Online Theory Supplement to

"A One-Covariate at a Time, Multiple Testing Approach to Variable Selection in High-Dimensional Linear Regression Models"

A. Chudik
Federal Reserve Bank of Dallas

G. Kapetanios
King's College, London

M. Hashem Pesaran
University of Southern California, USA and Trinity College, Cambridge, UK

This online theory supplement is organised as follows: Section A provides lemmas for the Appendix of the main paper. Section B provides a proof of Theorem 3. Section C provides a discussion of various results related to the case where both signal and noise variables are mixing processes. Section D presents lemmas for regressions with covariates that are mixing processors. Section E provides lemmas for the case where the regressors are deterministic, while Section F provides some further supplementary lemmas needed for Sections B and C of this supplement.

A. Lemmas

Before presenting the lemmas and their proofs we give an outline of their use. Lemmas A1 and A2 are technical auxiliary lemmas. Lemmas A3-A5 provide extensions to existing results in the literature that form the building blocks for our exponential probability inequalities. Lemmas A6 and A7 provide exponential probability inequalities for squares and cross-products of sums of random variables. Lemmas A8 and A9 provide results that help handle the denominator of a t-statistic in the context of exponential inequalities. Lemma A10 is a key lemma that provides exponential inequalities for t-statistics. Lemmas A11-A21 are further auxiliary lemmas.

Lemma A1 *Let y_t, for $t = 1, 2, ..., T$, be given by DGP (6) and define $\boldsymbol{x}_i = (x_{i1}, x_{i2}, ..., x_{iT})'$, for $i = 1, 2, ..., k$, and $\mathbf{X}_k = (\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_k)$, and suppose that Assumption 1 holds. Moreover, let $\mathbf{q}_{i\cdot} = (q_{i1}, q_{i2}, ..., q_{iT})'$, for $i = 1, 2, ..., l_T$, $\mathbf{Q} = (\mathbf{q}_{1\cdot}, \mathbf{q}_{2\cdot}, ..., \mathbf{q}_{l_T\cdot})'$, and assume $\mathbf{M}_q = \mathbf{I}_T - \mathbf{Q}(\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}'$ exists. Further, assume that $\boldsymbol{\tau}_T = (1, 1, ..., 1)'$ is included in $\mathbf{Q}$, a $(0 \leq a < k)$ column vectors of $\mathbf{X}_k$ belong to $\mathbf{Q}$, and the remaining $b = k - 1 > 0$ columns of $\mathbf{X}_k$ that do not belong in $\mathbf{Q}$ are collected in the $T \times b$ matrix $\mathbf{X}_b$. The slope coefficients that correspond to regressors in $\mathbf{X}_b$ are collected in the $b \times 1$ vector $\boldsymbol{\beta}_{b,T}$. Define $\boldsymbol{\theta}_{b,T} = \boldsymbol{\Omega}_{b,T}\boldsymbol{\beta}_{b,T}$, where $\boldsymbol{\Omega}_{b,T} = E(T^{-1}\mathbf{X}_b'\mathbf{M}_q\mathbf{X}_b)$. If $\boldsymbol{\Omega}_{b,T}$ is nonsingular, and $\boldsymbol{\beta}_{k,T} = (\beta_{1,T}, \beta_{2,T}, ..., \beta_{k,T})' \neq \mathbf{0}$, then at least one element of the $b \times 1$ vector $\boldsymbol{\theta}_{b,T}$ is nonzero.*

Proof. Since $\boldsymbol{\Omega}_{b,T}$ is nonsingular and $\boldsymbol{\beta}_{b,T} \neq \mathbf{0}$, it follows that $\boldsymbol{\theta}_{b,T} \neq \mathbf{0}$; otherwise $\boldsymbol{\beta}_{b,T} = \boldsymbol{\Omega}_{b,T}^{-1}\boldsymbol{\theta}_{b,T} = \mathbf{0}$, which contradicts the assumption that $\boldsymbol{\beta}_{b,T} \neq \mathbf{0}$. ■

Lemma A2 *Consider the critical value function* $c_p(n,\delta)$ *defined by (15), with* $0 < p < 1$ *and* $f(n,\delta) = cn^{\delta}$, *for some* $c,\delta > 0$. *Moreover, let* $a > 0$ *and* $0 < b \leq 1$. *Then: (i)* $c_p(n,\delta) = O\left([\delta \ln(n)]^{1/2}\right)$, *(ii)* $n^a \exp\left[-bc_p^2(n,\delta)\right] = \ominus\left(n^{a-2b\delta}\right)$.

Proof. Results follow from Lemma 3 of the Supplementary Appendix A of Bailey et al. (2018). ■

Lemma A3 *Let* z_t *be a martingale difference sequence with respect to the filtration* $\mathcal{F}_{t-1}^z = \sigma\left(\{z_s\}_{s=1}^{t-1}\right)$, *and suppose that there exist finite positive constants* C_0 *and* C_1, *and* $s > 0$ *such that* $\sup_t \Pr(|z_t| > \alpha) \leq C_0 \exp(-C_1\alpha^s)$, *for all* $\alpha > 0$. *Let* $\sigma_{zt}^2 = E(z_t^2 \left| \mathcal{F}_{t-1}^z \right.)$ *and* $\sigma_z^2 = \frac{1}{T}\sum_{t=1}^T \sigma_{zt}^2$. *Suppose that* $\zeta_T = \ominus(T^{\lambda})$, *for some* $0 < \lambda \leq (s+1)/(s+2)$. *Then, for any* π *in the range* $0 < \pi < 1$, *we have*

$$\Pr(|\textstyle\sum_{t=1}^T z_t| > \zeta_T) \leq \exp[-(1-\pi)^2 \zeta_T^2 T^{-1}\sigma_z^{-2}/2]. \tag{B.1}$$

If $\lambda > (s+1)/(s+2)$, *then for some finite positive constant* C_3,

$$\Pr(|\textstyle\sum_{t=1}^T z_t| > \zeta_T) \leq \exp[-C_3\zeta_T^{s/(s+1)}]. \tag{B.2}$$

Proof. We proceed to prove (B.1) first and then prove (B.2). Decompose z_t as $z_t = w_t + v_t$, where $w_t = z_t I(|z_t| \leq D_T)$ and $v_t = z_t I(|z_t| > D_T)$, and note that

$$\begin{aligned}\Pr\{|\textstyle\sum_{t=1}^T [z_t - E(z_t)]| > \zeta_T\} \leq & \Pr\{|\textstyle\sum_{t=1}^T [w_t - E(w_t)]| > (1-\pi)\zeta_T\} \\ & + \Pr\{|\textstyle\sum_{t=1}^T [v_t - E(v_t)]| > \pi\zeta_T\},\end{aligned} \tag{B.3}$$

for any $0 < \pi < 1$.[1] Further, it is easily verified that $w_t - E(w_t)$ is a martingale difference process, and since $|w_t| \leq D_T$ then by setting $b = T\sigma_z^2$ and $a = (1-\pi)\zeta_T$ in Proposition 2.1 of Freedman (1975), for the first term on the RHS of (B.3) we obtain

$$\Pr\{|\textstyle\sum_{t=1}^T [w_t - E(w_t)]| > (1-\pi)\zeta_T\} \leq \exp\{-\zeta_T^2 \left[T\sigma_z^2 + (1-\pi)D_T\zeta_T\right]^{-1}(1-\pi)^2/2\}.$$

Consider now the second term on the RHS of (B.3) and first note that

$$\Pr\{|\textstyle\sum_{t=1}^T [v_t - E(v_t)]| > \pi\zeta_T\} \leq \Pr[\textstyle\sum_{t=1}^T |v_t - E(v_t)| > \pi\zeta_T], \tag{B.4}$$

and by Markov's inequality,

$$\Pr\{\textstyle\sum_{t=1}^T |[v_t - E(v_t)]| > \pi\zeta_T\} \leq \pi^{-1}\zeta_T^{-1}\textstyle\sum_{t=1}^T E|v_t - E(v_t)| \leq 2\pi^{-1}\zeta_T^{-1}\textstyle\sum_{t=1}^T E|v_t|. \tag{B.5}$$

[1] Let $A_T = \sum_{t=1}^T [z_t - E(z_t)] = B_{1,T} + B_{2,T}$, where $B_{1,T} = \sum_{t=1}^T [w_t - E(w_t)]$ and $B_{2,T} = \sum_{t=1}^T [v_t - E(v_t)]$. We have $|A_T| \leq |B_{1,T}| + |B_{2,T}|$ and, therefore, $\Pr(|A_T| > \zeta_T) \leq \Pr(|B_{1,T}| + |B_{2,T}| > \zeta_T)$. Equation (B.3) now readily follows using the same steps as in the proof of (B.59).

But by Holder's inequality, for any finite $p, q \geq 1$ such that $p^{-1}+q^{-1}=1$ we have $E\left|v_t\right| = E\left(\left|z_t I\left[\left|z_t\right|>D_T\right]\right|\right) \leq \left(E\left|z_t\right|^p\right)^{1/p}\left\{E\left[\left|I\left(\left|z_t\right|>D_T\right)\right|^q\right]\right\}^{1/q} = \left(E\left|z_t\right|^p\right)^{1/p}\left\{E\left[I\left(\left|z_t\right|>D_T\right)\right]\right\}^{1/q}$, and therefore

$$E\left|v_t\right| \leq \left(E\left|z_t\right|^p\right)^{1/p}\left[\Pr\left(\left|z_t\right|>D_T\right)\right]^{1/q}. \tag{B.6}$$

Also, for any finite $p \geq 1$ there exists a finite positive constant C_2 such that $E\left|z_t\right|^p \leq C_2 < \infty$, by Lemma A15. Further, by assumption $\sup_t \Pr\left(\left|z_t\right|>D_T\right) \leq C_0 \exp\left(-C_1 D_T^s\right)$. Using this upper bound in (B.6) together with the upper bound on $E\left|z_t\right|^p$, we have $\sup_t E\left|v_t\right| \leq C_2^{1/p} C_0^{1/q}\left[\exp\left(-C_1 D_T^s\right)\right]^{1/q}$. Therefore, using (B.4)-(B.5), $\Pr\{\left|\sum_{t=1}^T\left[v_t-E(v_t)\right]\right| > \pi\zeta_T\} \leq (2/\pi) C_2^{1/p} C_0^{1/q} \zeta_T^{-1} T\left[\exp\left(-C_1 D_T^s\right)\right]^{1/q}$. We need to determine D_T such that

$$(2/\pi) C_2^{1/p} C_0^{1/q} \zeta_T^{-1} T\left[\exp\left(-C_1 D_T^s\right)\right]^{1/q} \leq \exp\{-\zeta_T^2\left[T\sigma_z^2+(1-\pi) D_T\zeta_T\right]^{-1}(1-\pi)^2/2\}. \tag{B.7}$$

Taking logs, we have $\ln[(2/\pi) C_2^{1/p} C_0^{1/q}]+\ln\left(\zeta_T^{-1} T\right)-\left(C_1/q\right) D_T^s \leq -(1-\pi)^2 \zeta_T^2/\{2[T\sigma_z^2+(1-\pi) D_T\zeta_T]\}$, or $C_1 q^{-1} D_T^s \geq \ln[(2/\pi) C_2^{1/p} C_0^{1/q}]+\ln\left(\zeta_T^{-1} T\right)+(1-\pi)^2 \zeta_T^2/\left\{2\left[T\sigma_z^2+(1-\pi) D_T\zeta_T\right]\right\}$. Post-multiplying by $2\left[T\sigma_z^2+(1-\pi) D_T\zeta_T\right]>0$ we have

$$\begin{aligned}&\left(2\sigma_z^2 C_1 q^{-1}\right) T D_T^s+\left(2C_1 q^{-1}\right)(1-\pi) D_T^{s+1}\zeta_T-2(1-\pi) D_T\zeta_T\{\ln\left(\zeta_T^{-1} T\right)+\ln[(2/\pi) C_2^{1/p} C_0^{1/q}]\} \\ &\geq 2\sigma_z^2 T\ln[(2/\pi) C_2^{1/p} C_0^{1/q}]+2\sigma_z^2 T\ln\left(\zeta_T^{-1} T\right)+(1-\pi)^2\zeta_T^2.\end{aligned} \tag{B.8}$$

The above expression can now be simplified for values of $T\to\infty$, by dropping the constants and terms that are asymptotically dominated by other terms on the same side of the inequality.[2] Since $\zeta_T=\ominus\left(T^\lambda\right)$, for some $0<\lambda\leq(s+1)/(s+2)$, and considering values of D_T such that $D_T=\ominus\left(T^\psi\right)$, for some $\psi>0$, implies that the third and fourth term on the LHS of (B.8), which have the orders $\ominus\left[\ln(T)T^{\lambda+\psi}\right]$ and $\ominus\left(T^{\lambda+\psi}\right)$, respectively, are dominated by the second term on the LHS of (B.8) which is of order $\ominus\left(T^{\lambda+\psi+s\psi}\right)$. Further the first term on the RHS of (B.8) is dominated by the second term. Note that for $\zeta_T=\ominus\left(T^\lambda\right)$, we have $T\ln\left(\zeta_T^{-1}T\right)=\ominus[T\ln(T)]$, whilst the order of the first term on the RHS of (B.8) is $\ominus(T)$. Result (B.7) follows if we show that there exists D_T such that

$$\left(C_1 q^{-1}\right)\left[2\sigma_z^2 T D_T^s+2(1-\pi) D_T^{s+1}\zeta_T\right] \geq 2\sigma_z^2 T\ln\left(\zeta_T^{-1}T\right)+(1-\pi)^2\zeta_T^2. \tag{B.9}$$

Set $\left(C_1 q^{-1}\right) D_T^{s+1}=(1-\pi)\zeta_T/2$, or $D_T=\left(C_1^{-1} q(1-\pi)\zeta_T/2\right)^{1/(s+1)}$, and note that (B.9) can be written as $2\sigma_z^2\left(C_1 q^{-1}\right) T\left(C_1^{-1} q(1-\pi)\zeta_T/2\right)^{s/(s+1)}+(1-\pi)^2\zeta_T^2 \geq 2\sigma_z^2 T\ln\left(\zeta_T^{-1}T\right)+(1-\pi)^2\zeta_T^2$. Hence, the required condition is met if $\lim_{T\to\infty}[\left(C_1 q^{-1}\right)\left(C_1^{-1} q(1-\pi)\zeta_T/2\right)^{s/(s+1)}-\ln\left(\zeta_T^{-1}T\right)] \geq 0$. This condition is clearly satisfied noting that for values of $\zeta_T=\ominus\left(T^\lambda\right)$, $q>0$, $C_1>0$ and $0<\pi<1$,

$$\left(C_1 q^{-1}\right)\left(C_1^{-1} q(1-\pi)\zeta_T/2\right)^{s/(s+1)}-\ln\left(\zeta_T^{-1}T\right)=\ominus(T^{\frac{\lambda s}{1+s}})-\ominus[\ln(T)],$$

[2] A term A is said to be asymptotically dominant compared to a term B if both tend to infinity and $A/B\to\infty$.

since $s > 0$ and $\lambda > 0$, the first term on the RHS, which is positive, dominates the second term. Finally, we require that $D_T \zeta_T = o(T)$, since then the denominator of the fraction inside the exponential on the RHS of (B.7) is dominated by T which takes us back to the Exponential inequality with bounded random variables and proves (B.1). Consider $T^{-1} D_T \zeta_T = [C_1^{-1} q (1-\pi)/2]^{1/(s+1)} T^{-1} \zeta_T^{(2+s)/(1+s)}$, and since $\zeta_T = \ominus(T^\lambda)$ then $D_T \zeta_T = o(T)$, as long as $\lambda < (s+1)/(s+2)$, as required.

If $\lambda > (s+1)/(s+2)$, it follows that $D_T \zeta_T$ dominates T in the denominator of the fraction inside the exponential on the RHS of (B.7). So the bound takes the form $\exp[-(1-\pi)\zeta_T^2/(C_4 D_T \zeta_T)]$, for some finite positive constant C_4. Noting that $D_T = \ominus(\zeta_T^{1/(s+1)})$, gives a bound of the form $\exp[-C_3 \zeta_T^{s/(s+1)}]$ proving (B.2). ■

Lemma A4 *Let x_t and u_t be sequences of random variables and suppose that there exist $C_0, C_1 > 0$, and $s > 0$ such that $\sup_t \Pr(|x_t| > \alpha) \leq C_0 \exp(-C_1 \alpha^s)$ and $\sup_t \Pr(|u_t| > \alpha) \leq C_0 \exp(-C_1 \alpha^s)$, for all $\alpha > 0$. Let $\mathcal{F}_{t-1}^{(1)} = \sigma(\{u_s\}_{s=1}^{t-1}, \{x_s\}_{s=1}^{t-1})$ and $\mathcal{F}_t^{(2)} = \sigma(\{u_s\}_{s=1}^{t-1}, \{x_s\}_{s=1}^{t})$. Then, assume either that (i) $E(u_t|\mathcal{F}_t^{(2)}) = 0$ or (ii) $E(x_t u_t - \mu_t|\mathcal{F}_{t-1}^{(1)}) = 0$, where $\mu_t = E(x_t u_t)$. Let $\zeta_T = \ominus(T^\lambda)$, for some λ such that $0 < \lambda \leq (s/2+1)/(s/2+2)$. Then, for any π in the range $0 < \pi < 1$ we have*

$$\Pr(|\sum_{t=1}^T (x_t u_t - \mu_t)| > \zeta_T) \leq \exp[-(1-\pi)^2 \zeta_T^2/(2T\sigma_{(T)}^2)], \tag{B.10}$$

where $\sigma_{(T)}^2 = T^{-1}\sum_{t=1}^T \sigma_t^2$ and $\sigma_t^2 = E[(x_t u_t - \mu_t)^2 |\mathcal{F}_{t-1}^{(1)}]$. If $\lambda > (s/2+1)/(s/2+2)$, then for some finite positive constant C_2,

$$\Pr(|\sum_{t=1}^T (x_t u_t - \mu_t)| > \zeta_T) \leq \exp[-C_2 \zeta_T^{s/(s+2)}]. \tag{B.11}$$

Proof. Let $\tilde{\mathcal{F}}_{t-1} = \sigma(\{x_s u_s\}_{s=1}^{t-1})$ and note that under (i), $E(x_t u_t|\tilde{\mathcal{F}}_{t-1}) = E[E(u_t|\mathcal{F}_t^{(2)}) x_t|\tilde{\mathcal{F}}_{t-1}] = 0$. Therefore, $x_t u_t$ is a martingale difference process. Under (ii) we simply note that $x_t u_t - \mu_t$ is a martingale difference process by assumption. Next, for any $\alpha > 0$ we have (using (B.60) with C_0 set equal to α and C_1 set equal to $\sqrt{\alpha}$)

$$\Pr[|x_t u_t| > \alpha] \leq \Pr\left[|x_t| > \alpha^{1/2}\right] + \Pr\left[|u_t| > \alpha^{1/2}\right]. \tag{B.12}$$

But, under the assumptions of the lemma, $\sup_t \Pr\left[|x_t| > \alpha^{1/2}\right] \leq C_0 e^{-C_1 \alpha^{s/2}}$, and $\sup_t \Pr\left[|u_t| > \alpha^{1/2}\right] \leq C_0 e^{-C_1 \alpha^{s/2}}$. Hence $\sup_t \Pr[|x_t u_t| > \alpha] \leq 2C_0 e^{-C_1 \alpha^{s/2}}$. Therefore, the process $x_t u_t$ satisfies the conditions of Lemma A3 and the results of the lemma apply. ■

Lemma A5 *Let $\boldsymbol{x} = (x_1, x_2, ..., x_T)'$ and $\boldsymbol{q}_{\cdot t} = (q_{1,t}, q_{2,t}, ..., q_{l_T,t})'$ be sequences of random variables and suppose that there exist finite positive constants C_0 and C_1, and $s > 0$ such that $\sup_t \Pr(|x_t| > \alpha) \leq C_0 \exp(-C_1 \alpha^s)$ and $\sup_{i,t} \Pr(|q_{i,t}| > \alpha) \leq C_0 \exp(-C_1 \alpha^s)$, for all $a > 0$. Consider the linear projection $x_t = \sum_{j=1}^{l_T} \beta_j q_{jt} + u_{x,t}$, and assume that only a finite number of slope coefficients β's are nonzero and bounded, and the remaining β's are zero. Then, there exist finite positive constants C_2 and C_3, such that $\sup_t \Pr(|u_{x,t}| > \alpha) \leq C_2 \exp(-C_3 \alpha^s)$.*

Proof. We assume without any loss of generality that the $|\beta_i| < C_0$ for $i = 1, 2, ..., M$, M is a finite positive integer and $\beta_i = 0$ for $i = M+1, M+2, ..., l_T$. Note that for some $0 < \pi < 1$, $\sup_t \Pr(|u_{x,t}| > \alpha) \leq \sup_t \Pr(|x_t - \sum_{j=1}^{M} \beta_j q_{jt}| > \alpha) \leq \sup_t \Pr(|x_t| > (1-\pi)\alpha) + \sup_t \Pr(|\sum_{j=1}^{M} \beta_j q_{jt}| > \pi\alpha) \leq \sup_t \Pr(|x_t| > (1-\pi)\alpha) + \sup_t \sum_{j=1}^{M} \Pr(|\beta_j q_{jt}| > \pi\alpha/M)$, and since $|\beta_j| > 0$, then $\sup_t \Pr(|u_{x,t}| > \alpha) \leq \sup_t \Pr(|x_t| > (1-\pi)\alpha) + M \sup_{j,t} \Pr[|q_{jt}| > \pi\alpha/(M|\beta_j|)]$. But $\sup_{j,t} \Pr[|q_{jt}| > \pi\alpha/(M|\beta_j|)] \leq \sup_{j,t} \Pr[|q_{jt}| > \pi\alpha/(M\beta_{\max})] \leq C_0 \exp\{-C_1[\pi\alpha/(M\beta_{\max})]^s\}$, and, for fixed M, the probability bound condition is clearly met. ■

Lemma A6 *Let x_{it}, $i = 1, 2, ..., n$, $t = 1, 2, ..., T$, and η_t be processes that satisfy exponential tail probability bounds of the form (9) and (10), with tail exponents s_x and s_η, where $s = \min(s_x, s_\eta) > 0$. Further, let $x_{it}\eta_t$, $i = 1, 2, ..., n$, be martingale difference processes. Let $\boldsymbol{q}_{\cdot t} = (q_{1,t}, q_{2,t}, ..., q_{l_T,t})'$ contain a constant and a subset of $\boldsymbol{x}_{nt} = (x_{1t}, x_{2t}, ..., x_{nt})'$. Let $\boldsymbol{\Sigma}_{qq} = T^{-1}\sum_{t=1}^{T} E(\boldsymbol{q}_{\cdot t}\boldsymbol{q}_{\cdot t}')$ and $\hat{\boldsymbol{\Sigma}}_{qq} = \mathbf{Q}'\mathbf{Q}/T$ be both invertible, where $\mathbf{Q} = (\boldsymbol{q}_{1\cdot}, \boldsymbol{q}_{2\cdot}, ..., \boldsymbol{q}_{l_T\cdot})$ and $\boldsymbol{q}_{i\cdot} = (q_{i1}, q_{i2}, ..., q_{iT})'$, for $i = 1, 2, ..., l_T$. Suppose that Assumption 5 holds for x_{it} and $\boldsymbol{q}_{\cdot t}$, $i = 1, 2, ..., n$, and for η_t and $\boldsymbol{q}_{\cdot t}$, and denote the corresponding projection residuals defined by (11) as $u_{x_i,t} = x_{it} - \boldsymbol{\gamma}_{qx_i,T}'\boldsymbol{q}_{\cdot t}$ and $u_{\eta,t} = \eta_t - \boldsymbol{\gamma}_{q\eta,T}'\boldsymbol{q}_{\cdot t}$, respectively. Let $\hat{\boldsymbol{u}}_{x_i} = (\hat{u}_{x_i,1}, \hat{u}_{x_i,2}, ..., \hat{u}_{x_i,T})' = \mathbf{M}_q\boldsymbol{x}_i$, $\boldsymbol{x}_i = (x_{i1}, x_{i2}, ..., x_{iT})'$, $\hat{\boldsymbol{u}}_\eta = (\hat{u}_{\eta,1}, \hat{u}_{\eta,2}, ..., \hat{u}_{\eta,T})' = \mathbf{M}_q\boldsymbol{\eta}$, $\boldsymbol{\eta} = (\eta_1, \eta_2, ..., \eta_T)'$, $\mathbf{M}_q = \mathbf{I}_T - \mathbf{Q}(\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}$, $\mathcal{F}_t = \mathcal{F}_t^\eta \cup \mathcal{F}_t^x$, $\mu_{x_i\eta,t} = E(u_{x_i,t}u_{\eta,t}|\mathcal{F}_{t-1})$, $\omega_{x_i\eta,1,T}^2 = \frac{1}{T}\sum_{t=1}^{T} E\left[(x_{it}\eta_t - E(x_{it}\eta_t|\mathcal{F}_{t-1}))^2\right]$, and $\omega_{x_i\eta,T}^2 = \frac{1}{T}\sum_{t=1}^{T} E\left[(u_{x_i,t}u_{\eta,t} - \mu_{x_i\eta,t})^2\right]$. Let $\zeta_T = \ominus(T^\lambda)$. Then, for any π in the range $0 < \pi < 1$, we have,*

$$\Pr[|\textstyle\sum_{t=1}^{T} x_{it}\eta_t - E(x_{it}\eta_t|\mathcal{F}_{t-1})| > \zeta_T] \leq \exp[-(1-\pi)^2 \zeta_T^2/(2T\omega_{x_i\eta,1,T}^2)], \tag{B.13}$$

if $0 < \lambda \leq (s/2+1)/(s/2+2)$. Further, if $\lambda > (s/2+1)/(s/2+2)$, we have,

$$\Pr[|\textstyle\sum_{t=1}^{T} x_{it}\eta_t - E(x_{it}\eta_t|\mathcal{F}_{t-1})| > \zeta_T] \leq \exp[-C_0\zeta_T^{s/(s+2)}], \tag{B.14}$$

for some finite positive constant C_0. If it is further assumed that $l_T = \ominus(T^d)$, such that $0 \leq d < 1/3$, then, if $3d/2 < \lambda \leq (s/2+1)/(s/2+2)$,

$$\Pr[|\textstyle\sum_{t=1}^{T}(\hat{u}_{x_i,t}\hat{u}_{\eta,t} - \mu_{x_i\eta,t})| > \zeta_T] \leq C_0\exp[-(1-\pi)^2\zeta_T^2/(2T\omega_{x_i\eta,T}^2)] + \exp\left[-C_1T^{C_2}\right]. \tag{B.15}$$

for some finite positive constants C_0, C_1 and C_2, and, if $\lambda > (s/2+1)/(s/2+2)$ we have

$$\Pr[|\textstyle\sum_{t=1}^{T}(\hat{u}_{x_i,t}\hat{u}_{\eta,t} - \mu_{x_i\eta,t})| > \zeta_T] \leq C_0\exp[-C_3\zeta_T^{s/(s+2)}] + \exp\left[-C_1T^{C_2}\right], \tag{B.16}$$

for some finite positive constants C_0, C_1, C_2 and C_3.

Proof. Note that all the results in the proofs below hold both for sequences and for triangular arrays of random variables. If $\boldsymbol{q}_{\cdot t}$ contains x_{it}, all results follow trivially, so, without loss of generality, we assume that, if this is the case, the relevant column of $\mathbf{Q}$ is removed. (B.13) and (B.14) follow immediately given our assumptions and Lemma A4. We proceed to prove the rest of the lemma. Let $\boldsymbol{u}_{x_i} = (u_{x_i,1}, u_{x_i,2}, ..., u_{x_i,T})'$ and $\boldsymbol{u}_\eta = (u_{\eta,1}, u_{\eta,2}, ..., u_{\eta,T})'$. We first note that $\sum_{t=1}^{T} (\hat{u}_{x_i,t}\hat{u}_{\eta,t} - \mu_{x_i\eta,t}) = \hat{\boldsymbol{u}}'_{x_i}\hat{\boldsymbol{u}}_\eta - \sum_{t=1}^{T} \mu_{x_i\eta,t} = \boldsymbol{u}'_{x_i}\mathbf{M}_q\boldsymbol{u}_\eta - \sum_{t=1}^{T} \mu_{x_i\eta,t}$, and

$$\sum_{t=1}^{T} (\hat{u}_{x_i,t}\hat{u}_{\eta,t} - \mu_{x_i\eta,t}) = \sum_{t=1}^{T} (u_{x_i,t}u_{\eta,t} - \mu_{x_i\eta,t}) - \left(T^{-1}\boldsymbol{u}'_{x_i}\mathbf{Q}\right)\hat{\boldsymbol{\Sigma}}_{qq}^{-1}(\mathbf{Q}'\boldsymbol{u}_\eta), \tag{B.17}$$

where $\hat{\boldsymbol{\Sigma}}_{qq} = T^{-1}(\mathbf{Q}'\mathbf{Q})$. The second term of the above expression can now be decomposed as

$$\left(T^{-1}\boldsymbol{u}'_{x_i}\mathbf{Q}\right)\hat{\boldsymbol{\Sigma}}_{qq}^{-1}(\mathbf{Q}'\boldsymbol{u}_\eta) = \left(T^{-1}\boldsymbol{u}'_{x_i}\mathbf{Q}\right)(\hat{\boldsymbol{\Sigma}}_{qq}^{-1} - \boldsymbol{\Sigma}_{qq}^{-1})(\mathbf{Q}'\boldsymbol{u}_\eta) + \left(T^{-1}\boldsymbol{u}'_{x_i}\mathbf{Q}\right)\boldsymbol{\Sigma}_{qq}^{-1}(\mathbf{Q}'\boldsymbol{u}_\eta). \tag{B.18}$$

By (B.59) and for any $0 < \pi_1, \pi_2, \pi_3 < 1$ such that $\sum_{i=1}^{3}\pi_i = 1$, we have $\Pr[|\sum_{t=1}^{T}(\hat{u}_{x_i,t}\hat{u}_{\eta,t} - \mu_{x_i\eta,t})| > \zeta_T] \leq \Pr[|\sum_{t=1}^{T}(u_{x_i,t}u_{\eta,t} - \mu_{x_i\eta,t})| > \pi_1\zeta_T] +$
$\Pr[|\left(T^{-1}\boldsymbol{u}'_{x_i}\mathbf{Q}\right)(\hat{\boldsymbol{\Sigma}}_{qq}^{-1} - \boldsymbol{\Sigma}_{qq}^{-1})(\mathbf{Q}'\boldsymbol{u}_\eta)| > \pi_2\zeta_T] + \Pr[|\left(T^{-1}\boldsymbol{u}'_{x_i}\mathbf{Q}\right)\boldsymbol{\Sigma}_{qq}^{-1}(\mathbf{Q}'\boldsymbol{u}_\eta)| > \pi_3\zeta_T]$. Also applying (B.60) to the last two terms of the above we obtain
$\Pr[|\left(T^{-1}\boldsymbol{u}'_{x_i}\mathbf{Q}\right)(\hat{\boldsymbol{\Sigma}}_{qq}^{-1} - \boldsymbol{\Sigma}_{qq}^{-1})(\mathbf{Q}'\boldsymbol{u}_\eta)| > \pi_2\zeta_T] \leq \Pr(||\hat{\boldsymbol{\Sigma}}_{qq}^{-1} - \boldsymbol{\Sigma}_{qq}^{-1}||_F \left\|T^{-1}\boldsymbol{u}'_{x_i}\mathbf{Q}\right\|_F \|\mathbf{Q}'\boldsymbol{u}_\eta\|_F > \pi_2\zeta_T) \leq$
$\Pr(||\hat{\boldsymbol{\Sigma}}_{qq}^{-1} - \boldsymbol{\Sigma}_{qq}^{-1}||_F > \zeta_T/\delta_T) + \Pr\left(T^{-1}\left\|\boldsymbol{u}'_{x_i}\mathbf{Q}\right\|_F \|\mathbf{Q}'\boldsymbol{u}_\eta\|_F > \pi_2\delta_T\right) \leq \Pr(||\hat{\boldsymbol{\Sigma}}_{qq}^{-1} - \boldsymbol{\Sigma}_{qq}^{-1}||_F > \zeta_T/\delta_T) +$
$\Pr[\left\|\boldsymbol{u}'_{x_i}\mathbf{Q}\right\|_F > (\pi_2\delta_T T)^{1/2}] + \Pr[\|\mathbf{Q}'\boldsymbol{u}_\eta\|_F > (\pi_2\delta_T T)^{1/2}]$, where $\delta_T > 0$ is a deterministic sequence. In what follows, we set $\delta_T = \ominus(\zeta_T^\alpha)$, for some $\alpha > 0$. Similarly,
$\Pr[|\left(T^{-1}\boldsymbol{u}'_{x_i}\mathbf{Q}\right)\boldsymbol{\Sigma}_{qq}^{-1}(\mathbf{Q}'\boldsymbol{u}_\eta)| > \pi_3\zeta_T] \leq \Pr(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F \left\|T^{-1}\boldsymbol{u}'_{x_i}\mathbf{Q}\right\|_F \|\mathbf{Q}'\boldsymbol{u}_\eta\|_F > \pi_3\zeta_T) \leq$
$\Pr[\left\|\boldsymbol{u}'_{x_i}\mathbf{Q}\right\|_F \|\mathbf{Q}'\boldsymbol{u}_\eta\|_F > \pi_3\zeta_T T/\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F] \leq \Pr(\left\|\boldsymbol{u}'_{x_i}\mathbf{Q}\right\|_F > \pi_3^{1/2}\zeta_T^{1/2}T^{1/2}\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-1/2}) +$
$\Pr(\|\mathbf{Q}'\boldsymbol{u}_\eta\|_F > \pi_3^{1/2}\zeta_T^{1/2}T^{1/2}\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-1/2})$. Overall

$$\begin{aligned}
&\Pr(|\textstyle\sum_{t=1}^{T}(\hat{u}_{x,t}\hat{u}_{\eta,t} - \mu_{x\eta,t})| > \zeta_T) \leq \Pr(|\textstyle\sum_{t=1}^{T}(u_{x,t}u_{\eta,t} - \mu_{x\eta,t})| > \pi_1\zeta_T) \\
&+ \Pr\left(\left\|\hat{\boldsymbol{\Sigma}}_{qq}^{-1} - \boldsymbol{\Sigma}_{qq}^{-1}\right\|_F > \zeta_T/\delta_T\right) + \Pr\left(\|\mathbf{Q}'\boldsymbol{u}_\eta\|_F > (\pi_2\delta_T T)^{1/2}\right) + \Pr\left(\|\boldsymbol{u}'_x\mathbf{Q}\|_F > (\pi_2\delta_T T)^{1/2}\right) \\
&+ \Pr(\|\boldsymbol{u}'_x\mathbf{Q}\|_F > \pi_3^{1/2}\zeta_T^{1/2}T^{1/2}\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-1/2}) + \Pr(\|\mathbf{Q}'\boldsymbol{u}_\eta\|_F > \pi_3^{1/2}\zeta_T^{1/2}T^{1/2}\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-1/2}).
\end{aligned} \tag{B.19}$$

First, since $u_{x,t}u_{\eta,t} - \mu_{x\eta,t}$ is a martingale difference process with respect to $\sigma(\{\eta_s\}_{s=1}^{t-1}, \{x_s\}_{s=1}^{t-1}, \{q_s\}_{s=1}^{t-1})$, by Lemma A4, we have, for any π in the range $0 < \pi < 1$,

$$\Pr[|\textstyle\sum_{t=1}^{T}(u_{x_i,t}u_{\eta,t} - \mu_{x_i\eta,t})| > \pi_1\zeta_T] \leq \exp[-(1-\pi)^2\zeta_T^2/(2T\omega_{x\eta,T}^2)], \tag{B.20}$$

if $0 < \lambda \leq (s/2+1)/(s/2+2)$, and

$$\Pr[|\textstyle\sum_{t=1}^{T}(u_{x_i,t}u_{\eta,t} - \mu_{x_i\eta,t})| > \pi_1\zeta_T] \leq \exp[-C_0\zeta_T^{s/(s+1)}], \tag{B.21}$$

if $\lambda > (s/2+1)/(s/2+2)$, for some finite positive constant C_0. We now show that the last five terms on the RHS of (B.19) are of order $\exp\left[-C_1T^{C_2}\right]$, for some finite positive constants C_1 and C_2. We will make use of Lemma A4 since by assumption $\{q_{it}u_{\eta,t}\}$ and $\{q_{it}u_{x_i,t}\}$ are martingale difference sequences. We note that some of the bounds of the last five terms exceed, in order, $T^{1/2}$. Since we do not know the value of s, we need to consider the possibility that either (B.10) or (B.11) of Lemma A4, apply. We start with (B.10). Then, for some finite positive constant C_0, we have[3]

$$\sup_i \Pr[\|\mathbf{q}'_i\boldsymbol{u}_\eta\| > (\pi_2\delta_T T)^{1/2}] \leq \exp(-C_0\delta_T). \tag{B.22}$$

[3] The required probability bound on u_{xt} follows from the probability bound assumptions on x_t and on q_{it}, for $i = 1, 2, ..., l_T$, even if $l_T \to \infty$. See also Lemma A5.

Also, using $\|\mathbf{Q}'\boldsymbol{u}_\eta\|_F^2 = \sum_{j=1}^{l_T}(\sum_{t=1}^{T} q_{jt}u_t)^2$ and (B.59), $\Pr[\|\mathbf{Q}'\boldsymbol{u}_\eta\|_F > (\pi_2\delta_T T)^{1/2}] = \Pr(\|\mathbf{Q}'\boldsymbol{u}_\eta\|_F^2 > \pi_2\delta_T T) \leq \sum_{j=1}^{l_T}\Pr[(\sum_{t=1}^{T} q_{jt}u_{\eta,t})^2 > \pi_2\delta_T T/l_T] = \sum_{j=1}^{l_T}\Pr[|\sum_{t=1}^{T} q_{jt}u_{\eta,t}| > (\pi_2\delta_T T/l_T)^{1/2}]$, which upon using (B.22) yields (for some finite positive constant C_0)

$$\Pr[\|\mathbf{Q}'\boldsymbol{u}_\eta\|_F > (\pi_2\delta_T T)^{1/2}] \leq l_T \exp\left(-C_0\delta_T/l_T\right), \;\; \Pr[\|\mathbf{Q}'\boldsymbol{u}_x\| > (\pi_2\delta_T T)^{1/2}] \leq l_T \exp\left(-C_0\delta_T/l_T\right). \tag{B.23}$$

Similarly,

$$\begin{aligned} \Pr(\|\mathbf{Q}'\boldsymbol{u}_\eta\|_F > \pi_3^{1/2}\zeta_T^{1/2}T^{1/2}\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-1/2}) &\leq l_T\exp[-C_0\zeta_T/(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F l_T)], \quad\quad (B.24)\\ \Pr(\|\mathbf{Q}'\boldsymbol{u}_x\| > \pi_3^{1/2}\zeta_T^{1/2}T^{1/2}\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-1/2}) &\leq l_T\exp[-C_0\zeta_T/(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F l_T)]. \end{aligned}$$

Turning to the second term of (B.19), since for all i and j, $\{q_{it}q_{jt} - E(q_{it}q_{jt})\}$ is a martingale difference process and q_{it} satisfy the required probability bound then

$$\sup_{ij}\Pr\{|T^{-1}\textstyle\sum_{t=1}^{T}\left[q_{it}q_{jt} - E(q_{it}q_{jt})\right]| > \pi_2\zeta_T/\delta_T\} \leq \exp(-C_0T\zeta_T^2/\delta_T^2). \tag{B.25}$$

Therefore, by Lemma A16, for some finite positive constant C_0, we have

$$\begin{aligned} \Pr(\|\hat{\boldsymbol{\Sigma}}_{qq}^{-1} - \boldsymbol{\Sigma}_{qq}^{-1}\| > \zeta_T/\delta_T) \leq\; & l_T^2\exp[-C_0T\zeta_T^2\delta_T^{-2}l_T^{-2}\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-2}(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F + \delta_T^{-1}\zeta_T)^{-2}] \\ & + l_T^2\exp(-C_0T\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-2}l_T^{-2}). \end{aligned} \tag{B.26}$$

Further by Lemma A14, $\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F = \ominus\left(l_T^{1/2}\right)$, and $T\zeta_T^2\delta_T^{-2}l_T^{-2}\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-2}(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F + \delta_T^{-1}\zeta_T)^{-2} = Tl_T^{-2}\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-2}(\delta_T\zeta_T^{-1}\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F + 1)^{-2}$. Consider now the different terms in the above expression and let $P_{11} = \delta_T/l_T$, $P_{12} = \zeta_T/(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F l_T)$, $P_{13} = Tl_T^{-2}\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-2}[\delta_T\zeta_T^{-1}\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F + 1]^{-2}$, and $P_{14} = T\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-2}l_T^{-2}$. Under $\delta_T = \ominus\left(\zeta_T^{\alpha}\right)$, $l_T = \ominus(T^d)$, and $\zeta_T = \ominus(T^{\lambda})$, we have $P_{11} = \delta_T/l_T = \ominus\left(T^{\alpha-d}\right)$,

$$P_{12} = \zeta_T/(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F l_T) = \ominus\left(T^{\lambda-3d/2}\right), \tag{B.27}$$

$P_{13} = Tl_T^{-2}\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-2}[\delta_T\zeta_T^{-1}\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F + 1]^{-2} = \ominus\left(T^{\max\{1+2\lambda-4d-2\alpha, 1+\lambda-7d/2-\alpha, 1-3d\}}\right)$, and $P_{14} = T\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-2}l_T^{-2} = \ominus\left(T^{1-3d}\right)$. Suppose that $d < 1/3$, and by (B.27) note that $\lambda \geq 3d/2$. Then, setting $\alpha = 1/3$, ensures that all the above four terms tend to infinity polynomially with T. Therefore, it also follows that they can be represented as terms of order $\exp\left[-C_1T^{C_2}\right]$, for some finite positive constants C_1 and C_2, and (B.15) follows. The same analysis can be repeated under (B.11). In this case, (B.23), (B.24), (B.25) and (B.26) are replaced by

$$\Pr\left(\|\mathbf{Q}'\boldsymbol{u}_\eta\|_F > (\pi_2\delta_T T)^{1/2}\right) \leq l_T \exp\left(-\frac{C_0\delta_T^{s/2(s+2)}T^{s/2(s+2)}}{l_T^{s/2(s+2)}}\right) = l_T \exp\left[-C_0\left(\frac{\delta_T T}{l_T}\right)^{s/2(s+2)}\right],$$

$$\Pr\left(\|\mathbf{Q}'\boldsymbol{u}_x\| > (\pi_2\delta_T T)^{1/2}\right) \leq l_T \exp\left(-\frac{C_0\delta_T^{s/2(s+2)}T^{s/2(s+2)}}{l_T^{s/2(s+2)}}\right) = l_T \exp\left[-C_0\left(\frac{\delta_T T}{l_T}\right)^{s/2(s+2)}\right],$$

$$\Pr\left(\|\mathbf{Q}'\boldsymbol{u}_\eta\|_F > \frac{\pi_3^{1/2}\zeta_T^{1/2}T^{1/2}}{\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{1/2}}\right) \leq l_T \exp\left(\frac{-C_0\zeta_T^{s/2(s+2)}T^{s/2(s+2)}}{\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{s/2(s+2)} l_T^{s/2(s+2)}}\right) = l_T \exp\left[-C_0\left(\frac{\zeta_T T}{\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F l_T}\right)^{s/2(s+2)}\right]$$

$$\Pr\left(\|\mathbf{Q}'\boldsymbol{u}_x\| > \frac{\pi_3^{1/2}\zeta_T^{1/2}T^{1/2}}{\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{1/2}}\right) \leq l_T \exp\left(\frac{-C_0\zeta_T^{s/2(s+2)}T^{s/2(s+2)}}{\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{s/2(s+2)} l_T^{s/2(s+2)}}\right) = l_T \exp\left[-C_0\left(\frac{\zeta_T T}{\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F l_T}\right)^{s/2(s+2)}\right],$$

$\sup_{ij}\Pr\{|T^{-1}\sum_{t=1}^T[q_{it}q_{jt} - E(q_{it}q_{jt})]| > \pi_2\zeta_T/\delta_T\} \leq \exp[-C_0T^{s/(s+2)}\zeta_T^{s/(s+2)}\delta_T^{-s/(s+2)}]$, and, using Lemma A17, $\Pr[||(\hat{\boldsymbol{\Sigma}}_{qq}^{-1} - \boldsymbol{\Sigma}_{qq}^{-1}|| > \pi_2\zeta_T/\delta_T] \leq$
$l_T^2\exp[-C_0T^{s/(s+2)}\zeta_T^{s/(s+2)}\delta_T^{-s/(s+2)}l_T^{-s/(s+2)}\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-s/(s+2)}(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F + \delta_T^{-1}\zeta_T)^{-s/(s+2)}]+$
$l_T^2\exp[-C_0T^{s/(s+2)}\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-s/(s+2)}l_T^{-s/(s+2)}] = l_T^2\exp\left(-C_0\{T\zeta_T/[\delta_Tl_T\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F + \delta_T^{-1}\zeta_T)]\}^{s/(s+2)}\right)+$
$l_T^2\exp[-C_0(T\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-1}l_T^{-1})^{s/(s+2)}]$, respectively. Once again, we need to derive conditions that imply that $P_{21} = \delta_TT/l_T$, $P_{22} = \zeta_TT\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-1}l_T^{-1}$, $P_{23} = T\zeta_T[\delta_Tl_T\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F+\delta_T^{-1}\zeta_T)]^{-1}$ and $P_{24} = T\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-1}l_T^{-1}$ are terms that tend to infinity polynomially with T. If that is the case then, as before, the relevant terms are of order $\exp\left[-C_1T^{C_2}\right]$, for some finite positive constants C_1 and C_2, and (B.16) follows, completing the proof of the lemma. P_{22} dominates P_{23} so we focus on P_{21}, P_{23} and P_{24}. We have $\delta_TT/l_T = \ominus\left(T^{1+\alpha-d/2}\right)$, $T\zeta_T[\delta_Tl_T\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F+\delta_T^{-1}\zeta_T)]^{-1} = \ominus\left[T^{\max(1+\lambda-\alpha-2d,1-3d/2)}\right]$, and $T\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{-1}l_T^{-1} = \ominus\left(T^{1-3d/2}\right)$. It immediately follows that under the conditions set when using (B.10), which were that $\alpha = 1/3$, $d < 1/3$ and $\lambda > 3d/2$, and as long as $s > 0$, P_{21} to P_{24} tend to infinity polynomially with T, proving the lemma.[4] ■

Lemma A7 *Let x_{it}, $i = 1, 2, ..., n$, be processes that satisfy exponential tail probability bounds of the form (9), with positive tail exponent s. Let $\boldsymbol{q}_{\cdot t} = (q_{1,t}, q_{2,t}, ..., q_{l_T,t})'$ contain a constant and a subset of $\boldsymbol{x}_{nt} = (x_{1t}, x_{2t}, ..., x_{nt})'$. Suppose that Assumption 5 holds for x_{it} and $\boldsymbol{q}_{\cdot t}$, $i = 1, 2, ..., n$, and denote the corresponding projection residuals defined by (11) as $u_{x_it} = x_{it} - \boldsymbol{\gamma}_{qx_i,T}'\boldsymbol{q}_{\cdot t}$. Let $\boldsymbol{\Sigma}_{qq} = T^{-1}\sum_{t=1}^T E\left(\boldsymbol{q}_{\cdot t}\boldsymbol{q}_{\cdot t}'\right)$ and $\hat{\boldsymbol{\Sigma}}_{qq} = \mathbf{Q}'\mathbf{Q}/T$ be both invertible, where $\mathbf{Q} = (\boldsymbol{q}_{1\cdot}, \boldsymbol{q}_{2\cdot}, ..., \boldsymbol{q}_{l_T\cdot})$ and $\boldsymbol{q}_{i\cdot} = (q_{i1}, q_{i2}, .., q_{iT})'$, for $i = 1, 2, ..., l_T$. Let $\hat{\boldsymbol{u}}_{x_i} = (\hat{u}_{x_i,1}, \hat{u}_{x_i,2}, ..., \hat{u}_{x_i,T})' = \mathbf{M}_q\boldsymbol{x}_i$, where $\boldsymbol{x}_i = (x_{i1}, x_{i2}, ..., x_{iT})'$ and $\mathbf{M}_q = \mathbf{I}_T - \mathbf{Q}\left(\mathbf{Q}'\mathbf{Q}\right)^{-1}\mathbf{Q}$. Moreover, suppose that $E\left(u_{x_i,t}^2 - \sigma_{x_it}^2|\mathcal{F}_{t-1}\right) = 0$,*

[4] It is important to highlight one particular feature of the above proof. In (B.23), $q_{it}u_{x,t}$ needs to be a martingale difference process. Note that if q_{it} is a martingale difference process distributed independently of $u_{x,t}$, then $q_{it}u_{x,t}$ is also a martingale difference process irrespective of the nature of $u_{x,t}$. This implies that one may not need to impose a martingale difference assumption on $u_{x,t}$ if x_{it} is a noise variable. Unfortunately, a leading case for which this lemma is used is one where $q_{it} = 1$. It is then clear that one needs to impose a martingale difference assumption on $u_{x,t}$, to deal with covariates that cannot be represented as martingale difference processes. We relax this assumption in Section C of the online theory supplement where we allow noise variables to be mixing processes.

where $\mathcal{F}_t = \mathcal{F}_t^x$ and $\sigma^2_{x_it} = E(u^2_{x_i,t})$. Let $\zeta_T = \ominus(T^\lambda)$. Then, if $0 < \lambda \leq (s/2+1)/(s/2+2)$, for any π in the range $0 < \pi < 1$, and some finite positive constant C_0, we have,

$$\Pr\left[\left|\sum_{t=1}^T \left(x^2_{it} - \sigma^2_{x_it}\right)\right| > \zeta_T\right] \leq C_0 \exp\left[-\left(1-\pi\right)^2 \zeta_T^2 T^{-1} \omega^{-2}_{i,1,T}/2\right]. \tag{B.28}$$

Otherwise, if $\lambda > (s/2+1)/(s/2+2)$, for some finite positive constant C_0, we have

$$\Pr\left[\left|\sum_{t=1}^T \left(x^2_{it} - \sigma^2_{x_it}\right)\right| > \zeta_T\right] \leq \exp\left[-C_0 \zeta_T^{s/(s+2)}\right]. \tag{B.29}$$

If it is further assumed that $l_T = \ominus\left(T^d\right)$, such that $0 \leq d < 1/3$, then, if $3d/2 < \lambda \leq (s/2+1)/(s/2+2)$,

$$\Pr\left[\left|\sum_{t=1}^T \left(\hat{u}^2_{x_i,t} - \sigma^2_{x_it}\right)\right| > \zeta_T\right] \leq C_0 \exp\left[-\left(1-\pi\right)^2 \zeta_T^2 T^{-1} \omega^{-2}_{i,T}/2\right] + \exp\left[-C_1 T^{C_2}\right], \tag{B.30}$$

for some finite positive constants C_0, C_1 and C_2, and, if $\lambda > (s/2+1)/(s/2+2)$,

$$\Pr\left[\left|\sum_{t=1}^T \left(\hat{u}^2_{x_i,t} - \sigma^2_{x_it}\right)\right| > \zeta_T\right] \leq C_0 \exp\left[-C_3 \zeta_T^{s/(s+2)}\right] + \exp\left[-C_1 T^{C_2}\right], \tag{B.31}$$

for some finite positive constants C_0, C_1, C_2 and C_3, where $\omega^2_{i,1,T} = T^{-1}\sum_{t=1}^T E\left[\left(x^2_{it} - \sigma^2_{x_it}\right)^2\right]$ and $\omega^2_{i,T} = T^{-1}\sum_{t=1}^T E\left[\left(u^2_{x_i,t} - \sigma^2_{x_it}\right)^2\right]$.

Proof. If $\boldsymbol{q}_{\cdot t}$ contains x_{it}, all results follow trivially, so, without loss of generality, we assume that, if this is the case, the relevant column of $\mathbf{Q}$ is removed. (B.28) and (B.29) follow similarly to (B.13) and (B.14). For (B.30) and (B.31), we first note that $\left|\sum_{t=1}^T \left(\hat{u}^2_{x_i,t} - \sigma^2_{x_it}\right)\right| \leq \left|\sum_{t=1}^T \left(u^2_{x_i,t} - \sigma^2_{x_it}\right)\right| + \left|\left(T^{-1}\boldsymbol{u}'_{x_i}\mathbf{Q}\right)\left(T^{-1}\mathbf{Q}'\mathbf{Q}\right)^{-1}\left(\mathbf{Q}'\boldsymbol{u}_{x_i}\right)\right|$. Since $\left\{u^2_{x_i,t} - \sigma^2_{x_it}\right\}$ is a martingale difference process and for $\alpha > 0$ and $s > 0$, $\sup_t \Pr\left(\left|u^2_{x_i,t}\right| > \alpha^2\right) = \sup_t \Pr\left(\left|u_{x_i,t}\right| > \alpha\right) \leq C_0 \exp\left(-C_1\alpha^s\right)$, by Lemma A5, then the conditions of Lemma A3 are met and we have $\Pr[|\sum_{t=1}^T \left(u^2_{x_i,t} - \sigma^2_{x_it}\right)| > \zeta_T] \leq \exp[-\left(1-\pi\right)^2 \zeta_T^2 T^{-1}\omega^{-2}_{i,T}/2]$, if $0 < \lambda \leq (s/2+1)/(s/2+2)$, and $\Pr[|\sum_{t=1}^T \left(u^2_{x_i,t} - \sigma^2_{x_it}\right)| > \zeta_T] \leq \exp[-C_0\zeta_T^{s/(s+2)}]$, if $\lambda > (s/2+1)/(s/2+2)$. Then, using the same line of reasoning as in the proof of Lemma A6 we establish the desired result. ■

Lemma A8 *Let y_t, for $t = 1, 2, ..., T$, be given by the data generating process (6) and suppose that u_t and $\boldsymbol{x}_{nt} = (x_{1t}, x_{2t}, ..., x_{nt})'$ satisfy Assumptions 2-4, with $s = \min(s_x, s_u) > 0$. Let $\boldsymbol{q}_{\cdot t} = (q_{1,t}, q_{2,t}, ..., q_{l_T,t})'$ contain a constant and a subset of $\boldsymbol{x}_{nt}$. Assume that $\boldsymbol{\Sigma}_{qq} = \frac{1}{T}\sum_{t=1}^T E\left(\boldsymbol{q}_{\cdot t}\boldsymbol{q}'_{\cdot t}\right)$ and $\hat{\boldsymbol{\Sigma}}_{qq} = \mathbf{Q}'\mathbf{Q}/T$ are both invertible, where $\mathbf{Q} = (\boldsymbol{q}_{1\cdot}, \boldsymbol{q}_{2\cdot}, ..., \boldsymbol{q}_{l_T\cdot})$ and $\boldsymbol{q}_{i\cdot} = (q_{i1}, q_{i2}, ..., q_{iT})'$, for $i = 1, 2, ..., l_T$. Moreover, suppose that Assumption 5 holds for x_t and $\boldsymbol{q}_{\cdot t}$, where x_t is a generic element of $\{x_{1t}, x_{2t}, ..., x_{nt}\}$ that does not belong to $\boldsymbol{q}_{\cdot t}$. Denote the corresponding projection residuals defined by (11) as $u_{x,t} = x_t - \boldsymbol{\gamma}'_{qx,T}\boldsymbol{q}_{\cdot t}$, and the projection residuals of y_t on $(\boldsymbol{q}'_{\cdot t}, x_t)'$ as $e_t = y_t - \boldsymbol{\gamma}'_{yqx,T}(\boldsymbol{q}'_{\cdot t}, x_t)'$. Define $\boldsymbol{x} = (x_1, x_2, ..., x_T)'$, and $\mathbf{M}_q = \mathbf{I}_T - \mathbf{Q}(\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}'$, and let $a_T = \ominus\left(T^{\lambda-1}\right)$. Then, for any π in the range $0 < \pi < 1$, and as long as $l_T = \ominus\left(T^d\right)$, such that $0 \leq d < 1/3$, we have, that, if $3d/2 < \lambda \leq (s/2+1)/(s/2+2)$,*

$$\Pr\left(\left|T^{-1}\sigma^{-2}_{x,(T)}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x} - 1\right| > a_T\right) \leq \exp\left[-\sigma^4_{x,(T)}\left(1-\pi\right)^2 T a_T^2 \omega^{-2}_{x,(T)}/2\right] + \exp\left[-C_0 T^{C_1}\right], \text{ and} \tag{B.32}$$

$$\Pr[|(T^{-1}\sigma_{x,(T)}^{-2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x})^{-1/2}-1|>a_T]\leq\exp[-\sigma_{x,(T)}^{4}\left(1-\pi\right)^2Ta_T^2\omega_{x,(T)}^{-2}/2]+\exp\left[-C_0T^{C_1}\right], \tag{B.33}$$

where

$$\sigma_{x,(T)}^2=T^{-1}\textstyle\sum_{t=1}^{T}E\left(u_{x,t}^2\right),\ \omega_{x,(T)}^2=T^{-1}\sum_{t=1}^{T}E\left[\left(u_{x,t}^2-\sigma_{xt}^2\right)^2\right]. \tag{B.34}$$

If $\lambda>(s/2+1)/(s/2+2)$,

$$\Pr(|T^{-1}\sigma_{x,(T)}^{-2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}-1|>a_T)\leq\exp[-C_0\left(Ta_T\right)^{s/(s+2)}]+\exp\left[-C_1T^{C_2}\right], \tag{B.35}$$

and

$$\Pr[|(T^{-1}\sigma_{x,(T)}^{-2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x})^{-1/2}-1|>a_T]\leq\exp[-C_0\left(Ta_T\right)^{s/(s+2)}]+\exp\left[-C_1T^{C_2}\right]. \tag{B.36}$$

Also, if $3d/2<\lambda\leq(s/2+1)/(s/2+2)$,

$$\Pr(|T^{-1}\sigma_{u,(T)}^{-2}\boldsymbol{e}'\boldsymbol{e}-1|>a_T)\leq\exp[-\sigma_{u,(T)}^{4}\left(1-\pi\right)^2Ta_T^2\omega_{u,(T)}^{-2}/2]+\exp\left[-C_0T^{C_1}\right], \tag{B.37}$$

and

$$\Pr[|(\sigma_{u,(T)}^{-2}\boldsymbol{e}'\boldsymbol{e}/T)^{-1/2}-1|>a_T]\leq\exp[-\sigma_{u,(T)}^{4}\left(1-\pi\right)^2Ta_T^2\omega_{u,T}^{-2}/2]+\exp\left[-C_0T^{C_1}\right], \tag{B.38}$$

where $\boldsymbol{e}=(e_1,e_2,...,e_T)'$,

$$\sigma_{u,(T)}^2=T^{-1}\textstyle\sum_{t=1}^{T}\sigma_t^2,\ \textit{and}\ \omega_{u,T}^2=T^{-1}\sum_{t=1}^{T}E[\left(u_t^2-\sigma_t^2\right)^2]. \tag{B.39}$$

If $\lambda>(s/2+1)/(s/2+2)$,

$$\Pr(|T^{-1}\sigma_{u,(T)}^{-2}\boldsymbol{e}'\boldsymbol{e}-1|>a_T)\leq\exp[-C_0\left(Ta_T\right)^{s/(s+2)}]+\exp\left[-C_1T^{C_2}\right],\ \textit{and} \tag{B.40}$$

$$\Pr[|(\sigma_{u,(T)}^{-2}\boldsymbol{e}'\boldsymbol{e}/T)^{-1/2}-1|>a_T]\leq\exp[-C_0\left(Ta_T\right)^{s/(s+2)}]+\exp\left[-C_1T^{C_2}\right], \tag{B.41}$$

Proof. First note that $T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}-\sigma_{x,(T)}^2=T^{-1}\sum_{t=1}^{T}\left(\hat{u}_{x,t}^2-\sigma_{xt}^2\right)$**,** where $\hat{u}_{x,t}$, for $t=1,2,...,T$,. is the t-th element of $\hat{\mathbf{u}}_x=\mathbf{M}_q\boldsymbol{x}$. Now applying Lemma A7 to $\sum_{t=1}^{T}\left(\hat{u}_{x,t}^2-\sigma_{xt}^2\right)$ with $\zeta_T=Ta_T$ we have $\Pr(|\sum_{t=1}^{T}\left(\hat{u}_{x,t}^2-\sigma_{xt}^2\right)|>\zeta_T)\leq\exp[-\left(1-\pi\right)^2\zeta_T^2\omega_{x,(T)}^{-2}/(2T)]+\exp\left[-C_0T^{C_1}\right]$, if $3d/2<\lambda\leq(s/2+1)/(s/2+2)$, and $\Pr(|\sum_{t=1}^{T}\left(\hat{u}_{x,t}^2-\sigma_{xt}^2\right)|>\zeta_T)\leq\exp[-C_0\zeta_T^{s/(s+2)}]+\exp\left[-C_1T^{C_2}\right]$, if $\lambda>(s/2+1)/(s/2+2)$, where $\omega_{x,(T)}^2$ is defined by (B.34). Also $\Pr[|T^{-1}\sigma_{x,(T)}^{-2}\sum_{t=1}^{T}\left(\hat{u}_{x,t}^2-\sigma_{xt}^2\right)|>T^{-1}\sigma_{x,(T)}^{-2}\zeta_T]\leq\exp[-\left(1-\pi\right)^2\zeta_T^2\omega_{x,(T)}^{-2}T^{-1}/2]+\exp\left[-C_0T^{C_1}\right]$, if $3d/2<\lambda\leq(s/2+1)/(s/2+2)$, and $\Pr[|T^{-1}\sigma_{x,(T)}^{-2}\sum_{t=1}^{T}\left(\hat{u}_{x,t}^2-\sigma_{xt}^2\right)|>\zeta_TT^{-1}\sigma_{x,(T)}^{-2}]\leq\exp[-C_0\zeta_T^{s/(s+2)}]+\exp\left[-C_1T^{C_2}\right]$, if $\lambda>(s/2+1)/(s/2+2)$. Therefore, setting $a_T=\zeta_T/T\sigma_{x,(T)}^2=\ominus\left(T^{\lambda-1}\right)$, we have

$$\Pr(|\sigma_{x,(T)}^{-2}T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}-1|>a_T)\leq\exp[-\sigma_{x,(T)}^{4}\left(1-\pi\right)^2Ta_T^2\omega_{x,(T)}^{-2}/2]+\exp\left[-C_0T^{C_1}\right], \tag{B.42}$$

if $3d/2<\lambda\leq(s/2+1)/(s/2+2)$, and

$$\Pr(|\sigma_{x,(T)}^{-2}T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}-1|>a_T)\leq\exp[-C_0\zeta_T^{s/(s+2)}]+\exp\left[-C_1T^{C_2}\right],$$

if $\lambda > (s/2+1)/(s/2+2)$, as required. Now setting $\omega_T = \sigma_{x,(T)}^{-2} T^{-1} \boldsymbol{x}'\mathbf{M}_q \boldsymbol{x}$, and using Lemma A13, we have $\Pr[|(\sigma_{x,(T)}^{-2} T^{-1} \boldsymbol{x}'\mathbf{M}_q \boldsymbol{x})^{-1/2} - 1| > a_T] \leq \Pr(|\sigma_{x,(T)}^{-2} T^{-1} \boldsymbol{x}'\mathbf{M}_q \boldsymbol{x} - 1| > a_T)$, and hence

$$\Pr[|(\sigma_{u,(T)}^{-2} T^{-1} \boldsymbol{x}'\mathbf{M}_q \boldsymbol{x})^{-1/2} - 1| > a_T] \leq \exp[-\sigma_{x,(T)}^4 (1-\pi)^2 T a_T^2 \omega_{x,(T)}^{-2}] + \exp\left[-C_0 T^{C_1}\right], \quad \text{(B.43)}$$

if $3d/2 < \lambda \leq (s/2+1)/(s/2+2)$, and

$$\Pr[|(\sigma_{u,(T)}^{-2} T^{-1} \boldsymbol{x}'\mathbf{M}_q \boldsymbol{x})^{-1/2} - 1| > a_T] \leq \exp[-C_0 \zeta_T^{s/(s+2)}] + \exp\left[-C_1 T^{C_2}\right],$$

if $\lambda > (s/2+1)/(s/2+2)$. Furthermore

$$\Pr[|(\sigma_{x,(T)}^{-2} T^{-1} \boldsymbol{x}'\mathbf{M}_q \boldsymbol{x})^{1/2} - 1| > a_T] = \Pr\left[\frac{|(\sigma_{x,(T)}^{-2} T^{-1} \boldsymbol{x}'\mathbf{M}_q \boldsymbol{x}) - 1|}{(\sigma_{x,(T)}^{-2} T^{-1} \boldsymbol{x}'\mathbf{M}_q \boldsymbol{x})^{1/2} + 1} > a_T\right],$$

and using Lemma A11 for some finite positive constant C, we have $\Pr[|(\sigma_{x,(T)}^{-2} T^{-1} \boldsymbol{x}'\mathbf{M}_q \boldsymbol{x})^{1/2} - 1| > a_T] \leq \Pr[|\sigma_{x,(T)}^{-2} T^{-1} \boldsymbol{x}'\mathbf{M}_q \boldsymbol{x} - 1| > a_T C^{-1}] + \Pr[(\sigma_{x,(T)}^{-2} T^{-1} \boldsymbol{x}'\mathbf{M}_q \boldsymbol{x})^{1/2} + 1 < C^{-1}]$. Let $C = 1$, and note that for this choice of C, $\Pr[(\sigma_{x,(T)}^{-2} T^{-1} \boldsymbol{x}'\mathbf{M}_q \boldsymbol{x})^{1/2} + 1 < C^{-1}] = \Pr[(\sigma_{x,(T)}^{-2} T^{-1} \boldsymbol{x}'\mathbf{M}_q \boldsymbol{x})^{1/2} < 0] = 0$. Hence $\Pr[|(\sigma_{x,(T)}^{-2} T^{-1} \boldsymbol{x}'\mathbf{M}_q \boldsymbol{x})^{1/2} - 1| > a_T] \leq \Pr[|(\sigma_{x,(T)}^{-2} T^{-1} \boldsymbol{x}'\mathbf{M}_q \boldsymbol{x}) - 1| > a_T]$, and using (B.42),

$$\Pr[|(\sigma_{x,(T)}^{-2} T^{-1} \boldsymbol{x}'\mathbf{M}_q \boldsymbol{x})^{1/2} - 1| > a_T] \leq \exp[-\sigma_{x,(T)}^4 (1-\pi)^2 T a_T^2 \omega_{x,(T)}^{-2}/2] + \exp\left[-C_0 T^{C_1}\right], \quad \text{(B.44)}$$

if $3d/2 < \lambda \leq (s/2+1)/(s/2+2)$, and

$$\Pr[|(\sigma_{x,(T)}^{-2} T^{-1} \boldsymbol{x}'\mathbf{M}_q \boldsymbol{x})^{1/2} - 1| > a_T] \leq \exp[-C_0 \zeta_T^{s/(s+2)}] + \exp\left[-C_1 T^{C_2}\right],$$

if $\lambda > (s/2+1)/(s/2+2)$. Consider now $\boldsymbol{e}'\boldsymbol{e} = \sum_{t=1}^T e_t^2$ and note that $|\sum_{t=1}^T (e_t^2 - \sigma_t^2)| \leq |\sum_{t=1}^T (u_t^2 - \sigma_t^2)| + |(T^{-1}\boldsymbol{u}'\mathbf{W})(T^{-1}\mathbf{W}'\mathbf{W})^{-1}(\mathbf{W}'\boldsymbol{u})|$, where $\mathbf{W} = (\mathbf{Q}, \boldsymbol{x})$. As before, applying Lemma A7 to $\sum_{t=1}^T (e_t^2 - \sigma_t^2)$, and following similar lines of reasoning we have

$$\Pr[|\textstyle\sum_{t=1}^T \left(e_t^2 - \sigma_t^2\right)| > \zeta_T] \leq \exp[-(1-\pi)^2 \zeta_T^2 T^{-1} \omega_{u,(T)}^{-2}/2] + \exp\left[-C_0 T^{C_1}\right],$$

if $3d/2 < \lambda \leq (s/2+1)/(s/2+2)$, and

$$\Pr[|\textstyle\sum_{t=1}^T \left(e_t^2 - \sigma_t^2\right)| > \zeta_T] \leq \exp[-C_0 \zeta_T^{s/(s+2)}] + \exp\left[-C_1 T^{C_2}\right],$$

if $\lambda > (s/2+1)/(s/2+2)$, which yield (B.37) and (B.40). Result (B.38) also follows along similar lines as used above to prove (B.33). ■

Lemma A9 *Let y_t, for $t = 1, 2, ..., T$, be given by the data generating process (6) and suppose that u_t and $\boldsymbol{x}_{nt} = (x_{1t}, x_{2t}, ..., x_{nt})'$ satisfy Assumptions 2-4. Let $\boldsymbol{q}_{\cdot t} = (q_{1,t}, q_{2,t}, ..., q_{l_T,t})'$ contain a constant and a subset of $\boldsymbol{x}_{nt} = (x_{1t}, x_{2t}, ..., x_{nt})'$, and $l_T = o(T^{1/3})$. Assume that $\mathbf{\Sigma}_{qq} = \frac{1}{T}\sum_{t=1}^T E\left(\boldsymbol{q}_{\cdot t}\boldsymbol{q}_{\cdot t}'\right)$ and $\hat{\mathbf{\Sigma}}_{qq} = \mathbf{Q}'\mathbf{Q}/T$ are both invertible, where $\mathbf{Q} = (\boldsymbol{q}_{1\cdot}, \boldsymbol{q}_{2\cdot}, ..., \boldsymbol{q}_{l_T\cdot})$ and $\boldsymbol{q}_{i\cdot} =$*

$(q_{i1}, q_{i2}, .., q_{iT})'$, *for* $i = 1, 2, ..., l_T$. *Suppose that Assumption 5 holds for* x_t *and* $\boldsymbol{q}_{\cdot t}$, *where* x_t *is a generic element of* $\{x_{1t}, x_{2t}, ..., x_{nt}\}$ *that does not belong to* $\boldsymbol{q}_{\cdot t}$. *Denote the projection residuals of* y_t *on* $(\boldsymbol{q}'_{\cdot t}, x_t)'$ *as* $e_t = y_t - \boldsymbol{\gamma}'_{yqx,T}(\boldsymbol{q}'_{\cdot t}, x_t)'$. *Define* $\boldsymbol{x} = (x_1, x_2, ..., x_T)'$, $\boldsymbol{e} = (e_1, e_2, ..., e_T)'$, *and* $\mathbf{M}_q = \mathbf{I}_T - \mathbf{Q}(\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}'$. *Moreover, let* $E(\boldsymbol{e}'\boldsymbol{e}/T) = \sigma^2_{e,(T)}$ *and* $E(\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}/T) = \sigma^2_{x,(T)}$. *Then*

$$\Pr\left[\left|\frac{a_T}{\sqrt{(T^{-1}\boldsymbol{e}'\boldsymbol{e})(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x})}}\right| > c_p(n,\delta)\right] \leq \Pr\left(\left|\frac{a_T}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| > \frac{c_p(n,\delta)}{1+d_T}\right) + \exp\left[-C_0 T^{C_1}\right] \tag{B.45}$$

for any random variable a_T, *some finite positive constants* C_0 *and* C_1, *and some bounded sequence* $d_T > 0$, *where* $c_p(n,\delta)$ *is defined in (15). Similarly,*

$$\Pr\left[\left|\frac{a_T}{\sqrt{(T^{-1}\boldsymbol{e}'\boldsymbol{e})}}\right| > c_p(n,\delta)\right] \leq \Pr\left(\left|\frac{a_T}{\sigma_{e,(T)}}\right| > \frac{c_p(n,\delta)}{1+d_T}\right) + \exp\left[-C_0 T^{C_1}\right] \tag{B.46}$$

Proof. We prove (B.45). (B.46) follows similarly. Define

$$g_T = [\sigma^2_{e,(T)}/(T^{-1}\boldsymbol{e}'\boldsymbol{e})]^{1/2} - 1, \; h_T = [\sigma^2_{x,(T)}/(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x})]^{1/2} - 1.$$

Using results in Lemma A11, note that for any $d_T > 0$ bounded in T,

$$\Pr\left[\left|\frac{a_T}{\sqrt{(T^{-1}\boldsymbol{e}'\boldsymbol{e})(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x})}}\right| > c_p(n,\delta)\,|\theta = 0\right] \leq \Pr\left(\left|\frac{a_T}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| > \frac{c_p(n,\delta)}{1+d_T}\right) + \Pr\left(\left|(1+g_T)(1+h_T)\right| > 1+d_T\right).$$

Since $(1+g_T)(1+h_T) > 0$, then

$$\Pr\left(\left|(1+g_T)(1+h_T)\right| > 1+d_T\right) = \Pr\left[(1+g_T)(1+h_T) > 1+d_T\right] = \Pr\left(g_T h_T + g_T + h_T\right) > d_T).$$

Using (B.33), (B.36), (B.38) and (B.41),

$$\Pr\left[|h_T| > d_T\right] \leq \exp\left[-C_0 T^{C_1}\right], \; \Pr\left[|h_T| > c\right] \leq \exp\left[-C_0 T^{C_1}\right],$$
$$\Pr\left[|g_T| > d_T\right] \leq \exp\left[-C_0 T^{C_1}\right], \; \Pr\left[|g_T| > d_T/c\right] \leq \exp\left[-C_0 T^{C_1}\right],$$

for some finite positive constants C_0 and C_1. Using the above results, for some finite positive constants C_0 and C_1, we have,

$$\Pr\left[\left|\frac{a_T}{\sqrt{(T^{-1}\boldsymbol{e}'\boldsymbol{e})(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x})}}\right| > c_p(n,\delta)\,|\theta = 0\right] \leq \Pr\left(\left|\frac{a_T}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| > \frac{c_p(n,\delta)}{1+d_T}\right) + \exp\left[-C_0 T^{C_1}\right],$$

which establishes the desired result. ■

Lemma A10 *Let y_t, for $t = 1, 2, ..., T$, be given by the data generating process (6) and suppose that u_t and $\boldsymbol{x}_{nt} = (x_{1t}, x_{2t}, ..., x_{nt})'$ satisfy Assumptions 2-4, with $s = \min(s_x, s_u) > 0$. Let $\boldsymbol{q}_{\cdot t} = (q_{1,t}, q_{2,t}, ..., q_{l_T,t})'$ contain a constant and a subset of $\boldsymbol{x}_{nt}$, and let $\eta_t = \boldsymbol{x}'_{b,t}\boldsymbol{\beta}_b + u_t$, where $\boldsymbol{x}_{b,t}$ is $k_b \times 1$ dimensional vector of signal variables that do not belong to $\boldsymbol{q}_{\cdot t}$, with the associated coefficients, $\boldsymbol{\beta}_b$. Assume that $\boldsymbol{\Sigma}_{qq} = \frac{1}{T}\sum_{t=1}^{T} E\left(\boldsymbol{q}_{\cdot t}\boldsymbol{q}'_{\cdot t}\right)$ and $\hat{\boldsymbol{\Sigma}}_{qq} = \mathbf{Q}'\mathbf{Q}/T$ are both invertible, where $\mathbf{Q} = (\boldsymbol{q}_{1\cdot}, \boldsymbol{q}_{2\cdot}, ..., \boldsymbol{q}_{l_T\cdot})$ and $\boldsymbol{q}_{i\cdot} = (q_{i1}, q_{i2}, ..., q_{iT})'$, for $i = 1, 2, ..., l_T$. Moreover, let $l_T = o(T^{1/3})$ and suppose that Assumption 5 holds for x_{it} and $\boldsymbol{q}_{\cdot t}$, $i = 1, 2, ..., n$, where x_t is a generic element of $\{x_{1t}, x_{2t}, ..., x_{nt}\}$ that does not belong to $\boldsymbol{q}_{\cdot t}$. Denote the corresponding projection residuals defined by (11) as $u_{x,t} = x_t - \boldsymbol{\gamma}'_{qx,T}\boldsymbol{q}_{\cdot t}$, and the projection residuals of y_t on $(\boldsymbol{q}'_{\cdot t}, x_t)'$ as $e_t = y_t - \boldsymbol{\gamma}'_{yqx,T}(\boldsymbol{q}'_{\cdot t}, x_t)'$. Define $\boldsymbol{x} = (x_1, x_2, ..., x_T)'$, $\boldsymbol{y} = (y_1, y_2, ..., y_T)'$, $\boldsymbol{e} = (e_1, e_2, ..., e_T)'$, $\mathbf{M}_q = \mathbf{I}_T - \mathbf{Q}(\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}'$, and $\theta_T = E\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\mathbf{X}_b\right)\boldsymbol{\beta}_b$, where $\mathbf{X}_b$ is $T \times k_b$ matrix of observations on $\boldsymbol{x}_{b,t}$. Finally, $c_p(n, \delta)$ is given by (15) with $0 < p < 1$ and $f(n, \delta) = cn^{\delta}$, for some $c, \delta > 0$, and there exists $\kappa_1 > 0$ such that $T = \ominus(n^{\kappa_1})$. Then, for any π in the range $0 < \pi < 1$, any $d_T > 0$ and bounded in T, and for some finite positive constants C_0 and C_1,*

$$\Pr\left[|t_x| > c_p(n, \delta) | \theta_T = 0\right] \leq \exp\left[\frac{-(1-\pi)^2 \sigma^2_{e,(T)}\sigma^2_{x,(T)}c_p^2(n, \delta)}{2(1+d_T)^2\omega^2_{xe,T}}\right] + \exp\left[-C_0 T^{C_1}\right], \quad \text{(B.47)}$$

where

$$t_x = \frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sqrt{\left(T^{-1}\boldsymbol{e}'\boldsymbol{e}\right)\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}\right)}}, \quad \text{(B.48)}$$

$$\sigma^2_{e,(T)} = E\left(T^{-1}\boldsymbol{e}'\boldsymbol{e}\right), \ \sigma^2_{x,(T)} = E\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}\right), \quad \text{(B.49)}$$

and

$$\omega^2_{xe,T} = T^{-1}\sum_{t=1}^{T} E\left[(u_{x,t}\eta_t)^2\right]. \quad \text{(B.50)}$$

Under $\sigma_t^2 = \sigma^2$ and/or $E\left(u^2_{x,t}\right) = \sigma^2_{xt} = \sigma^2_x$, for all $t = 1, 2, ..., T$,

$$\Pr\left[|t_x| > c_p(n, \delta) | \theta_T = 0\right] \leq \exp\left[-(1-\pi)^2 c_p^2(n, \delta)(1+d_T)^{-2}/2\right] + \exp\left(-C_0 T^{C_1}\right). \quad \text{(B.51)}$$

In the case where $\theta_T \neq 0$, let $\theta_T = \ominus\left(T^{-\vartheta}\right)$, for some $0 \leq \vartheta < 1/2$, where $c_p(n, \delta) = O\left(T^{1/2-\vartheta-C_8}\right)$, for some positive C_8. Then, for some bounded positive sequence d_T, and for some $C_2, C_3 > 0$, we have

$$\Pr\left[|t_x| > c_p(n, \delta) | \theta_T \neq 0\right] > 1 - \exp\left(-C_2 T^{C_3}\right). \quad \text{(B.52)}$$

Proof. The DGP, given by (7), can be written as $\boldsymbol{y} = a\boldsymbol{\tau}_T + \mathbf{X}_k\boldsymbol{\beta} + \boldsymbol{u} = a\boldsymbol{\tau}_T + \mathbf{X}_a\boldsymbol{\beta}_a + \mathbf{X}_b\boldsymbol{\beta}_b + \boldsymbol{u}$, where $\mathbf{X}_a$ is a subset of $\mathbf{Q}$. Let $\mathbf{Q}_x = (\mathbf{Q}, \boldsymbol{x})$, $\mathbf{M}_q = \mathbf{I}_T - \mathbf{Q}(\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}'$, $\mathbf{M}_{qx} = \mathbf{I}_T - \mathbf{Q}_x(\mathbf{Q}_x'\mathbf{Q}_x)^{-1}\mathbf{Q}_x'$. Then, $\mathbf{M}_q\mathbf{X}_a = \mathbf{0}$, and let $\mathbf{M}_q\mathbf{X}_b = (\boldsymbol{x}_{bq,1}, \boldsymbol{x}_{bq,2}, ..., \boldsymbol{x}_{bq,T})'$. Then,

$$t_x = \frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sqrt{\left(T^{-1}\boldsymbol{e}'\boldsymbol{e}/\right)\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}\right)}} = \frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\mathbf{X}_b\boldsymbol{\beta}_b}{\sqrt{\left(T^{-1}\boldsymbol{e}'\boldsymbol{e}\right)\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}\right)}} + \frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{u}}{\sqrt{\left(T^{-1}\boldsymbol{e}'\boldsymbol{e}\right)\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}\right)}}. \quad \text{(B.53)}$$

Let $\theta_T = E\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\mathbf{X}_b\right)\boldsymbol{\beta}_b$, $\boldsymbol{\eta} = \mathbf{X}_b\boldsymbol{\beta}_b + \boldsymbol{u}$, $\boldsymbol{\eta} = (\eta_1, \eta_2, ..., \eta_T)'$, and write (B.53) as

$$t_x = \frac{\sqrt{T}\theta_T}{\sqrt{\left(T^{-1}\boldsymbol{e}'\boldsymbol{e}/\right)\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}\right)}} + \frac{\sqrt{T}\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta} - \theta_T\right)}{\sqrt{\left(T^{-1}\boldsymbol{e}'\boldsymbol{e}/\right)\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}\right)}}. \tag{B.54}$$

First, consider the case where $\theta_T = 0$ and note that in this case
$t_x = \left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}\right)^{-1/2}\left(T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}\right)\left(T^{-1}\boldsymbol{e}'\boldsymbol{e}\right)^{-1/2}$. Now by Lemma A9, we have

$$\begin{aligned}\Pr\left[|t_x| > c_p(n,\delta)\,|\theta_T = 0\right] &= \Pr\left[\left|\frac{\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}\right)^{-1/2}\left(T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}\right)}{\left(T^{-1}\boldsymbol{e}'\boldsymbol{e}\right)^{1/2}}\right| > c_p(n,\delta)\,|\theta_T = 0\right] \\ &\leq \Pr\left(\left|\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| > \frac{c_p(n,\delta)}{1+d_T}\right) + \exp\left(-C_0T^{C_1}\right).\end{aligned}$$

where $\sigma^2_{e,(T)}$ and $\sigma^2_{x,(T)}$ are defined by (B.49). Hence, noting that $c_p(n,\delta) = o(T^{C_0})$, for all $C_0 > 0$, under Assumption 3, and by Lemma A6, we have

$$\Pr\left[|t_x| > c_p(n,\delta)\,|\theta_T = 0\right] \leq \exp\left[\frac{-(1-\pi)^2\sigma^2_{e,(T)}\sigma^2_{x,(T)}c_p^2(n,\delta)}{2(1+d_T)^2\omega^2_{xe,T}}\right] + \exp\left(-C_0T^{C_1}\right),$$

where $\omega^2_{xe,T} = T^{-1}\sum_{t=1}^T E[(u_{x,t}\eta_t)^2] = T^{-1}\sum_{t=1}^T E[u^2_{x,t}\left(\boldsymbol{x}'_{b,t}\boldsymbol{\beta}_b + u_t\right)^2]$, and $u_{x,t}$, being the error in the regression of x_t on $\mathbf{Q}$, is defined by (11). Since by assumption u_t are distributed independently of $u_{x,t}$ and $\boldsymbol{x}_{b,t}$, then

$$\omega^2_{xe,T} = T^{-1}\sum_{t=1}^T E[u^2_{x,t}(\boldsymbol{x}'_{bq,t}\boldsymbol{\beta}_b)^2] + T^{-1}\sum_{t=1}^T E\left(u^2_{xt}\right)E\left(u^2_t\right),$$

where $\boldsymbol{x}'_{bq,t}\boldsymbol{\beta}_b$ is the t-th element of $\mathbf{M}_q\mathbf{X}_b\boldsymbol{\beta}_b$. Furthermore, $E[u^2_{x,t}(\boldsymbol{x}'_{bq,t}\boldsymbol{\beta}_b)^2] = E\left(u^2_{x,t}\right)E(\boldsymbol{x}'_{bq,t}\boldsymbol{\beta}_b)^2 = E\left(u^2_{x,t}\right)\boldsymbol{\beta}'_bE(\boldsymbol{x}_{bq,t}\boldsymbol{x}'_{bq,t})\boldsymbol{\beta}_b$, noting that under $\theta = 0$, $u_{x,t}$ and $\boldsymbol{x}_{b,t}$ are independently distributed. Hence

$$\omega^2_{xe,T} = T^{-1}\sum_{t=1}^T E\left(u^2_{x,t}\right)\boldsymbol{\beta}'_bE(\boldsymbol{x}_{bq,t}\boldsymbol{x}'_{bq,t})\boldsymbol{\beta}_b + T^{-1}\sum_{t=1}^T E\left(u^2_{xt}\right)E\left(u^2_t\right). \tag{B.55}$$

Similarly

$$\begin{aligned}\sigma^2_{e,(T)} &= E\left(T^{-1}\boldsymbol{e}'\boldsymbol{e}\right) = E\left(T^{-1}\boldsymbol{\eta}'\mathbf{M}_{qx}\boldsymbol{\eta}\right) = E[T^{-1}\left(\mathbf{X}_b\boldsymbol{\beta}_b + \boldsymbol{u}\right)'\mathbf{M}_{qx}\left(\mathbf{X}_b\boldsymbol{\beta}_b + \boldsymbol{u}\right)] \\ &= \boldsymbol{\beta}'_bE\left(T^{-1}\mathbf{X}'_b\mathbf{M}_{qx}\mathbf{X}_b\right)\boldsymbol{\beta}_b + T^{-1}\sum_{t=1}^T E\left(u^2_t\right),\end{aligned}$$

and since under $\theta = 0$, $\boldsymbol{x}$ being a noise variable will be distributed independently of $\mathbf{X}_b$, then $E\left(T^{-1}\mathbf{X}'_b\mathbf{M}_{qx}\mathbf{X}_b\right) = E\left(T^{-1}\mathbf{X}'_b\mathbf{M}_q\mathbf{X}_b\right)$, and we have

$$\sigma^2_{e,(T)} = \boldsymbol{\beta}'_bE\left(T^{-1}\mathbf{X}'_b\mathbf{M}_q\mathbf{X}_b\right)\boldsymbol{\beta}_b + T^{-1}\sum_{t=1}^T E\left(u^2_t\right) = T^{-1}\sum_{t=1}^T\boldsymbol{\beta}'_bE\left(\boldsymbol{x}_{bq,t}\boldsymbol{x}'_{bq,t}\right)\boldsymbol{\beta}_b + T^{-1}\sum_{t=1}^T E\left(u^2_t\right). \tag{B.56}$$

Using (B.55) and (B.56), it is now easily seen that if either $E\left(u^2_{x,t}\right) = \sigma^2_{ux}$ or $E\left(u^2_t\right) = \sigma^2$, for all t, then we have $\omega^2_{xe,T} = \sigma^2_{e,(T)}\sigma^2_{x,(T)}$, and hence

$$\Pr\left[|t_x| > c_p(n,\delta)\,|\theta_T = 0\right] \leq \exp\left[-(1-\pi)^2c_p^2(n,\delta)(1+d_T)^{-2}/2\right] + \exp\left(-C_0T^{C_1}\right),$$

giving a rate that does not depend on error variances. Next, we consider $\theta_T \neq 0$. By (B.45) of Lemma A9, for $d_T > 0$ and bounded in T,

$$\Pr\left[\left|\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sqrt{(T^{-1}\boldsymbol{e}'\boldsymbol{e}/)\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}\right)}}\right| > c_p(n,\delta)\right] \leq \Pr\left(\left|\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| > \frac{c_p(n,\delta)}{1+d_T}\right) + \exp\left(-C_0T^{C_1}\right).$$

We then have

$$\begin{aligned}\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sigma_{e,(T)}\sigma_{x,(T)}} &= \frac{T^{1/2}\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\mathbf{X}_b\boldsymbol{\beta}_b - \theta_T\right)}{\sigma_{e,(T)}\sigma_{x,(T)}} + \frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{u}}{\sigma_{e,(T)}\sigma_{x,(T)}} + \frac{T^{1/2}\theta_T}{\sigma_{e,(T)}\sigma_{x,(T)}} \\ &= \frac{T^{1/2}\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta} - \theta_T\right)}{\sigma_{e,(T)}\sigma_{x,(T)}} + \frac{T^{1/2}\theta_T}{\sigma_{e,(T)}\sigma_{x,(T)}}.\end{aligned}$$

Then $\Pr[|T^{1/2}\sigma_{e,(T)}^{-1}\sigma_{x,(T)}^{-1}\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta} - \theta_T\right) + T^{1/2}\sigma_{e,(T)}^{-1}\sigma_{x,(T)}^{-1}\theta_T| > c_p(n,\delta)/(1+d_T)] = 1 - \Pr\left[\left|T^{1/2}\sigma_{e,(T)}^{-1}\sigma_{x,(T)}^{-1}\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta} - \theta_T\right) + T^{1/2}\sigma_{e,(T)}^{-1}\sigma_{x,(T)}^{-1}\theta_T\right| \leq c_p(n,\delta)/(1+d_T)\right]$. Note that since $c_p(n,\delta)$ is given by (15), then, $T^{1/2}\left|\theta_T\right|/(\sigma_{e,(T)}\sigma_{x,(T)}) - c_p(n,\delta)/(1+d_T) > 0$. Then by Lemma A12,

$$\begin{aligned}&\Pr\left[\left|\frac{T^{1/2}\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta} - \theta_T\right)}{\sigma_{e,(T)}\sigma_{x,(T)}} + \frac{T^{1/2}\theta_T}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| \leq \frac{c_p(n,\delta)}{1+d_T}\right] \\ &\leq \Pr\left[\left|\frac{T^{1/2}\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta} - \theta_T\right)}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| > \frac{T^{1/2}\left|\theta_T\right|}{\sigma_{e,(T)}\sigma_{x,(T)}} - \frac{c_p(n,\delta)}{1+d_T}\right].\end{aligned}$$

But, setting $\zeta_T = T^{1/2}\left[T^{1/2}\left|\theta_T\right|/[\sigma_{e,(T)}\sigma_{x,(T)}] - c_p(n,\delta)/(1+d_T)\right]$ and noting that $\theta_T = O(T^{-\vartheta})$, $0 \leq \vartheta < 1/2$, implies that this choice of ζ_T satisfies $\zeta_T = \ominus\left(T^{\lambda}\right)$ with $\lambda = 1-\vartheta$, (B.16) of Lemma A6 applies regardless of $s > 0$, which gives us

$$\begin{aligned}&\Pr\left[\left|\frac{T^{1/2}\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta} - \theta_T\right)}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| > \frac{T^{1/2}\left|\theta_T\right|}{\sigma_{e,(T)}\sigma_{x,(T)}} - \frac{c_p(n,\delta)}{1+d_T}\right] \\ &\leq C_4\exp\left\{-C_5\left[T^{1/2}\left(\frac{T^{1/2}\left|\theta_T\right|}{\sigma_{e,(T)}\sigma_{x,(T)}} - \frac{c_p(n,\delta)}{1+d_T}\right)\right]^{s/(s+2)}\right\} + \exp\left(-C_6T^{C_7}\right),\end{aligned} \tag{B.57}$$

for some C_4, C_5, C_6 and $C_7 > 0$. Hence, as long as the assumption that $c_p(n,\delta) = O\left(T^{1/2-\vartheta-C_8}\right)$ holds, for some positive C_8, there must exist positive finite constants C_2 and C_3, such that

$$\Pr\left[\left|\frac{T^{1/2}\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta} - \theta\right)}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| > \frac{T^{1/2}\left|\theta_T\right|}{\sigma_{e,(T)}\sigma_{x,(T)}} - \frac{c_p(n,\delta)}{1+d_T}\right] \leq \exp\left(-C_2T^{C_3}\right) \tag{B.58}$$

for any $s > 0$. So overall

$$\Pr\left[\left|\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sqrt{(T^{-1}\boldsymbol{e}'\boldsymbol{e}/)\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}\right)}}\right| > c_p(n,\delta)\right] > 1 - \exp\left(-C_2T^{C_3}\right).$$

■

Lemma A11 *Let X_{iT}, for $i = 1, 2, ..., l_T$, Y_T and Z_T be random variables. Then, for some finite positive constants C_0, C_1 and C_2, and any constants π_i, for $i = 1, 2, ..., l_T$, satisfying $0 < \pi_i < 1$ and $\sum_{i=1}^{l_T} \pi_i = 1$, we have*

$$\Pr\left(\sum_{i=1}^{l_T} |X_{iT}| > C_0\right) \leq \sum_{i=1}^{l_T} \Pr\left(|X_{iT}| > \pi_i C_0\right), \tag{B.59}$$

$$\Pr\left(|X_T| \times |Y_T| > C_0\right) \leq \Pr\left(|X_T| > C_0/C_1\right) + \Pr\left(|Y_T| > C_1\right), \tag{B.60}$$

and

$$\Pr\left(|X_T| \times |Y_T| \times |Z_T| > C_0\right) \leq \Pr\left(|X_T| > C_0/\left(C_1 C_2\right)\right) + \Pr\left(|Y_T| > C_1\right) + \Pr\left(|Z_T| > C_2\right). \tag{B.61}$$

Proof. Without loss of generality we consider the case $l_T = 2$. Consider the two random variables X_{1T} and X_{2T}. Then, for some finite positive constants C_0 and C_1, we have

$$\begin{aligned}\Pr\left(|X_{1T}| + |X_{2T}| > C_0\right) &\leq \Pr\left(\{|X_{1T}| > (1-\pi)C_0\} \cup \{|X_{2T}| > \pi C_0\}\right) \\ &\leq \Pr\left(|X_{1T}| > (1-\pi)C_0\right) + \Pr\left(|X_{2T}| > \pi C_0\right),\end{aligned}$$

proving the first result of the lemma.

Define events $\mathfrak{H} = \{|X_T| \times |Y_T| > C_0\}$, $\mathfrak{B}$= $\{|X_T| > C_0/C_1\}$ and $\mathfrak{C} = \{|Y_T| > C_1\}$. Then $\mathfrak{H} \subset (\mathfrak{B} \cup \mathfrak{C})$, namely $\mathfrak{H}$ must be contained in $\mathfrak{B} \cup \mathfrak{C}$. Hence $P(\mathfrak{H}) \leq P(\mathfrak{B} \cup \mathfrak{C})$. But $P(\mathfrak{B} \cup \mathfrak{C}) \leq P(\mathfrak{B}) + P(\mathfrak{C})$. Therefore, $P(\mathfrak{H}) \leq P(\mathfrak{B}) + P(\mathfrak{C})$, proving the second result of the lemma. The third result follows by a repeated application of the second result. ■

Lemma A12 *Consider the scalar random variable X, and the constants B and C. Then, if $|B| \geq C > 0$,*

$$\Pr\left(|X + B| \leq C\right) \leq \Pr\left(|X| > |B| - C\right). \tag{B.62}$$

Proof. We note that the event we are concerned with is of the form $\mathcal{A} = \{|X + B| \leq C\}$. Consider two cases: (i) $B > 0$. Then, $\mathcal{A}$ can occur only if $X < 0$ and $|X| > B - C = |B| - C$. (ii) $B < 0$. Then, $\mathcal{A}$ can occur only if $X > 0$ and $X = |X| > |B| - C$. It therefore follows that the event $\{|X| > |B| - C\}$ implies $\mathcal{A}$ proving (B.62). ■

Lemma A13 *Consider the scalar random variable, ω_T, and the deterministic sequence, $\alpha_T > 0$, such that $\alpha_T \to 0$ as $T \to \infty$. Then there exists $T_0 > 0$ such that for all $T > T_0$ we have*

$$\Pr\left(\left|\frac{1}{\sqrt{\omega_T}} - 1\right| > \alpha_T\right) \leq \Pr\left(|\omega_T - 1| > \alpha_T\right). \tag{B.63}$$

Proof. We first note that for $\alpha_T < 1/2$

$$\left|\frac{1}{\sqrt{\omega_T}} - 1\right| < |\omega_T - 1| \text{ for any } \omega_T \in [1-\alpha_T, 1+\alpha_T].$$

Also, since $a_T \to 0$ then there must exist a $T_0 > 0$ such that $a_T < 1/2$, for all $T > T_0$, and hence if event $A : |\omega_T - 1| \leq a_T$ is satisfied, then it must be the case that event $B : \left|\frac{1}{\sqrt{\omega_T}} - 1\right| \leq a_T$ is also satisfied for all $T > T_0$. Further, since $A \Rightarrow B$, then $B^c \Rightarrow A^c$, where A^c denotes the complement of A. Therefore, $\left|\frac{1}{\sqrt{\omega_T}} - 1\right| > a_T$ implies $|\omega_T - 1| > a_T$, for all $T > T_0$, and we have $\Pr\left(\left|\frac{1}{\sqrt{\omega_T}} - 1\right| > \alpha_T\right) \leq \Pr(|\omega_T - 1| > \alpha_T)$, as required. ■

Lemma A14 *Let $\mathbf{A}_T = (a_{ij,T})$ be a symmetric $l_T \times l_T$ matrix with eigenvalues $\mu_1 \leq \mu_2 \leq ... \leq \mu_{l_T}$. Let $\mu_i = \ominus(l_T)$, $i = l_T - M + 1, l_T - M + 2, ..., l_T$, for some finite M, and $\sup_{1\leq i \leq l_T - M} \mu_i < C_0 < \infty$, for some finite positive C_0. Then,*

$$\|\mathbf{A}_T\|_F = \ominus(l_T). \tag{B.64}$$

If, in addition, $\inf_{1 \leq i < l_T} \mu_i > C_1 > 0$, for some finite positive C_1, then

$$\left\|\mathbf{A}_T^{-1}\right\|_F = \ominus\left(\sqrt{l_T}\right). \tag{B.65}$$

Proof. We have

$$\|\mathbf{A}_T\|_F^2 = Tr\left(\mathbf{A}_T\mathbf{A}_T'\right) = Tr\left(\mathbf{A}_T^2\right) = \sum_{i=1}^{l_T} \mu_i^2,$$

where μ_i, for $i = 1, 2, ..., l_T$, are the eigenvalues of $\mathbf{A}_T$. But by assumption $\mu_i = \ominus(l_T)$, for $i = l_T - M + 1, l_T - M + 2, ..., l_T$, and $\sup_{1\leq i\leq l_T - M} \mu_i < C_0 < \infty$, then $\sum_{i=1}^{l_T} \mu_i^2 = M \ominus(l_T^2) + O(l_T - M) = \ominus(l_T^2)$, and since M is fixed then (B.64) follows. Note that $\mathbf{A}_T^{-1}$ is also symmetric, and using similar arguments as above, we have

$$\left\|\mathbf{A}_T^{-1}\right\|_F^2 = Tr\left(\mathbf{A}_T^{-2}\right) = \sum_{i=1}^{l_T} \mu_i^{-2},$$

but all eigenvalues of $\mathbf{A}_T$ are bounded away from zero under the assumptions of the lemma, which implies $\mu_i^{-2} = \ominus(1)$ and therefore $\left\|\mathbf{A}_T^{-1}\right\|_F = \ominus\left(\sqrt{l_T}\right)$, which establishes (B.65). ■

Lemma A15 *Let z be a random variable and suppose there exists finite positive constants C_0, C_1 and $s > 0$ such that*

$$\Pr(|z| > \alpha) \leq C_0 \exp\left(-C_1 \alpha^s\right), \text{ for all } \alpha > 0. \tag{B.66}$$

Then for any finite $p > 0$ and p/s finite, there exists $C_2 > 0$ such that

$$E|z|^p \leq C_2. \tag{B.67}$$

Proof. We have that

$$E|z|^p = \int_0^\infty \alpha^p d\Pr(|z| \le \alpha).$$

Using integration by parts, we get

$$\int_0^\infty \alpha^p d\Pr(|z| \le \alpha) = p\int_0^\infty \alpha^{p-1}\Pr(|z| > \alpha)\, d\alpha.$$

But, using (B.66), and a change of variables, implies

$$E|z|^p \le pC_0 \int_0^\infty \alpha^{p-1}\exp(-C_1\alpha^s)\, d\alpha = \frac{pC_0}{s}\int_0^\infty u^{\frac{p-s}{s}}\exp(-C_1 u)\, du = C_0 C_1^{-p/s}\left(\frac{p}{s}\right)\Gamma\left(\frac{p}{s}\right),$$

where $\Gamma(\cdot)$ is a gamma function. But for a finite positive p/s, $\Gamma(p/s)$ is bounded and (B.67) follows. ■

Lemma A16 *Let* $\boldsymbol{A}_T = (a_{ij,T})$ *be an* $l_T \times l_T$ *matrix and* $\hat{\boldsymbol{A}}_T = (\hat{a}_{ij,T})$ *be an estimator of* $\boldsymbol{A}_T$*. Suppose that* $\boldsymbol{A}_T$ *is invertible and there exists a finite positive* C_0*, such that*

$$\sup_{i,j}\Pr(|\hat{a}_{ij,T} - a_{ij,T}| > b_T) \le \exp\left(-C_0 T b_T^2\right), \tag{B.68}$$

for all $b_T > 0$*. Then*

$$\Pr\left(\left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F > b_T\right) \le l_T^2 \exp\left(-C_0\frac{Tb_T^2}{l_T^2}\right), \tag{B.69}$$

and

$$\begin{aligned}\Pr\left(\left\|\hat{\boldsymbol{A}}_T^{-1} - \boldsymbol{A}_T^{-1}\right\|_F > b_T\right) \le\ & l_T^2 \exp\left(\frac{-C_0 T b_T^2}{l_T^2\left\|\boldsymbol{A}_T^{-1}\right\|_F^2\left(\left\|\boldsymbol{A}_T^{-1}\right\|_F + b_T\right)^2}\right)\\ & + l_T^2\exp\left(-C_0\frac{T}{\left\|\boldsymbol{A}_T^{-1}\right\|_F^2 l_T^2}\right).\end{aligned} \tag{B.70}$$

Proof. First note that since $b_T > 0$, then

$$\begin{aligned}\Pr\left(\left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F > b_T\right) &= \Pr\left(\left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F^2 > b_T^2\right)\\ &= \Pr\left(\left[\sum_{j=1}^{l_T}\sum_{i=1}^{l_T}(\hat{a}_{ij,T} - a_{ij,T})^2 > b_T^2\right]\right),\end{aligned}$$

and using the probability bound result, (B.59), and setting $\pi_i = 1/l_T$, we have

$$\begin{aligned}\Pr\left(\left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F > b_T\right) &\le \sum_{j=1}^{l_T}\sum_{i=1}^{l_T}\Pr\left(|\hat{a}_{ij,T} - a_{ij,T}|^2 > l_T^{-2}b_T^2\right)\\ &= \sum_{j=1}^{l_T}\sum_{i=1}^{l_T}\Pr\left(|\hat{a}_{ij,T} - a_{ij,T}| > l_T^{-1}b_T\right)\\ &\le l_T^2 \sup_{ij=1,2,\ldots,l_T}\left[\Pr\left(|\hat{a}_{ij,T} - a_{ij,T}| > l_T^{-1}b_T\right)\right].\end{aligned}$$

Hence by (B.68) we obtain (B.69). To establish (B.70) define the events

$$\mathcal{A}_T = \left\{ \|\boldsymbol{A}_T^{-1}\|_F \left\| \hat{\boldsymbol{A}}_T - \boldsymbol{A}_T \right\|_F < 1 \right\} \text{ and } \mathcal{B}_T = \left\{ \left\| \hat{\boldsymbol{A}}_T^{-1} - \boldsymbol{A}_T^{-1} \right\|_F > b_T \right\}$$

and note that by (2.15) of Berk (1974) if $\mathcal{A}_T$ holds we have

$$\left\| \hat{\boldsymbol{A}}_T^{-1} - \boldsymbol{A}_T^{-1} \right\|_F \leq \frac{\|\boldsymbol{A}_T^{-1}\|_F^2 \left\| \hat{\boldsymbol{A}}_T - \boldsymbol{A}_T \right\|_F}{1 - \|\boldsymbol{A}_T^{-1}\|_F \left\| \hat{\boldsymbol{A}}_T - \boldsymbol{A}_T \right\|_F}. \tag{B.71}$$

Hence

$$\begin{aligned} \Pr\left(\mathcal{B}_T \left| \mathcal{A}_T \right.\right) &\leq \Pr\left(\frac{\|\boldsymbol{A}_T^{-1}\|_F^2 \left\| \hat{\boldsymbol{A}}_T - \boldsymbol{A}_T \right\|_F}{1 - \|\boldsymbol{A}_T^{-1}\|_F \left\| \hat{\boldsymbol{A}}_T - \boldsymbol{A}_T \right\|_F} > b_T \right) \\ &= \Pr\left(\left\| \hat{\boldsymbol{A}}_T - \boldsymbol{A}_T \right\|_F > \frac{b_T}{\|\boldsymbol{A}_T^{-1}\|_F \left(\|\boldsymbol{A}_T^{-1}\|_F + b_T \right)} \right). \end{aligned} \tag{B.72}$$

Note also that

$$\Pr\left(\mathcal{B}_T\right) = \Pr\left(\left\{ \mathcal{B}_T \cap \mathcal{A}_T \right\} \cup \left\{ \mathcal{B}_T \cap \mathcal{A}_T^C \right\} \right) = \Pr\left(\mathcal{B}_T | \mathcal{A}_T\right) \Pr\left(\mathcal{A}_T\right) + \Pr\left(\mathcal{B}_T | \mathcal{A}_T^C\right) \Pr\left(\mathcal{A}_T^C\right). \tag{B.73}$$

Furthermore

$$\begin{aligned} \Pr\left(\mathcal{A}_T^C\right) &= \Pr\left(\|\boldsymbol{A}_T^{-1}\|_F \left\| \hat{\boldsymbol{A}}_T - \boldsymbol{A}_T \right\|_F > 1 \right) \\ &= \Pr\left(\left\| \hat{\boldsymbol{A}}_T - \boldsymbol{A}_T \right\|_F > \|\boldsymbol{A}_T^{-1}\|_F^{-1} \right), \end{aligned}$$

and by (B.69) we have

$$\Pr\left(\mathcal{A}_T^C\right) \leq l_T^2 \exp\left(-C_0 \frac{T}{\|\boldsymbol{A}_T^{-1}\|_F^2 \, l_T^2} \right).$$

Using the above result and (B.72) in (B.73), we now have

$$\begin{aligned} \Pr\left(\mathcal{B}_T\right) \leq \Pr\left(\left\| \hat{\boldsymbol{A}}_T - \boldsymbol{A}_T \right\|_F > \frac{b_T}{\|\boldsymbol{A}_T^{-1}\|_F \left(\|\boldsymbol{A}_T^{-1}\|_F + b_T \right)} \right) \Pr\left(\mathcal{A}_T\right) \\ + \Pr\left(\mathcal{B}_T | \mathcal{A}_T^C\right) l_T^2 \exp\left(-C_0 \frac{T}{\|\boldsymbol{A}_T^{-1}\|_F^2 \, l_T^2} \right). \end{aligned}$$

Furthermore, since $\Pr\left(\mathcal{A}_T\right) \leq 1$ and $\Pr\left(\mathcal{B}_T | \mathcal{A}_T^C\right) \leq 1$ then

$$\begin{aligned} \Pr\left(\mathcal{B}_T\right) = \Pr\left(\left\| \hat{\boldsymbol{A}}_T^{-1} - \boldsymbol{A}_T^{-1} \right\|_F > b_T \right) \leq \Pr\left(\left\| \hat{\boldsymbol{A}}_T - \boldsymbol{A}_T \right\|_F > \frac{b_T}{\|\boldsymbol{A}_T^{-1}\|_F \left(\|\boldsymbol{A}_T^{-1}\|_F + b_T \right)} \right) \\ + l_T^2 \exp\left(-C_0 \frac{T}{\|\boldsymbol{A}_T^{-1}\|_F^2 \, l_T^2} \right). \end{aligned}$$

Result (B.70) now follows if we apply (B.69) to the first term on the RHS of the above. ■

Lemma A17 *Let $\boldsymbol{A}_T = (a_{ij,T})$ be a $l_T \times l_T$ matrix and $\hat{\boldsymbol{A}}_T = (\hat{a}_{ij,T})$ be an estimator of $\boldsymbol{A}_T$. Let $\left\|\boldsymbol{A}_T^{-1}\right\|_F > 0$ and suppose that for some $s > 0$, any $b_T > 0$ and some finite positive constant C_0,*

$$\sup_{i,j} \Pr\left(|\hat{a}_{ij,T} - a_{ij,T}| > b_T\right) \leq \exp\left[-C_0 \left(T b_T\right)^{s/(s+2)}\right].$$

Then

$$\Pr\left(\left\|\hat{\boldsymbol{A}}_T^{-1} - \boldsymbol{A}_T^{-1}\right\|_F > b_T\right) \leq l_T^2 \exp\left(\frac{-C_0 \left(T b_T\right)^{s/(s+2)}}{l_T^{s/(s+2)} \left\|\boldsymbol{A}_T^{-1}\right\|_F^{s/(s+2)} \left(\left\|\boldsymbol{A}_T^{-1}\right\|_F + b_T\right)^{s/(s+2)}}\right) \tag{B.74}$$
$$+ l_T^2 \exp\left(-C_0 \frac{T^{s/(s+2)}}{\left\|\boldsymbol{A}_T^{-1}\right\|_F^{s/(s+2)} l_T^{s/(s+2)}}\right).$$

Proof. First note that since $b_T > 0$, then

$$\Pr\left(\left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F > b_T\right) = \Pr\left(\left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F^2 > b_T^2\right)$$
$$= \Pr\left[\sum_{j=1}^{l_T} \sum_{i=1}^{l_T} \left(\hat{a}_{ij,T} - a_{ij,T}\right)^2 > b_T^2\right],$$

and using the probability bound result, (B.59), and setting $\pi_i = 1/l_T^2$, we have

$$\Pr\left(\left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F > b_T\right) \leq \sum_{j=1}^{l_T} \sum_{i=1}^{l_T} \Pr\left(\left|\hat{a}_{ij,T} - a_{ij,T}\right|^2 > l_T^{-2} b_T^2\right) \tag{B.75}$$
$$= \sum_{j=1}^{l_T} \sum_{i=1}^{l_T} \Pr\left(\left|\hat{a}_{ij,T} - a_{ij,T}\right| > l_T^{-1} b_T\right)$$
$$\leq l_T^2 \sup_{ij} \left[\Pr\left(\left|\hat{a}_{ij,T} - a_{ij,T}\right| > l_T^{-1} b_T\right)\right] = l_T^2 \exp\left(-C_0 T^{s/(s+1)} \frac{b_T^{s/(s+2)}}{l_t^{s/(s+2)}}\right).$$

To establish (B.74) define the events

$$\mathcal{A}_T = \left\{\left\|\boldsymbol{A}_T^{-1}\right\|_F \left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F < 1\right\} \text{ and } \mathcal{B}_T = \left\{\left\|\hat{\boldsymbol{A}}_T^{-1} - \boldsymbol{A}_T^{-1}\right\| > b_T\right\}$$

and note that by (2.15) of Berk (1974) if $\mathcal{A}_T$ holds we have

$$\left\|\hat{\boldsymbol{A}}_T^{-1} - \boldsymbol{A}_T^{-1}\right\| \leq \frac{\left\|\boldsymbol{A}_T^{-1}\right\|_F^2 \left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F}{1 - \left\|\boldsymbol{A}_T^{-1}\right\|_F \left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F}.$$

Hence

$$\Pr\left(\mathcal{B}_T \left|\mathcal{A}_T\right.\right) \leq \Pr\left(\frac{\left\|\boldsymbol{A}_T^{-1}\right\|_F^2 \left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F}{1 - \left\|\boldsymbol{A}_T^{-1}\right\|_F \left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F} > b_T\right)$$
$$= \Pr\left[\left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F > \frac{b_T}{\left\|\boldsymbol{A}_T^{-1}\right\|_F \left(\left\|\boldsymbol{A}_T^{-1}\right\|_F + b_T\right)}\right].$$

Note also that

$$\Pr(\mathcal{B}_T) = \Pr\left(\{\mathcal{B}_T \cap \mathcal{A}_T\} \cup \left\{\mathcal{B}_T \cap \mathcal{A}_T^C\right\}\right) = \Pr(\mathcal{B}_T|\mathcal{A}_T)\Pr(\mathcal{A}_T) + \Pr\left(\mathcal{B}_T|\mathcal{A}_T^C\right)\Pr\left(\mathcal{A}_T^C\right)$$

Furthermore

$$\begin{aligned}\Pr\left(\mathcal{A}_T^C\right) &= \Pr\left(\left\|\boldsymbol{A}_T^{-1}\right\|_F \left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F > 1\right) \\ &= \Pr\left(\left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F > \left\|\boldsymbol{A}_T^{-1}\right\|_F^{-1}\right),\end{aligned}$$

and by (B.75) we have

$$\Pr\left(\mathcal{A}_T^C\right) \leq l_T^2 \exp\left(-C_0 T^{s/(s+1)} \frac{b_T^{s/(s+2)}}{l_t^{s/(s+2)}}\right) = \exp\left(-C_0 \frac{T^{s/(s+2)}}{\left\|\boldsymbol{A}_T^{-1}\right\|_F^{s/(s+2)} l_T^{s/(s+2)}}\right).$$

Using the above result, we now have

$$\begin{aligned}\Pr(\mathcal{B}_T) \leq{}& \Pr\left(\left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F > \frac{b_T}{\left\|\boldsymbol{A}_T^{-1}\right\|_F \left(\left\|\boldsymbol{A}_T^{-1}\right\|_F + b_T\right)}\right)\Pr(\mathcal{A}_T) \\ &+ \Pr\left(\mathcal{B}_T|\mathcal{A}_T^C\right)\exp\left(-C_0 \frac{T^{s/(s+2)}}{\left\|\boldsymbol{A}_T^{-1}\right\|_F^{s/(s+2)} l_T^{s/(s+2)}}\right).\end{aligned}$$

Furthermore, since $\Pr(\mathcal{A}_T) \leq 1$ and $\Pr\left(\mathcal{B}_T|\mathcal{A}_T^C\right) \leq 1$ then

$$\begin{aligned}\Pr(\mathcal{B}_T) = \Pr\left(\left\|\hat{\boldsymbol{A}}_T^{-1} - \boldsymbol{A}_T^{-1}\right\| > b_T\right) \leq{}& \Pr\left(\left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F > \frac{b_T}{\left\|\boldsymbol{A}_T^{-1}\right\|_F \left(\left\|\boldsymbol{A}_T^{-1}\right\|_F + b_T\right)}\right) \\ &+ \exp\left(-C_0 \frac{T^{s/(s+2)}}{\left\|\boldsymbol{A}_T^{-1}\right\|_F^{s/(s+2)} l_T^{s/(s+2)}}\right).\end{aligned}$$

Result (B.74) now follows if we apply (B.75) to the first term on the RHS of the above. ■

Lemma A18 *Let* $\mathbf{S}_a$ *and* $\mathbf{S}_b$*, respectively, be* $T \times l_{a,T}$ *and* $T \times l_{b,T}$ *matrices of observations on* $s_{a,it}$*, and* $s_{b,it}$*, for* $i = 1, 2, ..., l_T$*,* $t = 1, 2, ..., T$*, and suppose that* $\{s_{a,it}, s_{b,it}\}$ *are either non-stochastic and bounded, or random with finite* 8^{th} *order moments. Consider the sample covariance matrix* $\hat{\boldsymbol{\Sigma}}_{ab} = T^{-1}\mathbf{S}_a'\mathbf{S}_b$ *and denote its expectations by* $\boldsymbol{\Sigma}_{ab} = T^{-1}E(\mathbf{S}_a'\mathbf{S}_b)$*. Let*

$$z_{ij,t} = s_{a,it}s_{b,jt} - E(s_{a,it}s_{b,jt}),$$

and suppose that

$$\sup_{i,j}\left[\sum_{t=1}^{T}\sum_{t'=1}^{T} E(z_{ij,t}z_{ij,t'})\right] = O(T). \tag{B.76}$$

Then,

$$E\left\|\hat{\boldsymbol{\Sigma}}_{ab} - \boldsymbol{\Sigma}_{ab}\right\|_F^2 = O\left(\frac{l_{a,T}l_{b,T}}{T}\right). \tag{B.77}$$

If, in addition,

$$\sup_{i,j,i',j'}\left[\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{s=1}^{T}\sum_{s'=1}^{T}E(z_{ij,t}z_{ij,t'}z_{i'j',s}z_{i'j',s'})\right]=O\left(T^{2}\right), \tag{B.78}$$

then

$$E\left\|\hat{\boldsymbol{\Sigma}}_{ab}-\boldsymbol{\Sigma}_{ab}\right\|_{F}^{4}=O\left(\frac{l_{a,T}^{2}l_{b,T}^{2}}{T^{2}}\right). \tag{B.79}$$

Proof. We first note that $E(z_{ij,t}z_{ij,t'})$ and $E\left(z_{ij,t}z_{ij,t'}z_{i'j',s}z_{i'j',s'}\right)$ exist since by assumption $\{s_{a,it},s_{b,it}\}$ have finite 8^{th} order moments. The (i,j) element of $\hat{\boldsymbol{\Sigma}}_{ab}-\boldsymbol{\Sigma}_{ab}$ is given by

$$a_{ij,T}=T^{-1}\sum_{t=1}^{T}z_{ij,t}, \tag{B.80}$$

and hence

$$\begin{aligned}E\left\|\hat{\boldsymbol{\Sigma}}_{ab}-\boldsymbol{\Sigma}_{ab}\right\|_{F}^{2}&=\sum_{i=1}^{l_{a,T}}\sum_{j=1}^{l_{b,T}}E\left(a_{ij,T}^{2}\right)=T^{-2}\sum_{i=1}^{l_{a,T}}\sum_{j=1}^{l_{b,T}}\sum_{t=1}^{T}\sum_{t'=1}^{T}E\left(z_{ij,t}z_{ij,t'}\right)\\&\leq\frac{l_{a,T}l_{b,T}}{T^{2}}\sup_{i,j}\left[\sum_{t=1}^{T}\sum_{t'=1}^{T}E(z_{ij,t}z_{ij,t'})\right],\end{aligned}$$

and (B.77) follows from (B.76). Similarly,

$$\begin{aligned}\left\|\hat{\boldsymbol{\Sigma}}_{ab}-\boldsymbol{\Sigma}_{ab}\right\|_{F}^{4}&=\left(\sum_{i=1}^{l_{a,T}}\sum_{j=1}^{l_{b,T}}a_{ij,T}^{2}\right)^{2}\\&=\sum_{i=1}^{l_{a,T}}\sum_{j=1}^{l_{b,T}}\sum_{i'=1}^{l_{a,T}}\sum_{j'=1}^{l_{b,T}}a_{ij,T}^{2}a_{i'j',T}^{2}.\end{aligned}$$

But using (B.80) we have

$$\begin{aligned}a_{ij,T}^{2}a_{i'j',T}^{2}&=T^{-4}\left(\sum_{t=1}^{T}\sum_{t'=1}^{T}z_{ij,t}z_{ij,t'}\right)\left(\sum_{s=1}^{T}\sum_{s'=1}^{T}z_{i'j',s}z_{i'j',s'}\right)\\&=T^{-4}\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{s=1}^{T}\sum_{s'=1}^{T}z_{ij,t}z_{ij,t'}z_{i'j',s}z_{i'j',s'},\end{aligned}$$

and

$$\begin{aligned}E\left\|\hat{\boldsymbol{\Sigma}}_{ab}-\boldsymbol{\Sigma}_{ab}\right\|_{F}^{4}&=T^{-4}\sum_{i=1}^{l_{a,T}}\sum_{j=1}^{l_{b,T}}\sum_{i'=1}^{l_{a,T}}\sum_{j'=1}^{l_{b,T}}\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{s=1}^{T}\sum_{s'=1}^{T}E\left(z_{ij,t}z_{ij,t'}z_{i'j',s}z_{i'j',s'}\right)\\&\leq\frac{l_{a,T}^{2}l_{b,T}^{2}}{T^{4}}\sup_{i,j,i',j'}\left[\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{s=1}^{T}\sum_{s'=1}^{T}E\left(z_{ij,t}z_{ij,t'}z_{i'j',s}z_{i'j',s'}\right)\right].\end{aligned}$$

Result (B.79) now follows from (B.78). ■

Remark 1 *It is clear that conditions (B.76) and (B.78) are met under Assumption 3 that requires z_{it} to be a martingale difference process. But it is easily seen that condition (B.76) also follows if we assume that $s_{a,it}$ and $s_{b,jt}$ are stationary processes with finite 8-th moments, since the product of stationary processes is also a stationary process under a certain additional cross-moment conditions (Wecker (1978)). The results of the lemma also follow readily if we assume that $s_{a,it}$ and $s_{b,jt'}$ are independently distributed for all $i \neq j$ and all t and t'.*

Lemma A19 *Consider the data generating process (6) with k signal variables, k^* pseudo-signal variables, and $n-k-k^*$ noise variables. Let $\hat{k}^o_{(s)}$ be the number of variables selected at the stage s of the OCMT procedure and suppose that conditions of Lemma A10 hold. Let $k^* = \ominus(n^\epsilon)$ for some $0 \leq \epsilon < \min\{1, \kappa_1/3\}$, where κ_1 is the positive constant that defines the rate for $T = \ominus(n^{\kappa_1})$ in Lemma A10. Let $\mathcal{D}_{s,T}$, be the event that the number of variables selected in the first s stages of OCMT is smaller than or equal to l_T, where $l_T = \ominus(n^\nu)$ and ν satisfies $\epsilon < \nu < \kappa_1/3$. Then there exist constants $C_0, C_1 > 0$ such that for any $0 < \varkappa < 1$, any $\delta_s > 0$, and any $j > 0$, it follows that*

$$\Pr\left(\hat{k}^o_{(s)} - k - k^* > j|\mathcal{D}_{s-1,T}\right) \leq \frac{n-k-k^*}{j}\left\{\exp\left[-\frac{\varkappa c_p^2(n,\delta_s)}{2}\right] + \exp(-C_0 T^{C_1})\right\}, \quad \text{(B.81)}$$

for $s = 1, 2, ..., k$.

Proof. By convention, the number of variables selected at the stage zero of OCMT is zero. Conditioning on $\mathcal{D}_{s-1,T}$ allows the application of Lemma A10. We drop the conditioning notation in the rest of the proof to simplify notations. Then, by Markov's inequality

$$\Pr\left(\hat{k}^o_{(s)} - k - k^* > j\right) \leq \frac{E\left(\hat{k}^o_{(s)} - k - k^*\right)}{j}. \quad \text{(B.82)}$$

But

$$\begin{aligned}
E\left(\hat{k}^o_{(s)}\right) &= \sum_{i=1}^{n} E\left[I_{(s)}\widehat{(\beta_i \neq 0)}\right] \\
&= \sum_{i=1}^{k+k^*} E\left[I_{(s)}\widehat{(\beta_i \neq 0)}\right] + \sum_{i=k+k^*+1}^{n} E\left[I_{(s)}\widehat{(\beta_i \neq 0)}\,|\theta_{i,(s)} = 0\right]. \\
&\leq k + k^* + \sum_{i=k+k^*+1}^{n} E\left[I_{(s)}\widehat{(\beta_i \neq 0)}\,|\theta_{i,(s)} = 0\right],
\end{aligned}$$

where we have used $I_{(s)}\widehat{(\beta_i \neq 0)} \leq 1$. Moreover,

$$E\left[I_{(s)}\widehat{(\beta_i \neq 0)}\,|\theta_{i,(s)} = 0\right] = \Pr\left(\left|t_{\hat{\phi}_{T,i,(s)}}\right| > c_p(n,\delta_s)\,|\theta_{i,(s)} = 0\right),$$

for $i = k+k^*+1, k+k^*+2, ..., n$, and using (B.51) of Lemma A10, we have (for some $0 < \varkappa < 1$ and $C_0, C_1 > 0$)

$$\sup_{i>k+k^*} \Pr\left(\left|t_{\hat{\phi}_{T,i,(s)}}\right| > c_p\left(n, \delta_s\right) |\theta_{i,(s)} = 0\right) \leq \exp\left[-\frac{\varkappa c_p^2\left(n, \delta_s\right)}{2}\right] + \exp(-C_0 T^{C_1}).$$

Hence,

$$E\left(\hat{k}^o_{(s)}\right) - k - k^* \leq (n - k - k^*)\left\{\exp\left[-\frac{\varkappa c_p^2\left(n, \delta_s\right)}{2}\right] + \exp(-C_0 T^{C_1})\right\},$$

and therefore (using this result in (B.82))

$$\Pr\left(\hat{k}^o_{(s)} - k - k^* > j\right) \leq \frac{n - k - k^*}{j}\left\{\exp\left[-\frac{\varkappa c_p^2\left(n, \delta_s\right)}{2}\right] + \exp(-C_0 T^{C_1})\right\},$$

as desired. ■

Lemma A20 *Consider the data generating process (6) with k signal, k^* pseudo-signal, and $n - k - k^*$ noise variables. Let $\mathcal{T}_k$ be the event that the OCMT procedure stops after k stages or less, and suppose that conditions of Lemma A10 hold. Let $k^* = \ominus\left(n^{\epsilon}\right)$ for some $0 \leq \epsilon <$ $\min\{1, \kappa_1/3\}$, where κ_1 is the positive constant that defines the rate for $T = \ominus\left(n^{\kappa_1}\right)$ in Lemma A10. Moreover, let $\delta > 0$ and $\delta^* > 0$ denote the critical value exponents for stage 1 and subsequent stages of the OCMT procedure, respectively. Then,*

$$\Pr\left(\mathcal{T}_k\right) = 1 + O\left(n^{1-\nu-\varkappa\delta}\right) + O\left(n^{1-\varkappa\delta^*}\right) + O\left[n \exp\left(-C_0 n^{C_1 \kappa_1}\right)\right], \tag{B.83}$$

for some $C_0, C_1 > 0$, any $\varkappa$ in $0 < \varkappa < 1$, and any ν in $\epsilon < \nu < \kappa_1/3$.

Proof. Consider the event $\mathcal{D}_{k,T} = \{\hat{k}_{(j)} \leq l_T, j = 1, 2, ..., k\}$ for $k \geq 1$, which is the event that the number of variables selected in the first k stages of OCMT is smaller than or equal to $l_T = \ominus\left(n^{\nu}\right)$, where ν lies in the interval $\epsilon < \nu < \kappa_1/3$. Such a ν exists since by assumption $0 \leq \epsilon < \min\{1, \kappa_1/3\}$. We have $\Pr\left(\mathcal{T}_k\right) = 1 - \Pr\left(\mathcal{T}_k^c\right)$, and

$$\begin{aligned}
\Pr\left(\mathcal{T}_k^c\right) &= \Pr\left(\mathcal{T}_k^c | \mathcal{D}_{k,T}\right) \Pr\left(\mathcal{D}_{k,T}\right) + \Pr\left(\mathcal{T}_k^c | \mathcal{D}_{k,T}^c\right) \Pr\left(\mathcal{D}_{k,T}^c\right) \\
&\leq \Pr\left(\mathcal{T}_k^c | \mathcal{D}_{k,T}\right) + \Pr\left(\mathcal{D}_{k,T}^c\right),
\end{aligned}$$

Therefore,

$$\Pr\left(\mathcal{T}_k\right) \geq 1 - \Pr\left(\mathcal{T}_k^c | \mathcal{D}_{k,T}\right) - \Pr\left(\mathcal{D}_{k,T}^c\right). \tag{B.84}$$

We note that

$$\Pr\left(\mathcal{D}_{k,T}\right) \geq \Pr\left[\left(\hat{k}^o_{(1)} \leq \frac{l_T}{k}\right) \cap \left(\hat{k}^o_{,(2)} \leq \frac{l_T}{k} \middle| \mathcal{D}_{1,T}\right) \cap ... \cap \left(\hat{k}^o_{(k)} \leq \frac{l_T}{k} \middle| \mathcal{D}_{k-1,T}\right)\right],$$

where $\hat{k}^o_{(s)}$ is the number of variables selected in the s-th stage of OCMT and $\mathcal{D}_{s,T} = \{\hat{k}_{(j)} \leq l_T, j = 1, 2, ..., s\}$ for $s = 1, 2, ..., k$. Hence

$$\Pr\left(\mathcal{D}^c_{k,T}\right) \leq \Pr\left\{\left[\begin{array}{c}\left(\hat{k}^o_{(1)} \leq \frac{l_T}{k}\right) \cap \left(\hat{k}^o_{(2)} \leq \frac{l_T}{k}\middle| \mathcal{D}_{1,T}\right) \cap ... \\ \cap\left(\hat{k}^o_{(k)} \leq \frac{l_T}{k}\middle| \mathcal{D}_{k-1,T}\right)\end{array}\right]^c\right\}.$$

Furthermore

$$\begin{aligned}
&\Pr\left\{\left[\begin{array}{c}\left(\hat{k}^o_{(1)} \leq \frac{l_T}{k}\right) \cap \left(\hat{k}^o_{(2)} \leq \frac{l_T}{k}\middle| \mathcal{D}_{1,T}\right) \cap ... \\ \cap\left(\hat{k}^o_{(k)} \leq \frac{l_T}{k}\middle| \mathcal{D}_{k-1,T}\right)\end{array}\right]^c\right\} \\
&= \Pr\left\{\left[\begin{array}{c}\left(\hat{k}^o_{(1)} > \frac{l_T}{k}\right) \cup \left(\hat{k}^o_{(2)} > \frac{l_T}{k}\middle| \mathcal{D}_{1,T}\right) \cup ... \\ \cup\left(\hat{k}^o_{(k)} > \frac{l_T}{k}\middle| \mathcal{D}_{k-1,T}\right)\end{array}\right]\right\} \\
&\leq \Pr\left(\hat{k}^o_{(1)} > \frac{l_T}{k}\right) + \sum_{s=2}^{k} \Pr\left(\hat{k}^o_{(s)} > \frac{l_T}{k}\middle| \mathcal{D}_{s-1,T}\right).
\end{aligned}$$

Since k is finite and $0 \leq \epsilon < \nu$, there exists T_0 such that for all $T > T_0$ we have $l_T/k > k + k^*$, and we can apply (B.81) of Lemma A19 (for $j = l_T/k - k - k^* > 0$), to obtain

$$\begin{aligned}
\Pr\left(\hat{k}^o_{(1)} > \frac{l_T}{k}\right) &= \Pr\left(\hat{k}^o_{(1)} - k - k^* > \frac{l_T}{k} - k - k^*\right) \\
&\leq \frac{n - k - k^*}{\frac{l_T}{k} - k - k^*}\left\{\exp\left[-\frac{\varkappa c_p^2\left(n,\delta\right)}{2}\right] + \exp(-C_0 T^{C_1})\right\},
\end{aligned}$$

for some $C_0, C_1 > 0$ and any $0 < \varkappa < 1$. Noting that for $0 \leq \epsilon < \nu$,

$$\frac{n - k - k^*}{\frac{l_T}{k} - k - k^*} = \ominus\left(n^{1-\nu}\right), \tag{B.85}$$

and using also result (ii) of Lemma A2, we obtain

$$\Pr\left(\hat{k}^o_{(1)} > \frac{l_T}{k}\right) = O\left(n^{1-\nu-\varkappa\delta}\right) + O\left[n^{1-\nu}\exp\left(-C_0 T^{C_1}\right)\right].$$

Similarly,

$$\begin{aligned}
\Pr\left(\hat{k}^o_{(s)} > \frac{l_T}{k}\middle| \mathcal{D}_{s-1,T}\right) &= \Pr\left(\hat{k}^o_{(s)} - k - k^* > \frac{l_T}{k} - k - k^*\middle| \mathcal{D}_{s-1,T}\right) \\
&\leq \frac{n - k - k^*}{\frac{l_T}{k} - k - k^*}\left\{\exp\left[-\frac{\varkappa c_p^2\left(n,\delta^*\right)}{2}\right] + \exp(-C_0 T^{C_1})\right\} \\
&= O\left(n^{1-\nu-\varkappa\delta^*}\right) + O\left[n^{1-\nu}\exp\left(-C_0 T^{C_1}\right)\right],
\end{aligned}$$

where the critical value exponent in the higher stages ($s > 1$) of OCMT (δ^*) could differ from the one in the first stage (δ). So, overall

$$\begin{aligned}
\Pr\left(\mathcal{D}^c_{k,T}\right) &\leq \Pr\left(\hat{k}^o_{(1)} > \frac{l_T}{k}\right) + \sum_{s=2}^{k} \Pr\left(\hat{k}^o_{(s)} > \frac{l_T}{k}\middle| \mathcal{D}_{s-1,T}\right) \\
&= O\left(n^{1-\nu-\varkappa\delta}\right) + O\left(n^{1-\nu-\varkappa\delta^*}\right) + O\left[n^{1-\nu}\exp\left(-C_0 T^{C_1}\right)\right],
\end{aligned} \tag{B.86}$$

for some $C_0, C_1 > 0$, any $\varkappa$ in $0 < \varkappa < 1$, and any ν in $\epsilon < \nu < \kappa_1/3$. Next, consider $\Pr(\mathcal{T}_k^c|\mathcal{D}_{k,T})$, and note that

$$\Pr(\mathcal{T}_k^c|\mathcal{D}_{k,T}) = \Pr(\mathcal{T}_k^c|\mathcal{D}_{k,T},\mathcal{L}_k)\Pr(\mathcal{L}_k|\mathcal{D}_{k,T}) + \Pr(\mathcal{T}_k^c|\mathcal{D}_{k,T},\mathcal{L}_k^c)\Pr(\mathcal{L}_k^c|\mathcal{D}_{k,T})$$
$$\leq \Pr(\mathcal{T}_k^c|\mathcal{D}_{k,T},\mathcal{L}_k) + \Pr(\mathcal{L}_k^c|\mathcal{D}_{k,T}), \tag{B.87}$$

where $\Pr(\mathcal{T}_k^c|\mathcal{D}_{k,T},\mathcal{L}_k)$ is the probability that a noise variable will be selected in a stage of OCMT that includes as regressors all signal variables, conditional on the event that fewer than l_T variables are selected in the first k steps of OCMT. Note that the event $\mathcal{T}_k^c|\mathcal{D}_{k,T},\mathcal{L}_k$ can only occur if OCMT selects some pseudo-signal and/or some noise variables in stage $k+1$. But the net effect coefficient of signal variables in stage $k+1$ must be zero when all signal variables were selected in earlier stages ($s = 1, 2, ..., k$), namely $\theta_{i,(k+1)} = 0$ for $i = k+1, k+2, ..., k+k^*$. Moreover, $\theta_{i,(k+1)} = 0$ also for $i = k+k^*+1, k+k^*+2, ..., n$, since the net effect coefficient of noise variables is always zero (in any stage). Therefore, we have

$$\Pr(\mathcal{T}_k^c|\mathcal{D}_{k,T},\mathcal{L}_k) \leq \sum_{i=k+1}^{n} \Pr\left[\left|t_{\hat{\phi}_{i,(k+1)}}\right| > c_p(n,\delta^*)\,|\theta_{i,(k+1)} = 0, \mathcal{D}_{k,T}\right].$$

Note that the number of regressors in the regressions involving the t statistics $t_{\hat{\theta}_{i,(k+1)}}$, does not exceed $l_T = \ominus(n^{\nu})$, for ν in the interval $0 \leq \epsilon < \nu < \kappa_1/3$ and hence $l_T = o(T^{1/3})$ as required by the conditions of Lemma A10. Using (B.51) of Lemma A10, we have

$$\Pr(\mathcal{T}_k^c|\mathcal{D}_{k,T},\mathcal{L}_k) \leq (n-k)\exp\left[\frac{-\varkappa c_p^2(n,\delta^*)}{2}\right]$$
$$+ (n-k)\exp\left(-C_0 T^{C_1}\right). \tag{B.88}$$

for some $C_0, C_1 > 0$ and any $0 < \varkappa < 1$. By Lemma A2, $\exp\left[-\varkappa c_p^2(n,\delta^*)/2\right] = \ominus\left(n^{-\varkappa\delta^*}\right)$, for any $0 < \varkappa < 1$, and noting that $n - k \leq n$ we obtain

$$\Pr(\mathcal{T}_k^c|\mathcal{D}_{k,T},\mathcal{L}_k) = O\left(n^{1-\varkappa\delta^*}\right) + O\left[n\exp\left(-C_0 T^{C_1}\right)\right]. \tag{B.89}$$

Consider next the second term of (B.87), $\Pr(\mathcal{L}_k^c|\mathcal{D}_{k,T})$, and recall that $\mathcal{L}_k = \cap_{i=1}^k \mathcal{L}_{i,k}$ where $\mathcal{L}_{i,k} = \cup_{j=1}^k \mathcal{B}_{i,j}$, $i = 1, 2, ..., k$. Hence $\mathcal{L}_{i,k}^c = \cap_{j=1}^k \mathcal{B}_{i,j}^c$, and

$$\Pr\left(\mathcal{L}_{i,k}^c \middle| \mathcal{T}_k, \mathcal{D}_{k,T}\right) = \Pr\left(\cap_{j=1}^k \mathcal{B}_{i,j}^c \middle| \mathcal{T}_k, \mathcal{D}_{k,T}\right) =$$
$$\Pr\left(\mathcal{B}_{i,1}^c \middle| \mathcal{T}_k, \mathcal{D}_{k,T}\right)\Pr\left(\mathcal{B}_{i,2}^c \middle| \mathcal{B}_{i,1}^c, \mathcal{T}_k, \mathcal{D}_{k,T}\right)$$
$$\Pr\left(\mathcal{B}_{i,3}^c \middle| \mathcal{B}_{i,2}^c \cap \mathcal{B}_{i,1}^c, \mathcal{T}_k, \mathcal{D}_{k,T}\right) \times ... \times$$
$$\Pr\left(\mathcal{B}_{i,k}^c \middle| \mathcal{B}_{i,k-1}^c \cap ... \cap \mathcal{B}_{i,1}^c, \mathcal{T}_k, \mathcal{D}_{k,T}\right).$$

But by Proposition 1 we are guaranteed that for some $1 \leq j \leq k$, $\theta_{i,(j)} \neq 0$. Therefore,

$$\Pr\left(\mathcal{B}_{i,j}^c \middle| \mathcal{B}_{i,j-1}^c \cap ... \cap \mathcal{B}_{i,1}^c, \mathcal{T}_k, \mathcal{D}_{k,T}\right) = \Pr\left(\mathcal{B}_{i,j}^c \middle| \mathcal{B}_{i,j-1}^c \cap ... \cap \mathcal{B}_{i,1}^c, \theta_{i,(j)} \neq 0, \mathcal{T}_k, \mathcal{D}_{k,T}\right),$$

and by (B.52) of Lemma A10,

$$\Pr\left(\mathcal{B}_{i,j}^{c}\middle|\,\mathcal{B}_{i,j-1}^{c}\cap...\cap\mathcal{B}_{i,1}^{c},\theta_{i,(j)}\neq 0,\mathcal{T}_{k},\mathcal{D}_{k,T}\right)=O\left[\exp\left(-C_{0}T^{C_{1}}\right)\right],$$

for some $C_0, C_1 > 0$. Therefore, for some $j \in \{1, 2, ..., k\}$ and $C_0, C_1 > 0$,

$$\begin{aligned}\Pr\left(\mathcal{L}_{i,k}^{c}\middle|\,\mathcal{T}_{k},\mathcal{D}_{k,T}\right) &\leq \Pr\left(\mathcal{B}_{i,j}^{c}\middle|\,\mathcal{B}_{i,j-1}^{c}\cap...\cap\mathcal{B}_{i,1}^{c},\theta_{i,(j)}\neq 0,\mathcal{T}_{k},\mathcal{D}_{k,T}\right)\\ &= O\left[\exp\left(-C_{0}T^{C_{1}}\right)\right]. \end{aligned} \tag{B.90}$$

Noting that k is finite and

$$\begin{aligned}\Pr\left(\mathcal{L}_{k}^{c}\middle|\,\mathcal{T}_{k},\mathcal{D}_{k,T}\right) &= \Pr\left(\cup_{i=1}^{k}\mathcal{L}_{ik}^{c}\middle|\,\mathcal{T}_{k},\mathcal{D}_{k,T}\right)\\ &\leq \sum_{i=1}^{k}\Pr\left(\mathcal{L}_{ik}^{c}\middle|\,\mathcal{T}_{k},\mathcal{D}_{k,T}\right),\end{aligned}$$

it follows, using (B.90), that

$$\Pr\left(\mathcal{L}_{k}^{c}\middle|\,\mathcal{T}_{k},\mathcal{D}_{k,T}\right)=O\left[\exp\left(-C_{0}T^{C_{1}}\right)\right], \tag{B.91}$$

for some $C_0, C_1 > 0$. Using (B.89) and (B.91) in (B.87) now gives[5]

$$\Pr\left(\mathcal{T}_{k}^{c}|\mathcal{D}_{k,T}\right)=O\left(n^{1-\varkappa\delta^{*}}\right)+O\left[n\exp\left(-C_{0}T^{C_{1}}\right)\right]. \tag{B.92}$$

Using (B.86) and (B.92) in (B.84), yields

$$\begin{aligned}\Pr\left(\mathcal{T}_{k}\right) =& 1+O\left(n^{1-\nu-\varkappa\delta}\right)+O\left(n^{1-\nu-\varkappa\delta^{*}}\right)+O\left[n^{1-\nu}\exp\left(-C_{0}T^{C_{1}}\right)\right]\\ &+O\left(n^{1-\varkappa\delta^{*}}\right)+O\left[n\exp\left(-C_{2}T^{C_{3}}\right)\right],\end{aligned}$$

for some $C_0, C_1, C_2, C_3 > 0$ and any $\varkappa$ in $0 < \varkappa < 1$, and any ν in $\epsilon < \nu < \kappa_1/3$. But $O\left(n^{1-\nu-\varkappa\delta^{*}}\right)$ is dominated by $O\left(n^{1-\varkappa\delta^{*}}\right)$, and $O\left[n^{1-\nu}\exp\left(-C_{0}T^{C_{1}}\right)\right]$ is dominated by $O\left[n\exp\left(-C_{2}T^{C_{3}}\right)\right]$, since $\nu > \epsilon \geq 0$. Hence,

$$\Pr\left(\mathcal{T}_{k}\right)=1+O\left(n^{1-\nu-\varkappa\delta}\right)+O\left(n^{1-\varkappa\delta^{*}}\right)+O\left[n\exp\left(-C_{0}T^{C_{1}}\right)\right],$$

for some $C_0, C_1 > 0$, any $\varkappa$ in $0 < \varkappa < 1$, and any ν in $\epsilon < \nu < \kappa_1/3$. This result in turn establishes (B.83), noting that $T = \ominus\left(n^{\kappa_1}\right)$. ■

Lemma A21 *Suppose that the data generating process (DGP) is given by*

$$\underset{T\times 1}{\boldsymbol{y}} = \underset{T\times k+1}{\mathbf{X}} \cdot \underset{k+1\times 1}{\boldsymbol{\beta}} + \underset{T\times 1}{\boldsymbol{u}}, \tag{B.93}$$

where $\boldsymbol{u} = (u_1, u_2, ..., u_T)'$, $E(\boldsymbol{u}) = \mathbf{0}$, $E(\boldsymbol{u}\boldsymbol{u}') = \sigma^2\mathbf{I}_T$, $0 < \sigma^2 < \infty$, $\mathbf{I}_T$ *is a* $T \times T$ *identity matrix,* $\mathbf{X} = (\boldsymbol{\tau}_T, \mathbf{X}_k) = (\boldsymbol{\tau}_T, \boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_k)$ *includes a* $T \times 1$ *column of ones,* $\boldsymbol{\tau}_T$, *and* $T \times 1$

[5] We have dropped the term $O\left[\exp\left(-C_{0}T^{C_{1}}\right)\right]$, which is dominated by $O\left[n\exp\left(-C_{0}T^{C_{1}}\right)\right]$.

vectors of observations, $\boldsymbol{x}_i = (x_{i1}, x_{i2}, ..., x_{iT})'$*, on the signal variables* $i = 1, 2, ..., k$*, and the elements of* $\boldsymbol{\beta}$ *are bounded. Consider the regression model*

$$\underset{T\times 1}{\boldsymbol{y}} = \underset{T\times l_T}{\mathbf{S}} \cdot \underset{l_T\times 1}{\boldsymbol{\delta}} + \underset{T\times 1}{\boldsymbol{\varepsilon}}, \tag{B.94}$$

where $\mathbf{S} = (s_{it}) = (\boldsymbol{s}_1, \boldsymbol{s}_2..., \boldsymbol{s}_{l_T})$*, with* $\boldsymbol{s}_j = (s_{j1}, s_{j2}, ..., s_{jT})'$*, for* $j = 1, 2, ..., l_T$*, Denote the least squares estimator of* $\boldsymbol{\delta}$ *in the regression model (B.94), by* $\hat{\boldsymbol{\delta}}$ *, and the associated* $T \times 1$ *vector of least squares residuals, by* $\tilde{\mathbf{u}} = \boldsymbol{y} - \mathbf{S}\hat{\boldsymbol{\delta}}$*, and set* $\boldsymbol{\beta}_0 = \left(\boldsymbol{\beta}', \mathbf{0}'_{l_T-k-1}\right)'$*. Denote the eigenvalues of* $\boldsymbol{\Sigma}_{ss} = E\left(T^{-1}\mathbf{S}'\mathbf{S}\right)$ *by* $\mu_1 \leq \mu_2 \leq ... \leq \mu_{l_T}$*, and assume that the following conditions hold:*

i. $\mu_i = O(l_T)$*,* $i = l_T - M + 1, l_T - M + 2, ..., l_T$*, for some finite* M*,* $\sup_{1\leq i\leq l_T - M} \mu_i < C_0 < \infty$*, for some* $C_0 > 0$*, and* $\inf_{1\leq i<l_T} \mu_i > C_1 > 0$*, for some* $C_1 > 0$*.*

ii. Regressors are uncorrelated with the errors, $E(s_{jt}u_t) = 0 = E(x_{it}u_t)$*, for all* $t = 1, 2..., T$*,* $i = 1, 2, ..., k$*, and* $j = 1, 2, ..., l_T$*,* s_{it} *have finite* 8^{th} *order moments, and* $z_{ij,t} = s_{it}s_{jt} - E(s_{it}s_{jt})$ *satisfies conditions (B.76) and (B.78) of Lemma A18. Moreover,* $z^*_{ij,t} = s_{it}x_{jt} - E(s_{it}x_{jt})$ *satisfies condition (B.76) of Lemma A18.*

Suppose that $l_T^3/T \to 0$*, as* l_T *and* $T \to \infty$*, Then, if* $\mathbf{S}$ *contains* $\mathbf{X}$

$$F_{\tilde{u}} = T^{-1}\left\|\tilde{\mathbf{u}}\right\|^2 = \sigma^2 + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{l_T^3}{T^{3/2}}\right) + O_p\left(\frac{l_T^{3/2}}{T}\right), \tag{B.95}$$

and

$$\left\|\hat{\boldsymbol{\delta}} - \boldsymbol{\beta}_0\right\|_F = O_p\left(\frac{l_T}{\sqrt{T}}\right) + O_p\left(\frac{l_T^{5/2}}{T}\right). \tag{B.96}$$

But if one or more columns of $\mathbf{X}$ *are not contained in* $\mathbf{S}$*, then*

$$F_{\tilde{u}} = \sigma^2 + O_p(1), \tag{B.97}$$

and

$$\left\|\hat{\boldsymbol{\delta}} - \boldsymbol{\beta}_0\right\|_F = O(l_T) + O_p\left(\frac{l_T^{5/2}}{T}\right) + O_p\left(\frac{l_T^{5/2}}{\sqrt{T}}\right) + O_p\left(\frac{l_T}{\sqrt{T}}\right). \tag{B.98}$$

Proof. Let $\hat{\boldsymbol{\Sigma}}_{ss} = \mathbf{S}'\mathbf{S}/T$, and recall that by assumption matrices $\boldsymbol{\Sigma}_{ss} = E\left(T^{-1}\mathbf{S}'\mathbf{S}\right)$ and $\hat{\boldsymbol{\Sigma}}_{ss}$ are positive definite. Let $\hat{\boldsymbol{\Delta}}_{ss} = \hat{\boldsymbol{\Sigma}}_{ss}^{-1} - \boldsymbol{\Sigma}_{ss}^{-1}$ and using (2.15) of Berk (1974), note that

$$\left\|\hat{\boldsymbol{\Delta}}_{ss}\right\|_F \leq \frac{\|\boldsymbol{\Sigma}_{ss}^{-1}\|_F^2 \left\|\hat{\boldsymbol{\Sigma}}_{ss} - \boldsymbol{\Sigma}_{ss}\right\|_F}{1 - \|\boldsymbol{\Sigma}_{ss}^{-1}\|_F \left\|\hat{\boldsymbol{\Sigma}}_{ss} - \boldsymbol{\Sigma}_{ss}\right\|_F}. \tag{B.99}$$

We focus on the individual terms on the right side of (B.99) to establish a bound, in probability, for $\left\|\hat{\boldsymbol{\Delta}}_{ss}\right\|_F$. The assumptions on eigenvalues of $\boldsymbol{\Sigma}_{ss}$ in this lemma are the same as in Lemma A14 with the only exception that $O(.)$ terms are used instead of $\ominus(.)$. Using the same arguments as in the proof of (B.64) and (B.65) of Lemma A14, it follows that

$$\left\|\boldsymbol{\Sigma}_{ss}\right\|_F = O(l_T), \tag{B.100}$$

and

$$\left\|\boldsymbol{\Sigma}_{ss}^{-1}\right\|_F = O\left(\sqrt{l_T}\right). \tag{B.101}$$

Moreover, note that (i,j)-th element of $\left(\hat{\boldsymbol{\Sigma}}_{ss} - \boldsymbol{\Sigma}_{ss}\right)$, $z_{ijt} = s_{it}s_{jt} - E(s_{it}s_{jt})$, satisfies the conditions of Lemma A18, which establishes

$$E\left(\left\|\hat{\boldsymbol{\Sigma}}_{ss} - \boldsymbol{\Sigma}_{ss}\right\|_F^2\right) = O\left(\frac{l_T^2}{T}\right), \tag{B.102}$$

and therefore, using $E\left\|\hat{\boldsymbol{\Sigma}}_{ss} - \boldsymbol{\Sigma}_{ss}\right\|_F \leq \left[E\left(\left\|\hat{\boldsymbol{\Sigma}}_{ss} - \boldsymbol{\Sigma}_{ss}\right\|_F^2\right)\right]^{1/2}$, and the fact that L_1-convergence implies convergence in probability, we have.

$$\left\|\hat{\boldsymbol{\Sigma}}_{ss} - \boldsymbol{\Sigma}_{ss}\right\|_F = O_p\left(\frac{l_T}{\sqrt{T}}\right). \tag{B.103}$$

Using (B.101) and (B.103), it now follows that

$$\left\|\boldsymbol{\Sigma}_{ss}^{-1}\right\|_F \left\|\hat{\boldsymbol{\Sigma}}_{ss} - \boldsymbol{\Sigma}_{ss}\right\|_F = O_p\left(\frac{l_T^{3/2}}{\sqrt{T}}\right),$$

and since by assumption $\frac{l_T^{3/2}}{\sqrt{T}} \to 0$, then

$$\frac{1}{\left(1 - \left\|\boldsymbol{\Sigma}_{ss}^{-1}\right\|_F \left\|\hat{\boldsymbol{\Sigma}}_{ss} - \boldsymbol{\Sigma}_{ss}\right\|_F\right)^2} = O_p(1). \tag{B.104}$$

Now using (B.103), (B.104), and (B.101) in (B.99), we have

$$\left\|\hat{\boldsymbol{\Delta}}_{ss}\right\|_F = O(l_T)\, O_p\left(\frac{l_T}{\sqrt{T}}\right) O_p(1) = O_p\left(\frac{l_T^2}{\sqrt{T}}\right), \tag{B.105}$$

and hence

$$\left\|\left(\frac{\mathbf{S}'\mathbf{S}}{T}\right)^{-1}\right\|_F = \left\|\hat{\boldsymbol{\Sigma}}_{ss}^{-1}\right\|_F \leq \left\|\hat{\boldsymbol{\Delta}}_{ss}\right\|_F + \left\|\boldsymbol{\Sigma}_{ss}^{-1}\right\|_F = O_p\left(\frac{l_T^2}{\sqrt{T}}\right) + O_p\left(\sqrt{l_T}\right). \tag{B.106}$$

Further, since by the assumption $E(\boldsymbol{s}_t u_t) = 0$, then $\left\|\frac{\mathbf{S}'\boldsymbol{u}}{T}\right\|_F^2 = O_p\left(\frac{l_T}{T}\right)$, and

$$\left\|\frac{\mathbf{S}'\boldsymbol{u}}{T}\right\|_F = O_p\left(\sqrt{\frac{l_T}{T}}\right). \tag{B.107}$$

Consider now the $T \times 1$ vector of residuals, $\tilde{\mathbf{u}}$ from the regression model (B.94) and note that under (B.93) it can be written as

$$\tilde{\mathbf{u}} = \mathbf{M}_s \boldsymbol{y} = \mathbf{M}_s \boldsymbol{u} + \mathbf{M}_s \mathbf{X} \beta, \text{ where } \mathbf{M}_s = \mathbf{I}_T - \mathbf{S} \left(\mathbf{S}'\mathbf{S} \right)^{-1} \mathbf{S}'. \tag{B.108}$$

In the case where $\mathbf{X}$ is a sub-set of $\mathbf{S}$, $\mathbf{M}_s \mathbf{X} \beta = \mathbf{0}$, and

$$F_{\tilde{u}} = T^{-1} \left\| \tilde{\mathbf{u}} \right\|^2 = T^{-1} \boldsymbol{u}' \mathbf{M}_s \boldsymbol{u} = T^{-1} \boldsymbol{u}' \boldsymbol{u} - \left(T^{-1} \boldsymbol{u}' \mathbf{S} \right) \left(T^{-1} \mathbf{S}'\mathbf{S} \right)^{-1} \left(T^{-1} \mathbf{S}' \boldsymbol{u} \right). \tag{B.109}$$

Also since u_t are serially uncorrelated with zero means and variance σ^2, we have

$$T^{-1} \boldsymbol{u}' \boldsymbol{u} = \sigma^2 + O_p \left(T^{-1/2} \right),$$

and

$$\left\| \left(T^{-1} \boldsymbol{u}' \mathbf{S} \right) \left(T^{-1} \mathbf{S}'\mathbf{S} \right)^{-1} \left(T^{-1} \mathbf{S}' \boldsymbol{u} \right) \right\|_F \leq \left\| \frac{\mathbf{S}' \boldsymbol{u}}{T} \right\|_F^2 \left\| \left(\frac{\mathbf{S}'\mathbf{S}}{T} \right)^{-1} \right\|_F,$$

which in view of (B.106) and (B.107) yields

$$\left(T^{-1} \boldsymbol{u}' \mathbf{S} \right) \left(T^{-1} \mathbf{S}'\mathbf{S} \right)^{-1} \left(T^{-1} \mathbf{S}' \boldsymbol{u} \right) = O_p \left(\frac{l_T^3}{T^{3/2}} \right) + O_p \left(\frac{l_T^{3/2}}{T} \right).$$

The result (B.95) now follows using the above results in (B.109). Now consider the case where $\mathbf{S}$ does not contain $\mathbf{X}$, and note from (B.108) that

$$F_{\tilde{u}} = T^{-1} \boldsymbol{u}' \mathbf{M}_s \boldsymbol{u} + T^{-1} \boldsymbol{\beta}' \mathbf{X}' \mathbf{M}_s \mathbf{X} \boldsymbol{\beta} + 2 T^{-1} \boldsymbol{\beta}' \mathbf{X}' \mathbf{M}_s \boldsymbol{u}. \tag{B.110}$$

Since $\mathbf{M}_s$ is an idempotent matrix then

$$\left\| T^{-1} \boldsymbol{\beta}' \mathbf{X}' \mathbf{M}_s \mathbf{X} \boldsymbol{\beta} \right\|_F \leq \boldsymbol{\beta}' \left(\frac{\mathbf{X}'\mathbf{X}}{T} \right) \boldsymbol{\beta} = \boldsymbol{\beta}' \boldsymbol{\Sigma}_{xx} \boldsymbol{\beta} + O_p \left(T^{-1/2} \right) = O_p(1).$$

Similarly,

$$\begin{aligned} T^{-1} \boldsymbol{\beta}' \mathbf{X}' \mathbf{M}_s \boldsymbol{u} &= T^{-1} \boldsymbol{\beta}' \mathbf{X}' \boldsymbol{u} - \left(T^{-1} \boldsymbol{\beta}' \mathbf{X}' \mathbf{S} \right) \left(T^{-1} \mathbf{S}'\mathbf{S} \right)^{-1} \left(T^{-1} \mathbf{S}' \boldsymbol{u} \right) \\ &= O_p \left(T^{-1/2} \right) + O_p \left(\frac{l_T}{\sqrt{T}} \right) + O_p \left(\frac{l_T^{5/2}}{T} \right). \end{aligned}$$

The result (B.97) now follows if we use the above results in (B.110) and recalling that the probability order of $T^{-1} \boldsymbol{u}' \mathbf{M}_s \boldsymbol{u}$ is given by (B.95). Consider now the least squares estimator of $\hat{\boldsymbol{\delta}}$ and note that under (B.93) it can be written as

$$\hat{\boldsymbol{\delta}} = \left(\mathbf{S}'\mathbf{S} \right)^{-1} \mathbf{S}' \boldsymbol{y} = \left(\mathbf{S}'\mathbf{S} \right)^{-1} \mathbf{S}' \mathbf{X} \boldsymbol{\beta} + \left(\mathbf{S}'\mathbf{S} \right)^{-1} \mathbf{S}' \boldsymbol{u}. \tag{B.111}$$

Suppose that $\mathbf{X}$ is included as the first $k+1$ columns of $\mathbf{S}$, and denote the remaining $l_T - k - 1$ columns of $\mathbf{S}$ by $\mathbf{W}$. Also partition $\hat{\boldsymbol{\delta}}$ accordingly as $\left(\hat{\boldsymbol{\delta}}_x', \hat{\boldsymbol{\delta}}_w' \right)'$, where $\hat{\boldsymbol{\delta}}_x$ is the $(k+1) \times 1$ vector of estimated coefficients associated with $\mathbf{X}$. Note also that in this case $\mathbf{S}\left(\mathbf{S}'\mathbf{S}\right)^{-1}\mathbf{S}'\mathbf{X}=\mathbf{X}$, and we have

$$\mathbf{S}\hat{\boldsymbol{\delta}}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{S}\left(\mathbf{S}'\mathbf{S}\right)^{-1}\mathbf{S}'\boldsymbol{u},$$

or

$$\mathbf{X}\left(\hat{\boldsymbol{\delta}}_x-\boldsymbol{\beta}\right)+\mathbf{W}\left(\hat{\boldsymbol{\delta}}_w-\mathbf{0}_{l_T-k-1}\right)=\boldsymbol{S}\left(\mathbf{S}'\mathbf{S}\right)^{-1}\mathbf{S}'\boldsymbol{u},$$

which can be written more compactly as $\mathbf{S}\left(\hat{\boldsymbol{\delta}}-\boldsymbol{\beta}_0\right)=\boldsymbol{S}\left(\mathbf{S}'\mathbf{S}\right)^{-1}\mathbf{S}'\boldsymbol{u}$, where $\boldsymbol{\beta}_0=\left(\boldsymbol{\beta}',\mathbf{0}'_{l_T-k-1}\right)'$. Premultiplying both sides by $\mathbf{S}'$, and noting that $\mathbf{S}'\mathbf{S}$ is invertible yields

$$\hat{\boldsymbol{\delta}}-\boldsymbol{\beta}_0=\left(\mathbf{S}'\mathbf{S}\right)^{-1}\mathbf{S}'\boldsymbol{u},$$

with the norm of $\hat{\boldsymbol{\delta}}-\boldsymbol{\beta}_0$ given by

$$\left\|\hat{\boldsymbol{\delta}}-\boldsymbol{\beta}_0\right\|_F=\left\|\left(\frac{\mathbf{S}'\mathbf{S}}{T}\right)^{-1}\left(\frac{\mathbf{S}'\boldsymbol{u}}{T}\right)\right\|_F\leq\left\|\left(\frac{\mathbf{S}'\mathbf{S}}{T}\right)^{-1}\right\|_F\left\|\left(\frac{\mathbf{S}'\boldsymbol{u}}{T}\right)\right\|_F.$$

Now using (B.106) and (B.107) it readily follows that

$$\left\|\hat{\boldsymbol{\delta}}-\boldsymbol{\beta}_0\right\|_F=O_p\left(\frac{l_T}{\sqrt{T}}\right)+O_p\left(\frac{l_T^{5/2}}{T}\right), \tag{B.112}$$

as required. Finally, in the case where one or more columns of $\mathbf{X}$ are not included in $\mathbf{S}$, consider the decomposition

$$\hat{\boldsymbol{\delta}}-\boldsymbol{\beta}_0=\left(\hat{\boldsymbol{\delta}}-\boldsymbol{\delta}_*\right)+\left(\boldsymbol{\delta}_*-\boldsymbol{\beta}_0\right), \tag{B.113}$$

where $\boldsymbol{\delta}_*=\boldsymbol{\Sigma}_{ss}^{-1}\boldsymbol{\Sigma}_{sx}\boldsymbol{\beta}$, and $\boldsymbol{\Sigma}_{sx}=E\left(T^{-1}\mathbf{S}'\mathbf{X}\right)$. When at least one of the columns of $\mathbf{X}$ does not belong to $\mathbf{S}$, then $\boldsymbol{\delta}_*\neq\boldsymbol{\beta}_0$. To investigate the probability order of the first term of the above, using (B.111), we note that

$$\hat{\boldsymbol{\delta}}-\boldsymbol{\delta}_*=\left(\hat{\boldsymbol{\Sigma}}_{ss}^{-1}\hat{\boldsymbol{\Sigma}}_{sx}-\boldsymbol{\Sigma}_{ss}^{-1}\boldsymbol{\Sigma}_{sx}\right)\boldsymbol{\beta}+\left(\mathbf{S}'\mathbf{S}\right)^{-1}\mathbf{S}'\boldsymbol{u},$$

where $\hat{\boldsymbol{\Sigma}}_{sx}=T^{-1}\mathbf{S}'\mathbf{X}$. But $\hat{\boldsymbol{\Sigma}}_{ss}^{-1}\hat{\boldsymbol{\Sigma}}_{sx}-\boldsymbol{\Sigma}_{ss}^{-1}\boldsymbol{\Sigma}_{sx}=\hat{\boldsymbol{\Delta}}_{ss}\hat{\boldsymbol{\Delta}}_{sx}+\hat{\boldsymbol{\Delta}}_{ss}\boldsymbol{\Sigma}_{sx}+\boldsymbol{\Sigma}_{ss}^{-1}\hat{\boldsymbol{\Delta}}_{sx}$, where $\hat{\boldsymbol{\Delta}}_{sx}=\hat{\boldsymbol{\Sigma}}_{sx}-\boldsymbol{\Sigma}_{sx}$, and, as before, $\hat{\boldsymbol{\Delta}}_{ss}=\hat{\boldsymbol{\Sigma}}_{ss}^{-1}-\boldsymbol{\Sigma}_{ss}^{-1}$. Hence

$$\begin{aligned}\left\|\left(\hat{\boldsymbol{\Sigma}}_{ss}^{-1}\hat{\boldsymbol{\Sigma}}_{sx}-\boldsymbol{\Sigma}_{ss}^{-1}\boldsymbol{\Sigma}_{sx}\right)\boldsymbol{\beta}\right\|_F\leq{}&\left\|\hat{\boldsymbol{\Delta}}_{ss}\right\|_F\left\|\hat{\boldsymbol{\Delta}}_{sx}\right\|_F\|\boldsymbol{\beta}\|+\left\|\hat{\boldsymbol{\Delta}}_{ss}\right\|_F\|\boldsymbol{\Sigma}_{sx}\|_F\|\boldsymbol{\beta}\|\\&+\left\|\boldsymbol{\Sigma}_{ss}^{-1}\right\|_F\left\|\hat{\boldsymbol{\Delta}}_{sx}\right\|_F\|\boldsymbol{\beta}\|\end{aligned}$$

Using Lemma A18 by setting $\mathbf{S}_a=\mathbf{S}$ $(l_{a,T}=l_T)$ and $\mathbf{S}_b=\mathbf{X}$ $(l_{b,T}=k+1)$, we also have, by (B.77),

$$\left\|\hat{\boldsymbol{\Delta}}_{sx}\right\|_F=\left\|\hat{\boldsymbol{\Sigma}}_{sx}-\boldsymbol{\Sigma}_{sx}\right\|_F=O_p\left(\sqrt{\frac{l_T}{T}}\right). \tag{B.114}$$

Also $\left\|\hat{\mathbf{\Delta}}_{ss}\right\|_F = O_p\left(l_T^2/\sqrt{T}\right)$ by (B.105), $\|\mathbf{\Sigma}_{ss}^{-1}\|_F = O\left(\sqrt{l_T}\right)$, by (B.101), $\|\mathbf{\Sigma}_{sx}\|_F = O\left(\sqrt{l_T}\right)$, $\|\boldsymbol{\beta}\| = O(1)$. Therefore

$$\left\|\left(\hat{\mathbf{\Sigma}}_{ss}^{-1}\hat{\mathbf{\Sigma}}_{sx} - \mathbf{\Sigma}_{ss}^{-1}\mathbf{\Sigma}_{sx}\right)\boldsymbol{\beta}\right\|_F = O_p\left(l_T^2/\sqrt{T}\right)O_p\left(\sqrt{\frac{l_T}{T}}\right) + O_p\left(l_T^2/\sqrt{T}\right)O\left(\sqrt{l_T}\right) + O\left(\sqrt{l_T}\right)O_p\left(\sqrt{\frac{l_T}{T}}\right)$$

$$= O_p\left(\frac{l_T^{5/2}}{T}\right) + O_p\left(\frac{l_T^{5/2}}{\sqrt{T}}\right) + O_p\left(\frac{l_T}{\sqrt{T}}\right).$$

Therefore, also using (B.112), overall we have

$$\left\|\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}_*\right\|_F = O_p\left(\frac{l_T^{5/2}}{T}\right) + O_p\left(\frac{l_T^{5/2}}{\sqrt{T}}\right) + O_p\left(\frac{l_T}{\sqrt{T}}\right).$$

Finally, using (B.113)

$$\left\|\hat{\boldsymbol{\delta}} - \boldsymbol{\beta}_0\right\|_F \leq \left\|\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}_*\right\|_F + \|\boldsymbol{\delta}_*\|_F + \|\boldsymbol{\beta}_0\|_F,$$

where $\|\boldsymbol{\beta}_0\| = O(1)$, since $\boldsymbol{\beta}_0$ contains finite $(k+1)$ number of bounded nonzero elements, and

$$\begin{aligned}\|\boldsymbol{\delta}_*\|_F &= \left\|\mathbf{\Sigma}_{ss}^{-1}\mathbf{\Sigma}_{sx}\right\|_F \\ &\leq \left\|\mathbf{\Sigma}_{ss}^{-1}\right\|_F \|\mathbf{\Sigma}_{sx}\|_F.\end{aligned}$$

$\|\mathbf{\Sigma}_{ss}^{-1}\|_F = O\left(\sqrt{l_T}\right)$ by (B.101), and $\|\mathbf{\Sigma}_{sx}\|_F = O\left(\sqrt{l_T}\right)$. Hence, in the case where at least one of the columns of $\mathbf{X}$ does not belong to $\mathbf{S}$, we have

$$\left\|\hat{\boldsymbol{\delta}} - \boldsymbol{\beta}_0\right\|_F = O(l_T) + O_p\left(\frac{l_T^{5/2}}{T}\right) + O_p\left(\frac{l_T^{5/2}}{\sqrt{T}}\right) + O_p\left(\frac{l_T}{\sqrt{T}}\right).$$

which completes the proof of (B.98). ■

B. Proof of Theorem 3

We proceed as in the proof of (B.52) in Lemma A10. We have that

$$\Pr\left[\left|\frac{T^{-1/2}\boldsymbol{x}_i'\mathbf{M}_q\boldsymbol{y}}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)\left(\frac{\boldsymbol{x}_i'\mathbf{M}_q\boldsymbol{x}_i}{T}\right)}}\right| > c_p(n,\delta)\right] \leq \Pr\left(\left|\frac{T^{1/2}\left(\frac{\boldsymbol{x}_i'\mathbf{M}_q\boldsymbol{\eta}}{T} - \theta\right)}{\sigma_{e,(T)}\sigma_{x_i,(T)}} + \frac{T^{1/2}\theta_i}{\sigma_{e,(T)}\sigma_{x_i,(T)}}\right| > \frac{c_p(n,\delta)}{1+d_T}\right).$$

We distinguish two cases: $\frac{T^{1/2}|\theta_i|}{\sigma_{e,(T)}\sigma_{x_i,(T)}} > \frac{c_p(n,\delta)}{1+d_T}$ and $\frac{T^{1/2}|\theta_i|}{\sigma_{e,(T)}\sigma_{x_i,(T)}} \leq \frac{c_p(n,\delta)}{1+d_T}$. If $\frac{T^{1/2}|\theta_i|}{\sigma_{e,(T)}\sigma_{x_i,(T)}} > \frac{c_p(n,\delta)}{1+d_T}$,

$$\Pr\left(\left|\frac{T^{1/2}\left(\frac{\boldsymbol{x}_i'\mathbf{M}_q\boldsymbol{\eta}}{T} - \theta\right)}{\sigma_{e,(T)}\sigma_{x_i,(T)}} + \frac{T^{1/2}\theta_i}{\sigma_{e,(T)}\sigma_{x_i,(T)}}\right| > \frac{c_p(n,\delta)}{1+d_T}\right) =$$

$$1 - \Pr\left(\left|\frac{T^{1/2}\left(\frac{\boldsymbol{x}_i'\mathbf{M}_q\boldsymbol{\eta}}{T} - \theta\right)}{\sigma_{e,(T)}\sigma_{x_i,(T)}} + \frac{T^{1/2}\theta_i}{\sigma_{e,(T)}\sigma_{x_i,(T)}}\right| \leq \frac{c_p(n,\delta)}{1+d_T}\right),$$

and, by Lemma A12

$$\Pr\left(\left|\frac{T^{1/2}\left(\frac{\boldsymbol{x}_i'\mathbf{M}_q\boldsymbol{\eta}}{T}-\theta\right)}{\sigma_{e,(T)}\sigma_{x_i,(T)}}+\frac{T^{1/2}\theta_i}{\sigma_{e,(T)}\sigma_{x_i,(T)}}\right|\leq\frac{c_p\left(n,\delta\right)}{1+d_T}\right)$$
$$\leq\Pr\left(\left|\frac{T^{1/2}\left(\frac{\boldsymbol{x}_i'\mathbf{M}_q\boldsymbol{\eta}}{T}-\theta\right)}{\sigma_{e,(T)}\sigma_{x_i,(T)}}\right|>\frac{T^{1/2}\left|\theta_i\right|}{\sigma_{e,(T)}\sigma_{x_i,(T)}}-\frac{c_p\left(n,\delta\right)}{1+d_T}\right)$$

while, if $\frac{T^{1/2}|\theta_i|}{\sigma_{e,(T)}\sigma_{x_i,(T)}}\leq\frac{c_p(n,\delta)}{1+d_T}$, by (B.150) of Lemma F4,

$$\Pr\left(\left|\frac{T^{1/2}\left(\frac{\boldsymbol{x}_i'\mathbf{M}_q\boldsymbol{\eta}}{T}-\theta\right)}{\sigma_{e,(T)}\sigma_{x_i,(T)}}+\frac{T^{1/2}\theta_i}{\sigma_{e,(T)}\sigma_{x_i,(T)}}\right|>\frac{c_p\left(n,\delta\right)}{1+d_T}\right)$$
$$\leq\Pr\left(\left|\frac{T^{1/2}\left(\frac{\boldsymbol{x}_i'\mathbf{M}_q\boldsymbol{\eta}}{T}-\theta\right)}{\sigma_{e,(T)}\sigma_{x_i,(T)}}\right|>\frac{c_p\left(n,\delta\right)}{1+d_T}-\frac{T^{1/2}\left|\theta_i\right|}{\sigma_{e,(T)}\sigma_{x_i,(T)}}\right)$$

We further note that since $c_p\left(n,\delta\right)\rightarrow\infty$, $\frac{T^{1/2}|\theta_i|}{\sigma_{e,(T)}\sigma_{x_i,(T)}}>\frac{c_p(n,\delta)}{1+d_T}$ implies $T^{1/2}\left|\theta_i\right|>C_2$, for some $C_2>0$. Then, noting that $\frac{\boldsymbol{x}_i'\mathbf{M}_q\boldsymbol{\eta}}{T}-\theta$ is the average of a martingale difference process, by Lemma A6, for some positive constants, C_1,C_2,C_3,C_4,C_5, and, for any $\psi>0$, we have

$$\sum_{i=k+1}^{n}\Pr\left[\left|\frac{T^{-1/2}\boldsymbol{x}_i'\mathbf{M}_q\boldsymbol{y}}{\sqrt{\left(\boldsymbol{e}'\boldsymbol{e}/T\right)\left(\frac{\boldsymbol{x}_i'\mathbf{M}_q\boldsymbol{x}_i}{T}\right)}}\right|>c_p\left(n,\delta\right)\right]\leq C_1\sum_{i=k+1}^{n}I\left(\sqrt{T}\theta_i>C_2\right)$$
$$+C_3\sum_{i=k+1}^{n}I\left(\sqrt{T}\theta_i\leq C_4\right)\exp\left[-\ln(n)^{C_5}\right],$$
$$=C_1\sum_{i=k+1}^{n}I\left(\sqrt{T}\theta_i>C_2\right)+o(n^{1-\psi})+O\left[\exp(-CT^{C_5})\right],\tag{B.115}$$

since $\exp\left[-\ln(n)^{C_5}\right]=o(n^{\psi})$, which follows by noting that $C_0\ln(n)^{1/2}=o\left(C_1\ln(n)\right)$, for any $C_0,C_1>0$. As a result, the crucial term for the behaviour of $FPR_{n,T}$ is the first term on the RHS of (B.115). Consider now the above probability bound under the two specifications assumed for θ_i as given by (4) and (5). Under (4), for any $\psi>0$,

$$\sum_{i=k+1}^{n}\Pr\left[\left|\frac{T^{-1/2}\boldsymbol{x}_i'\mathbf{M}_q\boldsymbol{y}}{\sqrt{\left(\boldsymbol{e}'\boldsymbol{e}/T\right)\left(\frac{\boldsymbol{x}_i'\mathbf{M}_q\boldsymbol{x}_i}{T}\right)}}\right|>c_p\left(n,\delta\right)\right]\leq C_0\sum_{i=k+1}^{n}I\left(\sqrt{T}\varrho^i>C_i\right)+o(n^{1-\psi}).$$

for some $C_0,C_i>0$, $i=k+1,...,n$. So we need to determine the limiting property of $\sum_{i=k+1}^{n}I\left(\sqrt{T}\varrho^i>C_i\right)$. Then, without loss of generality, consider $i=[n^{\zeta}]$, $T=n^{\kappa_1}$, $\zeta\in[0,1]$,

$\kappa_1 > 0$. Then, $\sqrt{T}\varrho^i = \sqrt{T}\varrho^{T^{(1/\kappa_1)\zeta}} = o(1)$ for all $\kappa_1, \zeta > 0$. Therefore,

$$C_a \sum_{i=k+1}^{n} I\left(\sqrt{T}\varrho^i > C_b/C_i\right) = o(n^{\zeta}),$$

for all $\zeta > 0$. This implies that under (4), $\theta_i = C_i\varrho^i$, $|\varrho| < 1$, and $c_p(n,\delta) = O\left[\ln(n)^{1/2}\right]$, we have

$$E\left|FPR_{n,T}\right| = o(n^{\zeta-1}) + O\left[\exp(-n^{C_0})\right],$$

for all $\zeta > 0$. Similarly, under (5), $\theta_i = C_i i^{-\gamma}$, and setting $i = [n^{\zeta}]$, $T = n^{\kappa_1}$, $\zeta, \kappa_1 > 0$, we have $\sqrt{T}\theta_i = T^{-(1/\kappa_1)\zeta\gamma+1/2}$. We need $-(1/\kappa_1)\zeta\gamma + 1/2 < 0$ or $\zeta > \frac{1}{2\kappa_1^{-1}\gamma}$. Then,

$$\frac{C_a}{n} \sum_{i=k+1}^{n} I\left(\sqrt{T}\theta_i > C_b/C_i\right) = O\left(T^{\frac{1}{2\kappa_1^{-1}\gamma}-\kappa_1^{-1}}\right) = O\left(n^{\frac{1}{2\kappa_1^{-2}\gamma}-1}\right)$$

So

$$E\left|FPR_{n,T}\right| = o(1), \tag{B.116}$$

as long as $2\kappa_1^{-2}\gamma > 1$ or if $\gamma > \frac{1}{2\kappa_1^{-2}}$.

Remark B1 *Note that if $\kappa_1 = 1$, then the condition for (B.116) requires that $\gamma > \frac{1}{2}$.*

C. Some results for the case where either noise variables are mixing, or both signal/pseudo-signal and noise variables are mixing

When only noise variables are mixing, all the results of the main paper go through since we can use the results obtained under (D1)-(D3) of Lemma D2 to replace Lemma A6.

As discussed in Section 4.2, some weak results can be obtained if both signal/pseudo-signal and noise variables are mixing processes, but only if $c_p(n)$ is allowed to grow faster than under the assumption of a martingale difference. This case is covered under (D4) of Lemma D2 and (B.140)-(B.141) of Lemma D3. There, it is shown that, for sufficiently large constants $C_0 - C_3$ for Assumption 4, the martingale difference bound which is given by $\exp\left[-\frac{1}{2}\varkappa c_p^2(n)\right]$ in Lemma A6 is replaced by the bound $\exp\left[-C_4 c_p(n)^{s/(s+2)}\right]$, for some $C_4 > 0$, where s is the exponent in the probability tail in Assumption 4. It is important to note here that this bound seems to be relatively sharp (see, e.g., Roussas (1996)), under our assumptions, and so we need to understand its implications for our analysis. We abstract from the constant C_4 which can further deteriorate rates. Given (see result (i) of Lemma A2),

$$c_p(n) = O\left\{\left[\ln\left(\frac{f(n)}{2p}\right)\right]^{1/2}\right\},$$

it follows that

$$\exp\left[-c_p(n)^{s/(s+2)}\right] = O\left[\exp\left\{-\left[\ln\left(\frac{f(n)}{2p}\right)\right]^{s/2(s+2)}\right\}\right]$$

Let $f(n) = 2p\exp(n^{a_n})$. Then,

$$\exp\left\{-\left[\ln\left(\frac{f(n)}{2p}\right)\right]^{s/2(s+2)}\right\} = \exp\left[-n^{a_n s/2(s+2)}\right]$$

To obtain the same bound as for the martingale difference case, we need to find a sequence $\{a_n\}$, such that $n^{Ca_n} = O(\ln(n))$. Setting $n^{Ca_n} = \ln(n)$, it follows that $a_n = \ln(\ln(n))/C\ln n$. Further, setting $C = s/2(s+2)$, we have $a_n = \frac{2(s+2)\ln(\ln(n))}{s\ln n}$, which leads to the following choice for $f(n)$

$$f(n) = 2p\exp\left(n^{\frac{2(s+2)\ln(\ln(n))}{s\ln n}}\right) \sim 2p\exp\left(\ln(n)^{\frac{2(s+2)}{s}}\right).$$

Then,

$$c_p(n) = O\left[\ln\left(\exp\left(\ln(n)^{\frac{2(s+2)}{s}}\right)\right)\right] = O\left(\ln(n)^{\frac{2(s+2)}{s}}\right),$$

which for $n = O\left(T^{C_1}\right)$, $C_1 > 0$, implies that $c_p(n) = O\left(\ln(T)^{\frac{2(s+2)}{s}}\right)$, and so, $c_p(n) = o\left(T^{C_2}\right)$, for all $C_2 > 0$, as long as $s > 0$.

We need to understand the implications of this result. For example, setting $s = 2$ which corresponds to the normal case gives $\exp(\ln(n)^4)$ which makes the calculation of $\Phi^{-1}\left(1 - \frac{p}{2f(n)}\right)$ numerically problematic for $n > 25$. The fast rate at which $f(n)$ grows basically implies that we need $s \to \infty$ which corresponds to $f(n) = 2p\exp(\ln(n)^2)$. Even then, the analysis becomes problematic for large n. $s \to \infty$ corresponds for all practical purposes to assuming boundedness for x_{it}. As a result, while the case of mixing x_{it} can be analysed theoretically, its practical implications are limited. On the other hand our Monte Carlo study in Section 5 suggests that setting $f(n) = f(n, \delta) = n^{\delta}$, $\delta \geq 1$ provides quite good results for autoregressive x_{it} in small samples.

D. Lemmas for mixing results

We consider the following assumptions that replace Assumption 3.

Assumption D1 *x_{it}, $i = 1, 2, ..., k + k^*$, are martingale difference processes with respect to $\mathcal{F}_{t-1}^{xs} \cup \mathcal{F}_t^{xn}$, where $\mathcal{F}_{t-1}^{xs}$ and $\mathcal{F}_t^{xn}$ are defined in Assumption 3. x_{it}, $i = 1, 2, ..., k + k^*$ are independent of x_{it}, $i = k + k^* + 1, k + k^* + 2, ..., n$. $E\left(x_{it}x_{jt} - E(x_{it}x_{jt}) \left| \mathcal{F}_{t-1}^{xs}\right.\right) = 0$, $i, j = 1, 2, ..., k + k^*$. x_{it}, $i = k + k^* + 1, k + k^* + 2, ..., n$, are heterogeneous strongly mixing processes with mixing coefficients given by $\alpha_{i\ell} = C_{i\ell}\xi^{\ell}$ for some $C_{i\ell}$ such that $\sup_{i,\ell} C_{i\ell} < \infty$ and some $0 < \xi < 1$. $E[x_{it}u_t | \mathcal{F}_{t-1}] = 0$, for $i = 1, 2, ..., n$, and all t.*

Assumption D2 *x_{it}, $i = 1, 2, ..., k + k^*$ are independent of x_{it}, $i = k + k^* + 1, k + k^* + 2, ..., n$. x_{it}, $i = 1, 2, ..., n$, are heterogeneous strongly mixing processes with mixing coefficients given by $\alpha_{i\ell} = C_{i\ell}\xi^{\ell}$ for some $C_{i\ell}$ such that $\sup_{i,\ell} C_{i\ell} < \infty$ and some $0 < \xi < 1$. $E[x_{it}u_t | \mathcal{F}_{t-1}] = 0$, for $i = 1, 2, ..., n$, and all t.*

Lemma D1 *Let* ξ_t *be a sequence of zero mean, mixing random variables with exponential mixing coefficients given by* $\phi_k = a_{0k}\varphi^k$, $0 < \varphi < 1$, $a_{0k} < \infty$, $k = 1, \ldots$. *Assume, further, that* $\Pr(|\xi_t| > \alpha) \leq C_0 \exp[-C_1\alpha^s]$, $s \geq 1$. *Then, for some* $C_2, C_3 > 0$, *each* $0 < \delta < 1$ *and* $v_T \geq \epsilon T^\lambda$, $\lambda > (1+\delta)/2$,

$$\Pr\left(\left|\sum_{t=1}^{T}\xi_t\right| > v_T\right) \leq C_2 \exp\left[-\left(C_3 v_T T^{-(1+\delta)/2}\right)^{s/(s+1)}\right]$$

Proof. We reconsider the proof of Theorem 3.5 of White and Wooldridge (1991) relaxing the assumption of stationarity. Define $w_t = \xi_t I(z_t \leq D_T)$ and $v_t = \xi_t - w_t$ where D_T will be defined below. Using Theorem 3.4 of White and Wooldridge (1991), which does not assume stationarity, we have that constants C_0 and C_1 in the statement of the present Lemma can be chosen sufficiently large such that

$$\Pr\left(\left|\sum_{t=1}^{T} w_t - E(w_t)\right| > v_T\right) \leq C_4 \exp\left[\frac{-C_5 v_T T^{-(1+\delta)/2}}{D_T}\right] \tag{B.117}$$

for some $C_4, C_5 > 0$, rather than

$$\Pr\left(\left|\sum_{t=1}^{T} w_t - E(w_t)\right| > v_T\right) \leq C_6 \exp\left[\frac{-C_7 v_T T^{-1/2}}{D_T}\right]$$

for some $C_6, C_7 > 0$, which uses Theorem 3.3 of White and Wooldridge (1991). We explore the effects this change has on the final rate. We revisit the analysis of the bottom half of page 489 of White and Wooldridge (1991). We need to determine D_T such that

$$v_T^{-1} T\left[\exp\left(-C_1\left(\frac{D_T}{2}\right)^s\right)\right]^{1/q} \leq \exp\left[\frac{-C v_T T^{-(1+\delta)/2}}{D_T}\right]$$

for some $C > 0$. Take logs and we have

$$\ln\left(v_T^{-1}T\right) - \left(\frac{1}{q}\right) C_1 \left(\frac{D_T}{2}\right)^s \leq \frac{-C v_T T^{-(1+\delta)/2}}{D_T}$$

or

$$D_T^s \geq 2^p \left(\frac{q}{C_1}\right) \ln\left(v_T^{-1}T\right) + \frac{2^s q C v_T}{C_1 T^{(1+\delta)/2} D_T}$$

For this it suffices that

$$\frac{2^s q C v_T}{T^{(1+\delta)/2} D_T} \geq 2^p q \ln\left(v_T^{-1}T\right) \tag{B.118}$$

and

$$D_T^s \geq \frac{2^s q C v_T}{C_1 T^{(1+\delta)/2} D_T}. \tag{B.119}$$

Set

$$D_T = \left(\frac{2^s q C v_T}{C_1 T^{(1+\delta)/2}}\right)^{1/(s+1)},$$

so that (B.119) holds with equality. But since $v_T \geq \epsilon T^{\lambda}$, $\lambda > (1+\delta)/2$, (B.118) holds. Therefore,

$$\frac{2^s qCv_T}{C_1 T^{(1+\delta)/2} D_T} = \left(\frac{2^s qCv_T}{C_1 T^{(1+\delta)/2}}\right)^{s/(s+1)},$$

and the desired result follows. ■

Remark D1 *The above lemma shows how one can relax the boundedness assumption in Theorem 3.4 of White and Wooldridge (1991) to obtain an exponential inequality for heterogeneous mixing processes with exponentially declining tail probabilities. Note that neither Theorem 3.4 of White and Wooldridge (1991) which deals with heterogeneity nor Theorem 3.5 of White and Wooldridge (1991) which deals with stationary mixing processes is sufficient for handling the heterogeneous mixing processes we consider.*

Remark D2 *It is important for the rest of the lemmas in this supplement, and in particular, the results obtained under (D4) of Lemma D2, to also note that Lemma 2 of Dendramis et al. (2015) provides the result of Lemma D1 when* $\delta = 0$.

Lemma D2 *Let* x_t, $\boldsymbol{q}_{\cdot t} = (q_{1,t}, q_{2,t}, ..., q_{l_T,t})'$, *and* u_t *be sequences of random variables and suppose that there exist finite positive constants* C_0 *and* C_1, *and* $s > 0$ *such that* $\sup_t \Pr(|x_t| > \alpha) \leq C_0 \exp(-C_1\alpha^s)$, $\sup_{i,t} \Pr(|q_{i,t}| > \alpha) \leq C_0 \exp(-C_1\alpha^s)$, *and* $\sup_t \Pr(|u_t| > \alpha) \leq C_0 \exp(-C_1\alpha^s)$, *for all* $\alpha > 0$. *Let* $\boldsymbol{\Sigma}_{qq} = \frac{1}{T}\sum_{t=1}^{T} E(\boldsymbol{q}_{\cdot t}\boldsymbol{q}'_{\cdot t})$ *be a nonsingular matrix such that* $0 < \left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F$. *Suppose that Assumption 5 holds for* x_t *and* $\boldsymbol{q}_{\cdot t}$, *and denote the corresponding projection residuals defined by (11) as* $u_{x,t} = x_t - \boldsymbol{\gamma}'_{qx,T}\boldsymbol{q}_{\cdot t}$. *Let* $\hat{\boldsymbol{u}}_x = (\hat{u}_{x,1}, \hat{u}_{x,2}, ..., \hat{u}_{x,T})'$ *denote the* $T \times 1$ *LS residual vector of the regression of* x_t *on* $\boldsymbol{q}_{\cdot t}$. *Let* $\mathcal{F}_t = \mathcal{F}_t^x \cup \mathcal{F}_t^u$, $\mathcal{F}_t^q = \sigma\left(\{\boldsymbol{q}_{\cdot t}\}_{s=1}^{t}\right)$ *and assume either (D1)* $E\left(u_{x,t}u_t - \mu_{xu,t}|\mathcal{F}_{t-1} \cup \mathcal{F}_{t-1}^q\right) = 0$, *where* $\mu_{xu,t} = E(u_{x,t}u_t)$, x_t *and* u_t *are martingale difference processes,* $\boldsymbol{q}_{\cdot t}$ *is an exponentially mixing process, and* $\zeta_T = o(T^{\lambda})$, *for all* $\lambda > 1/2$, *or (D2)* $E\left(u_{x,t}u_t - \mu_{xu,t}|\mathcal{F}_{t-1} \cup \mathcal{F}_{t-1}^q\right) = 0$, *where* $\mu_{xu,t} = E(u_{x,t}u_t)$, u_t *is a martingale difference processes,* x_t *and* $\boldsymbol{q}_{\cdot t}$ *are exponentially mixing processes, and* $\zeta_T = o(T^{\lambda})$, *for all* $\lambda > 1/2$, *or (D3)* x_t, u_t *and* $\boldsymbol{q}_{\cdot t}$ *are exponentially mixing processes, and* $\zeta_T = o(T^{\lambda})$, *for all* $\lambda > 1$, *or (D4)* x_t, u_t *and* $\boldsymbol{q}_{\cdot t}$ *are exponentially mixing processes, and* $\zeta_T = o(T^{\lambda})$, *for all* $\lambda > 1/2$. *Then, we have the following. If (D1) or (D2) hold, then, for any* π *in the range* $0 < \pi < 1$, *there exist finite positive constants* C_0 *and* C_1, *such that*

$$\Pr\left(\left|\sum_{t=1}^{T} x_t u_t - E(x_t u_t)\right| > \zeta_T\right) \leq \exp\left[\frac{-(1-\pi)^2\zeta_T^2}{2T\omega_{xu,1,T}^2}\right] + \exp\left[-C_0 T^{C_1}\right] \tag{B.120}$$

and

$$\Pr\left(\left|\sum_{t=1}^{T} \hat{u}_{x,t}u_t - \mu_{xu,t}\right| > \zeta_T\right) \leq \exp\left[\frac{-(1-\pi)^2\zeta_T^2}{2T\omega_{xu,T}^2}\right] + \exp\left[-C_0 T^{C_1}\right], \tag{B.121}$$

as long as $l_T = o(T^{1/3})$, *where* $\omega^2_{xu,1,T} = \frac{1}{T}\sum_{t=1}^{T} E\left[(x_t u_t - E(x_t u_t))^2\right]$, $\omega^2_{xu,T} = \frac{1}{T}\sum_{t=1}^{T} E\left[(u_{x,t} u_t - \mu_{xu,t})^2\right]$. *If (D3) holds*

$$\Pr\left(\left|\sum_{t=1}^{T} x_t u_t - E(x_t u_t)\right| > \zeta_T\right) \le \exp\left[-C_0 T^{C_1}\right], \tag{B.122}$$

for some $C_0, C_1 > 0$, *and*

$$\Pr\left(\left|\sum_{t=1}^{T} \hat{u}_{x,t} u_t - \mu_{xu,t}\right| > \zeta_T\right) \le \exp\left[-C_0 T^{C_1}\right], \tag{B.123}$$

for some $C_0, C_1 > 0$, *as long as* $l_T = o(T^{1/3})$. *Finally, if (D4) holds,*

$$\Pr\left(\left|\sum_{t=1}^{T} x_t u_t - E(x_t u_t)\right| > \zeta_T\right) \le C_1 \exp\left[-C_0 \left(\zeta_T T^{-1/2}\right)^{s/(s+2)}\right], \tag{B.124}$$

for some $C_0, C_1 > 0$, *and*

$$\Pr\left(\left|\sum_{t=1}^{T} \hat{u}_{x,t} u_t - \mu_{xu,t}\right| > \zeta_T\right) \le C_2 \exp\left[-C_3 \left(\zeta_T T^{-1/2}\right)^{s/(s+2)}\right] + \exp\left[-C_0 T^{C_1}\right], \tag{B.125}$$

for some $C_0, C_1, C_2, C_3 > 0$, *as long as* $l_T = o(T^{1/3})$.

Proof. We first prove the lemma under (D1) and then modify the derivations to establish that the results also hold under (D2)-(D4). The assumptions of the lemma state that there exists a regression model underlying $\hat{u}_{x,t}$ which is denoted by

$$x_t = \boldsymbol{\beta}_q' \boldsymbol{q}_{\cdot t} + u_{x,t}$$

for some $l \times 1$ vector, $\boldsymbol{\beta}_q$. Denoting $\boldsymbol{u}_x = (u_{x,1}, u_{x,2}, ..., u_{x,T})'$, $\boldsymbol{u} = (u_1, u_2, ..., u_T)'$, $\hat{\boldsymbol{\Sigma}}_{qq} = T^{-1}(\mathbf{Q}'\mathbf{Q})$, $\mathbf{Q} = (\boldsymbol{q}_1, \boldsymbol{q}_2, ..., \boldsymbol{q}_l)$, and $\boldsymbol{q}_{i\cdot} = (q_{i1}, q_{i2}, ..., q_{iT})'$, we have

$$\hat{\boldsymbol{u}}_x' \boldsymbol{u} = \boldsymbol{u}_x' \boldsymbol{u} - \left(T^{-1}\boldsymbol{u}_x'\mathbf{Q}\right)\hat{\boldsymbol{\Sigma}}_{qq}^{-1}\left(\mathbf{Q}'\boldsymbol{u}\right) = \boldsymbol{u}_x' \boldsymbol{u} - \left(T^{-1}\boldsymbol{u}_x'\mathbf{Q}\right)\left(\hat{\boldsymbol{\Sigma}}_{qq}^{-1} - \boldsymbol{\Sigma}_{qq}^{-1}\right)\left(\mathbf{Q}'\boldsymbol{u}\right) + \left(T^{-1}\boldsymbol{u}_x'\mathbf{Q}\right)\boldsymbol{\Sigma}_{qq}^{-1}\left(\mathbf{Q}'\boldsymbol{u}\right)$$

Noting that, since u_t is a martingale difference process with respect to $\sigma\left(\{u_s\}_{s=1}^{t-1}, \{u_{x,s}\}_{s=1}^{t}, \{q_s\}_{s=1}^{t}\right)$, by Lemma A4,

$$\Pr\left(|\boldsymbol{u}_x'\boldsymbol{u}| > \zeta_T\right) \le \exp\left[\frac{-(1-\pi)^2\zeta_T^2}{2T\omega^2_{xu,T}}\right]. \tag{B.126}$$

It therefore suffices to show that

$$\Pr\left(\left|\left(\frac{1}{T}\boldsymbol{u}_x'\mathbf{Q}\right)\left(\hat{\boldsymbol{\Sigma}}_{qq}^{-1} - \boldsymbol{\Sigma}_{qq}^{-1}\right)\left(\mathbf{Q}'\boldsymbol{u}\right)\right| > \zeta_T\right) \le \exp\left[-C_0 T^{C_1}\right] \tag{B.127}$$

and

$$\Pr\left(\left|\left(\frac{1}{T}\boldsymbol{u}_x'\mathbf{Q}\right)\boldsymbol{\Sigma}_{qq}^{-1}\left(\mathbf{Q}'\boldsymbol{u}\right)\right| > \zeta_T\right) \le \exp\left[-C_0 T^{C_1}\right] \tag{B.128}$$

We explore (B.126) and (B.127). We start with (B.126). We have by Lemma A11 that, for some sequence δ_T,[6]

$$\Pr\left(\left|\left(\frac{1}{T}\boldsymbol{u}_x'\mathbf{Q}\right)\left(\hat{\boldsymbol{\Sigma}}_{qq}^{-1}-\boldsymbol{\Sigma}_{qq}^{-1}\right)(\mathbf{Q}'\boldsymbol{u})\right|>\zeta_T\right)\leq$$

$$\Pr\left(\left\|\frac{1}{T}\boldsymbol{u}_x'\mathbf{Q}\right\|\left\|\left(\hat{\boldsymbol{\Sigma}}_{qq}^{-1}-\boldsymbol{\Sigma}_{qq}^{-1}\right)\right\|\left\|\mathbf{Q}'\boldsymbol{u}\right\|_F>\zeta_T\right)\leq\Pr\left(\left\|\left(\hat{\boldsymbol{\Sigma}}_{qq}^{-1}-\boldsymbol{\Sigma}_{qq}^{-1}\right)\right\|>\frac{\zeta_T}{\delta_T}\right)+$$

$$\Pr\left(\left\|\boldsymbol{u}_x'\mathbf{Q}\right\|_F\left\|\mathbf{Q}'\boldsymbol{u}\right\|_F>\delta_T T\right)\tag{B.130}$$

We consider the first term of the RHS of (B.130). Note that for all $1\leq i,j\leq l$.

$$\Pr\left(\left|\frac{1}{T}\sum_{t=1}^{T}\left[q_{it}q_{jt}-E(q_{it}q_{jt})\right]\right|>\zeta_T\right)\leq\exp(-C_0\left(T^{1/2}\zeta_T\right)^{s/(s+2)}),\tag{B.131}$$

since $q_{it}q_{jt}-E(q_{it}q_{jt})$ is a mixing process and $\sup_i\Pr\left(|q_{i,t}|>\alpha\right)\leq C_0\exp\left(-C_1\alpha^s\right)$, $s>0$. Then, by Lemma F3,

$$\Pr\left(\left\|\left(\hat{\boldsymbol{\Sigma}}_{qq}^{-1}-\boldsymbol{\Sigma}_{qq}^{-1}\right)\right\|>\frac{\zeta_T}{\delta_T}\right)\leq l_T^2\exp\left(\frac{-C_0T^{s/2(s+2)}\zeta_T^{s/(s+2)}}{\delta_T^{s/(s+2)}l_T^{s/(s+2)}\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{s/(s+1)}\left(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F+\frac{\zeta_T}{\delta_T}\right)^{s/(s+1)}}\right)+$$

$$l_T^2\exp\left(-C_0\frac{T^{s/2(s+2)}}{\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F^{s/(s+2)}l_T^{s/(s+2)}}\right)=$$

$$l_T^2\exp\left(-C_0\left(\frac{T^{1/2}\zeta_T}{\delta_T l_T\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F\left(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F+\frac{\zeta_T}{\delta_T}\right)}\right)^{s/(s+2)}\right)+$$

$$l_T^2\exp\left(-C_0\left(\frac{T^{1/2}}{\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F l_T}\right)^{s/(s+2)}\right).$$

We now consider the second term of the RHS of (B.130). By (B.12), we have

$$\Pr\left(\left\|\boldsymbol{u}_x'\mathbf{Q}\right\|_F\left\|\mathbf{Q}'\boldsymbol{u}\right\|_F>\delta_T T\right)\leq\Pr\left(\left\|\boldsymbol{u}_x'\mathbf{Q}\right\|_F>\delta_T^{1/2}T^{1/2}\right)+\Pr\left(\left\|\mathbf{Q}'\boldsymbol{u}\right\|_F>\delta_T^{1/2}T^{1/2}\right).$$

[6]In what follows we use

$$\Pr\left(|AB|>c\right)\leq\Pr\left(|A|\,|B|>c\right)\tag{B.129}$$

where A and B are random variables. To see this note that $|AB|\leq|A|\,|B|$. Further note that for any random variables $A_1>0$ and $A_2>0$ for which $A_2>A_1$ the occurrence of the event $\{A_1>c\}$, for any constant $c>0$, implies the occurrence of the event $\{A_2>c\}$. Therefore, $\Pr\left(A_2>c\right)\geq\Pr\left(A_1>c\right)$ proving the result.

Note that $\|\mathbf{Q}'\boldsymbol{u}\|_F^2 = \sum_{j=1}^{l_T}\left(\sum_{t=1}^{T} q_{jt}u_t\right)^2$, and

$$\Pr\left(\|\mathbf{Q}'\boldsymbol{u}\|_F > (\delta_T T)^{1/2}\right) = \Pr\left(\|\mathbf{Q}'\boldsymbol{u}\|_F^2 > \delta_T T\right)$$
$$\leq \sum_{j=1}^{l_T} \Pr\left[\left(\sum_{t=1}^{T} q_{jt}u_t\right)^2 > \frac{\delta_T T}{l_T}\right]$$
$$= \sum_{j=1}^{l_T} \Pr\left[\left|\sum_{t=1}^{T} q_{jt}u_t\right| > \left(\frac{\delta_T T}{l_T}\right)^{1/2}\right].$$

Noting further that $q_{it}u_t$ and $q_{it}u_{xt}$ are martingale difference processes satisfying a result of the usual form we obtain

$$\Pr\left(\|\boldsymbol{u}_x'\mathbf{Q}\|_F > \delta_T^{1/2}T^{1/2}\right) \leq l_T \Pr\left(|\boldsymbol{u}_x'\mathbf{q}_i| > \frac{\delta_T^{1/2}T^{1/2}}{l_T^{1/2}}\right) \leq l_T \exp\left(\frac{-C\delta_T}{l_T}\right),$$

or

$$\Pr\left(\|\boldsymbol{u}_x'\mathbf{Q}\|_F > \delta_T^{1/2}T^{1/2}\right) \leq l_T \Pr\left(|\boldsymbol{u}_x'\mathbf{q}_i| > \frac{\delta_T^{1/2}T^{1/2}}{l_T^{1/2}}\right) \leq l_T \exp\left(\left(\frac{-\delta_T T}{l_T}\right)^{s/2(s+2)}\right),$$

depending on the order of magnitude of $\frac{\delta_T^{1/2}T^{1/2}}{l_T^{1/2}}$, and a similar result for $\Pr\left(\|\mathbf{Q}'\boldsymbol{u}\|_F > \delta_T^{1/2}T^{1/2}\right)$. Therefore,

$$\Pr\left(\|\boldsymbol{u}_x'\mathbf{Q}\|_F \|\mathbf{Q}'\boldsymbol{u}\|_F > \delta_T T\right) \leq \exp\left[-C_0 T^{C_1}\right]. \tag{B.132}$$

We wish to derive conditions for l_T under which $\frac{T^{1/2}\zeta_T}{\delta_T l_T \left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F\left(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F + \frac{\zeta_T}{\delta_T}\right)}$, $\frac{T^{1/2}}{\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F l_T}$, and $\frac{\delta_T}{l_T}$ are of larger, polynomial in T, order than $\frac{\zeta_T^2}{T}$. Then, the factors in l_T in (B.26) and (B.132) are negligible. We let $\zeta_T = T^{\lambda}$, $l_T = T^d$, $\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F = l_T^{1/2} = T^{d/2}$ and $\delta_T = T^{\alpha}$, where $\alpha \geq 0$, can be chosen freely. This is a complex analysis and we simplify it by considering relevant values for our setting and, in particular, $\lambda \geq 1/2$, $\lambda < 1/2 + c$, for all $c > 1/2$, and $d < 1$. We have

$$\frac{T^{1/2}\zeta_T}{\delta_T l_T \left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F\left(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F + \frac{\zeta_T}{\delta_T}\right)} = O\left(T^{1/2+\lambda-\alpha-2d}\right) + O\left(T^{1/2-3d/2}\right), \tag{B.133}$$

$$\frac{T^{1/2}}{\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F l_T} = O\left(T^{1/2-3d/2}\right), \tag{B.134}$$

$$\frac{\delta_T}{l_T} = O\left(T^{\alpha-d}\right), \tag{B.135}$$

and

$$\frac{\zeta_T^2}{T} = O\left(T^{2\lambda-1}\right) = O\left(c\ln T\right). \tag{B.136}$$

Clearly $d < 1/3$. Setting $\alpha = 1/3$, ensures all conditions are satisfied. Since $\boldsymbol{\Sigma}_{qq}^{-1}$ is of lower norm order than $\hat{\boldsymbol{\Sigma}}_{qq}^{-1} - \boldsymbol{\Sigma}_{qq}^{-1}$, (B.128) follows similarly proving the result under (D1). For (D2) and (D3) we proceed as follows. Under (D3), noting that u_t is a mixing process, then by Lemma D1, we have that (B.126) is replaced by

$$\Pr\left(\left|\boldsymbol{u}_x'\boldsymbol{u}\right| > \zeta_T\right) \leq \exp\left[-C_0\left(T^{-(1+\vartheta)/2}\zeta_T\right)^{s/(s+2)}\right], \tag{B.137}$$

else, under (D2), we have again that (B.126) holds. Further, by a similar analysis to that above, it is easily seen that, under (D2),

$$\Pr\left(\left\|\boldsymbol{u}_x'\mathbf{Q}\right\|_F\left\|\mathbf{Q}'\boldsymbol{u}\right\|_F > \delta_T T\right) \leq l_T\exp\left(\frac{-C\delta_T}{l_T}\right) + l_T\exp\left[-C_0\left(\frac{T^{-\vartheta/2}\delta_T^{1/2}}{l_T^{1/2}}\right)^{s/(s+2)}\right],$$

and under (D3),

$$\Pr\left(\left\|\boldsymbol{u}_x'\mathbf{Q}\right\|_F\left\|\mathbf{Q}'\boldsymbol{u}\right\|_F > \delta_T T\right) \leq 2l_T\exp\left[-C_0\left(\frac{T^{-\vartheta/2}\delta_T}{l_T}\right)^{s/2(s+2)}\right].$$

Under (D2), we wish to derive conditions for l_T under which $\frac{T^{1/2}\zeta_T}{\delta_T l_T\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F\left(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F+\frac{\zeta_T}{\delta_T}\right)}$, $\frac{T^{1/2}}{\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F l_T}$, and $\frac{\delta_T}{l_T}$ are of larger, polynomial in T, order than $\frac{\zeta_T^2}{T}$. But this is the same requirement to that under (D1). Under (D3), we wish to derive conditions for l_T under which $\frac{T^{1/2}\zeta_T}{\delta_T l_T\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F\left(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F+\frac{\zeta_T}{\delta_T}\right)}$, $\frac{T^{1/2}}{\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F l_T}$, $\frac{\delta_T}{l_T}$ and $\left(T^{-1/2}\zeta_T\right)^{s/(s+2)}$ are of positive polynomial in T, order. But again the same conditions are needed as for (D1) and (D2). Finally, we consider (D4). But, noting Remark D2, the only difference to (D3) is that $\zeta_T \geq T^{1/2}$, rather than $\zeta_T \geq T$. Then, as long as $\left(T^{-1/2}\zeta_T\right)^{s/(s+2)} \to \infty$ the result follows. ■

Lemma D3 *Let y_t, for $t = 1, 2, ..., T$, be given by the data generating process (6) and suppose that u_t and $\boldsymbol{x}_{nt} = (x_{1t}, x_{2t}, ..., x_{nt})'$ satisfy Assumptions 2-4. Let $\boldsymbol{q}_{\cdot t} = (q_{1,t}, q_{2,t}, ..., q_{l_T,t})'$ contain a constant and a subset of $\boldsymbol{x}_{nt}$, and let $\eta_t = \boldsymbol{x}_{b,t}'\boldsymbol{\beta}_b + u_t$, where $\boldsymbol{x}_{b,t}$ is $k_b \times 1$ dimensional vector of signal variables that do not belong to $\boldsymbol{q}_{\cdot t}$, with the associated coefficients, $\boldsymbol{\beta}_b$. Assume that $\boldsymbol{\Sigma}_{qq} = \frac{1}{T}\sum_{t=1}^{T} E\left(\boldsymbol{q}_{\cdot t}\boldsymbol{q}_{\cdot t}'\right)$ and $\hat{\boldsymbol{\Sigma}}_{qq} = \mathbf{Q}'\mathbf{Q}/T$ are both invertible, where $\mathbf{Q} = (\boldsymbol{q}_{1\cdot}, \boldsymbol{q}_{2\cdot}, ..., \boldsymbol{q}_{l_T\cdot})$ and $\boldsymbol{q}_{i\cdot} = (q_{i1}, q_{i2}, ..., q_{iT})'$, for $i = 1, 2, ..., l_T$. Moreover, let $l_T = o(T^{1/4})$ and suppose that Assumption 5 holds for x_t and $\boldsymbol{q}_{\cdot t}$, where x_t is a generic element of $\{x_{1t}, x_{2t}, ..., x_{nt}\}$ that does not belong to $\boldsymbol{q}_{\cdot t}$. Denote the corresponding projection residuals defined by (11) as $u_{x,t} = x_t - \boldsymbol{\gamma}_{qx,T}'\boldsymbol{q}_{\cdot t}$, and the projection residuals of y_t on $(\boldsymbol{q}_{\cdot t}', x_t)'$ as $e_t = y_t - \boldsymbol{\gamma}_{yqx,T}'(\boldsymbol{q}_{\cdot t}', x_t)'$. Define $\boldsymbol{x} = (x_1, x_2, ..., x_T)'$, $\boldsymbol{y} = (y_1, y_2, ..., y_T)'$, $\boldsymbol{e} = (e_1, e_2, ..., e_T)'$, $\mathbf{M}_q = \mathbf{I}_T - \mathbf{Q}(\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}'$, and $\theta = E\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\mathbf{X}_b\right)\boldsymbol{\beta}_b$, where $\mathbf{X}_b$ is $T \times k_b$ matrix of observations on $\boldsymbol{x}_{b,t}$. Finally, $c_p(n, \delta)$ is such that $c_p(n, \delta) = o\left(\sqrt{T}\right)$. Then, under Assumption D1,for any π in the range $0 < \pi < 1$, $d_T > 0$ and bounded in T, and for some $C_i, c > 0$ for $i = 0, 1$,*

$$\Pr\left[|t_x| > c_p(n, \delta)\,|\theta = 0\right] \leq \exp\left[\frac{-(1-\pi)^2\sigma_{e,(T)}^2\sigma_{x,(T)}^2 c_p^2(n, \delta)}{2(1+d_T)^2\omega_{xe,T}^2}\right] + \exp\left(-C_0 T^{C_1}\right), \tag{B.138}$$

where

$$t_x = \frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}},$$

$$\sigma^2_{e,(T)} = E\left(T^{-1}\boldsymbol{e}'\boldsymbol{e}\right),\ \sigma^2_{x,(T)} = E\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}\right),$$

and

$$\omega^2_{xe,T} = \frac{1}{T}\sum_{t=1}^{T} E\left[\left(u_{x,t}\eta_t\right)^2\right].$$

Under $\sigma_t^2 = \sigma^2$ *and/or* $E\left(u^2_{x,t}\right) = \sigma^2_{xt} = \sigma^2_x$, *for all* $t = 1, 2, ..., T$,

$$\Pr\left[|t_x| > c_p(n,\delta)\,|\theta = 0\right] \leq \exp\left[\frac{-(1-\pi)^2 c_p^2(n,\delta)}{2(1+d_T)^2}\right] + \exp\left(-C_0 T^{C_1}\right).$$

In the case where $\theta > 0$, *and assuming that there exists* T_0 *such that for all* $T > T_0$, $\lambda_T - c_p(n,\delta)/\sqrt{T} > 0$, *where* $\lambda_T = \theta/\left(\sigma_{x,(T)}\sigma_{e,(T)}\right)$, *then for* $d_T > 0$ *and bounded in* T *and some* $C_i > 0$, $i = 0, 1, 2$, *we have*

$$\Pr\left[|t_x| > c_p(n,\delta)\,|\theta \neq 0\right] > 1 - \exp\left(-C_0 T^{C_1}\right). \tag{B.139}$$

Under Assumption D2, for some $C_0, C_1, C_2 > 0$,

$$\Pr\left[|t_x| > c_p(n,\delta)\,|\theta = 0\right] \leq \exp\left[-C_2 c_p(n,\delta)^{s/(s+2)}\right] + \exp\left(-C_0 T^{C_1}\right), \tag{B.140}$$

and

$$\Pr\left[|t_x| > c_p(n,\delta)\,|\theta \neq 0\right] > 1 - \exp\left(-C_0 T^{C_1}\right). \tag{B.141}$$

Proof. We start under Assumption D1 and in the end note the steps that differ under Assumption D2. We recall that the DGP, given by (7), can be written as

$$\boldsymbol{y} = a\boldsymbol{\tau}_T + \mathbf{X}_k\boldsymbol{\beta} + \boldsymbol{u} = a\boldsymbol{\tau}_T + \mathbf{X}_a\boldsymbol{\beta}_a + \mathbf{X}_b\boldsymbol{\beta}_b + \boldsymbol{u}$$

where $\mathbf{X}_a$ is a subset of $\mathbf{Q}$. Recall that $\mathbf{Q}_x = (\mathbf{Q}, \boldsymbol{x})$, $\mathbf{M}_q = \mathbf{I}_T - \mathbf{Q}(\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}'$, $\mathbf{M}_{qx} = \mathbf{I}_T - \mathbf{Q}_x(\mathbf{Q}_x'\mathbf{Q}_x)^{-1}\mathbf{Q}_x'$. Then, $\mathbf{M}_q\mathbf{X}_a = \mathbf{0}$, and let $\mathbf{M}_q\mathbf{X}_b = (\boldsymbol{x}_{bq,1}, \boldsymbol{x}_{bq,2}, ..., \boldsymbol{x}_{bq,T})'$. Then,

$$t_x = \frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}} = \frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\mathbf{X}_b\boldsymbol{\beta}_b}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}} + \frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{u}}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}}.$$

Let $\theta = E\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\mathbf{X}_b\right)\boldsymbol{\beta}_b$, $\boldsymbol{\eta} = \mathbf{X}_b\boldsymbol{\beta}_b + \boldsymbol{u}$, $\boldsymbol{\eta} = (\eta_1, \eta_2, ..., \eta_T)'$, and write (B.53) as

$$t_x = \frac{\sqrt{T}\theta}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}} + \frac{T^{1/2}\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}}{T} - \theta\right)}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}}.$$

First consider the case where $\theta = 0$, and note that in this case

$$t_x = \frac{T^{1/2}\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)^{-1/2}\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}}{T}}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)}}.$$

Now by (B.46) of Lemma A9 and (B.121) of Lemma D2, we have

$$\Pr\left[|t_x| > c_p(n,\delta)\,|\theta=0\right] = \Pr\left[\left|\frac{T^{1/2}\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)^{-1/2}\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}}{T}}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)}}\right| > c_p(n,\delta)\,|\theta=0\right] \leq \tag{B.142}$$

$$\Pr\left(\left|\frac{T^{1/2}\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)^{-1/2}\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}}{T}}{\sigma_{e,(T)}}\right| > \frac{c_p(n,\delta)}{1+d_T}\right) + \exp\left(-C_0T^{C_1}\right).$$

Then, by Lemma F1, under Assumption D1 and defining $\boldsymbol{\alpha}(\boldsymbol{X}_T) = \left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)^{-1/2}\boldsymbol{x}'\mathbf{M}_q$ where $\boldsymbol{\alpha}(\boldsymbol{X}_T)$ is exogenous to y_t, $\boldsymbol{\alpha}(\boldsymbol{X}_T)'\boldsymbol{\alpha}(\boldsymbol{X}_T) = 1$ and by (B.121) of Lemma D2, we have,

$$\Pr\left[|t_x| > c_p(n,\delta)\,|\theta=0\right] \leq \exp\left[\frac{-(1-\pi)^2\sigma^2_{e,(T)}\sigma^2_{x,(T)}c_p^2(n,\delta)}{2(1+d_T)^2\omega^2_{xe,T}}\right] + \exp\left(-C_0T^{C_1}\right) \tag{B.143}$$

where

$$\omega^2_{xe,T} = \frac{1}{T}\sum_{t=1}^{T}E\left[(u_{x,t}\eta_t)^2\right] = \frac{1}{T}\sum_{t=1}^{T}E\left[u^2_{x,t}\left(\boldsymbol{x}'_{b,t}\boldsymbol{\beta}_b + u_t\right)^2\right],$$

and $u_{x,t}$, being the error in the regression of x_t on $\mathbf{Q}$, is defined by (11). Since by assumption u_t are distributed independently of $u_{x,t}$ and $\boldsymbol{x}_{b,t}$, then

$$\omega^2_{xe,T} = \frac{1}{T}\sum_{t=1}^{T}E\left[u^2_{x,t}\left(\boldsymbol{x}'_{bq,t}\boldsymbol{\beta}_b\right)^2\right] + \frac{1}{T}\sum_{t=1}^{T}E\left(u^2_{xt}\right)E\left(u^2_t\right),$$

where $\boldsymbol{x}'_{bq,t}\boldsymbol{\beta}_b$ is the t-th element of $\mathbf{M}_q\mathbf{X}_b\boldsymbol{\beta}_b$. Furthermore $E\left[u^2_{x,t}\left(\boldsymbol{x}'_{bq,t}\boldsymbol{\beta}_b\right)^2\right] = E\left(u^2_{x,t}\right)E\left(\boldsymbol{x}'_{bq,t}\boldsymbol{\beta}_b\right)^2 = E\left(u^2_{x,t}\right)\boldsymbol{\beta}'_bE\left(\boldsymbol{x}_{bq,t}\boldsymbol{x}'_{bq,t}\right)\boldsymbol{\beta}_b$, noting that under $\theta = 0$, $u_{x,t}$ and $\boldsymbol{x}_{b,t}$ are independently distributed. Hence

$$\omega^2_{xe,T} = \frac{1}{T}\sum_{t=1}^{T}E\left(u^2_{x,t}\right)\boldsymbol{\beta}'_bE\left(\boldsymbol{x}_{bq,t}\boldsymbol{x}'_{bq,t}\right)\boldsymbol{\beta}_b + \frac{1}{T}\sum_{t=1}^{T}E\left(u^2_{xt}\right)E\left(u^2_t\right)$$

Similarly

$$\begin{aligned}\sigma^2_{e,(T)} &= E\left(T^{-1}\boldsymbol{e}'\boldsymbol{e}\right) = E\left(T^{-1}\boldsymbol{\eta}'\mathbf{M}_{qx}\boldsymbol{\eta}\right) = E\left[T^{-1}\left(\mathbf{X}_b\boldsymbol{\beta}_b + \boldsymbol{u}\right)'\mathbf{M}_{qx}\left(\mathbf{X}_b\boldsymbol{\beta}_b + \boldsymbol{u}\right)\right] \\ &= \boldsymbol{\beta}'_bE\left(T^{-1}\mathbf{X}'_b\mathbf{M}_{qx}\mathbf{X}_b\right)\boldsymbol{\beta}_b + \frac{1}{T}\sum_{t=1}^{T}E\left(u^2_t\right),\end{aligned}$$

and since under $\theta = 0$, $\boldsymbol{x}$ being a noise variable will be distributed independently of $\mathbf{X}_b$, then $E\left(T^{-1}\mathbf{X}_b'\mathbf{M}_{qx}\mathbf{X}_b\right) = E\left(T^{-1}\mathbf{X}_b'\mathbf{M}_q\mathbf{X}_b\right)$, and we have

$$\begin{aligned}\sigma^2_{e,(T)} &= \boldsymbol{\beta}_b' E\left(T^{-1}\mathbf{X}_b'\mathbf{M}_q\mathbf{X}_b\right)\boldsymbol{\beta}_b + \frac{1}{T}\sum_{t=1}^{T} E\left(u_t^2\right) \\ &= \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{\beta}_b' E\left(\boldsymbol{x}_{bq,t}\boldsymbol{x}_{bq,t}'\right)\boldsymbol{\beta}_b + \frac{1}{T}\sum_{t=1}^{T} E\left(u_t^2\right).\end{aligned}$$

Using (B.55) and (B.56), it is now easily seen that if either $E\left(u_{x,t}^2\right) = \sigma^2_{ux}$ or $E\left(u_t^2\right) = \sigma^2$, for all t, then we have $\omega^2_{xe,T} = \sigma^2_{e,(T)}\sigma^2_{x,(T)}$, and hence

$$\Pr\left[|t_x| > c_p(n,\delta)\,|\theta = 0\right] \le \exp\left[\frac{-(1-\pi)^2 c_p^2(n,\delta)}{2(1+d_T)^2}\right] + \exp\left(-C_0 T^{C_1}\right).$$

giving a rate that does not depend on error variances. Next, we consider $\theta \neq 0$. By (B.45) of Lemma A9, for $d_T > 0$,

$$\Pr\left[\left|\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}}\right| > c_p(n,\delta)\right] \le \Pr\left(\left|\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| > \frac{c_p(n,\delta)}{1+d_T}\right) + \exp\left(-C_0 T^{C_1}\right).$$

We then have

$$\begin{aligned}\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sigma_{e,(T)}\sigma_{x,(T)}} &= \frac{T^{1/2}\left(\frac{\boldsymbol{x}'\mathbf{M}_q\mathbf{X}_b\boldsymbol{\beta}_b}{T} - \theta\right)}{\sigma_{e,(T)}\sigma_{x,(T)}} + \frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{u}}{\sigma_{e,(T)}\sigma_{x,(T)}} + \frac{T^{1/2}\theta}{\sigma_{e,(T)}\sigma_{x,(T)}} \\ &= \frac{T^{1/2}\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}}{T} - \theta\right)}{\sigma_{e,(T)}\sigma_{x,(T)}} + \frac{T^{1/2}\theta}{\sigma_{e,(T)}\sigma_{x,(T)}}.\end{aligned}$$

Then

$$\begin{aligned}&\Pr\left(\left|\frac{T^{1/2}\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}}{T} - \theta\right)}{\sigma_{e,(T)}\sigma_{x,(T)}} + \frac{T^{1/2}\theta}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| > \frac{c_p(n,\delta)}{1+d_T}\right) \\ &= 1 - \Pr\left(\left|\frac{T^{1/2}\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}}{T} - \theta\right)}{\sigma_{e,(T)}\sigma_{x,(T)}} + \frac{T^{1/2}\theta}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| \le \frac{c_p(n,\delta)}{1+d_T}\right).\end{aligned}$$

We note that, by Lemma A12,

$$\begin{aligned}&\Pr\left(\left|\frac{T^{1/2}\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}}{T} - \theta\right)}{\sigma_{e,(T)}\sigma_{x,(T)}} + \frac{T^{1/2}\theta}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| \le \frac{c_p(n,\delta)}{1+d_T}\right) \\ &\le \Pr\left(\left|\frac{T^{1/2}\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}}{T} - \theta\right)}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| > \frac{T^{1/2}|\theta|}{\sigma_{e,(T)}\sigma_{x,(T)}} - \frac{c_p(n,\delta)}{1+d_T}\right).\end{aligned}$$

But $(T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}-\theta)$ is the average of a martingale difference process and so

$$\Pr\left(\left|\frac{T^{1/2}\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}}{T}-\theta\right)}{\sigma_{e,(T)}\sigma_{x,(T)}}\right|>\frac{T^{1/2}\left|\theta\right|}{\sigma_{e,(T)}\sigma_{x,(T)}}-\frac{c_p\left(n,\delta\right)}{1+d_T}\right) \tag{B.144}$$
$$\leq \exp\left[-C_1\left(T^{1/2}\left(\frac{T^{1/2}\left|\theta\right|}{\sigma_{e,(T)}\sigma_{x,(T)}}-\frac{\theta c_p\left(n,\delta\right)}{1+d_T}\right)\right)^{s/(s+2)}\right].$$

So overall

$$\Pr\left[\left|\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sqrt{\left(\boldsymbol{e}'\boldsymbol{e}/T\right)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}}\right|>c_p\left(n,\delta\right)\right]>1-\exp\left(-C_0T^{C_1}\right)$$
$$-\exp\left[-C_1\left(T^{1/2}\left(\frac{T^{1/2}\left|\theta\right|}{\sigma_{e,(T)}\sigma_{x,(T)}}-\frac{\theta c_p\left(n,\delta\right)}{1+d_T}\right)\right)^{s/(s+2)}\right].$$

Finally, we note the changes needed to the above arguments when Assumption D2 holds, rather than D1. (B.140) follows if in (B.142) we use (B.125) of Lemma D2 rather than (B.121) and, in (B.143), we use Lemma F2 rather than Lemma F1 and, again, we use (B.125) of Lemma D2 rather than (B.121). (B.140) follows again by using (B.125) of Lemma D2 rather than (B.121). ■

Remark D3 *We note that the above proof makes use of Lemmas F1 and F2. Alternatively one can use (B.45) of Lemma A9 in (B.142)-(B.143), rather that (B.46) of Lemma A9 and use the same line of proof as that provided in Lemma A10. However, we consider this line of proof as Lemmas F1 and F2 are of independent interest.*

E. Lemmas for the deterministic case

Lemmas E1 and E2 provide the necessary justification for the case where x_{it} are bounded deterministic sequences, by replacing Lemmas A6 and A10.

Lemma E1 *Let x_{it}, $i=1,2,...,n$, be a set of bounded deterministic sequences and u_t satisfy Assumption 2 and condition (10) of Assumption 4, and consider the data generating process (6) with k signal variables $x_{1t},x_{2t},...,x_{kt}$. Let $\boldsymbol{q}_{\cdot t}=(q_{1,t},q_{2,t},...,q_{l_T,t})'$ contain a constant and a subset of $\boldsymbol{x}_{nt}=(x_{1t},x_{2t},...,x_{nt})'$. Let $\eta_t=\boldsymbol{x}_{b,t}\boldsymbol{\beta}_b+u_{\eta,t}$, where $\boldsymbol{x}_{b,t}$ contains all signals that do not belong to $\boldsymbol{q}_{\cdot t}$. Let $\boldsymbol{\Sigma}_{qq}=\mathbf{Q}'\mathbf{Q}/T$ be invertible for all T, and $\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_{FF}=O\left(\sqrt{l_T}\right)$, where $\mathbf{Q}=(\boldsymbol{q}_{1\cdot},\boldsymbol{q}_{2\cdot},...,\boldsymbol{q}_{l_T\cdot})$ and $\boldsymbol{q}_{i\cdot}=(q_{i1},q_{i2},...,q_{iT})'$, for $i=1,2,...,l_T$. Suppose that Assumption 5 holds for x_{it} and $\boldsymbol{q}_{\cdot t}$, and u_t and $\boldsymbol{q}_{\cdot t}$. Let $u_{x_i,T}$ be as in (11), such that $\sup_{i,j}\lim_{T\to\infty}\frac{\left\|\mathbf{q}_i'\boldsymbol{u}_{x_j,T}\right\|}{T^{1/2}}<C<\infty$, and let $\hat{\boldsymbol{u}}_{x_i}=(\hat{u}_{x_i,1},\hat{u}_{x_i,2},...,\hat{u}_{x_i,T})'=\mathbf{M}_q\boldsymbol{x}_i$, $\boldsymbol{x}_i=(x_{i1},x_{i2},...,x_{iT})'$, $\hat{\boldsymbol{u}}_\eta=(\hat{u}_{\eta,1},\hat{u}_{\eta,2},...,\hat{u}_{\eta,T})'=\mathbf{M}_q\boldsymbol{\eta}$, $\boldsymbol{\eta}=(\eta_1,\eta_2,...,\eta_T)'$, $\mathbf{M}_q=\mathbf{I}_T-\mathbf{Q}\left(\mathbf{Q}'\mathbf{Q}\right)^{-1}\mathbf{Q}$,*

$\mathcal{F}_t = \mathcal{F}_t^x \cup \mathcal{F}_t^u$, $\mu_{x_i\eta,t} = E\left(u_{x_i,t}u_{\eta,t}\left|\mathcal{F}_{t-1}\right.\right)$, $\omega^2_{x_i\eta,1,T} = \frac{1}{T}\sum_{t=1}^T E\left[\left(x_{it}\eta_t - E\left(x_{it}\eta_t\left|\mathcal{F}_{t-1}\right.\right)\right)^2\right]$ *and* $\omega^2_{x_i\eta,T} = \frac{1}{T}\sum_{t=1}^T E\left[\left(u_{x_i,t}u_{\eta,t} - \mu_{x_i\eta,t}\right)^2\right]$. *Then, for any* π *in the range* $0 < \pi < 1$, *we have, under Assumption 3,*

$$\Pr\left(\left|\sum_{t=1}^T x_{it}\eta_t - E\left(x_{it}\eta_t\left|\mathcal{F}_{t-1}\right.\right)\right| > \zeta_T\right) \leq \exp\left[\frac{-\left(1-\pi\right)^2\zeta_T^2}{2T\omega^2_{x_i\eta,1,T}}\right], \tag{B.145}$$

where $\zeta_T = O\left(T^\lambda\right)$, *and* $(s+1)/(s+2) \geq \lambda$. *If* $(s+1)/(s+2) < \lambda$,

$$\Pr\left(\left|\sum_{t=1}^T x_{it}\eta_t - E\left(x_{it}\eta_t\left|\mathcal{F}_{t-1}\right.\right)\right| > \zeta_T\right) \leq \exp\left[-C_0\zeta_T^{s/(s+1)}\right],$$

for some $C_0 > 0$. *If it is further assumed that* $l_T = O\left(T^d\right)$, *for some* λ *and* d *such that* $d < 1/3$, *and* $1/2 \leq \lambda \leq (s+1)/(s+2)$, *then*

$$\Pr\left(\left|\sum_{t=1}^T \left(\hat{u}_{x_i,t}u_{\eta,t} - \mu_{x_i\eta,t}\right)\right| > \zeta_T\right) \leq C_2\exp\left[\frac{-\left(1-\pi\right)^2\zeta_T^2}{2T\omega^2_{x_i\eta,T}}\right] + \exp\left(-C_0T^{C_1}\right).$$

for some $C_0, C_1, C_2 > 0$. *Otherwise, if* $\lambda > (s+1)/(s+2)$,

$$\Pr\left(\left|\sum_{t=1}^T \left(\hat{u}_{x_i,t}u_{\eta,t} - \mu_{x_i\eta,t}\right)\right| > \zeta_T\right) \leq \exp\left[-C_2\zeta_T^{s/(s+1)}\right] + \exp\left(-C_0T^{C_1}\right),$$

for some $C_0, C_1, C_2 > 0$.

Proof. Note that all results used in this proof hold both for sequences and triangular arrays. (B.145) follows immediately given our assumptions and Lemma A3. We proceed to prove the rest of the lemma. Note that now $\hat{\boldsymbol{u}}_{x_i}$ is a bounded deterministic vector and $\boldsymbol{u}_{x_i} = (u_{x_i,1}, u_{x_i,2}, ..., u_{x_i,T})'$ a segment of dimension T of its limit. We first note that

$$\sum_{t=1}^T \left(\hat{u}_{x_i,t}\hat{u}_{\eta,t} - \mu_{x_i\eta,t}\right) = \hat{\boldsymbol{u}}'_{x_i}\hat{\boldsymbol{u}}_\eta - \sum_{t=1}^T \mu_{x_i\eta,t} = \boldsymbol{u}'_{x_i}\mathbf{M}_q\boldsymbol{u}_\eta - \sum_{t=1}^T \mu_{x_i\eta,t}$$
$$= \sum_{t=1}^T \left(u_{x_i,t}u_{\eta,t} - \mu_{x_i\eta,t}\right) - \left(T^{-1}\boldsymbol{u}'_{x_i}\mathbf{Q}\right)\boldsymbol{\Sigma}_{qq}^{-1}\left(\mathbf{Q}'\boldsymbol{u}_\eta\right),$$

where $\boldsymbol{u}_x = (u_{x,1}, u_{x,2}, ..., u_{x,T})'$ and $\boldsymbol{u}_\eta = (u_{\eta,1}, u_{x,2}, ..., u_{\eta,T})'$. By (B.59) and for any $0 < \pi_i < 1$ such that $\sum_{i=1}^2 \pi_i = 1$,we have

$$\Pr\left(\left|\sum_{t=1}^T \left(\hat{u}_{x_i,t}\hat{u}_{\eta,t} - \mu_{x_i\eta,t}\right)\right| > \zeta_T\right) \leq \Pr\left(\left|\sum_{t=1}^T \left(u_{x_i,t}u_{\eta,t} - \mu_{x_i\eta,t}\right)\right| > \pi_1\zeta_T\right)$$
$$+ \Pr\left(\left|\left(T^{-1}\boldsymbol{u}'_{x_i}\mathbf{Q}\right)\boldsymbol{\Sigma}_{qq}^{-1}\left(\mathbf{Q}'\boldsymbol{u}_\eta\right)\right| > \pi_2\zeta_T\right).$$

Also applying (B.60) to the last term of the above we obtain

$$
\begin{aligned}
&\Pr\left(\left|\left(T^{-1}\boldsymbol{u}_{x_i}'\mathbf{Q}\right)\boldsymbol{\Sigma}_{qq}^{-1}\left(\mathbf{Q}'\boldsymbol{u}_\eta\right)\right|>\pi_2\zeta_T\right)\\
&\leq \Pr\left(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F\left\|T^{-1}\boldsymbol{u}_{x_i}'\mathbf{Q}\right\|_F\left\|\mathbf{Q}'\boldsymbol{u}_\eta\right\|_F>\pi_2\zeta_T\right)\\
&\leq \Pr\left(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F>\frac{\pi_2\zeta_T}{\delta_T}\right)+\Pr\left(T^{-1}\left\|\boldsymbol{u}_{x_i}'\mathbf{Q}\right\|_F\left\|\mathbf{Q}'\boldsymbol{u}_\eta\right\|_F>\pi_2\delta_T\right)\\
&\leq \Pr\left(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F>\frac{\pi_2\zeta_T}{\delta_T}\right)+\Pr\left(\left\|\boldsymbol{u}_{x_i}'\mathbf{Q}\right\|_F>\left(\pi_2\delta_T T\right)^{1/2}\right)\\
&+\Pr\left(\left\|\mathbf{Q}'\boldsymbol{u}_\eta\right\|_F>\left(\pi_2\delta_T T\right)^{1/2}\right),
\end{aligned}
$$

where $\delta_T>0$ is a deterministic sequence. In what follows we set $\delta_T=O\left(\zeta_T^\alpha\right)$, with $0<\alpha<\lambda$, so that ζ_T/δ_T is rising in T. Overall

$$
\begin{aligned}
&\Pr\left(\left|\sum_{t=1}^{T}\left(\hat{u}_{x,t}u_{\eta,t}-\mu_{x\eta,t}\right)\right|>\zeta_T\right) \qquad \text{(B.146)}\\
&\leq \Pr\left(\left|\sum_{t=1}^{T}\left(u_{x,t}u_{\eta,t}-\mu_{x\eta,t}\right)\right|>\pi_1\zeta_T\right)+\Pr\left(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F>\frac{\pi_2\zeta_T}{\delta_T}\right)\\
&+\Pr\left(\left\|\mathbf{Q}'\boldsymbol{u}_\eta\right\|_F>\left(\pi_2\delta_T T\right)^{1/2}\right)+\Pr\left(\left\|\boldsymbol{u}_x'\mathbf{Q}\right\|_F>\left(\pi_2\delta_T T\right)^{1/2}\right).
\end{aligned}
$$

We consider the four terms of the above, and note that since by assumption $\{q_{it}u_{\eta,t}\}$ are martingale difference sequences and satisfy the required probability bound conditions of Lemma A4, and $\{q_{it}u_{x_i,t}\}$ are bounded sequences, then for some $C,c>0$ we have[7]

$$
\sup_i \Pr\left(\left\|\mathbf{q}_i'\boldsymbol{u}_\eta\right\|>\left(\pi_2\delta_T T\right)^{1/2}\right)\leq\exp\left(-C_0T^{C_1}\right)
$$

and as long as $l_T=o\left(\delta_T\right)$,

$$
\Pr\left(\left\|\boldsymbol{u}_x'\mathbf{Q}\right\|_F>\left(\pi_2\delta_T T\right)^{1/2}\right)=0
$$

Also, since $\left\|\mathbf{Q}'\boldsymbol{u}_\eta\right\|_F^2=\sum_{j=1}^{l_T}\left(\sum_{t=1}^{T}q_{jt}u_t\right)^2$,

$$
\begin{aligned}
&\Pr\left(\left\|\mathbf{Q}'\boldsymbol{u}_\eta\right\|_F>\left(\pi_2\delta_T T\right)^{1/2}\right)\\
&=\Pr\left(\left\|\mathbf{Q}'\boldsymbol{u}_\eta\right\|_F^2>\pi_2\delta_T T\right)\\
&\leq\sum_{j=1}^{l_T}\Pr\left[\left(\sum_{t=1}^{T}q_{jt}u_{\eta,t}\right)^2>\frac{\pi_2\delta_T T}{l_T}\right]\\
&=\sum_{j=1}^{l_T}\Pr\left[\left|\sum_{t=1}^{T}q_{jt}u_{\eta,t}\right|>\left(\frac{\pi_2\delta_T T}{l_T}\right)^{1/2}\right],
\end{aligned}
$$

[7] The required probability bound on u_{xt} follows from the probability bound assumptions on x_t and on q_{it}, for $i=1,2,...,l_T$, even if $l_T\to\infty$. See also Lemma A5.

which upon using (B.22) yields (for some $C, c > 0$)

$$\Pr\left(\|\mathbf{Q}'\boldsymbol{u}_\eta\|_F > (\pi_2\delta_T T)^{1/2}\right) \leq l_T \exp\left(-CT^c\right), \ \Pr\left(\|\mathbf{Q}'\boldsymbol{u}_x\| > (\pi_2\delta_T T)^{1/2}\right) = 0.$$

Further, it is easy to see that

$$\Pr\left(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F > \frac{\pi_2\zeta_T}{\delta_T}\right) = 0$$

as long as $\frac{\zeta_T}{\delta_T l_T^{1/2}} \to \infty$. But as long as $l_T = o\left(T^{1/3}\right)$, there exists a sequence δ_T such that $\zeta_T/\delta_T \to \infty$, $l_T = o\left(\delta_T\right)$ and $\frac{\zeta_T}{\delta_T l_T^{1/2}} \to \infty$ as required, establishing the required result. ■

Lemma E2 *Let y_t, for $t = 1, 2, ..., T$, be given by the data generating process (6) and suppose that $\boldsymbol{x}_{nt} = (x_{1t}, x_{2t}, ..., x_{nt})'$ are bounded deterministic sequences, and u_t satisfy Assumption 2 and condition (10) of Assumption 4. Let $\boldsymbol{q}_{\cdot t} = (q_{1,t}, q_{2,t}, ..., q_{l_T,t})'$ contain a constant and a subset of $\boldsymbol{x}_{nt} = (x_{1t}, x_{2t}, ..., x_{nt})'$, and let $\eta_t = \boldsymbol{x}_{b,t}\boldsymbol{\beta}_b + u_t$, where $\boldsymbol{x}_{b,t}$ is $k_b \times 1$ dimensional vector of signal variables that do not belong to $\boldsymbol{q}_{\cdot t}$. Assume that $\boldsymbol{\Sigma}_{qq} = \mathbf{Q}'\mathbf{Q}/T$ is invertible for all T, and $\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F = O\left(\sqrt{l_T}\right)$, where $\mathbf{Q} = (\boldsymbol{q}_{1\cdot}, \boldsymbol{q}_{2\cdot}, ..., \boldsymbol{q}_{l_T\cdot})$ and $\boldsymbol{q}_{i\cdot} = (q_{i1}, q_{i2}, ..., q_{iT})'$, for $i = 1, 2, ..., l_T$. Moreover, let $l_T = o(T^{1/4})$ and suppose that Assumption 5 holds for x_{it} and $\boldsymbol{q}_{\cdot t}$, and u_t and $\boldsymbol{q}_{\cdot t}$. Define $\boldsymbol{x} = (x_1, x_2, ..., x_T)'$, $\boldsymbol{y} = (y_1, y_2, ..., y_T)'$, $\mathbf{M}_q = \mathbf{I}_T - \mathbf{Q}(\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}'$, and $\theta = T^{-1}\boldsymbol{x}'\mathbf{M}_q\mathbf{X}_b\boldsymbol{\beta}_b$, where $\mathbf{X}_b$ is $T \times k_b$ matrix of observations on $\boldsymbol{x}_{b,t}$. Let $u_{x_i,T}$ be as in (11), such that $\sup_{i,j}\lim_{T\to\infty}\frac{\left\|\mathbf{q}_i'\boldsymbol{u}_{x_j,T}\right\|}{T^{1/2}} < C < \infty$. Let $\boldsymbol{e} = (e_1, e_2, ..., e_T)'$ be the $T \times 1$ vector of residuals in the linear regression model of y_t on $\boldsymbol{q}_{\cdot t}$ and x_t. Then, for any π in the range $0 < \pi < 1$, $d_T > 0$ and bounded in T, and for some $C_i > 0$ for $i = 0, 1$,*

$$\Pr\left[|t_x| > c_p\left(n, \delta\right)|\theta = 0\right] \leq \exp\left[\frac{-\left(1-\pi\right)^2\sigma_{u,(T)}^2\sigma_{x,(T)}^2 c_p^2\left(n,\delta\right)}{2\left(1+d_T\right)^2\omega_{xu,T}^2}\right] + \exp\left(-C_0T^{C_1}\right),$$

where

$$t_x = \frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sqrt{\left(\boldsymbol{e}'\boldsymbol{e}/T\right)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}},$$

$\sigma_{u,(T)}^2$ and $\sigma_{x,(T)}^2$ are defined by (B.39) and (B.34), and

$$\omega_{xu,T}^2 = \frac{1}{T}\sum_{t=1}^{T}\sigma_{xt}^2\sigma_t^2,$$

Under $\sigma_t^2 = \sigma^2$ and/or $\sigma_{xt}^2 = \sigma_x^2$ for all $t = 1, 2, ..., T$,

$$\Pr\left[|t_x| > c_p\left(n, \delta\right)|\theta = 0\right] \leq \exp\left[\frac{-\left(1-\pi\right)^2 c_p^2\left(n,\delta\right)}{2\left(1+d_T\right)^2}\right] + \exp\left(-C_0T^{C_1}\right).$$

for some $C_0, C_1 > 0$. In the case where $\theta > 0$, and assuming that $c_p(n,\delta) = o(\sqrt{T})$, then for $d_T > 0$ and some $C_i > 0$, $i = 0, 1$, we have

$$\Pr\left[|t_x| > c_p(n,\delta)\,|\theta \neq 0\right] > 1 - \exp\left(-C_0 T^{C_1}\right).$$

Proof. The model for $\boldsymbol{y}$ can be written as

$$\boldsymbol{y} = a\boldsymbol{\tau}_T + \mathbf{X}_k\boldsymbol{\beta} + \boldsymbol{u} = a\boldsymbol{\tau}_T + \mathbf{X}_a\boldsymbol{\beta}_a + \mathbf{X}_b\boldsymbol{\beta}_b + \boldsymbol{u}$$

where $\boldsymbol{\tau}_T$ is a $T\times 1$ vector of ones, $\mathbf{X}_a$ is a subset of $\mathbf{Q}$. Let $\mathbf{Q}_x = (\mathbf{Q}, \boldsymbol{x})$, $\mathbf{M}_q = \mathbf{I}_T - \mathbf{Q}(\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}'$, $\mathbf{M}_{qx} = \mathbf{I}_T - \mathbf{Q}_x(\mathbf{Q}_x'\mathbf{Q}_x)^{-1}\mathbf{Q}_x'$. Then, $\mathbf{M}_q\mathbf{X}_a = \mathbf{0}$. $\mathbf{M}_q\mathbf{X}_b = (\boldsymbol{x}_{bq,1}, \boldsymbol{x}_{bq,2}, ..., \boldsymbol{x}_{bq,T})'$. Then,

$$t_x = \frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}} = \frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\mathbf{X}_b\boldsymbol{\beta}_b}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}} + \frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{u}}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}}.$$

Let

$$\boldsymbol{\eta} = \mathbf{X}_b\boldsymbol{\beta}_b + \boldsymbol{u}, \quad \boldsymbol{\eta} = (\eta_1, \eta_2, ..., \eta_T)'$$

$$\theta = T^{-1}\boldsymbol{x}'\mathbf{M}_q\mathbf{X}_b\boldsymbol{\beta}_b,$$
$$\sigma^2_{e,(T)} = E(\boldsymbol{e}'\boldsymbol{e}/T) = E\left(\frac{\boldsymbol{\eta}'\mathbf{M}_{qx}\boldsymbol{\eta}}{T}\right), \ \sigma^2_{x,(T)} = E\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right),$$

and write (B.53) as

$$t_x = \frac{\sqrt{T}\theta}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}} + \frac{T^{-1/2}\left[\boldsymbol{x}'\mathbf{M}_q\,\boldsymbol{\eta} - E(\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta})\right]}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}}.$$

$$\boldsymbol{x}'\mathbf{M}_q\,\boldsymbol{\eta} - E(\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}) = \left[\boldsymbol{x}'\mathbf{M}_q\,\boldsymbol{u} - E(\boldsymbol{x}'\mathbf{M}_q\boldsymbol{u})\right],$$

$$\frac{(\mathbf{M}_q\mathbf{X}_b\boldsymbol{\beta}_b)'(\mathbf{M}_q\mathbf{X}_b\boldsymbol{\beta}_b)}{T} = \frac{1}{T}\sum_{t=1}^{T}\left(\boldsymbol{x}'_{bq,1}\boldsymbol{\beta}_b\right)^2 = \frac{1}{T}\sum_{t=1}^{T}\sigma^2_{xbt} = \sigma^2_{b,(T)}.$$

Then, we consider two cases: $\frac{\boldsymbol{x}'\mathbf{M}_q\mathbf{X}_b\boldsymbol{\beta}_b}{T} := \theta = 0$ and $\theta \neq 0$. We consider each in turn. First, we consider $\theta = 0$ and note that

$$t_x = \frac{T^{-1/2}\left[\boldsymbol{x}'\mathbf{M}_q\,\boldsymbol{u} - E(\boldsymbol{x}'\mathbf{M}_q\boldsymbol{u})\right]}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}}.$$

By Lemma A9, we have

$$\Pr\left[|t_x| > c_p(n,\delta)\,|\theta = 0\right] = \Pr\left[\left|\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}}\right| > c_p(n,\delta)\,|\theta = 0\right] \leq$$
$$\Pr\left(\left|\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}}{\sigma_{x,(T)}\sigma_{e,(T)}}\right| > \frac{c_p(n,\delta)}{1+d_T}\right) + \exp\left(-C_0T^{C_1}\right).$$

By Lemma E1, it then follows that,

$$\Pr\left[|t_x| > c_p(n,\delta)\,|\theta=0\right] \leq \exp\left[\frac{-(1-\pi)^2\sigma^2_{e,(T)}\sigma^2_{x,(T)}c_p^2(n,\delta)}{2(1+d_T)^2\omega^2_{xe,T}}\right] + \exp\left(-C_0T^{C_1}\right)$$

where $\omega^2_{xe,T} = \frac{1}{T}\sum_{t=1}^{T} E\left[(u_{x,t}\eta_t)^2\right]$. Note that, by independence of u_t with $u_{x,t}$ and $\boldsymbol{x}_{bq,t}$ we have

$$\omega^2_{xe,T} = \frac{1}{T}\sum_{t=1}^{T} E\left[(u_{x,t}\eta_t)^2\right] = \frac{1}{T}\sum_{t=1}^{T} E\left[u^2_{x,t}\left(\boldsymbol{x}'_{bq,1}\boldsymbol{\beta}_b\right)^2\right] + E\left(u^2_{xt}\right)E\left(u^2_t\right).$$

By the deterministic nature of x_{it}, and under homoscedasticity for η_t, it follows that $\sigma^2_{e,(T)}\sigma^2_{x,(T)} = \omega^2_{xe,T}$, and so

$$\Pr\left[|t_x| > c_p(n,\delta)\,|\theta=0\right] \leq \exp\left[\frac{-(1-\pi)^2 c_p^2(n,\delta)}{2(1+d_T)^2}\right] + \exp\left(-C_0T^{C_1}\right).$$

giving a rate that does not depend on variances. Next, we consider $\theta \neq 0$. By Lemma A9, for $d_T > 0$,

$$\Pr\left[\left|\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sqrt{(\boldsymbol{e}'\boldsymbol{e}/T)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}}\right| > c_p(n,\delta)\right] \leq \Pr\left(\left|\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| > \frac{c_p(n,\delta)}{1+d_T}\right) + \exp\left(-C_0T^{C_1}\right).$$

We then have

$$\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sigma_{e,(T)}\sigma_{x,(T)}} = \frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{u}}{\sigma_{e,(T)}\sigma_{x,(T)}} + \frac{T^{1/2}\theta}{\sigma_{e,(T)}\sigma_{x,(T)}}$$

Then,

$$\Pr\left(\left|\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{u}}{\sigma_{e,(T)}\sigma_{x,(T)}} + \frac{T^{1/2}\theta}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| > \frac{c_p(n,\delta)}{1+d_T}\right) = 1 - \Pr\left(\left|\frac{T^{1/2}T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{u}}{\sigma_{e,(T)}\sigma_{x,(T)}} + \frac{T^{1/2}\theta}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| \leq \frac{c_p(n,\delta)}{1+d_T}\right).$$

We note that

$$\Pr\left(\left|\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{u}}{\sigma_{e,(T)}\sigma_{x,(T)}} + \frac{T^{1/2}\theta}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| \leq \frac{c_p(n,\delta)}{1+d_T}\right) \leq \Pr\left(\left|\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{u}}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| > \frac{T^{1/2}|\theta|}{\sigma_{e,(T)}\sigma_{x,(T)}} - \frac{c_p(n,\delta)}{1+d_T}\right).$$

But $T^{-1}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{u}$ is the average of a martingale difference process and so

$$\Pr\left(\left|\frac{T^{1/2}\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{u}}{T}\right)}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| > \frac{T^{1/2}\left|\theta\right|}{\sigma_{e,(T)}\sigma_{x,(T)}} - \frac{c_p\left(n,\delta\right)}{1+d_T}\right)$$
$$\leq \exp\left(-C_0T^{C_1}\right) + \exp\left[-C\left(T^{1/2}\left(\frac{T^{1/2}\left|\theta\right|}{\sigma_{e,(T)}\sigma_{x,(T)}} - \frac{c_p\left(n,\delta\right)}{1+d_T}\right)\right)^{s/(s+2)}\right].$$

So overall,

$$\Pr\left[\left|\frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\boldsymbol{y}}{\sqrt{\left(\boldsymbol{e}'\boldsymbol{e}/T\right)\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{x}}{T}\right)}}\right| > c_p\left(n,\delta\right)\right] > 1 - \exp\left(-C_0T^{C_1}\right)$$
$$- \exp\left[-C\left(T^{1/2}\left(\frac{T^{1/2}\left|\theta\right|}{\sigma_{e,(T)}\sigma_{x,(T)}} - \frac{c_p\left(n,\delta\right)}{1+d_T}\right)\right)^{s/(s+2)}\right].$$

■

F. Supplementary lemmas for Sections B and C of the online theory supplement

Lemma F1 *Suppose that u_t, $t = 1, 2, ..., T$, is a martingale difference process with respect to $\mathcal{F}_{t-1}^u$ and with constant variance σ^2, and there exist constants $C_0, C_1 > 0$ and $s > 0$ such that $\Pr\left(|u_t| > \alpha\right) \leq C_0\exp\left(-C_1\alpha^s\right)$, for all $\alpha > 0$. Let $\boldsymbol{X}_T = (\boldsymbol{x}_{l_T,1}, \boldsymbol{x}_{l_T,2}, , .., \boldsymbol{x}_{l_T,T})$, where $\boldsymbol{x}_{l_T,t}$ is an $l_T \times 1$ dimensional vector of random variables, with probability measure given by $P(\boldsymbol{X}_T)$, and assume*

$$E\left(u_t\left|\mathcal{F}_T^x\right.\right) = 0, \text{ for all } t = 1, 2, ..., T, \tag{B.147}$$

where $\mathcal{F}_T^x = \sigma\left(\boldsymbol{x}_{l_T,1}, \boldsymbol{x}_{l_T,2}, , .., \boldsymbol{x}_{l_T,T}\right)$. Further assume that there exist functions $\boldsymbol{\alpha}(\boldsymbol{X}_T) = [\alpha_1(\boldsymbol{X}_T), \alpha_2(\boldsymbol{X}_T), ..., \alpha_T(\boldsymbol{X}_T)]'$ such that $0 < \sup_{\boldsymbol{X}_T}\boldsymbol{\alpha}(\boldsymbol{X}_T)'\boldsymbol{\alpha}(\boldsymbol{X}_T) \leq g_T$, for some sequence $g_T > 0$. Then,

$$\Pr\left(\left|\sum_{t=1}^T\alpha_t(\boldsymbol{X}_T)u_t\right| > \zeta_T\right) \leq \exp\left(\frac{-\zeta_T^2}{2g_T\sigma^2}\right).$$

Proof. Define $\mathcal{A}_T = \left\{\left|\sum_{t=1}^T\alpha_t(\boldsymbol{X}_T)u_t\right| > \zeta_T\right\}$. Then,

$$\Pr\left(\mathcal{A}_T\right) = \int_{\boldsymbol{X}_T}\Pr\left(\mathcal{A}_T\left|\mathcal{F}_T^x\right.\right)P(\boldsymbol{X}_T) \leq \sup_{\boldsymbol{X}_T}\Pr\left(\mathcal{A}_T\left|\mathcal{F}_T^x\right.\right)\int_{\boldsymbol{X}_T}P(\boldsymbol{X}_T) = \sup_{\boldsymbol{X}_T}\Pr\left(\mathcal{A}_T\left|\mathcal{F}_T^x\right.\right)$$

■

But, by (B.147) and Lemma A3

$$\Pr\left(\mathcal{A}_T\left|\mathcal{F}_T^x\right.\right) \leq \exp\left(\frac{-\zeta_T^2}{2\sigma^2\sum_{t=1}^T\alpha_t^2(\boldsymbol{X}_T)}\right)$$

But

$$\sup_{\boldsymbol{X}_T} \exp\left(\frac{-\zeta_T^2}{2\sigma^2 \sum_{t=1}^T \alpha_t^2(\boldsymbol{X}_T)}\right) \leq \exp\left(\frac{-\zeta_T^2}{2g_T\sigma^2}\right),$$

proving the result.

Lemma F2 *Suppose that u_t, $t = 1, 2, ..., T$, is a zero mean mixing random variable with exponential mixing coefficients given by $\phi_k = a_{0k}\varphi^k$, $0 < \varphi < 1$, $a_{0k} < \infty$, $k = 1, ...,$ with constant variance σ^2, and there exist sufficiently large constants $C_0, C_1 > 0$ and $s > 0$ such that $\Pr(|u_t| > \alpha) \leq C_0 \exp(-C_1\alpha^s)$, for all $\alpha > 0$. Let $\boldsymbol{X}_T = (\boldsymbol{x}_{l_T,1}, \boldsymbol{x}_{l_T,2}, , .., \boldsymbol{x}_{l_T,T})$, where $\boldsymbol{x}_{l_T,t}$ is an $l_T \times 1$ dimensional vector of random variables, with probability measure given by $P(\boldsymbol{X}_T)$. Further assume that there exist functions $\boldsymbol{\alpha}(\boldsymbol{X}_T) = [\alpha_1(\boldsymbol{X}_T), \alpha_2(\boldsymbol{X}_T), ..., \alpha_T(\boldsymbol{X}_T)]'$ such that $0 < \sup_{\boldsymbol{X}_T} \boldsymbol{\alpha}(\boldsymbol{X}_T)'\boldsymbol{\alpha}(\boldsymbol{X}_T) \leq g_T$, for some sequence $g_T > 0$. Then,*

$$\Pr\left(\left|\sum_{t=1}^T \alpha_t(\boldsymbol{X}_T)u_t\right| > \zeta_T\right) \leq \exp\left(-\left(\frac{\zeta_T}{g_T^{1/2}\sigma}\right)^{s/(s+1)}\right).$$

Proof. Define $\mathcal{A}_T = \left\{\left|\sum_{t=1}^T \alpha_t(\boldsymbol{X}_T)u_t\right| > \zeta_T\right\}$ and consider $\mathcal{F}_T^x = \sigma\left(\boldsymbol{x}_{l_T,1}, \boldsymbol{x}_{l_T,2}, , .., \boldsymbol{x}_{l_T,T}\right)$. Then,

$$\Pr(\mathcal{A}_T) = \int_{\boldsymbol{X}_T} \Pr\left(\mathcal{A}_T \left|\mathcal{F}_T^x\right.\right) P(\boldsymbol{X}_T) \leq \sup_{\boldsymbol{X}_T} \Pr\left(\mathcal{A}_T \left|\mathcal{F}_T^x\right.\right) \int_{\boldsymbol{X}_T} P(\boldsymbol{X}_T) = \sup_{\boldsymbol{X}_T} \Pr\left(\mathcal{A}_T \left|\mathcal{F}_T^x\right.\right)$$

But, using Lemma 2 of Dendramis et al. (2015) we can choose C_0, C_1 such that

$$\Pr\left(\mathcal{A}_T \left|\mathcal{F}_T^x\right.\right) \leq \exp\left[-\left(\frac{-\zeta_T}{\sigma\sqrt{\sum_{t=1}^T \alpha_t^2(\boldsymbol{X}_T)}}\right)^{s/(s+1)}\right],$$

and

$$\sup_{\boldsymbol{X}_T} \exp\left[-\left(\frac{-\zeta_T}{\sigma\sqrt{\sum_{t=1}^T \alpha_t^2(\boldsymbol{X}_T)}}\right)^{s/(s+1)}\right] \leq \exp\left[-\left(\frac{\zeta_T}{g_T^{1/2}\sigma}\right)^{s/(s+1)}\right],$$

thus establishing the desired result. ■

Lemma F3 *Let $\boldsymbol{A}_T = (a_{ij,T})$ be a $l_T \times l_T$ matrix and $\hat{\boldsymbol{A}}_T = (\hat{a}_{ij,T})$ be an estimator of $\boldsymbol{A}_T$. Let $\left\|\boldsymbol{A}_T^{-1}\right\|_F > 0$ and suppose that for some $s > 0$, any $b_T > 0$ and $C_0 > 0$*

$$\sup_{i,j} \Pr\left(|\hat{a}_{ij,T} - a_{ij,T}| > b_T\right) \leq \exp\left(-C_0\left(T^{1/2}b_T\right)^{s/(s+2)}\right).$$

Then

$$\Pr\left(\left\|\hat{\boldsymbol{A}}_T^{-1}-\boldsymbol{A}_T^{-1}\right\|>b_T\right)\leq l_T^2\exp\left(\frac{-C_0\left(T^{1/2}b_T\right)^{s/(s+2)}}{l_T^{s/(s+2)}\left\|\boldsymbol{A}_T^{-1}\right\|_F^{s/(s+2)}\left(\left\|\boldsymbol{A}_T^{-1}\right\|_F+b_T\right)^{s/(s+2)}}\right) \tag{B.148}$$

$$+\,l_T^2\exp\left(-C_0\frac{T^{s/2(s+2)}}{\left\|\boldsymbol{A}_T^{-1}\right\|_F^{s/(s+2)}l_T^{s/(s+2)}}\right),$$

where $\|\mathbf{A}\|$ *denotes the Frobenius norm of* $\mathbf{A}$*.*

Proof. First note that since $b_T>0$, then

$$\Pr\left(\left\|\hat{\boldsymbol{A}}_T-\boldsymbol{A}_T\right\|_F>b_T\right)=\Pr\left(\left\|\hat{\boldsymbol{A}}_T-\boldsymbol{A}_T\right\|_F^2>b_T^2\right)$$
$$=\Pr\left(\left[\sum_{j=1}^{l_T}\sum_{i=1}^{l_T}\left(\hat{a}_{ij,T}-a_{ij,T}\right)^2>b_T^2\right]\right),$$

and using the probability bound result, (B.59), and setting $\pi_i=1/l_T$, we have

$$\Pr\left(\left\|\hat{\boldsymbol{A}}_T-\boldsymbol{A}_T\right\|_F>b_T\right)\leq\sum_{j=1}^{l_T}\sum_{i=1}^{l_T}\Pr\left(\left|\hat{a}_{ij,T}-a_{ij,T}\right|^2>l_t^{-2}b_T^2\right) \tag{B.149}$$
$$=\sum_{j=1}^{l_T}\sum_{i=1}^{l_T}\Pr\left(\left|\hat{a}_{ij,T}-a_{ij,T}\right|>l_t^{-1}b_T\right)$$
$$\leq l_T^2\sup_{ij}\left[\Pr\left(\left|\hat{a}_{ij,T}-a_{ij,T}\right|>l_t^{-1}b_T\right)\right]=l_T^2\exp\left(-C_0T^{s/2(s+1)}\frac{b_T^{s/(s+2)}}{l_t^{s/(s+2)}}\right).$$

To establish (B.148) define the events

$$\mathcal{A}_T=\left\{\left\|\boldsymbol{A}_T^{-1}\right\|_F\left\|\hat{\boldsymbol{A}}_T-\boldsymbol{A}_T\right\|_F<1\right\}\text{ and }\mathcal{B}_T=\left\{\left\|\hat{\boldsymbol{A}}_T^{-1}-\boldsymbol{A}_T^{-1}\right\|>b_T\right\}$$

and note that by (2.15) of Berk (1974) if $\mathcal{A}_T$ holds we have

$$\left\|\hat{\boldsymbol{A}}_T^{-1}-\boldsymbol{A}_T^{-1}\right\|\leq\frac{\left\|\boldsymbol{A}_T^{-1}\right\|_F^2\left\|\hat{\boldsymbol{A}}_T-\boldsymbol{A}_T\right\|_F}{1-\left\|\boldsymbol{A}_T^{-1}\right\|_F\left\|\hat{\boldsymbol{A}}_T-\boldsymbol{A}_T\right\|_F}.$$

Hence

$$\Pr\left(\mathcal{B}_T\,|\mathcal{A}_T\right)\leq\Pr\left(\frac{\left\|\boldsymbol{A}_T^{-1}\right\|_F^2\left\|\hat{\boldsymbol{A}}_T-\boldsymbol{A}_T\right\|_F}{1-\left\|\boldsymbol{A}_T^{-1}\right\|_F\left\|\hat{\boldsymbol{A}}_T-\boldsymbol{A}_T\right\|_F}>b_T\right)$$
$$=\Pr\left(\left\|\hat{\boldsymbol{A}}_T-\boldsymbol{A}_T\right\|_F>\frac{b_T}{\left\|\boldsymbol{A}_T^{-1}\right\|_F\left(\left\|\boldsymbol{A}_T^{-1}\right\|_F+b_T\right)}\right).$$

Note also that

$$\Pr(\mathcal{B}_T) = \Pr\left(\{\mathcal{B}_T \cap \mathcal{A}_T\} \cup \left\{\mathcal{B}_T \cap \mathcal{A}_T^C\right\}\right) = \Pr(\mathcal{B}_T|\mathcal{A}_T)\Pr(\mathcal{A}_T) + \Pr\left(\mathcal{B}_T|\mathcal{A}_T^C\right)\Pr\left(\mathcal{A}_T^C\right).$$

Furthermore

$$\begin{aligned}\Pr\left(\mathcal{A}_T^C\right) &= \Pr\left(\left\|\boldsymbol{A}_T^{-1}\right\|_F \left\|\hat{\mathbf{A}}_T - \boldsymbol{A}_T\right\|_F > 1\right) \\ &= \Pr\left(\left\|\hat{\mathbf{A}}_T - \boldsymbol{A}_T\right\|_F > \left\|\boldsymbol{A}_T^{-1}\right\|_F^{-1}\right),\end{aligned}$$

and by (B.149) we have

$$\Pr\left(\mathcal{A}_T^C\right) \leq l_T^2 \exp\left(-C_0 T^{s/2(s+2)} \frac{b_T^{s/(s+2)}}{l_t^{s/(s+2)}}\right) = \exp\left(-C_0 \frac{T^{s/2(s+2)}}{\left\|\boldsymbol{A}_T^{-1}\right\|_F^{s/(s+2)} l_T^{s/(s+2)}}\right).$$

Using the above result, we now have

$$\begin{aligned}\Pr(\mathcal{B}_T) \leq{}& \Pr\left(\left\|\hat{\mathbf{A}}_T - \boldsymbol{A}_T\right\|_F > \frac{b_T}{\left\|\boldsymbol{A}_T^{-1}\right\|_F \left(\left\|\boldsymbol{A}_T^{-1}\right\|_F + b_T\right)}\right)\Pr(\mathcal{A}_T) \\ &+ \Pr\left(\mathcal{B}_T|\mathcal{A}_T^C\right)\exp\left(-C_0 \frac{T^{s/2(s+2)}}{\left\|\boldsymbol{A}_T^{-1}\right\|_F^{s/(s+2)} l_T^{s/(s+2)}}\right).\end{aligned}$$

Furthermore, since $\Pr(\mathcal{A}_T) \leq 1$ and $\Pr\left(\mathcal{B}_T|\mathcal{A}_T^C\right) \leq 1$ then

$$\begin{aligned}\Pr(\mathcal{B}_T) = \Pr\left(\left\|\hat{\mathbf{A}}_T^{-1} - \boldsymbol{A}_T^{-1}\right\| > b_T\right) \leq{}& \Pr\left(\left\|\hat{\mathbf{A}}_T - \boldsymbol{A}_T\right\|_F > \frac{b_T}{\left\|\boldsymbol{A}_T^{-1}\right\|_F \left(\left\|\boldsymbol{A}_T^{-1}\right\|_F + b_T\right)}\right) \\ &+ \exp\left(-C_0 \frac{T^{s/2(s+2)}}{\left\|\boldsymbol{A}_T^{-1}\right\|_F^{s/(s+2)} l_T^{s/(s+2)}}\right).\end{aligned}$$

Result (B.148) now follows if we apply (B.149) to the first term on the RHS of the above. ■

Lemma F4 *Consider the scalar random variable X_T, and the constants B and C. Then, if $C > |B| > 0$,*

$$\Pr(|X + B| > C) \leq \Pr(|X| > C - |B|). \tag{B.150}$$

Proof. The result follows by noting that $|X + B| \leq |X| + |B|$. ■

References

Bailey, N., M. H. Pesaran, and L. V. Smith (2018). A multiple testing approach to the regularisation of large sample correlation matrices. *CAFE Research Paper No. 14.05, revised 2018*.

Berk, K. N. (1974). Consistent autoregressive spectral estimates. *Annals of Statistics 2*, 489–502.

Dendramis, Y., L. Giraitis, and G. Kapetanios (2015). Estimation of random coefficient time varying covariance matrices for large datasets. *Mimeo*.

Freedman, D. A. (1975). On tail probabilities for martingales. *Annals of Probability 3*, 100–118.

Roussas, G. (1996). Exponential probability inequalities with some applications. *Statistica, Probability and Game Theory, IMS Lecture Notes - Monograph Series 30*, 303–319.

Wecker, W. E. (1978). A note on the time series which is the product of two stationary time series. *Stochastic Processes and their Applications 8*, 153–157.

White, H. and J. M. Wooldridge (1991). Some results on sieve estimation with dependent observations. In W. J. Barnett, J. Powell, and G. Tauchen (Eds.), *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, pp. 459–493. New York: Cambridge University Press.