Online Theoretical Supplement

for Paper

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

by A. Chudik, G. Kapetanios and M. Hashem Pesaran

This online Supplement is organised as follows: Section A provides supplementary 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. Section D presents lemmas related to mixing regressors. 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. Supplementary lemmas

Lemma A1 *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*

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

$$\Pr\left(|X_T| \times |Y_T| > C_0\right) \leq \Pr\left(|X_T| > C_0/C_1\right) + \Pr\left(|Y_T| > C_1\right), \tag{B.2}$$

and

$$\Pr\left(|X_T| \times |Y_T| \times |Z_T| > C_0\right) \leq \Pr\left(|X_T| > C_0/\left(C_1 C_2\right)\right) + \Pr\left(|Y_T| > C_1\right) + \Pr\left(|Z_T| > C_2\right). \tag{B.3}$$

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. Also

$$\begin{aligned}\Pr\left(|X_T| \times |Y_T| > C_0\right) = &\Pr\left(|X_T| \times |Y_T| > C_0 \;|\{|Y_T| > C_1\}\right)\Pr\left(|Y_T| > C_1\right) + \\ &\Pr\left(|X_T| \times |Y_T| > C_0 \;|\{|Y_T| \leq C_1\}\right)\Pr\left(|Y_T| \leq C_1\right),\end{aligned}$$

and since

$$\Pr\left(|X_T| \times |Y_T| > C_0 \ \ |\{|Y_T| > C_1\}\right) \leq \Pr\left(|X_T| > C_0/C_1\right),$$

and

$$0 \leq \Pr\left(|X_T| \times |Y_T| > C_0 \ \ |\{|Y_T| \leq C_1\}\right) \leq 1,$$

then

$$\Pr\left(|X_T| \times |Y_T| > C_0\right) \leq \Pr\left(|X_T| > C_0/C_1\right) + \Pr\left(|Y_T| > C_1\right),$$

proving the second result of the lemma. The third result follows by a repeated application of the second result. ■

Lemma A2 *Consider the scalar random variable X_T, and the constants B and C. Then, if $|B| \geq C > 0$,*

$$\Pr\left(|X + B| \leq C\right) \leq \Pr\left(|X| > |B| - C\right). \tag{B.4}$$

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.4). ■

Lemma A3 *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*

$$\Pr\left(\left|\frac{1}{\sqrt{\omega_T}} - 1\right| > \alpha_T\right) \leq \Pr\left(|\omega_T - 1| > \alpha_T\right). \tag{B.5}$$

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 $\Pr\left(\left|\frac{1}{\sqrt{\omega_T}} - 1\right| > \alpha_T\right) \leq \Pr\left(|\omega_T - 1| > \alpha_T\right)$, as required. ■

Lemma A4 *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,*

$$\left\|\mathbf{A}_T\right\|_F = \ominus(l_T). \tag{B.6}$$

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

$$\left\|\mathbf{A}_T^{-1}\right\|_F = \ominus\left(\sqrt{l_T}\right). \tag{B.7}$$

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 $\sup_{1 \le i \le l_T - M} \mu_i < C_0 < \infty$, 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.6) 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.7). ■

Lemma A5 *Let z be a random variable and suppose there exists finite positive constants C_0, C_1 and $s > 0$ such that*

$$\Pr\left(|z| > \alpha\right) \le C_0 \exp\left(-C_1 \alpha^s\right), \text{ for all } \alpha > 0. \tag{B.8}$$

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

$$E\left|z\right|^p \le C_2. \tag{B.9}$$

Proof. We have that

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

Using integration by parts, we get

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

But, using (B.8), and a change of variables, implies

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

where $\Gamma(\cdot)$ is a gamma function. But for a finite positive p/s, $\Gamma(p/s)$ is bounded and (B.9) follows. ■

Lemma A6 *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*

$$\sup_{i,j} \Pr\left(|\hat{a}_{ij,T} - a_{ij,T}| > b_T\right) \le \exp\left(-C_0 T b_T^2\right), \tag{B.10}$$

for all $b_T > 0$. Then

$$\Pr\left(\left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F > b_T\right) \le l_T^2 \exp\left(-C_0 \frac{T b_T^2}{l_T^2}\right), \tag{B.11}$$

and

$$\Pr\left(\left\|\hat{\boldsymbol{A}}_T^{-1} - \boldsymbol{A}_T^{-1}\right\|_F > b_T\right) \le l_T^2 \exp\left(\frac{-C_0 T b_T^2}{l_T^2 \left\|\boldsymbol{A}_T^{-1}\right\|_F^2 \left(\left\|\boldsymbol{A}_T^{-1}\right\|_F + b_T\right)^2}\right)$$
$$+ l_T^2 \exp\left(-C_0 \frac{T}{\left\|\boldsymbol{A}_T^{-1}\right\|_F^2 l_T^2}\right). \tag{B.12}$$

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.1), and setting $\pi_i = 1/l_T$, we have

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

Hence by (B.10) we obtain (B.11). To establish (B.12) define the sets

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

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

$$\left\|\hat{\boldsymbol{A}}_T^{-1} - \boldsymbol{A}_T^{-1}\right\|_F \le \frac{\left\|\boldsymbol{A}_T^{-1}\right\|_F^2 \left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F}{1 - \left\|\boldsymbol{A}_T^{-1}\right\|_F \left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F}. \tag{B.13}$$

Hence

$$\Pr\left(\mathcal{B}_T \left|\mathcal{A}_T\right.\right) \le \Pr\left(\frac{\left\|\boldsymbol{A}_T^{-1}\right\|_F^2 \left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F}{1 - \left\|\boldsymbol{A}_T^{-1}\right\|_F \left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F} > b_T\right)$$
$$= \Pr\left(\left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F > \frac{b_T}{\left\|\boldsymbol{A}_T^{-1}\right\|_F \left(\left\|\boldsymbol{A}_T^{-1}\right\|_F + b_T\right)}\right). \tag{B.14}$$

Note also that

$$\Pr\left(\mathcal{B}_T\right)=\Pr\left(\left\{\mathcal{B}_T\cap\mathcal{A}_T\right\}\cup\left\{\mathcal{B}_T\cap\mathcal{A}_T^C\right\}\right)=\Pr\left(\mathcal{B}_T|\mathcal{A}_T\right)\Pr\left(\mathcal{A}_T\right)+\Pr\left(\mathcal{B}_T|\mathcal{A}_T^C\right)\Pr\left(\mathcal{A}_T^C\right). \tag{B.15}$$

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.11) we have

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

Using the above result and (B.14) in (B.15), we now have

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

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

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

Result (B.12) now follows if we apply (B.11) to the first term on the RHS of the above. ■

Lemma A7 *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,*

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

Then

$$\begin{aligned}\Pr\left(\left\|\hat{\boldsymbol{A}}_T^{-1}-\boldsymbol{A}_T^{-1}\right\|_F>b_T\right)\le{}&l_T^2\exp\left(\frac{-C_0\left(Tb_T\right)^{s/(s+2)}}{l_T^{s/(s+2)}\left\|\boldsymbol{A}_T^{-1}\right\|_F^{s/(s+2)}\left(\left\|\boldsymbol{A}_T^{-1}\right\|_F+b_T\right)^{s/(s+2)}}\right)\\&+l_T^2\exp\left(-C_0\frac{T^{s/(s+2)}}{\left\|\boldsymbol{A}_T^{-1}\right\|_F^{s/(s+2)}l_T^{s/(s+2)}}\right).\end{aligned}\tag{B.16}$$

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.1), and setting $\pi_i = 1/l_T^2$, we have

$$\Pr\left(\left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F > b_T\right) \leq \sum_{j=1}^{l_T}\sum_{i=1}^{l_T}\Pr\left(\left|\hat{a}_{ij,T} - a_{ij,T}\right|^2 > l_T^{-2}b_T^2\right) \tag{B.17}$$
$$= \sum_{j=1}^{l_T}\sum_{i=1}^{l_T}\Pr\left(\left|\hat{a}_{ij,T} - a_{ij,T}\right| > l_T^{-1}b_T\right)$$
$$\leq l_T^2 \sup_{ij}\left[\Pr\left(\left|\hat{a}_{ij,T} - a_{ij,T}\right| > l_T^{-1}b_T\right)\right] = l_T^2\exp\left(-C_0 T^{s/(s+1)}\frac{b_T^{s/(s+2)}}{l_t^{s/(s+2)}}\right).$$

To establish (B.16) define the sets

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

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

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

Furthermore

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

and by (B.17) 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

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

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

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

Result (B.16) now follows if we apply (B.17) to the first term on the RHS of the above. ■

B. Proof of Theorem 3

We proceed as in the proof of (A.108) of Lemma 10. We have that

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

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 A2

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

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.57) 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 $c_p\left(n,\delta\right)\to\infty$, $\frac{T^{1/2}|\theta_i|}{\sigma_{e,(T)}\sigma_{x_i,(T)}} > \frac{c_p(n,\delta)}{1+d_T}$ implies $T^{1/2}\left|\theta_i\right| > C_2$, for some C_2. Then, noting that $\frac{\boldsymbol{x}_i'\mathbf{M}_q\boldsymbol{\eta}}{T}-\theta$ is the average of a martingale difference process, by Lemma 6, for some positive constants, C_1, C_2, C_3, C_4, C_5, and, for any $\psi > 0$, we have

$$\sum_{i=k+1}^{n}\Pr\left[\left|\frac{T^{-1/2}\boldsymbol{x}_i'\mathbf{M}_q\mathbf{y}}{\sqrt{\left(\mathbf{e}'\mathbf{e}/T\right)\left(\frac{\mathbf{x}_i'\mathbf{M}_q\mathbf{x}_i}{T}\right)}}\right|>c_p\left(n,\delta\right)\right]\leq C_1\sum_{i=k+1}^{n}I\left(\sqrt{T}\theta_i>C_2\right)$$
$$+C_3\sum_{i=k+1}^{n}I\left(\sqrt{T}\theta_i\leq C_4\right)\exp\left[-\ln(n)^{C_5}\right],$$
$$=C_1\sum_{i=k+1}^{n}I\left(\sqrt{T}\theta_i>C_2\right)+o(n^{1-\psi})+O\left[\exp(-CT^{C_5})\right], \tag{B.18}$$

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.18). Consider now the above probability bound under the two specifications assumed for θ_i as given by (38) and (39). Under (38), for any $\psi>0$,

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

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 (38), $\theta_i=C_i\varrho^i$, $|\varrho|<1$, and $c_p\left(n,\delta\right)=O\left[\ln(n)^{1/2}\right]$, 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 (39), $\theta_i = C_i i^{-\gamma}$, 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

$$E\left|FPR_{n,T}\right| = o(1), \tag{B.19}$$

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.19) 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 6.

As discussed in Section 4, 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.47)-(B.48) of Lemma D3. There, it is shown that, for sufficiently large constants $C_1 - C_4$ for Assumption 4, the martingale difference bound which is given by $\exp\left[-\frac{1}{2}\varkappa c_p^2(n)\right]$ in Lemma 6 is replaced by the bound $\exp\left[-c_p(n)^{s/(s+1)}\right]$ where s is the exponent in the probability tail in Assumption 4. It is important to note here that this bound seems to be sharp (see, e.g., Roussas (1996)) and so we need to understand its implications for our analysis. Given (see result (i) of Lemma 2),

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

it follows that

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

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

$$\exp\left\{-\left[\ln\left(\frac{f(n)}{2p}\right)\right]^{s/2(s+1)}\right\} = \exp\left[-n^{a_n s/2(s+1)}\right]$$

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 $C = s/2(s+1)$, we have $a_n = \frac{2(s+1)\ln(\ln(n))}{s\ln n}$, which leads to the following choice for $f(n)$

$$f(n) = 2p\exp\left(n^{\frac{2(s+1)\ln(\ln(n))}{s\ln n}}\right) \sim 2p\exp\left(\ln(n)^{\frac{2(s+1)}{s}}\right).$$

Then,

$$c_p(n) = O\left[\ln\left(\exp\left(\ln(n)^{\frac{2(s+1)}{s}}\right)\right)\right] = O\left(\ln(n)^{\frac{2(s+1)}{s}}\right),$$

which for $n = O\left(T^{C_1}\right)$, $C_1 > 0$, implies that $c_p(n) = O\left(\ln(T)^{\frac{2(s+1)}{s}}\right)$, 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 $\exp\left(\ln(n)^3\right)$ 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 $f(n) = 2p\exp\left(\ln(n)^2\right)$. 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}, $i = k + k^* + 1, k + k^* + 2, ..., n$. $E\left(x_{it}x_{jt} - E\left(x_{it}x_{jt}\right) \middle| \mathcal{F}_{t-1}^{xs}\right) = 0$, $i, j = 1, 2, ..., k + k^*$. x_{it}, $i = k + k^* + 1, k + k^* + 2, ..., n$, are heterogeneous strongly mixing processes with mixing coefficients given by $\alpha_{ik} = C_{ik}\xi^k$ for some C_{ik} such that $\sup_{i,k} C_{ik} < \infty$ and some $0 < \xi < 1$.*

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 $\alpha_{ik} = C_{ik}\xi^k$ for some C_{ik} such that $\sup_{i,k} C_{ik} < \infty$ and some $0 < \xi < 1$.*

Lemma D1 *Let ξ_t be a sequence of zero mean, mixing random variables with exponential mixing coefficients given by $\alpha_k = a_0\varphi^k$, $0 < \varphi < 1$. Assume, further, that $\Pr\left(|\xi_t| > \alpha\right) \leq C_0\exp\left[-C_1\alpha^s\right]$, $s \geq 1$. Then, for some $C > 0$, each $0 < \delta < 1$ and $v_T \geq \epsilon T^{\lambda}$, $\lambda > (1+\delta)/2$, ,*

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

Proof. We reconsider the proof of Theorem 3.5 of White and Wooldridge (1991). 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), we have that constants $C_1 - C_4$ in Assumption 4 can be chosen sufficiently large such that

$$\Pr\left(\left|\sum_{t=1}^{T} w_t - E(w_t)\right| > v_T\right) \leq \exp\left[\frac{-v_T T^{-(1+\delta)/2}}{D_T}\right] \tag{B.20}$$

rather than

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

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

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

Take logs and we have

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

or

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

For this it suffices that

$$\frac{2^s q v_T}{T^{(1+\delta)/2} D_T} \geq 2^p q \ln\left(v_T^{-1} T\right) \tag{B.21}$$

and

$$D_T^s \geq \frac{2^s q v_T}{T^{(1+\delta)/2} D_T}. \tag{B.22}$$

Set

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

so that (B.22) holds with equality. But since $v_T \geq \epsilon T^{\lambda}$, $\lambda > (1+\delta)/2$, (B.21) holds. Therefore,

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

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 mixing processes with exponentially declining tail probabilities. It is important for the rest of the lemmas in this Appendix, and in particular, the results obtained under (D4) of Lemma D2, to also note that Lemma 2 of Dendramis, Giraitis, and Kapetanios (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 the pairs x_t and $\boldsymbol{q}_{\cdot t}$, and denote the corresponding projection residuals defined by (20) 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 $T \times 1$ OLS 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 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*

$$\Pr\left(\left|\sum_{t=1}^{T} x_t u_t - E(x_t u_t)\right| > \zeta_T\right) \leq \exp\left[\frac{-(1-\pi)^2 \zeta_T^2}{2T\omega_{xu,1,T}^2}\right] + \exp\left[-C_0 T^{C_1}\right] \tag{B.23}$$

and

$$\Pr\left(\left|\sum_{t=1}^{T} \hat{u}_{x,t} u_t - \mu_{xu,t}\right| > \zeta_T\right) \leq \exp\left[\frac{-(1-\pi)^2 \zeta_T^2}{2T\omega_{xu,T}^2}\right] + \exp\left[-C_0 T^{C_1}\right], \tag{B.24}$$

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

$$\Pr\left(\left|\sum_{t=1}^{T} x_t u_t - E(x_t u_t)\right| > \zeta_T\right) \leq \exp\left[-C_0 T^{C_1}\right], \tag{B.25}$$

and

$$\Pr\left(\left|\sum_{t=1}^{T} \hat{u}_{x,t} u_t - \mu_{xu,t}\right| > \zeta_T\right) \leq \exp\left[-C_0 T^{C_1}\right], \tag{B.26}$$

as long as $l_T = o(T^{1/3})$. Finally, if (D4) holds,

$$\Pr\left(\left|\sum_{t=1}^{T} x_t u_t - E(x_t u_t)\right| > \zeta_T\right) \leq \exp\left[-C_0\left(\zeta_T T^{-1/2}\right)^{s/(s+2)}\right], \tag{B.27}$$

and

$$\Pr\left(\left|\sum_{t=1}^{T} \hat{u}_{x,t} u_t - \mu_{xu,t}\right| > \zeta_T\right) \leq \exp\left[-\left(\zeta_T T^{-1/2}\right)^{s/(s+2)}\right] + \exp\left[-C_0 T^{C_1}\right], \tag{B.28}$$

as long as $l_T = o(T^{1/3})$.

Proof. We first prove the lemma under (D1) and then modify arguments to show results 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

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

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 4,

$$\Pr\left(|\boldsymbol{u}_x'\boldsymbol{u}| > \zeta_T\right) \leq \exp\left[\frac{-(1-\pi)^2\zeta_T^2}{2T\omega_{xu,T}^2}\right]. \tag{B.29}$$

It therefore suffices to show that

$$\Pr\left(\left|\left(\frac{1}{T}\boldsymbol{u}_x'\mathbf{Q}\right)\left(\hat{\boldsymbol{\Sigma}}_{qq}^{-1} - \boldsymbol{\Sigma}_{qq}^{-1}\right)(\mathbf{Q}'\boldsymbol{u})\right| > \zeta_T\right) \leq \exp\left[-C_0 T^{C_1}\right] \tag{B.30}$$

and

$$\Pr\left(\left|\left(\frac{1}{T}\boldsymbol{u}_x'\mathbf{Q}\right)\boldsymbol{\Sigma}_{qq}^{-1}(\mathbf{Q}'\boldsymbol{u})\right| > \zeta_T\right) \leq \exp\left[-C_0 T^{C_1}\right] \tag{B.31}$$

We explore (B.29) and (B.30). We start with (B.29). We have by Lemma A1 that, for some sequence δ_T,[1]

$$\begin{aligned}
&\Pr\left(\left|\left(\frac{1}{T}\boldsymbol{u}_x'\mathbf{Q}\right)\left(\hat{\boldsymbol{\Sigma}}_{qq}^{-1} - \boldsymbol{\Sigma}_{qq}^{-1}\right)(\mathbf{Q}'\boldsymbol{u})\right| > \zeta_T\right) \leq \\
&\Pr\left(\left\|\frac{1}{T}\boldsymbol{u}_x'\mathbf{Q}\right\|\left\|\left(\hat{\boldsymbol{\Sigma}}_{qq}^{-1} - \boldsymbol{\Sigma}_{qq}^{-1}\right)\right\|\|\mathbf{Q}'\boldsymbol{u}\|_F > \zeta_T\right) \leq \Pr\left(\left\|\left(\hat{\boldsymbol{\Sigma}}_{qq}^{-1} - \boldsymbol{\Sigma}_{qq}^{-1}\right)\right\| > \frac{\zeta_T}{\delta_T}\right) + \\
&\Pr\left(\|\boldsymbol{u}_x'\mathbf{Q}\|_F\|\mathbf{Q}'\boldsymbol{u}\|_F > \delta_T T\right)
\end{aligned} \tag{B.33}$$

We consider the first term of the RHS of (B.33). Note that for all $1 \leq i, j \leq l$.

$$\Pr\left(\left|\frac{1}{T}\sum_{t=1}^{T}\left[q_{it}q_{jt} - E(q_{it}q_{jt})\right]\right| > \zeta_T\right) \leq \exp(-C_0\left(T^{1/2}\zeta_T\right)^{s/(s+2)}), \tag{B.34}$$

[1] In what follows we use

$$\Pr\left(|AB| > c\right) \leq \Pr\left(|A|\,|B| > c\right) \tag{B.32}$$

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, $\Pr(A_2 > c) \geq \Pr(A_1 > c)$ proving the result.

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

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

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

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

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

We now consider the second term of the RHS of (B.33). By (A.63), we have

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

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

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

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

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

or

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

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

$$\Pr\left(\left\|\boldsymbol{u}_x'\mathbf{Q}\right\|_F \left\|\mathbf{Q}'\boldsymbol{u}\right\|_F > \delta_T T\right) \le \exp\left[-C_0 T^{C_1}\right]. \tag{B.35}$$

We wish to derive conditions for l_T under which $\frac{T^{1/2}\zeta_T}{\delta_T l_T \left\|\mathbf{\Sigma}_{qq}^{-1}\right\|_F\left(\left\|\mathbf{\Sigma}_{qq}^{-1}\right\|_F+\frac{\zeta_T}{\delta_T}\right)}$, $\frac{T^{1/2}}{\left\|\mathbf{\Sigma}_{qq}^{-1}\right\|_F l_T}$, 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 (A.78) and (B.35) are negligible. We let $\zeta_T = T^\lambda$, $l_T = T^d$, $\left\|\mathbf{\Sigma}_{qq}^{-1}\right\|_F = l_T^{1/2} = T^{d/2}$ 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

$$\frac{T^{1/2}\zeta_T}{\delta_T l_T \left\|\mathbf{\Sigma}_{qq}^{-1}\right\|_F\left(\left\|\mathbf{\Sigma}_{qq}^{-1}\right\|_F+\frac{\zeta_T}{\delta_T}\right)} = O\left(T^{1/2+\lambda-\alpha-2d}\right) + O\left(T^{1/2-3d/2}\right) \tag{B.36}$$

$$\frac{T^{1/2}}{\left\|\mathbf{\Sigma}_{qq}^{-1}\right\|_F l_T} = O\left(T^{1/2-3d/2}\right) \tag{B.37}$$

$$\frac{\delta_T}{l_T} = O\left(T^{\alpha-d}\right) \tag{B.38}$$

and

$$\frac{\zeta_T^2}{T} = O\left(T^{2\lambda-1}\right) = O\left(c\ln T\right) \tag{B.39}$$

Clearly $d < 1/3$. Setting $\alpha = 1/3$, ensures all conditions are satisfied. Since $\mathbf{\Sigma}_{qq}^{-1}$ is of lower norm order than $\hat{\mathbf{\Sigma}}_{qq}^{-1} - \mathbf{\Sigma}_{qq}^{-1}$, (B.31) 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.29) is replaced by

$$\Pr\left(\left|\boldsymbol{u}_x'\boldsymbol{u}\right| > \zeta_T\right) \leq \exp\left[-C_0\left(T^{-(1+\vartheta)/2}\zeta_T\right)^{s/(s+2)}\right], \tag{B.40}$$

else, under (D2), we have again that (B.29) 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 $\frac{T^{1/2}\zeta_T}{\delta_T l_T \left\|\mathbf{\Sigma}_{qq}^{-1}\right\|_F\left(\left\|\mathbf{\Sigma}_{qq}^{-1}\right\|_F+\frac{\zeta_T}{\delta_T}\right)}$, $\frac{T^{1/2}}{\left\|\mathbf{\Sigma}_{qq}^{-1}\right\|_F l_T}$, 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 $\frac{T^{1/2}\zeta_T}{\delta_T l_T \left\|\mathbf{\Sigma}_{qq}^{-1}\right\|_F\left(\left\|\mathbf{\Sigma}_{qq}^{-1}\right\|_F+\frac{\zeta_T}{\delta_T}\right)}$, $\frac{T^{1/2}}{\left\|\mathbf{\Sigma}_{qq}^{-1}\right\|_F l_T}$, $\frac{\delta_T}{l_T}$ 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 D1, 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 (1) 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 all the pairs x_t and $\boldsymbol{q}_{\cdot t}$, and y_t and $(\boldsymbol{q}'_{\cdot t}, x_t)$, where x_t is a generic element of $\{x_{1t}, x_{2t}, ..., x_{nt}\}$ that does not belong to $\boldsymbol{q}_{\cdot t}$, and denote the corresponding projection residuals defined by (20) as $u_{x,t} = x_t - \boldsymbol{\gamma}'_{qx,T}\boldsymbol{q}_{\cdot t}$ and $e_t = y_t - \boldsymbol{\gamma}'_{yqx,T}(\boldsymbol{q}'_{\cdot t}, x_t)'$. Define $\boldsymbol{x} = (x_1, x_2, ..., x_T)'$, $\mathbf{y} = (y_1, y_2, ..., y_T)'$, $\boldsymbol{e} = (e_1, e_2, ..., e_T)'$, $\mathbf{M}_q = \mathbf{I}_T - \mathbf{Q}(\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}'$, 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 $i = 0, 1$,*

$$\Pr\left[|t_x| > c_p(n, \delta) |\theta = 0\right] \leq \exp\left[\frac{-(1-\pi)^2 \sigma^2_{e,(T)}\sigma^2_{x,(T)}c_p^2(n, \delta)}{2(1+d_T)^2 \omega^2_{xe,T}}\right] + \exp\left(-C_0 T^{C_1}\right), \tag{B.41}$$

where

$$t_x = \frac{T^{-1/2}\boldsymbol{x}'\mathbf{M}_q\mathbf{y}}{\sqrt{(\mathbf{e}'\mathbf{e}/T)\left(\frac{\mathbf{x}'\mathbf{M}_q\mathbf{x}}{T}\right)}}, \tag{B.42}$$

$$\sigma^2_{e,(T)} = E\left(T^{-1}\mathbf{e}'\mathbf{e}\right),\ \sigma^2_{x,(T)} = E\left(T^{-1}\mathbf{x}'\mathbf{M}_q\mathbf{x}\right), \tag{B.43}$$

and

$$\omega^2_{xe,T} = \frac{1}{T}\sum_{t=1}^{T} E\left[(u_{x,t}\eta_t)^2\right]. \tag{B.44}$$

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

$$\Pr\left[|t_x| > c_p(n, \delta) |\theta = 0\right] \leq \exp\left[\frac{-(1-\pi)^2 c_p^2(n, \delta)}{2(1+d_T)^2}\right] + \exp\left(-C_0 T^{C_1}\right). \tag{B.45}$$

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

$$\Pr\left[|t_x| > c_p(n, \delta) |\theta \neq 0\right] > 1 - \exp\left(-C_0 T^{C_1}\right). \tag{B.46}$$

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

$$\Pr\left[|t_x| > c_p(n, \delta) |\theta = 0\right] \leq \exp\left[-c_p(n, \delta)^{s/(s+2)}\right] + \exp\left(-C_0 T^{C_1}\right), \tag{B.47}$$

and

$$\Pr\left[|t_x| > c_p\left(n,\delta\right)|\theta \neq 0\right] > 1 - \exp\left(-C_0 T^{C_1}\right). \tag{B.48}$$

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 (2), can be written as

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

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,

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

Let $\theta = E\left(T^{-1}\boldsymbol{x}'\mathbf{M}_q\mathbf{X}_b\right)\boldsymbol{\beta}_b$, $\boldsymbol{\eta} = \mathbf{X}_b\boldsymbol{\beta}_b + \mathbf{u}$, $\boldsymbol{\eta} = (\eta_1, \eta_2, ..., \eta_T)'$, and write (A.109) as

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

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

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

Now by (A.102) of Lemma 9 and (B.24) of Lemma D2, we have

$$\Pr\left[|t_x| > c_p\left(n,\delta\right)|\theta = 0\right] = \Pr\left[\left|\frac{T^{1/2}\left(\frac{\mathbf{x}'\mathbf{M}_q\mathbf{x}}{T}\right)^{-1/2}\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}}{T}}{\sqrt{(\mathbf{e}'\mathbf{e}/T)}}\right| > c_p\left(n,\delta\right)|\theta = 0\right] \leq \tag{B.49}$$

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

Then, by Lemma F1, under Assumption D1 and defining $\boldsymbol{\alpha}(\boldsymbol{X}_T) = \left(\frac{\mathbf{x}'\mathbf{M}_q\mathbf{x}}{T}\right)^{-1/2}\boldsymbol{x}'\mathbf{M}_q$ 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.24) of Lemma D2, we have,

$$\Pr\left[|t_x| > c_p\left(n,\delta\right)|\theta = 0\right] \leq \exp\left[\frac{-\left(1-\pi\right)^2\sigma^2_{e,(T)}\sigma^2_{x,(T)}c_p^2\left(n,\delta\right)}{2\left(1+d_T\right)^2\omega^2_{xe,T}}\right] \tag{B.50}$$
$$+\exp\left(-C_0T^{C_1}\right)$$

where

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

and $u_{x,t}$, being the error in the regression of x_t on $\mathbf{Q}$, is defined by (20). Since by assumption u_t are distributed independently of $u_{x,t}$ and $\boldsymbol{x}_{b,t}$, then

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

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

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

Similarly

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

and since under $\theta = 0$, $\mathbf{x}$ being a pure 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

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

Using (A.111) and (A.112), it is now easily seen that if either $E\left(u_{x,t}^2\right) = \sigma_{ux}^2$ or $E\left(u_t^2\right) = \sigma^2$, for all t, then we have $\omega_{xe,T}^2 = \sigma_{e,(T)}^2\sigma_{x,(T)}^2$, and hence

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

giving a rate that does not depend on error variances. Next, we consider $\theta \neq 0$. By (A.101) of Lemma 9, for $d_T > 0$,

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

We then have

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

Then

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

We note that, by Lemma A2,

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

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

$$\Pr\left(\left|\frac{T^{1/2}\left(\frac{\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}}{T}-\theta\right)}{\sigma_{e,(T)}\sigma_{x,(T)}}\right| > \frac{T^{1/2}|\theta|}{\sigma_{e,(T)}\sigma_{x,(T)}} - \frac{c_p(n,\delta)}{1+d_T}\right) \tag{B.51}$$
$$\le \exp\left[-C_1\left(T^{1/2}\left(\frac{T^{1/2}|\theta|}{\sigma_{e,(T)}\sigma_{x,(T)}} - \frac{\theta c_p(n,\delta)}{1+d_T}\right)\right)^{s/(s+2)}\right].$$

So overall

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

Finally, we note the changes needed to the above arguments when Assumption D2 holds, rather than D1. (B.47) follows if in (B.49) we use (B.28) of Lemma D2 rather than (B.24) and, in (B.50), we use Lemma F2 rather than Lemma F1 and, again, we use (B.28) of Lemma D2 rather than (B.24). (B.47) follows again by using (B.28) of Lemma D2 rather than (B.24).

■

Remark D2 *We note that the above proof makes use of Lemmas F1 and F2. Alternatively one can use (A.101) of Lemma 9 in (B.49)-(B.50), rather that (A.102) of Lemma 9 and use the same line of proof as that provided in Lemma 10. 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 6 and 10.

Lemma E1 *Let x_{it}, $i = 1, 2, ..., n$, be a set of bounded deterministic sequences and u_t satisfy Assumptions 2-4 and 3, and consider the data generating process (1) 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 $\boldsymbol{x}_t = (x_{1t}, x_{2t}, ..., x_{nt})'$. Let $\eta_t = \boldsymbol{x}_{b,t}\boldsymbol{\beta}_b + u_{\eta,t}$, where $\boldsymbol{x}_{b,t}$ contains all signal variables 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 all the pairs x_{it} and $\boldsymbol{q}_{\cdot t}$, u_t and $\boldsymbol{q}_{\cdot t}$, and y_t and $(\boldsymbol{q}'_{\cdot t}, x_t)$, where x_t is a generic element of $\{x_{1t}, x_{2t}, ..., x_{nt}\}$ that does not belong to $\boldsymbol{q}_{\cdot t}$. Let $u_{x_i,T}$ be as in (20), 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,*

$$\Pr\left(\left|\sum_{t=1}^T x_{it}\eta_t - E\left(x_{it}\eta_t\left|\mathcal{F}_{t-1}\right.\right)\right| > \zeta_T\right) \leq \exp\left[\frac{-\left(1-\pi\right)^2\zeta_T^2}{2T\omega^2_{x_i\eta,1,T}}\right], \tag{B.52}$$

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

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

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

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

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

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

Proof. Note that all results used in this proof hold both for sequences and triangular arrays. (B.52) follows immediately given our assumptions and Lemma 3. 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.1) 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.2) to the last term of the above we obtain

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

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

$$\Pr\left(\left|\sum_{t=1}^{T} \left(\hat{u}_{x,t}u_{\eta,t} - \mu_{x\eta,t}\right)\right| > \zeta_T\right) \tag{B.53}$$
$$\leq \Pr\left(\left|\sum_{t=1}^{T} \left(u_{x,t}u_{\eta,t} - \mu_{x\eta,t}\right)\right| > \pi_1\zeta_T\right) + \Pr\left(\left\|\boldsymbol{\Sigma}_{qq}^{-1}\right\|_F > \frac{\pi_2\zeta_T}{\delta_T}\right)$$
$$+ \Pr\left(\left\|\mathbf{Q}'\boldsymbol{u}_{\eta}\right\|_F > \left(\pi_2\delta_T T\right)^{1/2}\right) + \Pr\left(\left\|\boldsymbol{u}_{x}'\mathbf{Q}\right\|_F > \left(\pi_2\delta_T T\right)^{1/2}\right).$$

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 4, and $\{q_{it}u_{x_i,t}\}$ are bounded sequences, then for some $C, c > 0$ we have[2]

$$\sup_i \Pr\left(\left\|\mathbf{q}_i'\boldsymbol{u}_{\eta}\right\| > \left(\pi_2\delta_T T\right)^{1/2}\right) \leq \exp\left(-C_0 T^{C_1}\right)$$

[2]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 5.

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

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

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

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

which upon using (A.74) 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 $\zeta_T/\delta_T \to \infty$, $l_T = o(\delta_T)$ 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 (1) and suppose that $\boldsymbol{x}_t = (x_{1t}, x_{2t}, ..., x_{nt})'$ are bounded deterministic sequences, and u_t satisfy Assumptions 2-4, and either Assumption 3 or Assumption 3 hold. Let $\boldsymbol{q}_{\cdot t} = (q_{1,t}, q_{2,t}, ..., q_{l_T,t})'$ contain a constant and a subset of $\boldsymbol{x}_t = (x_{1t}, x_{2t}, ..., x_{nt})'$, 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 all the pairs x_{it} and $\boldsymbol{q}_{\cdot t}$, and u_t and $\boldsymbol{q}_{\cdot t}$. Define $\boldsymbol{x} = (x_1, x_2, ..., x_T)'$, $\mathbf{y} = (y_1, y_2, ..., y_T)'$, $\mathbf{M}_q = \mathbf{I}_T - \mathbf{Q}(\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}'$, 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 (20), 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 $i = 0, 1$,*

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

where

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

$\sigma^2_{u,(T)}$ *and* $\sigma^2_{x,(T)}$ *are defined by (A.95) and (A.90), and*

$$\omega^2_{xu,T} = \frac{1}{T}\sum_{t=1}^{T}\sigma^2_{xt}\sigma^2_t,$$

Under $\sigma^2_t = \sigma^2$ *and/or* $\sigma^2_{xt} = \sigma^2_x$ *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* $c_p(n,\delta) = o(\sqrt{T})$*, then for* $d_T > 0$ *and some* $C_i > 0$*,* $i = 0, 1, 2, 3$*, we have*

$$\Pr\left[|t_x| > c_p(n,\delta)\,|\theta \neq 0\right] > 1 - C_0\exp\left(-C_1 T^{C_3}\right).$$

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

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

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,

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

Let

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

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

and write (A.109) as

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

$$\boldsymbol{x}'\mathbf{M}_q\,\boldsymbol{\eta} - E\left(\boldsymbol{x}'\mathbf{M}_q\boldsymbol{\eta}\right) = \left[\boldsymbol{x}'\mathbf{M}_q\mathbf{u} - E\left(\boldsymbol{x}'\mathbf{M}_q\boldsymbol{u}\right)\right],$$
$$\frac{\left(\mathbf{M}_q\mathbf{X}_b\boldsymbol{\beta}_b\right)'\left(\mathbf{M}_q\mathbf{X}_b\boldsymbol{\beta}_b\right)}{T} = \frac{1}{T}\sum_{t=1}^{T}\left(\boldsymbol{x}'_{bq,1}\boldsymbol{\beta}_b\right)^2 = \frac{1}{T}\sum_{t=1}^{T}\sigma_{xbt}^2 = \sigma_{b,(T)}^2.$$

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

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

By Lemma 9, we have

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

By Lemma E1, it then follows that,

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

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

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

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

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

giving a rate that does not depend on variances. Next, we consider $\theta \neq 0$. By Lemma 9, for $d_T > 0$,

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

We then have

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

Then,

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

We note that

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

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

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

So overall,

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

■

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 $\Pr(|u_t| > \alpha) \leq C_0\exp(-C_1\alpha^s)$, for all $\alpha > 0$. Let $\boldsymbol{X}_T = (\boldsymbol{x}_1, \boldsymbol{x}_2, .., \boldsymbol{x}_T)$, where $\boldsymbol{x}_t$ is an $l_T \times 1$ dimensional vector of random variables, with probability measure given by $P(\boldsymbol{X}_T)$, and assume*

$$E\left(u_t\left|\mathcal{F}_T^x\right.\right) = 0, \text{ for all } t = 1,2,...,T, \tag{B.54}$$

where $\mathcal{F}_T^x = \sigma\left(\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_T\right)$*. 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,*

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

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.54) and Lemma 3

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

But

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

proving the result.

Lemma F2 *Suppose that* u_t*,* $t = 1, 2, ..., T$*, is a mixing random variable with exponential mixing coefficients given by* $\alpha_k = a_0 \varphi^k$*,* $0 < \varphi < 1$*. , with constant variance* σ^2*, and there exist sufficiently large 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* $\boldsymbol{X}_T = (\boldsymbol{x}_1, \boldsymbol{x}_2, .., \boldsymbol{x}_T)$*, where* $\boldsymbol{x}_t$ *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 $\mathcal{F}_T^x = \sigma\left(\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_T\right)$. 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, using Lemma 2 of Dendramis, Giraitis, and Kapetanios (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

$$\Pr\left(\left\|\hat{\boldsymbol{A}}_T^{-1} - \boldsymbol{A}_T^{-1}\right\| > b_T\right) \leq l_T^2 \exp\left(\frac{-C_0\left(T^{1/2}b_T\right)^{s/(s+2)}}{l_T^{s/(s+2)}\left\|\boldsymbol{A}_T^{-1}\right\|_F^{s/(s+2)}\left(\left\|\boldsymbol{A}_T^{-1}\right\|_F + b_T\right)^{s/(s+2)}}\right) + l_T^2 \exp\left(-C_0\frac{T^{s/2(s+2)}}{\left\|\boldsymbol{A}_T^{-1}\right\|_F^{s/(s+2)} l_T^{s/(s+2)}}\right), \quad \text{(B.55)}$$

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

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

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

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

$$\begin{aligned}
\Pr\left(\left\|\hat{\boldsymbol{A}}_T - \boldsymbol{A}_T\right\|_F > b_T\right) &\leq \sum_{j=1}^{l_T}\sum_{i=1}^{l_T} \Pr\left(\left|\hat{a}_{ij,T} - a_{ij,T}\right|^2 > l_t^{-2}b_T^2\right) \qquad \text{(B.56)} \\
&= \sum_{j=1}^{l_T}\sum_{i=1}^{l_T} \Pr\left(\left|\hat{a}_{ij,T} - a_{ij,T}\right| > l_t^{-1}b_T\right) \\
&\leq l_T^2 \sup_{ij}\left[\Pr\left(\left|\hat{a}_{ij,T} - a_{ij,T}\right| > l_t^{-1}b_T\right)\right] = l_T^2 \exp\left(-C_0 T^{s/2(s+1)}\frac{b_T^{s/(s+2)}}{l_t^{s/(s+2)}}\right).
\end{aligned}$$

To establish (B.55) define the sets

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

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

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

Furthermore

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

and by (B.56) 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

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

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

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

Result (B.55) now follows if we apply (B.56) 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 $C > |B| > 0$,*

$$\Pr\left(|X + B| > C\right) \leq \Pr\left(|X| > C - |B|\right). \tag{B.57}$$

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