Supplemental Appendix

We state primitive conditions that are sufficient for fixed-b asymptotic theory in Section 3.1 and the asymptotic validity of the bootstrap in Section 4 with proofs for the random missing process case. For the non-random missing process case, primitive conditions are made about the latent process. Hence the results of Gonçalves and Vogelsang (2011) directly apply and no proof is required.

We derive results under the assumption that the latent processes is near epoch dependent (NED) on an underlying mixing process similar to Gonçalves and Vogelsang (2011) and that the missing process is strong mixing. We follow the definitions in Davidson (2002). Let the L_p norm of x be defined as $\|x\|_p = (E|x|^p)^{1/p}$. Also, let $|\bullet|$ denote the Euclidean norm of the corresponding vector or matrix. For a stochastic sequence $\{\varepsilon_t\}_{-\infty}^{\infty}$, on a probability space $(\Omega, \mathcal{F}, P)$, let $\mathcal{F}_{t-m}^{t+m} = \sigma(\varepsilon_{t-m}, \ldots, \varepsilon_{t+m})$, such that $\{\mathcal{F}_{t-m}^{t+m}\}_{m=0}^{\infty}$ is an increasing sequence of σ-fields. We say that a sequence of integrable random variables $\{w_t\}_{-\infty}^{\infty}$ is L_p-NED on $\{\varepsilon_t\}_{-\infty}^{\infty}$ if, for $p > 0$, $\|w_t - E(w_t|\mathcal{F}_{t-m}^{t+m})\|_p < d_t \nu_m$, where $\nu_m \to 0$ and $\{d_t\}_{-\infty}^{\infty}$ is a sequence of positive constants. For a sequence $\{a_t\}_{-\infty}^{\infty}$, let $\mathcal{F}_{-\infty}^{t} = \sigma(\ldots, a_{t-1}, a_t)$, and similarly define $\mathcal{F}_{t+m}^{\infty} = \sigma(a_{t+m}, a_{t+m+1}, \ldots)$. The sequence is said to be $\alpha-$mixing if $\lim_{m\to\infty} \alpha_m = 0$, where $\alpha_m = \sup_t \sup_{G\in\mathcal{F}_{-\infty}^{t}, H\in\mathcal{F}_{t+m}^{\infty}} |P(G\cap H) - P(G)P(H)|$. A sequence is $\alpha-$mixing of size $-\psi_0$ if $\alpha_m = O(m^{-\psi})$ for some $\psi > \psi_0$. Similarly, a sequence is L_p-NED of size $-\phi_0$ if $\nu_m = O(m^{-\phi})$ for some $\phi > \phi_0$.

We first state the primitive conditions that are sufficient for fixed-b asymptotic theory when the missing process is random and the AM approach is used (Lemma SA1). Recall that Assumption R is sufficient for fixed-b asymptotic theory to go through when the missing process is random and the AM approach is used (Section 3.1).

Assumption R.

1. $T^{-1}\sum_{t=1}^{[rT]} x_t x_t' \Rightarrow rQ,\ \forall r \in [0,1]$.

2. $T^{-1/2}\sum_{t=1}^{[rT]} v_t \Rightarrow \Lambda \mathcal{W}_k(r),\ \forall r \in [0,1]$.

The following Assumption R′ is sufficient for Assumption R.

Assumption R′.

1. *For some* $r > 2$, $\|x_t^*\|_{2r} \leq \Delta < \infty$ *for all* $t = 1, \ldots$.
2. $\{x_t^*\}$ *is a weakly stationary sequence* L_2-*NED on* $\{\varepsilon_t\}$ *with* NED *coefficient of size* $-\frac{2(r-1)}{r-2}$.
3. $\|v_t^*\|_r \leq \Delta < \infty$, *and* $E(v_t^*) = 0$ *for all* $t = 1, 2, \ldots$.
4. $\{v_t^*\}$ *is a mean zero weakly stationary sequence* L_2*-NED on* $\{\varepsilon_t\}$ *with* NED *coefficient of size* $-\frac{1}{2}$.
5. $\{(a_t, \varepsilon_t)\}$ *is a* $\alpha-$*mixing sequence with* $\alpha-$*mixing coefficient of size* $-\frac{2r}{r-2}$.
6. $\{a_t\}$ *is a weakly stationary process that is independent of* $\{(x_t^*, u_t^*)\}$.
7. $\Omega = \lim_{T\to\infty} Var\left(T^{-1/2}\sum_{t=1}^{T} a_t v_t^*\right)$ *is positive definite.*

Lemma SA1. *Assumption R′ is sufficient for Assumption R.*

Proof: Gonçalves and Vogelsang (2011, Assumption 1) is sufficient for Assumption R and we show that Assumption R′ is sufficient for Assumption R by showing that when Assumption R′ is satisfied the AM series satisfy Gonçalves and Vogelsang (2011, Assumption 1).

Define $\boldsymbol{\epsilon}_t = (a_t, \varepsilon_t)$. With Assumption R′, the AM series satisfy the following conditions (Gonçalves and Vogelsang (2011, Assumption 1)):

1. For some $r > 2$, $\|x_t\|_{2r} \leq \Delta < \infty$ for all $t = 1, 2, \ldots$.

2. $\{x_t\}$ is a weakly stationary sequence L_2-NED on $\{\boldsymbol{\epsilon_t}\}$ with NED coefficients of size $-\frac{2(r-1)}{r-2}$.
3. $\|v_t\|_r \leq \Delta < \infty$, and $E(v_t) = 0$ for all $t = 1, 2, \ldots$.
4. $\{v_t\}$ is a weakly stationary sequence L_2-NED on $\{\boldsymbol{\epsilon_t}\}$ with NED coefficients of size $-\frac{1}{2}$.
5. $\{\boldsymbol{\epsilon_t}\}$ is an $\alpha-$mixing sequence of size $-\frac{2r}{r-2}$.
6. $\Omega = lim_{T\to\infty} \text{Var}\left(T^{-1/2}\sum_{t=1}^{T} v_t\right)$ is positive definite.

1: Note that

$$\|x_t\|_{2r} = \|a_t x_t^*\|_{2r} \leq \|x_t^*\|_{2r} \leq \Delta < \infty, t = 1, \ldots, r > 2.$$

The first inequality follows form the fact that $\{a_t\}$ is a binary sequence. The second inequality is Assumption $R'1$.

2: Because $\{a_t\}$ and $\{x_t^*\}$ are weakly stationary, $\{x_t\}$ is also weakly stationary. To show that $\{x_t\}$ is L_2-NED, we first define the following notation. Let $\mathcal{F}_s^t = \sigma\left(\boldsymbol{\epsilon_s}, \boldsymbol{\epsilon_{s+1}}, \ldots, \boldsymbol{\epsilon_t}\right)$ and $\mathcal{G}_s^t = \sigma\left(\varepsilon_s, \varepsilon_{s+1}, \ldots, \varepsilon_t\right)$. Note that we can write

$$\begin{aligned}
\left\|a_t x_t^* - E(a_t x_t^*|\mathcal{F}_{t-m}^{t+m})\right\|_p &= \left\|a_t\left(x_t^* - E(x_t^*|\mathcal{F}_{t-m}^{t+m})\right)\right\|_p \\
&\leq \left\|\left(x_t^* - E(x_t^*|\mathcal{F}_{t-m}^{t+m})\right)\right\|_p \\
&\leq 2\left\|\left(x_t^* - E(x_t^*|\mathcal{G}_{t-m}^{t+m})\right)\right\|_p \\
&\leq 2 d_t \nu_m.
\end{aligned}$$

The first equality follows from the fact that $\{a_t\}$ is $\mathcal{F}_{t-m}^{t+m}$ measurable. The first inequality is straightforward because $\{a_t\}$ is a binary sequence. The second inequality uses Davidson (2002, 10.28, p157). The last inequality uses the fact that $\{x_t^*\}$ is L_2-NED on $\{\varepsilon_t\}$ with NED coefficient of size $-2(r-1)/(r-2)$ (Assumption R$'$2). Therefore we have

$$\left\|a_t x_t^* - E(a_t x_t^*|\mathcal{F}_{t-m}^{t+m})\right\|_p \leq d_t' \nu_m, \quad d_t' = 2d_t,$$

where ν_m is of size $-2(r-1)/(r-2)$.

3: Note that we can write

$$\|v_t\|_r = \|a_t v_t^*\|_r \leq \|v_t^*\|_r \leq \Delta < \infty, \quad r > 2.$$

The first inequality uses the fact that $\{a_t\}$ is a binary sequence. The second inequality is Assumption R$'$3.

4: The proof of the fourth condition is identical to that of the second condition. we can write

$$\begin{aligned}
\left\|a_t v_t^* - E(a_t v_t^*|\mathcal{F}_{t-m}^{t+m})\right\|_p &= \left\|a_t\left(v_t^* - E(v_t^*|\mathcal{F}_{t-m}^{t+m})\right)\right\|_p \\
&\leq \left\|\left(v_t^* - E(v_t^*|\mathcal{F}_{t-m}^{t+m})\right)\right\|_p \\
&\leq 2\left\|\left(v_t^* - E(v_t^*|\mathcal{G}_{t-m}^{t+m})\right)\right\|_p \\
&\leq 2 d_t \nu_m.
\end{aligned}$$

The first equality follows from the fact that $\{a_t\}$ is $\mathcal{F}_{t-m}^{t+m}$ measurable. The first inequality is straightforward because $\{a_t\}$ is a binary sequence. The second inequality uses Davidson (2002, 10.28, p157). The last inequality uses the fact that $\{v_t^*\}$ is L_2-NED on $\{\varepsilon_t\}$ with NED coefficient of size $-1/2$ (Assumption R$'$4). Therefore we have

$$\left\|a_t v_t^* - E(a_t v_t^*|\mathcal{F}_{t-m}^{t+m})\right\|_p \leq d_t' \nu_m, \quad d_t' = 2d_t,$$

where ν_m is of size $-1/2$.

5: The fifth condition is identical to Assumption R$'$5.

6: The sixth condition is identical to Assumption R$'$7. $\square$

Next, we prove that when the missing process is random and Assumption R$'$ with Assumption R$'$ 3-5 strengthened to Assumption R$''$ 3-5 is satisfied, the moving block bootstrap (MBB) HAR *Wald* test, $W_T^{\bullet}$, defined in Section 4 has the usual fixed-b limit in Kiefer and Vogelsang (2005). This result is stated in Theorem SA1.

Assumption R$''$.

3. $\|v_t^*\|_{r+\delta} < \infty, r > 2.$

4. $\{v_t^*\}$ *is a weakly stationary* $L_{2+\delta}-$*NED on* $\{\varepsilon_t\}$ *with* ν_m *of size* -1.

5. $\{(a_t, \varepsilon_t)\}$ *is a* $\alpha-$*mixing sequence with* α_m *of size* $-\frac{(2+\delta)(r+\delta)}{r-2}$.

Theorem SA1. *Let* $W_T^{\bullet}$ *and* $t_T^{\bullet}$ *be naive bootstrap test statistics obtained from the moving block bootstrap resamples as defined in Section 4. Suppose that the block size* l *is either fixed as* $T \to \infty$ *or* $l \to \infty$ *as* $T \to \infty$ *such that* $l^2/T \to 0$. *Let* $b \in (0,1]$ *be fixed and suppose* $M = bT$. *Then, under Assumption R$'$ with Assumption R$'$3-5 strengthened to Assumption R$''$3-5, as* $T \to \infty$,

$$W_T^{\bullet} \overset{p^{\bullet}}{\Rightarrow} \mathcal{W}_q'(1) P(b, \tilde{B}_q)^{-1} \mathcal{W}_q(1)$$

and

$$t_T^{\bullet} \overset{p^{\bullet}}{\Rightarrow} \frac{\mathcal{W}_1(1)}{\sqrt{P(b, \tilde{B}_1)}}.$$

For the proof of Theorem SA1, we start by three lemmas (Lemmas SA2-SA4) which are the building blocks for proving the required weak dependence of the functions of AM series - $\{x_t x_t'\}$, $\{v_t\}$, $\{v_t v_{t+j}'\}$ (Results 1-3). With these required weak dependence results of the functions of AM series, we prove three lemmas (Lemmas SA5-SA7). These lemmas in turn would be used to prove that Assumption R' with Assumption R$'$ 3-5 strengthened to Assumption R$''$ 3-5 is sufficient for conditions (a) and (b) in Section 4, (a) $T^{-1}\sum_{t=1}^{[rT]} x_t^{\bullet} x_t^{\bullet\prime} \overset{p^{\bullet}}{\Rightarrow} rQ^{\bullet}$ and (b) $T^{-1/2}\sum_{t=1}^{[rT]} v_t^{\bullet} \overset{p^{\bullet}}{\Rightarrow} \Lambda^{\bullet}\mathcal{W}_k(r)$, which completes the proof of Theorem SA1.

Lemma SA2 shows that under Assumption R$'$ the mean zero AM series are mixingales (see, e.g., Davidson (2002, p247) for a definition of mixingale). Lemma SA3 and Lemma SA4 show properties of NED and mixingale sequence. With these three lemmas we show in Results 1-3 that the functions of AM series - $\{x_t x_t'\}$, $\{v_t\}$, $\{v_t v_{t+j}'\}$ - satisfy the required weak dependence conditions.

Lemma SA2. *Let* $r \geq p \geq 1$. *Suppose* $\|w_t\|_r \leq \Delta < \infty$. *Let* $\{a_t\}$ *be a random sequence which takes values either* 0 *or* 1. *If* $\{(a_t, \varepsilon_t)\}$ *is a* $\alpha-$*mixing sequence with* α_m *of size* $-a$ *and* $\{w_t\}$ *is* L_p-*NED on* $\{\varepsilon_t\}$ *with* ν_m *of size* $-b$, *then* $\{a_t w_t - E(a_t w_t), \mathcal{F}^t\}$ *is* L_p-*mixingale of size* $-\min\{b, a\frac{r-2}{2r}\}$ *with uniformly bounded mixingale constants where* $\mathcal{F}^t$ *is a nondecreasing sequence of* σ*-fields,* $\sigma(\boldsymbol{X_t}, \boldsymbol{X_{t-1}}, \dots)$, $\boldsymbol{X_t} = (a_t, \varepsilon_t)$.

Proof: We start by defining the following notation. Let $\boldsymbol{X_t} = (a_t, \varepsilon_t)$, $\mathcal{F}_s^t = \sigma(\boldsymbol{X_s}, \boldsymbol{X_{s+1}}, \dots, \boldsymbol{X_t})$, $\mathcal{G}_s^t = \sigma(\varepsilon_s, \varepsilon_{s+1}, \dots, \varepsilon_t)$. Proving that $\{a_t w_t - E(a_t w_t)\}$ is L_p-mixingale is equivalent to proving

$$\left\| E\left[a_t w_t - E(a_t w_t) | \mathcal{F}_{-\infty}^{t-m}\right] \right\|_p \leq c_t \psi_m \quad \text{(SA.1)}$$

$$\left\| a_t w_t - E(a_t w_t) - E\left[a_t w_t - E(a_t w_t) | \mathcal{F}_{-\infty}^{t+m}\right] \right\|_p \leq c_t \psi_{m+1}. \quad \text{(SA.2)}$$

Proof of (SA.1): Let $m \geq 1$ and let $k = \left[\frac{m}{2}\right]$ be the largest integer not exceeding $\frac{m}{2}$. By the Minkowski inequality (Davidson (2002, 9.27, p139)) we can rewrite (SA.1) as

$$
\begin{aligned}
&\left\| E\left[a_t w_t - E(a_t w_t) | \mathcal{F}_{-\infty}^{t-m}\right] \right\|_p \\
=&\left\| E\left[a_t w_t - a_t E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right] + a_t E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right] - E\left(a_t E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right]\right) + E\left(a_t E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right]\right) - E(a_t w_t) \right| \mathcal{F}_{-\infty}^{t-m}\right] \right\|_p \\
\leq&\left\| E\left[a_t \left(w_t - E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right]\right) \right| \mathcal{F}_{-\infty}^{t-m}\right] \right\|_p + \left\| E\left[a_t E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right] - E\left(a_t E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right]\right) \right| \mathcal{F}_{-\infty}^{t-m}\right] \right\|_p \\
&+ \left\| E\left(a_t \left(E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right] - w_t\right)\right) \right\|_p \\
\equiv&\Pi_1 + \Pi_2 + \Pi_3.
\end{aligned}
$$

We can bound each of the three terms as follows. Π_1 can be rewritten as

$$
\begin{aligned}
\Pi_1 &\leq \left\| a_t \left(w_t - E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right]\right) \right\|_p \\
&\leq \left\| w_t - E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right] \right\|_p \\
&\leq d_t \nu_k.
\end{aligned}
$$

The first inequality uses the conditional Jensen's inequality and law of iterated expectations. The second inequality is straightforward because a_t is a binary process. Third inequality is using the fact that w_t is L_p-NED on $\{\varepsilon_t\}$ with NED coefficient ν_m.

Next we bound Π_2. Note that $E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right]$ is a finite-lag measurable function of $\varepsilon_{t-k}, \ldots, \varepsilon_{t+k}$ for finite k. Because $\{(a_t, \varepsilon_t)\}$ is an $\alpha-$mixing sequence with α_m of size $-a$, $E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right]$ is $\alpha-$mixing of size $-a$. This in turn implies that $a_t E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right]$ is $\alpha-$mixing of size $-a$ (see Davidson (2002, Theorem 14.1, p210)). Then, using a mixing inequality (Davidson (2002, Theorem 14.2, p211)), we can write

$$
\begin{aligned}
\Pi_2 &\leq 2\left(2^{\frac{1}{p}} + 1\right) \alpha_k^{\frac{1}{p} - \frac{1}{r}} \left\| a_t E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right] \right\|_r \\
&\leq 6 \alpha_k^{\frac{1}{p} - \frac{1}{r}} \left\| a_t E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right] \right\|_r \\
&\leq 6 \alpha_k^{\frac{1}{p} - \frac{1}{r}} \left\| E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right] \right\|_r \\
&\leq 6 \alpha_k^{\frac{1}{p} - \frac{1}{r}} \left\| w_t \right\|_r.
\end{aligned}
$$

The second and the third inequalities are straightforward by noting that $p \geq 1$ and a_t is a binary process. The last inequality follows from the conditional Jensen's inequality and law of iterated expectations.

Finally, we bound Π_3. Π_3 can be rewritten as

$$
\begin{aligned}
\Pi_3 &= \left| E\left(a_t \left(E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right] - w_t\right)\right) \right| \\
&\leq \left\| a_t \left(E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right] - w_t\right) \right\|_1 \\
&\leq \left\| E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right] - w_t \right\|_1 \\
&\leq \left\| w_t - E\left[w_t | \mathcal{G}_{t-k}^{t+k}\right] \right\|_P \\
&\leq d_t \nu_k.
\end{aligned}
$$

The first inequality uses Jensen's inequality. Because a_t is a binary process the second inequality is straightforward. Because $p \geq 1$, by Liapunov's inequality (Davidson (2002, 9.23, p138)), the third inequality is also straightforward. The last inequality follows from the fact that w_t is L_p-NED on $\{\varepsilon_t\}$ with NED coefficient ν_m. Hence combining the inequality results for all three terms, we have

$$
\begin{aligned}
\left\| E\left[a_t w_t - E(a_t w_t)|\mathcal{F}_{-\infty}^{t-m}\right]\right\|_p &\leq A1_1 + A1_2 + A1_3 \\
&\leq 2d_t \nu_k + 6\alpha_k^{\frac{1}{p}-\frac{1}{r}} \|w_t\|_r \\
&\leq \max\{d_t, \|w_t\|_r\}\left(2\nu_k + 6\alpha_k^{\frac{1}{p}-\frac{1}{r}}\right) \equiv c_t\psi_m.
\end{aligned}
$$

Proof of (SA.2): We can rewrite (SA.2) as

$$
\begin{aligned}
\left\|(a_t w_t - E(a_t w_t)) - E\left[a_t w_t - E(a_t w_t)|\, \mathcal{F}_{-\infty}^{t+m}\right]\right\|_p &= \left\|a_t w_t - E\left[a_t w_t|\, \mathcal{F}_{-\infty}^{t+m}\right]\right\|_p \\
&\leq 2\left\|a_t w_t - E\left[a_t w_t|\, \mathcal{F}_{t-m}^{t+m}\right]\right\|_p \\
&= 2\left\|a_t w_t - a_t E\left[w_t|\, \mathcal{F}_{t-m}^{t+m}\right]\right\|_p \quad \because a_t \text{ is } \mathcal{F}_{t-m}^{t+m}-\text{measurable} \\
&\leq 2\left\|w_t - E\left[w_t|\, \mathcal{F}_{t-m}^{t+m}\right]\right\|_p \\
&\leq 2d_t\nu_m \leq 2d_t\nu_{\left[\frac{m+1}{2}\right]} \leq c_t\psi_{m+1}.
\end{aligned}
$$

The first inequality follows from Davidson (2002, 10.28, p157). The second inequality is straightforward because a_t is a binary process. The third inequality is using the fact that w_t is L_p-NED on $\{\varepsilon_t\}$ with NED coefficient ν_m. The fourth inequality is straightforward because without loss of generality we can consider $\{\nu_m\}_{m=1}^{\infty}$ as a decreasing sequence. Recall that ν_m is of size $-b$ and α_m is of size $-a$. Therefore $\{a_t w_t - E(a_t w_t)\}$ is L_p-mixingale with ψ_m of size $-\min\left\{b, a\frac{r-p}{pr}\right\}$ with $c_t << \max\{d_t, \|w_t\|_r\}$.

Now we are only left with proving that the mixingale constants are uniformly bounded. According to the Minkowski inequality (Davidson (2002, 9.27, p139)) and conditional Jensen's inequality,

$$
\begin{aligned}
\left\|w_t - E\left[w_t|\mathcal{G}_{t-m}^{t+m}\right]\right\|_p &\leq \|w_t\|_p + \left\|E\left[w_t|\mathcal{G}_{t-k}^{t+k}\right]\right\|_p \\
&\leq \|w_t\|_p + \|w_t\|_p \\
&= 2\|w_t\|_p.
\end{aligned}
$$

Since $\|w_t\|_p \leq \|w_t\|_r$ by the norm inequality (Davidson (2002, 9.23, p138))) and $\|w_t\|_r$ is uniformly bounded, we can set d_t equal to a finite constant for all t. Thus, mixingale constant, $c_t << \max\{d_t, \|w_t\|_r\} \leq \max\left\{2\|w_t\|_p, \|w_t\|_r\right\}$, is uniformly bounded. □

Lemma SA3. *Let x_t and w_t be L_p-NED on $\{\varepsilon_t\}$ with ν_m^x and ν_m^w of respective sizes $-\phi_x$ and $-\phi_w$. Then $x_t w_t$ is $L_{p/2}$-NED of size $-\min\{\phi_x, \phi_w\}$.*

Proof: We follow the proof of Davidson (2002, Theorem 17.9, p268). Define $\mathcal{F}_s^t = \sigma(\varepsilon_s, \varepsilon_{s+1}, \ldots, \varepsilon_t)$. By the Minkowski inequality (Davidson (2002, 9.27, p139)), we can write

$$
\begin{aligned}
\left\|x_t w_t - E\left[x_t w_t|\, \mathcal{F}_{t-m}^{t+m}\right]\right\|_{\frac{p}{2}} &= \Big\|x_t w_t - x_t E\left[w_t|\, \mathcal{F}_{t-m}^{t+m}\right] + x_t E\left[w_t|\, \mathcal{F}_{t-m}^{t+m}\right] - E\left[x_t|\, \mathcal{F}_{t-m}^{t+m}\right] E\left[w_t|\, \mathcal{F}_{t-m}^{t+m}\right] \\
&\quad + E\left[x_t|\, \mathcal{F}_{t-m}^{t+m}\right] E\left[w_t|\, \mathcal{F}_{t-m}^{t+m}\right] - E\left[x_t w_t|\, \mathcal{F}_{t-m}^{t+m}\right]\Big\|_{\frac{p}{2}} \\
&\leq \left\|x_t w_t - x_t E\left[w_t|\, \mathcal{F}_{t-m}^{t+m}\right]\right\|_{\frac{p}{2}} + \left\|x_t E\left[w_t|\, \mathcal{F}_{t-m}^{t+m}\right] - E\left[x_t|\, \mathcal{F}_{t-m}^{t+m}\right] E\left[w_t|\, \mathcal{F}_{t-m}^{t+m}\right]\right\|_{\frac{p}{2}} \\
&\quad + \left\|E\left[x_t|\, \mathcal{F}_{t-m}^{t+m}\right] E\left[w_t|\, \mathcal{F}_{t-m}^{t+m}\right] - E\left[x_t w_t|\, \mathcal{F}_{t-m}^{t+m}\right]\right\|_{\frac{p}{2}} \\
&\equiv \Pi_1 + \Pi_2 + \Pi_3.
\end{aligned}
$$

First consider Π_1. By Hölder's inequality (Davidson (2002, 9.21, p138)) we can write

$$
\begin{aligned}
\Pi_1 = \left\|x_t\left(w_t - E\left[w_t|\, \mathcal{F}_{t-m}^{t+m}\right]\right)\right\|_{\frac{p}{2}} &\leq \|x_t\|_p \left\|w_t - E\left[w_t|\, \mathcal{F}_{t-m}^{t+m}\right]\right\|_p \\
&\leq \|x_t\|_p\, d_t^w \nu_m^w.
\end{aligned}
$$

The second inequality is straightforward because w_t is L_p-NED with NED coefficient ν_m^w.

Next we consider Π_2. By Hölder's inequality (Davidson (2002, 9.21, p138)), the conditional Jensen's inequality, and the law of iterated expectations, we can write

$$\Pi_2 = \left\| \left(x_t - E\left[x_t | \mathcal{F}_{t-m}^{t+m}\right]\right) E\left[w_t | \mathcal{F}_{t-m}^{t+m}\right] \right\|_{\frac{p}{2}} \leq \left\| x_t - E\left[x_t | \mathcal{F}_{t-m}^{t+m}\right] \right\|_p \|w_t\|_p$$
$$\leq d_t^x \nu_m^x \|w_t\|_p .$$

The second inequality is straightforward because x_t is L_p-NED with NED coefficient ν_m^x.

For Π_3, using the conditional Jensen's inequality we can write

$$\Pi_3 = \left\| E\left[\left(x_t - E\left[x_t | \mathcal{F}_{t-m}^{t+m}\right]\right) \left(w_t - E\left[w_t | \mathcal{F}_{t-m}^{t+m}\right]\right) \middle| \mathcal{F}_{t-m}^{t+m}\right] \right\|_{\frac{p}{2}} \leq \left\| \left(x_t - E\left[x_t | \mathcal{F}_{t-m}^{t+m}\right]\right) \left(w_t - E\left[w_t | \mathcal{F}_{t-m}^{t+m}\right]\right) \right\|_{\frac{p}{2}}$$
$$\leq \left\| x_t - E\left[x_t | \mathcal{F}_{t-m}^{t+m}\right] \right\|_p \left\| w_t - E\left[w_t | \mathcal{F}_{t-m}^{t+m}\right] \right\|_p$$
$$\leq d_t^x \nu_m^x d_t^w \nu_m^w .$$

The second inequality uses Hölder's inequality (Davidson (2002, 9.21, p138)). The third inequality follows from the fact that both x_t and w_t are L_p-NED on $\{\varepsilon_t\}$. Combining the three inequality results for Π_1, Π_2, and Π_3,

$$\left\| x_t w_t - E\left[x_t w_t | \mathcal{F}_{t-m}^{t+m}\right] \right\|_{\frac{p}{2}} \leq \|x_t\|_p d_t^w \nu_m^w + d_t^x \nu_m^x \|w_t\|_p + d_t^x \nu_m^x d_t^w \nu_m^w$$
$$\leq \max\left\{ \|x_t\|_p d_t^w, \|w_t\|_p d_t^x, d_t^x d_t^w \right\} \left(\nu_m^w + \nu_m^x + \nu_m^x \nu_m^w\right) \equiv d_t \nu_m .$$

In other words, $x_t w_t$ is $L_{p/2}-$NED on $\{\varepsilon_t\}$ with NED coefficients $\nu_m = \nu_m^w + \nu_m^x + \nu_m^x \nu_m^w$. This completes the proof because $\nu_m = O\left(m^{-\min\{\phi_x, \phi_w\}}\right)$. □

Lemma SA4. *For some nondecreasing sequence of σ-fields $\{\mathcal{F}^t\}$ and for some $p > 1$, let $\{w_t, \mathcal{F}^t\}$ be an L_p-mixingale with mixingale coefficients ψ_m and mixingale constants c_t. Then letting $S_j = \sum_{t=1}^{j} w_t$ and $\Psi = \sum_{m=1}^{\infty} \psi_m$, it follows that*

$$\left\| \max_{j \leq T} |S_j| \right\|_p \leq K\Psi \left(\sum_{t=1}^{T} c_t^{\beta} \right)^{\frac{1}{\beta}}, \quad \beta = \min\{p, 2\}$$

for some generic constant K.

Proof: See Hansen (1991), Hansen (1992). □

Result 1. *Under Assumption R', $\{x_t x_t' - E(x_t x_t')\}$ is L_2-mixingale of size -1 with uniformly bounded mixingale constants.*

Proof: First, we can show that under Assumption R', $\{x_t^* x_t^{*\prime}\}$ is L_2-NED on $\{\varepsilon_t\}$ of size -1 (see Davidson (2002, Example 17.17, p273)). Also note that $\|x_t^* x_t^{*\prime}\|_r \leq \Delta < \infty$ by Assumption R$'$1 and Hölder's inequality (Davidson (2002, 9.21, p138)). Therefore using Lemma SA2, $\{a_t x_t^{*\prime} x_t^{*\prime} - E(a_t x_t^{*\prime} x_t^{*\prime})\}$ is L_2-mixingale of size $-\min\{1, (2r/(r-2)) \times ((r-2)/2r)\}$ with uniformly bounded mixingale constants. In other words, $\{x_t x_t' - E(x_t x_t')\}$ is L_2-mixingale of size -1 with uniformly bounded mixingale constants. □

Result 2. *Under Assumption R'', v_t is $L_{2+\delta}-$mixingale of size -1 with uniformly bounded mixingale constants.*

Proof: Using Lemma SA2, $a_t v_t^* - E(a_t v_t^*)$ is $L_{2+\delta}-$mixingale of size $-\min\{1, ((2+\delta)(r+\delta)/(r-2)) \times ((r-2)/2r)\} = -\min\{1, (2+\delta)(r+\delta)/2r\} = -1$ with uniformly bounded mixingale constants. Note that $E(a_t v_t^*) = 0$. Hence $a_t v_t^*$ is $L_{2+\delta}-$mixingale of size -1. In other words, v_t is $L_{2+\delta}$-mixingale of size -1 with uniformly bounded mixingale constants. □

Result 3. *Under Assumption R'', $\left\{v_t v'_{t+j} - E\left(v_t v'_{t+j}\right)\right\}$ is $L_{(2+\delta)/2}-$mixingale of size -1 with uniformly bounded mixingale constants.*

Proof: Note that under Assumption R$''$4, $\{v_t^*\}$ is $L_{2+\delta}-$NED on $\{\varepsilon_t\}$ of size -1, which implies that $\{v_{t+j}^*\}$ is $L_{2+\delta}-$NED on $\{\varepsilon_t\}$ of size -1 as well (see Davidson (2002, Theorem 17.10, p268)). Then $\left\{v_t^* v_{t+j}^{*\prime}\right\}$ is $L_{(2+\delta)/2}-$NED on $\{\varepsilon_t\}$ of size -1 by Lemma SA3. Also note that under Assumption $R''5$, $\{(a_t, \varepsilon_t)\}$ is α-mixing of size $-(2+\delta)(r+\delta)/(r-2)$ which implies that the binary process $a_t a_{t+j}$ is also $\alpha-$mixing of the same size, $-(2+\delta)(r+\delta)/(r-2)$. By the application of Lemma SA2, $\left\{a_t a_{t+j} v_t^* v_{t+j}^{*\prime} - E\left(a_t a_{t+j} v_t^* v_{t+j}^{*\prime}\right)\right\}$ is $L_{(2+\delta)/2}-$mixingale of size $-\min\{1, ((2+\delta)(r+\delta)/(r-2)) \times ((r-2)/2r)\} = -\min\{1, (2+\delta)(r+\delta)/2r\}$ with uniformly bounded mixingale constants. In other words, $\left\{v_t v'_{t+j} - E\left(v_t v_{t+j}\right)\right\}$ is $L_{(2+\delta)/2}-$mixingale of size -1 with uniformly bounded mixingale constants. □

Using Results 1-3 above, we prove Lemmas SA5-SA7. Lemma SA5 establishes a LLN for the MBB sample mean. Lemma SA6 gives the probability limits of the MBB variance of the scaled bootstrap sample mean. Lemma SA7 establishes a FCLT for the MBB partial sum process. These will be used to prove Theorem SA1.

Our proofs and notation are similar to those of Gonçalves and Vogelsang (2011). We use the following notation. Define the vector $\omega_t = (y_t, x_t')'$ that collects dependent and explanatory variables. Let $l \in \mathbb{N}(1 \le l \le T)$ be a block length and let $B_{t,l} = \{\omega_t, \omega_{t+1}, \ldots, \omega_{t+l-1}\}$ be the block of l consecutive observations starting at ω_t. Draw $k_0 = T/l$ blocks randomly with replacement from the set of overlapping blocks $\{B_{1,l}, \ldots, B_{T-l+1,l}\}$ to obtain a bootstrap resample denoted as $\omega_t^\bullet = (y_t^\bullet, x_t^{\bullet\prime})'$, $t = 1, \ldots, T$. Given MBB resample $\omega_t^\bullet = (y_t^\bullet, x_t^{\bullet\prime})'$, we let $v_{0t}^\bullet = x_t^\bullet(y_t^\bullet - x_t^{\bullet\prime}\beta) \equiv x_t^\bullet u_{0t}^\bullet$ and $v_t^\bullet = x_t^\bullet(y_t^\bullet - x_t^{\bullet\prime}\hat{\beta}) \equiv x_t^\bullet u_t^\bullet$. $p^\bullet$ denotes the probability measure induced by the bootstrap resampling, conditional on a realization of the original time series. Let $Z_T^\bullet$ be bootstrap statistics. Then, we write $Z_T^\bullet = o_{p^\bullet}(1)$ in probability or $Z_T^\bullet \xrightarrow{p^\bullet} 0$ if for any $\varepsilon > 0$, $\delta > 0$, $\lim_{T\to\infty} p[p^\bullet(|Z_T^\bullet| > \delta) > \varepsilon] = 0$. Similarly we say that $Z_T^\bullet = O_{p^\bullet}(1)$ in probability if for all $\varepsilon > 0$ there exists an $M_\varepsilon < \infty$ such that $\lim_{T\to\infty} p[p^\bullet(|Z_T^\bullet| > M_\varepsilon) > \varepsilon] = 0$. Finally, we write $Z_T^\bullet \overset{p^\bullet}{\Rightarrow} Z$ in probability if conditional on the sample, $Z_T^\bullet$ weakly converges to Z under $p^\bullet$, for all samples contained in a set with probability converging to one. Specifically, we write $Z_T^\bullet \overset{p^\bullet}{\Rightarrow} Z$ in probability if and only if $E^\bullet[f(Z_T^\bullet)] \to E[f(Z)]$ in probability for any bounded and uniformly continuous function f.

Lemma SA5. *Suppose that $\{w_t - E(w_t)\}$ is a weakly stationary L_2-mixingale with $\|w_t\|_p \le \Delta < \infty$ for some $p > 2$ such that its mixingale coefficients ψ_m satisfy $\sum_{m=1}^{\infty} \psi_m < \infty$ and its mixingale constants are uniformly bounded. Let $\{w_t^\bullet : t = 1, \ldots, T\}$ denote an moving block bootstrap resample of $\{w_t : t = 1, \ldots, T\}$ with block size l satisfying either of the two following conditions: (a) l is fixed as $T \to \infty$, or (b) $l \to \infty$ as $T \to \infty$ with $l = o(T)$. Then, for any $\eta > 0$, as $T \to \infty$,*

$$p^\bullet\left(\sup_{r\in[0,1]}\left|T^{-1}\sum_{t=1}^{[rT]}\left(w_t^\bullet - E^\bullet\left(w_t^\bullet\right)\right)\right| > \eta\right) = o_p(1).$$

Proof: We follow Gonçalves and Vogelsang (2011, Proof of Lemma A.4). Note that we can write

$$\frac{1}{T}\sum_{t=1}^{[rT]}\left(w_t^\bullet - E^\bullet\left(w_t^\bullet\right)\right) = \frac{1}{T}\sum_{m=1}^{M_r}\sum_{s=1}^{B}\left(w_{I_m+s} - E^\bullet\left(w_{I_m+s}\right)\right),$$

where $M_r = [([rT]-1)/l] + 1$ and $B = \min\{l, [rT] - (m-1)l\}$. Note that $I_1, \ldots, I_{k_0}$ are *i.i.d.* uniformly distributed on $\{0, \ldots, T-l\}$ and for $r \in [0,1]$, $M_r \in \{1, \ldots, k_0\}$ and $B \in \{1, \ldots, l\}$. We can further write

$$\begin{aligned}\frac{1}{T}\sum_{m=1}^{M_r}\sum_{s=1}^{B}\left(w_{I_m+s} - E^\bullet\left(w_{I_m+s}\right)\right) &= \frac{1}{T}\sum_{m=1}^{M_r}\sum_{s=1}^{l}\left(w_{I_m+s} - E^\bullet\left(w_{I_m+s}\right)\right) - \frac{1}{T}\sum_{s=B_{M_R}+1}^{l}\left(w_{I_{M_r}+s} - E^\bullet\left(w_{I_{M_r}+s}\right)\right)\\ &\equiv \Pi_{1T}(r) + \Pi_{2T}(r),\end{aligned}$$

where $B_{M_r} = [rT] - (M_r - 1)l$. By the Markov inequality it is sufficient to show that

$$E^{\bullet}\left(\sup_{r\in[0,1]}|\Pi_{1T}(r)|^2\right) = E^{\bullet}\left(\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{m=1}^{M_r}\sum_{s=1}^{l}\left(w_{I_m+s} - E^{\bullet}\left(w_{I_m+s}\right)\right)\right|^2\right) = o_p(1) \quad \text{and} \tag{SA.3}$$

$$E^{\bullet}\left(\sup_{r\in[0,1]}|\Pi_{2T}(r)|\right) = E^{\bullet}\left(\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{s=B_{M_r}+1}^{l}\left(w_{I_{M_r}+s} - E^{\bullet}\left(w_{I_{M_r}+s}\right)\right)\right|\right) = o_p(1). \tag{SA.4}$$

First we consider (SA.3). Note that $\mathbb{M}_j = \sum_{m=1}^{j}\sum_{s=1}^{l}\left(w_{I_m+s} - E^{\bullet}\left(w_{I_m+s}\right)\right)$ is a martingale array with respect to $\mathcal{F}_{T,j} = \sigma\left(I_1, \dots, I_j\right)$. By Doob's inequality (see Davidson (2002, 15.15, p241)), it follows that

$$\begin{aligned}E^{\bullet}\left(\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{m=1}^{M_r}\sum_{s=1}^{l}\left(w_{I_m+s} - E^{\bullet}\left(w_{I_m+s}\right)\right)\right|^2\right) &= \frac{1}{T^2}E^{\bullet}\left(\max_{1\le j\le k_0}\left|\mathbb{M}_j\right|^2\right)\\ &\le \frac{1}{T^2}\frac{2}{2-1}E^{\bullet}\left(\left|\mathbb{M}_{k_0}\right|^2\right) = \frac{2}{T^2}E^{\bullet}\left(\left|\mathbb{M}_{k_0}\right|^2\right).\end{aligned}$$

We can further write as

$$\begin{aligned}E^{\bullet}\left(\left|\mathbb{M}_{k_0}\right|^2\right) &= k_0 E^{\bullet}\left(\left|\sum_{s=1}^{l}\left(w_{I_1+s} - E^{\bullet}\left(w_{I_1+s}\right)\right)\right|^2\right)\\ &= k_0 E^{\bullet}\left(\left|\sum_{s=1}^{l}\left(w_{I_1+s} - E(w_t) + E(w_t) - E^{\bullet}\left(w_{I_1+s}\right)\right)\right|^2\right)\\ &= k_0 E^{\bullet}\left(\left|\sum_{s=1}^{l}\left(w_{I_1+s} - E(w_t)\right) - E^{\bullet}\left(w_{I_1+s} - E(w_t)\right)\right|^2\right)\\ &\le 2k_0 E^{\bullet}\left(\left|\sum_{s=1}^{l}\left(w_{I_1+s} - E(w_t)\right)\right|^2\right).\end{aligned}$$

The first equality follows from the fact that $\sum_{s=1}^{l}\left(w_{I_m+s} - E^{\bullet}\left(w_{I_m+s}\right)\right)$ is a sum of independent random variables with respect to the probability measure $p^{\bullet}$ and $I_j \sim i.i.d.$ uniform $\{0, \dots, T-l\}$. The inequality follows immediately from Jensen's and the triangle inequalities.

Using Lemma SA4 we can show that

$$E\left(2k_0 E^{\bullet}\left(\left|\sum_{s=1}^{l}\left(w_{I_1+s} - E(w_t)\right)\right|^2\right)\right) = O(T).$$

Hence,

$$E\left(E^{\bullet}\left(\sup_{r\in[0,1]}|\Pi_{1T}(r)|^2\right)\right) = \frac{1}{T^2}O(T) = O(T^{-1}) = o(1). \tag{SA.5}$$

By the Markov inequality (SA.5) implies (SA.3) and we are left with proving (SA.4).

Notice that we can write

$$
\begin{aligned}
E^{\bullet}\left(\sup_{r\in[0,1]}|\Pi_{2T}(r)|\right) &= E^{\bullet}\left(\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{s=B_{M_r}+1}^{l}\left(w_{I_{M_r}+s}-E^{\bullet}\left(w_{I_{M_r}+s}\right)\right)\right|\right) \\
&= E^{\bullet}\left(\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{s=B_{M_r}+1}^{l}\left(w_{I_{M_r}+s}-E(w_t)+E(w_t)-E^{\bullet}\left(w_{I_{M_r}+s}\right)\right)\right|\right) \\
&= E^{\bullet}\left(\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{s=B_{M_r}+1}^{l}\left(w_{I_{M_r}+s}-E(w_t)\right)-E^{\bullet}\left(w_{I_{M_r}+s}-E(w_t)\right)\right|\right) \\
&\leq 2E^{\bullet}\left(\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{s=1}^{l}\left(w_{I_{M_r}+s}-E(w_t)\right)\right|\right) \\
&= \frac{1}{T-l+1}\sum_{j=0}^{T-l}\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{s=1}^{l}\left(w_{j+s}-E(w_t)\right)\right|.
\end{aligned}
$$

The inequality follows immediately from Jensen's and the triangle inequalities. The last equality follows from the fact that $I_{M_r} \sim i.i.d.$ uniform $\{0,\ldots,T-l\}$. Note that

$$
\begin{aligned}
E\left(\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{s=B_{M_r}+1}^{l}\left(w_{j+s}-E(w_t)\right)\right|\right) &\leq \frac{1}{T}E\left(\max_{1\leq i\leq l}\left|\sum_{s=j+i}^{j+l}\left(w_{j+s}-E(w_t)\right)\right|\right) \\
&\leq \frac{1}{T}\left\|\max_{1\leq i\leq l}\left|\sum_{s=j+i}^{j+l}\left(w_{j+s}-E(w_t)\right)\right|\right\|_2.
\end{aligned}
$$

Recall that for any $r\in[0,1]$, $B_{M_r}\in\{1,\ldots,l\}$. Hence the first inequality follows immediately. By the norm inequality (Davidson (2002, 9.23, p138)), the second inequality is also straightforward. Since $\{w_{j+s}-E(w_t)\}$ is L_2-mixingale, applying Lemma SA4, we can write

$$
\left\|\max_{1\leq i\leq l}\left|\sum_{s=j+i}^{j+l}\left(w_{j+s}-E(w_t)\right)\right|\right\|_2 \leq K\Psi\left(\sum_{t=1}^{l}c_t^2\right)^{1/2},
$$

where $\{c_t\}$ are mixingale constants and $\Psi=\sum_{m=1}^{\infty}\psi_m<\infty$. Since the mixingale constants are uniformly bounded, $K\Psi\left(\sum_{t=1}^{l}c_t^2\right)^{1/2}=O(l^{1/2})$, which implies that

$$
E\left(\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{s=B_{M_r}+1}^{l}\left(w_{j+s}-E(w_t)\right)\right|\right) \leq \frac{1}{T}\left\|\max_{1\leq i\leq l}\left|\sum_{s=j+i}^{j+l}\left(w_{j+s}-E(w_t)\right)\right|\right\|_2 = O\left(\frac{l^{1/2}}{T}\right).
$$

Therefore,

$$
E\left(E^{\bullet}\left(\sup_{r\in[0,1]}|\Pi_{2T}(r)|\right)\right) \leq \frac{1}{T}\left\|\max_{1\leq i\leq l}\left|\sum_{s=j+i}^{j+l}\left(w_{j+s}-E(w_t)\right)\right|\right\|_2 = O\left(\frac{l^{1/2}}{T}\right) = o(1). \tag{SA.6}
$$

Note that $O\left(\frac{l^{1/2}}{T}\right)=o(1)$ for both l fixed and $l\to\infty, l/T\to 0$. By Markov inequality (SA.6) implies (SA.4) which completes the proof. □

Lemma SA6. *Let $\Omega_T^{\bullet}=Var^{\bullet}\left(T^{-1/2}\sum_{t=1}^{T}v_{0t}^{\bullet}\right)$. Suppose that Assumption R′ is satisfied.*

(a) *For any fixed* l *such that* $1 \leq l < T$, $T \to \infty$,

$$p \lim_{T\to\infty} \Omega_T^{\bullet} = \Gamma_0 + \sum_{j=1}^{l} \left(1 - \frac{j}{l}\right) \left(\Gamma_j + \Gamma_j'\right) \equiv \Omega_l,$$

where $\Gamma_j = E\left(v_t v_{t-j}'\right)$.

(b) *Let* $l = l_T \to \infty$ *as* $T \to \infty$ *such that* $l^2/T \to 0$. *Then,*

$$p \lim_{T\to\infty} \Omega_T^{\bullet} = \Gamma_0 + \sum_{j=1}^{\infty} \left(\Gamma_j + \Gamma_j'\right) \equiv \Omega.$$

Proof: From Lemma SA1, we know that Assumption R$'$ is sufficient for Gonçalves and Vogelsang (2011, Assumption 1) which is sufficient for Gonçalves and Vogelsang (2011, Assumption A). This in turn is sufficient for proving this lemma. See Gonçalves and Vogelsang (2011) for details. □

Lemma SA7. *Suppose that Assumption R$''$ holds. Define* $Z_T^{\bullet} = T^{-1/2} \sum_{T=1}^{[rT]} \left(v_{0t}^{\bullet} - E^{\bullet}\left(v_{0t}^{\bullet}\right)\right)$. *Let* Ω_l *and* Ω *as defined in Lemma SA6 be positive definite matrices. It follows that*

(a) *For any fixed* l *such that* $1 \leq l < T$ *as* $T \to \infty$,

$$Z_T^{\bullet}(r) \Rightarrow^{p^{\bullet}} \Lambda_l \mathcal{W}_k(r),$$

in probability where Λ_l *is the square root matrix of* Ω_l.

(b) *Let* $l = l_T \to \infty$ *as* $T \to \infty$ *such that* $l^2/T \to 0$. *Then*

$$Z_T^{\bullet}(r) \Rightarrow^{p^{\bullet}} \Lambda \mathcal{W}_k(r),$$

in probability where Λ *is the square root matrix of* Ω.

Proof: We follow the proofs in Gonçalves and Vogelsang (2011, Lemma A.3). In fact, a sufficient condition for the proof is that v_t is $L_{2+\delta}-$mixingale with size -1 with uniformly bounded mixing coefficients, which is implied by Assumption R'' (see Result 2).

We will show $\lambda' \Omega_T^{\bullet -1/2} T^{-1/2} \sum_{t=1}^{[rT]} \left(v_{0t}^{\bullet} - E^{\bullet}(v_{0t}^{\bullet})\right) \Rightarrow^{p^{\bullet}} \lambda' \mathcal{W}_k(r)$ in probability for any λ such that $\lambda'\lambda = 1$. For any $r \in [0, 1]$, we can write

$$\begin{aligned}
&\lambda' \Omega_T^{\bullet -1/2} T^{-1/2} \sum_{t=1}^{[rT]} \left(v_{0t}^{\bullet} - E^{\bullet}(v_{0t}^{\bullet})\right) \\
=&\lambda' \Omega_T^{\bullet -1/2} T^{-1/2} \sum_{m=1}^{M_r} \sum_{s=1}^{B} \left(v_{I_m+s} - E^{\bullet}(v_{I_m+s})\right) \\
=&\lambda' \Omega_T^{\bullet -1/2} T^{-1/2} \sum_{m=1}^{M_r} \sum_{s=1}^{l} \left(v_{I_m+s} - E^{\bullet}(v_{I_m+s})\right) - \lambda' \Omega_T^{\bullet -1/2} T^{-1/2} \sum_{s=B_{M_r}+1}^{l} \left(v_{I_{M_r}+s} - E^{\bullet}(v_{I_m+s})\right) \\
\equiv&\Pi_{1T}(r) - \Pi_{2T}(r),
\end{aligned}$$

where $M_r = [(\lbrack rT \rbrack - 1)/l] + 1$, $B = \min\{l, [rT] - (m-1)l\}$, and $B_{M_r}[rT] - (M_r - 1)l$. Recall that $I_1, \ldots, I_{k_0}$ are *i.i.d.* uniformly distributed on $\{0, \ldots, T-l\}$ and for any $r \in [0, 1]$, $M_r \in \{1, \ldots, k_0\}$ and $B \in \{1, \ldots, l\}$. We first show that $\sup_{r\in(0,1]} |\Pi_{2T}(r)| = O_{p^{\bullet}}(k_0^{-1/2}) = o_{p^{\bullet}}(1)$ in probability and then show that $\Pi_{1T}(r) \Rightarrow^{p^{\bullet}} \mathcal{W}_1(r)$.

To show $\sup_{r\in(0,1]} |\Pi_{2T}(r)| = O_{p^{\bullet}}(k_0^{-1/2})$, it is sufficient to show $E^{\bullet}\left(\sup_{r\in(0,1]} |\Pi_{2T}(r)|\right) = O_p(k_0^{-1/2})$ by the Markov inequality. Notice that $\Omega_T^{\bullet -1/2} = O_p(1)$ because by Lemma SA6, plim $\Omega_T^{\bullet} = \Omega^{\bullet}$ and $\Omega^{\bullet}$ is p.d.. Therefore, it is sufficient to show that $E^{\bullet}\left(\sup_{r\in(0,1]}\left|T^{-1/2}\sum_{s=B_{M_r}+1}^{l}\left(v_{I_m+s}-E^{\bullet}(v_{I_m+s})\right)\right|\right)=O_p(k_0^{-1/2})$. Using Jensen's and the triangle inequalities, it follows that

$$
\begin{aligned}
E^{\bullet}\left(\sup_{r\in(0,1]}\left|T^{-1/2}\sum_{s=B_{M_r}+1}^{l}\left(v_{I_m+s}-E^{\bullet}(v_{I_m+s})\right)\right|\right) &\leq 2E^{\bullet}\left(\sup_{r\in(0,1]}\left|T^{-1/2}\sum_{s=B_{M_r}+1}^{l}v_{I_m+s}\right|\right)\\
&=\frac{2T^{-1/2}}{T-l+1}\sum_{j=0}^{T-l}\sup_{r\in(0,1]}\left|\sum_{s=B_{M_r}+1}^{l}v_{j+s}\right|\\
&\leq\frac{2T^{-1/2}}{T-l+1}\sum_{j=0}^{T-l}\max_{1\leq i\leq l}\left|\sum_{s=j+i}^{j+l}v_s\right|.
\end{aligned}
$$

The equality is straightforward because $I_1,\ldots,I_{k_0}$ are *i.i.d.* uniformly distributed on $\{0,\ldots,T-l\}$. The second inequality is obvious because for $r\in[0,1]$, $B\in\{1,\ldots,l\}$. Hence to show that $E^{\bullet}\left(\sup_{r\in(0,1]}|\Pi_{2T}(r)|\right)=O_p(k_0^{-1/2})$, it is sufficient to show that $E\left(k_0^{1/2}\frac{2T^{-1/2}}{T-l+1}\sum_{j=0}^{T-l}\max_{1\leq i\leq l}\left|\sum_{s=j+i}^{j+l}v_s\right|\right)=O(1)$ by the Markov inequality. Using the norm inequality (Davidson (2002, 9.23, p138)), we can write

$$
\begin{aligned}
E\left(k_0^{1/2}\frac{2T^{-1/2}}{T-l+1}\sum_{j=0}^{T-l}\max_{1\leq i\leq l}\left|\sum_{s=j+i}^{j+l}v_s\right|\right) &\leq k_0^{1/2}\frac{2T^{-1/2}}{T-l+1}\sum_{j=0}^{T-l}\left\|\max_{1\leq i\leq l}\left|\sum_{s=j+i}^{j+l}v_s\right|\right\|_{2+\delta}\\
&\leq k_0^{1/2}\frac{2T^{-1/2}}{T-l+1}\sum_{j=0}^{T-l}K\Psi l^{1/2}\\
&=2(k_0 l)^{1/2}T^{-1/2}K\Psi=2K\Psi=O(1).
\end{aligned}
$$

The second inequality follows from the fact that $\{v_t\}$ is a $L_{2+\delta}-$mixingale of size -1 with uniformly bounded mixingale constants (see Result 2). Note that $\Psi=\sum_{m=1}^{\infty}\psi_m<\infty$ because ψ_m, the mixingale coefficient, is of size -1.

Next we show that $\Pi_{1T}(r)\Rightarrow^{p^{\bullet}}\mathcal{W}_1(r)$. Note that $\lambda'\Omega_T^{\bullet-1/2}T^{-1/2}\sum_{m=1}^{M_r}\sum_{s=1}^{l}\left(v_{I_m+s}-E^{\bullet}(v_{I_m+s})\right)$ is asymptotically equivalent to $\lambda'\Omega_T^{\bullet-1/2}(lk_0)^{-1/2}\sum_{m=1}^{[rk_0]+1}\sum_{s=1}^{l}\left(v_{I_m+s}-E^{\bullet}(v_{I_m+s})\right)$. By rearranging the terms, we can write

$$
k_0^{-1/2}\sum_{m=1}^{[rk_0]+1}\lambda'\Omega_T^{\bullet-1/2}\left(l^{-1/2}\sum_{s=1}^{l}\left(v_{I_m+s}-E^{\bullet}(v_{I_m+s})\right)\right)\equiv\sum_{m=1}^{[rk_0]+1}k_0^{-1/2}\mathbb{V}_m,
$$

where $\mathbb{V}_m\equiv\lambda'\Omega_T^{\bullet-1/2}\left(l^{-1/2}\sum_{s=1}^{l}\left(v_{I_m+s}-E^{\bullet}(v_{I_m+s})\right)\right)$. Here $\mathbb{V}_m$ is an array of independent variables with $E^{\bullet}(\mathbb{V}_m)=0$ and

$$
\begin{aligned}
Var^{\bullet}(\mathbb{V}_m) &= \lambda'\Omega_T^{\bullet-1/2}Var^{\bullet}\left(l^{-1/2}\sum_{s=1}^{l}\left(v_{I_m+s}-E^{\bullet}(v_{I_m+s})\right)\right)\Omega_T^{\bullet-1/2}\lambda\\
&=\lambda'\Omega_T^{\bullet-1/2}\Omega_T^{\bullet}\Omega_T^{\bullet-1/2}\lambda=1. \qquad \text{(SA.7)}
\end{aligned}
$$

We use a FCLT for martingale difference arrays. Note that $k_0^{-1/2}\mathbb{V}_m$ is a martingale array with respect to the σ-field $\mathcal{F}_{T,m-1}=\sigma(I_1,\ldots,I_{m-1})$ given the independence of $\mathbb{V}_m$. First, we can show that as $k_0\to\infty$,

$$
Var^{\bullet}\left(\sum_{m=1}^{[rk_0]+1}k_0^{-1/2}\mathbb{V}_m\right)=\frac{[rk_0]+1}{k_0}\to r. \qquad \text{(SA.8)}
$$

This is straightforward by (SA.7) and the fact that $\mathbb{V}_m$ is independent. Next, we show that

$$
p\lim_{k_0\to\infty}\sum_{m=1}^{[rk_0]+1}E^{\bullet}\left|k_0^{-1/2}\mathbb{V}_m\right|^{2+\delta}=0. \qquad \text{(SA.9)}
$$

(SA.9) implies that the Lindeberg condition holds in probability. Since $\Omega_T^\bullet = O_p(1)$, it is sufficient to show that

$$E\left(\sum_{m=1}^{[rk_0]+1} E^\bullet \left| k_0^{-1/2} l^{-1/2} \sum_{s=1}^{l} \left(v_{I_m+s} - E^\bullet(v_{I_m+s})\right)\right|^{2+\delta}\right) \to 0$$

by the Markov inequality. Note that

$$\begin{aligned}
E\left(\sum_{m=1}^{[rk_0]+1} E^\bullet \left| k_0^{-1/2} l^{-1/2} \sum_{s=1}^{l} \left(v_{I_m+s} - E^\bullet(v_{I_m+s})\right)\right|^{2+\delta}\right) &= T^{-(2+\delta)/2} E\left(\sum_{m=1}^{[rk_0]+1} E^\bullet \left| \sum_{s=1}^{l} \left(v_{I_m+s} - E^\bullet(v_{I_m+s})\right)\right|^{2+\delta}\right) \\
&\leq 2T^{-(2+\delta)/2} E\left(\sum_{m=1}^{[rk_0]+1} E^\bullet \left| \sum_{s=1}^{l} v_{I_m+s}\right|^{2+\delta}\right) \\
&= \frac{2T^{-(2+\delta)/2}}{T-l+1} \sum_{m=1}^{[rk_0]+1} \sum_{j=0}^{T-l} E\left(\left| \sum_{s=1}^{l} v_{j+s}\right|^{2+\delta}\right) \\
&\leq \frac{2T^{-(2+\delta)/2}}{T-l+1} \sum_{m=1}^{[rk_0]+1} \sum_{j=0}^{T-l} K^{2+\delta} \Psi^{2+\delta} \left(\sum_{t=1}^{l} c_t^2\right)^{(2+\delta)/2} \\
&\leq \frac{2T^{-(2+\delta)/2}}{T-l+1} \sum_{m=1}^{[rk_0]+1} \sum_{j=0}^{T-l} K' l^{(2+\delta)/2} \\
&= 2k_0^{-(2+\delta)/2} ([rk_0]+1) K' \\
&= O\left(k_0^{-\delta/2}\right) = O\left(\left(\frac{l}{T}\right)^{\delta/2}\right) = o(1).
\end{aligned}$$

The first inequality follows from Jensen's and the triangle inequalities. The second inequality is straightforward because $\{v_t\}$ is a $L_{2+\delta}-$mixingale of size -1 with uniformly bounded mixingale constants (see Result 2) and therefore Lemma SA4 applies. $\Psi = \sum_{m=1}^{\infty} \psi_m < \infty$ because $\{v_t\}$ is mixingale of size -1 which implies that $K' < \infty$. Therefore under given assumptions (SA.8) and (SA.9) are satisfied. By applying a FCLT for martingale difference arrays, it follows that $\sum_{m=1}^{[rk_0]+1} k_0^{-1/2} \mathbb{V}_m \Rightarrow \mathcal{W}(r)$. □

Proof of Theorem SA1: If we show that

1. $T^{-1} \sum_{t=1}^{[rT]} x_t^\bullet x_t^{\bullet\prime} \Rightarrow^{p^\bullet} rQ^\bullet$ for some $Q^\bullet$ and
2. $T^{-1/2} \sum_{t=1}^{[rT]} v_t^\bullet \Rightarrow^{p^\bullet} \Lambda^\bullet \mathcal{W}_k(r)$ for some $\Lambda^\bullet$

are true under Assumption R$'$ with Assumption R$'$ 3-5 strengthened to Assumption R$''$ 3-5, $W_T^\bullet$ will have the usual fixed-b limit given by Kiefer and Vogelsang (2005). Especially we want to prove that (1) $T^{-1} \sum_{t=1}^{[rT]} x_t^\bullet x_t^{\bullet\prime} \Rightarrow^{p^\bullet} rQ$ and (2) $T^{-1/2} \sum_{t=1}^{[rT]} v_t^\bullet \Rightarrow^{p^\bullet} \Lambda^\bullet \mathcal{W}_k(r)$ where $\Lambda^\bullet = \Lambda_l$ when l is fixed and $\Lambda^\bullet = \Lambda$ when $l \to \infty$, $l^2/T \to 0$. Λ and Λ_l are defined in Lemma SA7.

First we show that $T^{-1} \sum_{t=1}^{[rT]} x_t^\bullet x_t^{\bullet\prime} \Rightarrow^{p^\bullet} rQ$. We can write

$$\begin{aligned}
\left|\frac{1}{T}\sum_{t=1}^{[rT]} x_t^\bullet x_t^{\bullet\prime} - rQ\right| &= \left|\frac{1}{T}\sum_{t=1}^{[rT]} \left(x_t^\bullet x_t^{\bullet\prime} - E^\bullet\left(x_t^\bullet x_t^{\bullet\prime}\right) + E^\bullet\left(x_t^\bullet x_t^{\bullet\prime}\right) - x_t x_t' + x_t x_t'\right) - rQ\right| \\
&\leq \left|\frac{1}{T}\sum_{t=1}^{[rT]} \left(x_t^\bullet x_t^{\bullet\prime} - E^\bullet\left(x_t^\bullet x_t^{\bullet\prime}\right)\right)\right| + \left|\frac{1}{T}\sum_{t=1}^{[rT]} \left(E^\bullet\left(x_t^\bullet x_t^{\bullet\prime}\right) - x_t x_t'\right)\right| + \left|\frac{1}{T}\sum_{t=1}^{[rT]} x_t x_t' - rQ\right|.
\end{aligned}$$

We can show that the first term converges to 0 in probability uniformly in r using Lemma SA5. Since x_t^* and a_t are weakly stationary x_t is also weakly stationary. Note that under Assumption R', $\{x_t x_t' - E(x_t x_t')\}$ is L_2-mixingale of size -1 with uniformly bounded mixingale constants (see Result 1). Since the mixingale coefficient is of size -1, $\sum_{m=1}^{\infty} \psi_m < \infty$. Also Assumption R′1 implies that $\|x_t x_t'\|_r \leq \Delta, r > 2$ by Hölder's inequality (Davidson (2002, 9.21, p138)). Therefore the conditions required for Lemma SA5 are satisfied and we have

$$p^{\bullet}\left(\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{t=1}^{[rT]}\left(x_t^{\bullet}x_t^{\bullet\prime} - E^{\bullet}\left(x_t^{\bullet}x_t^{\bullet\prime}\right)\right)\right| > \eta\right) = o_p(1).$$

Assumption R' implies Assumption R. Therefore we have

$$p^{\bullet}\left(\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{t=1}^{[rT]} x_t x_t' - rQ\right| > \eta\right) = o_p(1).$$

We are left with proving

$$p^{\bullet}\left(\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{t=1}^{[rT]}\left(E^{\bullet}\left(x_t^{\bullet}x_t^{\bullet\prime}\right) - x_t x_t'\right)\right| > \eta\right) = o_p(1).$$

To show this, we write as

$$\begin{aligned}
\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{t=1}^{[rT]}\left(E^{\bullet}\left(x_t^{\bullet}x_t^{\bullet\prime}\right) - x_t x_t'\right)\right| &= \sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{t=1}^{[rT]}\left(E^{\bullet}\left(x_t^{\bullet}x_t^{\bullet\prime} - E(x_t x_t')\right) - \left(x_t x_t' - E\left(x_t x_t'\right)\right)\right)\right| \\
&\leq \sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{t=1}^{[rT]} E^{\bullet}\left(x_t^{\bullet}x_t^{\bullet\prime} - E(x_t x_t')\right)\right| + \sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{t=1}^{[rT]}\left(x_t x_t' - E(x_t x_t')\right)\right| \\
&\equiv \Pi_{1T}(r) + \Pi_{2T}(r)
\end{aligned}$$

using the triangle inequality. Note that

$$\begin{aligned}
\Pi_{1T}(r) &= \sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{t=1}^{[rT]} E^{\bullet}\left(x_t^{\bullet}x_t^{\bullet\prime} - E(x_t x_t')\right)\right| \\
&= \sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{m=1}^{M_r}\sum_{s=1}^{l} E^{\bullet}\left(x_{I_m+s}x_{I_m+s}' - E(x_t x_t')\right) - \frac{1}{T}\sum_{s=B_{M_r}+1}^{l} E^{\bullet}\left(x_{I_{M_r}+s}x_{I_{M_r}+s}' - E(x_t x_t')\right)\right| \\
&\leq \sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{m=1}^{M_r}\sum_{s=1}^{l} E^{\bullet}\left(x_{I_m+s}x_{I_m+s}' - E(x_t x_t')\right)\right| + \sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{s=B_{M_r}+1}^{l} E^{\bullet}\left(x_{I_{M_r}+s}x_{I_{M_r}+s}' - E(x_t x_t')\right)\right|.
\end{aligned}$$

Using Jensen's inequality, we can write the second term as

$$\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{s=B_{M_r}+1}^{l} E^{\bullet}\left(x_{I_{M_r}+s}x_{I_{M_r}+s}' - E(x_t x_t')\right)\right| \leq E^{\bullet}\left(\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{s=B_{M_r}+1}^{l}\left(x_{I_{M_r}+s}x_{I_{M_r}+s}' - E(x_t x_t')\right)\right|\right),$$

which in turn can be shown to be

$$E\left(E^{\bullet}\left(\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{s=B_{M_r}+1}^{l}\left(x_{I_{M_r}+s}x_{I_{M_r}+s}' - E(x_t x_t')\right)\right|\right)\right) = O\left(\frac{l^{1/2}}{T}\right) = o(1).$$

For details, see (SA.6) in the proof of Lemma SA5.

For the first term, we can write

$$\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{m=1}^{M_r}\sum_{s=1}^{l}E^{\bullet}\left(x_{I_m+s}x_{I_m+s}'-E(x_tx_t')\right)\right| = \sup_{r\in[0,1]}\left|\frac{1}{k_0}\sum_{m=1}^{M_r}\frac{1}{l}\sum_{s=1}^{l}E^{\bullet}\left(x_{I_m+s}x_{I_m+s}'-E(x_tx_t')\right)\right|$$
$$\leq \frac{1}{k_0}\sum_{m=1}^{M_r}\sup_{r\in[0,1]}\left|\frac{1}{l}\sum_{s=1}^{l}E^{\bullet}\left(x_{I_m+s}x_{I_m+s}'-E(x_tx_t')\right)\right|$$
$$\leq \left|\frac{1}{l}\sum_{s=1}^{l}E^{\bullet}\left(x_{I_1+s}x_{I_1+s}'-E(x_tx_t')\right)\right|.$$

The first inequality uses the triangle inequality. The second inequality follows from the fact that $M_r \leq k_0$, $\{I_j\}$ is *i.i.d.* uniformly distributed on $\{0,\ldots,T-l\}$. Note that we can write (see Fitzenberger (1997))

$$E^{\bullet}\left(\frac{1}{l}\sum_{s=1}^{l}\left(x_{I_m+s}x_{I_m+s}'-E(x_tx_t')\right)\right) = \frac{1}{T}\sum_{t=1}^{T}\left(x_tx_t-E(x_tx_t')\right)+O_p\left(\frac{l}{T}\right).$$

Because $x_tx_t - E(x_tx_t')$ is L_2-mixingale of size -1 with uniformly bounded mixingale constants (see Result 1), applying Lemma SA4, we can write

$$E\left|\sum_{t=1}^{T}\left(x_tx_t-E(x_tx_t')\right)\right| \leq \left\|\sum_{t=1}^{T}\left(x_tx_t-E(x_tx_t')\right)\right\|_2 \leq K\Psi\left(\sum_{t=1}^{T}c_t^2\right)^{1/2} = O\left(T^{1/2}\right).$$

Therefore

$$\frac{1}{T}\sum_{t=1}^{T}\left(x_tx_t-E(x_tx_t')\right) = O_p(T^{-1/2}) \tag{SA.10}$$

by Markov inequality and we have

$$E^{\bullet}\left(\frac{1}{l}\sum_{s=1}^{l}\left(x_{I_m+s}x_{I_m+s}'-E(x_tx_t')\right)\right) = O_p(T^{-1/2})+O_p\left(\frac{l}{T}\right) = o_p(1).$$

Hence, $\Pi_{1T}(r) = o_{p^{\bullet}}(1)$ in probability. By (SA.10), it is straightforward to show that $\Pi_{2T}(r) = O_p(T^{-1/2}) = o_p(1)$ which completes the proof of the first condition.

Now we prove the second condition. Given our definitions for $v_{0t}^{\bullet}$ and $v_t^{\bullet}$, we can write

$$v_t^{\bullet} = v_{0t}^{\bullet} - x_t^{\bullet}x_t^{\bullet\prime}\left(\hat{\beta}-\beta\right),$$

which implies that

$$T^{-\frac{1}{2}}\sum_{t=1}^{[rT]}v_t^{\bullet} = T^{-\frac{1}{2}}\sum_{t=1}^{[rT]}\left(v_{0t}^{\bullet}-E^{\bullet}\left(v_{0t}^{\bullet}\right)\right)+T^{-\frac{1}{2}}\sum_{t=1}^{[rT]}E^{\bullet}\left(v_{0t}^{\bullet}\right)-T^{-\frac{1}{2}}\sum_{t=1}^{[rT]}x_t^{\bullet}x_t^{\bullet\prime}\left(\hat{\beta}-\beta\right)$$
$$\equiv Z_T^{\bullet}(r)+\Pi_{1T}^{\bullet}(r)-\Pi_{2T}^{\bullet}(r).$$

First note that $Z_T^{\bullet}(1) \Rightarrow^{p^{\bullet}} \Lambda^{\bullet}\mathcal{W}_k(r)$ by Lemma SA7. Thus we are done if we show that $\sup_{r\in[0,1]}\left|\Pi_{1T}^{\bullet}(r)-\Pi_{2T}^{\bullet}(r)\right| =$

$o_{p^\bullet}(1)$ in probability. Note that

$$\begin{aligned}
\Pi_{1T}^\bullet(r) - \Pi_{2T}^\bullet(r) &= T^{-\frac{1}{2}} \sum_{t=1}^{[rT]} E^\bullet \left(x_t^\bullet \left(y_t^\bullet - x_t^{\bullet\prime}\hat{\beta} + x_t^{\bullet\prime}\hat{\beta} - x_t^{\bullet\prime}\beta \right) \right) - T^{-\frac{1}{2}} \sum_{t=1}^{[rT]} x_t^\bullet x_t^{\bullet\prime} \left(\hat{\beta} - \beta \right) \\
&= T^{-\frac{1}{2}} \sum_{t=1}^{[rT]} E^\bullet \left(x_t^\bullet \left(y_t^\bullet - x_t^{\bullet\prime}\hat{\beta} \right) \right) + T^{-\frac{1}{2}} \sum_{t=1}^{[rT]} E^\bullet \left(x_t^\bullet x_t^{\bullet\prime}\hat{\beta} - x_t^\bullet x_t^{\bullet\prime}\beta \right) - T^{-\frac{1}{2}} \sum_{t=1}^{[rT]} x_t^\bullet x_t^{\bullet\prime} \left(\hat{\beta} - \beta \right) \\
&= T^{-\frac{1}{2}} \sum_{t=1}^{[rT]} E^\bullet \left(v_t^\bullet \right) - T^{-\frac{1}{2}} \sum_{t=1}^{[rT]} \left(x_t^\bullet x_t^{\bullet\prime} - E^\bullet \left(x_t^\bullet x_t^{\bullet\prime} \right) \right) \left(\hat{\beta} - \beta \right) \\
&\equiv \Gamma_{1T}^\bullet(r) - \Gamma_{2T}^\bullet(r).
\end{aligned}$$

It is sufficient to show that $\sup_{r\in[0,1]} \left|\Gamma_{1T}^\bullet(r)\right| = o_p(1)$ and $\sup_{r\in[0,1]} \left|\Gamma_{2T}^\bullet(r)\right| = o_{p^\bullet}(1)$ in probability by the triangle inequality. We first prove that $\sup_{r\in[0,1]} \left|\Gamma_{1T}^\bullet(r)\right| = o_p(1)$. We can write

$$\begin{aligned}
\Gamma_{1T}^\bullet(r) &= T^{-\frac{1}{2}} \sum_{t=1}^{[rT]} E^\bullet \left(v_t^\bullet \right) \\
&= T^{-\frac{1}{2}} \sum_{m=1}^{M_r} \sum_{s=1}^{B} E^\bullet \left(\hat{v}_{I_m+s} \right) \\
&= T^{-\frac{1}{2}} \sum_{m=1}^{M_r} \sum_{s=1}^{l} E^\bullet \left(\hat{v}_{I_m+s} \right) - T^{-\frac{1}{2}} \sum_{s=B_{M_r}+1}^{l} E^\bullet \left(\hat{v}_{I_{M_r}+s} \right) \\
&\equiv \mu_{1T}^\bullet - \mu_{2T}^\bullet,
\end{aligned}$$

where $M_r = \left[([rT]-1)/l \right] + 1$, $B = \min\{l, [rT] - (m-1)l\}$, and $B_{M_r} = [rT] - (M_r - 1)l$. Recall that $M_r \in \{1, \ldots, k_0\}$, $B \in \{1, \ldots, l\}$, and $I_1, \ldots, I_{k_0}$ are *i.i.d.* uniformly distributed on $\{0, 1, \ldots, T-l\}$. To show $\sup_{r\in[0,1]} |\mu_{1T}(r)| = o_p(1)$, we write as

$$\begin{aligned}
\sup_{r\in[0,1]} \left|\mu_{1T}^\bullet(r)\right| &= \sup_{r\in[0,1]} \left| T^{-\frac{1}{2}} \sum_{m=1}^{M_r} \sum_{s=1}^{l} E^\bullet \left(\hat{v}_{I_m+s} \right) \right| \\
&= \sup_{r\in[0,1]} \left| l^{1/2} k_0^{-\frac{1}{2}} \sum_{m=1}^{M_r} E^\bullet \left(\frac{1}{l} \sum_{s=1}^{l} \hat{v}_{I_m+s} \right) \right| \\
&\le l^{1/2} k_0^{-\frac{1}{2}} \sup_{r\in[0,1]} \sum_{m=1}^{M_r} \left| E^\bullet \left(\frac{1}{l} \sum_{s=1}^{l} \hat{v}_{I_m+s} \right) \right| \\
&\le l^{1/2} k_0^{1/2} \left| E^\bullet \left(\frac{1}{l} \sum_{s=1}^{l} \hat{v}_{I_1+s} \right) \right|.
\end{aligned}$$

The first inequality uses the triangle inequality. The last inequality follows from the fact that $M_r \le k_0$. Note that

$$E^\bullet \left(\frac{1}{l} \sum_{s=1}^{l} \hat{v}_{I_1+s} \right) = \frac{1}{T} \sum_{t=1}^{T} \hat{v}_t + O_p\left(\frac{l}{T} \right) = O_p\left(\frac{l}{T} \right).$$

See Fitzenberger (1997, MBB-lemma A.1) for details. The second equality follows from the fact that $\sum_{t=1}^{T} \hat{v}_t = 0$. Hence,

$$\sup_{r\in[0,1]} \left|\mu_{1T}^\bullet(r)\right| = O_p\left(\frac{l}{T^{1/2}} \right) = o_p(1).$$

Note that $O_p\left(\frac{l}{T^{1/2}} \right) = o_p(1)$ when l is either fixed or $l \to \infty$, $l^2/T \to 0$.

Next we show that $\sup_{r\in[0,1]} \left|\mu_{2T}^\bullet\right| = o_p(1)$. In fact we will show that $\sup_{r\in[0,1]} \left|\mu_{2T}^\bullet\right| = O_p(k_0^{-1/2})$ which implies $\sup_{r\in[0,1]} \left|\mu_{2T}^\bullet\right| = o_p(1)$ for both l fixed and $l^2/T \to 0$, $l \to \infty$. By the the Markov inequality, it is sufficient to show that $E\left(\sup_{r\in[0,1]}|\mu^{\bullet}_{2T}|\right) = O(k_0^{-1/2})$. First note that we can write

$$\begin{aligned}
|\mu^{\bullet}_{2T}| &= \left| T^{-\frac{1}{2}} \sum_{s=B_{M_r}+1}^{l} E^{\bullet}\left(\hat{v}_{I_{M_r}+s}\right)\right| \\
&= \left| T^{-\frac{1}{2}} \sum_{s=B_{M_r}+1}^{l} E^{\bullet}\left(v_{I_{M_r}+s} - x_{I_{M_r}+s}x'_{I_{M_r}+s}(\hat{\beta}-\beta)\right)\right| \\
&\leq \left| T^{-\frac{1}{2}} \sum_{s=B_{M_r}+1}^{l} E^{\bullet}\left(v_{I_{M_r}+s}\right)\right| + \left|\frac{1}{T}\sum_{s=B_{M_r}+1}^{l} E^{\bullet}\left(x_{I_{M_r}+s}x'_{I_{M_r}+s}\right)\right| \left|\sqrt{T}(\hat{\beta}-\beta)\right|.
\end{aligned}$$

Therefore,

$$\sup_{r\in[0,1]}|\mu^{\bullet}_{2T}| \leq \sup_{r\in[0,1]}\left| T^{-\frac{1}{2}} \sum_{s=B_{M_r}+1}^{l} E^{\bullet}\left(v_{I_{M_r}+s}\right)\right| + \sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{s=B_{M_r}+1}^{l} E^{\bullet}\left(x_{I_{M_r}+s}x'_{I_{M_r}+s}\right)\right| \left|\sqrt{T}(\hat{\beta}-\beta)\right|.$$

To show that $\sup_{r\in[0,1]}\left|\mu^{\bullet}_{2T}\right| = O_p(k_0^{-1/2})$, it is sufficient to show that $\sup_{r\in[0,1]}\left| T^{-\frac{1}{2}} \sum_{s=B_{M_r}+1}^{l} E^{\bullet}\left(v_{I_{M_r}+s}\right)\right| = O_p(k_0^{-1/2})$ and $\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{s=B_{M_r}+1}^{l} E^{\bullet}\left(x_{I_{M_r}+s}x_{I_{M_r}+s'}\right)\right| = O_p(k_0^{-1/2})$ because $\sqrt{T}(\hat{\beta}-\beta) = O_p(1)$. By Markov inequality, it is sufficient to show that

$$E\left(k_0^{1/2}\sup_{r\in[0,1]}\left| T^{-\frac{1}{2}} \sum_{s=B_{M_r}+1}^{l} E^{\bullet}\left(v_{I_{M_r}+s}\right)\right|\right) = O(1) \quad \text{and} \tag{SA.11}$$

$$E\left(k_0^{1/2}\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{s=B_{M_r}+1}^{l} E^{\bullet}\left(x_{I_{M_r}+s}x'_{I_{M_r}+s}\right)\right|\right) = O(1). \tag{SA.12}$$

First we show (SA.11). Note that

$$\begin{aligned}
E\left(k_0^{1/2}\sup_{r\in[0,1]}\left| T^{-\frac{1}{2}} \sum_{s=B_{M_r}+1}^{l} E^{\bullet}\left(v_{I_{M_r}+s}\right)\right|\right) &= E\left(k_0^{1/2}\sup_{r\in[0,1]}\left| T^{-\frac{1}{2}} \sum_{s=B_{M_r}+1}^{l} \frac{1}{T-l+1}\sum_{j=0}^{T-l} v_{j+s}\right|\right) \\
&\leq E\left(\frac{k_0^{1/2}T^{-\frac{1}{2}}}{T-l+1}\sum_{j=0}^{T-l}\sup_{r\in[0,1]}\left|\sum_{s=B_{M_r}+1}^{l} v_{j+s}\right|\right) \\
&\leq E\left(\frac{k_0^{\frac{1}{2}}T^{-\frac{1}{2}}}{T-l+1}\sum_{j=0}^{T-l}\max_{1\leq i\leq l}\left|\sum_{s=j+i}^{j+l} v_s\right|\right) \\
&\leq \frac{k_0^{\frac{1}{2}}T^{-\frac{1}{2}}}{T-l+1}\sum_{j=0}^{T-l}\left\|\max_{1\leq i\leq l}\left|\sum_{s=j+i}^{j+l} v_s\right|\right\|_{2+\delta} \\
&\leq \frac{k_0^{\frac{1}{2}}T^{-\frac{1}{2}}}{T-l+1}\sum_{j=0}^{T-l} K\Psi l^{1/2} = O(1).
\end{aligned} \tag{SA.13}$$

The first equality is by the definition of $E^{\bullet}$. Recall that $\{I_j\}$ is *i.i.d.* uniformly distributed on $\{0,\ldots,T-l\}$. The first inequality is trivial by the triangle inequality. The second inequality is obvious because for $r\in[0,1]$, $B\in\{1,\ldots,l\}$. The third inequality follows from the norm inequality (Davidson (2002, 9.23, p138)). Also note that under Assumption R'', $\{v_t\}$ is a $L_{2+\delta}-$mixingale of size -1 with uniformly bounded mixingale constants (see Result 2). Then, applying Lemma SA4, the fourth inequality immediately follows. Furthermore, the mixingale coefficient being of size -1 implies that $\Psi = \sum_{m=1}^{\infty}\psi_m < \infty$.

Next we show (SA.12). Following the same steps used to show (SA.13), we can write

$$
\begin{aligned}
& E\left(k_0^{1/2}\sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{s=B_{M_r}+1}^{l}E^{\bullet}\left(x_{I_{M_r}+s}x_{I_{M_r}+s'}\right)\right|\right) \\
\leq & \frac{k_0^{\frac{1}{2}}}{T(T-l+1)}\sum_{j=0}^{T-l}\left\|\max_{1\leq i\leq l}\left|\sum_{s=j+i}^{j+l}x_{j+s}x'_{j+s}\right|\right\|_2 \\
= & \frac{k_0^{\frac{1}{2}}}{T(T-l+1)}\sum_{j=0}^{T-l}\left\|\max_{1\leq i\leq l}\left|\sum_{s=j+i}^{j+l}\left(x_{j+s}x'_{j+s}-E(x_tx'_t)+E(x_tx'_t)\right)\right|\right\|_2 \\
\leq & \frac{k_0^{\frac{1}{2}}}{T(T-l+1)}\sum_{j=0}^{T-l}\left\|\max_{1\leq i\leq l}\left|\sum_{s=j+i}^{j+l}\left(x_{j+s}x'_{j+s}-E(x_tx'_t)\right)\right|\right\|_2+\frac{k_0^{\frac{1}{2}}l}{T(T-l+1)}\sum_{j=0}^{T-l}E(x_tx'_t) \\
\leq & \frac{k_0^{\frac{1}{2}}}{T(T-l+1)}\sum_{j=0}^{T-l}K\Psi l^{\frac{1}{2}}+\frac{k_0^{\frac{1}{2}}l}{T(T-l+1)}\sum_{j=0}^{T-l}E(x_tx'_t) \\
= & K\Psi\frac{(k_0l)^{1/2}}{T}+\frac{k_0^{1/2}l}{T}E(x_tx'_t) \\
= & O(T^{-1/2})+O\left(\left(\frac{l}{T}\right)^{1/2}\right).
\end{aligned}
$$

The second inequality follows from the Minkowski inequality (Davidson (2002, 9.27, p139)). Note that $\{x_tx'_t-E(x_tx'_t)\}$ is L_2-mixingale of size -1 with uniformly bounded mixingale constants (see Result 1). Thus, using Lemma SA4, the third inequality follows immediately. The first term is $O(T^{-1/2})=o(1)$. The second term is $O((1/T)^{1/2})=o(1)$ because l is either fixed or increasing slower than T. Hence we have shown that $\sup_{r\in[0,1]}\left|\mu^{\bullet}_{2T}\right|=o_p(1)$.

So far we have $\sup_{r\in[0,1]}\left|\mu^{\bullet}_{1T}(r)\right|=o_p(1)$ and $\sup_{r\in[0,1]}\left|\mu^{\bullet}_{2T}(r)\right|=o_p(1)$, which implies that $\sup_{r\in[0,1]}\left|\Gamma^{\bullet}_{1T}(r)\right|=o_p(1)$. We are left with proving $\sup_{r\in[0,1]}\left|\Gamma^{\bullet}_{2T}(r)\right|=o_{p^{\bullet}}(1)$. We can write

$$
\begin{aligned}
\sup_{r\in[0,1]}\left|\Gamma^{\bullet}_{2T}(r)\right| &= \sup_{r\in[0,1]}\left|T^{-\frac{1}{2}}\sum_{t=1}^{[rT]}\left(x^{\bullet}_tx^{\bullet\prime}_t-E^{\bullet}\left(x^{\bullet}_tx^{\bullet\prime}_t\right)\right)\right|\left|\hat{\beta}-\beta\right| \\
&= \sup_{r\in[0,1]}\left|\frac{1}{T}\sum_{t=1}^{[rT]}\left(x^{\bullet}_tx^{\bullet\prime}_t-E^{\bullet}\left(x^{\bullet}_tx^{\bullet\prime}_t\right)\right)\right|\left|\sqrt{T}(\hat{\beta}-\beta)\right|.
\end{aligned}
$$

We know that $\left|\sqrt{T}(\hat{\beta}-\beta)\right|=O_p(1)$. From Lemma SA5, $\sup_{r\in[0,1]}\left|T^{-1}\sum_{t=1}^{[rT]}\left(x^{\bullet}_tx^{\bullet\prime}_t-E^{\bullet}\left(x^{\bullet}_tx^{\bullet\prime}_t\right)\right)\right|=o_{p^{\bullet}}(1)$. Hence $\sup_{r\in[0,1]}\left|\Gamma^{\bullet}_{2T}(r)\right|=o_{p^{\bullet}}(1)$, which completes the proof of Theorem SA1. □

References

Davidson, J. (2002). *Stochastic Limit Theory*. Advanced Texts in Econometrics. Oxford University Press.

Fitzenberger, B. (1997). The moving blocks bootstrap and robust inference for linear least squares and quantile regressions. *Journal of Econometrics*, 82:235–287.

Gonçalves, S. and Vogelsang, T. J. (2011). Block bootstrap HAC robust tests: the sophistication of the naive bootstrap. *Econometric Theory*, 27:745–791.

Hansen, B. E. (1991). Strong laws for dependent heterogeneous processes. *Econometric Theory*, 7:213–221.

Hansen, B. E. (1992). Erratum. *Econometric Theory*, 8:421–422.

Kiefer, N. M. and Vogelsang, T. J. (2005). A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. *Econometric Theory*, 21:1130–1164.