November 18, 2003

Detailed Proofs to Accompany

Testing for the Cointegrating Rank of a VAR Process with Level Shifts at Unknown Time

by Helmut Lütkepohl, Pentti Saikkonen and Carsten Trenkler

In the proofs we use the following notational conventions in addition to the notation defined earlier. Right hand side and left hand side will be abbreviated by r.h.s. and l.h.s., respectively. The smallest and largest eigenvalues of a matrix are denoted by $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$, respectively. The complement of a set B is signified by B^c. A general notational convention is that the dependence of quantities on the sample size T is not indicated. The symbol $\Rightarrow$ signifies weak convergence in a product space of $D([\underline{\lambda}, \bar{\lambda}])$ or $D([0, 1])$. The former is relevant for random functions depending on the parameter λ, whereas the latter is used when the weak limit is a Brownian motion. Unless otherwise stated, all limits assume that $T \to \infty$. When obtaining weak convergences in a product space of $D([\underline{\lambda}, \bar{\lambda}])$ we frequently make use of results given in Appendix A.1 of Gregory & Hansen (1996). It is straightforward to check that these results are applicable despite the differences in assumptions.

In the proofs we assume the model described in Sections 2 and 3, where the parameters $\mu_0, \mu_1, \delta \in \mathbf{R}^n$ and the true α, β, Π and Γ_j $(j = 1, \ldots, p-1)$ satisfy the restrictions which ensure that the observed variables are at most $I(1)$ whereas these restrictions are not imposed in the estimation. Moreover, $\tau \in \mathcal{T}$, the white noise covariance matrix Ω is positive definite and the components of ε_t have finite moments of order b, where $b > 4$, as specified in Section 2. For convenience it is assumed that the initial values of the DGP are such that $\beta' x_t$ and Δx_t are stationary.

The true DGP is one specific process from our model class. It is occasionally helpful to be more explicit about its particular parameter values. In these cases they will be indicated with a subscript 'o' (e.g., μ_{0o}, μ_{1o}, τ_o etc.). We begin by proving the first part of Theorem 3.1.

Proof of Theorem 3.1(i)

Instead of the series y_t it will be convenient to use the mean adjusted series

$$x_t = y_t - \mu_{0o} - \mu_{1o}t - \delta_o d_{t\tau_o}, \quad t = 1, 2, \ldots$$

Solving the above equation for y_t and inserting the result into (3.1) yields

$$\Delta x_t = \nu_0^{(0)} + \nu_1^{(0)} t + \delta_1 d_{t\tau} + \underline{\gamma}\,\underline{d}_{t\tau} - \delta_1^{(0)} d_{t\tau_o} - \underline{\gamma}^{(0)} \underline{d}_{t\tau_o} + \Pi x_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta x_{t-j} + \varepsilon_t, \quad t = p+1, p+2, \ldots \qquad (A.1)$$

Here

$$\begin{aligned}
\nu_0^{(0)} &= \nu_0 + \Pi\mu_{0o} - \Psi\mu_{1o} - \Pi\mu_{1o} \\
\nu_1^{(0)} &= \nu_1 + \Pi\mu_{1o} \\
\underline{\gamma} &= [\gamma_0 : \cdots : \gamma_{p-1}] \\
\underline{d}_{t\tau} &= [\Delta d_{t\tau} : \cdots : \Delta d_{t-p+1,\tau}]' \\
\delta_1^{(0)} &= -\Pi\delta_o
\end{aligned}$$

and

$$\underline{\gamma}^{(0)} = [\gamma_0^{(0)} : \cdots : \gamma_{p-1}^{(0)}]$$

with

$$\gamma_j^{(0)} = \begin{cases} \delta_o - \delta_1^{(0)}, & j = 0 \\ -\Gamma_j \delta_o, & j = 1, \ldots, p-1. \end{cases}$$

Note that the true values of $\nu_0^{(0)}$ and $\nu_1^{(0)}$ are zero.

It will also be convenient to use the transformation

$$\Pi x_t = \alpha^{(0)} u_{t-1}^{(0)} + \rho^{(0)} v_{t-1}^{(0)},$$

where $u_{t-1}^{(0)} = \beta_o' x_{t-1}$, $v_{t-1}^{(0)} = \beta_{o\perp}' x_{t-1}$, $\alpha^{(0)} = \alpha\beta'\beta_o(\beta_o'\beta_o)^{-1}$ and $\rho^{(0)} = \alpha\beta'\beta_{o\perp}(\beta_{o\perp}'\beta_{o\perp})^{-1}$. Clearly, the true values of $\alpha^{(0)}$ and $\rho^{(0)}$ are α_o and zero, respectively. With this transformation the preceding error correction form can be expressed as

$$\Delta x_t = \nu_0^{(0)} + \nu_1^{(0)} t + \delta_1 d_{t\tau} + \underline{\gamma}\,\underline{d}_{t\tau} - \delta_1^{(0)} d_{t\tau_o} - \underline{\gamma}^{(0)} \underline{d}_{t\tau_o} + \alpha^{(0)} u_{t-1}^{(0)} + \rho^{(0)} v_{t-1}^{(0)} + \sum_{j=1}^{p-1} \Gamma_j \Delta x_{t-j} + \varepsilon_t,$$

$$t = p+1, p+2, \ldots \qquad (A.2)$$

Denote

$$w_t^{(0)} = \left[1 : \frac{t}{T} : T^{-1/2} v_{t-1}^{(0)\prime} : u_{t-1}^{(0)\prime} : \Delta x_{t-1}' : \cdots : \Delta x_{t-p+1}'\right]'$$

and

$$q_{t\tau} = [d_{t\tau} : \underline{d}_{t\tau}']'.$$

With this notation $(A.2)$ becomes

$$\Delta x_t = \Phi w_t^{(0)} + \Xi q_{t\tau} - \Xi^{(0)} q_{t\tau_o} + \varepsilon_t, \quad t = p+1, p+2, \ldots, \qquad (A.3)$$

where $\Phi = [\nu_0^{(0)} : T\nu_1^{(0)} : T^{1/2}\rho^{(0)} : \alpha^{(0)} : \Gamma_1 : \cdots : \Gamma_{p-1}]$, $\Xi = [\delta_1 : \underline{\gamma}]$ and $\Xi^{(0)} = [\delta_1^{(0)} : \underline{\gamma}^{(0)}]$.

Let $\Theta = [\Phi : \Xi]$ contain the freely varying parameters in $(A.3)$ or $(A.2)$. ($\Xi^{(0)}$ is not a freely varying parameter because it is determined by $\alpha^{(0)}$, $\rho^{(0)}$ and $\Gamma_1, \ldots, \Gamma_{p-1}$.) Set

$$\varepsilon_{t\tau}(\Theta) = \Delta x_t - \Phi w_t^{(0)} - \Xi q_{t\tau} + \Xi^{(0)} q_{t\tau_o}.$$

Then

$$l_T(\Theta, \tau, \Omega) = (T-p)\log\det\Omega + \operatorname{tr}\left(\Omega^{-1}\sum_{t=p+1}^{T}\varepsilon_{t\tau}(\Theta)\varepsilon_{t\tau}(\Theta)'\right)$$

is -2 times the (conditional) Gaussian log-likelihood function of the parameters in $(A.3)$. Minimizing this function yields Gaussian ML estimators of the parameters Θ, τ and Ω. It is not difficult to see that the resulting estimators of Θ and τ can alternatively be obtained by minimizing the concentrated counterpart of $l_T(\Theta, \tau, \Omega)$, that is,

$$l_T^{(c)}(\Theta, \tau) = (T-p)\log\det\left(\sum_{t=p+1}^{T}\varepsilon_{t\tau}(\Theta)\varepsilon_{t\tau}(\Theta)'\right).$$

The definition of $\varepsilon_{t\tau}(\Theta)$ (and the fact that $\Xi^{(0)}$ is not a freely varying parameter) makes it clear that the value of τ that minimizes the function $l_T^{(c)}(\Theta, \tau)$ is identical to $\hat{\tau}$ defined by (3.2). Thus, (asymptotic) properties of $\hat{\tau}$ can be studied by using the Gaussian ML estimator of τ discussed above. Before turning to this issue we note that the above discussion also makes clear that a minimizer of $l_T(\Theta, \tau, \Omega)$ exists (for every T larger than some constant).

The proof of Theorem 3.1 consists of several steps. In the first one we consider a subset of the parameter space of (Θ, Ω) defined by

$$0 < \underline{\omega} \leq \lambda_{\min}(\Omega) \leq \lambda_{\max}(\Omega) \leq \bar{\omega} < \infty \qquad (A.4)$$

and

$$\|\Phi\|^2 + \|\delta_1 - \delta_1^{(0)}\|^2 \le \bar{M} < \infty. \qquad (A.5)$$

Note that here $\bar{M}$ does not depend on T although Φ does. We use B^c to signify the complement of the set B and prove

Lemma A.1. Let $B_1 = B_1(\bar{M}, \underline{\omega}, \bar{\omega})$ be the part of the parameter space of (Θ, τ, Ω) in which conditions $(A.4)$ and $(A.5)$ hold. Then there exist choices of $\bar{M}$, $\underline{\omega}$ and $\bar{\omega}$ such that

$$\inf_{(\Theta,\tau,\Omega)\in B_1^c} l_T(\Theta,\tau,\Omega) - l_T(\Theta_o,\tau_o,\Omega_o) > 0$$

with probability approaching one.

Proof: First note that

$$T^{-1} l_T(\Theta_o, \tau_o, \Omega_o) = \left(1 - \frac{p}{T}\right) \log\det\Omega_o + \mathrm{tr}\left(\Omega_o^{-1} T^{-1} \sum_{t=p+1}^{T} \varepsilon_t \varepsilon_t{}'\right) = O_p(1), \qquad (A.6)$$

where the latter equality is justified by the weak law of large numbers.

Next, since $[T\underline{\lambda}] \le \tau, \tau_o \le [T\bar{\lambda}]$, we find from the definitions that

$$\varepsilon_{t\tau}(\Theta) = \Delta x_t - \Phi w_t^{(0)}, \quad t = p+1, \ldots, [T\underline{\lambda}] - 1,$$

and

$$\varepsilon_{t\tau}(\Theta) = \Delta x_t - \Phi w_t^{(0)} - (\delta_1 - \delta_1^{(0)}), \quad t = [T\bar{\lambda}] + p, \ldots, T.$$

Hence,

$$\begin{aligned} T^{-1} l_T(\Theta,\tau,\Omega) \ \ge\ & \left(1 - \tfrac{p}{T}\right) \log\det\Omega \\ & + \mathrm{tr}\left(\Omega^{-1} T^{-1} \textstyle\sum_{t=p+1}^{[T\underline{\lambda}]-1} [\Delta x_t - \Phi w_t^{(0)}][\Delta x_t - \Phi w_t^{(0)}]'\right) \\ & + \mathrm{tr}\left(\Omega^{-1} T^{-1} \textstyle\sum_{t=[T\underline{\lambda}]+p}^{T} [\Delta x_t - \Phi^{(0)} w_t^{(0)}][\Delta x_t - \Phi^{(0)} w_t^{(0)}]'\right), \end{aligned} \qquad (A.7)$$

where $\Phi^{(0)} = \Phi + [\delta_1 - \delta_1^{(0)} : 0]$. Here

$$\lambda_{\min}\left(T^{-1} \sum_{t=p+1}^{[T\underline{\lambda}]-1} \begin{bmatrix} \Delta x_t \\ w_t^{(0)} \end{bmatrix} \left[\Delta x_t' : w_t^{(0)\prime}\right]'\right) \ge \epsilon_* \qquad (A.8)$$

where $\epsilon_* > 0$ is a suitable real number and the inequality holds with probability approaching one. This fact can be justified in the same way as Lemma A.4 of Saikkonen (2001). A similar result is also obtained by changing the range of summation on the l.h.s. of $(A.8)$ to $t = [T\bar{\lambda}] + p, \ldots, T$. When these two eigenvalue conditions are assumed arguments entirely similar to those in Saikkonen (2001, pp. 320-321) show that, with suitable choices of $\bar{M}$, $\underline{\omega}$ and $\bar{\omega}$, the r.h.s of $(A.7)$ can be made arbitrarily large whenever $(\Theta, \tau, \Omega) \notin B_1(\bar{M}, \underline{\omega}, \bar{\omega})$. The assertion of the lemma follows from this and $(A.6)$. □

Lemma A.1 implies that a minimizer of $l_T(\Theta, \tau, \Omega)$ will asymptotically satisfy inequality restrictions of the form $(A.4)$ and $(A.5)$. In what follows, the set B_1 is always assumed to be defined in such a way that the conclusion of Lemma A.1 holds. We shall now proceed in the same way as in Saikkonen (2001) and express the function $l_T(\Theta, \tau, \Omega)$ as a sum of two components. To this end, define

$$w_{1t}^{(0)} = \left[1 : \frac{t}{T} : T^{-1/2} v_{t-1}^{(0)\prime}\right]'$$

and

$$w_{2t}^{(0)} = \left[u_{t-1}^{(0)\prime} : \Delta x_{t-1}' : \cdots : \Delta x_{t-p+1}'\right]'.$$

Then $w_t^{(0)} = [w_{1t}^{(0)\prime} : w_{2t}^{(0)\prime}]'$ and we also partition the parameter matrix Φ conformably as $\Phi = [\Phi_1 : \Phi_2]$ where $\Phi_1 = [\nu_0^{(0)} : T\nu_1^{(0)} : T^{1/2}\rho^{(0)}]$ and $\Phi_2 = [\alpha^{(0)} : \Gamma_1 : \cdots : \Gamma_{p-1}]$. With these definitions,

$$\varepsilon_{t\tau}(\Theta) = \varepsilon_{1t\tau}(\Theta) + \varepsilon_{2t}(\Phi_2),$$

where

$$\varepsilon_{1t\tau}(\Theta) = -\Phi_1 w_{1t}^{(0)} - \Xi q_{t\tau} + \Xi^{(0)} q_{t\tau_o}$$

and

$$\varepsilon_{2t}(\Phi_2) = \Delta x_t - \Phi_2 w_{2t}^{(0)}.$$

Clearly, $\varepsilon_{1t\tau_o}(\Theta_o) = 0$ and

$$l_T(\Theta, \tau, \Omega) = l_{1T}(\Theta, \tau, \Omega) + l_{2T}(\Phi_2, \Omega),$$

where

$$l_{1T}(\Theta, \tau, \Omega) = \operatorname{tr}\left(\Omega^{-1} \sum_{t=p+1}^{T} \varepsilon_{1t\tau}(\Theta)\varepsilon_{1t\tau}(\Theta)'\right) + 2\operatorname{tr}\left(\Omega^{-1} \sum_{t=p+1}^{T} \varepsilon_{1t\tau}(\Theta)\varepsilon_{2t}(\Phi_2)'\right)$$

and

$$l_{2T}(\Phi_2, \Omega) = (T-p)\log\det\Omega + \mathrm{tr}\left(\Omega^{-1}\sum_{t=p+1}^{T}\varepsilon_{2t}(\Phi_2)\varepsilon_{2t}(\Phi_2)'\right).$$

For $l_{2T}(\Phi_2, \Omega)$ we have the following result.

Lemma A.2.

$$\inf_{(\Phi_2,\Omega)} l_{2T}(\Phi_2, \Omega) - l_{2T}(\Phi_{2o}, \Omega_o) = O_p(1),$$

where the infimum is over unrestricted values of Φ_2 and $\Omega > 0$.

Proof: Because we can treat Δx_t as a zero mean stationary process and because $l_{2T}(\Phi_2, \Omega)$ can be interpreted as -2 times the logarithm of the Gaussian likelihood function associated with the regression model $\Delta x_t = \Phi_2 w_{2t}^{(0)} + \varepsilon_t$, the stated result follows from standard regression theory (cf. Saikkonen (2001, p. 321)). □

Next consider the function $l_{1T}(\Theta, \tau, \Omega)$. Our treatment will be divided into several steps in which the time index t is suitably restricted. This means considering the function $l_{1T}(\Theta, \tau, \Omega)$ with the sample size T replaced by appropriate quantities smaller than T. Most of the subsequent results will explicitly be formulated for $\tau \leq \tau_o$ and only briefly discussed in the case $\tau \geq \tau_o$. Due to the occurrence of impulse dummies the situation is in this respect somewhat more complicated than in previous cases where the break date parameter is not affected by impulse dummies.

In the following results about the function $l_{1T}(\Theta, \tau, \Omega)$, $c_1, c_2, \ldots$ denote positive constants and $a_{1T}, a_{2T}, \ldots$ are nonnegative random variables which depend on the sample size but not on the parameters Θ, τ or Ω. First we prove

Lemma A.3. There exists a constant $c_1 > 0$ such that, with probability approaching one and uniformly in $[T\underline{\lambda}] \leq \tau \leq \tau_o$ and $(\Theta, \tau, \Omega) \in B_1$,

$$l_{1,\tau-1}(\Theta, \tau, \Omega) \geq c_1 \|T^{1/2}\Phi_1\|^2 - a_{1T}\|T^{1/2}\Phi_1\|,$$

where $a_{1T} \geq 0$ and $a_{1T} = O_p(1)$.

Proof: For $t \leq \tau - 1$, $\varepsilon_{1t\tau}(\Theta) = -\Phi_1 w_{1t}^{(0)}$ and, consequently,

$$l_{1,\tau-1}(\Theta, \tau, \Omega) = \text{tr}\left(\Omega^{-1}\Phi_1 \sum_{t=p+1}^{\tau-1} w_{1t}^{(0)} w_{1t}^{(0)\prime} \Phi_1'\right) - 2\text{tr}\left(\Omega^{-1}\Phi_1 \sum_{t=p+1}^{\tau-1} w_{1t}^{(0)} \varepsilon_{2t}(\Phi_2)'\right) \stackrel{def}{=} L_1 + L_2.$$

For L_1 we have

$$L_1 \geq \lambda_{\min}(\Omega^{-1})\text{tr}\left(\Phi_1 \sum_{t=p+1}^{\tau-1} w_{1t}^{(0)} w_{1t}^{(0)\prime} \Phi_1{}'\right) \geq \lambda_{\min}(\Omega^{-1})\lambda_{\min}\left(T^{-1} \sum_{t=p+1}^{\tau-1} w_{1t}^{(0)} w_{1t}^{(0)\prime}\right) \|T^{1/2}\Phi_1\|^2.$$

For $(\Theta, \tau, \Omega) \in B_1$, the first eigenvalue in the last expression is bounded away from zero. That the same holds with probability approaching one and uniformly in $[T\underline{\lambda}] \leq \tau \leq \tau_o$ for the second eigenvalue, can be seen by using an analog of (A.3) of Gregory & Hansen (1996, p. 118). Thus, we have shown that $L_1 \geq c_1\|\Phi_1\|^2$, $c_1 \geq 0$, with probability approaching one.

It remains to show that $L_2 \geq -a_{1T}\|T^{1/2}\Phi_1\|$ with a_{1T} having the properties stated in the lemma. To demonstrate this, notice that

$$\begin{aligned} |L_2| &\leq 2\|\Omega^{-1}\| \left\|\Phi_1 \sum_{t=p+1}^{\tau-1} w_{1t}^{(0)} [\Delta x_t - \Phi_2 w_{2t}^{(0)}]'\right\| \\ &\leq 2\|\Omega^{-1}\| \left\|T^{-1/2} \sum_{t=p+1}^{\tau-1} w_{1t}^{(0)} [\Delta x_t - \Phi_2 w_{2t}^{(0)}]'\right\| \|T^{1/2}\Phi_1\|. \end{aligned}$$

Here we have used the definition of $\varepsilon_{2t}(\Phi_2)$, the Cauchy-Schwarz inequality and the norm inequality. By an analog of (A.4) of Gregory & Hansen (1996, p. 118), the norm in the middle of the last expression is of order $O_p(1)$ uniformly in $[T\underline{\lambda}] \leq \tau \leq \tau_0$ and for any fixed value of Φ_2. Thus, since the parameters Φ_2 and Ω belong to bounded sets when $(\Theta, \tau, \Omega) \in B_1$ the last expression as a whole has an upper bound $a_{1T}\|T^{1/2}\Phi_1\|$ with a_{1T} as required. This completes the proof. □

Our next result deals with the contribution of $l_{1,\tau_o-1}(\Theta, \tau, \Omega) - l_{1,\tau-1}(\Theta, \tau, \Omega)$ to $l_{1T}(\Theta, \tau, \Omega)$. Here the relevant expression of $\varepsilon_{1t\tau}(\Theta)$ is

$$\varepsilon_{1t\tau}(\Theta) = -\Psi_1 w_{1t}^{(0)} - \underline{\gamma} \underline{d}_{t\tau}, \quad t = \tau, \ldots, \tau_o - 1,$$

where $\Psi_1 = \Phi_1 + [\delta_1 : 0]$.

Lemma A.4. Let ϵ be any real number with the property $0 < \epsilon < \lambda_o - \underline{\lambda}$. Then, for $\underline{\lambda} \leq \lambda \leq \lambda_o - \epsilon$ there exists a constant $c_2 > 0$ such that, with probability approaching one and uniformly in $[T\underline{\lambda}] \leq \tau \leq [T(\lambda_o - \epsilon)]$ and $(\Theta, \tau, \Omega) \in B_1$,

$$l_{1,\tau_0-1}(\Theta,\tau,\Omega) - l_{1,\tau-1}(\Theta,\tau,\Omega) \geq c_2\|T^{1/2}\Psi_1\|^2 + c_2\|\underline{\gamma}\|^2 - a_{2T}\|T^{1/2}\Psi_1\| - a_{3T}\|\underline{\gamma}\|,$$

where $a_{iT} \geq 0$ $(i = 2, 3)$, $a_{2T} = O_p(1)$ and $a_{3T} = o_p(T^{\eta})$ with $\frac{1}{b} < \eta < \frac{1}{4}$.

Proof: By the definitions,

$$\begin{aligned} & l_{1,\tau_o-1}(\Theta,\tau,\Omega) - l_{1,\tau-1}(\Theta,\tau,\Omega) \\ &= \operatorname{tr}\left(\Omega^{-1}\textstyle\sum_{t=\tau}^{\tau_o-1}\varepsilon_{1t\tau}(\Theta)\varepsilon_{1t\tau}(\Theta)'\right) + 2\operatorname{tr}\left(\Omega^{-1}\textstyle\sum_{t=\tau}^{\tau_o-1}\varepsilon_{1t\tau}(\Theta)\varepsilon_{2t}(\Phi_2)'\right) \\ &\overset{def}{=} L_3 + L_4. \end{aligned}$$

First consider L_3 and for simplicity denote $\underline{\Psi}_1 = [\Psi_1 : \underline{\gamma}]$ and $\underline{z}_{1t\tau}^{(0)} = [w_{1t}^{(0)\prime} : \underline{d}_{t\tau}{}']'$. Then

$$L_3 = \operatorname{tr}\left(\Omega^{-1}\underline{\Psi}_1\sum_{t=\tau}^{\tau_o-1}\underline{z}_{1t\tau}^{(0)}\underline{z}_{1t\tau}^{(0)\prime}\underline{\Psi}_1'\right) \geq \lambda_{\min}(\Omega^{-1})\operatorname{tr}\left(\underline{\Psi}_1 D_{1T}\left(D_{1T}^{-1}\sum_{t=\tau}^{\tau_o-1}\underline{z}_{1t\tau}^{(0)}\underline{z}_{1t\tau}^{(0)\prime}D_{1T}^{-1}\right)D_{1T}\underline{\Psi}_1'\right), \tag{A.9}$$

where $D_{1T} = \operatorname{diag}[T^{-1/2}I_{n-r+2} : I_p]$.

Next note that

$$T^{-1/2}\sum_{t=\tau}^{\tau_o-1} w_{1t}^{(0)}\Delta d_{t-i,\tau} = O_p(T^{-1/2}), \quad i = 0,\ldots,p-1, \tag{A.9a}$$

uniformly in $[T\underline{\lambda}] \leq \tau < \tau_o$. Because $w_{1t}^{(0)} = [1 : \frac{t}{T} : T^{-1/2}v_{t-1}^{(0)\prime}]'$ this is obvious for the first and second components of $w_{1t}^{(0)}$. For the third component the same is true because $T^{-1/2}\max_{1\leq t\leq\tau_o}\|v_{t-1}^{(0)}\| \leq T^{-1/2}\max_{1\leq t\leq T}\|\beta_{o\perp}'x_{t-1}\| = O_p(1)$, where the equality follows form the fact that $T^{-1/2}\beta_{o\perp}'x_{[Ts]}$ obeys an invariance principle. Thus, we can conclude that

$$D_{1T}^{-1}\sum_{t=\tau}^{\tau_o-1}\underline{z}_{1t\tau}^{(0)}\underline{z}_{1t\tau}^{(0)\prime}D_{1T}^{-1} = \operatorname{diag}\left[T^{-1}\sum_{t=\tau}^{\tau_o-1}w_{1t}^{(0)}w_{1t}^{(0)\prime} : I_p\right] + o_p(1) \tag{A.10}$$

uniformly in $[T\underline{\lambda}] \leq \tau < \tau_o$.

The next step is to observe that

$$T^{-1}\sum_{t=[T\lambda]}^{[T\lambda_o]-1} w_{1t}^{(0)}w_{1t}^{(0)\prime} \Rightarrow M_{11}(\lambda_o) - M_{11}(\lambda), \quad \underline{\lambda} \leq \lambda < \lambda_o,$$

where $M_{11}(\lambda)$ is the weak limit of $T^{-1}\sum_{t=p+1}^{[T\lambda]-1}w_{1t}^{(0)}w_{1t}^{(0)\prime}$ (cf. (A.3) of Gregory & Hansen (1996, p. 118)). It is straightforward to check that the difference $M_{11}(\lambda_o) - M_{11}(\lambda)$ is positive definite and its smallest eigenvalue is bounded from below by a positive constant when $\underline{\lambda} \leq \lambda \leq \lambda_o - \epsilon$.

The above discussion implies that, with probability approaching one, the smallest eigenvalue of the matrix on the l.h.s. of $(A.10)$ is bounded away from zero uniformly in $[T\underline{\lambda}] \leq \tau \leq [T(\lambda_o - \epsilon)]$. Thus, it follows from $(A.9)$ that, for $(\Theta, \tau, \Omega) \in B_1$ and with probability approaching one,

$$L_3 \geq c_2 \mathrm{tr}(\underline{\Psi}_1 D_{1T} D_{1T} \underline{\Psi}_1') = c_2 \|T^{1/2}\Psi_1\|^2 + c_2 \|\underline{\gamma}\|^2,$$

where $c_2 > 0$ is a (small) constant. This implies that it only remains to show that $L_4 \geq -a_{2T}\|T^{1/2}\Psi_1\| - a_{3T}\|\underline{\gamma}\|$ with a_{2T} and a_{3T} stated in the lemma.

To show the above mentioned inequality about L_4, conclude from the definitions that

$$L_4 = -2\mathrm{tr}\left(\Omega^{-1}\Psi_1 \sum_{t=\tau}^{\tau_o - 1} w_{1t}^{(0)} \varepsilon_{2t}(\Phi_2)'\right) - 2\mathrm{tr}\left(\Omega^{-1}\underline{\gamma} \sum_{t=\tau}^{\tau_0 - 1} \underline{d}_{t\tau} \varepsilon_{2t}(\Phi_2)'\right) \stackrel{def}{=} L_{41} + L_{42}.$$

Arguments similar to those already used in the proof of Lemma A.3 show that

$$|L_{41}| \leq 2\|\Omega^{-1}\| \|T^{1/2}\Psi_1\| \left\| T^{-1/2} \sum_{t=\tau}^{\tau_o - 1} w_{1t}^{(0)} [\Delta x_t - \Phi_2 w_{2t}^{(0)}]' \right\| \leq a_{2T} \|T^{1/2}\Psi_1\|,$$

where $a_{2T} = O_p(1)$ in the required uniform sense.

Regarding L_{42}, one similarly obtains

$$|L_{42}| \leq 2\|\Omega^{-1}\| \|\underline{\gamma}\| \left\| \sum_{t=\tau}^{\tau_o - 1} \underline{d}_{t\tau} [\Delta x_t - \Phi_2 w_{2t}^{(0)}]' \right\| \leq a_{3T} \|\underline{\gamma}\|,$$

where $a_{3T} = o_p(T^\eta)$, $\frac{1}{b} < \eta < \frac{1}{4}$, in the required uniform sense. The latter inequality follows if the last norm in the preceding expression can be replaced by $o_p(T^\eta)$. To justify this, recall that Δx_t and $w_{2t}^{(0)}$ can be treated as stationary processes with finite moments of order $b > 4$ and that Φ_2 can be assumed to belong to a bounded set. Thus, it suffices to show that $\max_{1 \leq t \leq T} \|\Delta x_t\| = o_p(T^\eta)$ and similarly with Δx_t replaced by $w_{2t}^{(0)}$. This, however, can be done by using an argument entirely similar to that in (A.14) of Saikkonen & Lütkepohl (2002). The inequalities obtained for $|L_{41}|$ and $|L_{42}|$ above show that L_2 has the required lower bound and the proof is complete. □

Our next result describes the contribution of $l_{1,\tau_o+p-1}(\Theta, \tau, \Omega) - l_{1,\tau_o-1}(\Theta, \tau, \Omega)$ to $l_{1T}(\Theta, \tau, \Omega)$. We introduce the notation

$$\zeta_{t\tau}^{(0)} = (d_{t\tau} - d_{t\tau_o})\delta_1^{(0)} + \underline{\gamma}\underline{d}_{t\tau} - \underline{\gamma}^{(0)}\underline{d}_{t\tau_o}.$$

In the following lemma the relevant values of $\varepsilon_{1t\tau}(\Theta)$ can then be written as

$$\varepsilon_{1t\tau}(\Theta) = -\Psi_2 w_{1t}^{(0)} - \zeta_{t\tau}^{(0)}, \quad t = \tau_o, \dots, \tau_o + p - 1,$$

where $\Psi_2 = \Phi_1 + [\delta_1 - \delta_1^{(0)} : 0]$. Note that here the first term in the definition of $\zeta_{t\tau}^{(0)}$ vanishes but the general definition is convenient in later derivations. Now we can formulate

Lemma A.5. There exists a constant $c_3 > 0$ such that, with probability approaching one and uniformly in $[T\underline{\lambda}] \leq \tau \leq \tau_o$ and $(\Theta, \tau, \Omega) \in B_1$,

$$l_{1,\tau_o+p-1}(\Theta, \tau, \Omega) - l_{1,\tau_o-1}(\Theta, \tau, \Omega) \geq c_3 \sum_{t=\tau_o}^{\tau_o+p-1} \|\zeta_{t\tau}^{(0)}\|^2 - a_{4T} \left(\sum_{t=\tau_o}^{\tau_o+p-1} \|\zeta_{t\tau}^{(0)}\|^2 \right)^{1/2} - a_{5T},$$

where $a_{iT} \geq 0$ and $a_{iT} = O_p(1)$ $(i = 4, 5)$.

Proof: By the definitions,

$$\begin{aligned} l_{1,\tau_o+p-1}(\Theta, \tau, \Omega) - l_{1,\tau_o-1}(\Theta, \tau, \Omega) &= \operatorname{tr}\left(\Omega^{-1} \sum_{t=\tau_o}^{\tau_o+p-1} \varepsilon_{1t\tau}(\Theta)\varepsilon_{1t\tau}(\Theta)' \right) \\ &\quad + 2\operatorname{tr}\left(\Omega^{-1} \sum_{t=\tau_o}^{\tau_o+p-1} \varepsilon_{1t\tau}(\Theta)\varepsilon_{2t}(\Phi_2)' \right) \\ &\stackrel{def}{=} L_5 + L_6. \end{aligned}$$

Assuming $(\Theta, \tau, \Omega) \in B_1$ we find that

$$\begin{aligned} L_5 &\geq \lambda_{\min}(\Omega^{-1}) \sum_{t=\tau_o}^{\tau_o+p-1} \|\varepsilon_{1t\tau}(\Theta)\|^2 \\ &\geq \bar{\omega}^{-1} \sum_{t=\tau_o}^{\tau_o+p-1} \|\Psi_2 w_{1t}^{(0)}\|^2 + \bar{\omega}^{-1} \sum_{t=\tau_o}^{\tau_o+p-1} \|\zeta_{t\tau}^{(0)}\|^2 + 2\bar{\omega}^{-1} \sum_{t=\tau_o}^{\tau_o+p-1} \zeta_{t\tau}^{(0)\prime} \Psi_2 w_{1t}^{(0)}. \end{aligned}$$

Because we can here assume that Ψ_2 is bounded (see $(A.5)$), an application of the triangle inequality and the Cauchy-Schwarz inequality shows that the absolute value of the third term in the last expression can be bounded from above by

$$const. \times \left(\sum_{t=\tau_o}^{\tau_o+p-1} \|\zeta_{t\tau}^{(0)}\|^2 \right)^{1/2} \left(\sum_{t=\tau_o}^{\tau_o+p-1} \|w_{1t}^{(0)}\|^2 \right)^{1/2}.$$

Here the latter square root is of order $O_p(1)$ (see the argument leading to $(A.10)$). Hence, we can conclude that

$$L_5 \geq c_3 \sum_{t=\tau_o}^{\tau_o+p-1} \|\zeta_{t\tau}^{(0)}\|^2 - a_{41T} \left(\sum_{t=\tau_o}^{\tau_o+p-1} \|\zeta_{t\tau}^{(0)}\|^2 \right)^{1/2}, \tag{A.11}$$

where $c_3 = \bar{\omega}^{-1} > 0$ and $a_{41T} = O_p(1)$ in the required uniform sense.

Now consider L_6. Arguments similar to those used in previous derivations combined with the present definition of $\varepsilon_{1t\tau}(\Theta)$ yield

$$|L_6| \leq 2\|\Omega^{-1}\| \left\| \Psi_2 \sum_{t=\tau_o}^{\tau_o+p-1} w_{1t}^{(0)} \varepsilon_{2t}(\Phi_2)' \right\| + 2\|\Omega^{-1}\| \left\| \sum_{t=\tau_o}^{\tau_o+p-1} \zeta_{t\tau}^{(0)} \varepsilon_{2t}(\Phi_2)' \right\| .$$

It is easy to see that the first term on the r.h.s. can be used to define the term a_{5T} in the lemma. The arguments needed are similar to those used to obtain $(A.11)$ and they can also be applied to the second term so that we can write

$$|L_6| \leq a_{5T} + a_{42T} \left(\sum_{t=\tau_o}^{\tau_o+p-1} \|\zeta_{t\tau}^{(0)}\|^2 \right)^{1/2}, \qquad (A.12)$$

where also $a_{42T} = O_p(1)$ in the required uniform sense. The result of the lemma now follows from $(A.11)$ and $(A.12)$ be defining $a_{4T} = a_{41T} + a_{42T}$. □

The next lemma is concerned with the contribution of $l_{1T}(\Theta, \tau, \Omega) - l_{1,\tau_o+p-1}(\Theta, \tau, \Omega)$ to $l_{1T}(\Theta, \tau, \Omega)$. Here $\varepsilon_{1t\tau}(\Theta)$ is given by

$$\varepsilon_{1t\tau}(\Theta) = -\Psi_2 w_{1t}^{(0)}, \quad t = \tau_o + p, \ldots, T.$$

Lemma A.6. There exists a constant $c_4 > 0$ such that, with probability approaching one and uniformly in $[T\underline{\lambda}] \leq \tau \leq \tau_o$ and $(\Theta, \tau, \Omega) \in B_1$,

$$l_{1T}(\Theta, \tau, \Omega) - l_{1,\tau_o+p-1}(\Theta, \tau, \Omega) \geq c_4 \|T^{1/2}\Psi_2\|^2 - a_{6T}\|T^{1/2}\Psi_2\|,$$

where $a_{6T} \geq 0$ and $a_{6T} = O_p(1)$.

Proof: The proof is similar to that of Lemma A.3 except for being simpler because the considered quantities are independent of τ and uniformity over this parameter is therefore of no concern. Details are omitted. □

Our next lemma is used as an alternative to Lemma A.4 in some of the subsequent derivations. The formulation of this lemma makes use of the notation $\zeta_{t\tau}^{(0)}$ employed in Lemma A.5.

Lemma A.7. There exists a constant $c_5 > 0$ such that with probability approaching one and uniformly in $[T\underline{\lambda}] \leq \tau \leq \tau_o - 1$ and $(\Theta, \tau, \Omega) \in B_1$,

$$
\begin{aligned}
& l_{1,\tau_o-1}(\Theta,\tau,\Omega) - l_{1,\tau-1}(\Theta,\tau,\Omega) \\
& \geq c_5 \sum_{t=\tau}^{\tau_o-1} \|\zeta_{t\tau}^{(0)}\|^2 - \left(a_{7T}(\tau_o-\tau)^{\eta} + a_{8T}\left(\tfrac{\tau_o-\tau}{T}\right)^{1/2} \|T^{1/2}\Psi_2\|\right)\left(\sum_{t=\tau}^{\tau_o-1} \|\zeta_{t\tau}^{(0)}\|^2\right)^{1/2} \\
& \quad -a_{9T}\|T^{1/2}\Psi_2\|,
\end{aligned}
$$

where $\frac{1}{b} < \eta < \frac{1}{4}$, $a_{iT} \geq 0$ and $a_{iT} = O_p(1)$ $(i = 7, 8, 9)$.

Proof: By the definitions,

$$
\begin{aligned}
& l_{1,\tau_o-1}(\Theta,\tau,\Omega) - l_{1,\tau-1}(\Theta,\tau,\Omega) \\
& = \mathrm{tr}\left(\Omega^{-1}\sum_{t=\tau}^{\tau_o-1} \varepsilon_{1t\tau}(\Theta)\varepsilon_{1t\tau}(\Theta)'\right) + 2\mathrm{tr}\left(\Omega^{-1}\sum_{t=\tau}^{\tau_o-1} \varepsilon_{1t\tau}(\Theta)\varepsilon_{2t}(\Phi_2)'\right) \\
& \overset{def}{=} L_7 + L_8.
\end{aligned}
$$

Recall that $\Psi_1 = \Phi_1 + [\delta_1 : 0]$ and $\Psi_2 = \Phi_1 + [\delta_1 - \delta_1^{(0)} : 0]$. For $t = \tau, \ldots, \tau_o - 1$, we thus have

$$
\varepsilon_{1t\tau}(\Theta) = -\Psi_1 w_{1t}^{(0)} - \underline{\gamma}\underline{d}_{t\tau} = -\Psi_2 w_{1t}^{(0)} - \zeta_{t\tau}^{(0)}.
$$

Hence,

$$
\begin{aligned}
L_7 &= \mathrm{tr}\left(\Omega^{-1}\Psi_2 \sum_{t=\tau}^{\tau_o-1} w_{1t}^{(0)} w_{1t}^{(0)\prime}\Psi_2'\right) + \mathrm{tr}\left(\Omega^{-1}\sum_{t=\tau}^{\tau_o-1} \zeta_{t\tau}^{(0)}\zeta_{t\tau}^{(0)\prime}\right) \\
& \quad +2\mathrm{tr}\left(\Omega^{-1}\sum_{t=\tau}^{\tau_o-1} \zeta_{t\tau}^{(0)} w_{1t}^{(0)\prime}\Psi_2'\right) \\
& \overset{def}{=} L_{71} + L_{72} + L_{73}.
\end{aligned}
$$

Assume that $(\Theta, \tau, \Omega) \in B_1$. An application of the Cauchy-Schwarz inequality, the norm inequality and the triangle inequality yields

$$
\begin{aligned}
|L_{73}| &\leq 2\|\Omega^{-1}\|\|T^{1/2}\Psi_2\| \left\| T^{-1/2}\sum_{t=\tau}^{\tau_o-1} \zeta_{t\tau}^{(0)} w_{1t}^{(0)\prime} \right\| \\
&\leq 2\|\Omega^{-1}\| \left(\frac{\tau_o-\tau}{T}\right)^{1/2} \|T^{1/2}\Psi_2\| \left((\tau_o-\tau)^{-1}\sum_{t=\tau}^{\tau_o-1} \|w_{1t}^{(0)}\|^2\right)^{1/2} \left(\sum_{t=\tau}^{\tau_0-1} \|\zeta_{t\tau}^{(0)}\|^2\right)^{1/2}.
\end{aligned}
$$

Because $\max_{1\leq t\leq T} \|w_{1t}^{(0)}\| = O_p(1)$ (see the arguments leading to $(A.10)$), the second square root in the last expression is of order $O_p(1)$ uniformly in $[T\underline{\lambda}] \leq \tau < \tau_o$. Hence,

$$
|L_{73}| \leq a_{8T}\left(\frac{\tau_o-\tau}{T}\right)^{1/2} \|T^{1/2}\Psi_2\| \left(\sum_{t=\tau}^{\tau_o-1} \|\zeta_{t\tau}^{(0)}\|^2\right)^{1/2}, \qquad (A.13)
$$

where $a_{8T} = O_p(1)$ in the required uniform sense.

Next note that $L_{71} \geq 0$ and $\lambda_{\min}(\Omega^{-1}) \geq \bar{\omega}^{-1}$ for $(\Theta, \tau, \Omega) \in B_1$. Consequently,

$$L_{71} + L_{72} \geq \bar{\omega}^{-1} \sum_{t=\tau}^{\tau_o - 1} \|\zeta_{t\tau}^{(0)}\|^2. \tag{A.14}$$

Now consider L_8 for which we have

$$L_8 = -2\text{tr}\left(\Omega^{-1}\Psi_2 \sum_{t=\tau}^{\tau_o-1} w_{1t}^{(0)} \varepsilon_{2t}(\Phi_2)'\right) - 2\text{tr}\left(\Omega^{-1} \sum_{t=\tau}^{\tau_o-1} \zeta_{t\tau}^{(0)} \varepsilon_{2t}(\Phi_2)'\right) \stackrel{def}{=} L_{81} + L_{82}.$$

Arguments similar to those used for L_{73} show that

$$|L_{81}| \leq 2\|\Omega^{-1}\|\|T^{1/2}\Psi_2\| \left\| T^{-1/2} \sum_{t=\tau}^{\tau_o-1} w_{1t}^{(0)} \varepsilon_{2t}(\Phi_2)' \right\| \leq a_{9T}\|T^{1/2}\Psi_2\|, \tag{A.15}$$

where $a_{9T} = O_p(1)$ in the required uniform sense. The latter inequality is obtained because, for $(\Theta, \tau, \Omega) \in B_1$, the last norm in the second expression can be replaced by $O_p(1)$ by an analog of (A.4) of Gregory & Hansen (1996, p. 118).

As for L_{82}, assume first that $\tau < \tau_o - p$ and use the Cauchy-Schwarz inequality to conclude that

$$\begin{aligned}
|L_{82}| &\leq 2\|\Omega^{-1}\| \left\| \sum_{t=\tau}^{\tau_o-1} \zeta_{t\tau}^{(0)} \varepsilon_{2t}(\Phi_2)' \right\| \\
&\leq 2\|\Omega^{-1}\| \left\| \sum_{t=\tau}^{\tau+p-1} (\delta_1^{(0)} + \underline{\gamma}\underline{d}_{t\tau})\varepsilon_{2t}(\Phi_2)' \right\| + 2\|\Omega^{-1}\| \left\| \delta_1^{(0)} \sum_{t=\tau+p}^{\tau_0-1} \varepsilon_{2t}(\Phi_2)' \right\| \\
&\leq 2\|\Omega^{-1}\| \left(\sum_{t=\tau}^{\tau+p-1} \|\delta_1^{(0)} + \underline{\gamma}\underline{d}_{t\tau}\|^2 \right)^{1/2} \left(\sum_{t=\tau}^{\tau+p-1} \|\varepsilon_{2t}(\Phi_2)\|^2 \right)^{1/2} \\
&\quad +2\|\Omega^{-1}\|\|\delta_1^{(0)}\| \left\| \sum_{t=\tau+p}^{\tau_o-1} \varepsilon_{2t}(\Phi_2) \right\|.
\end{aligned}$$

Here the second inequality is based on the definitions and the triangle inequality whereas the third one also makes use of the Cauchy-Schwarz inequality and the norm inequality.

In the last expression

$$(\tau_o - \tau - p)^{-2\eta} \sum_{t=\tau}^{\tau+p-1} \|\varepsilon_{2t}(\Phi_2)\|^2 = O_p(1), \quad \frac{1}{b} < \eta < \frac{1}{4},$$

and

$$(\tau_o - \tau - p)^{-1/2} \left\| \sum_{t=\tau+p}^{\tau_o-1} \varepsilon_{2t}(\Phi_2) \right\| = O_p(1)$$

uniformly in $[T\underline{\lambda}] \leq \tau < \tau_o - p$ and $(\Theta, \tau, \Omega) \in B_1$. Here the latter result can be concluded from the Hájek-Rényi inequality given in Proposition 1 of Bai (1994). The former can be obtained by an argument similar to that used to prove (A.14) of Saikkonen & Lütkepohl (2002).

Combining the above discussion on L_{82} shows that

$$\begin{aligned} |L_{82}| &\leq a_{71T} \left[\left(\sum_{t=\tau}^{\tau+p-1} \|\delta_1^{(0)} + \underline{\gamma} \underline{d}_{t\tau})\|^2 \right)^{1/2} + (\tau_o - \tau - p)^{1/2} \|\delta_1^{(0)}\| \right] \\ &= a_{71T} \left[\left(\sum_{t=\tau}^{\tau+p-1} \|\zeta_{t\tau}^{(0)}\|^2 \right)^{1/2} + \left(\sum_{t=\tau+p}^{\tau_o-1} \|\zeta_{t\tau}^{(0)}\|^2 \right)^{1/2} \right], \end{aligned}$$

where $a_{71T} = O_p((\tau_o - \tau)^\eta)$ in the required uniform sense and the equality follows from definitions. Since for any real numbers $a \geq 0$ and $b \geq 0$ we have $a + b \leq \sqrt{2}(a^2 + b^2)^{1/2}$ it follows that

$$|L_{82}| \leq \sqrt{2} a_{71T} \left(\sum_{t=\tau}^{\tau_o-1} \|\zeta_{t\tau}^{(0)}\|^2 \right)^{1/2}. \qquad (A.16)$$

In the proof of this result it was assumed that $\tau < \tau_o - p$ but it also holds for $\tau_o - p \leq \tau < \tau_o$. In that case arguments similar to those used for L_{73} give

$$|L_{82}| \leq 2\|\Omega^{-1}\| \left(\sum_{t=\tau}^{\tau_o-1} \|\zeta_{t\tau}^{(0)}\|^2 \right)^{1/2} \left(\sum_{t=\tau}^{\tau_o-1} \|\varepsilon_{2t}(\Phi_2)\|^2 \right)^{1/2}$$

and $(A.16)$ holds with $a_{71T} = O_p(1)$. The result of the lemma is obtained from the definitions of L_7 and L_8 in conjunction with $(A.13)-(A.16)$ by defining $c_5 = \bar{\omega}^{-1}$, $a_{7T} = \sqrt{2} a_{71T}/(\tau_o - \tau)^\eta$ and a_{8T} and a_{9T} as done in $(A.13)$ and $(A.15)$, respectively. □

In the proof of the next lemma as well as in subsequent proofs, frequent use will be made of the elementary inequality

$$a_2 x^2 - a_1 x - a_0 \geq -\frac{a_1^2}{4a_2} - a_0, \quad x \geq 0, \qquad (A.17)$$

which holds for $a_0, a_1 \geq 0$ and $a_2 > 0$.

Lemma A.8. Let $\epsilon > 0$ and $B_2 = \{(\Theta, \tau, \Omega) : \|T^{1/2-\eta}\Phi_1\|^2 + \|T^{1/2-\eta}\Psi_2\|^2 \leq \epsilon^2\}$, where $\frac{1}{b} < \eta < \frac{1}{4}$ is the same as in Lemma A.7. Then,

$$\inf_{(\Theta,\tau,\Omega) \in B_2^c} l_T(\Theta, \tau, \Omega) - l_T(\Theta_o, \tau_o, \Omega_o) > 0$$

with probability approaching one and uniformly in $[T\underline{\lambda}] \le \tau \le \tau_o$.

Proof: By the definitions and Lemma A.2,

$$\begin{aligned} l_T(\Theta,\tau,\Omega) - l_T(\Theta_o,\tau_o,\Omega_o) &= l_{1T}(\Theta,\tau,\Omega) + l_{2T}(\Phi_2,\Omega) - l_{2T}(\Phi_{2o},\Omega_o) \\ &\ge l_{1T}(\Theta,\tau,\Omega) + \inf_{(\Phi_2,\Omega)} l_{2T}(\Phi_2,\Omega) - l_{2T}(\Phi_{2o},\Omega_o) \qquad (A.18) \\ &= l_{1T}(\Theta,\tau,\Omega) + O_p(1) \end{aligned}$$

Thus, it suffices to show that, for some $\epsilon_* > 0$,

$$\inf_{(\Theta,\tau,\Omega)\in B_2^c} T^{-2\eta} l_{1T}(\Theta,\tau,\Omega) \ge \epsilon_* \qquad (A.19)$$

with probability approaching one and uniformly in $[T\underline{\lambda}] \le \tau \le \tau_o$.

From Lemma A.1 it follows that we only need to prove $(A.19)$ with the set B_2^c replaced by $B_1 \cap B_2^c$. Let $0 < \epsilon_1 \le \lambda_o - \underline{\lambda}$ and define the sets

$$B_{21} = B_1 \cap B_2^c \cap \{(\Theta,\tau,\Omega) : [T\underline{\lambda}] \le \tau < [T(\lambda_o - \epsilon_1)]\}$$

and

$$B_{22} = B_1 \cap B_2^c \cap \{(\Theta,\tau,\Omega) : [T(\lambda_o - \epsilon_1)] < \tau \le \tau_o\}.$$

According to what was said above, it suffices to establish $(A.19)$ separately with B_2^c replaced by B_{21} and B_{22}. Here we are free to choose the value of ϵ_1. Whatever our choice, Lemma A.4 can be applied on the set B_{21} on which we shall first concentrate.

From Lemmas A.4, A.5 and $(A.17)$ we first find that, uniformly in B_{21},

$$T^{-2\eta} l_{1,\tau_o-1}(\Theta,\tau,\Omega) - T^{-2\eta} l_{1,\tau-1}(\Theta,\tau,\Omega) \ge -(a_{2T}^2 + a_{3T}^2)/4c_2 T^{2\eta} = o_p(1)$$

and

$$T^{-2\eta} l_{1,\tau_o+p-1}(\Theta,\tau,\Omega) - l_{1,\tau_o-1}(\Theta,\tau,\Omega) \ge -\frac{a_{4T}^2}{4c_3 T^{2\eta}} - \frac{a_{5T}}{T^{2\eta}} = o_p(1).$$

Combining these inequalities with those obtained from Lemmas A.3 and A.6 shows that, uniformly in B_{21},

$$\begin{aligned} T^{-2\eta} l_{1T}(\Theta,\tau,\Omega) \ \ge\ & c_1 \|T^{1/2-\eta}\Phi_1\|^2 - T^{-\eta} a_{1T} \|T^{1/2-\eta}\Phi_1\| \\ & + c_4 \|T^{1/2-\eta}\Psi_2\|^2 - T^{-\eta} a_{6T} \|T^{1/2-\eta}\Psi_2\| + o_p(1). \end{aligned} \qquad (A.20)$$

Denote $c^* = \min(c_1, c_4)$ and $a_T^* = \sqrt{2}\max(a_{1T}, a_{6T})$. Then the preceding inequality implies that, uniformly in B_{21},

$$\begin{aligned} T^{-2\eta} l_{1T}(\Theta, \tau, \Omega) \;\; \geq \;\; & c^*(\|T^{1/2-\eta}\Phi_1\|^2 + \|T^{1/2-\eta}\Psi_2\|^2) \\ & -\frac{1}{\sqrt{2}} T^{-\eta} a_T^* (\|T^{1/2-\eta}\Phi_1\| + \|T^{1/2-\eta}\Psi_2\|) + o_p(1). \end{aligned}$$

For simplicity, denote $\varphi_T^2 = \|T^{1/2-\eta}\Phi_1\|^2 + \|T^{1/2-\eta}\Psi_2\|^2$ and note that the sum of the two norms in the last expression above is at most $\sqrt{2}\varphi_T$. Thus, uniformly in B_{21},

$$T^{-2\eta} l_{1T}(\Theta, \tau, \Omega) \geq c^* \varphi_T^2 - T^{-\eta} a_T^* \varphi_T + o_p(1) = c^* \varphi_T^2 \left(1 - \frac{a_T^*}{c^* T^\eta \varphi_T}\right) + o_p(1). \qquad (A.21)$$

Because $\varphi_T > \epsilon$ on B_{21} and $a_T^* = O_p(1)$ uniformly in B_{21}, this shows that $(A.19)$ holds with B_2^c replaced by B_{21}.

Now consider proving $(A.19)$ with B_2^c replaced by B_{22}. Here we can use Lemmas A.3, A.5, A.6 and A.7 to conclude that, with probability approaching one and uniformly in B_{22},

$$\begin{aligned} T^{-2\eta} l_{1,T}(\Theta, \tau, \Omega) \;\; \geq \;\; & c_1 \|T^{1/2-\eta}\Phi_1\| - T^{-\eta} a_{1T} \|T^{1/2-\eta}\Phi_1\| \\ & + c_4 \|T^{1/2-\eta}\Psi_2\|^2 - T^{-\eta}(a_{6T} + a_{9T}) \|T^{1/2-\eta}\Psi_2\| \\ & + c_3 T^{-2\eta} \textstyle\sum_{t=\tau_o}^{\tau_o+p+1} \|\zeta_{t\tau}^{(0)}\|^2 - T^{-\eta} a_{4T} \left(T^{-2\eta} \textstyle\sum_{t=\tau_o}^{\tau_o+p+1} \|\zeta_{t\tau}^{(0)}\|^2\right)^{1/2} - T^{-2\eta} a_{5T} \\ & + c_5 T^{-2\eta} \textstyle\sum_{t=\tau}^{\tau_o-1} \|\zeta_{t\tau}^{(0)}\|^2 \\ & - \left[a_{7T} \left(\tfrac{\tau_o-\tau}{T}\right)^\eta + a_{8T} \left(\tfrac{\tau_o-\tau}{T}\right)^{1/2} \|T^{1/2-\eta}\Psi_2\|\right] \left(T^{-2\eta} \textstyle\sum_{t=\tau}^{\tau_o-1} \|\zeta_{t\tau}^{(0)}\|^2\right)^{1/2}. \end{aligned} \qquad (A.22)$$

By $(A.17)$ the sum of the fifth, sixth and seventh terms on the r.h.s. is of order $o_p(1)$ uniformly in B_{22} and the sum of the last two terms can be bounded from below by

$$-\frac{1}{4c_5} \left[a_{7T} \left(\frac{\tau_o - \tau}{T}\right)^\eta + a_{8T} \left(\frac{\tau_o - \tau}{T}\right)^{1/2} \|T^{1/2-\eta}\Psi_2\|\right]^2.$$

Thus, expanding the square and inserting the result to the r.h.s. of the preceding inequality yields, uniformly in B_{22},

$$\begin{aligned} T^{-2\eta} l_{1T}(\Theta, \tau, \Omega) \;\; \geq \;\; & c_1 \|T^{1/2-\eta}\Phi_1\|^2 - T^{-\eta} a_{1T} \|T^{1/2-\eta}\Phi_1\| \\ & + c_{4T}(\tau) \|T^{1/2-\eta}\Psi_2\|^2 - a_{10T}(\tau) \|T^{1/2-\eta}\Psi_2\| - a_{11T}(\tau) + o_p(1), \end{aligned} \qquad (A.23)$$

where

$$c_{4T}(\tau) = c_4 - \frac{a_{8T}^2}{4c_5} \left(\frac{\tau_0 - \tau}{T}\right),$$

$$a_{10T}(\tau) = T^{-\eta}a_{6T} + T^{-\eta}a_{9T} + \frac{a_{7T}a_{8T}}{2c_5}\left(\frac{\tau_o - \tau}{T}\right)^{1/2+\eta}$$

and

$$a_{11T}(\tau) = \frac{a_{7T}^2}{4c_5}\left(\frac{\tau_o - \tau}{T}\right)^{2\eta}.$$

Note that here $a_{6T}, \ldots, a_{9T}$ are of order $O_p(1)$ uniformly in B_{22} and that, on B_{22}, $(\tau_o - \tau)/T \leq 2\epsilon_1$, say. Since we are here free to choose the value of ϵ_1 we can choose it so small that the following two conditions hold with probability approaching one and uniformly in B_{22}: (i) $c_{4T}(\tau) \geq c_4/2$ and (ii) $a_{10T}(\tau)$ and $a_{11T}(\tau)$ become smaller than any preassigned positive number. Taking these facts into account and comparing the inequality $(A.23)$ with $(A.20)$ shows that there are only two points which make the previous proof based on inequality $(A.20)$ directly inapplicable in the present context. These points are that instead of the terms $T^{-\eta}a_{6T} = o_p(1)$ and $o_p(1)$ we have in $(A.23)$ $a_{10T}(\tau)$ and $a_{11T}(\tau) + o_p(1)$, respectively, which are not of order $o_p(1)$ but can only be replaced by an arbitrarily small positive number independent of parameters. However, this is sufficient for the application of essentially the same proof as previously. Indeed, we can conclude that, uniformly in B_{22}, an analog of $(A.21)$ holds except that in the last expression T^{η} is replaced by a fixed positive number which can be assumed as large as we wish and $o_p(1)$ is replaced by a fixed negative number which, in absolute value, can be assumed as small as we wish. In particular, we can assume that T^{η} and $o_p(1)$ in $(A.21)$ are replaced by M/ϵ and $-\epsilon/M$, respectively, where M can be chosen arbitrarily large. This shows that we can make the r.h.s. of the present version of $(A.21)$ larger than some $\epsilon_* > 0$ with probability approaching one. Thus, there is a choice of ϵ_1 such that $(A.19)$ holds with B_2^c replaced by B_{21} and B_{22}. This completes the proof. □

The next lemma is similar to Lemma A.8 except that it deals with the short-run parameter Φ_2.

Lemma A.9. Let $\epsilon > 0$ and $B_3 = \{(\Theta, \tau, \Omega) : \|T^{1/2-\eta}(\Phi_2 - \Phi_{2o})\| \leq \epsilon\}$, where $\frac{1}{b} < \eta < \frac{1}{4}$ is the same as in Lemma A.7. Then,

$$\inf_{(\Theta,\tau,\Omega)\in B_3^c} l_T(\Theta, \tau, \Omega) - l_T(\Theta_o, \tau_o, \Omega_o) > 0$$

with probability approaching one and uniformly in $[T\underline{\lambda}] \leq \tau \leq \tau_o$.

Proof: By Lemma A.1 it suffices to prove the result with B_3^c replaced by $B_1 \cap B_3^c$. First consider the break dates $[T\underline{\lambda}] \leq \tau \leq [T(\lambda_o - \epsilon_1)]$ and note that the derivation of the inequality in $(A.21)$ is valid for these break dates and for all $(\Theta, \tau, \Omega) \in B_1 \cap B_3^c$. It is also valid for every $\epsilon_1 > 0$. Thus, an application of $(A.17)$ shows that in this part of the parameter space $T^{-2\eta} l_{1T}(\Theta, \tau, \Omega) \geq o_p(1)$ holds uniformly. Next note that the inequality $(A.23)$ is valid for $[T(\lambda_o - \epsilon)] < \tau \leq \tau_o$ and for all $(\Theta, \tau, \Omega) \in B_1 \cap B_3^c$. Moreover, as the discussion after that inequality reveals, we can, with a suitable (small) choice of ϵ_1, obtain an analog of $(A.21)$ and conclude that, with probability approaching one and uniformly in the considered part of the parameter space, $T^{-2\eta} l_{1T}(\Theta, \tau, \Omega) \geq -\epsilon_2$, where $\epsilon_2 > 0$ can be chosen arbitrarily small. From the above discussion and the first equality in $(A.18)$ it thus follows that we need to show that, for some $\epsilon_* > 0$,

$$\inf_{(\Theta,\tau,\Omega)\in B_3^c} T^{-2\eta} l_{2T}(\Phi_2, \Omega) - T^{-2\eta} l_{2T}(\Phi_2, \Omega) \geq \epsilon_*$$

with probability approaching one. Arguments needed to show this are similar to those used in previous proofs and also very similar to those used to prove the consistency of the LS estimators of the parameters Φ_2 and Ω in the standard regression model $\Delta x_t = \Phi_2 w_t^{(0)} + \varepsilon_t$. Details are straightforward and are omitted. □

The next lemma again makes use of the notation $\zeta_{t\tau}^{(0)}$ introduced for Lemma A.5.

Lemma A.10. Let $B_4 = \{(\Theta, \tau, \Omega) : (\tau_o - \tau)^{-2\eta} \sum_{t=\tau}^{\tau_o+p-1} \|\zeta_{t\tau}^{(0)}\|^2 \leq M^2\}$, where $\tau < \tau_o$ and $\frac{1}{b} < \eta < \frac{1}{4}$ is the same as in Lemma A.7. Then, there exists a real number $M_0 > 0$ such that, for all $M \geq M_0$,

$$\inf_{(\Theta,\tau,\Omega)\in B_4^c} l_T(\Theta, \tau, \Omega) - l_T(\Theta_o, \tau_o, \Omega_o) > 0$$

with probability approaching one and uniformly in $[T\underline{\lambda}] \leq \tau \leq \tau_o - 1$.

Proof: From $(A.18)$ it follows that it suffices to show that there exists a real number $M_0 > 0$ such that, for all $M \geq M_0$ and any $M_1 > 0$,

$$\inf_{(\Theta,\tau,\Omega)\in B_4^c} l_{1T}(\Theta, \tau, \Omega) > M_1 \qquad (A.24)$$

with probability approaching one and uniformly in $[T\underline{\lambda}] \leq \tau \leq \tau_o - 1$. From Lemmas A.1, A.8 and A.9 it further follows that here the set B_4^c can be replaced by $B_1 \cap B_2 \cap B_3 \cap B_4^c$.

From $(A.19)$ it can be seen that the value of ϵ in the definition of B_2 can be chosen arbitrarily small.

We wish to apply Lemmas A.3, A.5, A.6 and A.7 to obtain a lower bound for $l_{1T}(\Theta, \tau, \Omega)$. This lower bound can be obtained by multiplying both sides of the inequality $(A.22)$ by $T^{2\eta}$. By $(A.17)$ the contribution of the first four terms to the r.h.s. of the resulting inequality can be replaced by $O_p(1)$. This is also the case for the seventh term. Hence, we can write

$$\begin{aligned} l_{1T}(\Theta,\tau,\Omega) \quad \geq \quad & c_3 \sum_{t=\tau_o}^{\tau_o+p-1} \|\zeta_{t\tau}^{(0)}\|^2 - a_{4T}\left(\sum_{t=\tau_o}^{\tau_o+p-1} \|\zeta_{t\tau}^{(0)}\|^2\right)^{1/2} + c_5 \sum_{t=\tau}^{\tau_o-1} \|\zeta_{t\tau}^{(0)}\|^2 \\ & - \left(a_{7T}(\tau_o-\tau)^{\eta} + a_{8T}(\tau_o-\tau)^{1/2}\|\Psi_2\|\right)\left(\sum_{t=\tau}^{\tau_o-1} \|\zeta_{t\tau}^{(0)}\|^2\right)^{1/2} + O_p(1). \end{aligned}$$

This holds uniformly in $B_1 \cap B_2 \cap B_3 \cap B_4^c$ and $[T\underline{\lambda}] \leq \tau \leq \tau_o - 1$. In this part of the parameter space we also have

$$(\tau_o-\tau)^{1/2}\|\Psi_2\| = (\tau_o-\tau)^{\eta}\left(\frac{\tau_o-\tau}{T}\right)^{1/2-\eta}\|T^{1/2-\eta}\Psi_2\| \leq \epsilon(\tau_o-\tau)^{\eta}$$

and $a_{4T} \leq a_{4T}(\tau_o-\tau)^{\eta}$ (see Lemma A.8). Denote $c^* = \min(c_3, c_5)$, $a_T^* = \max(a_{4T}, a_{7T}+\epsilon a_{8T})$ and for simplicity,

$$\xi_\tau^2 = \sum_{t=\tau}^{\tau_o+p-1} \|\zeta_{t\tau}^{(0)}\|^2.$$

From the lower bound obtained for $l_{1T}(\Theta, \tau, \Omega)$ above we can then further obtain

$$l_{1T}(\Theta,\tau,\Omega) \geq c^*\xi_\tau^2 - a_T^*(\tau_o-\tau)^{\eta}\xi_\tau + O_p(1) = c^*\xi_\tau^2\left(1 - \frac{a_T^*(\tau_o-\tau)^{\eta}}{c^*\xi_\tau}\right) + O_p(1). \quad (A.25)$$

Again, this holds uniformly in $B_1 \cap B_2 \cap B_3 \cap B_4^c$ and $[T\underline{\lambda}] \leq \tau \leq \tau_o - 1$. Now, on B_4^c, $\xi_\tau > M(\tau_o-\tau)^{\eta}$ so that, for all M large enough and with probability approaching one, we can make the r.h.s. of $(A.25)$ larger than any preassigned number $M_1 > 0$. Thus, we have established $(A.24)$ and thereby the assertion of the lemma. □

Before proceeding to new proofs we discuss how Lemmas A.3-A.10 are formulated when $\tau \geq \tau_o$.

The counterpart of Lemma A.3 is concerned with the time points $t = p+1, \ldots, \tau_o - 1$ and break dates $\tau_o \leq \tau \leq [T\bar{\lambda}]$ but is otherwise similar to Lemma A.3 and its proof is similar to the proof of Lemma A.6 in that uniformity in τ is of no concern.

The next time points of interest are now $t = \tau_o, \ldots, \tau_o + p - 1$ so that we need to consider a counterpart of Lemma A.5. Here we write

$$\begin{aligned} \varepsilon_{1t\tau}(\Theta) &= -\Phi_1 w_{1t}^{(0)} - d_{t\tau_o}(\delta_1 - \delta_1^{(0)}) - (d_{t\tau} - d_{t\tau_o})\delta_1 - \underline{\gamma}\underline{d}_{t\tau} + \underline{\gamma}^{(0)}\underline{d}_{t\tau_o} \\ &= -\Psi_2 w_{1t}^{(0)} - \zeta_{t\tau}, \qquad t = \tau_o, \ldots, \tau_o + p - 1, \end{aligned}$$

where $\Psi_2 = \Phi_1 + [\delta_1 - \delta_1^{(0)} : 0]$ as before and

$$\zeta_{t\tau} = (d_{t\tau} - d_{t\tau_o})\delta_1 + \underline{\gamma}\underline{d}_{t\tau} - \underline{\gamma}^{(0)}\underline{d}_{t\tau_o}.$$

In other words, in place of $\zeta_{t\tau}^{(0)}$ we now use an analogous variable defined by using the parameter δ_1 instead of $\delta_1^{(0)}$. However, replacing $\zeta_{t\tau}^{(0)}$ in Lemma A.5 by $\zeta_{t\tau}$ is clearly possible, as can be seen from the given proof.

Instead of the time points $t = \tau_o + p, \ldots, \tau - 1$ it is next reasonable to consider the time points $t = \tau_o + p, \ldots, \tau + p - 1$. Then the number of time points is the same as in Lemmas A.4 and A.7. Changes in parameters have to be made, though. Now

$$\varepsilon_{1t\tau}(\Theta) = -\Phi_1 w_{1t}^{(0)} + d_{t\tau_o}\delta_1^{(0)} - d_{t\tau}\delta_1 - \underline{\gamma}\underline{d}_{t\tau} = -\Psi_1^{(0)} w_{1t}^{(0)} - (d_{t\tau}\delta_1 + \underline{\gamma}\underline{d}_{t\tau}), \quad t = \tau_o + p, \ldots, \tau + p - 1,$$

where $\Psi_1^{(0)} = \Phi_1 - [\delta_1^{(0)} : 0]$. Thus, we now have the matrix $\Psi_1^{(0)}$ in place of Ψ_1 used in Lemma A.4 and, as above, the former is defined by using $\delta_1^{(0)}$ instead of δ_1 in Ψ_1. The parameter $\underline{\gamma}$ used in Lemma A.4 is also changed by adding δ_1 to its columns. With these replacements the counterpart of Lemma A.4 applies with $[T(\lambda_o + \epsilon)] \le \tau \le [T\bar{\lambda}]$.

Next consider the counterpart of Lemma A.7 which is also concerned with time points $t = \tau_o + p, \ldots, \tau_o + p - 1$. Here the preceding expression of $\varepsilon_{1t\tau}(\Theta)$ is modified to the form

$$\varepsilon_{1t\tau}(\Theta) = -\Phi_2 w_{1t}^{(0)} - \zeta_{t\tau}, \quad t = \tau_o + p, \ldots, \tau + p - 1.$$

In the counterpart of Lemma A.7 we then have $\zeta_{t\tau}$ in place of $\zeta_{t\tau}^{(0)}$ and $\tau_o + 1 \le \tau \le [T\bar{\lambda}]$. The proof can again be basically obtained by following the previous proof.

The counterpart of Lemma A.6 is straightforward. The relevant time points are $t = \tau, \ldots, T$ and the obtained lower bound is as before except for the obvious change in the values of τ which become $\tau_o \le \tau \le [T\bar{\lambda}]$. The proof is also changed and becomes similar to the proof of Lemma A.3.

It is not difficult to check that the modified versions of Lemmas A.3 - A.7 can be used to show that the results of Lemmas A.8 and A.9 also apply for $\tau_o \leq \tau \leq [T\bar{\lambda}]$. Regarding Lemma A.10, when $\tau_o + 1 \leq \tau \leq [T\bar{\lambda}]$, the set B_4 is defined as

$$B_4 = \left\{ (\Theta, \tau, \Omega) : (\tau_o - \tau)^{-2\eta} \sum_{t=\tau_o}^{\tau+p-1} \|\zeta_{t\tau}\|^2 \leq M^2 \right\}$$

but otherwise the same result is obtained.

Now we can turn to our next lemma which is central in studying asymptotic properties of the break date estimator. Recall that $\delta_{1o} = -\Pi_o \delta_o = -\alpha_o \beta_o' \delta_o$. Thus, $\delta_{1o} \neq 0$ if and only if ${\beta_o}' \delta_o \neq 0$. Note also that we shall use the convention that the infimum over an empty set is ∞.

Lemma A.11. Let $M > 0$. Assume that $\delta_{1o} \neq 0$ and define $B_5 = \{(\Theta, \tau, \Omega) : (|\tau_o - \tau| - p)\|\delta_{1o}\|^{2/(1-2\eta)} \leq M\}$, where $\frac{1}{b} < \eta < \frac{1}{4}$ is the same as in Lemma A.7 or its counterpart when $\tau > \tau_o$. Then there exists a real number $M_0 > 0$ such that, for all $M \geq M_0$,

$$\inf_{(\Theta,\tau,\Omega) \in B_5^c} l_T(\Theta, \tau, \Omega) - l_T(\Theta_o, \tau_o, \Omega_o) > 0$$

with probability approaching one.

Proof: Assume first that $\tau < \tau_o - p$. From Lemmas A.1, A.8 and A.9 it follows that we can replace the set B_5^c by $B_1 \cap B_2 \cap B_3 \cap B_5^c$.

By the definitions,

$$\delta_1^{(0)} = -\Pi \delta_o = -\alpha^{(0)} \beta_o' \delta_o - \rho^{(0)} \beta_{o\perp}' \delta_o,$$

where $\beta_o' \delta_o \neq 0$. On B_3, $\|\alpha^{(0)} - \alpha_o\| \leq \epsilon T^{\eta - 1/2}$ and, on B_2, $\|\rho^{(0)}\| \leq \epsilon T^{\eta-1}$. Thus, since $\delta_{1o} = -\alpha_o \beta_o' \delta_o$,

$$\begin{aligned} \|\delta_1^{(0)} - \delta_{1o}\| &\leq \|\alpha^{(0)} - \alpha_o\| \|\beta_o' \delta_o\| + \|\rho^{(0)}\| \|\beta_{o\perp}' \delta_o\| \\ &\leq c\epsilon T^{\eta - 1/2} \end{aligned}$$

for some positive and finite constant c. Hence, because $\zeta_{t\tau}^{(0)} = (d_{t\tau} - d_{t\tau_o})\delta_1^{(0)}$ for $t = \tau + p, \ldots, \tau_o - 1$, we have on $B_1 \cap B_2 \cap B_3 \cap B_5^c$ and when $\tau < \tau_o - p$,

$$\begin{aligned}
\left((\tau_o - \tau)^{-2\eta} \textstyle\sum_{t=\tau}^{\tau_o+p-1} \|\zeta_{t\tau}^{(0)}\|^2\right)^{1/2} &\geq \left((\tau_o - \tau)^{-2\eta} \textstyle\sum_{t=\tau+p}^{\tau_o-1} \|\zeta_{t\tau}^{(0)}\|^2\right)^{1/2} \\
&= (\tau_o - \tau)^{-\eta}(\tau_o - \tau - p)^{1/2}\|\delta_1^{(0)}\| \\
&= \left(1 - \tfrac{p}{\tau_o-\tau}\right)^{\eta} (\tau_o - \tau - p)^{1/2-\eta}\|\delta_1^{(0)}\| \\
&\geq \left(\tfrac{1}{p+1}\right)^{\eta} \left((\tau_o - \tau - p)\|\delta_{1o}\|^{2/(1-2\eta)}\right)^{1/2-\eta} \left(1 - \tfrac{\|\delta_1^{(0)}-\delta_{1o}\|}{\|\delta_{1o}\|}\right) \\
&\geq \tfrac{M^{1/2-\eta}}{(p+1)^{\eta}} \left(1 - \tfrac{c\epsilon T^{\eta-1/2}}{\|\delta_{1o}\|}\right).
\end{aligned} \qquad (A.26)$$

Here the fourth relation makes use of the triangle inequality. For all T and M large enough the last expression can be made larger than the real number M_0 in Lemma A.10. Thus, the stated result follows from Lemma A.10.

Now consider the case $\tau > \tau_o + p$. Then, using the counterparts of Lemmas A.8 and A.9 we can proceed in the same way as in the case $\tau < \tau_o - p$ until the relations $(A.26)$ which start now as

$$\left((\tau - \tau_o)^{-2\eta} \sum_{t=\tau_o}^{\tau+p-1} \|\zeta_{t\tau}\|^2\right)^{1/2} \geq \left((\tau - \tau_o)^{-2\eta} \sum_{t=\tau_o+p}^{\tau-1} \|\zeta_{t\tau}\|^2\right)^{1/2} = (\tau-\tau_o)^{-\eta}(\tau-\tau_o-p)^{1/2}\|\delta_1\|.$$

Thus, in place of $\delta_1^{(0)}$ we have now δ_1. However, from the counterpart of Lemma A.8 we find that, on B_2, $\|\delta_1 - \delta_1^{(0)}\| \leq \epsilon T^{\eta-1/2}$ and a straightforward modification of the arguments in the latter part of $(A.26)$ combined with the present version of Lemma A.10 give the desired result. This completes the proof of the lemma. □

As discussed earlier, the estimator $\hat{\tau}$ can also be obtained by minimizing -2 times the Gaussian log-likelihood function $l_T(\Theta, \tau, \Omega)$. Thus, Theorem 3.1(i) follows directly from Lemma A.11.

Proof of Theorem 3.1(ii)

The estimator $\tilde{\tau}$ can be obtained by minimizing the Gaussian likelihood function $l_T(\Theta, \tau, \Omega)$ subject to the restriction $\underline{\gamma} = 0$. Thus, we can consider minimizing the objective function

$$l_T^*(\Theta_1, \tau, \Omega) = (T - p) \log \det \Omega + \operatorname{tr}\left(\Omega^{-1} \sum_{t=p+1}^{T} \varepsilon_{t\tau}^*(\Theta_1)\varepsilon_{t\tau}^*(\Theta_1)'\right),$$

where $\Theta_1 = [\Phi : \delta_1]$ and

$$\varepsilon^*_{t\tau}(\Theta_1) = \Delta x_t - \Phi w^{(0)}_t - \delta_1 d_{t\tau} + \delta^{(0)}_1 d_{t\tau_o} + \underline{\gamma}^{(0)} \underline{d}_{t\tau_o}.$$

Denote $y^{(*)}_t = x_t + \sum_{j=0}^{p-1} \gamma_{jo} d_{t-j,\tau_o}$. Then $\Delta y^{(*)}_t = \Delta x_t + \underline{\gamma}_o \underline{d}_{t\tau_o}$ and we can express $\varepsilon^*_{t\tau}(\Theta_1)$ as

$$\varepsilon^*_{t\tau}(\Theta_1) = \varepsilon^*_{1t\tau}(\Theta_1) + \varepsilon^*_{2t}(\Phi_2),$$

where

$$\varepsilon^*_{1t\tau}(\Theta_1) = -\Phi_1 w^{(0)}_{1t} - \delta_1 d_{t\tau} + \delta^{(0)}_1 d_{t\tau_o} + (\underline{\gamma}^{(0)} - \underline{\gamma}_o)\underline{d}_{t\tau_o}$$

and

$$\varepsilon^*_{2t}(\Phi_2) = \Delta y^{(*)}_t - \Phi_2 w^{(0)}_{2t}.$$

Analogously to $l_T(\Theta, \tau, \Omega)$ we can decompose $l^*_T(\Theta_1, \tau, \Omega)$ as

$$l^*_T(\Theta_1, \tau, \Omega) = l^*_{1T}(\Theta_1, \tau, \Omega) + l^*_{2T}(\Phi_2, \Omega),$$

where

$$l^*_{1T}(\Theta_1, \tau, \Omega) = \operatorname{tr}\left(\Omega^{-1} \sum_{t=p+1}^{T} \varepsilon^*_{1t\tau}(\Theta_1)\varepsilon^*_{1t\tau}(\Theta_1)'\right) + 2\operatorname{tr}\left(\Omega^{-1} \sum_{t=p+1}^{T} \varepsilon^*_{1t\tau}(\Theta_1)\varepsilon^*_{2t}(\Phi_2)'\right)$$

and

$$l^*_{2T}(\Phi_2, \Omega) = (T - p)\log\det\Omega + \operatorname{tr}\left(\Omega^{-1} \sum_{t=p+1}^{T} \varepsilon^*_{2t}(\Phi_2)\varepsilon^*_{2t}(\Phi_2)'\right).$$

We shall now discuss modifications of Lemmas A.1 - A.10 based on the objective function $l^*_T(\Theta_1, \tau, \Omega)$. Some of them are minor and are therefore only briefly mentioned. For Lemmas A.5 and A.7 - A.10, new statements of the results will be provided which will be furnished with a superscript '*' to indicate the correspondence to the previous results. We will refer to all the modified versions of the lemmas by attaching a superscript '*' to the number even if no explicit statement of the result is presented.

In what follows frequent use will be made of the facts that $\varepsilon^*_{1t\tau_o}(\Theta_{1o}) = 0$ and $\varepsilon^*_{2t}(\Phi_2) = \varepsilon_{2t}(\Phi_2) + \underline{\gamma}_o \underline{d}_{t\tau_o}$ so that, in particular, $\varepsilon^*_{2t}(\Phi_{2o}) = \varepsilon_t + \underline{\gamma}_o \underline{d}_{t\tau_o} = \varepsilon^*_t$. These are immediate consequences of the definitions. We begin by discussing the required modifications of Lemmas A.1 - A.4 and A.6.

The result of <u>Lemma A.1</u> holds with $l_T(\Theta, \tau, \Omega)$ replaced by $l^*_T(\Theta_1, \tau, \Omega)$ and the set B_1 redefined by replacing the parameter Θ by Θ_1. In subsequent discussions this redefinition of

B_1 will be denoted by B_1^*. To see that this modification of Lemma A.1 holds, notice that from the proof of that lemma it can be seen that we first need to show that $T^{-1}l_T^*(\Theta_{1o}, \tau_o, \Omega_o) = O_p(1)$. However, because $T^{-1}l_T^*(\Theta_{1o}, \tau_o, \Omega_o)$ differs from the expression in the middle of $(A.6)$ only in that ε_t is replaced by $\varepsilon_t^* = \varepsilon_t + \underline{\gamma}_o \underline{d}_{t\tau_o}$, this follows by straightforward application of the weak law of large numbers. The proof can be completed by repeating the latter part of the proof of Lemma A.1 because therein only time points are involved for which $\varepsilon_{t\tau}(\Theta) = \varepsilon_{t\tau}^*(\Theta_1)$ holds.

The result of <u>Lemma A.2</u> holds with $l_{2T}(\Phi_2, \Omega)$ replaced by $l_{2T}^*(\Phi_2, \Omega)$. To see this, notice that $l_{2T}^*(\Phi_2, \Omega)$ can be interpreted as -2 times the logarithm of the Gaussian likelihood function associated with the (misspecified) regression model $\Delta y_t^{(*)} = \Phi_2 w_{2t}^{(0)} + \varepsilon_t^*$. Because $\varepsilon_t^* = \varepsilon_t + \underline{\gamma}_o \underline{d}_{t\tau_o}$, it is straightforward to check that the LS estimator of Φ_2 is consistent of order $O_p(T^{-1/2})$. Thus, the desired result can be seen by using standard arguments.

The result of <u>Lemma A.3</u> holds with $l_{1,\tau-1}(\Theta, \tau, \Omega)$ replaced by $l_{1,\tau-1}^*(\Theta_1, \tau, \Omega)$ and the set B_1 replaced by B_1^*. This is obvious because $l_{1,\tau-1}(\Theta, \tau, \Omega) = l_{1,\tau-1}^*(\Theta_1, \tau, \Omega)$ when $\tau < \tau_o$.

The result of <u>Lemma A.4</u> holds with the inequality replaced by

$$l_{1,\tau_o-1}^*(\Theta_1, \tau, \Omega) - l_{1,\tau-1}^*(\Theta_1, \tau, \Omega) \geq c_2 \|T^{1/2}\Psi_1\|^2 - a_{2T}\|T^{1/2}\Psi_1\|$$

and the set B_1 replaced by B_1^*. (Here c_2 and a_{2T} have the same properties as in Lemma A.4.) A proof of this result is obtained by following the proof of Lemma A.4 with the restriction $\underline{\gamma} = 0$ imposed. Because this means that only the upper left hand corner of the matrix on the l.h.s. of $(A.10)$ needs to be analyzed and the term L_{42} can be ignored, the desired result readily follows. (Note that $\varepsilon_{2t}^*(\Phi_2) = \varepsilon_{2t}(\Phi_2)$ for the time points considered.)

The result of <u>Lemma A.6</u> applies with $l_{1T}(\Theta, \tau, \Omega) - l_{1,\tau_o+p-1}(\Theta, \tau, \Omega)$ replaced with $l_{1T}^*(\Theta_1, \tau, \Omega) - l_{1,\tau_o+p-1}^*(\Theta_1, \tau, \Omega)$ and the set B_1 replaced by B_1^*. This is obvious because the two differences have identical values.

Because more substantial modifications are required for the remaining lemmas, we formulate new versions of them. For an analog of Lemma A.5 we introduce the notation

$$\zeta_{t\tau}^{(*)} = (d_{t\tau} - d_{t\tau_o})\delta_1^{(0)} - (\underline{\gamma}^{(0)} - \underline{\gamma}_o)\underline{d}_{t\tau_o}.$$

Clearly,

$$\varepsilon_{1t\tau}^*(\Theta_1) = -\Psi_2 w_{1t}^{(0)} - \zeta_{t\tau}^{(*)}, \quad t = \tau_o, \ldots, \tau_o + p - 1.$$

Lemma A.5*. There exists a constant $c_3 > 0$ such that, with probability approaching one and uniformly in $[T\underline{\lambda}] \leq \tau \leq \tau_o$ and $(\Theta_1, \tau, \Omega) \in B_1^*$,

$$l^*_{1,\tau_o+p-1}(\Theta_1, \tau, \Omega) - l^*_{1,\tau_o-1}(\Theta_1, \tau, \Omega) \geq c_3 \sum_{t=\tau_o}^{\tau_o+p-1} \|\zeta_{t\tau}^{(*)}\|^2 - a_{4T} \left(\sum_{t=\tau_o}^{\tau_o+p-1} \|\zeta_{t\tau}^{(*)}\|^2 \right)^{1/2} - a_{5T} - a^*_{5T},$$

where a_{4T} and a_{5T} are as in Lemma A.5 and $a^*_{5T} = c_3^*(\|\rho^{(0)}\| + \|\Phi_2 - \Phi_{2o}\|)\|\delta_o\|^2$ with $0 < c_3^* < \infty$.

Proof: In the same way as in the proof of Lemma A.5 we can write

$$l^*_{1,\tau_o+p-1}(\Theta_1, \tau, \Omega) - l^*_{1,\tau_o-1}(\Theta_1, \tau, \Omega) = L_5^* + L_6^*,$$

where L_5^* and L_6^* are as L_5 and L_6, respectively, except for being defined in terms of $\varepsilon^*_{1t\tau}(\Theta_1)$ and $\varepsilon^*_{2t}(\Phi_2)$. The analysis given for L_5 in the proof of Lemma A.5 applies also here with $\zeta_{t\tau}^{(0)}$ replaced by $\zeta_{t\tau}^{(*)}$. Thus, as an analog of $(A.11)$ we have

$$L_5^* \geq c_3 \sum_{t=\tau_o}^{\tau_o+p-1} \|\zeta_{t\tau}^{(*)}\|^2 - a_{41T} \left(\sum_{t=\tau_o}^{\tau_o+p-1} \|\zeta_{t\tau}^{(*)}\|^2 \right)^{1/2},$$

where $c_3 = \bar{\omega}^{-1} > 0$ and $a_{41T} = O_p(1)$ in the required uniform sense.

As for L_6^*, arguments similar to those used for L_6 give

$$|L_6^*| \leq 2\|\Omega^{-1}\| \left\| \Psi_2 \sum_{t=\tau_o}^{\tau_o+p-1} w_{1t}^{(0)} \varepsilon^*_{2t}(\Phi_2)' \right\| + 2\|\Omega^{-1}\| \left\| \sum_{t=\tau_o}^{\tau_o+p-1} \zeta_{t\tau}^{(*)} \varepsilon^*_{2t}(\Phi_2)' \right\| \overset{def}{=} L^*_{61} + L^*_{62}.$$

Using the identity $\varepsilon^*_{2t}(\Phi_2) = \varepsilon_{2t}(\Phi_2) + \underline{\gamma}_o \underline{d}_{t\tau_o}$ and arguments already used, it is seen that $L^*_{61} = O_p(1)$ in the required uniform sense. Thus, L^*_{61} can be used to define the quantity a_{5T} and it remains to study L^*_{62}. The latter norm in the definition of L^*_{62} can be expressed as

$$\left\| \sum_{t=\tau_o}^{\tau_o+p-1} \zeta_{t\tau}^{(*)} \varepsilon^*_{2t}(\Phi_2)' \right\| = \left\| \sum_{t=\tau_o}^{\tau_o+p-1} \zeta_{t\tau}^{(*)} (\varepsilon_{2t}(\Phi_2) + \underline{\gamma}_o d_{t\tau_o})' \right\| \leq \left\| \sum_{t=\tau_o}^{\tau_o+p-1} \zeta_{t\tau}^{(*)} \varepsilon_{2t}(\Phi_2)' \right\| + \left\| \sum_{t=\tau_o}^{\tau_o+p-1} \zeta_{t\tau}^{(*)} \underline{d}'_{t\tau_o} \underline{\gamma}'_o \right\|.$$

The first term in the last expression can be bounded from above in the same way as its counterpart in the proof of Lemma A.5. Thus, combining the preceding discussion we find that

$$|L_6^*| \leq a_{5T} + a_{42T} \left(\sum_{t=\tau_o}^{\tau_o+p-1} \|\zeta_{t\tau}^{(*)}\|^2 \right)^{1/2} + \left\| \sum_{t=\tau_o}^{\tau_o+p-1} \zeta_{t\tau}^{(*)} \underline{d}'_{t\tau_o} \underline{\gamma}'_o \right\|,$$

where $a_{42T} = O_p(1)$ in the required uniform sense. From this inequality and the lower bound obtained for L_5^* it can be seen that it suffices to show that the third term can be used to define a_{5T}^*. To justify this conclude from the definitions that

$$\left\| \sum_{t=\tau_o}^{\tau_o+p-1} \zeta_{t\tau}^{(*)} \underline{d}'_{t\tau_o} \underline{\gamma}'_o \right\| = \left\| (\underline{\gamma}^{(0)} - \underline{\gamma}_o) \sum_{t=\tau_o}^{\tau_o+p-1} \underline{d}_{t\tau_o} \underline{d}'_{t\tau_o} \underline{\gamma}'_o \right\| = \left\| (\underline{\gamma}^{(0)} - \underline{\gamma}_o) \underline{\gamma}'_o \right\|.$$

The definitions of $\underline{\gamma}^{(0)}$ and $\underline{\gamma}_o$ and simple calculations show that the last norm can be bounded by a quantity of the form a_{5T}^*. This completes the proof. □

Lemma A.7*. There exists a constant $c_5 > 0$ such that, with probability approaching one and uniformly in $[T\underline{\lambda}] \le \tau \le \tau_o - 1$ and $(\Theta_1, \tau, \Omega) \in B_1^*$,

$$\begin{aligned} &l_{1,\tau_o-1}^*(\Theta_1, \tau, \Omega) - l_{1,\tau-1}^*(\Theta_1, \tau, \Omega) \\ &\ge c_5 \textstyle\sum_{t=\tau}^{\tau_o-1} \|\zeta_{t\tau}^{(*)}\|^2 - \left(a_{7T} + a_{8T} \left(\frac{\tau_o-\tau}{T}\right)^{1/2} \|T^{1/2}\Psi_2\|\right) \left(\sum_{t=\tau}^{\tau_o-1} \|\zeta_{t\tau}^{(*)}\|^2\right)^{1/2} \\ &\quad -a_{9T}\|T^{1/2}\Psi_2\|, \end{aligned}$$

where $a_{iT} \ge 0$ and $a_{iT} = O_p(1)$ $(i = 7, 8, 9)$.

Proof: In the same way as in the proof of Lemma A.7 we can write,

$$l_{1,\tau_o-1}^*(\Theta_1, \tau, \Omega) - l_{1,\tau-1}^*(\Theta_1, \tau, \Omega) = L_7^* + L_8^*,$$

where L_7^* and L_8^* differ from L_7 and L_8, respectively, only in that $\varepsilon_{1t\tau}^*(\Theta_1)$ is used in place of $\varepsilon_{1t\tau}(\Theta)$. (Note that here $\varepsilon_{2t}^*(\Phi_2) = \varepsilon_{2t}(\Phi_2)$.) By the definitions we can also write

$$\varepsilon_{1t\tau}^*(\Theta_1) = -\Psi_2 w_{1t}^{(0)} - \zeta_{t\tau}^{(*)}, \quad t = \tau, \ldots, \tau_o - 1,$$

(cf. the corresponding representation of $\varepsilon_{1t\tau}(\Theta)$). An inspection of the proof of Lemma A.7 reveals that the analysis given for L_7 applies to L_7^* if only the quantity $\zeta_{t\tau}^{(0)}$ is replaced by $\zeta_{t\tau}^{(*)}$. Thus, we can write $L_7^* = L_{71}^* + L_{72}^* + L_{73}^*$, where $|L_{73}^*|$ and $L_{71}^* + L_{72}^*$ satisfy inequality $(A.13)$ and $(A.14)$, respectively, if $\zeta_{t\tau}^{(0)}$ on the r.h.s. is replaced by $\zeta_{t\tau}^{(*)}$.

Regarding L_8^* we can write $L_8^* = L_{81} + L_{82}^*$, where L_{81} satisfies $(A.15)$. Thus, we can conclude from the preceding discussion that it only remains to show that

$$|L_{82}^*| \le a_{7T} \left(\sum_{t=\tau}^{\tau_o-1} \|\zeta_{t\tau}^{(*)}\|^2 \right)^{1/2},$$

where a_{7T} is as stated in Lemma A.7*. To see this, notice that $\zeta_{t\tau}^{(*)} = \delta_1^{(0)}$, $t = \tau, \ldots, \tau_o - 1$, and therefore (cf. the definition of L_{82})

$$\begin{aligned} |L_{82}^*| &\leq 2\|\Omega^{-1}\| \left\| \sum_{t=\tau}^{\tau_o-1} \zeta_{t\tau}^{(*)} \varepsilon_{2t}(\Phi_2)' \right\| \\ &\leq 2\|\Omega^{-1}\| \|\delta_1^{(0)}\| \left\| \sum_{t=\tau}^{\tau_o-1} \varepsilon_{2t}(\Phi_2) \right\| \\ &= 2\|\Omega^{-1}\| \left(\sum_{t=\tau}^{\tau_o-1} \|\zeta_{t\tau}^{(*)}\|^2 \right)^{1/2} (\tau_o - \tau)^{-1/2} \left\| \sum_{t=\tau}^{\tau_o-1} \varepsilon_{2t}(\Phi_2) \right\|. \end{aligned}$$

This gives the desired result because the last norm is of order $O_p((\tau_o - \tau)^{1/2})$, as discussed in the proof of Lemma A.7. □

Because the quantity η does not appear in the results of Lemmas A.4* and A.7* the subsequent modifications of Lemmas A.8, A.9 and A.10 can be generalized accordingly.

Lemma A.8*. Let $\epsilon > 0$ and $B_2^* = \{(\Theta_1, \tau, \Omega) : \|T^{1/2-\eta}\Phi_1\|^2 + \|T^{1/2-\eta}\Psi_2\|^2 \leq \epsilon^2\}$, where $0 < \eta < \frac{1}{2}$. Then,

$$\inf_{(\Theta_1,\tau,\Omega)\in B_2^{*c}} l_T^*(\Theta_1, \tau, \Omega) - l_T^*(\Theta_{1o}, \tau_o, \Omega_o) > 0$$

with probability approaching one and uniformly in $[T\underline{\lambda}] \leq \tau \leq \tau_o$.

Proof: We can follow the proof of Lemma A.8. First, from Lemma A.2* and an analog of $(A.18)$ we can conclude that it suffices to show that, for some $\epsilon_* > 0$,

$$\inf_{(\Theta_1,\tau,\Omega)\in B_2^{*c}} T^{-2\eta} l_{1T}^*(\Theta_1, \tau, \Omega) \geq \epsilon_* \qquad (A.27)$$

with probability approaching one and uniformly in $[T\underline{\lambda}] \leq \tau \leq \tau_o$. From Lemma A.1* it then follows that $(A.27)$ can be established with B_2^{*c} replaced by $B_1^* \cap B_2^{*c}$. This in turn can be done separately for the sets B_{21}^* and B_{22}^* defined by modifying the definitions of B_{21} and B_{22}, respectively, in an obvious way.

Consider the l.h.s. of $(A.27)$ with B_2^{*c} replaced by B_{21}^*. From Lemmas A.4*, A.5* and $(A.17)$ we then find that $T^{-2\eta} l_{1T}^*(\Theta_1, \tau, \Omega)$ has a lower bound exactly of the same form as obtained for $T^{-2\eta} l_{1T}(\Theta, \tau, \Omega)$ in $(A.20)$. The only thing which may be worth noting here is that from Lemma A.5* we obtain the term a_{5T}^* which is not present in Lemma A.5 and the derivation of $(A.21)$. However, the stated result is still obtained because $T^{-2\eta}a^*_{5T} = o(1)$ uniformly in B^*_{21}. Because we have an analog of $(A.21)$ we can proceed in the same way as in the proof of Lemma A.8 and show that $(A.27)$ holds when B^{*c}_2 is replaced by B^*_{21}.

Now consider proving $(A.27)$ with B^{*c}_2 replaced by B^*_{22}. Here we can obtain an analog of $(A.22)$ by using Lemmas A.3*, A.5*, A.6* and A.7*. This yields a lower bound for $T^{-2\eta}l^*_{1T}(\Theta_1, \tau, \Omega)$ which differs from that obtained for $T^{-2\eta}l_{1T}(\Theta, \tau, \Omega)$ in $(A.22)$ only in that (i) $\zeta^{(0)}_{t\tau}$ is replaced by $\zeta^{(*)}_{t\tau}$, (ii) $a_{7T}\left(\frac{\tau_o-\tau}{T}\right)^\eta$ is replaced by $a_{7T}T^{-\eta}$ and (iii) the additional term $-T^{-2\eta}a^*_{5T}$ is included. However, observing that this additional term can be replaced by $o_p(1)$ and proceeding in the same way as after $(A.22)$ we can obtain the following analog of $(A.23)$:

$$\begin{aligned} T^{-2\eta}l^*_{1T}(\Theta_1, \tau, \Omega) \;\geq\;& c_1\|T^{1/2-\eta}\Phi_1\|^2 - T^{-\eta}a_{1T}\|T^{1/2-\eta}\Phi_1\| \\ & + c_{4T}(\tau)\|T^{1/2-\eta}\Psi_2\|^2 - a^*_{10T}(\tau)\|T^{1/2-\eta}\Psi_2\| - a^*_{11T} + o_p(1), \end{aligned}$$

where $c_{4T}(\tau)$ is as in $(A.23)$,

$$a^*_{10T}(\tau) = T^{-\eta}a_{6T} + T^{-\eta}a_{9T} + T^{-\eta}\frac{a_{7T}a_{8T}}{2c_5}\left(\frac{\tau_o-\tau}{T}\right)^{1/2}$$

and

$$a^*_{11T} = \frac{a^2_{7T}}{T^{2\eta}4c_5}.$$

Thus, it follows that we have, with probability approaching one and uniformly in B^*_{22}, $c_{4T}(\tau) \geq c_4/2$, $a^*_{10T}(\tau) = o_p(1)$ and $a^*_{11T} = o_p(1)$, and the situation becomes exactly the same as in the case of the set B^*_{21}. This completes the proof of Lemma A.8*. □

Lemma A.9*. Let $\epsilon > 0$ and $B^*_3 = \{(\Theta_1, \tau, \Omega) : \|T^{1/2-\eta}(\Phi_2 - \Phi_{2o})\| \leq \epsilon\}$, where η is the same as in Lemma A.8*. Then,

$$\inf_{(\Theta_1,\tau,\Omega)\in B^{*c}_3} l^*_T(\Theta_1, \tau, \Omega) - l^*_T(\Theta_{1o}, \tau_o, \Omega_o) > 0$$

with probability approaching one and uniformly in $[T\underline{\lambda}] \leq \tau \leq \tau_o$.

Proof: The proof can be obtained in the same way as that of Lemma A.9. Instead of using results obtained in the proof of Lemma A.8 we, of course, use corresponding results discussed in the proof of Lemma A.8*. Details are omitted. □

Lemma A.10*. Let $B_4^* = \{(\Theta_1, \tau, \Omega) : (\tau_o - \tau)\|\delta_1^{(0)}\|^{2/(1-2\eta)} \leq M\}$, where $\tau < \tau_o$ and η is the same as in Lemma A.8*. Then, there exists a real number M_0 such that, for all $M \geq M_0$,

$$\inf_{(\Theta_1,\tau,\Omega)\in B_4^{*c}} l_T^*(\Theta_1, \tau, \Omega) - l_T^*(\Theta_{1o}, \tau_o, \Omega_o) > 0$$

with probability approaching one and uniformly in $[T\underline{\lambda}] \leq \tau \leq \tau_o - 1$.

Proof: In the same way as in the proof of Lemma A.10 it can first be seen that it suffices to show that $l_{1T}^*(\Theta_1, \tau, \Omega)$ satisfies an analog of $(A.24)$. This in turn can be done by using Lemmas A.3*, A.5*, A.6* and A.7* to obtain a lower bound for $l_{1T}^*(\Theta_1, \tau, \Omega)$. This lower bound divided by $T^{2\eta}$ was discussed in the proof of Lemma A.8*. From that discussion it can be concluded that

$$\begin{aligned} l_{1T}^*(\Theta_1, \tau, \Omega) \;\; \geq \;\; & c_3 \sum_{t=\tau_o}^{\tau_o+p-1} \|\zeta_{t\tau}^{(*)}\|^2 - a_{4T}\left(\sum_{t=\tau_o}^{\tau_o+p-1} \|\zeta_{t\tau}^{(*)}\|^2\right)^{1/2} + c_5 \sum_{t=\tau}^{\tau_o-1} \|\zeta_{t\tau}^{(*)}\|^2 \\ & - \left(a_{7T} + a_{8T}(\tau_o - \tau)^{1/2}\|\Psi_2\|\right)\left(\sum_{t=\tau}^{\tau_o-1} \|\zeta_{t\tau}^{(*)}\|^2\right)^{1/2} - a_{5T}^* + O_p(1). \end{aligned}$$

This holds uniformly in $B_1^* \cap B_2^* \cap B_3^* \cap B_4^{*c}$ and $[T\underline{\lambda}] \leq \tau \leq \tau_o - 1$. By $(A.17)$, the first two terms on the r.h.s. can be bounded from below by $-a_{4T}^2/4c_3$ and in what follows they will be absorbed in the term $O_p(1)$. The term $(\tau_o - \tau)^{1/2}\|\Psi_2\|$ can be bounded from above by $\epsilon_1(\tau_o - \tau)^\eta$, where $\epsilon_1 > 0$ (cf. the proof of Lemma A.10). Further, from the definition we have $a_{5T}^* = O(1)$ (see Lemma A.5*) and, hence, we can absorb a_{5T}^* in the term $O_p(1)$. Taking all these facts into account and observing that $\zeta_{t\tau}^{(*)} = \delta_1^{(0)}$ for $t = \tau, \ldots, \tau_o - 1$, we can write the preceding inequality as

$$\begin{aligned} l_{1T}^*(\Theta_1, \tau, \Omega) \;\; & \geq \;\; c_5(\tau_o - \tau)\|\delta_1^{(0)}\|^2 - a_T^*(\tau_o - \tau)^{\eta+1/2}\|\delta_1^{(0)}\| + O_p(1) \\ & = \;\; (\tau_o - \tau)\|\delta_1^{(0)}\|^2\left(1 - \frac{a_T^*(\tau_o - \tau)^{\eta-1/2}}{c_5\|\delta_1^{(0)}\|}\right) + O_p(1). \end{aligned}$$

Here $a_T^* = O_p(1)$, which as well as the whole result, holds uniformly in $B_1^* \cap B_2^* \cap B_3^* \cap B_4^{*c}$ (cf. $(A.25)$). Thus, the rest of the proof is similar to that of Lemma A.10. □

The results of the adjusted Lemmas can be modified in a straightforward way to the case $\tau \geq \tau_o$. Using Lemma A.10* and its modification in conjunction with arguments similar to those in the proof of Lemma A.11 we can now prove Theorem 3.1(ii). In the same way as in $(A.26)$, we first find that

$$(\tau_o-\tau)^{(1/2)-\eta}\|\delta_1^{(0)}\| \geq \left((\tau_o-\tau)\|\delta_{1o}\|^{2/(1-2\eta)}\right)^{(1/2)-\eta}\left(1-\frac{\|\delta_1^{(0)}-\delta_{1o}\|}{\|\delta_{1o}\|}\right)$$

and furthermore that an analog of Lemma A.11 holds with the set B_5 replaced by $B_5^* = \{(\Theta_1,\tau,\Omega) : |\tau-\tau|\|\delta_{1o}\|^{2/(1-2\eta)} \leq M\}$, where η is the same as in Lemma A.8*. Theorem 3.1(ii) then follows from this in the same way as Theorem 3.1(i) followed from Lemma A.11.

Proof of Theorem 4.1

Because we do not need values from $\mathcal{T}$ other than $\hat{\tau}$ and the true value in the following, we denote the latter again by τ and hence we drop the subscript 'o' for convenience.

Properties of RR Estimators

We shall first show that the RR estimators of the parameters α, β, Γ_j $(j=1,\dots,p-1)$ and Ω discussed in Section 2 satisfy appropriate consistency properties (cf. Lemma 2.1 of S&L). These estimators are based on equation (2.5) with the unknown break date τ replaced by the estimator $\hat{\tau}$. This replacement changes the error term ε_t to

$$\varepsilon_{t\hat{\tau}} = \varepsilon_t + \Pi\delta(d_{t-1,\hat{\tau}} - d_{t-1,\tau}) - \sum_{j=0}^{p-1}\gamma_j^*(\Delta d_{t-j,\hat{\tau}} - \Delta d_{t-j,\tau}). \qquad (A.28)$$

Defining

$$y_{t\hat{\tau}} = y_t - \mu_0 - \mu_1 t - \delta d_{t\hat{\tau}} = x_t - \delta(d_{t\hat{\tau}} - d_{t\tau}) \qquad (A.28a)$$

we can consider the following counterpart of (2.5):

$$\Delta y_{t\hat{\tau}} = \nu+\alpha(\beta' y_{t-1,\hat{\tau}} - \phi(t-1) - \theta d_{t-1,\hat{\tau}}) + \sum_{j=1}^{p-1}\Gamma_j\Delta y_{t-j,\hat{\tau}} + \sum_{j=0}^{p-1}\gamma_j^*\Delta d_{t-j,\hat{\tau}} + \varepsilon_{t\hat{\tau}},\ t=p+1,p+2,\dots, \qquad (A.28b)$$

where $\nu=0$, $\phi=0$, $\theta=0$ and $\gamma_j^*=0$ $(j=0,\dots,p-1)$. It will be shown that using the estimator $\hat{\tau}$ in $(A.28b)$ instead of the true break date τ has no effect on the asymptotic properties of the RR estimators. When $\delta_1 \neq 0$, the proof makes use of the following general results.

Lemma A. Let J_{tT} $(t = 1, \ldots, T)$ be a (possibly) random vector such that $\max_{1 \le t \le T} \|J_{tT}\| = O_p(1)$ and J_t a vector valued stochastic process satisfying $\sup_t E\|J_t\| < \infty$. Then, if $\hat{\tau} = \tau + O_p(1)$,

(i) $T^{-1/2} \sum_{t=p+1}^{T} J_{tT}(d_{t\hat{\tau}} - d_{t\tau}) = o_p(1)$

(ii) $T^{-1/2} \sum_{t=p+1}^{T} J_t(d_{t\hat{\tau}} - d_{t\tau}) = o_p(1)$.

Proof: The first result is an immediate consequence of the assumptions which imply that the l.h.s. of the equality is dominated by $O_p(T^{-1/2})|\hat{\tau} - \tau| = o_p(1)$. To prove the second result, let $\mathbf{I}(\cdot)$ stand for the indicator function and, for $\epsilon > 0$, write

$$\begin{aligned} T^{-1/2} \sum_{t=p+1}^{T} J_t(d_{t\hat{\tau}} - d_{t\tau}) &= T^{-1/2} \sum_{t=p+1}^{T} J_t(d_{t\hat{\tau}} - d_{t\tau})\mathbf{I}(|\hat{\tau} - \tau| > T^{1/4}\epsilon) \\ &\quad + T^{-1/2} \sum_{t=p+1}^{T} J_t(d_{t\hat{\tau}} - d_{t\tau})\mathbf{I}(|\hat{\tau} - \tau| \le T^{1/4}\epsilon) \\ &\overset{def}{=} A_{1T} + A_{2T}. \end{aligned}$$

Now, since $\hat{\tau} - \tau = O_p(1)$, we have for every $\epsilon > 0$,

$$Pr\{\|A_{1T}\| > \epsilon\} \le Pr\{|\hat{\tau} - \tau| > T^{1/4}\epsilon\} \to 0.$$

Thus, $A_{1T} = o_p(1)$. As for A_{2T}, notice that for every $\epsilon > 0$,

$$\|A_{2T}\| \le T^{-1/2} \sum_{t=p+1}^{T} \|J_t\| |d_{t\hat{\tau}} - d_{t\tau}| \mathbf{I}(|\hat{\tau} - \tau| \le T^{1/4}\epsilon) \le T^{-1/2} \sum_{|t-\tau| \le [T^{1/4}\epsilon]} \|J_t\|.$$

Hence,

$$E\|A_{2T}\| \le const. \times T^{(1/4)-(1/2)}\epsilon \to 0.$$

This implies that $A_{2T} = o_p(1)$ and the stated result follows. □

Lemma A can be used to show that, when $\delta_1 \neq 0$, the replacement of the true break date τ by the estimator $\hat{\tau}$ has no effect on the asymptotic properties of various second sample moments. The first result can be used with J_{tT} given by $(T-p)^{-1/2} \beta_{\perp}' y_{t\hat{\tau}}$ and $(t-1)/(T-p)$, whereas in the second result typical choices of J_t are $\Delta d_{t-j,\hat{\tau}}$, $\beta' y_{t\hat{\tau}}$ and $d_{t-1,\hat{\tau}}$ (see $(A.28)$). When $\delta_1 = 0$, a different proof is needed to obtain the desired consistency results. In this case use is made of the fact that the step dummy in equation (2.5) becomes redundant and the effect of the impulse dummies vanishes asymptotically. Now we can prove the necessary consistency properties of the RR estimators discussed above.

Lemma B. Let $\tilde{\beta}_\xi = \tilde{\beta}(\xi'\tilde{\beta})^{-1}$ and $\tilde{\alpha}_\xi = \tilde{\alpha}\tilde{\beta}'\xi$ where $\xi' = (\beta'\beta)^{-1}\beta'$. Then, $\tilde{\beta}_\xi = \beta + O_p(T^{-1})$, $\tilde{\alpha}_\xi = \alpha + O_p(T^{-1/2})$, $\tilde{\Gamma}_j = \Gamma_j + O_p(T^{-1/2})$ $(j = 1, \ldots, p-1)$ and $\tilde{\Omega} = \Omega + O_p(T^{-1/2})$.

Proof: First we introduce some notation. Define $\mathbf{x}_{t\hat{\tau}} = [y'_{t-1,\hat{\tau}} : (t-1) : d_{t-1,\hat{\tau}}]'$ and $\mathbf{p}_{1t\hat{\tau}} = [1 : \Delta y'_{t-1,\hat{\tau}} : \cdots : \Delta y'_{t-p+1,\hat{\tau}}]'$. Then equation $(A.28b)$ can be expressed as

$$\Delta y_{t\hat{\tau}} = \alpha\psi'\mathbf{x}_{t\hat{\tau}} + \Xi\mathbf{p}_{t\hat{\tau}} + \varepsilon_{t\hat{\tau}}, \quad t = p+1, p+2, \ldots, \qquad (A.29)$$

where $\mathbf{p}_{t\hat{\tau}} = [\mathbf{p}'_{1t\hat{\tau}} : \mathbf{d}'_{t\hat{\tau}}]'$, $\psi' = [\beta' : -\phi : -\theta]$ and $\Xi = [\Xi_1 : \boldsymbol{\gamma}^*]$ with $\Xi_1 = [\nu : \Gamma_1 : \ldots : \Gamma_{p-1}]$. Here $\mathbf{d}_{t\tau} = [\Delta d_{t\tau}, \ldots, \Delta d_{t-p+1,\tau}]$ and $\boldsymbol{\gamma}^* = [\gamma_0^*, \ldots, \gamma_{p-1}^*]$. The RR estimators of α, ψ and Ω can be obtained as follows. Define

$$S_{00\hat{\tau}} = T^{-1}\sum_{t=p+1}^{T} \Delta y_{t\hat{\tau}}\Delta y'_{t\hat{\tau}} - T^{-1}\sum_{t=p+1}^{T} \Delta y_{t\hat{\tau}}\mathbf{p}'_{t\hat{\tau}}\left(\sum_{t=p+1}^{T}\mathbf{p}_{t\hat{\tau}}\mathbf{p}'_{t\hat{\tau}}\right)^{-1}\sum_{t=p+1}^{T}\mathbf{p}_{t\hat{\tau}}\Delta y'_{t\hat{\tau}},$$

$$S_{01\hat{\tau}} = S'_{10\hat{\tau}} = T^{-1}\sum_{t=p+1}^{T} \Delta y_{t\hat{\tau}}\mathbf{x}'_{t\hat{\tau}} - T^{-1}\sum_{t=p+1}^{T} \Delta y_{t\hat{\tau}}\mathbf{p}'_{t\hat{\tau}}\left(\sum_{t=p+1}^{T}\mathbf{p}_{t\hat{\tau}}\mathbf{p}'_{t\hat{\tau}}\right)^{-1}\sum_{t=p+1}^{T}\mathbf{p}_{t\hat{\tau}}\mathbf{x}'_{t\hat{\tau}}$$

and

$$S_{11\hat{\tau}} = T^{-1}\sum_{t=p+1}^{T} \mathbf{x}_{t\hat{\tau}}\mathbf{x}'_{t\hat{\tau}} - T^{-1}\sum_{t=p+1}^{T} \mathbf{x}_{t\hat{\tau}}\mathbf{p}'_{t\hat{\tau}}\left(\sum_{t=p+1}^{T}\mathbf{p}_{t\hat{\tau}}\mathbf{p}'_{t\hat{\tau}}\right)^{-1}\sum_{t=p+1}^{T}\mathbf{p}_{t\hat{\tau}}\mathbf{x}'_{t\hat{\tau}}.$$

As is well-known, the RR estimator of ψ is based on the eigenvectors corresponding to the r largest eigenvalues of the determinantal equation

$$\det(\lambda S_{11\hat{\tau}} - S_{10\hat{\tau}}S_{00\hat{\tau}}^{-1}S_{01\hat{\tau}}) = 0. \qquad (A.30)$$

When the RR estimator of ψ is available, those of the other parameters can be obtained by replacing ψ by its estimator in $(A.29)$ and applying LS to the obtained auxiliary regression model.

First consider the case $\delta_1 \neq 0$. The idea of the following proof is to show that the consistency properties of the RR estimators of ψ, α and Ω discussed above are the same as in the case where the true value of the break date is used instead of the estimator $\hat{\tau}$. Once this has been shown, the desired results follow from S&L.

In the same way as in S&L, we follow the proof of Lemma 13.1 of Johansen (1995) and transform equation $(A.30)$ to

$$\det(\lambda A_T' S_{11\hat{\tau}} A_T - A_T' S_{10\hat{\tau}} S_{00\hat{\tau}}^{-1} S_{01\hat{\tau}} A_T) = 0, \tag{A.31}$$

where

$$A_T = \begin{bmatrix} \beta & T^{-1/2}\beta_\perp(\beta_\perp'\beta_\perp)^{-1} & 0 & 0 \\ 0 & 0 & T^{-1} & 0 \\ 0 & 0 & 0 & (\frac{T}{T-\tau})^{1/2} \end{bmatrix}$$

and, consequently,

$$A_T'\mathbf{x}_{t\hat{\tau}} = \begin{bmatrix} \beta' y_{t-1,\hat{\tau}} \\ T^{-1/2}(\beta_\perp'\beta_\perp)^{-1}\beta_\perp' y_{t-1,\hat{\tau}} \\ T^{-1}(t-1) \\ (\frac{T}{T-\tau})^{1/2} d_{t-1,\hat{\tau}} \end{bmatrix}.$$

Note that

$$T^{-1/2}\sum_{t=\tau}^{\tau_o-1} w_t^{(0)} \Delta d_{t-i,\tau} = o_p\left(T^{\eta-1/2}\right), \quad i = 0,\dots,p-1, \tag{A.31a}$$

uniformly in $[T\underline{\lambda}] \leq \tau < [T\bar{\lambda}]$. For the first component of $w_t^{(0)}$, that is, for $w_{1t}^{(0)}$, this is justified in the proof of Lemma A.4 (see the Equation $(A.9a)$). Regarding $w_{2t}^{(0)}$, consider its first component $u_{t-1}^{(0)}$. Since $u_{t-1}^{(0)} = \beta' x_{t-1}$ is stationary we can use an argument similar to that for (A.14) of Saikkonen & Lütkepohl (2002) and conclude that $\max_{1\leq t\leq T}\left\|u_{t-1}^{(0)}\right\| = o_p\left(T^{\eta}\right)$, which gives the desired result. A similar reasoning applies to the remaining components of $w_{2t}^{(0)}$, that is, to Δx_{t-j} $(j = 1,...,p-1)$, and, hence, we have established $(A.31a)$.

Now, using $(A.31a)$, $(A.28a)$ and Lemma A we can show that

$$S_{00\hat{\tau}} = S_{00\tau} + o_p(1), \tag{A.32}$$

$$S_{01\hat{\tau}} A_T = S_{01\tau} A_T + o_p(1) \tag{A.33}$$

and

$$A_T' S_{11\hat{\tau}} A_T = A_T' S_{11\tau} A_T + o_p(1). \tag{A.34}$$

Details are straightforward but somewhat tedious and will be omitted.

From $(A.32) - (A.34)$ and the proof of Lemma 2.1 of S&L we can conclude that the estimators $\tilde{\beta}_\xi$, $\tilde{\phi}$ and $\tilde{\theta}$ are consistent of orders $o_p(T^{-1/2})$, $o_p(T^{-1})$ and $o_p(1)$, respectively. Because $\psi'\mathbf{x}_{t\hat{\tau}} = \beta' y_{t-1,\hat{\tau}}$, arguments similar to those used to obtain $(A.32)-(A.34)$ combined with these consistency results also readily show that the consistency proof for the estimators $\tilde{\alpha}_\xi$, $\tilde{\Gamma}_j$ $(j = 0, \dots, p-1)$ and $\tilde{\Omega}$ can be reduced to that with τ known and, hence, the stated result follows from S&L (cf. the definition of these estimators after $(A.30)$).

To obtain the stated orders of consistency, we follow S&L and note that the first order conditions for $\tilde{\psi}$, the RR estimator of ψ, can be expressed as

$$0 = \tilde{\alpha}_\xi' \tilde{\Omega}^{-1} \left(S_{\varepsilon 1 \hat{\tau}} B_T - \tilde{\alpha}_\xi U_T' [T^{-1} B_T' S_{11\hat{\tau}} B_T] - (\tilde{\alpha}_\xi - \alpha) \psi' S_{11\hat{\tau}} B_T \right), \qquad (A.35)$$

where $S_{\varepsilon 1 \hat{\tau}} = S_{01\hat{\tau}} - \alpha \psi' S_{11\hat{\tau}}$,

$$U_T = \begin{bmatrix} T \beta_\perp' \tilde{\beta}_\xi \\ T^{3/2} \tilde{\phi}' \\ (T-\tau)^{1/2} \tilde{\theta}' \end{bmatrix}$$

and

$$B_T = \begin{bmatrix} \beta_\perp (\beta_\perp' \beta_\perp)^{-1} & 0 & 0 \\ 0 & T^{-1/2} & 0 \\ 0 & 0 & \frac{T}{(T-\tau)^{1/2}} \end{bmatrix}.$$

Now, notice that B_T is formed by the last $(n - r + 2)$ columns of $T^{1/2} A_T$ and recall that $\psi' \mathbf{x}_{t\hat{\tau}} = \beta' y_{t-1,\hat{\tau}}$. Using these facts, $(A.31a)$, $(A.28a)$ and Lemma A, it is straightforward to proceed in the same way as in the above consistency proof and show that replacing $\hat{\tau}$ by τ on the r.h.s. of $(A.35)$ causes an error of order $o_p(1)$. Thus, from the proof of Lemma 2.1 of S&L we can conclude that $U_T = O_p(1)$ and furthermore that $\tilde{\beta}_\xi = \beta + O_p(T^{-1})$.

To obtain the remaining orders of consistency, notice that $U_T = O_p(1)$ also implies that the estimator $\tilde{\psi}$ satisfies $A_T^{-1}(\tilde{\psi} - \psi) = O_p(T^{-1/2})$ (cf. S&L). Using this fact, the expression for the error term $\varepsilon_{t\hat{\tau}}$ given in $(A.28)$ and arguments used in connection with equations $(A.32) - (A.35)$ it can be seen that replacing the true break date in (A.11) of S&L by the estimator $\hat{\tau}$ has no effect on the orders of consistency obtained for the estimators $\tilde{\alpha}_\xi$, $\tilde{\Gamma}_j$ $(j = 0, \dots, p-1)$, and $\tilde{\Omega}$. Thus, the desired result follows from Lemma 2.1 of S&L. This completes the proof when $\delta_1 \neq 0$.

Now consider the case $\delta_1 = 0$ for which we cannot use Lemma A because $\hat{\tau} = \tau + O_p(1)$ does not hold and even $\hat{\tau}/N$ is not a consistent estimator of the sample fraction λ. However, we still have $(A.32)$ and $(A.33)$, as we shall now demonstrate. The former is simple because, by $(A.28a)$ and the definition of $\mathbf{p}_{t\hat{\tau}}$, the matrix $S_{00\hat{\tau}}$ only depends on the impulse dummies $\Delta d_{t-j,\hat{\tau}}$ and $\Delta d_{t-j,\tau}$ $(j = 0, \dots, p-1)$ but not on the corresponding step dummies. Thus, because $(A.31a)$ still holds, we can proceed in the same way as in the proof of Lemma 2.1 of S&L and show that replacing $\mathbf{p}_{t\hat{\tau}}$ in the definition of $S_{00\hat{\tau}}$ first by $\mathbf{p}_{1t\hat{\tau}}$ and then further by $\mathbf{p}_{1t\tau}$ causes an error of order $o_p(1)$. Because in $S_{00\tau}$ the vector $\mathbf{p}_{t\tau}$ can similarly be replaced by $\mathbf{p}_{1t\tau}$ (see S&L), $(A.32)$ follows.

Now consider $(A.33)$. Using $(A.31a)$ and $(A.28a)$ we can again readily show that replacing $\mathbf{p}_{t\hat{\tau}}$ and $\mathbf{p}_{t\tau}$ in the definitions of $S_{01\hat{\tau}}A_T$ and $S_{01\tau}A_T$ by $\mathbf{p}_{1t\hat{\tau}}$ and $\mathbf{p}_{1t\tau}$, respectively, causes an error of order $o_p(1)$ (cf. the justification of $(A.38)$ below). Next, an application of $(A.31a)$ and $(A.28a)$ gives

$$T^{-1}\sum_{t=p+1}^{T}\mathbf{p}_{1t\hat{\tau}}\mathbf{p}_{1t\hat{\tau}}' = T^{-1}\sum_{t=p+1}^{T}\mathbf{p}_{1t\tau}\mathbf{p}_{1t\tau}' + o_p(1) \tag{A.36}$$

and

$$T^{-1}\sum_{t=p+1}^{T}\Delta y_{t\hat{\tau}}\mathbf{p}_{1t\hat{\tau}}' = T^{-1}\sum_{t=p+1}^{T}\Delta y_{t\tau}\mathbf{p}_{1t\tau}' + o_p(1). \tag{A.37}$$

In addition to these results, we have

$$T^{-1}\sum_{t=p+1}^{T}\Delta y_{t-j,\hat{\tau}}\mathbf{x}_{t\hat{\tau}}'A_T = T^{-1}\sum_{t=p+1}^{T}\Delta y_{t-j,\tau}\mathbf{x}_{t\tau}'A_T + o_p(1) \quad (j=0,\dots,p-1). \tag{A.38}$$

To justify this relation, notice that now $\beta' y_{t-1,\hat{\tau}} = \beta' x_{t-1}$ which shows that $(A.38)$ holds for the first r columns of the involved matrices. For the next $n-r$ columns the result is simple because the sums are divided by $T^{3/2}$. The same can be said about the $(n+1)$th column. That the stated result holds for the last column can be seen from

$$\begin{aligned} T^{-1}\textstyle\sum_{t=p+1}^{T}\Delta y_{t-j,\hat{\tau}}d_{t-1,\tau} &= T^{-1}\textstyle\sum_{t=\tau+1}^{T}\Delta x_{t-j} + o_p(1) \\ &= o_p(1) \qquad (j=0,\dots,p-1) \end{aligned} \tag{A.39}$$

uniformly in $[T\underline{\lambda}] \le \tau \le [T\bar{\lambda}]$. Here the former equality is an immediate consequence of $(A.28a)$ and the latter follows because an invariance principle applies to partial sums of Δx_t.

In addition to $(A.36)$ - $(A.38)$, we also need to consider the matrix $T^{-1}\sum_{t=p+1}^{T}\mathbf{p}_{1t\hat{\tau}}\mathbf{x}_{t\hat{\tau}}'A_T$. Arguments used above show that replacing $\hat{\tau}$ here by τ causes an error of order $o_p(1)$ except for the element in the first row and last column which is

$$\left(\frac{T}{T-\tau}\right)^{1/2} T^{-1}\sum_{t=p+1}^{T} d_{t-1,\hat{\tau}} = O_p(1).$$

However, the contribution of this element to the matrix on the l.h.s. of $(A.33)$ is of order $o_p(1)$. Because in the definitions of $S_{01\hat{\tau}}$ and $S_{01\tau}$ we can replace $\mathbf{p}_{t\hat{\tau}}$ and $\mathbf{p}_{t\tau}$ by $\mathbf{p}_{1t\hat{\tau}}$ and

$\mathbf{p}_{1t\tau}$, respectively, this follows from the following two facts. (i) The first matrix on the r.h.s. of $(A.36)$ is asymptotically block diagonal with blocks defined after the first row and first column. (ii) The first column of the first matrix on the r.h.s. of $(A.37)$ is of order $o_p(1)$. Both of these results follow from $(A.28a)$ and the fact that Δx_t obeys a weak law of large numbers. Taking these results together, we can thus conclude that $(A.33)$ also holds when $\delta_1 = 0$.

Since $(A.32)$ and $(A.33)$ hold we can write

$$A_T' S_{10\hat{\tau}} S_{00\hat{\tau}}^{-1} S_{01\hat{\tau}} A_T = A_T' S_{10\tau} S_{00\tau}^{-1} S_{01\tau} A_T + o_p(1) = \begin{bmatrix} \Sigma_{\beta 0}\Sigma_{00}^{-1}\Sigma_{0\beta} & 0 \\ 0 & 0 \end{bmatrix} + o_p(1), \qquad (A.40)$$

where the latter equality is justified by (A.5) of S&L and the notation is as explained therein. The partition is after the first r rows and columns.

Now consider the matrix $A_T' S_{11\hat{\tau}} A_T$. In the present case we do not have $(A.34)$ which, given $(A.40)$, would be sufficient for the consistency of the RR estimator of ψ. However, $(A.34)$ is not necessary. Write

$$A_T' S_{11\hat{\tau}} A_T = \begin{bmatrix} \bar{S}_{11\hat{\tau}}^{11} & \bar{S}_{11\hat{\tau}}^{12} \\ \bar{S}_{11\hat{\tau}}^{21} & \bar{S}_{11\hat{\tau}}^{22} \end{bmatrix},$$

where the partition is after the first r rows and columns. Because $\beta' y_{t-1,\hat{\tau}} = \beta' x_{t-1}$ by $(A.28a)$, we can show that $\bar{S}_{11\hat{\tau}}^{11} = \Sigma_{\beta\beta} + o_p(1)$ with $\Sigma_{\beta\beta}$ as in S&L and $\bar{S}_{11\hat{\tau}}^{12} = o_p(1)$. The required arguments are based on $(A.31a)$ and $(A.28a)$ in the same way as in the case of $(A.32)$, $(A.33)$ and $(A.39)$. Thus, because we also have $\bar{S}_{11\hat{\tau}}^{21} = o_p(1)$, the above discussion and $(A.40)$ show that equation $(A.31)$ is to order $o_p(1)$ identical to

$$\det(\lambda\Sigma_{\beta\beta} - \Sigma_{\beta 0}\Sigma_{00}^{-1}\Sigma_{0\beta}) \det(\lambda \bar{S}_{11\hat{\tau}}^{22}) = 0. \qquad (A.41)$$

This implies that the consistency proof given in Johansen (1995, pp. 180 - 181) applies if, with probability approaching one, $\lambda_{\min}(\bar{S}_{11\hat{\tau}}^{22}) \geq \epsilon$ for some $\epsilon > 0$. This, however, is the case because arguments similar to those used below $(A.10)$ show that $\bar{S}_{11[T\lambda]}^{22}$ converges weakly in $D([\underline{\lambda}, \bar{\lambda}])$ to a (a.s.) positive definite limit. In fact, arguments used to arrive at $(A.41)$ show that $\bar{S}_{11[T\lambda]}^{22}$ is to order $o_p(1)$ identical to a demeaned version of the matrix of second sample moments formed from the last $n - r + 2$ components of $A_T' \mathbf{x}_{t[T\lambda]}$.

Thus, in the same way as in the case $\delta_1 \neq 0$ we can conclude that the estimators $\tilde{\beta}_\xi, \tilde{\phi}$ and $\tilde{\theta}$ are consistent of orders $o_p(T^{-1/2}), o_p(T^{-1})$ and $o_p(1)$, respectively. After this, the consistency of the estimators $\tilde{\alpha}_\xi$, $\tilde{\Gamma}_j$ $(j = 1, \ldots, p-1)$ and $\tilde{\Omega}$ can be proved in the same way as in S&L. Since the second term in the definition of $\varepsilon_{t\hat{\tau}}$ in $(A.28)$ now vanishes the employed arguments consist of similar analyses of second sample moments already carried out at previous steps of the proof. Details are therefore omitted.

To obtain the orders of consistency, consider the first order conditions $(A.35)$ and notice that now $\psi'\mathbf{x}_{t\hat{\tau}} = \beta' x_{t-1}$. Using this fact it can be seen that $\psi' S_{11\hat{\tau}} B_T = O_p(1)$ and $S_{\varepsilon 1\hat{\tau}} B_T = O_p(1)$. The employed arguments are similar to those used in the above consistency proof and in the proof of Theorem 3.1. To give an example of the latter, note that in order to prove $\psi' S_{11\hat{\tau}} B_T = O_p(1)$ we need to show that

$$(T-\tau)^{-1/2} \sum_{t=p+1}^{T} \beta' x_{t-1} d_{t-1,\hat{\tau}} = (T-\tau)^{-1/2} \sum_{t=\hat{\tau}+1}^{T} \beta' x_{t-1} = O_p(1).$$

This, however, follows from the Hájek-Rényi inequality given in Proposition 1 of Bai (1994). Omitting other details we need to note that $\lambda_{\min}(T^{-1} B_T S_{11\hat{\tau}} B_T) = \lambda_{\min}(\bar{S}^{22}_{11\hat{\tau}})$, which is asymptotically bounded away from zero, as noticed after $(A.41)$. From $(A.35)$ and what has been said above we now find that $U_T = O_p(1)$. In the same way as in the case $\delta_1 \neq 0$ this implies $\tilde{\beta}_\xi = \beta + O_p(T^{-1})$ and more generally $A_T^{-1}(\tilde{\psi} - \psi) = O_p(T^{-1/2})$. After this the remaining orders of consistency can be obtained in the same way as in S&L. Details are again based on arguments already employed in the foregoing. This completes the proof of Lemma B. □

Properties of Estimators Considered by S&L

If the estimators of the parameters of the deterministic part proposed by S&L are used with the known break date replaced by an estimator $\hat{\tau} = \tau + O_p(1)$, it can be shown that Assumption 1 is satisfied.

S&L propose to estimate the parameters μ_i $(i = 0, 1)$ and δ by a feasible GLS approach. Defining

$$a_{0t} = \begin{cases} 1 & \text{for} \quad t \geq 1 \\ 0 & \text{for} \quad t \leq 0 \end{cases} \qquad \text{and} \qquad a_{1t} = \begin{cases} t & \text{for} \quad t \geq 1 \\ 0 & \text{for} \quad t \leq 0 \end{cases}$$

and multiplying (2.1) from the left by $QA(L)$ gives

$$QA(L)y_t = QH_{0t}\mu_0 + QH_{1t}\mu_1 + QK_{t\tau}\delta + Q\varepsilon_t, \qquad t = 1, 2, \ldots,$$

where $H_{it} = A(L)a_{it}$ $(i = 0, 1)$, $K_{t\tau} = A(L)d_{t\tau}$ and

$$Q = [\Omega^{-1}\alpha(\alpha'\Omega^{-1}\alpha)^{-1/2} : \alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1/2}]$$

has the property $QQ' = \Omega^{-1}$. Therefore the error term $Q\varepsilon_t$ has an identity covariance matrix. Thereby we have a transformation which results in a (multivariate) regression model with standard properties of the error term, as required in GLS estimation.

Of course, the above transformation is not feasible because it involves unknown parameters. For a feasible transformation S&L propose estimators of the parameters α, β, Γ_j $(j = 1, \dots, p-1)$ and Ω obtained by a RR regression of (2.5) using the rank which is specified in the null hypothesis of the cointegrating rank test. We denote these estimators by $\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{\Gamma}_j$ and $\tilde{\Omega}$. They can be used to compute estimators $\tilde{A}_j$ for the coefficient matrices A_j via the formulas given before Eq. (2.5). We define $\tilde{A}(L) = I_n - \tilde{A}_1 L - \cdots - \tilde{A}_p L^p$, $\tilde{H}_{it} = \tilde{A}(L)a_{it}$ $(i = 0, 1)$ and $\tilde{K}_{t\hat{\tau}} = \tilde{A}(L)d_{t\hat{\tau}}$. A suitable estimator of Q can be obtained by forming $\tilde{\alpha}_\perp$ from $\tilde{\alpha}$ and replacing Ω, α and $\alpha_\perp$ in the definition of Q by their estimators. Denoting by $\tilde{Q}$ the resulting estimator of Q leads to the multivariate auxiliary regression model

$$\tilde{Q}'\tilde{A}(L)y_t = \tilde{Q}'\tilde{H}_{0t}\mu_0 + \tilde{Q}'\tilde{H}_{1t}\mu_1 + \tilde{Q}'\tilde{K}_{t\hat{\tau}}\delta + \eta_{t\hat{\tau}}, \qquad t = 1, \dots, T. \tag{A.41a}$$

The LS estimators of the parameters μ_i $(i = 0, 1)$ and δ from this auxiliary model will be denoted by $\hat{\mu}_i$ $(i = 0, 1)$ and $\hat{\delta}$, respectively, in the following. Some of their properties are stated next.

Lemma C.

$$\beta'(\hat{\mu}_0 - \mu_0) = O_p(T^{-1/2}), \tag{A.42}$$

$$\beta_\perp'(\hat{\mu}_0 - \mu_0) = O_p(1), \tag{A.43}$$

$$\beta'(\hat{\delta} - \delta) = O_p(T^{-1/2}), \tag{A.44}$$

$$\beta_\perp'(\hat{\delta} - \delta) = O_p(1), \tag{A.45}$$

$$\beta'(\hat{\mu}_1 - \mu_1) = O_p(T^{-3/2}), \tag{A.46}$$

$$T^{1/2}\beta_\perp'(\hat{\mu}_1 - \mu_1) \xrightarrow{d} N(0, \beta_\perp' C\Omega C'\beta_\perp). \tag{A.47}$$

Here $C = \beta_\perp(\alpha_\perp'\Psi\beta_\perp)^{-1}\alpha_\perp'$ as before and all quantities converge jointly in distribution upon appropriate standardization.

Proof: The idea is to show how the stated results can be obtained from the proof of Theorem 2.1 of S&L. In doing this we ignore the fact that S&L had additional impulse dummies in their model because this feature has no essential effect on the arguments.

We shall first introduce a reparameterized form of the auxiliary regression model $(A.41a)$ which is a counterpart of equation (A.12) of S&L. To this end, define

$$k_1 = \begin{bmatrix} k_{11} \\ k_{12} \end{bmatrix} = \begin{bmatrix} \tilde{\beta}_\perp' \delta \\ \tilde{\beta}_\perp' \mu_0 \end{bmatrix} \quad \text{and} \quad k_2 = \begin{bmatrix} k_{21} \\ k_{22} \\ k_{23} \\ k_{24} \end{bmatrix} = \begin{bmatrix} \tilde{\beta}' \delta \\ \tilde{\beta}' \mu_0 \\ \tilde{\beta}' \mu_1 \\ \tilde{\beta}_\perp' \mu_1 \end{bmatrix}.$$

Next transform the regressors in $(A.41a)$ accordingly as $\tilde{F}_{11t} = \tilde{Q}'\tilde{K}_{t\hat{\tau}}\tilde{\beta}_\perp(\tilde{\beta}_\perp'\tilde{\beta}_\perp)^{-1}$, $\tilde{F}_{12t} = \tilde{Q}'\tilde{H}_{0t}\tilde{\beta}_\perp(\tilde{\beta}_\perp'\tilde{\beta}_\perp)^{-1}$ and $\tilde{F}_{21t} = \tilde{Q}'\tilde{K}_{t\hat{\tau}}\tilde{\beta}(\tilde{\beta}'\tilde{\beta})^{-1}$, $\tilde{F}_{22t} = \tilde{Q}'\tilde{H}_{0t}\tilde{\beta}(\tilde{\beta}'\tilde{\beta})^{-1}$, $\tilde{F}_{23t} = \tilde{Q}'\tilde{H}_{1t}\tilde{\beta}(\tilde{\beta}'\tilde{\beta})^{-1}$, $\tilde{F}_{24t} = \tilde{Q}'\tilde{H}_{1t}\tilde{\beta}_\perp(\tilde{\beta}_\perp'\tilde{\beta}_\perp)^{-1}$. Then, setting $\tilde{F}_{1t} = [\tilde{F}_{11t} : \tilde{F}_{12t}]$ and $\tilde{F}_{2t} = [\tilde{F}_{21t} : \tilde{F}_{22t} : \tilde{F}_{23t} : \tilde{F}_{24t}]$, we can write $(A.41a)$ as

$$\tilde{Q}'\tilde{A}(L)y_t = \tilde{F}_{1t}k_1 + \tilde{F}_{2t}k_2 + \eta_{t\hat{\tau}}, \qquad t = 1\ldots,T. \tag{A.48}$$

As in S&L, $\tilde{F}_{1t}$ takes nonzero values only for a fixed number of time indices t.

Because $(A.48)$ is obtained from (2.1) by first replacing $d_{t\tau}$ by $d_{t\hat{\tau}}$ and then premultiplying by $\tilde{Q}'\tilde{A}(L)$, the error term $\eta_{t\hat{\tau}}$ differs from its counterpart η_t in S&L by the additive term $-\tilde{Q}'\tilde{A}(L)\delta(d_{t\hat{\tau}} - d_{t\tau})$. Thus, using

$$A(L) = I_n - A_1L - \cdots - A_pL^p = I_n\Delta - \Pi L - \Gamma_1\Delta L - \cdots - \Gamma_{p-1}\Delta L^{p-1}.$$

we have

$$\begin{aligned} \eta_{t\hat{\tau}} &= \eta_{t\tau} + \tilde{Q}'\tilde{\alpha}\tilde{\beta}'\delta(d_{t-1,\hat{\tau}} - d_{t-1,\tau}) - \tilde{Q}'\delta(\Delta d_{t\hat{\tau}} - \Delta d_{t\tau}) \\ &\quad + \tilde{Q}'\textstyle\sum_{j=1}^{p-1}\tilde{\Gamma}_j\delta(\Delta d_{t-j,\hat{\tau}} - \Delta d_{t-j,\tau}), \end{aligned} \tag{A.49}$$

where $\eta_{t\tau} = \tilde{Q}'\varepsilon_t - \tilde{Q}'\tilde{\alpha}(\tilde{\beta} - \beta)'x_{t-1} - \tilde{Q}'(\tilde{\alpha} - \alpha)\beta'x_{t-1} - \tilde{Q}'\sum_{j=1}^{p-1}(\tilde{\Gamma}_j - \Gamma_j)\Delta x_{t-j}$ is identical to η_t defined in (A.13) of S&L.[1]

Now assume that $\delta_1 \neq 0$. Then Lemmas A and B in conjunction with the arguments used in S&L to justify (A.14) and (A.15) of that paper show that

$$\sum_{t=1}^{T}\tilde{F}_{1t}'\eta_{t\hat{\tau}} = \sum_{t=1}^{T}\tilde{F}_{1t}'\eta_{t\tau} + O_p(1) = O_p(1) \tag{A.50}$$

[1]Note that there is a typo in (A.13) of S&L because in the third term on the r.h.s. there is $\tilde{\beta}$ in place of β.

and

$$\Upsilon_T^{-1}\sum_{t=1}^{T}\tilde{F}_{2t}'\eta_{t\hat{\tau}} = \Upsilon_T^{-1}\sum_{t=1}^{T}\tilde{F}_{2t}'\eta_{t\tau} + o_p(1) = O_p(1), \qquad (A.51)$$

where $\Upsilon_T^{-1} = \text{diag}[(T-\tau)^{-1/2}I_r : T^{-1/2}I_r : T^{-3/2}I_r : T^{-1/2}I_{n-r}]$. Standardizing the moment matrix of the auxiliary regression model $(A.48)$ to conform to the standardizations of the sums in $(A.50)$ and $(A.51)$ and using Lemmas A and B shows that the resulting matrix is asymptotically equivalent to its counterpart obtained when the break date is known. This combined with $(A.50)$ and $(A.51)$ implies that we have reduced the problem to that treated in Theorem 2.1 of S&L. Because we can write $O_p(T^{-1/2}) = O_p((T-\tau)^{-1/2})$ in $(A.44)$, this gives the stated results when $\delta_1 \neq 0$.

To complete the proof, assume that $\delta_1 = 0$ and observe that in the second term on the r.h.s. of $(A.49)$ we have $\tilde{\alpha}\tilde{\beta}'\delta = \tilde{\alpha}_\xi(\tilde{\beta}_\xi - \beta)'\delta = O_p(T^{-1})$ by Lemma B. Using this fact, it is easy to check that $(A.50)$ also holds in the present case. Regarding subsequent derivations it appears convenient to replace Υ_T in $(A.51)$ by $\hat{\Upsilon}_T$ obtained by using $\hat{\tau}$ in place of τ. This change does not affect rates of convergence because $[T\underline{\lambda}] \leq \hat{\tau} \leq [T\bar{\lambda}]$ by assumption. Arguments used to justify $(A.50)$ and to prove Theorem 3.1 combined with Lemma B now straightforwardly show that $(A.51)$ holds with Υ_T^{-1} replaced by $\hat{\Upsilon}_T^{-1}$.

Next consider the standardized moment matrix related to the LS estimation of the parameters k_1 and k_2 in $(A.48)$. Here we use arguments similar to those used in the discussion paper version of S&L to prove Theorem 2.1 therein. First note that arguments already used above give $\hat{\Upsilon}_T^{-1}\sum_{t=1}^{T}\tilde{F}_{2t}'\tilde{F}_{1t} = o_p(1)$ so that the standardized moment matrix is asymptotically block diagonal. From the definition of $\tilde{F}_{2t}$ it can further be shown that

$$\hat{\Upsilon}_T^{-1}\sum_{t=1}^{T}\tilde{F}_{2t}'\tilde{F}_{2t}\hat{\Upsilon}_T^{-1} = \sum_{t=1}^{T}\begin{pmatrix} -(T-\hat{\tau})^{-1/2}\tilde{\alpha}'\tilde{Q}d_{t\hat{\tau}} \\ -T^{-1/2}\tilde{\alpha}'\tilde{Q} \\ -T^{-3/2}(t-1)\tilde{\alpha}'\tilde{Q} \\ -T^{-1/2}(\tilde{\beta}_\perp'\tilde{\beta}_\perp)^{-1}\tilde{\beta}_\perp'\tilde{\Psi}'\tilde{Q}\end{pmatrix}\begin{pmatrix} -(T-\hat{\tau})^{-1/2}\tilde{\alpha}'\tilde{Q}d_{t\hat{\tau}} \\ -T^{-1/2}\tilde{\alpha}'\tilde{Q} \\ -T^{-3/2}(t-1)\tilde{\alpha}'\tilde{Q} \\ -T^{-1/2}(\tilde{\beta}_\perp'\tilde{\beta}_\perp)^{-1}\tilde{\beta}_\perp'\tilde{\Psi}'\tilde{Q}\end{pmatrix}' + o_p(1)$$

$$= \begin{bmatrix} \hat{E}_T \otimes \tilde{A}_{11} & \hat{e}_T \otimes \tilde{A}_{12} \\ \hat{e}_T' \otimes \tilde{A}_{21} & \tilde{A}_{22}\end{bmatrix} + o_p(1). \qquad (A.52)$$

Here the notation is defined as follows: $\tilde{A}_{11} = \tilde{\alpha}'\tilde{\Omega}\tilde{\alpha}$, $\tilde{A}_{12} = \tilde{A}_{21}' = \tilde{\alpha}'\tilde{\Omega}^{-1}\tilde{\Psi}\tilde{\beta}_\perp(\tilde{\beta}_\perp'\tilde{\beta}_\perp)^{-1}$ and $\tilde{A}_{22} = (\tilde{\beta}_\perp'\tilde{\beta}_\perp)^{-1}\tilde{\beta}_\perp'\tilde{\Psi}'\tilde{\Omega}^{-1}\tilde{\Psi}\tilde{\beta}_\perp(\tilde{\beta}_\perp'\tilde{\beta}_\perp)^{-1} = \tilde{A}_{21}\tilde{A}_{11}^{-1}\tilde{A}_{12} + \tilde{B}^{-1}$, where

$$\tilde{B} = \tilde{\beta}_\perp'\tilde{\beta}_\perp(\tilde{\alpha}_\perp'\tilde{\Psi}\tilde{\beta}_\perp)^{-1}\tilde{\alpha}_\perp'\tilde{\Omega}\tilde{\alpha}_\perp(\tilde{\beta}_\perp'\tilde{\Psi}'\tilde{\alpha}_\perp)^{-1}\tilde{\beta}_\perp'\tilde{\beta}_\perp = \tilde{\beta}_\perp'\tilde{C}\tilde{\Omega}\tilde{C}'\tilde{\beta}_\perp$$

with $\tilde{C} = \tilde{\beta}_\perp(\tilde{\alpha}'_\perp\tilde{\Psi}\tilde{\beta}_\perp)^{-1}\tilde{\alpha}'_\perp$. The matrix $\hat{E}_T$ is defined by

$$\hat{E}_T = \begin{bmatrix} 1 & \left(\frac{T-\hat{\tau}}{T}\right)^{1/2} & T^{-3/2}(T-\hat{\tau})^{-1/2}\sum_{t=\hat{\tau}+1}^{T}(t-1) \\ \left(\frac{T-\hat{\tau}}{T}\right)^{1/2} & 1 & T^{-2}\sum_{t=1}^{T}(t-1) \\ T^{-3/2}(T-\hat{\tau})^{-1/2}\sum_{t=\hat{\tau}+1}^{T}(t-1) & T^{-2}\sum_{t=1}^{T}(t-1) & T^{-3}\sum_{t=1}^{T}(t-1)^2 \end{bmatrix},$$

while the vector $\hat{e}_T$ is the second column of $\hat{E}_T$, that is,

$$\hat{e}_T = \left[\left(\frac{T-\hat{\tau}}{T}\right)^{1/2} : 1 : T^{-2}\sum_{t=1}^{T}(t-1)\right]'.$$

Because now $\hat{\tau}/T$ does not converge, the same is true for $\hat{E}_T$ and $\hat{e}_T$ and, consequently, for the first term in the last expression of $(A.52)$. However, using the definitions and Lemma B, it can be shown that the smallest eigenvalue of this matrix is asymptotically bounded away from zero. Because it is easy to check that this is also the case for the smallest eigenvalue of the matrix $\sum_{t=1}^{T}\tilde{F}'_{1t}\tilde{F}_{1t}$, we can conclude that the smallest eigenvalue of the standardized moment matrix related to the LS estimation of k_1 and k_2 in $(A.48)$ is similarly asymptotically bounded away from zero. Thus, because we also have $(A.50)$ and $(A.51)$ with Υ_T replaced by $\hat{\Upsilon}_T$, the LS estimators obtained from $(A.48)$ satisfy $\hat{k}_1 = k_1 + O_p(1)$ and $\hat{\Upsilon}_T(\hat{k}_2 - k_2) = O_p(1)$. Because here $\hat{\Upsilon}_T$ can be replaced by Υ_T without changing rates of convergence, we can conclude in the same way as in the proof of Theorem 2.1 of S&L that the results of the lemma hold except for the weak limit in $(A.47)$.

To see that $(A.47)$ holds, notice that, according to our previous derivations, we can write

$$\hat{\Upsilon}_T(\hat{k}_2 - k_2) = \left(\hat{\Upsilon}_T^{-1}\sum_{t=1}^{T}\tilde{F}'_{2t}\tilde{F}_{2t}\hat{\Upsilon}_T^{-1}\right)^{-1}\hat{\Upsilon}_T\sum_{t=1}^{T}\tilde{F}'_{2t}\eta_{t\tau} + o_p(1).$$

We have to consider the last $n-r$ components of the vector on the r.h.s. and therefore calculate the last $n-r$ rows of the inverse on the r.h.s.. For this purpose we apply the well-known formula for the partitioned inverse to the first term in the last expression of $(A.52)$. Because it is easy to check that $\hat{e}'_T\hat{E}_T^{-1} = [0:1:0]$ and $\hat{e}'_T\hat{E}_T^{-1}\hat{e}_T = 1$, this yields

$$\left(\hat{\Upsilon}_T^{-1}\sum_{t=1}^{T}\tilde{F}'_{2t}\tilde{F}_{2t}\hat{\Upsilon}_T^{-1}\right)^{-1} = \begin{bmatrix} * & * & * & * \\ 0 & -\tilde{B}\tilde{A}_{21}\tilde{A}_{11}^{-1} & 0 & \tilde{B} \end{bmatrix} + o_p(1)$$

where the blocks denoted by '*' are not needed and the partition on the r.h.s. conforms to the four components of the estimator vector $\hat{k}_2$. Thus, we can write

$$T^{1/2}(\hat{k}_{24} - k_{24}) = -\tilde{B}\tilde{A}_{21}\tilde{A}_{11}^{-1}T^{-1/2}\sum_{t=1}^{T}\tilde{F}'_{22t}\eta_{t\tau} + \tilde{B}T^{-1/2}\sum_{t=1}^{T}\tilde{F}'_{24t}\eta_{t\tau} + o_p(1).$$

This is asymptotically equivalent to what was obtained in S&L so that $(A.45)$ follows from that paper. $\square$

Proof of Theorem 4.1

Now we can prove Theorem 4.1. First note that

$$\hat{x}_t = x_t - (\hat{\mu}_0 - \mu_0) - (\hat{\mu}_1 - \mu_1)t - (\hat{\delta} - \delta)d_{t\tau} - \hat{\delta}(d_{t\hat{\tau}} - d_{t\tau}). \qquad (A.53)$$

This expression differs from its counterpart in the proof of Theorem 3.1 of S&L by the last term if we again ignore the unessential impulse dummies in S&L. This implies that the error term $e_{t\hat{\tau}}$ in (4.2) can be written as

$$e_{t\hat{\tau}} = e_{t\tau} - \Pi\hat{\delta}(d_{t\hat{\tau}} - d_{t\tau}) - \sum_{j=0}^{p-1} \gamma_j(\Delta d_{t-j,\hat{\tau}} - \Delta d_{t-j,\tau}), \qquad (A.54)$$

where $e_{t\tau}$ equals the error term e_t defined in the proof of Theorem 3.1 of S&L except that it is based on different estimators and some unessential impulse dummies are ignored. The difference in estimators has no consequences, however, because the asymptotic properties required in Assumption 1 are the same as those given in Theorem 2.1 of S&L or Lemma C. Thus, when $\delta_1 \neq 0$, we can use Lemma A and Assumption 1 in conjunction with $(A.53)$ and $(A.54)$ to conclude that the standardized second sample moments of $\beta'\hat{x}_{t-1}$, $\beta'_\perp\hat{x}_{t-1}$ and $\Delta\hat{x}_{t-j}$ $(j = 1, \ldots, p-1)$ are asymptotically equivalent to their counterparts in S&L. By similar arguments it can be further shown that the standardized second sample moments between $e_{t\hat{\tau}}$ and $[\hat{x}'_{t-1}\beta : \Delta\hat{x}'_{t-1} : \cdots : \Delta\hat{x}'_{t-p+1}]'$ as well as between $e_{t\hat{\tau}}$ and $\beta'_\perp\hat{x}_{t-1}$ are asymptotically equivalent to their counterparts in S&L. The discussion given at the end of the proof of Theorem 3.1 of S&L then implies that the RR estimators of α, β and Ω based on (4.2) have the same consistency properties as the corresponding estimators in Lemma B. This means that we can proceed in the same way as in the proof of Theorem 3.1 of S&L with $\tilde{\alpha}$, $\tilde{\beta}$ and $\tilde{\Omega}$ interpreted as the RR estimators based on (4.2).

It can be seen analogously that the same conclusion also holds when $\delta_1 = 0$, except that instead of Lemma A we now use the fact that $\beta'\hat{\delta} = O_p(T^{-1/2})$ by $(A.44)$. In this case it should be noted that the standardized second sample moments between $e_{t\hat{\tau}}$ and $[\hat{x}'_{t-1}\beta : \Delta x'_{t-1} : \cdots : \Delta x'_{t-p+1}]'$ and between $e_{t\hat{\tau}}$ and $\beta'_\perp\hat{x}_{t-1}$ are not asymptotically equivalent to their counterparts in S&L. However, they are of order $O_p(1)$ which is sufficient for the RR estimators of α, β and Ω based on (4.2) to have the same consistency properties as the corresponding estimators in Lemma B. Thus we can proceed in the same way as in the proof of Theorem 3.1 of S&L with $\tilde{\alpha}$, $\tilde{\beta}$ and $\tilde{\Omega}$ interpreted as the RR estimators based on (4.2).

The proof can be completed by showing that the problem can be reduced to that treated in the proof of Theorem 3.1 of S&L. This means showing that the last term on the r.h.s. of $(A.53)$ and the two last terms on the r.h.s. of $(A.54)$ have no effect on the limiting distribution of the test statistic or on the asymptotic behavior of the counterparts of the quantities defined in (A.18) - (A.22) of S&L. This can be done by using arguments already used in previous proofs in conjunction with Lemma A when $\delta_1 \neq 0$ and the fact $\tilde{\beta}_\xi'\hat{\delta} = \beta'\hat{\delta} + O_p(T^{-1}) = O_p(T^{-1/2})$ when $\delta_1 = 0$. Details are straightforward but tedious and will be omitted.