilde{\omega})'F^{-1}\boldsymbol{\delta} - \tfrac{1}{2}\boldsymbol{\delta}'D_\Gamma(\boldsymbol{\delta} - \mathbf{e}\bar{\delta})] \xrightarrow{p} 0$$

in P_G (where we used $\tilde{S}(1) = 0$).

By Theorem 21, p. 64, of Protter (2005), and the CMT,

$$\begin{aligned} & \exp[T^{1/2}\tilde{S}(1,\tilde{\omega})'(\delta_T - \bar{\delta}) - \bar{\mathbf{S}}(\tilde{\omega})'F^{-1}\boldsymbol{\delta} - \tfrac{1}{2}\boldsymbol{\delta}'D_\Gamma(\boldsymbol{\delta} - \mathbf{e}\bar{\delta})] \\ \Rightarrow \ & \exp[-\textstyle\int \tilde{S}(l,\tilde{\omega})'dG(l) - \frac{1}{2}\int G'\Gamma G + \frac{1}{2}(\int \Gamma G)'(\int \Gamma)^{-1}(\int \Gamma G)] \end{aligned} \tag{5}$$

in P_G. Furthermore,

$$E_\delta(\widetilde{LR}_T(\boldsymbol{\delta}, \tilde{S}_T(\cdot,\tilde{\omega})))^2 \le E_\delta \exp[2T^{1/2}\tilde{S}_T(1,\tilde{\omega})'(\delta_T - \bar{\delta}) - 2\bar{\mathbf{S}}_T(\tilde{\omega})'F^{-1}\boldsymbol{\delta}]$$

which is uniformly bounded in T by Lemma 6 (iv), so that for all $\tilde{\omega} \in \tilde{A}$, $\widetilde{LR}_T(\boldsymbol{\delta}, \tilde{S}_T(\cdot,\tilde{\omega}))$ is uniformly integrable on $(\mathcal{F}_G, \mathfrak{F}_G, P_G)$. Hence (5) implies that also

$$\begin{aligned} E_\delta\widetilde{LR}_T(\boldsymbol{\delta}, \tilde{S}_T(\cdot,\tilde{\omega})) &\to E_G \exp[-\textstyle\int \tilde{S}(l,\tilde{\omega})'dG(l) - \frac{1}{2}\int G'\Gamma G + \frac{1}{2}(\int \Gamma G)'(\int \Gamma)^{-1}(\int \Gamma G)] \\ &= E_G \exp[-\textstyle\int \tilde{S}(l,\tilde{\omega})'dG^*(l) - \frac{1}{2}\int G^{*\prime}\Gamma G^*] \end{aligned}$$

and the equality follows from $\tilde{S}(1,\omega) = 0$ and the definition of G^*. But almost sure convergence implies convergence in distribution, so that in $(\tilde{\mathcal{F}}, \tilde{\mathfrak{F}}, \tilde{P})$

$$\begin{aligned} E_\delta\widetilde{LR}_T(\boldsymbol{\delta}, \tilde{S}_T) &\Rightarrow E_G \exp[-\textstyle\int \tilde{S}(l)'dG^*(l) - \frac{1}{2}\int G^{*\prime}\Gamma G^*] \\ &= E_G \exp[\textstyle\int G^*(l)'d\tilde{S}(l) - \frac{1}{2}\int G^{*\prime}\Gamma G^*] \\ &\sim E_G \exp[\textstyle\int G^{*\prime}\Gamma^{1/2}dW - \frac{1}{2}\int G^{*\prime}\Gamma G^*] \end{aligned}$$

where the equality follows from the integration by parts formula on p. 83 of Protter (2005), and the last line uses $\int \Gamma G^* = 0$ almost surely. ■

References

Davidson, J. (1994): *Stochastic Limit Theory*. Oxford University Press, New York.

Dudley, R. M. (2002): *Real Analysis and Probability*. Cambridge University Press, Camebridge, UK.

Hall, P., and C. C. Heyde (1980): *Martingale Limit Theory and its Applications*. Academic Press, New York.

McConnell, M. M., and G. Perez-Quiros (2000): "Output Fluctuations in the United States: What Has Changed Since the Early 1980's," *American Economic Review*, 90, 1464–1476.

McLeish, D. L. (1974): "Dependent Central Limit Theorems and Invariance Principles," *The Annals of Probability*, 2, 620–628.

Pollard, D. (2002): *A User's Guide to Measure Theoretic Probability*. Cambridge University Press, Cambridge, UK.

Protter, P. E. (2005): *Stochastic Integration and Differential Equations*. Springer Verlag, Berlin, 2nd edition edn.

Stock, J. H., and M. W. Watson (2002): "Has the Business Cycle Changed and Why?," in *NBER Macroeconomics Annual 2002*, ed. by M. Gertler, and K. S. Rogoff, pp. 159–218. MIT Press, Cambridge, MA.