Received: 9 October 2022 | Accepted: 22 March 2023

DOI: 10.1111/biom.13866

BIOMETRIC METHODOLOGY

Biometrics A JOURNAL OF THE INTERNATIONAL BIOMETRIC SOCIETY WILEY

Asynchronous and error-prone longitudinal data analysis via functional calibration

Xinyue Chang[1] | **Yehua Li**[2] | **Yi Li**[3]

[1]Eli Lilly and Company, Indianapolis, Indiana, USA

[2]Department of Statistics, University of California, Riverside, California, USA

[3]Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA

Correspondence
Yehua Li, Department of Statistics, University of California, Riverside, CA 92521, USA.
Email: yehuali@ucr.edu

Funding information
National Institute of Health, Grant/Award Numbers: 5R21AG058198, R01CA269398, R01CA249096

Abstract

In many longitudinal settings, time-varying covariates may not be measured at the same time as responses and are often prone to measurement error. Naive last-observation-carried-forward methods incur estimation biases, and existing kernel-based methods suffer from slow convergence rates and large variations. To address these challenges, we propose a new functional calibration approach to efficiently learn longitudinal covariate processes based on sparse functional data with measurement error. Our approach, stemming from functional principal component analysis, calibrates the unobserved synchronized covariate values from the observed asynchronous and error-prone covariate values, and is broadly applicable to asynchronous longitudinal regression with time-invariant or time-varying coefficients. For regression with time-invariant coefficients, our estimator is asymptotically unbiased, root-n consistent, and asymptotically normal; for time-varying coefficient models, our estimator has the optimal varying coefficient model convergence rate with inflated asymptotic variance from the calibration. In both cases, our estimators present asymptotic properties superior to the existing methods. The feasibility and usability of the proposed methods are verified by simulations and an application to the Study of Women's Health Across the Nation, a large-scale multisite longitudinal study on women's health during midlife.

KEYWORDS

functional principal component analysis, kernel smoothing, measurement error, regression calibration, sparse functional data, varying coefficient model

1 | INTRODUCTION

In many decade-long longitudinal studies, participants' heath information is repeatedly measured by diverse instruments, such as blood tests, physical examinations, nutritional evaluations, and psychological assessments. These tests and assessments usually follow different schedules, and are not synchronized in time. The resulting data structures create an asynchronous issue where the response variable and covariates are not measured at the same time. For example, in our motivating Study of Women's Health Across the Nation (SWAN, 1996) a multisite longitudinal study on women's health during their midlife years, a total of 3302 women were followed from 1996 to 2008 to study their physical, biological, psychological, and social changes that occurred during the menopausal transition. These health-related metrics were grouped into physical, hormone, and cardiovascular measurements. Figure 1(A) shows the measurement times for hormone, physical, and cardiovascular measurements for

 © 2023 The International Biometric Society.

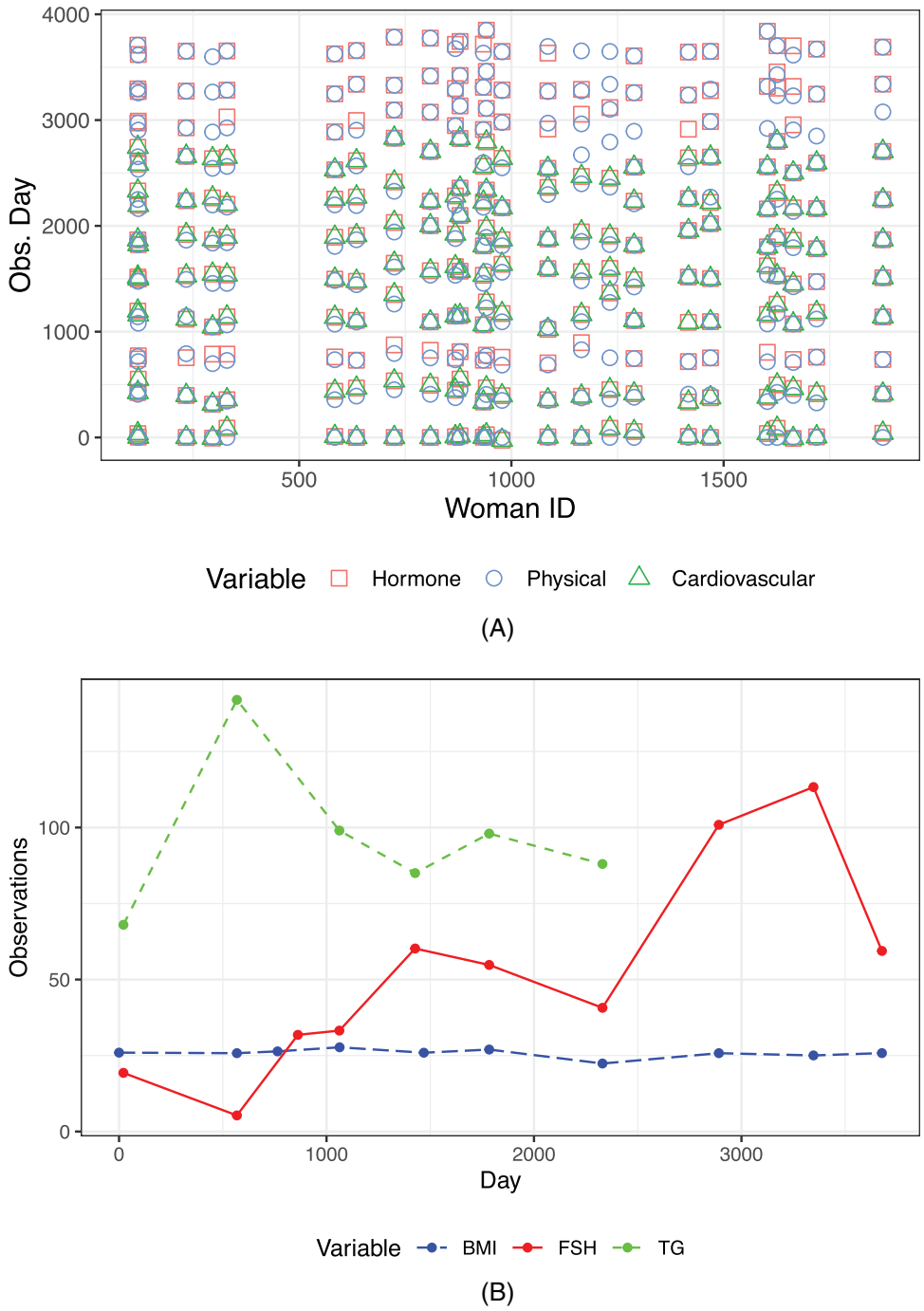


FIGURE 1 (A) Observation days for a randomly selected subset of SWAN participants. Each column corresponds to one woman, with points in different colors and shapes representing variables types: hormone, physical, and cardiovascular measures. (b) Longitudinal trajectories of follicle-stimulating hormone (FSH, red-solid), body mass index (BMI, blue-long dash), and triglycerides (TG, green-short dash) from a randomly selected SWAN participant (ID = 37,522). This figure appears in color in the electronic version of this article, and any mention of color refers to that version.

a random sample of SWAN participants. As seen, these measurements were taken following different schedules. During this important transition, of particular interest is the level of the follicle-stimulating hormone (FSH), our response variable. Two important physical and cardiovascular covariates, the body mass index (BMI) and triglycerides (TG), are also repeatedly measured but on different schedules. Another complication as manifested by Figure 1(B), a spaghetti plot for the longitudinal trajectories of these three variables from a randomly selected participant, is that these asynchronized variables also exhibit short-term fluctuations, which need to be modeled as

measurement error or nugget effect (Carroll et al., 2006). This figure appears in color in the electronic version of this article, and any mention of color refers to that version.

There has been some literature on analyzing incomplete longitudinal data using missing data techniques such as the inverse probability weighting: Robins et al. (1995) assumed the response and time-varying covariates must be missing or present simultaneously; Cook et al. (2004) assumed that the repeated measurements within a subject are complete before the subject dropout from the study. These methods rely on parametric modeling of the missing data mechanism and are not designed for data that are asynchronous by design.

More recently, Cao et al. (2015) modeled asynchronous longitudinal data under generalized linear models with either time-invariant or time-varying coefficients, by proposing kernel-weighted (KW) estimating equation methods to down-weight covariates that are further away in time from the response. These KW estimators are consistent and asymptotically normal, but with slow convergence rates. For time-invariant regression models, their estimated regression coefficients converge in a nonparametric kernel regression rate instead of the usual root-n parametric rate; for time-varying coefficient models, their estimator converges in a bivariate nonparametric smoothing rate, which is much slower than classic convergence rate of varying coefficient models in Cai et al. (2000) and sensitive to bandwidth selection, as shown in our simulation studies. Also, none of the existing methods adequately address the measurement error issue arising from the asynchronous variables.

To address these limitations, we propose to model the longitudinal trajectories of the covariates in the SWAN study as functional data (Ramsay & Silverman, 2005), and use the functional principal component analysis (FPCA) technique (Li & Hsing, 2010) to impute the missing synchronized covariate values from the observed asynchronous, error-prone covariate values. We then use the imputed values in second-stage regression analyses. This method is similar in spirit to the regression calibration method in the measurement error literature (Carroll et al., 2006), but is completely nonparametric. We therefore term our proposed methodology functional calibration for asynchronous regression (FCAR). The proposed method can be easily implemented using existing software, such as the "fdapace" package in R, and combined with other existing regression methodology such as the common linear regression with time-invariant regression coefficients and the time-varying coefficients (Hoover et al., 1998). We show that our estimators for time-invariant coefficient regression models are root-n consistent and asymptotic normal, whereas our estimators for time-varying coefficient models enjoy the optimal convergent rate as Hoover et al. (1998), which is one order of magnitude faster than the existing methods such as Cao et al. (2015). We also show that our method can be extended to accommodate multiple asynchronous longitudinal covariate processes using multivariate FPCA (Dai et al., 2021; Happ & Greven, 2018).

We are aware of the related work on asynchronous longitudinal regression using functional data analysis approaches. For example, Şentürk and Müller (2010) and Şentürk et al. (2013) proposed to estimate the time-varying coefficient function by estimating the covariance function of the time-varying covariate and cross-covariance function between the covariate and response processes using bivariate kernel smoothing. However, their method is associated with the same slower bivariate smoothing convergence rate as Cao et al. (2015). Our simulation study shows our method outperforms Şentürk et al. (2013) and Cao et al. (2015) in efficiency and numerical stability.

The paper is organized as follows. We introduce our model assumptions in Section 2, propose the new functional calibration method, and apply it to longitudinal regression models with time-invariant or time-varying coefficients in Section 3. The asymptotic properties of the proposed estimators are established in Section 4, whereas the practical performance of the proposed methods is illustrated by simulation studies in Section 5. We apply the proposed methods to analyze the SWAN data and investigate the potential effects of BMI and TGs on FSH changes during menopausal transition in Section 6. We provide concluding remarks in Section 7. We present technical proofs and additional numerical results (tables and graphs) and extend the proposed method to multivariate time-varying covariate processes in the Online Appendix.

2 | MODEL ASSUMPTIONS

Let $\{X_i(t), Y_i(t)\}$, $i = 1, \ldots, n$, be independent and identically distributed (iid) bivariate longitudinal processes defined on a compact time interval $\mathcal{T} \subset \mathbb{R}$, where $Y_i(t)$ is the response of the ith subject at time t and $X_i(t)$ is the time-varying covariate process. For simplicity, we focus on the case where $X_i(t)$ is a univariate process, and present its multivariate extension in Online Appendix B. We will consider both the time-invariant coefficient model

$$Y_i(t) = \beta_0 + \beta_1 X_i(t) + \epsilon_i(t), \tag{1}$$

and the time-varying coefficient model

$$Y_i(t) = \beta_0(t) + \beta_1(t) X_i(t) + \epsilon_i(t). \tag{2}$$

In Model (1), $\boldsymbol{\beta} = (\beta_0, \beta_1)^{\mathrm{T}}$ are the time-invariant intercept and slope parameters, whereas $\beta_0(\cdot)$ and $\beta_1(\cdot)$ are

the time-varying counterparts in Model (2). In both models, we assume that $\epsilon_i(t)$ are iid zero-mean error processes with covariance function $\Omega(s,t) = \mathrm{Cov}\{\epsilon(s), \epsilon(t)\}$. We also assume that $X_i(t)$ and $\epsilon_i(t)$ are independent of each other.

Longitudinal variables are observed on discrete time points. Denote by $\mathbf{T}_i = (T_{i1}, \ldots, T_{im_{y,i}})^{\mathrm{T}}$ the time points when $Y_i(\cdot)$ is observed, and by $\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{im_{y,i}})^{\mathrm{T}}$ the observed response vector, where $Y_{ij} = Y_i(T_{ij})$, $j = 1, \ldots, m_{y,i}$. On the other hand, in an asynchronous longitudinal design, $X(t)$ are observed on time points $\mathbf{S}_i = (S_{i1}, \ldots, S_{im_{x,i}})^{\mathrm{T}}$. Let $\mathbf{X}_i = (X_{i1}, \ldots, X_{im_{x,i}})^{\mathrm{T}}$ where $X_{ij} = X_i(S_{ij})$, $j = 1, \ldots, m_{x,i}$. As illustrated in our motivating example, $\mathbf{S}_i$ and $m_{x,i}$ can be completely different from $\mathbf{T}_i$ and $m_{y,i}$. In addition, these time-varying covariates are usually contaminated with measurement error (Liao et al., 2011). These measurement errors, not necessarily arising from instrument error, may be due to local variations. In the SWAN study, as the BMI and TGs level naturally fluctuate over days or even within the same day, it is reasonable to use their long-term or average values as the "true" values to predict the response; failing to take into account the measurement error can result in biased estimates and reduced statistical power (Carroll et al., 2006). To proceed, we relate X_{ij}, the "truth," to its observed error-contaminated surrogates, W_{ij}, via an additive measurement error model:

$$W_{ij} = X_{ij} + U_{ij}, \quad j = 1, \ldots, m_{x,i}, \quad i = 1, \ldots, n, \tag{3}$$

where U_{ij} are iid zero-mean measurement errors with variance σ_u^2 and independent of $\mathbf{X}_i$. Model (1) or (2), coupled with (3), is referred to as *asynchronous longitudinal regression with measurement error.*

Let $\mathbf{W}_i = (W_{i1}, \ldots, W_{im_{x,i}})^{\mathrm{T}}$ be the *observed surrogate* values that are *asynchronous* with $\mathbf{Y}_i$, whereas we denote by $\mathbf{X}_{*i} = \{X_i(T_{i1}), \ldots, X_i(T_{im_{y,i}})\}^{\mathrm{T}}$ the *unobserved true* covariate values that are *synchronized* with $\mathbf{Y}_i$. In the measurement error literature (Carroll et al., 2006), a commonly used technique to impute $\mathbf{X}_{*i}$ from observed surrogate $\mathbf{W}_i$ is regression calibration, which ignores longitudinal correlations, does not capture the dynamic changes in the time-varying covariates, and may incur efficiency loss. We instead propose to calibrate the value of $\mathbf{X}_{*i}$ using a more efficient functional data analysis approach. We assume that the time-varying covariate $X(t), t \in \mathcal{T}$ is a stochastic process defined on $\mathcal{T}$ with mean and covariance functions

$$\mu(t) = \mathrm{E}\{X_i(t)\}, \quad R(s,t) = \mathrm{Cov}\{X_i(s), X_i(t)\}, \quad s, t, \in \mathcal{T}.$$

The covariance function is a smooth, symmetric, positive semidefinite function with a spectral decomposition of $R(s,t) = \sum_{k=1}^{q} \omega_k \psi_k(s)\psi_k(t)$, where $\omega_1 \geq \omega_2 \geq \cdots \geq \omega_q > 0$ are the eigenvalues, and $\psi_k(\cdot)$ are the corresponding eigenfunctions (or principal components) such that $\int_{\mathcal{T}} \psi_k(t)\psi_{k'}(t)dt = I(k = k')$. By the Karhunen–Loève theorem,

$$X_i(t) = \mu(t) + \textstyle\sum_{k=1}^{q} \xi_{ik}\psi_k(t), \quad t \in \mathcal{T}, \tag{4}$$

for $i = 1, \ldots, n$, where $\xi_{ik} = \int_{\mathcal{T}} \{X_i(t) - \mu(t)\}\psi_k(t)dt$ are the principal component scores with mean zero and $\mathrm{Cov}(\xi_{ik}, \xi_{ik'}) = \omega_k I(k = k')$. The number of principal components q can be infinity in theory, but it is common to assume that $X_i(t)$ has a reduced rank representation with a finite q (Li & Hsing, 2010). This is suitable for longitudinal or sparse functional data, where the number of measurements on each trajectories is so small that one cannot realistically estimate a large number of principal components. In practice, q is chosen in a data-driven fashion (Li et al., 2013; Yao et al., 2005), which is to be detailed in Section 3.1.

3 | FUNCTIONAL CALIBRATION FOR ASYNCHRONOUS REGRESSION

3.1 | Calibration using functional principal component analysis

Let $\boldsymbol{\mu}_i = \{\mu(S_{i1}), \ldots, \mu(S_{im_{x,i}})\}^{\mathrm{T}}$ and $\boldsymbol{\psi}_{ik} = \{\psi_k(S_{i1}), \ldots, \psi_k(S_{im_{x,i}})\}^{\mathrm{T}}$, $k = 1, \ldots, q$, be the mean and eigenfunctions interpolated on the observed time points, and put $\boldsymbol{\Psi}_i = [\boldsymbol{\psi}_{i1}, \ldots, \boldsymbol{\psi}_{iq}]$. Under the reduced rank model (4) with a finite rank q, the within-subject covariance matrix for $\mathbf{W}_i$ is $\boldsymbol{\Sigma}_i = \mathrm{Cov}(\mathbf{W}_i) = \boldsymbol{\Psi}_i\boldsymbol{\Lambda}\boldsymbol{\Psi}_i^{\mathrm{T}} + \sigma_u^2 I$, where $\boldsymbol{\Lambda} = \mathrm{diag}(\omega_1, \ldots, \omega_q)$. If $\boldsymbol{\mu}_i$ and $\boldsymbol{\Lambda}$ were known, a roadmap for calibrating the unobserved, synchronized covariates $\mathbf{X}_{*i}$ would be as follows. First, the best linear unbiased predictors (BLUPs) for the FPC scores would be

$$\widetilde{\boldsymbol{\xi}}_i = (\widetilde{\xi}_{i1}, \ldots, \widetilde{\xi}_{iq})^{\mathrm{T}} = \boldsymbol{\Lambda}\boldsymbol{\Psi}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}(\mathbf{W}_i - \boldsymbol{\mu}_i). \tag{5}$$

Second, one could predict the functional trajectory of $X_i(t)$ by

$$\widetilde{X}_i(t) = \mu(t) + \sum_{k=1}^{q} \widetilde{\xi}_{ik}\psi_k(t). \tag{6}$$

Finally, interpolating these predicted trajectories on the observation times of Y, we could predict the unobserved, synchronized covariates $\mathbf{X}_{*i}$ by

$$\widetilde{\mathbf{X}}_{*i} = \boldsymbol{\mu}_{*i} + \sum_{k=1}^{q} \widetilde{\xi}_{ik}\boldsymbol{\psi}_{*ik} = \boldsymbol{\mu}_{*i} + \boldsymbol{\Psi}_{*i}\widetilde{\boldsymbol{\xi}}_i, \tag{7}$$

where $\boldsymbol{\mu}_{*i} = \{\mu(T_{i1}), \ldots, \mu(T_{im_{y,i}})\}^{\mathrm{T}}$, $\boldsymbol{\Psi}_{*i} = \{\boldsymbol{\psi}_{*i1}, \ldots, \boldsymbol{\psi}_{*iq}\}$, and $\boldsymbol{\psi}_{*ik} = \{\psi_k(T_{i1}), \ldots, \psi_k(T_{im_{y,i}})\}^{\mathrm{T}}$, $k = 1, \ldots, q$.

However, as $\boldsymbol{\mu}_i$ and $\boldsymbol{\Lambda}$ are unknown, and in order to complete this calibration roadmap, we need to estimate them based on the observed data $\{\mathbf{W}_i, i = 1, \dots, n\}$. Let $K(\cdot)$ be a kernel function, and denote by $K_h(u) = K(u/h)/h$ where h is the bandwidth. Following Yao et al. (2005) and Li and Hsing (2010), we use local linear smoothers to estimate mean and covariance functions. For any fixed t, we estimate $\mu(t)$ by $\hat{\mu}(t) = \hat{a}_0$, where

$$(\hat{a}_0, \hat{a}_1) = \underset{a_0,a_1}{\operatorname{argmin}} \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m_{x,i}}\sum_{j=1}^{m_{x,i}}\{W_{ij} - a_0 - a_1(S_{ij}-t)\}^2 K_{h_\mu}(S_{ij}-t),$$

with $h_\mu > 0$ being the bandwidth. We then estimate $R(s,t)$ by $\hat{R}(s,t) = \hat{a}_0$ with

$$(\hat{a}_0, \hat{a}_1, \hat{a}_2) = \underset{a_0,a_1,a_2}{\operatorname{argmin}} \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{M_{x,i}}\sum_{j\neq l}\{L_{ij}L_{il} - a_0 - a_1(S_{ij}-s) - a_2(S_{il}-t)\}^2 \times K_{h_R}(S_{ij}-s)K_{h_R}(S_{il}-t)\right],$$

where $L_{ij} = W_{ij} - \hat{\mu}(S_{ij})$, $M_{x,i} = m_{x,i}(m_{x,i}-1)$, and $h_R > 0$ is the bandwidth. Similarly, we can estimate the variance function $V(t) = \mathrm{Var}\{W(t)\} = R(t,t) + \sigma_u^2$ by $\hat{V}(t) = \hat{a}_0$ where

$$(\hat{a}_0, \hat{a}_1) = \underset{a_0,a_1}{\operatorname{argmin}} \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m_{x,i}}\sum_{j=1}^{m_{x,i}}\{L_{ij}^2 - a_0 - a_1(S_{ij}-t)\}^2 K_{h_V}(S_{ij}-t).$$

Then, we estimate σ_u^2 by

$$\hat{\sigma}_u^2 = \frac{1}{|\mathcal{T}|}\int_{\mathcal{T}}\{\hat{V}(t,t) - \hat{R}(t,t)\}dt.$$

To estimate the functional principal components, we take a spectral decomposition of $\hat{R}(s,t)$

$$\hat{R}(s,t) = \sum_k \hat{\omega}_k\hat{\psi}_k(s)\hat{\psi}_k(t),$$

which can be solved numerically by discretizing the smoothed covariance.

Let $\hat{\boldsymbol{\mu}}_i, \hat{\boldsymbol{\psi}}_{ik}, \hat{\boldsymbol{\Psi}}_i, \hat{\boldsymbol{\Lambda}}$, and $\hat{\boldsymbol{\Sigma}}_i$ be the estimated counterparts of $\boldsymbol{\mu}_i, \boldsymbol{\psi}_{ik}, \boldsymbol{\Psi}_i, \boldsymbol{\Lambda}$, and $\boldsymbol{\Sigma}_i$ using the kernel estimators described above. We adopt the PACE method of Yao et al. (2005) to estimate the principal component score ξ_{ik} by a sample version of BLUP (5), that is,

$$\hat{\boldsymbol{\xi}}_i = (\hat{\xi}_{i1}, \dots, \hat{\xi}_{iq})^{\mathrm{T}} = \hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{\Psi}}_i^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_i^{-1}\{\mathbf{W}_i - \hat{\boldsymbol{\mu}}_i\}. \quad (8)$$

With (4), we can recover the covariate process by

$$\hat{X}_i(t) = \hat{\mu}(t) + \sum_{k=1}^{q}\hat{\xi}_{ik}\hat{\psi}_k(t), \quad t \in \mathcal{T}. \quad (9)$$

The number of components q can be selected by minimizing the Akaike information criterion (AIC). There are two commonly used versions of AIC based on a marginal log-likelihood (Rice & Wu, 2001),

$$AIC_{marg}(q) = \sum_{i=1}^{n}\left\{m_i\log(2\pi) + \log(\det\hat{\boldsymbol{\Sigma}}_{iq}) + (\mathbf{W}_i - \hat{\boldsymbol{\mu}}_i)^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_{iq}^{-1}(\mathbf{W}_i - \hat{\boldsymbol{\mu}}_i)\right\} + 2q, \quad (10)$$

and a conditional log-likelihood (Li et al., 2013),

$$AIC_{cond}(q) = N\log\left(N^{-1}\sum_{i=1}^{n}||\hat{\sigma}_u^2\hat{\boldsymbol{\Sigma}}_{iq}^{-1}(\mathbf{W}_i - \hat{\boldsymbol{\mu}}_i)||_2^2\right) + N + 2nq, \quad (11)$$

where $N = \sum_{i=1}^{n} m_{x,i}$, $\hat{\boldsymbol{\Sigma}}_{iq} = \hat{\boldsymbol{\Psi}}_{iq}\hat{\boldsymbol{\Lambda}}_q\hat{\boldsymbol{\Psi}}_{iq}^{\mathrm{T}} + \hat{\sigma}_u^2 I$, $\hat{\boldsymbol{\Lambda}}_q = \mathrm{diag}(\hat{\omega}_1, \dots, \hat{\omega}_q)$, $\hat{\boldsymbol{\Psi}}_{iq} = [\hat{\boldsymbol{\psi}}_{i1}, \dots, \hat{\boldsymbol{\psi}}_{iq}]$, and the subscript '$q$' emphasizes the dependence on the number of FPC's.

3.2 | Asynchronous regression using calibrated covariates

With $\hat{\boldsymbol{\mu}}_{*i}$, $\hat{\boldsymbol{\psi}}_{*ik}$, and $\hat{\boldsymbol{\Psi}}_{*i}$ (the kernel estimates interpolated at $\mathbf{T}_i$ instead of $\mathbf{S}_i$), the empirical version of the calibrated covariate $\tilde{\mathbf{X}}_{*i}$ is $\hat{\mathbf{X}}_{*i} = \hat{\boldsymbol{\mu}}_{*i} + \hat{\boldsymbol{\Psi}}_{*i}\hat{\boldsymbol{\xi}}_i$, and the design matrices using calibrated covariates are $\mathbb{X}_i = (\mathbf{1}, \hat{\mathbf{X}}_{*i})$, $i = 1, \dots, n$. The regression coefficients in Model (1) can be estimated by

$$\hat{\boldsymbol{\beta}} := (\hat{\beta}_0, \hat{\beta}_1)^{\mathrm{T}} = (\mathbb{X}^{\mathrm{T}}\mathbb{X})^{-1}\mathbb{X}^{\mathrm{T}}\mathbf{Y}, \quad (12)$$

where $\mathbf{Y} = (\mathbf{Y}_1^{\mathrm{T}}, \dots, \mathbf{Y}_n^{\mathrm{T}})^{\mathrm{T}}$ and $\mathbb{X} = (\mathbb{X}_1^{\mathrm{T}}, \dots, \mathbb{X}_n^{\mathrm{T}})^{\mathrm{T}}$.

For Model (2), we estimate $\boldsymbol{\beta}(t) := \{\beta_0(t), \beta_1(t)\}^{\mathrm{T}}$ by using the local linear estimator of Hoover et al. (1998), with the unobserved $X_i(T_{ij})$ replaced by the calibrated value $\hat{X}_i(T_{ij})$ as defined in (9). That is, we estimate $\boldsymbol{\beta}(t)$ by $\hat{\mathbf{b}}_0$ from

$$(\hat{\mathbf{b}}_0, \hat{\mathbf{b}}_1) = \underset{\mathbf{b}_0,\mathbf{b}_1}{\operatorname{argmin}}\sum_{i=1}^{n}\sum_{j=1}^{m_{y,i}}\{Y_{ij} - \mathbb{X}_{ij}^{\mathrm{T}}\mathbf{b}_0 - \mathbb{X}_{ij}^{\mathrm{T}}\mathbf{b}_1(T_{ij}-t)\}^2 K_h(T_{ij}-t), \quad (13)$$

where $\mathbb{X}_{ij} = \{1, \widehat{X}_i(T_{ij})\}^{\mathrm{T}}$ and $h > 0$ is the bandwidth.

4 | ASYMPTOTIC THEORY

4.1 | Preliminaries

For ease of exposition, we assume in the asymptotic theory that both $X(t)$ and $Y(t)$ have been centered, such that $\mu(t) \equiv 0$ in (4), $\beta_0 = 0$ in (1), and $\beta_0(t) = 0$ in (2). We focus on estimating the slope parameter β_1 and the slope function $\beta_1(t)$ in Models (1) and (2), respectively; extensions to noncentered situations are straightforward but with more notation. For any positive constant sequences $\{a_n\}$ and $\{b_n\}$, denote by $a_n \prec b_n$ if $a_n/b_n \to 0$ as $n \to \infty$.

Recall $\mathbf{X}_{*i}$ is the unobserved covariate vector synchronized with the response $\mathbf{Y}_i$, $\widetilde{\mathbf{X}}_{*i} = \boldsymbol{\Psi}_{*i}\widetilde{\boldsymbol{\xi}}_i$ is the BLUP of $\mathbf{X}_{*i}$ as defined in (7), and $\widehat{\mathbf{X}}_{*i}$ is the empirical version of $\widetilde{\mathbf{X}}_{*i}$ replacing the unknown functions with their kernel estimators. Let $\boldsymbol{\epsilon}_i = \{\epsilon_i(T_{i1}), \ldots, \epsilon_i(T_{im_{y,i}})\}^{\mathrm{T}}$ be the vector of measurement error as defined in model (1) or (2), with the covariance matrix $\boldsymbol{\Omega}_i = \{\Omega(T_{ij}, T_{ij'})\}_{j,j'=1}^{m_{y,i}}$.

We assume that the numbers of observations $(m_{x,i}, m_{y,i})$ are random with $P(2 \le m_{x,i}, m_{y,i} \le M) = 1$ for a constant $M < \infty$. Given $m_{x,i}$, S_{ij} are iid copies of the random variable S with a density $f_S(\cdot)$; and given $m_{y,i}$, T_{ij} are iid copies of the random variable T with a density $f_T(\cdot)$. Both $f_T(\cdot)$ and $f_S(\cdot)$ are strictly greater than 0, with bounded derivatives on $\mathcal{T}$. Assume that $\{X_i(t), \epsilon_i(t), m_{x,i}, m_{y,i}, \mathbf{T}_i, \mathbf{S}_i\}$ are iid tuples across i. In addition, we assume the following conditions hold.

(C.1) The kernel function $K(\cdot)$ is a symmetric probability density function on $[-1, 1]$ with

$$\sigma_K^2 := \int_{-1}^{1} u^2 K(u)du < \infty, \quad \nu_0 := \int_{-1}^{1} K^2(u)du < \infty.$$

(C.2) All eigenfunctions $\psi_k(t)$, $k = 1, \ldots, q$, are twice differentiable, and the second derivatives $\psi_k''(t)$ are uniformly continuous on $\mathcal{T}$.

(C.3) There exists a constant $C > 4$ such that $\mathrm{E}(|U_{ij}|^C) + \mathrm{E}\{\sup_{t\in\mathcal{T}}|X_i(t)|^C\} < \infty$.

(C.4) Assume $h_R \to 0$, $h_V \to 0$ as $n \to \infty$, $(\log n/n)^{1/3} \prec h_R \prec n^{-1/4}$ and $(\log n)n^{-3/4} \prec h_V \prec n^{-1/4}$.

Remark 1. These conditions are common in functional data analysis. Under them, Li and Hsing (2010) proved that the FPCA estimators possess the following uniform convergence properties:

$$\sup_{s,t\in\mathcal{T}}|\widehat{R}(s,t) - R(s,t)| = O\{h_R^2 + \sqrt{\log n/(nh_R^2)}\} \quad a.s.,$$

$$\widehat{\sigma}_u^2 - \sigma_u^2 = O\{h_R^2 + \sqrt{\log n/(nh_R)} + h_V^2 + \log n/(nh_V)\} \quad a.s.,$$

$$\sup_{t\in\mathcal{T}}|\widehat{\psi}_k(t) - \psi_k(t)| = O\{h_R^2 + \sqrt{\log n/(nh_R)}\} \quad a.s.,$$

$$\widehat{\omega}_k - \omega_k = O(\sqrt{\log n/n}) \quad a.s. \quad k = 1, \ldots, q.$$

Recall $\widetilde{X}_i(t)$ defined in (6) is the BLUP for $X_i(t)$ and $\widehat{X}_i(t)$ in (9) is its empirical counterpart. By straightforward calculations,

$$\sup_{t\in\mathcal{T}}|\widehat{X}_i(t) - \widetilde{X}_i(t)| = O\{h_R^2 + \sqrt{\log n/(nh_R)} + h_V^2 + \log n/(nh_V)\} \quad a.s. \tag{14}$$

It is well known in the measurement error literature (Carroll et al., 2006), replacing $X_i(t)$ with the calibrated value $\mathrm{E}\{X_i(t) \mid \mathbf{W}_i\}$ will result in consistent but less efficient estimators. Equation (14) shows that our functional calibration $\widehat{X}_i(t)$ uniformly converges to the BLUP $\widetilde{X}_i(t)$, and hence, our estimators in (12) and (13) are consistent; however, the derivation of their asymptotic distributions needs much involved analysis.

4.2 | Asymptotic properties of FCAR estimator for the time-invariant regression model

The following theorem establishes the asymptotic property of the coefficient estimator (12) under the time-invariant regression model (1).

Theorem 1. *Under the assumptions above, the estimated slope parameter for model (1) has the following asymptotic distribution:*

$$\sqrt{n}(\widehat{\beta}_1 - \beta_1) \xrightarrow{d} Normal\{0, (\gamma_1 + \beta_1^2\gamma_2)/\gamma_x^2\},$$

where

$$\gamma_x = \mathrm{E}(\widetilde{\mathbf{X}}_{*i}^{\mathrm{T}}\widetilde{\mathbf{X}}_{*i}) = \mathrm{E}\left\{tr\left(\boldsymbol{\Psi}_{*i}\boldsymbol{\Lambda}\boldsymbol{\Psi}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\Psi}_i\boldsymbol{\Lambda}\boldsymbol{\Psi}_{*i}^{\mathrm{T}}\right)\right\},$$

$$\gamma_1 = \mathrm{E}\left\{tr(\boldsymbol{\Psi}_{*i}\boldsymbol{\Lambda}\boldsymbol{\Psi}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\Psi}_i\boldsymbol{\Lambda}\boldsymbol{\Psi}_{*i}^{\mathrm{T}}\boldsymbol{\Omega}_i)\right\},$$

$$\gamma_2 = \mathrm{Var}\left[\sum_{j=1}^{m_{y,i}}\widetilde{X}_i(T_{ij})\{X_i(T_{ij}) - \widetilde{X}_i(T_{ij})\}\right.$$

$$+ \frac{1}{M_{x,i}} \sum_{j \neq j'} u^*_{i,jj'} \mathcal{A}(S_{ij}, S_{ij'}) \Bigg], \tag{15}$$

$u^*_{i,jj'} = W_{ij}W_{ij'} - R(S_{ij}, S_{ij'})$, $\mathcal{A}(s_1, s_2)$ *is defined in Lemma 2 in the Online Appendix, and the expectations are taken over* $(\mathbf{X}_i, \boldsymbol{\epsilon}_i, m_{x,i}, m_{y,i}, \mathbf{T}_i, \mathbf{S}_i)$.

Remark 2. Under the special case that the BLUP $\widetilde{X}_i(t)$ is also the conditional mean $\mathrm{E}\{X_i(t)|\mathbf{W}_i\}$, for example, when $\boldsymbol{\xi}_i$ and $\mathbf{U}_i$ are jointly Gaussian, $\{\widetilde{X}_i(t) - X_i(t),\ t \in \mathcal{T}\}$ is uncorrelated with any function of $\mathbf{W}_i$. One can show, under such an circumstance, $\gamma_2 = \gamma_{21} + \gamma_{22}$, where $\gamma_{21} = \mathrm{Var}[\sum_{j=1}^{m_{y,i}} \widetilde{X}_i(T_{ij})\{X_i(T_{ij}) - \widetilde{X}_i(T_{ij})\}]$ and $\gamma_{22} = \mathrm{Var}\{M_{x,i}^{-1} \sum_{j \neq j'} u^*_{i,jj'} \mathcal{A}(S_{ij}, S_{i'j'})\}$. Under the additional Gaussian assumption on $\boldsymbol{\xi}_i$ and $\mathbf{U}_i$, we can also obtain

$$\gamma_{21} = \mathrm{E}\Big[\mathrm{tr}\Big\{\boldsymbol{\Psi}_{*i}\boldsymbol{\Lambda}\boldsymbol{\Psi}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\Psi}_i\boldsymbol{\Lambda}\boldsymbol{\Psi}_{*i}{}^{\mathrm{T}}\boldsymbol{\Psi}_{*i}\boldsymbol{\Lambda}(\mathbf{I} - \boldsymbol{\Psi}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\Psi}_i\boldsymbol{\Lambda})\boldsymbol{\Psi}_{*i}{}^{\mathrm{T}}\Big\}\Big].$$

As in the classic regression calibration literature (Carroll et al., 2006), one can define $\widetilde{\boldsymbol{\beta}}$ to be the counterpart of $\widehat{\boldsymbol{\beta}}$ in (12) replacing $\widehat{X}_i(t)$ with $\widetilde{X}_i(t)$, as if all the functional and scalar parameters in (3) and (4) are known, then $(\gamma_1 + \beta_1^2\gamma_{21})/\gamma_x^2$ is the asymptotic variance of $\widetilde{\beta}_1$. In our problem, $\beta_1^2\gamma_{22}/\gamma_x^2$ is the additional variation caused by the FPCA estimation errors, that is, those caused by substituting $\mu(t)$, $\psi_k(t)$, ω_k, and σ_u^2 with their functional estimators described in Section 3.1. .

While it is tempting to treat the calibrated values $\widehat{\mathbf{X}}_{*i}$ as the truth and use the naive standard error for linear regression to infer β_1, the decomposition of the asymptotic variance in Theorem 1 suggests that this approach ignores the extra variations caused by calibration of the covariate values as well as estimation errors from FPCA. As a result, the naive approach leads to an underestimated variation, a low coverage rate in confidence intervals and illegitimate inferences. We recommend estimating the standard error of $\widehat{\boldsymbol{\beta}}$ using bootstrap, where we resample the subjects and repeat the FPCA procedure to the bootstrap samples to properly account for these extra variations.

4.3 | Asymptotic properties of FCAR estimator for the time-varying regression model

Again, we assume that both $X(t)$ and $Y(t)$ are centered so that $\beta_0(t) = 0$ and we can focus on estimating $\beta_1(t)$ in Model (2). We also make the additional assumptions.

(C.5) The slope function $\beta_1(t)$ is twice continuously differentiable on $\mathcal{T}$.

(C.6) The kernel function in fitting the time-varying regression model in (13) is Lipschitz continuous and satisfies 4.1.

(C.7) The bandwidth h in (13) satisfies $h_R/h \to 0$, $\log(n)h^5/h_R \to 0$, $nh^7 \to 0$ and $nh \to \infty$.

Theorem 2. *Under the framework outlined in Section 4.1 and assumptions 4.1—4.3, the estimated slope function for model (2) has the following asymptotic distribution:*

$$\sqrt{nh}\Gamma_0(t)\{\widehat{\beta}_1(t) - \beta_1(t) - \frac{1}{2}\beta_1^{(2)}(t)\sigma_K^2 h^2\} \xrightarrow{d}$$

$$Normal[0, \Gamma_1(t) + \beta_1^2(t)\{\Gamma_2(t) + \Gamma_3(t)\}],$$

for any $t \in \mathcal{T}$, *where* $\bar{m}_y = \mathrm{E}(m_{y,i})$, $\Gamma_0(t) = \bar{m}_y f_T(t)\Gamma_x(t)$,

$$\Gamma_x(t) = \mathrm{Var}\{\widetilde{X}_i(t)\} = \boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\Lambda}\mathrm{E}(\boldsymbol{\Psi}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\Psi}_i)\boldsymbol{\Lambda}\boldsymbol{\psi}(t),$$

$$\Gamma_1(t) = \bar{m}_y\Gamma_x(t)\Omega(t,t)f_T(t)\nu_0,$$

$$\Gamma_2(t) = \bar{m}_y\mathrm{E}[\widetilde{X}^2(t)\{X(t) - \widetilde{X}(t)\}^2]\nu_0,$$

$$\Gamma_3(t) = \bar{m}_y^2 f_S(t)\nu_0 \Bigg[\mathrm{E}(M_{x,i}^{-1}) \int \Pi(t, s_2, s_2)\mathcal{Q}^2(s_2, t)f_S(s_2)ds_2$$

$$+ \ \mathrm{E}\{M_{x,i}^{-1}(m_{x,i} - 2)\} \int \Pi(t, s_2, s_3)$$

$$\mathcal{Q}(s_2, t)\mathcal{Q}(s_3, t)f_S(s_2)f_S(s_3)ds_2ds_3\Bigg],$$

$\mathcal{Q}(s,t) = \boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\Lambda}\mathrm{E}(\boldsymbol{\Psi}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\Psi}_i)\boldsymbol{\psi}(s)f_T(t)/\{f_S(s)f_S(t)\}$,
$\Pi(s_1, s_2, s_3) = \mathrm{E}\{X^2(s_1)X(s_2)X(s_3)\} + R(s_2, s_3)\sigma_u^2 - R(s_1, s_2)R(s_1, s_3) + I(s_2 = s_3)\{R(s_1, s_1)\sigma_u^2 + \sigma_u^4\}$, *and the expectations are taken over* $(m_{x,i}, m_{y,i}, \mathbf{T}_i, \mathbf{S}_i, \mathbf{X}_i, \boldsymbol{\epsilon}_i)$.

Remark 3. Theorem 2 suggests that our estimator enjoys the optimal convergence rate in varying coefficient models as established in Cai et al. (2000), which is much faster than those for the competing method of Cao et al. (2015) and Şentürk and Müller (2010). Analogous to Theorem 1, $\Gamma_0^{-2}(t)\{\Gamma_1(t) + \beta_1^2(t)\Gamma_2(t)\}$ is the asymptotic variance of $\widetilde{\beta}_1(t)$, obtained by using $\widetilde{\mathbf{X}}_{*i}$ as the predictors in the varying coefficient model (2), and $\Gamma_0^{-2}(t)\beta_1^2(t)\Gamma_3(t)$ is the extra variation caused by the FPCA errors. We therefore recommend to make inference on $\boldsymbol{\beta}(t)$ using a bootstrap procedure that accounts for the FPCA estimation error as described in Remark 2. Also Assumptions 4.1 and 4.3 require undersmoothing in the FPCA procedure; we need $h_R/h \to 0$ so that the biases caused by FPCA estimation are asymptotically negligible compared with the smoothing bias in varying coefficient models.

TABLE 1 Simulation 1: performance of $\widehat{\beta}_1$ under the proposed FCAR method under Settings 1 and 2. SD: standard deviation; Naive SE: mean of the naive standard error; Naive CP: coverage rate of a 95% confidence interval using the naive SE; Bootstrap SE: mean of the bootstrap standard error; Bootstrap CP: coverage rate of a 95% confidence interval using the bootstrap SE; IE: independent errors; DE: dependent errors; MEF: model-error free with dependent errors.

	Setting I			Setting II		
Error type	**IE**	**DE**	**MEF**	**IE**	**DE**	**MEF**
Bias	0.007	0.004	−0.002	−0.008	−0.013	0.025
SD	0.028	0.029	0.017	0.127	0.122	0.060
Naive SE	0.019	0.017	0.012	0.064	0.058	0.034
Naive CP	0.830	0.770	0.820	0.670	0.640	0.725
Bootstrap SE	0.030	0.030	0.019	0.117	0.119	0.064
Bootstrap CP	0.955	0.950	0.965	0.925	0.930	0.940

5 | SIMULATION STUDIES

We conduct simulations to examine the finite sample performances of the time-invariant regression model (1) and time-varying coefficient model (2), and compare them with those of various exiting methods.

5.1 | Simulation 1: FCAR for time-invariant coefficient model

Let the time domain be $\mathcal{T} = [0, 10]$, $X_i(t)$ be iid copies of a stochastic process described by model (4) with $q = 3$ principal components, and $\boldsymbol{\xi}_i \sim \text{Normal}\{\mathbf{0}, \text{diag}(4, 2, 1)\}$. Set $n = 200$ and generate $Y_i(t)$ from Model (1) with $\beta_0 = 1$ and $\beta_1 = 2$. Suppose there are $m = 5$ discrete observations on $X_i(t)$ and $Y_i(t)$, respectively, where $\{S_{i1}, \ldots, S_{im}\}$ and $\{T_{i1}, \ldots, T_{im}\}$ are generated independently from a uniform distribution on $\mathcal{T}$. Error-contaminated discrete observations $\mathbf{W}_i$ are generated from Model (3) with $U_{ij} \sim \text{Normal}(0, 1)$. We consider two settings for the mean and eigenfunctions of $X_i(t)$:

Setting I: $\mu(t) = t + \sin(t)$, $\psi_k(t) = (1/\sqrt{5})\sin(\pi k t/10)$, $t \in \mathcal{T}$, $k = 1, 2, 3$;

Setting II: $\mu(t) = \sin(t)$, $\psi_1(t) = \sqrt{10}$, $\psi_2(t) = \sqrt{5}\sin(2\pi t/10)$, $\psi_3(t) = \sqrt{5}\cos(2\pi t/10)$.

We generate residual $\epsilon_i(t)$ from a zero-mean Gaussian process with covariance function $\Omega(s, t) = \text{Cov}\{\epsilon_i(s), \epsilon_i(t)\}$, and consider two different covariance structures: (1) independent (IE) with $\Omega(s, t) = 1.5I(s = t)$ and (2) dependent (DE) with $\Omega(s, t) = 2^{-|t-s|/5}$. As an ideal case, we also consider a measurement-error free (MEF) scenario under the DE structure, where X_{ij}'s in (3) are correctly observed and the covariate measurement error $U_{ij} = 0$.

For each setting and each error correlation structure, we simulate 200 data sets and apply the proposed FCAR method to each simulated data set. Specifically, FPCA is performed using the `fdapace` package of R with its built-in bandwidth selector and q is selected by the marginal likelihood AIC. In Table 1, we summarize the performance of $\widehat{\beta}_1$ under both settings and all three measurement error structures (IE, DE, and MEF). The criteria include the bias, standard deviation, mean of the naive standard error pretending the calibrated values are the true covariates, coverage rate of a 95% confidence interval using the naive SE, mean of the bootstrap standard error, and coverage rate of a 95% confidence interval using the bootstrap SE. The results on $\widehat{\beta}_0$ are similar but less interesting and hence relegated to Online Appendix C. It appears that the bias of our estimator is much smaller than the standard deviation, corroborating Theorem 1 that $\widehat{\beta}_1$ is asymptotically unbiased. The results also support Remark 2 that the naive standard error estimator underestimates the standard error and results in confidence intervals with lower than nominal coverage rates. In contrast, bootstrap standard errors capture the extra variations caused by calibrating the covariate value and FPCA estimation errors, and as a result, the confidence intervals based on bootstrap standard errors yield coverage rates close to the nominal ones. As noted in Remark 2, we perform FPCA to each bootstrap sample, and Table 1 is based on 500 bootstrap samples.

Table 2 compares the proposed FCAR method with the KW method (Cao et al., 2015) on biases, Monte Carlo standard deviations, and the average estimated standard errors. For the KW method, we use the function `asynchTI` from the R package `AsynchLong` (Cao et al., 2015), which provides a built-in standard error estimator. For FCAR method, the standard error refers to the bootstrap standard error in Table 1. It is noteworthy that, under the IE and DE covariance structures, the magnitude of the KW biases still dominates that of the standard errors, yielding confidence intervals with a coverage rate close to 0. In contrast, the proposed FCAR estimator incurs negligible biases and produces confidence intervals with a coverage rate close to the nominal level. The coverage rate of the KW confidence intervals improves much under the ideal MEF (no measurement errors) scenario, but is still lower than that of the FCAR confidence intervals. For the coefficient

TABLE 2 Simulation 1: comparison of $\widehat{\beta}_1$ using the proposed FCAR method with the kernel weighted (KW) method of (Cao et al., 2015) on bias, standard deviation (SD), mean of standard error (SE), and coverage rate of a 95% confidence interval using standard error (CP) under two settings and three error structures (IE: independent residuals; DE: dependent residuals; MEF: measurement error free with dependent residuals).

		IE		DE		MEF	
		FCAR	KW	FCAR	KW	FCAR	KW
Setting I	Bias	0.007	−0.225	0.004	−0.213	−0.002	−0.024
	SD	0.028	0.067	0.029	0.067	0.017	0.045
	SE	0.030	0.049	0.030	0.051	0.019	0.032
	CP	0.955	0.060	0.950	0.065	0.965	0.830
Setting II	Bias	−0.008	−0.978	−0.013	−0.978	0.025	−0.060
	SD	0.127	0.097	0.122	0.092	0.060	0.108
	SE	0.117	0.076	0.119	0.077	0.064	0.069
	CP	0.925	0.000	0.930	0.000	0.940	0.735

estimation, FCAR and KW are on par in computational intensity, taking 16.08 and 14.76 s, respectively; the calculation of standard errors for FCAR takes 1.96 min in 10-cores parallel for each data set on average, a bit more than 0.13 s taken by KW. This is reasonable, as FCAR needs a bootstrap procedure to compute standard errors, whereas KW does not. Including an additional setting of $m_i = 15$ per subject under the DE structure, we investigate the impact of the sparsity level on the performance and find that the performance is fairly robust; see Online Appendix C.

5.2 | Simulation 2: FCAR for time-varying coefficient model

We simulate data from the time-varying coefficients regression model (2). As in Simulation 1, we set the time domain to be $\mathcal{T} = [0, 10]$ and simulate $n = 200$ subjects with $m_i = 5$ repeated measures on $X_i(t)$ and $Y_i(t)$ allowing the measuring time points to be asynchronous between X and Y. We simulate $X(t)$ using the Karhunen–Loève expansion (4), with mean function $\mu(t) = t + \sin(t)$, $q = 3$, $\boldsymbol{\xi}_i \sim \text{Normal}\{\mathbf{0}, \text{diag}(4, 2, 1)\}$ and $\psi_k(t) = (1/\sqrt{5})\sin(\pi kt/10)$, $k = 1, 2, 3$. We simulate discrete observations W_{ij} from (3) where U_{ij} are iid standard normal, and simulate Y_{ij} from (2), where the measurement error $\epsilon_i(t)$ is generated from a mean zero Gaussian process with covariance $\text{Cov}\{\epsilon_i(s), \epsilon_i(t)\} = 2^{-|t-s|/5}$. For each subject, the observation time points $\{S_{ij}\}$ and $\{T_{ij}\}$ are uniformly distributed on $\mathcal{T}$ and independent from each other. We consider the following two settings for the time-varying coefficients:

Setting I: $\beta_0(t) = 0.2t + 0.5$, $\beta_1(t) = \sin(\pi t/10)$;
Setting II: $\beta_0(t) = t^{1/2}$, $\beta_1(t) = \sin(\pi t/5)$.

We perform functional calibration using the `fdapace` package with *AIC* as the principal component selection criterion. To fit a time-varying coefficients model after the functional calibration, we used the `tvLM` function in the R package `tvReg` that implements the kernel smoothing method in Hoover et al. (1998) and its built-in cross-validation procedure to choose the bandwidth. As a comparison, we consider the following estimators, that is, the Oracle estimator with the known synchronized true values of X, the KW estimator (Cao et al., 2015), and the functional varying coefficients model (FVCM) (Şentürk et al., 2013). The Oracle estimator is implemented by using the `tvReg` package, and the KW method for time-varying coefficient model is implemented by using the authors's own `asynchTD` function in the `AsynchLong` package. The FVCM method requires estimation of the covariance function of $X(t)$ and the cross-covariance function between $X(\cdot)$ and $Y(\cdot)$, which are calculated using the `fdapace` package. Bandwidths for all methods are selected using the built-in options of the packages mentioned above: generalized cross-validation (GCV) of `fdapace`, cross-validation of `tvReg`, and adaptive selection procedure of `AsynchLong`.

We repeat the simulation 200 times for both settings and apply the proposed and competing methods to each data set. Following Şentürk and Müller (2010), we compare different methods using two evaluation criteria: the mean absolute deviation error (MADE) and the weighted average squared error (WASE)

$$\text{MADE} = \frac{1}{2|\mathcal{T}|}\sum_{r=0}^{1}\frac{\int_{\mathcal{T}}|\widehat{\beta}_r(t) - \beta_r(t)|dt}{\text{range}(\beta_r)}, \qquad \text{WASE} = \frac{1}{2|\mathcal{T}|}\sum_{r=0}^{1}\frac{\int_{\mathcal{T}}\{\widehat{\beta}_r(t) - \beta_r(t)\}^2 dt}{\text{range}^2(\beta_r)},$$

where $\text{range}(\beta_r)$ is the range of function $\beta_r(t)$, $r = 0, 1$.

As summarized in Table 3, the proposed FCAR method yields MADE and WASE that are close to the Oracle estimator. The FVCM and KW methods equipped with built-in

TABLE 3 Simulation 2: MADE and WASE of various methods. SD: standard deviation); 25%: 25% quantile; 75%: 75% quantile.

	Method	Criterion	Mean(SD)	Median	25%	75%
Setting I	FCAR	MADE	0.319(0.151)	0.302	0.207	0.388
	FVCM		1.494(2.521)	0.948	0.756	1.418
	KW		1.452(3.533)	1.026	0.868	1.297
	Oracle		0.209(0.091)	0.192	0.142	0.269
	FCAR	WASE	0.345(0.395)	0.224	0.103	0.402
	FVCM		461.495(5235.731)	3.433	1.576	9.553
	KW		1057.299(14234.414)	3.830	1.716	15.221
	Oracle		0.216(0.242)	0.111	0.058	0.294
Setting II	FCAR	MADE	0.263(0.104)	0.246	0.186	0.321
	FVCM		0.944(1.551)	0.616	0.440	0.913
	KW		1.153(1.669)	0.720	0.605	1.115
	Oracle		0.180(0.062)	0.172	0.137	0.220
	FCAR	WASE	0.316(0.351)	0.200	0.093	0.394
	FVCM		299.501(3746.442)	1.464	0.613	5.921
	KW		204.128(1538.798)	2.709	1.196	15.833
	Oracle		0.287(0.347)	0.183	0.088	0.330

tuning parameter selectors perform worse than FCAR, likely because both of them, by evoking bivariate kernel smoothing while estimating univariate coefficient functions in Model (2), are numerically unstable. Both FVCM and KW present large mean WASEs (Figure 2), which are further magnified by the square operator.

MADE and WASE are overall numerical summaries combining $\widehat{\beta}_0(\cdot)$ and $\widehat{\beta}_1(\cdot)$; we also provide graphical summaries of $\widehat{\beta}_0(\cdot)$ and $\widehat{\beta}_1(\cdot)$ separately. In Figure 2, we summarize $\widehat{\beta}_1(t)$ under Setting I by all four methods mentioned above; a similar graph (Figure C.1) under Setting I\negthinspace I is provided in the Online Appendix. The proposed FCAR estimator for $\beta_1(t)$ has negligible biases and overall performance comparable to the Oracle estimator. In contrast, with slower convergence rates and numerical instability of bivariate kernel smoothing, the KW and FVCM estimators for $\beta_1(t)$ are highly variable and affected by the boundary effect. Graphical summaries of $\widehat{\beta}_0(t)$ allude to the same message. We therefore relegate the graphs on $\widehat{\beta}_0(t)$ under these two settings to Figures C.2 and C.3 in Online Appendix C. Across both settings, FCAR takes an average of 0.45 min to estimate the time-varying coefficient functions. A typical run of the bootstrap procedure takes an additional 8.47 min under a 10-core parallel. In contrast, KW takes a combined 5.18 min for estimation and inference. The FVCM takes an average of 1.93 min in estimation and relies on a bootstrap procedure similar to ours to make inference.

As KW and FVCM are visually sensitive to the boundary effects, we furnish summary tables and plots with the 95% truncated domains in Online Appendix C for a more fair comparison. In addition, we consider the case of MEF with an added sparsity level of $m_i = 15$, which again demonstrates the fine performance of the proposed FCAR estimator; see Online Appendix C.

6 | REAL DATA ANALYSIS

We apply the proposed FCAR method to the SWAN data described in Section 1. The study admitted 3302 premenopausal or early perimenopausal women between 1996 and 1997, with the baseline age ranging from 42 to 53. These women were scheduled to have annual follow-ups up to 10 years, although various hormonal, physical, and cardiovascular biomarkers were measured according different schedules as illustrated in Figure 1 until the study ended in 2008. One of the most important biomarkers in menopausal studies is the FSH level, the outcome variable of our primary interest.

As declining follicular reserve is the immediate cause of the perimenopausal and menopausal transitions (Richardson et al., 1987), an increase in the serum FSH level was one of the major endocrine changes associated with menopausal transitions (Burger et al., 1995). FSH levels rise progressively before the final menses and will continue for 2–4 years, before remaining elevated postmenopause (Burger et al., 1999). Changes in FSH levels have been linked to or are precursors of various medical conditions. For example, abnormal variations of FSH levels are related to the depressive symptoms during the menopausal transition (Bromberger et al., 2010), and may also increase women's risk of developing cardiovascular disease after menopause (El Khoudary et al., 2016).

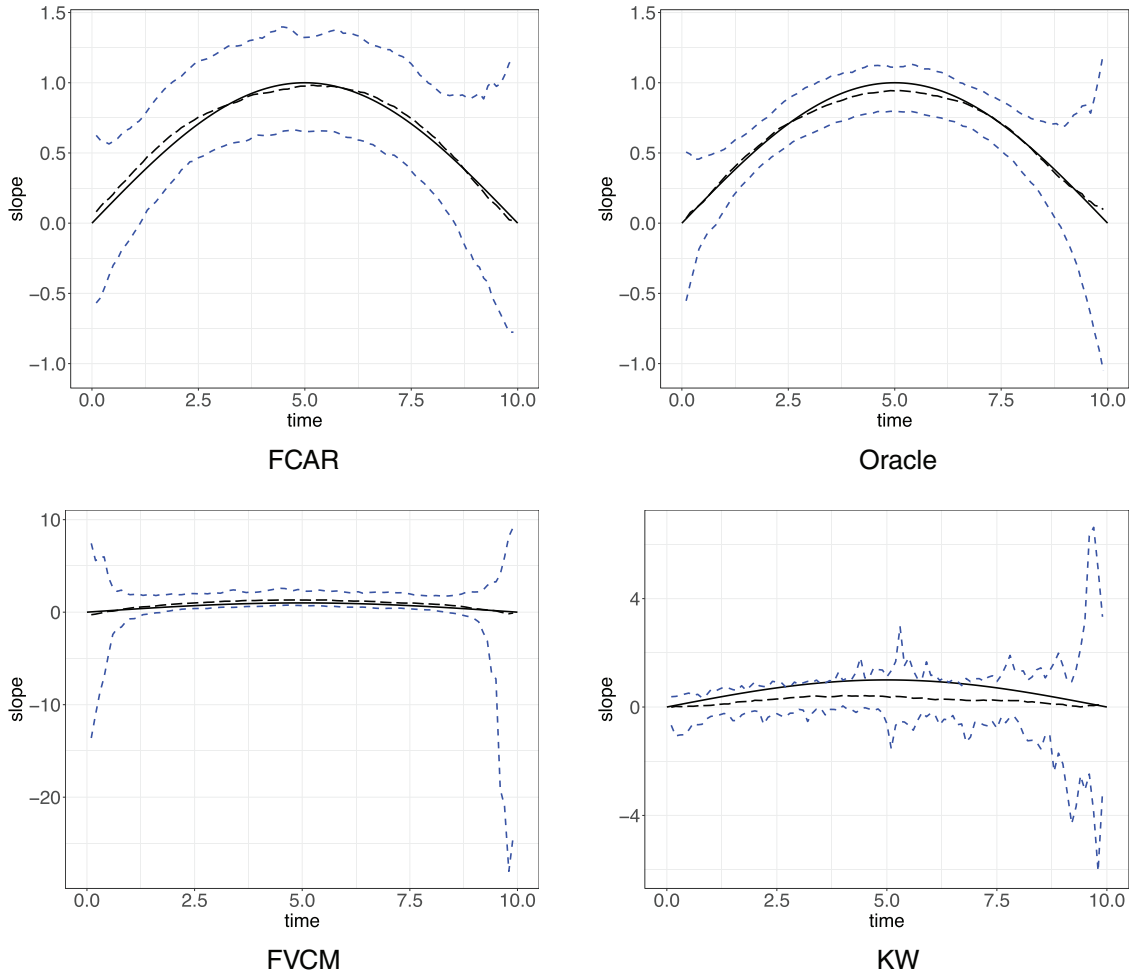


FIGURE 2 Summary of $\widehat{\beta}_1(t)$ under Simulation 2, Setting I using various methods. In each panel, dark solid: true $\beta_1(t)$; dark long dash: median of $\widehat{\beta}_1(t)$; blue short dash: 0.975 and 0.025 quantiles. This figure appears in color in the electronic version of this article, and any mention of color refers to that version.

Therefore, studying the dynamic relationship between FSH and other physiological measurements is of great importance to understand women's reproductive life and their midlife health (Bromberger et al., 2010). Following Wang et al. (2020), we study the association between FSH and TGs adjusting for age, income level, and BMI. FSH was measured every year for the SWAN participants following the hormone measurement schedule in Figure 1, whereas TG and BMI were following the cardiovascular and physical measurement schedules in Figure 1. Of note, TG was not collected in year 2 or beyond year 8, and 47.5% of BMI measurements were asynchronous with FSH. We also included the baseline age and income as time-invariant covariates, where the income was dichotomized (1 if annual income is more than \$50k and 0 otherwise). After removing subjects with missing incomes, there are 1634 high-income subjects and 1578 low-income subjects in the data. We focus our analysis on the first 8 years of the study, when FSH, TG, and BMI are all available. An added rationale behind this truncation is that all participants experienced the entire menopausal transition by year 8, becoming postmenopausal or late perimenopausal afterward (Bromberger et al., 2010).

Existing works, such as Wang et al. (2020), assume that the association between FSH and the covariates is time-invariant and ignores asynchronous issue in this data set, which may mask some intriguing time-varying associations. Instead, we apply the time-varying coefficients model to model the dynamic relationship between FSH and other time varying or invariant covariates. Among the competing methods described in Section 5.2, the KW estimator requires that all time-varying covariates are measured at the same time points, which is not applicable in our data since TG and BMI are measured on different time as well; the functional varying coefficient model (FVCM) of Şentürk and Müller (2010) was proposed for univariate time-varying covariates and is not readily applicable to multiple time-varying covariates in this data set.

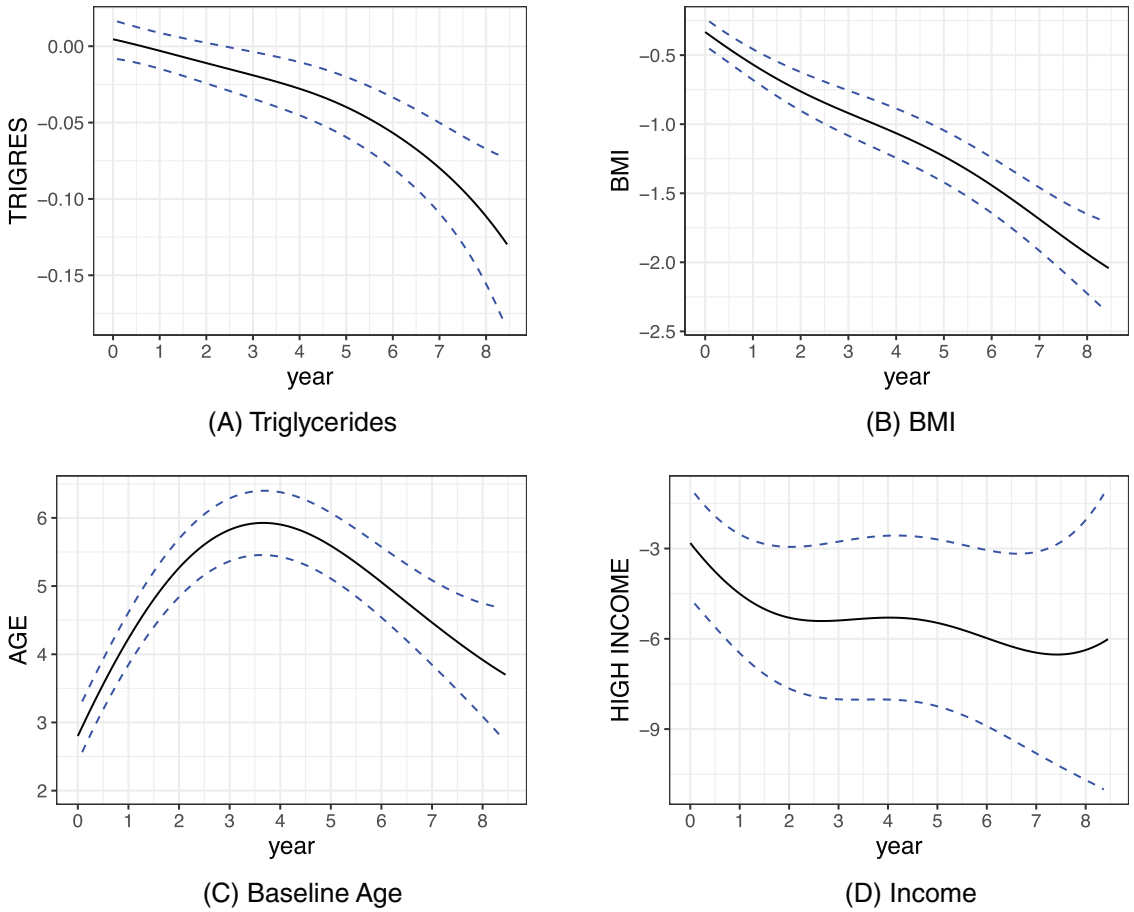


FIGURE 3 SWAN data analysis using FCAR: time-varying coefficient model on FSH against TG, BMI, age, and income. In each panel, the solid curve is the estimated coefficient function and the dashed curves are 95% pointwise confidence intervals obtained using bootstrap. This figure appears in color in the electronic version of this article, and any mention of color refers to that version.

To accommodate two time-varying covariates in our data, we slightly extend the proposed FPCA to a multivariate setting as described in Online Appendix B and implement it by using the `fdapace` package in R, where a built-in GCV procedure is used to select the bandwidths for mean and cross-covariance estimations. We then use the conditional AIC described in (11) to select the number of principal components for TG and BMI separately. To implement the undersmoothing scheme described in condition 4.3, we multiply the GCV-selected bandwidths by a factor of $n^{-1/10}$, refit FPCA using the undersmoothing bandwidths, and use the FPCA-calibrated values for the subsequent analyses.

We regress FSH against the calibrated TG and BMI values and adjust for time-invariant covariates age and income, using a multivariate time-varying coefficient model

$$Y_i(t) = \beta_0(t) + \boldsymbol{\beta}_z^{\mathrm{T}}(t)\mathbf{Z}_i + \boldsymbol{\beta}_x^{\mathrm{T}}(t)\mathbf{X}_i(t) + \epsilon_i(t).$$

Figure 3 shows the estimated coefficient functions for TG, BMI, age, and income using FCAR, respectively, where the 95% pointwise confidence intervals are obtained using bootstrap. As commented in Remarks 2 and 3, we resample the subjects and perform FPCA using the same bandwidth as in real data to every bootstrap sample in order to properly take into account the FPCA estimation errors. The pointwise confidence intervals in Figure 3 are based on a normal approximation suggested by Theorem 2, where the pointwise standard error is estimated based on 200 bootstrap replicates.

The estimated time-varying coefficients reveal that FSH is negatively associated with TG, which is consistent with the SWAN data analysis conducted in El Khoudary et al. (2016) using time-invariant regression models. Wang et al. (2020) studied the association between FSH and TG among postmenopausal and perimenopausal women separately using independent studies. The comparison between their results suggests a stronger negative association between FSH and TG among postmenopausal women than perimenopausal women, which supports our findings in Figure 3(A) that the negative association between FSH and

TG becomes stronger through the menopausal transition. Similarly, Figure 3(B) suggests that FSH is negatively associated with BMI. This is consistent with previous findings in the SWAN literature (Randolph Jr et al., 2004), which suggested that the negative association between FSH and BMI becomes stronger throughout the menopausal transition. This time-varying effect of BMI on FSH is not only confirmed by our study, but can be visualized as a continuous curve in Figure 3(B).

Time-invariant variables, such as age and income, are confounders, whose effects need to be adjusted for in the model. The effect of baseline age represents a cohort effect, reflecting different baseline FSH levels in different age cohorts. The positive association between age and FSH seen in Figure 3(C) is consistent with the literature that FSH is elevated as women age through the menopause transition (Burger et al., 1999). Also, we find that lower income women are more likely to present higher FSH, agreeing to the literature that links low socioeconomic status to high FSH (Wise et al., 2002), possibly because of poor health awareness (Burger et al., 1995), risk behaviors (Haddad et al., 2008), and inadequate access to health care (Barut et al., 2016).

7 | CONCLUSIONS

We have proposed a new functional calibration method, termed FCAR, for learning sparse asynchronous longitudinal data. The key idea behind the approach is to calibrate the missing synchronized covariates by the FPCA approach, which can be easily implemented using existing software. More broadly, our method is applicable to asynchronous longitudinal regression with time-invariant or time-varying coefficients, and addresses a serious limitation of the existing literature. Indeed, our FCAR estimator in a time-invariant regression model enjoys nice asymptotic properties, such as root-n consistency and asymptotic normality. By implementing an undersmoothing scheme in our functional calibration method, the FPCA estimation errors cause a negligible bias to the estimated model, but will inflate the asymptotic variance of the final estimator. Based on these theoretical findings, we recommend to use bootstrap standard error that takes into account FPCA errors, rather than using naive standard errors. Our theoretical analysis as well as our empirical studies show that our proposed method outperforms the existing methods, including the KW estimator of Cao et al. (2015) and the FVCM method of Şentürk and Müller (2010).

As a reviewer pointed out, the methods of Cao et al. (2015) and Şentürk et al. (2013) can handle generalized outcomes, whereas our investigation has been confined to Gaussian-type responses. This extension requires substantial theoretical work and we defer it to future work. As demonstrated in Online Appendix B and our real data analysis, the proposed FCAR method can be easily implemented for multiple asynchronous time-varying covariates, when the number of time-varying covariates is not too high. When the number is high, the computational load of multivariate FPCA (mFPCA) can rapidly escalate, leading to an unmanageable number of cross-covariance functions to estimate and causing the mFPCA estimators to become unreliable. In these situations, choosing the pertinent time-varying covariates presents a formidable challenge of "model selection with error-in-variable," as the calibrated covariate values are subject to estimation errors, which are not independent and possess complex structures. These difficulties deserve further exploration.

ACKNOWLEDGMENTS

We thank the Editor, the anonymous AE, and referee for their insightful comments and suggestions that have improved substantially the quality of the manuscript. The work is partially supported by grants from the National Institutes of Health.

DATA AVAILABILITY STATEMENT

The data that support the findings in this paper are openly available in Study of Women's Health Across the Nation at https://www.swanstudy.org.

ORCID

Yehua Li https://orcid.org/0000-0002-5945-378X
Yi Li https://orcid.org/0000-0003-1720-2760

REFERENCES

Barut, M.U., Agacayak, E., Bozkurt, M., Aksu, T. & Gul, T. (2016) There is a positive correlation between socioeconomic status and ovarian reserve in women of reproductive age. *Medical Science Monitor*, 22, 4386–4392.

Bromberger, J.T., Schott, L.L., Kravitz, H.M., Sowers, M., Avis, N.E., Gold, E.B. et al. (2010) Longitudinal change in reproductive hormones and depressive symptoms across the menopausal transition: results from the Study of Women's Health Across the Nation (Swan). *Archives of General Psychiatry*, 67(6), 598–607.

Burger, H.G., Dudley, E.C., Hopper, J.L., Groome, N., Guthrie, J.R., Green, A. et al. (1999) Prospectively measured levels of serum follicle-stimulating hormone, estradiol, and the dimeric inhibins during the menopausal transition in a population-based cohort of women. *The Journal of Clinical Endocrinology & Metabolism*, 84(11), 4025–4030.

Burger, H.G., Dudley, E.C., Hopper, J.L., Shelley, J.M., Green, A., Smith, A. et al. (1995) The endocrinology of the menopausal transition: a cross-sectional study of a population-based sample. *The Journal of Clinical Endocrinology & Metabolism*, 80(12), 3537–3545.

Cai, Z., Fan, J. & Li, R. (2000) Efficient estimation and inferences for varying-coefficient models. *Journal of the American Statistical Association*, 95(451), 888–902.

Cao, H., Zeng, D. & Fine, J.P. (2015) Regression analysis of sparse asynchronous longitudinal data. *Journal of the Royal Statistical Society, Series B*, 77, 755–776.

Carroll, R.J., Ruppert, D., Stefanski, L.A. & Crainiceanu, C.M. (2006) *Measurement error in nonlinear models: a modern perspective*. Boca Raton, FL: Chapman and Hall/CRC.

Cook, R.J., Zeng, L. & Yi, G.Y. (2004) Marginal analysis of incomplete longitudinal binary data: a cautionary note on locf imputation. *Biometrics*, 60, 820–838.

Dai, X., Lin, Z. & Müller, H.G. (2021) Modeling sparse longitudinal data on Riemannian manifolds. *Biometrics*, 77(4), 1328–1341.

El Khoudary, S.R., Santoro, N., Chen, H.Y., Tepper, P.G., Brooks, M.M., Thurston, R.C. et al. (2016) Trajectories of estradiol and follicle-stimulating hormone over the menopause transition and early markers of atherosclerosis after menopause. *European Journal of Preventive Cardiology*, 23(7), 694–703.

Haddad, R., Crum, C., Chen, Z., Krane, J., Posner, M., Li, Y. et al. (2008) Hpv16 transmission between a couple with hpv-related head and neck cancer. *Oral Oncology*, 44(8), 812–815.

Happ, C. & Greven, S. (2018) Multivariate functional principal component analysis for data observed on different (dimensional) domains. *Journal of the American Statistical Association*, 113(522), 649–659.

Hoover, D.R., Rice, J.A., Wu, C.O. & Yang, L.P. (1998) Nonparametric smoothing estimates of time-varying coefficient models with longitudinal data. *Biometrika*, 85(4), 809–822.

Li, Y. & Hsing, T. (2010) Uniform convergence rates for nonparametric regression and principal component analysis in functional/longitudinal data. *Annals of Statistics*, 38, 3321–3351.

Li, Y., Wang, N. & Carroll, R.J. (2013) Selecting the number of principal components in functional data. *Journal of the American Statistical Association*, 108(504), 1284–1294.

Liao, X., Zucker, D.M., Li, Y. & Spiegelman, D. (2011) Survival analysis with error-prone time-varying covariates: a risk set calibration approach. *Biometrics*, 67(1), 50–58.

Ramsay, J.O. & Silverman, B.W. (2005) *Functional data analysis* (2nd ed.). New York: Springer-Verlag.

Randolph Jr, J.F., Sowers, M., Bondarenko, I.V., Harlow, S.D., Luborsky, J.L. & Little, R.J. (2004) Change in estradiol and follicle-stimulating hormone across the early menopausal transition: effects of ethnicity and age. *The Journal of Clinical Endocrinology & Metabolism*, 89(4), 1555–1561.

Rice, J.A. & Wu, C.O. (2001) Nonparametric mixed effects models for unequally sampled noisy curves. *Biometrics*, 57(1), 253–259.

Richardson, S.J., Senikas, V. & Nelson, J.F. (1987) Follicular depletion during the menopausal transition: evidence for accelerated loss and ultimate exhaustion. *The Journal of Clinical Endocrinology & Metabolism*, 65(6), 1231–1237.

Robins, J.M., Rotnitzky, A. & Zhao, L.P. (1995) Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of American Statistical Association*, 90, 106–121.

Şentürk, D., Dalrymple, L.S., Mohammed, S.M., Kaysen, G.A. & Nguyen, D.V. (2013) Modeling time-varying effects with generalized and unsynchronized longitudinal data. *Statistics in Medicine*, 32(17), 2971–2987.

Şentürk, D. & Müller, H.G. (2010) Functional varying coefficient models for longitudinal data. *Journal of the American Statistical Association*, 105(491), 1256–1264.

SWAN (1996) Study of Women's Health across the Nation. https://www.swanstudy.org.

Wang, X., Zhang, H., Chen, Y., Du, Y., Jin, X. & Zhang, Z. (2020) Follicle stimulating hormone, its association with glucose and lipid metabolism during the menopausal transition. *Journal of Obstetrics and Gynaecology Research*, 46(8), 1419–1424.

Wise, L., Krieger, N., Zierler, S. & Harlow, B. (2002) Lifetime socioeconomic position in relation to onset of perimenopause. *Journal of Epidemiology & Community Health*, 56(11), 851–860.

Yao, F., Müller, H.G. & Wang, J.L. (2005) Functional data analysis for sparse longitudinal data. *Journal of the American Statistical Association*, 100(470), 577–590.

SUPPORTING INFORMATION

Web Appendices, Tables, and Figures, and data/code referenced in Sections 1, 2, 4, 5 and 6 are available with this paper at the Biometrics website on Wiley Online Library

How to cite this article: Chang, X., Li, Y. & Li, Y. (2023) Asynchronous and error-prone longitudinal data analysis via functional calibration. *Biometrics*, 79, 3374–3387. https://doi.org/10.1111/biom.13866