Supplementary Paper to: Fair Simultaneous Prediction and Confidence Bands for Concurrent Functional Regressions: Comparing Sprinters with Prosthetic versus Biological Legs

Michael L. Creutzinger
Department of Statistics, Colorado State University
Dominik Liebl
Institute of Finance and Statistics, University of Bonn
and
Julia L. Sharp
Mathematical Statistician, National Institute of Standards and Technology

February 7, 2024

1 Proofs of the Theoretical Results

Let $G_j^{(d)}$ denote the dth derivative of the jth, $j = 1, \ldots, K+1$, element of the $(K+1)$-dimensional vector of random functions,

$$G = (X_1, \ldots, X_K, \varepsilon)^T.$$

In the no derivative case, $G_j^{(0)}$, we usually drop the superscript and write $G_j^{(0)} = G_j$. Define

$$G_{jk}^{(d)} := G_j^{(d)} G_k^{(d)}, \quad j, k = 1, \ldots, K+1.$$

In our proofs, we make use of the following lemma:

Lemma 1.1 (Stochastic Lipschitz Continuity). *Under Assumption **A3**, we have for all $d \in \{0, 1\}$ and all $j, k = 1, \ldots, K+1$*

$$\left| G_{jk}^{(d)}(t) - G_{jk}^{(d)}(s) \right| \leq A_{jkd}\, \phi_{jkd}(|t - s|) \quad \textit{for all} \quad t, s \in [0, 1],$$

where ϕ_{jkd} is a deterministic nondecreasing continuous function on $[0, 1]$ with $\phi_{jkd}(0) = 0$, and where A_{jkd} is a real-valued random variable with $E(|A_{jkd}|^2) < \infty$.

Proof of Lemma 1.1

Under Assumption **A3**, $G_{jk}^{(d)}$ is continuously differentiable, and thus, by the Mean Value Theorem,

$$G_{jk}^{(d)}(t) - G_{jk}^{(d)}(s) = G_{jk}^{(d+1)}(\xi)(t-s) \quad \text{for some} \quad \xi \in (s,t)$$

and any $0 \leq s < t \leq 1$. This implies, for all $s, t \in [0,1]$,

$$\left(G_{jk}^{(d)}(t) - G_{jk}^{(d)}(s)\right)^2 \leq \sup_{\xi \in [0,1]} \left(G_{jk}^{(d+1)}(\xi)\right)^2 (t-s)^2$$

Now, if we take the expectation of the left-hand side and apply the uniform 2nd moment assumption in **A3**, it yields for all $s, t \in [0,1]$,

$$\mathbb{E}\left[\left(G_{jk}^{(d)}(t) - G_{jk}^{(d)}(s)\right)^2\right] \leq C_{jkd}(t-s)^2 =: f_{jkd}((t-s)^2), \tag{1}$$

where the constant $0 < C_{jkd} < \infty$, and thus also the deterministic function f_{jkd}, only depends on $j, k = 1, \dots, K+1$, and $d \in \{0, 1\}$. Now, observe that for all $j, k = 1, \dots, K+1$, and $d \in \{0, 1\}$

$$\int_0^1 x^{-3/2} f_{jkd}^{1/2}(x) dx = 2C_{jkd}^{1/2} < \infty \tag{2}$$

The result of Lemma 1.1 follows now directly from (1) and (2) by applying Theorem 2.3 in Hahn (1977) for the case of $r = $ 2nd moments. □

Proof of Theorem 2.1 (a) (Uniform Convergence of the OLS Estimator).

We first consider pointwise convergence for each $t \in [0,1]$ and then expand this to uniform convergence. Using standard arguments, we can express the vector-valued parameter function estimator $\hat{\beta}(t) = (\hat{\beta}_1(t), \dots, \hat{\beta}_K(t))^T$ as

$$\begin{aligned}\hat{\beta}(t) &= \left(n^{-1} \sum_{i=1}^n X_i(t) X_i^T(t)\right)^{-1} n^{-1} \sum_{i=1}^n X_i(t) Y_i(t) \\ &= \beta(t) + \left(n^{-1} \sum_{i=1}^n X_i(t) X_i^T(t)\right)^{-1} n^{-1} \sum_{i=1}^n X_i(t) \varepsilon_i(t).\end{aligned}$$

By Kolmogorov's strong law of large numbers (SLLN) and the continuous mapping theorem, the second summand converges (a.s.) to the K-dimensional zero vector

$$\left(n^{-1} \sum_{i=1}^n X_i(t) X_i^T(t)\right)^{-1} n^{-1} \sum_{i=1}^n X_i^T \varepsilon_i(t) \overset{a.s.}{\to} 0, \quad n \to \infty,$$

which implies that pointwise for each $t \in [0,1]$

$$\hat{\beta}_j(t) \overset{a.s.}{\to} \beta_j(t), \quad \text{for each} \quad j = 1, \dots, K. \tag{3}$$

Moreover, from $\mathbb{E}[\varepsilon(t)|X(t)] = 0$ and our iid assumption, it follows that the estimator $\hat{\beta}_j(t)$ is unbiased, since pointwise for each $t \in [0,1]$ and every n

$$\mathbb{E}\left[\left(n^{-1}\sum_{i=1}^{n} X_i(t)X_i^T(t)\right)^{-1} n^{-1}\sum_{i=1}^{n} X_i^T(t)\varepsilon_i(t)\right] = 0.$$

Now, we need to expand result (3) to uniform convergence across all $t \in [0,1]$. Let $G_{ijk}(t) = G_{ij}(t)G_{ik}(t)$ denote the iid copies of $G_{jk}(t) = G_j(t)G_k(t)$. Lemma 1.1 implies for all $j,k = 1,\ldots,K+1$ that

$$|G_{ijk}(t) - G_{ijk}(s)| \leq A_{ijk}\phi_{jk}(|t-s|) \quad \text{for all} \quad t,s \in [0,1]$$

where A_{ijk}, $i = 1,\ldots,n$ is iid as A_{jk} with $E(A_{jk}^2) < \infty$. Define

$$B_{jkn}(t) := n^{-1}\sum_{i=1}^{n} G_{ijk}(t) - E\left[G_{ijk}(t)\right].$$

By SLLN,

$$B_{jkn}(t) \overset{a.s.}{\to} 0 \quad \text{pointwise for each} \quad t \in [0,1]. \tag{4}$$

Then,

$$\begin{aligned}
&|B_{jkn}(t) - B_{jkn}(s)| \\
&= \left|\left(n^{-1}\sum_{i=1}^{n} G_{ijk}(t) - E\left[G_{ijk}(t)\right]\right) - \left(n^{-1}\sum_{i=1}^{n} G_{ijk}(s) - E\left[G_{ijk}(s)\right]\right)\right| \\
&\leq n^{-1}\sum_{i=1}^{n} |G_{ijk}(t) - G_{ijk}(s)| + E\left[|G_{ijk}(t) - G_{ijk}(s)|\right] \quad \text{(Triangle Inequality)} \\
&\leq n^{-1}\sum_{i=1}^{n} A_{ijk}\phi_{jk}(|t-s|) + E\left(A_{ijk}\right)\phi_{jk}(|t-s|) \quad \text{(Lemma 1.1)} \\
&\leq \left(n^{-1}\sum_{i=1}^{n} A_{ijk} + E\left(A_{ijk}\right)\right)\phi_{jk}(|t-s|),
\end{aligned} \tag{5}$$

where by SLLN and Lemma 1.1 ($E\left(A_{ijk}\right) < \infty$),

$$\left(n^{-1}\sum_{i=1}^{n} A_{ijk} + E\left(A_{ijk}\right)\right) \overset{a.s.}{\to} 2E\left(A_{ijk}\right) < \infty.$$

Result (5) implies, by Theorem 22.8 from Davidson (2021), that B_{jkn} is strongly stochastically equicontinuous for every $j,k \in \{1,\ldots,K\}$. This, together with the pointwise consistency (4) implies, by Theorem 22.10 from Davidson (2021), that

$$\sup_{t\in[0,1]} |B_{jkn}(t)| \overset{a.s.}{\to} 0, \quad \text{for all } j,k \in \{1,\ldots,K+1\}.$$

The latter elementwise result implies the following matrix-valued and vector-valued uniform convergence results,

$$\sup_{t\in[0,1]}\left|n^{-1}\sum_{i=1}^{n}X_i(t)X_i^T(t)-E\left(X_i(t)X_i^T(t)\right)\right|\overset{a.s.}{\to}\underset{(K\times K)}{0}$$
$$\sup_{t\in[0,1]}\left|n^{-1}\sum_{i=1}^{n}X_i(t)\varepsilon_i(t)-E\left(X_i(t)\varepsilon_i(t)\right)\right|\overset{a.s.}{\to}\underset{(K\times 1)}{0},$$

where $E\left(X_i(t)\varepsilon_i(t)\right)=E\left(X_i(t)E(\varepsilon_i(t)|X_i(t))\right)=0$ for all $t\in[0,1]$. Under our assumptions, $E\left[X_i(t)X_i^T(t)\right]$ is invertible. Thus, by the functional version of the uniform continuous mapping theorem, we also have that

$$\sup_{t\in[0,1]}\left|\left(n^{-1}\sum_{i=1}^{n}X_i(t)X_i^T(t)\right)^{-1}n^{-1}\sum_{i=1}^{n}X_i(t)\varepsilon_i(t)\right|\overset{a.s.}{\to}\underset{(K\times 1)}{0}$$

which implies that

$$\sup_{t\in[0,1]}\left|\hat{\beta}(t)-\beta(t)\right|\overset{a.s.}{\to}\underset{(K\times 1)}{0}$$

concluding the proof of Theorem 2.1 (a). □

Proof of Theorem 2.1 (b) (Uniform Convergence of the Variance Estimator).

The estimators $\hat{\sigma}_{\varepsilon}^{ml}$, $\hat{\sigma}_{\varepsilon}^{ub}$, and $\hat{\sigma}_{\varepsilon}^{mm}$ only differ with respect to the scaling parameters $\frac{1}{n}$, $\frac{1}{n-K}$, and $\frac{\nu_0-4}{(\nu_0-2)(n-K+2)}$ and thus are asymptotically equivalent. Therefore, it suffices to consider the ML estimator $\hat{\sigma}_{\varepsilon}^{ml}\equiv\hat{\sigma}_{\varepsilon}$. Note that the residuals, $e_i(t)$, can be defined as

$$\begin{aligned}
e_i(t)&=Y_i(t)-X_i^T(t)\hat{\beta}(t)\\
&=Y_i(t)-X_i^T(t)\beta(t)-X_i^T(t)\left(\hat{\beta}(t)-\beta(t)\right)\\
&=\varepsilon_i(t)-X_i^T(t)\left(\hat{\beta}(t)-\beta(t)\right),\\
e_i^2(t)&=\varepsilon_i^2(t)-2\left(\hat{\beta}(t)-\beta(t)\right)^T X_i(t)\varepsilon_i(t)\\
&\quad+\left(\hat{\beta}(t)-\beta(t)\right)^T X_i(t)X_i^T(t)\left(\hat{\beta}(t)-\beta(t)\right),
\end{aligned}$$

and the ML estimator of the variance is

$$
\begin{aligned}
\hat{\sigma}_{\varepsilon}(t,t) &= \frac{1}{n}\sum_{i=1}^{n}\left(e_i(t)\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(Y_i(t) - X_i^T(t)\hat{\beta}(t)\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\varepsilon_i^2(t) - 2\left(\hat{\beta}(t)-\beta(t)\right)^T X_i(t)\varepsilon_i(t) + \left(\hat{\beta}(t)-\beta(t)\right)^T X_i(t)X_i^T(t)\left(\hat{\beta}(t)-\beta(t)\right)\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2(t) - 2\left(\hat{\beta}(t)-\beta(t)\right)^T \frac{1}{n}\sum_{i=1}^{n}X_i(t)\varepsilon_i(t) \\
&\quad + \left(\hat{\beta}(t)-\beta(t)\right)^T \frac{1}{n}\sum_{i=1}^{n}\left(X_i(t)X_i^T(t)\right)\left(\hat{\beta}(t)-\beta(t)\right).
\end{aligned}
$$

By the SLLN and continuous mapping theorem, the second and third summand converge (a.s.) to zero. This implies that pointwise for each $t \in [0,1]$,

$$
\hat{\sigma}_{\varepsilon}(t,t) \overset{a.s.}{\to} \sigma_{\varepsilon}(t,t). \tag{6}
$$

Now, we need to expand result (6) to uniform convergence across all $t \in [0,1]$. The following uniform convergence results were proven in Theorem 2.1 (a):

$$
\begin{aligned}
&\sup_{t\in[0,1]}\left|\hat{\beta}(t)-\beta(t)\right| \overset{a.s.}{\to} \underset{(K\times 1)}{0}, \\
&\sup_{t\in[0,1]}\left|n^{-1}\sum_{i=1}^{n}X_i(t)X_i^T(t) - E\left(X_i(t)X_i^T(t)\right)\right| \overset{a.s.}{\to} \underset{(K\times K)}{0}, \text{ and} \\
&\sup_{t\in[0,1]}\left|n^{-1}\sum_{i=1}^{n}X_i(t)\varepsilon_i(t) - E\left(X_i(t)\varepsilon_i(t)\right)\right| \overset{a.s.}{\to} \underset{(K\times 1)}{0}.
\end{aligned}
$$

Then, by functional continuous mapping theorem,

$$
\begin{aligned}
&\sup_{t\in[0,1]}\left|\left(\hat{\beta}(t)-\beta(t)\right)^T \frac{1}{n}\sum_{i=1}^{n}X_i(t)\varepsilon_i(t)\right| \overset{a.s.}{\to} 0 \quad \text{and} \\
&\sup_{t\in[0,1]}\left|\left(\hat{\beta}(t)-\beta(t)\right)^T \frac{1}{n}\sum_{i=1}^{n}\left(X_i(t)X_i^T(t)\right)\left(\hat{\beta}(t)-\beta(t)\right)\right| \overset{a.s.}{\to} 0,
\end{aligned}
$$

which implies that

$$
\sup_{t\in[0,1]}\left|\hat{\sigma}_{\varepsilon}(t,t) - \sigma_{\varepsilon}(t,t)\right| \overset{a.s.}{\to} 0. \tag{7}
$$

Next, we need to show the covariance estimator converges uniformly. Let, without loss of generality, $s < t \in [0,1]$, then

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} e_i(s)e_i(t) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\varepsilon_i(s) - X_i^T(s)\left(\hat{\beta}(s)-\beta(s)\right)\right)\left(\varepsilon_i(t) - X_i^T(t)\left(\hat{\beta}(t)-\beta(t)\right)\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i(s)\varepsilon_i(t) \\
&\quad - \left(\hat{\beta}(s)-\beta(s)\right)^T \ \frac{1}{n}\sum_{i=1}^{n} X_i(s)\varepsilon_i(t) \\
&\quad - \left(\hat{\beta}(t)-\beta(t)\right)^T \ \frac{1}{n}\sum_{i=1}^{n} X_i(t)\varepsilon_i(s) \\
&\quad + \frac{1}{n}\sum_{i=1}^{n} X_i^T(s)\left(\hat{\beta}(s)-\beta(s)\right) X_i^T(t)\left(\hat{\beta}(t)-\beta(t)\right).
\end{aligned}$$

By the SLLN and continuous mapping theorem, the second, third, and fourth summand converge (a.s.) to zero. This implies that pointwise for each $s, t \in [0,1]$,

$$\hat{\sigma}_\varepsilon(s,t) \overset{a.s.}{\to} \sigma_\varepsilon(s,t). \tag{8}$$

Now, we need to establish uniform convergence of the covariance estimator. Observe that

$$X_i^T(s)\left(\hat{\beta}(s)-\beta(s)\right)\varepsilon_i(t) = \varepsilon_i(t)X_i^T(s)\left(\hat{\beta}(s)-\beta(s)\right) \qquad \text{and}$$
$$X_i^T(s)\left(\hat{\beta}(s)-\beta(s)\right)X_i^T(t)\left(\hat{\beta}(t)-\beta(t)\right) = \left(\hat{\beta}(s)-\beta(s)\right)^T X_i(s)X_i^T(t)\left(\hat{\beta}(t)-\beta(t)\right).$$

It was already proven in Theorem 2.1 (a) that

$$\sup_{t\in[0,1]}\left|\hat{\beta}(t)-\beta(t)\right| \overset{a.s.}{\to} \underset{(K\times 1)}{0}.$$

Thus, it suffices to show that $\varepsilon_i(s)X_i^T(t)$ and $X_i(s)X_i^T(t)$ converge uniformly.

Let $G_{ijk}(s,t) = G_{ij}(s)G_{ik}(t)$ for some predictor, j, such that

$$n^{-1}\sum_{i=1}^{n} X_{ij}(s)X_{ik}^T(t) = n^{-1}\sum_{i=1}^{n} G_{ij}(s)G_{ik}^T(t) = n^{-1}\sum_{i=1}^{n} G_{ijk}(s,t).$$

Note that by SLLN,

$$n^{-1}\sum_{i=1}^{n} G_{ijk}(s,t) \overset{a.s.}{\to} E\left[G_{ijk}(s,t)\right], \tag{9}$$

pointwise for each j, k, s, and t. Next, define

$$B_{jkn}(s,t) := n^{-1} \sum_{i=1}^{n} G_{ijk}(s,t) - E\left[G_{ijk}(s,t)\right].$$

We need to show that B_{jkn} is strongly stochastically equicontinuous for every $j \in \{1, \ldots, K+1\}$. Let $s, t, u, v \in [0,1]$, such that $s < u$ and $t < v$, without loss of generality. Then,

$$\begin{aligned}
|B_{jkn}(s,t) - B_{jkn}(u,v)| = &\Bigg| \left(n^{-1} \sum_{i=1}^{n} G_{ijk}(s,t) - E\left[G_{ijk}(s,t)\right] \right) \\
&- \left(n^{-1} \sum_{i=1}^{n} G_{ijk}(u,v) - E\left[G_{ijk}(u,v)\right] \right) \Bigg| \\
= &\Bigg| \left(n^{-1} \sum_{i=1}^{n} G_{ij}(s) G_{ik}(t) - n^{-1} \sum_{i=1}^{n} G_{ij}(u) G_{ik}(v) \right) \\
&+ \left(E\left[G_{ij}(s) G_{ik}(t)\right] - E\left[G_{ij}(u) G_{ik}(v)\right] \right) \Bigg| \\
\leq &\left| n^{-1} \sum_{i=1}^{n} \left(G_{ij}(s) G_{ik}(t) - G_{ij}(u) G_{ik}(v) \right) \right| \\
&+ \left| \left(E\left[G_{ij}(s) G_{ik}(t) - G_{ij}(u) G_{ik}(v)\right] \right) \right|,
\end{aligned}$$

by the Triangle Inequality. Now, focusing on the first quantity:

$$\begin{aligned}
&\left| n^{-1} \sum_{i=1}^{n} \big(G_{ij}(s) G_{ik}(t) - G_{ij}(u) G_{ik}(v) \big) \right| \\
&\qquad \leq \left| n^{-1} \sum_{i=1}^{n} \left(\left(G_{ij}(s) - G_{ij}(u) \right) G_{ik}(t) - G_{ij}(u) \left(G_{ik}(t) - G_{ik}(v) \right) \right) \right| \\
&\qquad \leq n^{-1} \sum_{i=1}^{n} \left| \left(G_{ij}(s) - G_{ij}(u) \right) \right| \left| G_{ik}(t) \right| + \left| G_{ij}(u) \right| \left| \left(G_{ik}(t) - G_{ik}(v) \right) \right|.
\end{aligned}$$

Again, we focus on the first part of the summation, since the second part is analogous to the first. Using the Cauchy-Schwarz inequality,

$$
\begin{aligned}
& n^{-1} \sum_{i=1}^{n} \left|(G_{ij}(s) - G_{ij}(u))\right| \left|G_{ik}(t)\right| \\
& \leq \left(n^{-1} \sum_{i=1}^{n} \left|(G_{ij}(s) - G_{ij}(u))\right|^2\right)^{1/2} \left(n^{-1} \sum_{i=1}^{n} \left|G_{ik}(t)\right|^2\right)^{1/2} \\
& \leq \left(n^{-1} \sum_{i=1}^{n} \left|(G_{ij}(s) - G_{ij}(u))\right|^2\right)^{1/2} \left(n^{-1} \sum_{i=1}^{n} \sup_t \left(G_{ik}(t)\right)^2\right)^{1/2} \\
& \leq \left(n^{-1} \sum_{i=1}^{n} \left[A_{ij} \phi_j \left(|s-u|\right)\right]^2\right)^{1/2} \left(n^{-1} \sum_{i=1}^{n} \sup_t \left(G_{ik}(t)\right)^2\right)^{1/2},
\end{aligned}
$$

where the last inequality holds by Lemma 1.1 and $A_{ij}, i = 1, \ldots, n$ is iid as A_j with $E\left(A_j^2\right) < \infty$. Then, by the SLLN and functional continuous mapping theorem,

$$
n^{-1} \sum_{i=1}^{n} \left|(G_{ij}(s) - G_{ij}(u))\right| \left|G_{ik}(t)\right| \overset{a.s.}{\to} E\left(A_j^2\right)^{1/2} \phi_j\left(|s-u|\right) \cdot E\left[\sup_t G_k(t)^2\right]^{1/2}, \tag{10}
$$

which is finite, by assumption A3. Similary,

$$
n^{-1} \sum_{i=1}^{n} \left|G_{ij}(u)\right| \left|(G_{ik}(t) - G_{ik}(v))\right| \overset{a.s.}{\to} E\left(A_k^2\right)^{1/2} \phi_k\left(|t-v|\right) \cdot E\left[\sup_u G_j(u)^2\right]^{1/2}, \tag{11}
$$

which is finite, by assumption A3. Therefore,

$$
\begin{aligned}
& \left|n^{-1} \sum_{i=1}^{n} \left(G_{ij}(s) G_{ik}(t) - G_{ij}(u) G_{ik}(v)\right)\right| \qquad (12) \\
& \overset{a.s.}{\to} E\left(A_j^2\right)^{1/2} \cdot E\left[\sup_t G_k(t)^2\right]^{1/2} \cdot \phi_j\left(|s-u|\right) + E\left(A_k^2\right)^{1/2} \cdot E\left[\sup_u G_j(u)^2\right]^{1/2} \cdot \phi_k\left(|t-v|\right),
\end{aligned}
$$

which is also finite. By a similar argument,

$$
\begin{aligned}
& \left|\left(E\left[G_{ij}(s) G_{ik}(t) - G_{ij}(u) G_{ik}(v)\right]\right)\right| \qquad (13) \\
& \overset{a.s.}{\to} E\left(A_j^2\right)^{1/2} \cdot E\left[\sup_t G_k(t)^2\right]^{1/2} \cdot \phi_j\left(|s-u|\right) + E\left(A_k^2\right)^{1/2} \cdot E\left[\sup_u G_j(u)^2\right]^{1/2} \cdot \phi_k\left(|t-v|\right).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
& \left|B_{jkn}(s,t) - B_{jkn}(u,v)\right| \qquad (14) \\
& \overset{a.s.}{\to} 2\left(E\left(A_j^2\right)^{1/2} \cdot E\left[\sup_t G_k(t)^2\right]^{1/2} \cdot \phi_j\left(|s-u|\right) + E\left(A_k^2\right)^{1/2} \cdot E\left[\sup_u G_j(u)^2\right]^{1/2} \cdot \phi_k\left(|t-v|\right)\right) \\
& < \infty.
\end{aligned}
$$

Result (14) implies, by Theorem 22.8 from Davidson (2021), that B_{jkn} is strongly stochastically equicontinuous for every $j,k \in \{1,\ldots,K+1\}$. This, together with the pointwise consistency (9) implies, by Theorem 22.10 from Davidson (2021), that

$$\sup_{s,t\in[0,1]} |B_{jkn}(s,t)| \overset{a.s.}{\to} 0, \quad \text{for all } j,k \in \{1,\ldots,K+1\}. \tag{15}$$

The latter elementwise result implies the following matrix-valued and vector-valued uniform convergence results,

$$\sup_{s,t\in[0,1]} \left| n^{-1} \sum_{i=1}^{n} X_i(s)X_i^T(t) - E\left[X_i(s)X_i^T(t)\right] \right| \overset{a.s.}{\to} \underset{(K\times K)}{0}. \tag{16}$$

$$\sup_{s,t\in[0,1]} \left| n^{-1} \sum_{i=1}^{n} \varepsilon_i(s)X_i^T(t) - E\left[\varepsilon_i(s)X_i^T(t)\right] \right| \overset{a.s.}{\to} \underset{(1\times K)}{0}. \tag{17}$$

Furthermore, by the functional continuous mapping theorem,

$$\sup_{s,t\in[0,1]} \left| n^{-1} \sum_{i=1}^{n} \varepsilon_i(t)X_i^T(s) \left(\hat{\beta}(s) - \beta(s)\right) \right| \overset{a.s.}{\to} 0, \quad \text{and}$$

$$\sup_{s,t\in[0,1]} \left| n^{-1} \sum_{i=1}^{n} \left(\hat{\beta}(s) - \beta(s)\right)^T X_i(s)X_i^T(t) \left(\hat{\beta}(t) - \beta(t)\right) \right| \overset{a.s.}{\to} 0.$$

The results of uniform convergence, combined with pointwise convergence (8), implies that

$$\sup_{t,s\in[0,1]} |\hat{\sigma}_\varepsilon(s,t) - \sigma_\varepsilon(s,t)| \overset{a.s.}{\to} 0, \tag{18}$$

concluding the proof of Theorem 2.1 (b).

□

Proof of Theorem 2.1 (c) (Uniform Convergence of the τ Estimator).

Applying functional continuous mapping theorem together with assumption A3 immediately gives the result.

□

Proof for Theorem 2.2 (Functional Central Limit Theorem for Concurrent, Functional Regression Parameter Estimation)

Given that $\hat{\beta}(t)$ is the OLS estimator of $\beta(t)$, we can write

$$\sqrt{n}\left(\hat{\beta}(t) - \beta(t)\right) = \left(n^{-1} \sum_{i=1}^{n} X_i(t)X_i^T(t)\right)^{-1} \left(\sqrt{n}\left[n^{-1} \sum_{i=1}^{n} X_i(t)\varepsilon_i(t)\right]\right).$$

As was shown in Theorem 2.1(a), we also know that

$$\sup_{t\in[0,1]} \left| n^{-1} \sum_{i=1}^{n} X_i(t) X_i^T(t) - E\left(X_i(t)X_i^T(t)\right) \right| \overset{a.s.}{\to} \underset{(K\times K)}{0}.$$

We need to show that $(\sqrt{n}\,[n^{-1}\sum_{i=1}^{n} X_i(t)\varepsilon_i(t)])$ converges in distribution to a Gaussian process. First, let $\lambda \in \mathbb{R}^K$. Then,

$$\left(\sqrt{n}\left[n^{-1}\sum_{i=1}^{n}\lambda^T X_i(t)\varepsilon_i(t)\right]\right) \in \mathbb{R},$$

with mean zero. By Lemma 3.1, $X_i(t)\varepsilon_i(t)$ is stochastic lipschitz continuous, and meets the integrability assumption (equation 2) of Theorem 2.5 in Hahn (1977) for $r = 2$. Therefore, $\sum_{i=1}^{n} X_i(t)\varepsilon_i(t)$ satisfies the CLT in $\mathcal{C}^1[0,1]$, since it is also mean zero. Specifically, we have

$$\left(\sqrt{n}\left[n^{-1}\sum_{i=1}^{n}\lambda^T X_i(t)\varepsilon_i(t)\right]\right) \to \mathcal{G}_p\left(0, \lambda^T c_{X\varepsilon}(s,t)\lambda\right), \tag{19}$$

where $\mathcal{G}_p$ is a mean-zero Gaussian process and variance term $c_{X\varepsilon}(s,t) = E\left[\left(X_i(s)\varepsilon_i(s)\right)\left(X_i(t)\varepsilon_i(t)\right)^T\right]$. By the Cramer-Wold device, we now have that

$$\left(\sqrt{n}\left[n^{-1}\sum_{i=1}^{n} X_i(t)\varepsilon_i(t)\right]\right) \to \mathcal{G}_p\left(0, c_{X\varepsilon}(s,t)\right). \tag{20}$$

Then, by the functional version of Slutsky's Theorem,

$$\sqrt{n}\left(\hat{\beta}(t) - \beta(t)\right) \to \mathcal{G}_p\left(0, E\left[X_i(s)X_i(t)\right]^{-1} c_{X\varepsilon}(s,t) E\left[X_i(s)X_i(t)\right]^{-1}\right). \tag{21}$$

This allows for heteroscedastic errors, where $c_{X\varepsilon}(s,t)$ is a function of X_i. If homoscedastic errors are assumed, then

$$c_{X\varepsilon}(s,t) = E\left[X_i(s)X_i(t)^T\right] c_\varepsilon(s,t),$$

where $c_\varepsilon(s,t) = E\left[\varepsilon_i(s)\varepsilon_i(t)\right]$. Substituting into the covariance for equation 21, for homoscedastic errors, we have

$$\sqrt{n}\left(\hat{\beta}(t) - \beta(t)\right) \to \mathcal{G}_p\left(0, c_\varepsilon(s,t) E\left[X_i(s)X_i(t)\right]^{-1}\right). \tag{22}$$

□

2 Supplemental Definitions

Definition 2.1. *A stochastic process* $Z = \{Z(t) \in \mathbb{R},\, t \in [0,1]\}$ *is a* $\mathcal{L}^b$*-Lipschitz process, if there exists a random variable* A *satisfying* $\mathbb{E}[|A|^b] < \infty$*, such that*

$$|Z(t) - Z(t')| \leq A\,|t - t'| \quad \textit{for all} \quad t, t' \in [0,1]$$

and

$$\int_0^1 \sqrt{\log(N([0,1],u))}du < \infty,$$

where $N([0,1],u)$ *is the minimal number of intervals of length* u *needed to cover the domain* $[0,1]$.

Side note: Since $\log(N([0,1],u)) \leq \log(C/u) \leq (C/u) - 1 \leq (C/u)$ for some $0 < C < \infty$, then

$$\begin{aligned}\int_0^1 \sqrt{\log(N([0,1],u))}du &\leq C \int_0^1 u^{-1/2} du \\ &= C[2u^{1/2}]_0^1 \\ &= 2C < \infty\end{aligned}$$

3 Supplemental Figures

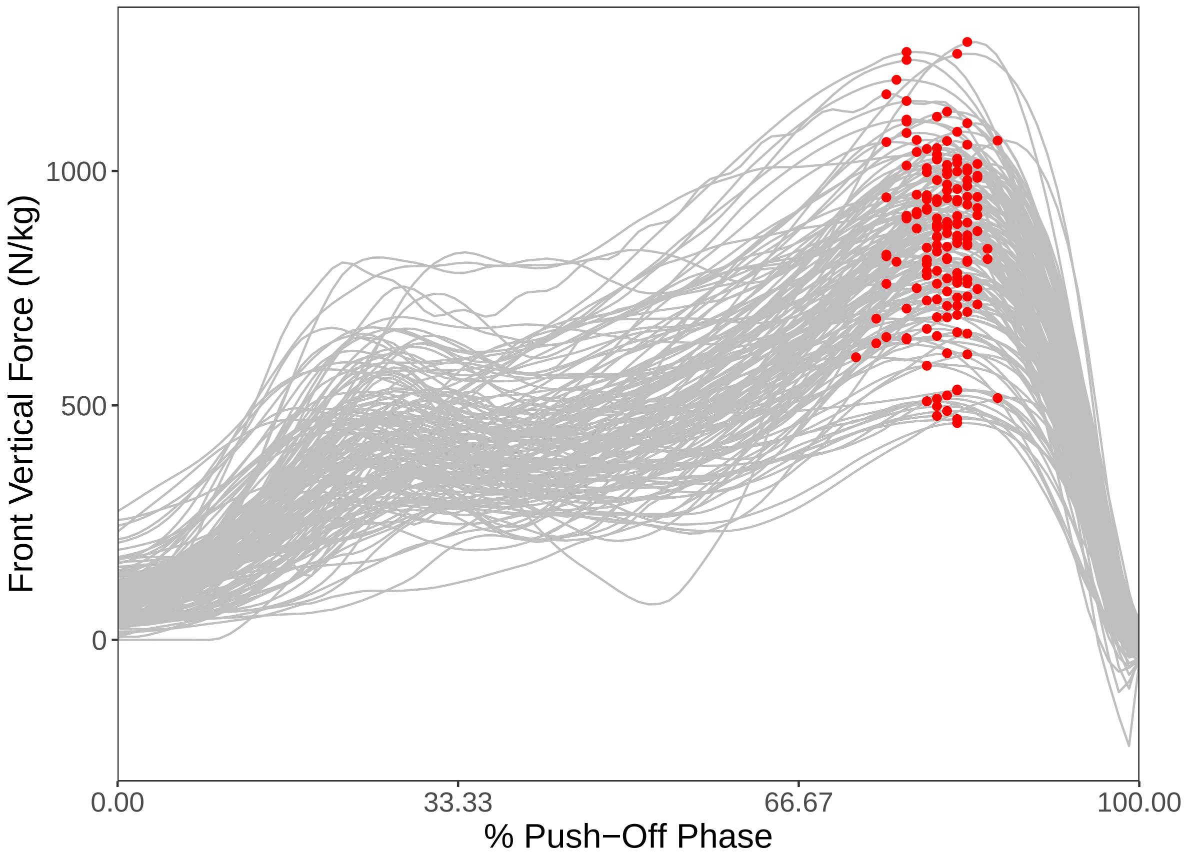

Figure 1: A plot of the original vertical force $Y(t)$ for each sprinter, with their maximum vertical force plotted in red.

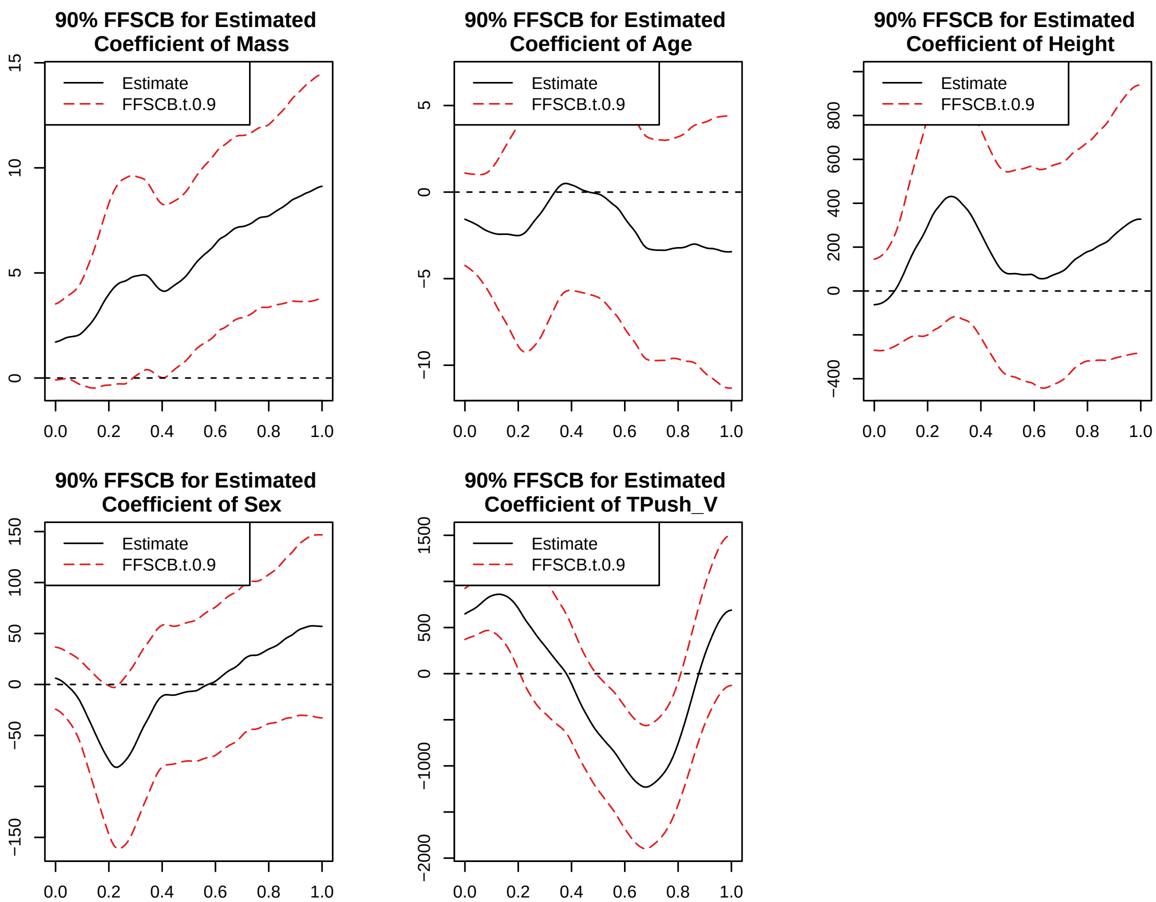

Figure 2: 90% simultaneous confidence bands for the estimated coefficients of the concurrent regression model. The bands are made by FFSCBs, using three intervals.

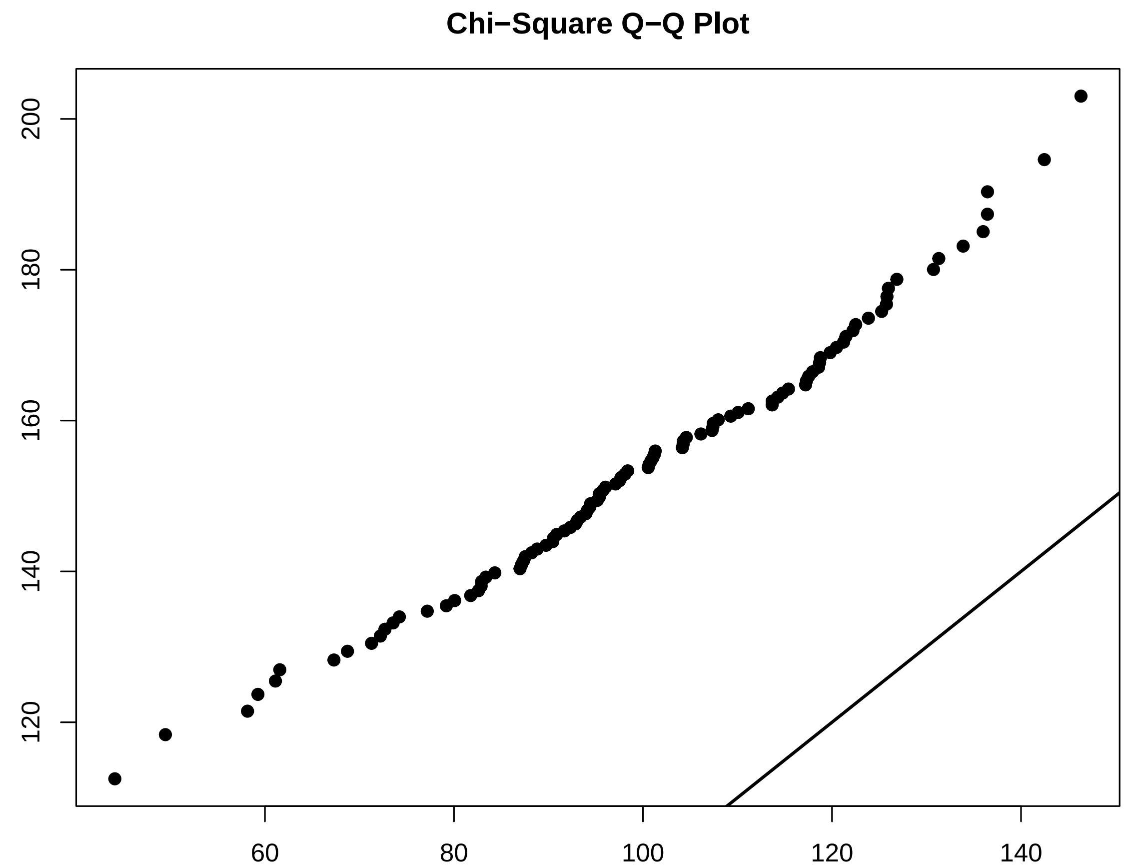

Figure 3: QQPlot obtained by the Mardia multivariate normality test implemented in R.

References

Davidson, J. (2021). *Stochastic Limit Theory: An Introduction for Econometricians* (Second edition ed.). Advanced Texts in Econometrics. Oxford University Press.

Hahn, M. G. (1977). Conditions for sample-continuity and the central limit theorem. *The Annals of Probability 5*(3).