Web Appendix of
On Varying-coefficient Independence Screening for High-dimensional Varying-coefficient Models

Rui Song, Feng Yi and Hui Zou

North Carolina State University and University of Minnesota

Summary

This web appendix contains two parts. In the first part, we provide the proofs of all the theoretical results in the main paper. In the second part, we provide figures for the real data analysis in Section 5.

1. Proofs

Proof of Lemma 1.

By the property of the least-squares, $E(Y(T)-\tilde{\beta}_l(T)X_l(T))\tilde{\beta}_l(T)X_l(T)=0$ and $E(Y(T)-\beta_{l0}(T)X_l(T))\tilde{\beta}_l(T)X_l(T)=0$, where T is the random observation time with distribution $F_T(\cdot)$ Therefore,

$$\begin{aligned}&E\tilde{\beta}_l(T)X_l(T)(\beta_{l0}(T)X_l(T)-\tilde{\beta}_l(T)X_l(T))\\&=E(Y(T)-\tilde{\beta}_l(T)X_l(T))\tilde{\beta}_l(T)X_l(T)-E(Y(T)-\beta_{l0}(T)X_l(T))\tilde{\beta}_l(T)X_l(T)=0.\end{aligned}$$

It follows from this and the orthogonal decomposition $\beta_{l0}(T)=\tilde{\beta}_l(T)+(\beta_{10}(T)-\tilde{\beta}_l(T))$ that

$$E\tilde{\beta}_l(T)^2=E\beta_{l0}(T)^2-E(\beta_{l0}(T)-\tilde{\beta}_l(T))^2.$$

By Conditions A, C and definition of ρ_n,

$$\begin{aligned}E\beta_{l0}(T)^2&\geq M_1\int_{\mathcal{T}}\beta_{l0}(t)^2dt\geq M_1L_1c_1n^{-2\kappa},\\E(\beta_{l0}(T)-\tilde{\beta}_l(T))^2&\leq M_2\int_{\mathcal{T}}(\beta_{l0}(t)-\tilde{\beta}_l(t))^2dt\leq M_2L_2\rho_n^2.\end{aligned}$$

The desired result follows from Triangle inequality. □

For functions $g_l^{(1)}(t)$ and $g_l^{(2)}(t)$, define the following inner product

$$\langle g_l^{(1)},g_l^{(2)}\rangle_n=\frac{1}{n}\sum_i w_i\sum_j\Big(X_{li}(t_{ij})g_l^{(1)}(t_{ij})\Big)\Big(X_{li}(t_{ij})g_l^{(2)}(t_{ij})\Big).$$

Its population version is

$$\langle g_l^{(1)},g_l^{(2)}\rangle=E\langle g_l^{(1)},g_l^{(2)}\rangle_n=E\Big(X_{li}(T)g_l^{(1)}(T)\Big)\Big(X_{li}(T)g_l^{(2)}(T)\Big),$$

where T is the random observation time with distribution $F_T(\cdot)$. We denote the corresponding norms by $\|\cdot\|_n$ and $\|\cdot\|$. Let $\mathbf{g}_{\boldsymbol{\gamma}}=\sum_k\gamma_{lk}B_{lk}$, note that $\boldsymbol{\gamma}'(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l/n)\boldsymbol{\gamma}=\|\mathbf{g}_{\boldsymbol{\gamma}}\|_n^2$.

The following result from Huang, Wu and Zhou (2004) will be used.

Lemma 2. (Lemma A.2 in Huang, Wu and Zhou (2004)) Let $\mathbb{G}$ denote the collection of vectors of functions $\mathbf{g}=(g_1,\ldots,g_p)'$ with $g_l\in\mathbb{G}_l$, then for $s>0$, there exist positive constants C_1, C_2 such that

$$P\Big(\sup_{\mathbf{g}_1,\mathbf{g}_2\in\mathbb{G}}\frac{|\langle\mathbf{g}_1,\mathbf{g}_2\rangle_n-\langle\mathbf{g}_1,\mathbf{g}_2\rangle|}{\|\mathbf{g}_1\|\|\mathbf{g}_2\|}>s\Big)\le C_1K_m^2\exp\Big(-C_2\frac{n}{K_m}\frac{s^2}{1+s}\Big).$$

Lemma 3. Under Conditions A and B, for $s>0$, there exist constants C_1 and C_2 such that

$$P\Big(\frac{\Big|\lambda_{\min}(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)-\lambda_{\min}(E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)\Big|}{\lambda_{\max}(E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)}>s\Big)\le C_1K_l^2\exp\Big(-C_2\frac{n}{K_l}\frac{s^2}{1+s}\Big). \quad (1.1)$$

More over, there exists some positive constants c_4 and c_5 such that

$$P\left\{\Big|\lambda_{\max}\left((\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-1}\right)-\lambda_{\max}\left((E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-1}\right)\Big|\ge\lambda_{\max}\left((E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-1}\right)\right\}$$
$$\le c_5K_l^2\exp\Big(-c_4nK_l^{-1}\Big). \quad (1.2)$$

Proof of Lemma 3. For $g_1=\sum_k g_{1k}B_{1k}(t)$ and $g_2=\sum_k g_{2k}B_{2k}(t)\in\mathcal{G}_l$, by Lemma 2, for $s>0$,

$$P\Big(\sup_{g_1,g_2\in\mathbb{G}_l}\frac{|\langle g_1,g_2\rangle_n-\langle g_1,g_2\rangle|}{\|g_1\|\|g_2\|}>s\Big)\le C_1K_l^2\exp\Big(-C_2\frac{n}{K_l}\frac{s^2}{1+s}\Big). \quad (1.3)$$

Since for any $g\in\mathbb{G}_l$,

$$\frac{|\|g\|_n^2-\|g\|^2|}{\|g\|^2}\ge\frac{|\|g\|_n^2-\|g\|^2|}{\sup_{g\in\mathbb{G}_l}\|g\|^2},$$

we have

$$\sup_{g_1,g_2\in\mathbb{G}_l}\frac{|\langle g_1,g_2\rangle_n-\langle g_1,g_2\rangle|}{\|g_1\|\|g_2\|}\ge\sup_{g\in\mathbb{G}_l}\frac{|\langle g,g\rangle_n-\langle g,g\rangle|}{\|g\|\|g\|}\ge\sup_{g\in\mathbb{G}_l}\frac{|\|g\|_n^2-\|g\|^2|}{\sup_{g\in\mathbb{G}_l}\|g\|^2}. \quad (1.4)$$

Since for any $g\in\mathcal{G}_l$,

$$\inf_{g\in\mathcal{G}_l}(\|g\|_n^2-\|g\|^2)\le\inf_{g\in\mathcal{G}_l}(\|g\|_n^2)-\inf_{g\in\mathcal{G}_l}(\|g\|^2).$$

By switching the roles of $\|g\|_n^2$ and $\|g\|^2$, we also have

$$\inf_{g\in\mathcal{G}_l}(\|g\|^2-\|g\|_n^2)\le\inf_{g\in\mathcal{G}_l}(\|g\|^2)-\inf_{g\in\mathcal{G}_l}(\|g\|_n^2).$$

In other words,

$$\left|\inf_{g\in\mathcal{G}_l}(\|g\|_n^2) - \inf_{g\in\mathcal{G}_l}(\|g\|^2)\right| \le \max\{|\inf_{g\in\mathcal{G}_l}(\|g\|_n^2 - \|g\|^2)|, |\inf_{g\in\mathcal{G}_l}(\|g\|^2 - \|g\|_n^2)|\}. \quad (1.5)$$

The right-hand-side of (1.5) is further bounded from above as

$$\max\{|\inf_{g\in\mathcal{G}_l}(\|g\|_n^2 - \|g\|^2)|, |\inf_{g\in\mathcal{G}_l}(\|g\|^2 - \|g\|_n^2)|\} \le \sup_{g\in\mathcal{G}_l}\left|\|g\|_n^2 - \|g\|^2\right|. \quad (1.6)$$

It follows from (1.4)–(1.6) that

$$\sup_{g_1,g_2\in\mathbb{G}_l} \frac{\left|\langle g_1, g_2\rangle_n - \langle g_1, g_2\rangle\right|}{\|g_1\|\|g_2\|} \ge \frac{\left|\inf_{g\in\mathcal{G}_l}(\|g\|_n^2) - \inf_{g\in\mathcal{G}_l}(\|g\|^2)\right|}{\sup_{g\in\mathbb{G}_l}\|g\|^2}. \quad (1.7)$$

For $g \in \mathcal{G}_l$, by the property of B-splines in de Boor (1978), there exist positive constants E_1, E_2 such that

$$\frac{E_1}{K_l}\sum_k \gamma_{lk}^2 \le \int_{\mathcal{T}}(\sum_k \gamma_{lk}B_{lk}(t))^2 dt \le \frac{E_2}{K_l}\sum_k \gamma_{lk}^2,\ \gamma_{lk}\in\mathbb{R},\ k = 1,\ldots,K_l.$$

Hence by Condition A, there exist positive D_1 and D_2 such that

$$D_1/K_l \le \lambda_{\min}(E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l/n) \le \sup_{g\in\mathcal{G}_l}\|g\|^2 = \lambda_{\max}(E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l/n) \le D_2/K_l. \quad (1.8)$$

Since $\|g\|_n^2 = \boldsymbol{\gamma}'(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l/n)\boldsymbol{\gamma}$,

$$\inf_{g\in\mathcal{G}_l}(\|g\|_n^2) = \min_{\boldsymbol{\gamma}}\boldsymbol{\gamma}'(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l/n)\boldsymbol{\gamma} = \lambda_{\min}(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l/n),\ \text{and}$$
$$\inf_{g\in\mathcal{G}_l}(\|g\|^2) = \min_{\boldsymbol{\gamma}}\boldsymbol{\gamma}'(E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l/n)\boldsymbol{\gamma} = \lambda_{\min}E(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l/n),$$

the right-hand-side of (1.7) can be expressed as

$$\frac{\left|\inf_{g\in\mathcal{G}_l}(\|g\|_n^2) - \inf_{g\in\mathcal{G}_l}(\|g\|^2)\right|}{\sup_{g\in\mathbb{G}_l}\|g\|^2} = \frac{\left|\lambda_{\min}(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l/n) - \lambda_{\min}(E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l/n)\right|}{\lambda_{\max}(E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l/n)}. \quad (1.9)$$

Now it follows from (1.9) and (1.3) that for $s > 0$,

$$\begin{aligned} & P\Big(\frac{\Big|\lambda_{\min}(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l/n) - \lambda_{\min}(E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l/n)\Big|}{\lambda_{\max}(E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l/n)} > s\Big) \\ \leq\ & C_1 K_l^2 \exp\Big(-C_2 \frac{n}{K_l}\frac{s^2}{1+s}\Big) \end{aligned} \tag{1.10}$$

To prove the second inequality, letting $s = 1/2D_1D_2^{-1}$ in (1.10), we have

$$\begin{aligned} & P\Big(\Big|\lambda_{\min}(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l) - \lambda_{\min}(E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)\Big| > 1/2D_1K_l^{-1}\Big) \\ \leq\ & C_1 K_l^2 \exp\Big(-c_4 n K_l^{-1}\Big), \end{aligned} \tag{1.11}$$

for some positive constant c_4. By (1.8), it follows that

$$\begin{aligned} & P(|\lambda_{\min}(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l/n) - \lambda_{\min}(E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l/n)| \geq 1/2\lambda_{\min}(E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l/n)) \\ \leq\ & C_1 K_l^2 \exp\Big(-c_4 n K_l^{-1}\Big), \end{aligned} \tag{1.12}$$

for some positive constants C_1 and c_4.

The second part of the lemma thus follows from the fact that $\lambda_{\min}(\mathbf{H})^{-1} = \lambda_{\max}(\mathbf{H}^{-1})$, if we establish that

$$\begin{aligned} & P\left(\left|\Big\{\lambda_{\min}(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)\Big\}^{-1} - \Big\{\lambda_{\min}(E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)\Big\}^{-1}\right| \geq \Big\{\lambda_{\min}(E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)\Big\}^{-1}\right) \\ \leq\ & C_1 K_m^2 \exp\Big(-c_4 n K_m^{-1}\Big), \end{aligned} \tag{1.13}$$

by using (1.12). A similar argument in Lemma 5 of Fan et al. (2011) applies hence we omit the details. The desired result follows be letting $c_5 = C_1$. □

Lemma 4. Under Conditions A, B, D and E, for any $\delta > 0$,

$$\begin{aligned} & P(\|\mathbf{U}_l'\mathbf{W}\mathbf{Y} - E\mathbf{U}_l'\mathbf{W}\mathbf{Y}\|^2 \geq K_l\delta^2) \\ \leq\ & 4K_l \exp(-\delta^2/2(c_7N^2\omega^2nK_l^{-1} + c_8N\omega\delta)), \end{aligned} \tag{1.14}$$

for some positive constants c_7 and c_8.

Proof of Lemma 4. . Recall that $\mathbf{U}_{li} = (\mathbf{U}_{li1}, \ldots, \mathbf{U}_{lin_i})'$, $\mathbf{W}_{li} = w_iI_{n_i}$ and $\boldsymbol{\varepsilon}_i = (\epsilon_{i1}, \ldots, \epsilon_{in_i})'$. $\tilde{Y}_{ij} = \mathbf{X}_i(t_{ij})'\boldsymbol{\alpha}(t_{ij})$, $\tilde{\mathbf{Y}}_i = (\tilde{Y}_{i1}, \ldots, \tilde{Y}_{in_i})'$, $\tilde{\mathbf{Y}} = (\tilde{\mathbf{Y}}_1, \ldots, \tilde{\mathbf{Y}}_n)'$.

We have

$$\mathbf{U}_l'\mathbf{W}\mathbf{Y} = \sum_i \mathbf{U}_{li}'\mathbf{W}_{li}\mathbf{Y}_i = \sum_i (\mathbf{U}_{li}'\mathbf{W}_{li}\tilde{\mathbf{Y}}_i + \mathbf{U}_{li}'\mathbf{W}_{li}\boldsymbol{\varepsilon}_i).$$

Note that

$$\mathbf{U}_{li}'\mathbf{W}_i\mathbf{Y}_i = w_i \sum_j \mathbf{U}_{lij}Y_{ij} = w_i \sum_j X_{lij}Y_{ij}(B_{l1}(t_{ij}), \ldots, B_{lK_l}(t_{ij}))', \text{ and}$$
$$\sum_i \mathbf{U}_{li}'\mathbf{W}_i\mathbf{Y}_i = \Big(\sum_i w_i \sum_j X_{lij}Y_{ij}B_{l1}(t_{ij}), \ldots, \sum_i w_i \sum_j X_{lij}Y_{ij}B_{lK_l}(t_{ij})\Big)'.$$

Let $T_{lkij} = B_{lk}(t_{ij})X_{lij}Y_{ij} - EB_{lk}(t_{ij})X_{lij}Y_{ij}$. Hence

$$\mathbf{U}_l'\mathbf{W}\mathbf{Y} - E\mathbf{U}_l'\mathbf{W}\mathbf{Y} = \Big(\sum_i w_i \sum_j T_{l1ij}, \ldots, \sum_i w_i \sum_j T_{lK_lij}\Big)'. \tag{1.15}$$

Since $Y_{ij} = \tilde{Y}_{ij} + \epsilon_{ij}$, we can write $T_{lkij} = T_{lkij1} + T_{lkij2}$, where

$$T_{lkij1} = B_{lk}(t_{ij})X_{lij}\tilde{Y}_{ij} - EB_{lk}(t_{ij})X_{lij}\tilde{Y}_{ij},$$

and $T_{lkij2} = B_{lk}(t_{ij})X_{lij}\epsilon_{ij}$.

By Properties of B-splines in de Boor (1978), $B_{lk}(t) \geq 0$, $\sum_{k=1}^{K_l} B_{lk}(t) = 1$. $EB_{lk}(t_{ij}) \leq C_3K_l^{-1}$ for some $C_3 > 0$. By Conditions A, B, E, we have

$$|\sum_j T_{lkij1}| \leq 2n_iB_1M_3,$$
$$\text{var}(\sum_j T_{lkij1}) \leq n_i^2EB_{lk}^2(t_{ij})X_{lij}^2\tilde{Y}_{ij}^2 \leq n_i^2C_3B_1^2M_3^2K_l^{-1}. \tag{1.16}$$

By Bernstein's inequality, for any $\delta_1 > 0$,

$$\begin{aligned} & P(\Big|\sum_i w_i \sum_j T_{lkij1}\Big| > \delta_1) \\ \leq \quad & 2\exp\Big(-\frac{1}{2}\frac{\delta_1^2}{nN^2w^2C_3M_3^2B_1^2K_l^{-1} + 2NwB_1M_3\delta_1/3}\Big), \end{aligned} \tag{1.17}$$

where $w = \max_i w_i$ and $N = \max_i n_i$. Now we bound the tail of T_{lkij2}. Note that for $k = 1, \dots, K_l$, $m \geq 2$, $E\Big|\sum_j T_{lkij2}\Big|^m = E\Big|\sum_j X_{lij}\epsilon_{ij}B_{lk}(t_{ij})\Big|^m$

$$\leq \sum_{\substack{\{b_r\}\geq 0 \\ \sum_{r=1}^{n_i} b_r = m}} \begin{pmatrix} m \\ b_1 \end{pmatrix} \begin{pmatrix} m - b_1 \\ b_2 \end{pmatrix} \cdots \begin{pmatrix} m - \sum_{r=1}^{n_i-2} b_r \\ b_{n_i-1} \end{pmatrix} E\Big|T_{lki12}\Big|^{b_1} \cdots E\Big|T_{lkin_i2}\Big|^{b_{n_i}}.$$

By Condition E, for any $b_r \geq 1$, $r = 1, \dots, n_i$, $\sum_{r=1}^{n_i} b_r = m \geq 2$,

$$\begin{aligned} &E\Big|T_{lkir2}\Big|^{b_r} = E\Big|X_{lir}\epsilon_{ir}B_{lk}(t_{ir})\Big|^{b_r} \leq E\Big(\Big|X_{lir}^{b_r}B_{lk}(t_{ir})^{b_r}\Big|E(|\epsilon_{ir}|^{b_r}|X_{lir})\Big) \\ &\leq b_r!B_2^{-b_r}C_3E\exp(B_2|\varepsilon_i||\mathbf{X}_i)M_3^{b_r}K_l^{-1} \leq B_3C_3b_r!B_2^{-b_r}M_3^{b_r}K_l^{-1}. \end{aligned}$$

Hence $E\Big|w_i\sum_j X_{lij}\epsilon_{ij}B_{lk}(t_{ij})\Big|^m \leq m!w_i^m n_i^m B_4K_l^{-1}B_2^{-m}M_3^m$, for some constant $B_4 > 0$.

By Bernstein's inequality, for any $\delta_2 > 0$,

$$\begin{aligned} &P\Big(\Big|\sum_{i=1}^n w_i\sum_j X_{lij}\epsilon_{ij}B_{lk}(t_{ij})\Big| > \delta_2\Big) \\ \leq\;& 2\exp\Big(-\frac{1}{2}\frac{\delta_2^2}{2nw^2N^2B_2^{-2}M_3^2B_4K_l^{-1} + wNB_2^{-1}M_3\delta_2}\Big). \end{aligned} \tag{1.18}$$

Combining (1.17) and (1.18), taking $c_7 = \max(C_3M_3^2B_1^2, 2B_2^{-2}M_3^2B_4)$ and $c_8 = \max(2/3B_1M_3, B_2^{-1}M_3)$,

$$P\Big(\Big|\sum_{i=1}^n w_i\sum_j T_{lkij}\Big| > \delta\Big) \leq 4\exp\Big(-\frac{1}{2}\frac{\delta^2}{c_7N^2\omega^2nK_l^{-1} + c_8N\omega\delta}\Big). \tag{1.19}$$

The desired result now follows from the union bound of probability,

$$P(\|\mathbf{U}_l'\mathbf{W}\mathbf{Y} - E\mathbf{U}_l'\mathbf{W}\mathbf{Y}\|^2 \geq K_l\delta^2) \leq 4K_l\exp(-\delta^2/2(c_7N^2\omega^2nK_l^{-1} + c_8N\omega\delta)).$$

□

Throughout the rest of the proof, for any matrix $\mathbf{A}$, let $\|\mathbf{A}\| = \sqrt{\lambda_{\max}(\mathbf{A}^T\mathbf{A})}$ be the operator norm and $\|\mathbf{A}\|_\infty = \max_{i,j}|A_{ij}|$ be the infinity norm. The next lemma is about the tail probability of the eigenvalues of the design matrix.

LEMMA 5. Under Conditions A, B, D and E, for any $\delta > 0$,

$$P(\Big|\|\hat{\gamma}_l\|^2 - \|\tilde{\gamma}_l\|^2\Big| \geq c_2 K_l n^{-2\kappa}) \leq$$
$$(8K_l + 2K_l^2)\exp(-c_3 N^{-2}\omega^{-2} n^{1-4\kappa} K_l^{-3}) + 6K_l^2 \exp\Big(-c_4 n K_l^{-1}\Big).$$

Proof of Lemma 5. Recall that

$$\tilde{\gamma}_l = \Big(E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l\Big)^{-1} E\mathbf{U}_l'\mathbf{W}\mathbf{Y}, \text{ and } \tilde{\beta}_l(t) = \mathbf{B}_l(t)'\tilde{\gamma}_l.$$

Let $\mathbf{a}_n = \mathbf{U}_l'\mathbf{W}\mathbf{Y}$, $\mathbf{B}_n = (\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2}$, $\mathbf{a} = E\mathbf{U}_l'\mathbf{W}\mathbf{Y}$ and $\mathbf{B} = (E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2}$. By some algebra,

$$\mathbf{a}_n'\mathbf{B}_n\mathbf{a}_n - \mathbf{a}'\mathbf{B}\mathbf{a} = (\mathbf{a}_n - \mathbf{a})'\mathbf{B}_n(\mathbf{a}_n - \mathbf{a}) + 2(\mathbf{a}_n - \mathbf{a})'\mathbf{B}_n\mathbf{a} + \mathbf{a}_n'(\mathbf{B}_n - \mathbf{B})\mathbf{a},$$

we have

$$\|\hat{\gamma}_l\|^2 - \|\tilde{\gamma}_l\|^2 = S_1 + S_2 + S_3, \tag{1.20}$$

where

$$S_1 = \Big(\mathbf{U}_l'\mathbf{W}\mathbf{Y} - E\mathbf{U}_l'\mathbf{W}\mathbf{Y}\Big)'(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2}\Big(\mathbf{U}_l'\mathbf{W}\mathbf{Y} - E\mathbf{U}_l'\mathbf{W}\mathbf{Y}\Big),$$
$$S_2 = 2\Big(\mathbf{U}_l'\mathbf{W}\mathbf{Y} - E\mathbf{U}_l'\mathbf{W}\mathbf{Y}\Big)'(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2}E\mathbf{U}_l'\mathbf{W}\mathbf{Y},$$
$$S_3 = (E\mathbf{U}_l'\mathbf{W}\mathbf{Y})'\Big((\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2} - (E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2}\Big)E\mathbf{U}_l'\mathbf{W}\mathbf{Y}.$$

Note that

$$S_1 \leq \lambda_{\max}((\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2}) \cdot \|\mathbf{U}_l'\mathbf{W}\mathbf{Y} - E\mathbf{U}_l'\mathbf{W}\mathbf{Y}\|^2. \tag{1.21}$$

By Lemma 4,

$$P(\|\mathbf{U}_l'\mathbf{W}\mathbf{Y} - E\mathbf{U}_l'\mathbf{W}\mathbf{Y}\|^2 \geq K_l\delta^2) \leq 4K_l \exp(-\delta^2/2(c_7 N^2\omega^2 n K_l^{-1} + c_8 N\omega\delta)).$$

Recall the result in Lemma 3 that,

$$P\left\{\Big|\lambda_{\max}\left((\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-1}\right) - \lambda_{\max}\left((E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-1}\right)\Big| \geq \lambda_{\max}\left((E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-1}\right)\right\}$$
$$\leq C_1 K_m^2 \exp\Big(-c_4 n K_m^{-1}\Big).$$

Since by the property of B-spline in de Boor (1978),

$$\lambda_{\max}\Big((E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2}\Big) \le n^{-2}D_1^{-2}K_m^2,$$

it follows that

$$P\left\{\lambda_{\max}\Big((\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2}\Big) \ge 4D_1^{-2}K_l^2 n^{-2}\right\} \le C_1 K_m^2 \exp\Big(-c_4 n K_m^{-1}\Big). \quad (1.22)$$

Combining (1.21)–(1.22) and the union bound of probability, we have

$$\begin{aligned} & P(S_1 \ge 4n^{-2}D_1^{-2}K_m^3\delta^2) \\ \le\ & 4K_l \exp(-\delta^2/2(c_7 N^2\omega^2 n K_m^{-1} + c_8 N\omega\delta)) \\ & + C_1 K_m^2 \exp\Big(-c_4 n K_m^{-1}\Big). \end{aligned} \quad (1.23)$$

To bound S_2, we note that

$$\begin{aligned} |S_2| &\le 2\|\mathbf{U}_l'\mathbf{W}\mathbf{Y} - E\mathbf{U}_l'\mathbf{W}\mathbf{Y}\| \cdot \|(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2}E\mathbf{U}_l'\mathbf{W}\mathbf{Y}\| \\ &\le 2\|\mathbf{U}_l'\mathbf{W}\mathbf{Y} - E\mathbf{U}_l'\mathbf{W}\mathbf{Y}\| \cdot \|(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2}\| \cdot \|E\mathbf{U}_l'\mathbf{W}\mathbf{Y}\|. \end{aligned} \quad (1.24)$$

Since by Condition D,

$$\|E\mathbf{U}_l'\mathbf{W}\mathbf{Y}\|^2 = \sum_{k=1}^{K_l}\Big(E\sum_i w_i \sum_j B_{lk}(t_{ij})X_{lij}Y_{ij}\Big)^2 \quad (1.25)$$

$$= \sum_{k=1}^{K_l}\Big(E\sum_i w_i \sum_j B_{lk}(t_{ij})X_{lij}\tilde{Y}_{ij}\Big)^2 \le n^2 B_1^2 M_3^2 N^2 w^2 \sum_{k=1}^{K_l} EB_{lk}^2 \le C_4 N^2\omega^2 n^2,$$

for some $C_4 > 0$, it follows from (1.14), (1.22), (1.24), (1.25) and the union bound of probability that

$$\begin{aligned} & P(|S_2| \ge 8D_1^{-2}C_4^{1/2}n^{-1}K_l^{5/2}\delta) \\ \le\ & 4K_l \exp(-\delta^2/2(c_7 N^2\omega^2 n K_l^{-1} + c_8 N\omega\delta)) \\ & + C_1 K_l^2 \exp\Big(-c_4 n K_l^{-1}\Big). \end{aligned} \quad (1.26)$$

Now we bound S_3. Note that

$$S_3 = (E\mathbf{U}_l'\mathbf{W}\mathbf{Y})'\Big((\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2} - (E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2}\Big)E\mathbf{U}_l'\mathbf{W}\mathbf{Y}.$$

It can be further expressed as

$$\begin{aligned} S_3 &= (E\mathbf{U}_l'\mathbf{W}\mathbf{Y})'(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2}\Big((E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^2 - (\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^2\Big)\cdot \\ &\quad (E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2}E\mathbf{U}_l'\mathbf{W}\mathbf{Y}. \end{aligned} \tag{1.27}$$

By the fact that $\|\mathbf{AB}\| \le \|\mathbf{A}\|\cdot\|\mathbf{B}\|$, we have

$$\begin{aligned} |S_3| &\le \|(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^2 - (E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^2\|\cdot\|(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2}\|\cdot \\ &\quad \|(E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2}\|\cdot\|E\mathbf{U}_l'\mathbf{W}\mathbf{Y}\|^2. \end{aligned} \tag{1.28}$$

Since for any symmetric matrix $\|\mathbf{A}+\mathbf{B}\| \le \|\mathbf{A}\| + \|\mathbf{B}\|$, and $\mathbf{U}_l'\mathbf{W}\mathbf{U}_l$ is symmetric,

$$\begin{aligned} &\|(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^2 - (E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^2\| \qquad (1.29)\\ &= \|(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l - E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l - E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l + 2E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)\| \\ &\le \|\mathbf{U}_l'\mathbf{W}\mathbf{U}_l - E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l\|^2 + 2\|\mathbf{U}_l'\mathbf{W}\mathbf{U}_l - E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l\|\cdot\|E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l\|. \end{aligned}$$

Since for K_l-dimensional square matrix $\mathbf{D}$, $\|\mathbf{D}\| \le K_l\|\mathbf{D}\|_\infty$, we have

$$\big\|\mathbf{U}_l'\mathbf{W}\mathbf{U}_l - E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l\big\| \le K_l\|\mathbf{U}_l'\mathbf{W}\mathbf{U}_l - E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l\|_\infty. \tag{1.30}$$

We now use Bernstein's inequality to bound the right-hand side of (1.30). Since $\|B_{lk}\|_\infty \le 1$, we have that for $k, m \le K_l$, $j \le n_i$,

$$\mathrm{var}\Big(\sum_{j=1}^{n_i} B_{lk}(t_{ij})B_{lm}(t_{ij})X_{lij}^2\Big) \le n_i M_3^2 E B_{lk}^2(t_{ij})B_{lm}^2(t_{ij}) \le n_i M_3^2 E B_{lk}^2(t_{ij}) \le N M_3^2 C_3 K_l^{-1}.$$

By Bernstein's inequality, for any $\delta > 0$,

$$P\Big(|\sum_{i=1}^{n}\omega_i\sum_{j=1}^{n_i}X_{lij}^2B_{lk}(t_{ij})B_{lm}(t_{ij}) - E\sum_{i=1}^{n}\omega_i\sum_{j=1}^{n_i}X_{lij}^2B_{lk}(t_{ij})B_{lm}(t_{ij})| > \delta\Big) \leq$$
$$2\exp\Big\{-\frac{\delta^2}{2(Nw^2M_3^2C_3nK_l^{-1} + 2wNM_3^2\delta/3)}\Big\}. \quad (1.31)$$

By (1.29) and (1.31), we have

$$P\Big(\|(\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^2 - (E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^2\| > K_l^2\delta^2 + 2D_2n\delta\Big) \leq$$
$$2\exp\Big\{-\frac{\delta^2}{2(Nw^2M_3^2C_3nK_l^{-1} + 2wNM_3^2\delta/3)}\Big\}. \quad (1.32)$$

By(1.22), (1.25), (1.28), (1.30)-(1.32) and the union bound of probability, it follows that

$$\begin{aligned} & P(|S_3| \geq 8C_4D_1^{-4}D_2K_l^4n^{-1}\delta + 4K_l^6D_1^{-4}n^{-2}\delta^2) \\ \leq\ & 2K_l^2\exp(-\delta^2/2(c_7N^2\omega^2nK_l^{-1} + c_8N\omega\delta)) \\ & +2K_l^2C_1\exp\Big(-c_4nK_l^{-1}\Big). \quad (1.33) \end{aligned}$$

It follows from (1.20), (1.23), (1.26), (1.33) and the union bound of probability that for some positive constants c_9, c_{10}, c_{11} and c_{12},

$$\begin{aligned} & P\left(\Big|\|\hat{\gamma}_l\|^2 - \|\tilde{\gamma}_l\|^2\Big| \geq c_9K_l^3\delta^2/n^2 + c_{10}K_l^{5/2}\delta/n + c_{11}K_l^4\delta/n + c_{12}K_l^6\delta^2/n^2\right) \\ \leq\ & (8K_l + 2K_l^2)\exp(-\delta^2/2(c_7N^2\omega^2nK_l^{-1} + c_8N\omega\delta)) \\ & +6K_l^2C_1\exp\Big(-c_4nK_l^{-1}\Big). \quad (1.34) \end{aligned}$$

In (1.34), let $c_9K_l^3\delta^2/n^2 + c_{10}K_l^{5/2}\delta/n + c_{11}K_l^4\delta/n + c_{12}K_l^6\delta^2/n^2 = c_2K_ln^{-2\kappa}$ for any given $c_2 > 0$, i.e., taking $\delta = n^{1-2\kappa}K_l^{-2}c_2/c_{11}$, there exist some positive constants c_3 and c_4 such that

$$\begin{aligned} & P(\Big|\|\hat{\gamma}_l\|^2 - \|\tilde{\gamma}_l\|^2\Big| \geq c_2K_ln^{-2\kappa}) \\ & \leq (8K_l + 2c_5K_l^2)\exp(-c_3n^{1-4\kappa}K_l^{-3}) + 6c_5K_l^2\exp\Big(-c_4nK_l^{-1}\Big). \ \square \end{aligned}$$

Proof of Theorem 1.

Note that $|\int_{\mathcal{T}} \hat{\beta}_l(t)^2 dt - \int_{\mathcal{T}} \tilde{\beta}_l(t)^2 dt| =$

$$\Big|\hat{\gamma}_l' \int_{\mathcal{T}} \mathbf{B}_l(t)\mathbf{B}_l(t)' dt \hat{\boldsymbol{\gamma}}_l - \tilde{\gamma}_l' \int_{\mathcal{T}} \mathbf{B}_l(t)\mathbf{B}_l(t)' dt \tilde{\boldsymbol{\gamma}}_l\Big| = \Big|(\hat{\gamma}_l - \tilde{\gamma}_l)' \int_{\mathcal{T}} \mathbf{B}_l(t)\mathbf{B}_l(t)' dt (\hat{\gamma}_l + \tilde{\gamma}_l)\Big|.$$

Since by B-splines property, $\lambda_{\max}\{\int_{\mathcal{T}} \mathbf{B}_l(t)\mathbf{B}_l(t)' dt\} \le C_5 K_l^{-1}$ for some positive constant C_5, the above can be bounded by

$$\le \lambda_{\max}\{\int_{\mathcal{T}} \mathbf{B}_l(t)\mathbf{B}_l(t)' dt\}\Big|(\hat{\boldsymbol{\gamma}}_l - \tilde{\boldsymbol{\gamma}}_l)'(\hat{\boldsymbol{\gamma}}_l + \tilde{\boldsymbol{\gamma}}_l)\Big| = C_5 K_l^{-1}\Big|\|\hat{\boldsymbol{\gamma}}_l\|^2 - \|\tilde{\boldsymbol{\gamma}}_l\|^2\Big|. \tag{1.35}$$

Therefore on the event

$$B_n \equiv \{\max_{l \in \mathcal{M}_\star}\Big|\|\hat{\boldsymbol{\gamma}}_l\|^2 - \|\tilde{\boldsymbol{\gamma}}_l\|^2\Big| \le c_1 C_5^{-1} L_1 \xi K_m n^{-2\kappa}/2\},$$

we have

$$\max_{l \in \mathcal{M}_\star} \frac{1}{|\mathcal{T}|}\Big|\int_{\mathcal{T}} \hat{\beta}_l(t)^2 dt - \int_{\mathcal{T}} \tilde{\beta}_l(t)^2 dt\Big| \le c_1 \xi n^{-2\kappa}/2.$$

Meanwhile by Lemma 1, we have

$$\frac{1}{|\mathcal{T}|}\Big|\int_{\mathcal{T}} \hat{\beta}_l(t)^2 dt \ge c_1 \xi n^{-2\kappa}/2, \quad \text{for all } l \in \mathcal{M}_\star. \tag{1.36}$$

Hence, by the choice of ν_n, we have $\mathcal{M}_\star \subset \widehat{\mathcal{M}}_{\nu_n}$. Let $c_2 = c_1 C_5^{-1} L_1 \xi$, the result now follows from Lemma 5 and a simple union bound:

$$P(B_n^c) \le s_n\Big\{(8K_m + 2c_5 K_m^2)\exp\Big(-c_3 N^{-2}\omega^{-2} n^{1-4\kappa} K_m^{-1}\Big) + 6c_5 K_m^2 \exp\Big(-c_4 n K_m^{-1}\Big)\Big\}. \ \square$$

Proof of Theorem 2.

The proof takes two steps. Let $\mathbf{U} = (\mathbf{U}_1, \ldots, \mathbf{U}_p)'$. In the first step we show that

$$\|E\mathbf{U}\mathbf{W}\mathbf{Y}\|^2 = O(\lambda_{\max}(E\mathbf{U}\mathbf{W}\mathbf{U}')). \tag{1.37}$$

Define the joint regression coefficients in the population

$$\boldsymbol{\alpha}_n = \operatorname{argmin}_{\boldsymbol{\alpha}} E(\mathbf{Y} - \mathbf{U}'\boldsymbol{\alpha})'\mathbf{W}(\mathbf{Y} - \mathbf{U}'\boldsymbol{\alpha}).$$

By the score equation of $\boldsymbol{\alpha}_n$, we get

$$E\mathbf{U}\mathbf{W}(\mathbf{Y} - \mathbf{U}'\boldsymbol{\alpha}_n) = 0.$$

Hence, $\|E\mathbf{U}\mathbf{W}\mathbf{Y}\|^2 = \boldsymbol{\alpha}_n' E\mathbf{U}\mathbf{W}\mathbf{U}' E\mathbf{U}\mathbf{W}\mathbf{U}'\boldsymbol{\alpha}_n \leq \lambda_{\max}(\boldsymbol{\Sigma})\boldsymbol{\alpha}_n' E\mathbf{U}\mathbf{W}\mathbf{U}'\boldsymbol{\alpha}_n$. It now follows from the orthogonal decomposition that

$$\mathrm{var}(\mathbf{Y}) = \mathrm{var}(\mathbf{U}'\boldsymbol{\alpha}_n) + \mathrm{var}(\mathbf{Y} - \mathbf{U}'\boldsymbol{\alpha}_n).$$

Since $\mathrm{var}(\mathbf{Y}) = O(1)$, we conclude that $\mathrm{var}(\mathbf{U}'\boldsymbol{\alpha}_n) = O(1)$, i.e.

$$\boldsymbol{\alpha}_n' E\mathbf{U}\mathbf{W}\mathbf{U}'\boldsymbol{\alpha}_n = O(1).$$

For the second step, by definition and the fact that $\|B_{lk}\|_\infty \leq 1$, we have

$$\sum_{l=1}^{p_n} \|\tilde{\gamma}_l\|^2 \leq \max_{1\leq l\leq p_n} \lambda_{\max}\{(E\mathbf{U}_l'\mathbf{W}\mathbf{U}_l)^{-2}\}\|E\mathbf{U}\mathbf{W}\mathbf{Y}\|^2 = O(K_m^2\lambda_{\max}(\boldsymbol{\Sigma})).$$

This implies that the number of $\{l : \|\tilde{\gamma}_l\|^2 > \varepsilon K_m n^{-2\kappa}\}$ can not exceed $O(n^{2\kappa}\lambda_{\max}(\boldsymbol{\Sigma}))$ for any $\varepsilon > 0$. Since on the set

$$B_n = \{\max_{1\leq l\leq p_n}\Big| \|\tilde{\gamma}_l\|^2 - \|\hat{\gamma}_l\|^2\Big| \leq \varepsilon K_m n^{-2\kappa}\},$$

the number of $\{l : \|\hat{\gamma}_l\|^2 > 2\varepsilon K_m n^{-2\kappa}\}$ can not exceed the number of $\{l : \|\tilde{\gamma}_l\|^2 > \varepsilon K_m n^{-2\kappa}\}$, which is bounded by $O\{n^{2\kappa}\lambda_{\max}(\boldsymbol{\Sigma})\}$ Moreover, on set B_n we have

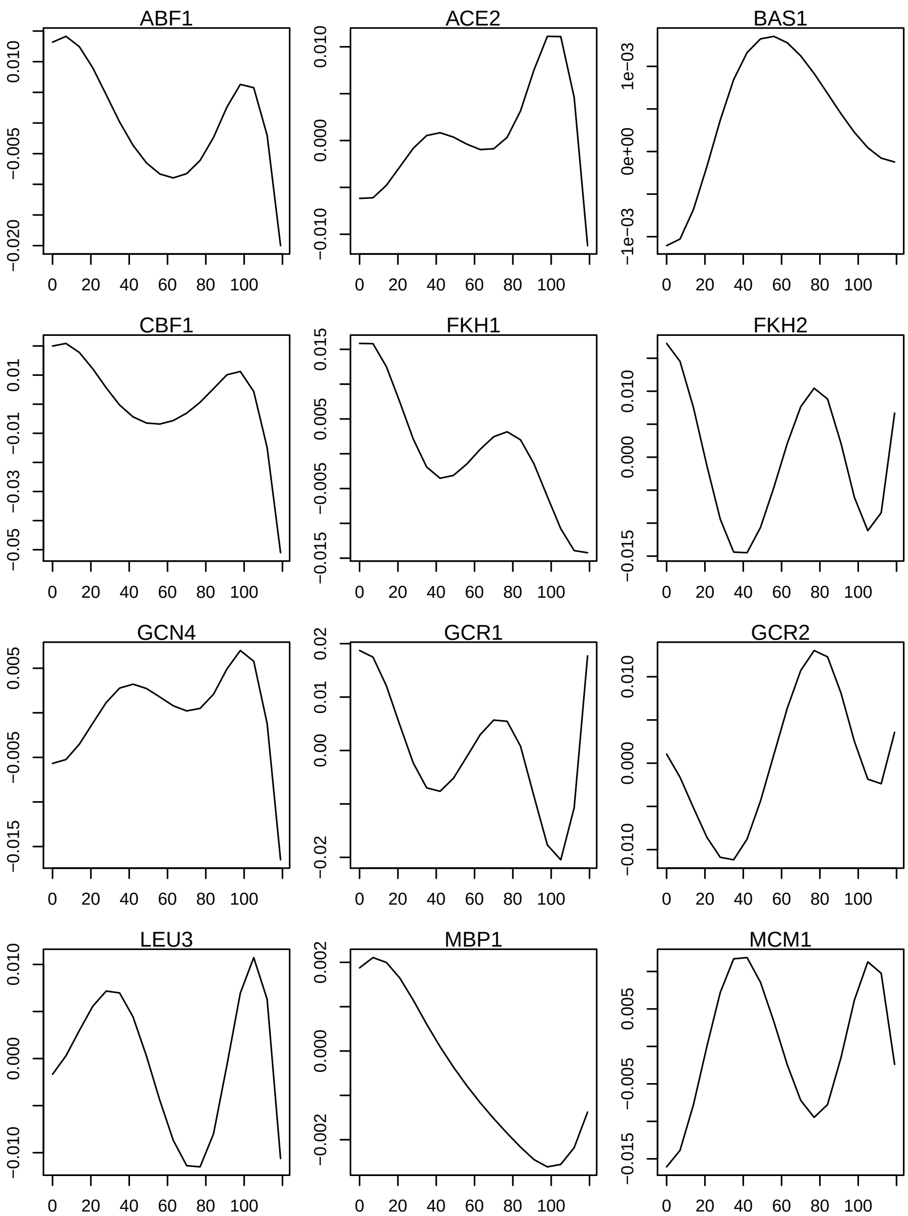
$$\max_{l\in\mathcal{M}_\star} \frac{1}{|\mathcal{T}|}\Big|\int_{\mathcal{T}} \hat{\beta}_l(t)^2dt - \int_{\mathcal{T}} \tilde{\beta}_l(t)^2dt\Big| \leq \varepsilon C_5 L_1^{-1} n^{-2\kappa}.$$

Let $\varepsilon = c_6 C_5^{-1} L_1/2$, the number of $\{l : \frac{1}{|\mathcal{T}|}\int_{\mathcal{T}} \hat{\beta}_l(t)^2dt > c_6 n^{-2\kappa}\}$ can not exceed the number of $\{l : \|\hat{\gamma}_l\|^2 > \varepsilon K_m n^{-2\kappa}\}$, which is bounded by $O\{n^{2\kappa}\lambda_{\max}(\boldsymbol{\Sigma})\}$. Thus we have

$$P[|\widehat{\mathcal{M}}_{\nu_n}| \leq O\{n^{2\kappa}\lambda_{\max}(\boldsymbol{\Sigma})\}] \geq P(B_n).$$

The conclusion follows from Theorem 1. This completes the proof. $\square$.

2. Figures



(Plot continued on the next page)

References

de Boor, C. (1978). *A Practical Guide to Splines.* Springer, New York.

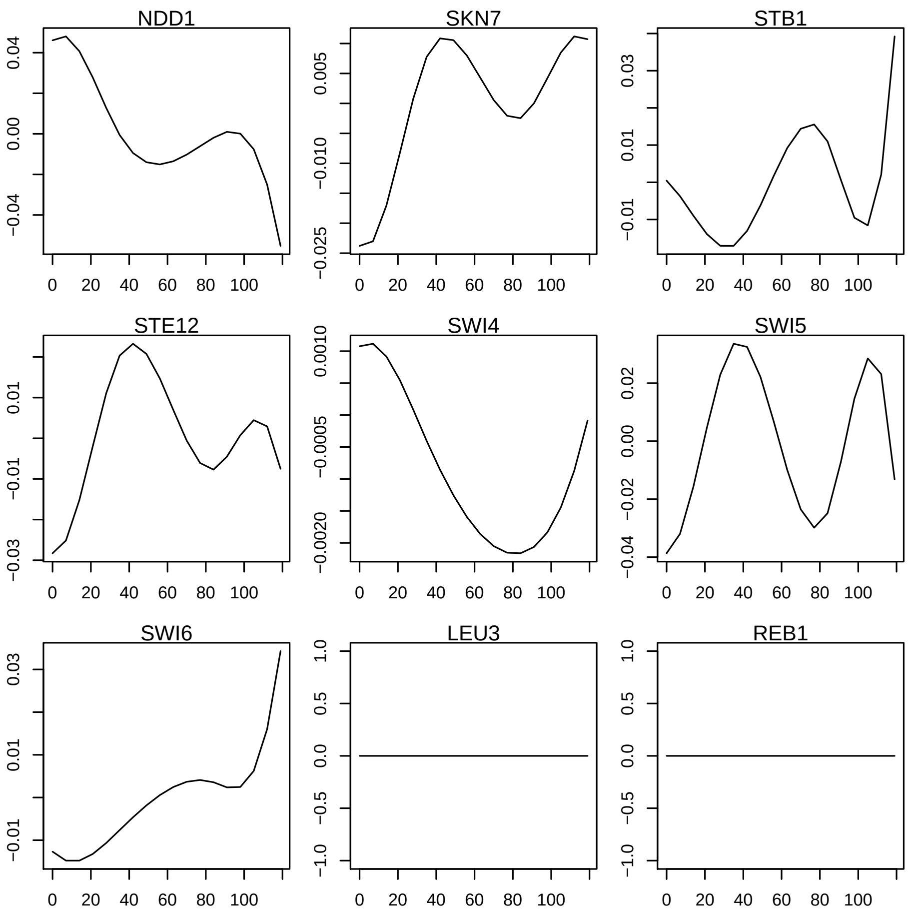
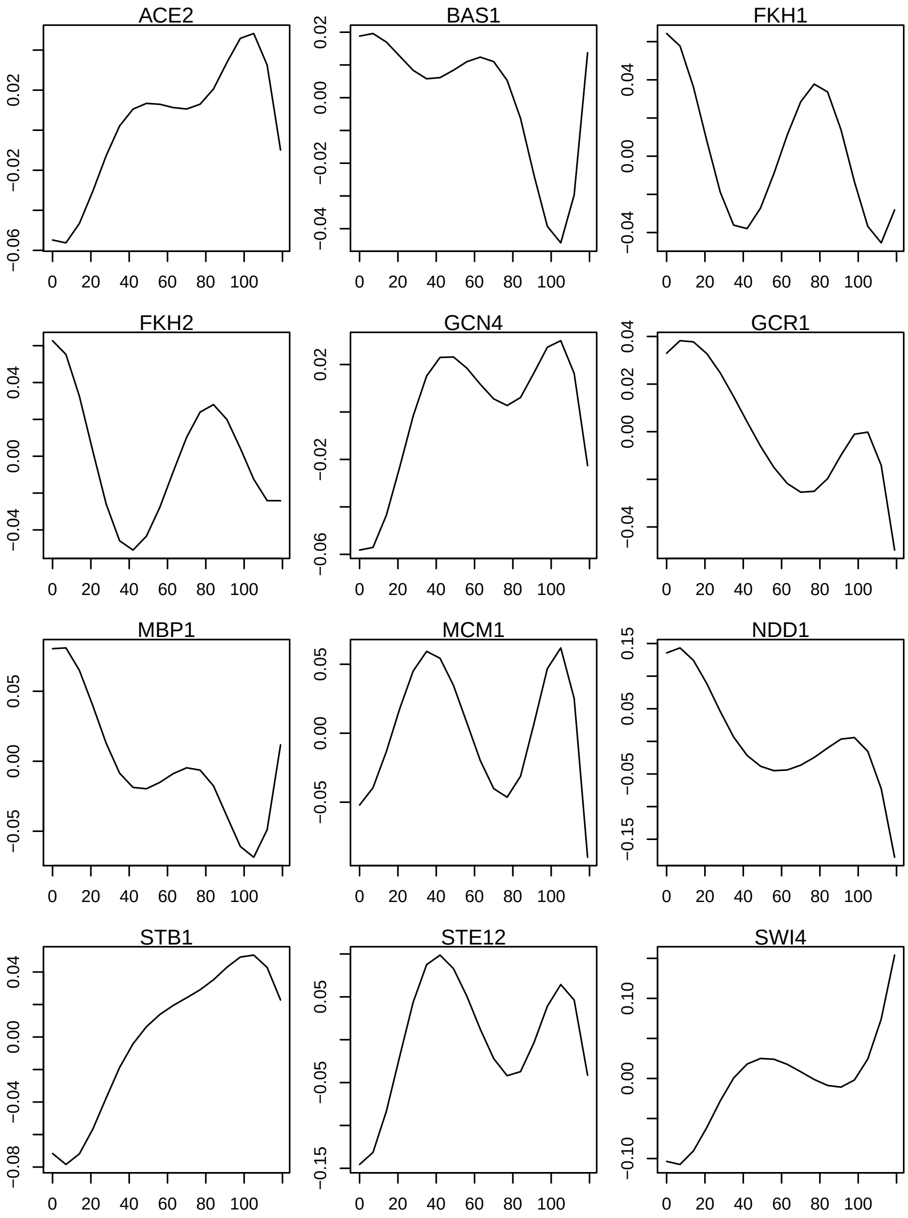


Figure 2.1: Estimated time-varying transcriptional effects for 21 known yeast TFs related to cell cycle process. LEU3 and REB1 are not selected, so there are no estimates for these two.

Fan, J., Feng, Y. and Song, R. (2011). Nonparametric independence screening in sparse ultra-high-dimensional additive models. *Journal of the American Statistical Association*, **106** 544–557.



(Plot continued on the next page)

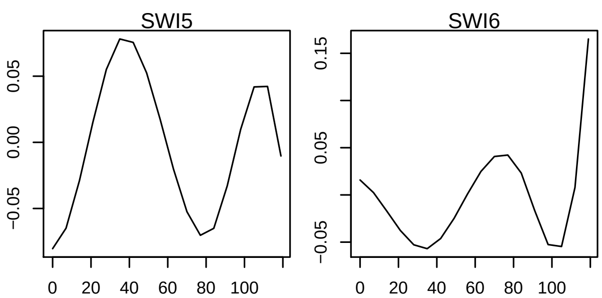


Figure 2.2: Estimated time-varying transcriptional effects for 14 TFs identified by IVIS on an augmented higher dimension dataset.