Semiparametric Estimation of Fixed Effects Panel Data Models with Smooth Coefficients

Yiguo Sun
Department of Economics
University of Guelph
Guelph, ON, Canada N1G2W1
yisun@uoguelph.ca

Raymond J. Carroll
Department of Statistics, Texas A&M University
College Station, TX 77843-3134,
carroll@stat.tamu.edu

Qi Li
Department of Economics, Texas A&M University
College Station, TX 77843-4228,
qi@econmail.tamu.edu

Abstract

In this paper we consider the problem of estimating semiparametric fixed-effects panel data models with smooth coefficients by local linear regression approach. We show that the proposed estimator has the usual nonparametric convergence rate and is asymptotically normally distributed under regular conditions. A modified least-squared cross-validatory method is used to find the optimal bandwidth automatically. Moreover, we propose a test statistic for testing the null hypothesis of random effects against fixed effects for semiparametric panel data regression models with smooth coefficients. Monte Carlo simulations are used to study finite sample performance of the proposed estimator and test.

Some key words: Panel data; Local least squares; Smooth coefficients; Bootstrap.

Short title: Fixed Effects Panel Data Models with Smooth Coefficients

1 Introduction

Panel data traces information on each individual unit across time. Economists often find that they could overcome econometrics difficulties and extract more economic inferences using panel data, which would not be possible using pure time-series data or cross-section data. With the increased availability of panel data both theoretical and applied econometrics work in panel data analysis has become ever popular in the recent years.

To avoid imposing inadequate parametric panel data model, many econometricians and statisticians have been working on theories of nonparametric and semiparametric panel data regression models. Among many contributions, we name a few here: Ke and Wang (2001), Li and Stengos (1996), and Ullah and Roy (1998) for semiparametric panel data models with random effects; Henderson and Ullah (2005), Lin and Carroll (2000, 2001, 2006), Lin, Wang, Welsh and Carroll (2004), Lin and Ying (2001), Ruckstuhl, Welsh and Carroll (1999), Wu and Zhang (2002) and Wang (2003) for estimation of nonparametric panel data models with random effects. Estimation of these types of models are appropriate when the individual effect is independent of the regressors.

However, random effects estimators are inconsistent if the true model is one with fixed effects, i.e. individual effects which are correlated with the regressors. Indeed, economists often view the assumptions for the random effects model as being unsupported by the data. Su and Ullah (2006) consider a fixed effects partially linear panel data model. This paper considers fixed effects panel data models with smooth coefficients. The model is assumed to apply to situations where we see smooth transitive changes across sections and/or across time.

The remainder of the paper is organized as follows: Section 2 introduces a semiparametric fixed-effects panel data regression model with smooth coefficients and estimation methodology. In Section 3, we propose a nonparametric estimator for unknown smoothing coefficients and derive its asymptotic results. In Section 4 we propose a test statistic for testing for random effects versus fixed effects in semiparametric panel data models with smooth coefficients. Section 5 examines finite sample properties with a small Monte Carlo study. Finally, Section 6 concludes the paper. We delay detailed mathematical proofs to Appendix A for fixed-effects estimation and Appendix B for random-effects estimation.

2 Fixed Effects Semiparametric Panel Data Models with Smooth Coefficients

We consider the following semiparametric fixed-effects panel data regression model with smooth coefficients

$$Y_{i,t} = X_{i,t}^T \theta(Z_{i,t}) + \mu_i + \nu_{i,t}, (i = 1, ..., n;\ t = 1, ..., m) \tag{1}$$

where the covariate $Z_{i,t} = (Z_{i,t,1}, ..., Z_{i,t,q})^T$ is of dimension q, $X_{i,t} = (X_{i,t,1}, ..., X_{i,t,k})^T$ is of dimension k, $\theta(\cdot) = (\theta_1(\cdot), \cdots, \theta_k(\cdot))^T$ contains k unknown functions, and all other variables are scalars. The random errors $\nu_{i,t}$ are assumed to be i.i.d. with a zero mean, finite variance σ_v^2 and independent of $Z_{i,t}$ and $X_{i,t}$ for all i and t. Further, μ_i has a finite mean and variance. We allow μ_i to be correlated with $Z_{i,t}$ and/or $X_{i,t}$ with an unknown correlation structure. Hence, model (1) is a fixed-effects model. Alternatively, when μ_i is uncorrelated with $Z_{i,t}$ and

$X_{i,t}$ model (1) is a random-effects model.

For a given fixed effects model, there are many ways of removing the unknown fixed effects from the model.

Example 1 *The usual first-differencing (FD) estimation method could deduct one equation from another to remove the time-invariant fixed effects. For example, deducting equation for time t from that for time $t-1$, we have for $t = 2, \cdots, m$*

$$\tilde{y}_{i,t} = y_{i,t} - y_{i,t-1} = X_{i,t}^T\theta(Z_{i,t}) - X_{i,t-1}^T\theta(Z_{i,t-1}) + \tilde{v}_{i,t}, \tilde{v}_{i,t} = v_{i,t} - v_{i,t-1}; \quad (2)$$

or deducting equation for time t from that for time 1, we obtain for $t = 2, \cdots, m$

$$\tilde{y}_{i,t} = y_{i,t} - y_{i1} = X_{i,t}^T\theta(Z_{i,t}) - X_{i,1}^T\theta(Z_{i1}) + \tilde{v}_{i,t}, \tilde{v}_{i,t} = v_{i,t} - v_{i,1}; \quad (3)$$

and so on.

Example 2 *The conventional fixed effects (FE) estimation method, on the other hand, removes the fixed effects by deducting each equation from the cross-time average of the system, and it gives for $t = 2, \cdots, m$*

$$\tilde{y}_{i,t} = X_{i,t}^T\theta(Z_{i,t}) - \frac{1}{m}\sum_{s=1}^{m} X_{i,s}^T\theta(Z_{i,s}) + \tilde{v}_{i,t} = \sum_{s=1}^{m} q_{t,s}X_{i,s}^T\theta(Z_{i,s}) + \tilde{v}_{i,t} \quad (4)$$

where $\bar{w}_{i.} = m^{-1}\sum_{t=1}^{m} w_{i,t}$ and $\tilde{w}_{i,t} = w_{i,t} - \bar{w}_{i.}$ with w being y or v. In addition, $q_{t,s} = -1/m$ if $s \neq t$ and $1 - 1/m$ otherwise, and $\sum_{s=1}^{m} q_{t,s} = 0$ for all t.

In general cases, we could introduce a constant $(m-1)\times m$ matrix P with full rank $m-1$ such that $Pe_m = 0_{m-1}$ where e_m is an $m \times 1$ vector of ones. Premultiplying P to the m equations for each i will remove the unknown fixed effects μ_i. For the FD estimator, each row of P only takes values of -1, 0, and 1. Denote $P = \left(p_1', \cdots, p_{m-1}'\right)'$, where $p_t = (p_{t,j})_{j=1}^m$ is an $m \times 1$ row vector. Model (2) assumes $p_{t,t} = -1, p_{t,t+1} = 1$, and $p_{t,j} = 0$ for $j \neq t, t+1$, $t = 1, \cdots, m-1$. Model (3) assumes $p_{t,1} = -1, p_{t,t+1} = 1$, and $p_{t,j} = 0$ for $j \neq 1, t+1$, $t = 1, \cdots, m-1$. For the FE estimator, $P = S_{m-1}\left(I_m - \frac{1}{m} e_m e_m'\right)$, where S_{m-1} is an $(m-1) \times m$ matrix containing the last $m-1$ rows of the $m \times m$ identity matrix I_m. When $m = 2$, FD estimator and FE estimator will be exactly the same; this is not true for $m > 2$.

Many nonparametric local smoothing methods can be used to estimate the unknown function $\theta(\cdot)$. However, for each i, the right-hand of equations (2)-(4) contain linear combination of $\left\{X_{i,t}^T \theta(Z_{i,t})\right\}_{t=1}^m$ over time. If X contains an intercept term, backfitting algorithm has to be used to recover the unknown functions,which brings not only computational burden but also complicate mathematical proofs.

Therefore, this paper proposes a new way of removing the unknown fixed effects, motived by least square dummy variable (LSDV) model in parametric panel data analysis. As the text goes along, we will describe how the proposed method removes fixed effects by deducting a smoothed version of average across time from each observation.

Rewriting model (1) in matrix format yields

$$Y = B\left(X, \theta\left(Z\right)\right) + D_0 \mu_0 + V, \tag{5}$$

where $Y = \left[Y_1^T, \cdots, Y_n^T\right]^T$ and $V = \left[v_1^T, \cdots, v_n^T\right]^T$ are $(nm) \times 1$ vectors, $B\left(X, \theta\left(Z\right)\right)$ stacks all $X_{i,t}^T \theta\left(Z_{i,t}\right)$ into an $(nm) \times 1$ vector in ascending order of

i first, then of t, $\mu_0 = [\mu_1, \cdots, \mu_n]^T$ is an $n \times 1$ vector, and $D_0 = I_{n\times n} \otimes e_m$ is a $(nm) \times n$ matrix with main diagonal blocks being e_m, where $\otimes$ refers to Kroneker product operation. However, we can not work on this model without further restriction on the fixed effects for cases like $X_{i,t}^T \theta (Z_{i,t}) = \theta_1 (Z_{i,t}) + X_{i,t,2}\theta_2 (Z_{i,t}) + \cdots$, since we can not identify $\theta_1 (\cdot)$ in the presents of unknown fixed effects. Therefore, we impose an identification condition like $\sum_{i=1}^n \mu_i = 0$, which has been used by Su and Ullah (2006). Define $\mu = [\mu_2, \cdots, \mu_n]^T$. With restriction $\sum_{i=1}^n \mu_i = 0$, we can rewrite (5) as

$$Y = B(X, \theta(Z)) + D\mu + V, \tag{6}$$

where $D = \left[-e_{n-1}\ I_{(n-1)\times(n-1)}\right]^T \otimes e_m$ is a $(nm) \times (n-1)$ matrix.

Define an $m\times m$ diagonal matrix $K_H (Z_i, z) = diag (K_H (Z_{i,1}, z), \cdots, K_H (Z_{i,m}, z))$ for $i = 1, 2, \cdots, n$, and a $(nm)\times(nm)$ diagonal matrix $W_H (z) = diag(K_H (Z_1, z), \cdots, K_H (Z_n, z))$, where $K_H (Z_{i,t}, z) = K \left(H^{-1} (Z_{i,t} - z)\right)$ for all i and t, and $H = diag (h_1, \cdots, h_q)$ is a $q \times q$ diagonal smoothing matrix. We then solve the following optimization problem

$$\min_{\theta(Z),\mu} (Y - B(X, \theta(Z)) - D\mu)^T W_H (z) (Y - B(X, \theta(Z)) - D\mu). \tag{7}$$

We use the local weight matrix $W_H (z)$ to take into account localness of our nonparametric fitting, and place no weight matrix for data variation since $\{v_{it}\}$ are i.i.d. across equations. Taking first-order condition with respect to μ gives

$$D^T W_H (z) (Y - B(X, \theta(Z)) - D\hat{\mu}(z)) = 0, \tag{8}$$

which yields

$$\hat{\mu}(z) = \left(D^T W_H (z) D\right)^{-1} D^T W_H (z) (Y - B(X, \theta(Z))). \tag{9}$$

Replacing μ in (7) by $\hat{\mu}(z)$, we obtain the concentrated weighted least squares

$$\min_{\theta(Z)} \left(Y - B\left(X, \theta\left(Z\right)\right)\right)^T S_H\left(z\right) \left(Y - B\left(X, \theta\left(Z\right)\right)\right), \tag{10}$$

where we define $S_H(z) = M_H(z)^T W_H(z) M_H(z)$ and $M_H(z) = I_{(nm)} - D\left(D^T W_H(z) D\right)^{-1} D^T W_H(z)$. Apparently, $M_H(z) D\mu \equiv 0_{(nm)\times 1}$ for all z effectively removes unknown fixed effects from model (1).

How does $M_H(z)$ transform the observed data? Simple calculations give

$$M_H(z) = I_{(nm)\times(nm)} - D\left(A^{-1} - A^{-1} e_{n-1} e_{n-1}^T A^{-1} / \sum_{i=1}^{n} c_H(Z_i, z)\right) D^T W_H(z),$$

where $c_H(Z_i, z)^{-1} = \sum_{t=1}^{m} K_H(Z_{i,t}, z)$ for $i = 1, \cdots, n$ and $A = diag\left(c_H(Z_2, z)^{-1}, \cdots, c_H(Z_n, z)^{-1}\right)$. We use the formula $(A + BCD)^{-1} = A^{-1} - A^{-1}B\left(DA^{-1}B + C^{-1}\right)^{-1} DA^{-1}$ to derive the inverse matrix, see Appendix B in Poirier (1995).

3 Nonparametric Estimator and Asymptotic Theory

Local linear regression approach is usually used to estimate non-/semi-parametric models. The basic idea of this method is to apply Taylor expansion up to the second-order derivative. That is, for each $l = 1, \cdots, k$, we have the following Taylor expansion around a point z :

$$\theta_l(z_{i,t}) \approx \theta_l(z) + \left[H\theta_l'(z)\right]^T \left[H^{-1}(z_{i,t} - z)\right] + \frac{1}{2} r_{H,l}(z_{i,t}, z), \left\|H^{-1}(z_{i,t} - z)\right\| = O(1) \text{ a.s.}, \tag{11}$$

where $r_{H,l}(z_{i,t}, z) = \left[H^{-1}(z_{i,t} - z)\right]^T \left[H \frac{\partial^2 \theta_l(z)}{\partial z \partial z^T} H\right] \left[H^{-1}(z_{i,t} - z)\right]$. $\theta_l(z)$ approximates $\theta_l(z_{i,t})$ and $\theta_l'(z)$ approximates $\theta_l'(z_{i,t})$ when $z_{i,t}$ is close to z. Define

$\beta_l(z) = \left[\theta_l(z), \left[H\theta_l'(z)\right]^T\right]^T$, a $(q+1)\times 1$ column vector for $l = 1, 2, \cdots, k$, and $\beta(z) = [\beta_1(z), \cdots, \beta_k(z)]^T$, a $k \times (q+1)$ parameter matrix. The first column of $\beta(z)$ is $\theta(z)$. Therefore, we will replace $\theta(Z_{i,t})$ in (1) by $\beta(z) G_{i,t}(z, H)$ for each i and t, where $G_{i,t}(z, H) = [1, \{H^{-1}(Z_{i,t} - z)\}^T]^T$ is a $(q+1)\times 1$ vector.

To make matrix operation simpler, we stack the parameter matrix $\beta(z)$ into a $[k(q+1)] \times 1$ column vector and denote it by $vec(\beta(z))$. Since $vec(ABC) = (C^T \otimes A)\, vec(B)$ and $(A \otimes B)^T = A^T \otimes B^T$, where $\otimes$ refers to Kronecker product, we have $X_{i,t}^T \beta(z) G_{i,t}(z, H) = (G_{i,t}(z, H) \otimes X_{i,t})^T vec(\beta(z))$ for all i and t. Thus, we consider the following minimization problem

$$\min_{\beta(z)} [Y - R(z, H)\, vec(\beta(z))]^T S_H(z) [Y - R(z, H)\, vec(\beta(z))] \tag{12}$$

where

$$R_i(z, H) = \begin{bmatrix} (G_{i,1}(z, H) \otimes X_{i,1})^T \\ \vdots \\ (G_{i,m}(z, H) \otimes X_{i,m})^T \end{bmatrix} \text{ is an } m \times [k(q+1)] \text{ matrix,} \tag{13}$$

$$R(z, H) = \left[R_1(z, H)^T, \cdots, R_n(z, H)^T\right]^T \text{ is a } (nm) \times [k(q+1)] \text{ matrix.} \tag{14}$$

Simple calculations give

$$vec\left(\widehat{\beta}(z)\right) = \left[R(z, H)^T S_H(z) R(z, H)\right]^{-1} R(z, H)^T S_H(z) Y \tag{15}$$

$$= vec(\beta(z)) + \left[R(z, H)^T S_H(z) R(z, H)\right]^{-1} (A_n/2 + B_n), \tag{16}$$

where $A_n = R(z, H)^T S_H(z) \Pi(z, H)$ and $B_n = R(z, H)^T S_H(z) V$. The $(t + (i-1)m)^{th}$ element of the column vector $\Pi(z, H)$ is $X_{i,t}^T r_H\left(\tilde{Z}_{i,t}, z\right)$, where $r_H(\cdot, \cdot) = [r_{H,1}(\cdot, \cdot), \cdots, r_{H,k}(\cdot, \cdot)]^T$ and $r_{H,l}\left(\tilde{Z}_{i,t}, z\right) = \left[H^{-1}(Z_{i,t} - z)\right]^T \left[H \frac{\partial^2 \theta_l(\tilde{Z}_{i,t})}{\partial z \partial z^T} H\right] \left[H^{-1}(Z_{i,t} - z)\right]$ with $\tilde{Z}_{i,t}$ lying between $Z_{i,t}$ and z for each i and t.

To derive the asymptotic distribution of $vec\left(\widehat{\beta}(z)\right)$, we first give some regularity conditions. We use $M > 0$ to stand for a finite constant.

Assumption 1: The random variables $(Y_{i,t}, X_{i,t}, Z_{i,t})$ are independent and identically distributed (iid) across the i index, and $E\left(\left\|X_{1,t}X_{1,s}^T X_{1,t}X_{1,s}^T\right\|\right) \leq M < \infty$ for all t and s. $Z_{i,t}$ are continuos random variables with a pdf $f_t(\cdot)$, conditional pdfs $f_t(\cdot|x_{i,t})$ and $f_t\left(z|x_{i,t}, x_{i,s}\right)$. Denote $f_{t,s}\left(z, z|x_{i,t}, x_{i,s}\right)$ to be the joint pdf of $(Z_{i,t}, Z_{i,s})$ conditional on $(X_{i,t}, X_{i,s}) = (x_{i,t}, x_{i,s})$. Also, all these (marginal or conditional) pdfs and $\theta(\cdot)$ are twice continuously differentiable functions and their second-order derivatives satisfies Lipschitz conditions. Let $\mathcal{S}_t$ denote the support of $Z_{i,t}$; then $f_t(z)$ is bounded away from zero in its domain $\mathcal{S}_t$.

Assumption 2: X has full rank k, and $X_{i,t,l} \neq X_{i,t,l'}^T Z_{i,t,j}$ for any $l \neq l'(l, l' = 1, \cdots, k)$, and $j = 1, \cdots, q$. In addition, $n-1$ out of n variables $\{\mu_i\}_{i=1}^n$ are independently distributed with zero mean and variance σ_μ^2 and the other one is determined by $\sum_{i=1}^n \mu_i = 0$. If $X_{i,t,l} \equiv X_{i,l}$ for some l, we assume $\sum_{i=1}^n X_{i,l} \neq 0$.

This is one identification condition. For cases that $X_{1,t} \equiv 1$, then $\theta_1(z)$ is estimable if $M_H(z)(e_n \otimes e_m) \neq 0$ for a given z. Simple calculations give $M_H(z)(e_n \otimes e_m) = \left[n^{-1}\sum_{i=1}^n c_H(Z_i, z)\right]^{-1} \left(c_H(Z_1, z), \cdots, c_H(Z_n, z)\right)^T \otimes e_m$; the proof of Lemma 6 in Appendix A can be used to show that $M_H(z)(e_n \otimes e_m) \neq 0$ for any given z with probability one.

In addition, for cases that $X_{i,t,l} \equiv X_{i,l}$ and $X_{i,l}$ is not constant for some l, i.e. the l^{th} variable in X is time invariant, $\theta_l(z)$ is estimable if $M_H(z)(b \otimes e_m) \neq 0$

for a given z, where b is a $n \times 1$ vector with $b_i = X_{i,l}$. Simple calculations give $M_H(z)(b \otimes e_m) = \frac{b_1 + \cdots + b_n}{n} M_H(z)(e_n \otimes e_m)$. Therefore, $\theta_l(z)$ is identifiable if $\sum_{i=1}^{n} X_{i,l} \neq 0$.

There could be other ways of introducing identification conditions. We use Su and Ullah's (2006) assumption, $\sum_{i=1}^{n} \mu_i = 0$, this restriction simplifies estimation and proofs and there is no need to do backfitting iteration.

Assumption 3: $K(v) = \prod_{s=1}^{q} k(v_s)$ is a product kernel, the univariate kernel function $k(\cdot)$ is a uniformly bounded, symmetric (around zero) probability density function and its first four moments are finite. In addition, define $|H| = h_1 \cdots h_q$ and $\|H\| = \sqrt{\sum_{j=1}^{q} h_j^2}$. As $n \to \infty$, $\|H\| \to 0$ and $n|H| \to \infty$.

$\theta(z)$ is to be estimated by $\widehat{\theta}(z)$ which is the first column of $\widehat{\beta}(z)$.

Theorem 3 *Under Assumptions 1-3, we have the following bias and variance for* $\widehat{\theta}(z)$*:*

$$\begin{aligned} Bais\left(\widehat{\theta}(z)\right) &= \frac{1}{2}\left(\int K(u)\, uu^T du\right)\Theta_H(z) + o\left(\|H\|^2\right), \\ Var\left(\widehat{\theta}(z)\right) &= 2\sigma_v^2 n^{-1}|H|^{-1}\left(\int K^2(u)\, du\right)\Phi^{-1} + o\left(n^{-1}|H|^{-1}\right), \end{aligned}$$

where $\Phi = \sum_{t=1}^{m}\left(1 - \sqrt{\frac{f_t(z)}{f(z)}}\right) E\left[f_t(z|X_{1,t}) X_{1,t} X_{1,t}^T\right]$, $f(z) = \sum_{t=1}^{m} f_t(z)$, *and* $\Theta_H(z) = \left[tr\left(H \frac{\partial^2 \theta_1(z)}{\partial z \partial z^T} H\right), \cdots, tr\left(H \frac{\partial^2 \theta_k(z)}{\partial z \partial z^T} H\right)\right]^T$.

Moreover, we have derived the following asymptotic normality results for $\widehat{\theta}(z)$.

Theorem 4 *Under Assumptions 1-3,* $E\|X_{i,t}\|^{2+\delta} < \infty$, $E|v_{i,t}|^{2+\delta} < \infty$ *for some* $\delta > 0$, *and* $\sqrt{n|H|}\,\|H\|^2 = o(1)$ *as* $n \to \infty$, *we have* $\sqrt{n|H|}\left(\widehat{\theta}(z) - \theta(z)\right) \xrightarrow{d}$

$N\left(0, \Sigma_{\theta(z)}\right)$, *where* $\Sigma_{\theta(z)} = 2\sigma_v^2 \left(\int K^2(u)\, du\right) \Phi^{-1}$. *Moreover, a consistent estimator for* $\Sigma_{\theta(z)}$ *is given as follows:*

$$\begin{aligned}
\hat{\Sigma}_{\theta(z)} &= S_k \hat{\Omega}(z,H)^{-1} \hat{J}(z,H)\, \hat{\Omega}(z,H)^{-1} S_k^T \xrightarrow{p} \Sigma_{\theta(z)}, \\
\hat{\Omega}(z,H) &= n^{-1} |H|^{-1} R(z,H)^T S_H(z)\, R(z,H) \\
\hat{J}(z,H) &= n^{-1} |H|^{-1} R(z,H)^T S_H(z)\, \hat{V}\hat{V}^T S_H(z)\, R(z,H)
\end{aligned}$$

where $\hat{V}$ *is the vector of estimated residuals and* S_k *includes the first* k *rows of the* $[k(q+1)] \times [k(q+1)]$ *identify matrix.*

When calculating $\hat{J}(z,H)$, we do not need to estimate μ, since $S_H(z)\, \hat{V}\hat{V}^T S_H(z)$ $= W_H(z) \left(M_H(z)\, \hat{V}\right) \left(M_H(z)\, \hat{V}\right)^T W_H(z)$ and $M_H(z)\, D = 0$; $M_H(z)\, \hat{V}$ are the estimated residuals from the local least squares estimation (12).

4 A Nonparametric Test: Random Effects vs Fixed Effects

In this section we discuss how to test for the presence of random effects versus fixed effects in the semiparametric panel data model with smoothing coefficients. The model remains as (1). Define $u_{i,t} = \mu_i + \nu_{i,t}$. The random effects specification assumes that μ_i is uncorrelated with the regressors X and Z, while for the fixed effects case, μ_i is allowed to be correlated with X and/or Z in an unknown way.

We are interested in testing the null hypothesis (H_0) that μ_i is a random effect versus the alternative hypothesis (H_1) that μ_i is a fixed effect. The null and alternative hypotheses can be written as

$$H_0 \quad : \quad \Pr\left[E(\mu_i|Z_{i,1},...,Z_{i,m},X_{i,1},\cdots,X_{i,m}) \equiv 0\right] = 1 \text{ for all } i, \qquad (17)$$

$$H_1 \quad : \quad \Pr\left[E(\mu_i|Z_{i,1},...,Z_{i,m},X_{i,1},\cdots,X_{i,m}) \neq 0\right] > 0 \text{ for some } i\ , \qquad (18)$$

while we keep the same setup given in model (1) under both H_0 and H_1.

Our statistic is based on the squared difference between the FE and RE estimators, which is asymptotically zero under H_0 and positive under H_1. To simplify the proofs and save computing time, we use local constant estimator instead of local linear estimator for the test. Then following the argument in Section 2 and Appendix B, we have

$$\begin{aligned}\hat{\theta}_{FE}(z) &= \left(X^T S_H(z) X\right)^{-1} X^T S_H(z) Y \\ \hat{\theta}_{RE}(z) &= \left(X^T W_H(z) X\right)^{-1} X^T W_H(z) Y\end{aligned}$$

where X is a $(nm)\times k$ matrix with $X = \left[X_1^T,\cdots,X_n^T\right]$, and for each i, $X_i = (X_{i,1},\cdots,X_{i,m})^T$ is an $m\times k$ matrix with $X_{i,t} = [X_{i,t,1},\cdots,X_{i,t,k}]^T$. Motivated by Li, Huang, Li, and Fu (2002), we remove the random denominator of $\hat{\theta}_{FE}(z)$ by multiplying $X^T S_H(z) X$ and the test statistic is defined as

$$\begin{aligned}T_n &= \int \left[\hat{\theta}_{FE}(z) - \hat{\theta}_{RE}(z)\right]^T \left(X^T S_H(z) X\right)^T \left(X^T S_H(z) X\right) \left[\hat{\theta}_{FE}(z) - \hat{\theta}_{RE}(z)\right] dz \\ &= \int \tilde{U}(z)^T S_H(z) X X^T S_H(z) \tilde{U}(z)\, dz\end{aligned}$$

since $\left(X^T S_H(z) X\right)\left[\hat{\theta}_{FE}(z) - \hat{\theta}_{RE}(z)\right] = X^T S_H(z)\left(Y - X\hat{\theta}_{RE}(z)\right) \triangleq X^T S_H(z)\tilde{U}(z)$. To simplify the statistic, we make several changes in T_n. Firstly, we simplify the integration calculation by replacing $\tilde{U}(z)$ by $\hat{U}$, where $\hat{U} = Y - B\left(X, \hat{\theta}_{RE}(Z)\right)$ and $B\left(X, \hat{\theta}_{RE}(Z)\right)$ stacks up $X_{i,t}^T\hat{\theta}_{RE}(Z_{i,t})$ in the increasing order of i first then of t. Secondly, to overcome the complexity caused by the random denominator in $M_H(z)$, we replace $M_H(z)$ by $M_D = I_{nm} - m^{-1} I_n \otimes \left(e_m e_m^T\right)$, and apparently, $M_D D = 0$. Now removing the term of $j = i$, we obtain

$$\tilde{T}_n = \sum_{i=1}^{n} \sum_{j \neq i} \hat{U}_i^T Q_m \int K_H(Z_i, z) X_i^T X_j K_H(Z_j, z)\, dz Q_m \hat{U}_j,$$

where $Q_m = I_m - m^{-1} e_m e_m^T$. If $|H| \to 0$ as $n \to \infty$ and $E\|X_{i,t}\|^4 < M < \infty$, we obtain

$$|H|^{-1} \int K_H(Z_i, z) X_i^T X_j K_H(Z_j, z)\, dz \approx \begin{bmatrix} \bar{K}_H(Z_{i,1}, Z_{j,1}) X_{i,1}^T X_{j,1} & \cdots & \bar{K}_H(Z_{i,1}, Z_{j,m}) X_{i,1}^T X_{j,m} \\ \vdots & \ddots & \vdots \\ \bar{K}_H(Z_{i,m}, Z_{j,1}) X_{i,m}^T X_{j,1} & \cdots & \bar{K}_H(Z_{i,m}, Z_{j,m}) X_{i,m}^T X_{j,m} \end{bmatrix}, \tag{19}$$

where $\bar{K}_H(Z_{i,t}, Z_{j,s}) = \int K\left(H^{-1}(Z_{i,t} - Z_{j,s}) + \omega\right) K(\omega)\, d\omega$. We then replace $\bar{K}_H(Z_{i,t}, Z_{j,s})$ by $K_H(Z_{i,t}, Z_{j,s})$; this replacement will not affect the essence of the test statistic since the local weight is untouched. Now, our proposed test statistic is given as

$$\hat{T}_n = \frac{1}{n^2 |H|} \sum_{i=1}^{n} \sum_{j \neq i} \hat{U}_i^T Q_m A_{i,j} Q_m \hat{U}_j \tag{20}$$

where $A_{i,j}$ equals to the right-hand side of equation (19) after replacing $\bar{K}_H(Z_{i,t}, Z_{j,s})$ by $K_H(Z_{i,t}, Z_{j,s})$. Finally, to remove the asymptotic bias term of the proposed test statistic, the random-effects estimator of $\theta(Z_{i,t})$ are leave-one-unit-out random effect estimators; for a given pair of (i,j), $i \neq j$, $\hat{\theta}_{RE}(Z_{i,t})$ is calculated without using observations on $\{(X_{j,t}, Z_{j,t}, Y_{j,t})\}_{t=1}^m$.

We present the asymptotic properties of this test below and delay the proofs to Appendix C.

Theorem 5 *Under Assumptions 1-3, and* $n\left|H\right|\left\|H\right\|^4 \to 0$ *as* $n \to \infty$*, we have under* H_0 *as* $n \to \infty$

$$J_n = n\sqrt{|H|}\hat{T}_n/\hat{\sigma}_0 \xrightarrow{d} N\left(0,1\right) \tag{21}$$

where

$$\hat{\sigma}_0^2 = \frac{1}{n^2\left|H\right|}\sum_{i=1}^{n}\sum_{j\neq i}\left(\hat{V}_i^T Q_m A_{i,j} Q_m \hat{V}_j\right)^2 \tag{22}$$

is a consistent estimator of

$$\sigma_0^2 = 2\left(1-\frac{1}{m}\right)^2 \sigma_v^4 \int K^2\left(u\right)du \sum_{t=1}^{m}\sum_{s=1}^{t} E\left[f_{2,t|1,s}\left(Z_{1,s}|Z_{1,s}, X_{1,s}, X_{2,t}\right)\left(X_{1,s}^T X_{2,t}\right)^2\right],$$

where $\hat{V}_{i,t} = Y_{i,t} - X_{i,t}^T\hat{\theta}_{FE}\left(Z_{i,t}\right) - \hat{\mu}_i$ *and* $\hat{\theta}_{FE}\left(Z_{i,t}\right)$ *is a leave-two-unit-out FE estimator without using the observations from the ith and jth units and* $\hat{\mu}_i = \bar{Y}_i - m^{-1}\sum_{t=1}^{m} X_{i,t}^T\hat{\theta}_{FE}\left(Z_{i,t}\right)$*. Under* H_1*,* $\Pr\left[J_n > B_n\right] \to 1$ *as* $n \to \infty$*, where* B_n *is any nonstochastic sequence with* $B_n = o\left(n\sqrt{|H|}\right)$*.*

Theorem 5 states that the test statistic $J_n = n\sqrt{|H|}\hat{T}_n/\hat{\sigma}_0$ is a consistent test for testing H_0 against H_1. It is a one-sided test. If J_n is greater than the critical values from the standard normal distribution, we will reject the null hypothesis at the corresponding significance levels.

5 Monte Carlo Simulations

In this section we report some Monte Carlo simulation results to examine the finite sample performance of the proposed estimator. The following data generating process is used:

$$Y_{i,t} = \theta_1(Z_{i,t}) + \theta_2(Z_{i,t})X_{i,t} + \mu_i + v_{i,t}, \tag{23}$$

where $\theta_1(z) = 1 + z + z^2$, $\theta_2(z) = sin(z\pi)$, $Z_{i,t} = w_{i,t} + w_{i,t-1}$, $w_{i,t}$ is i.i.d. uniformly distributed in $[0, \pi/2]$, $X_{i,t} = 0.5X_{i,t-1} + \xi_{i,t}$, $\xi_{i,t}$ is i.i.d. $N(0, 1)$. In addition, $\mu_i = c_0 Z_{i\cdot} + u_i$ for $i = 1, 2, \cdots, n-1$ and $\mu_n = -\sum_{i=1}^{n-1} \mu_i$ where $c_0 = 0$, 0.5,and 1.0, u_i is i.i.d. $N(0, 1)$. When $c_0 \neq 0$, μ_i and $Z_{i,t}$ are correlated; we use c_0 to control the correlation between μ_i and $Z_{i\cdot} = m^{-1}\sum_{t=1}^{m} Z_{i,t}$. Moreover, $v_{i,t}$ is i.i.d. $N(0, 1)$, $w_{i,t}$, $\xi_{i,t}$, u_i and $v_{i,t}$ are independent of each other.

We report estimation results for both the proposed fixed-effects (FE) estimator and the random-effects (RE) estimator (see Appendix B for the asymptotic results of the RE estimator). To learn how the two estimators perform when we have fixed-effects model and when we have random-effects model, we use the integrated squared error as a standard measure of estimation accuracy:

$$ISE\left(\hat{\theta}_l\right) = \int \left[\hat{\theta}_l\left(z\right) - \theta_l\left(z\right)\right]^2 f\left(z\right) dz, \tag{24}$$

which can be approximated by the average mean squared error $AMSE(\hat{\theta}_l) = nm)^{-1}\sum_{i=1}^{n}\sum_{t=1}^{m}[\hat{\theta}_{j,l}(Z_{i,t}) - \theta_{j,l}(Z_{i,t})]^2$ for $l = 1, 2$. Finally, in Table 1 we present the average value of $AMSE(\hat{\theta}_l)$ from 1000 Monte Carlo experiments. We take $m = 3$ and the sample size n=50, 100, and 200.

Since the bias and variance of the proposed FE estimator do not depend on the values of the fixed effects, our estimates are the same for different values of c_0; however, it is not true if random-effects model is true. Therefore, the results derived from the FE estimator are only reported once for each c_0 in Table 1.

It is well-known that the performance of non-/semi-parametric models depends on the choice of bandwidth. Therefore, we propose a leave-**one-unit**-out cross validation method to automatically find the optimal bandwidth for estimating both the FE and RE models. Specifically, when estimating $\theta(\cdot)$ at a point $Z_{i,t}$, we remove $\{(X_{i,t}, Y_{i,t}, Z_{i,t})\}_{t=1}^{m}$ from the data and only use the rest of $(n-1)\,m$ observations to calculate $\hat{\theta}_{(-i)}(Z_{i,t})$. In computing the RE estimate, the leave-one-unit-out cross validation method is just a trivial extension of the conventional leave-one-out cross validation method. However, such simple extension fails to provide satisfying result when we calculate a FE estimator due to the existence of unknown fixed effects. Therefore, based on the arguments made in Section 2, when calculating the FE estimator, we use the following modified leave-one-unit-out cross validation method:

$$\hat{H}_{opt} = \arg\min_{H} \left(Y - B\left(X, \hat{\theta}_{(-1)}(Z)\right)\right)^{T} M_D^T M_D \left(Y - B\left(X, \hat{\theta}_{(-1)}(Z)\right)\right), \tag{25}$$

where $M_D = I_{(nm)\times(nm)} - m^{-1} I_{n\times n} \otimes \left(e_m e_m^T\right)$ satisfies $M_D D = 0$ (this is used to remove the unknown fixed effects) and $B\left(X, \hat{\theta}_{(-1)}(Z)\right)$ stacks up $X_{i,t}^T \hat{\theta}_{(-i)}(Z_{i,t})$ in the increasing order of i first then of t. Simple calculations give

$$\begin{aligned}
& \left(Y - B\left(X, \hat{\theta}_{(-1)}(Z)\right)\right)^{T} M_D^T M_D \left(Y - B\left(X, \hat{\theta}_{(-1)}(Z)\right)\right) \\
= & \left(B(X, \theta(Z)) - B\left(X, \hat{\theta}_{(-1)}(Z)\right) + V\right)^{T} M_D^T M_D \left(B(X, \theta(Z)) - B\left(X, \hat{\theta}_{(-1)}(Z)\right) + V\right) \\
= & \left(B(X, \theta(Z)) - B\left(X, \hat{\theta}_{(-1)}(Z)\right)\right)^{T} M_D^T M_D \left(B(X, \theta(Z)) - B\left(X, \hat{\theta}_{(-1)}(Z)\right)\right) \\
& +2\left(B(X, \theta(Z)) - B\left(X, \hat{\theta}_{(-1)}(Z)\right)\right)^{T} M_D^T M_D V + V^T M_D M_D V,
\end{aligned} \tag{26}$$

where the last term does not depend on the bandwidth at all. If $v_{i,t}$ is independent of $\{(X_{j,s}, Z_{j,s}) : j = 1, \cdots, n; s = 1, \cdots m\}$ for all i and t, or $(X_{i,t}, Z_{i,t})$ is strictly exogenous variable, then the second term has zero expectation because the linear transformation matrix M_D removes a cross-time **not** cross-sectional average from each variable, e.g. $\tilde{Y}_{i,t} = Y_{i,t} - m^{-1}\sum_{t=1}^{m} Y_{i,t}$ for all i and t. Therefore, the first term is the dominate term in large sample and (25) is used to find an optimal smoothing matrix minimizing a weighted mean squared error of $\left\{\hat{\theta}\left(Z_{i,t}\right)\right\}_{i=1,n}^{t=1,m}$. Of course, we could use other weight matrix in (25) instead of M_D as long as the weight matrix can remove the fixed effects and does not trigger non-zero expectation of the second term in (26).

Table 1 shows that the RE estimator performs much better than the FE estimator when the true model is a random effects model, and the opposite is true when the true model is a fixed-effects model. This observation is consistent with what we have derived in a parametric panel data regression model analysis. Therefore, our simulation results indicate that a test for random effects against fixed effects will be always in demand when we analyze panel data models. In Table 2 we give Monte Carlo simulations of the proposed nonparametric test of random effects against fixed effects.

How to choose the bandwidth h for the test? For univariate case, Theorem 5 indicates that $nh \to \infty$ and $nh^{9/2} \to 0$ as $n \to \infty$; if we take $h \sim n^{-\alpha}$, Theorem 5 requires $\alpha \in \left(\frac{2}{9}, 1\right)$. If we balance the two conditions $nh \to \infty$ and $nh^{9/2} \to 0$ as $n \to \infty$, we have $\alpha = 2/7$. Therefore, in producing Table 2, we use $h = c\,(nm)^{-2/7}\,\hat{\sigma}_z$ to calculate the RE estimator with c being value of .8, 1.0, and 1.2. The results in Table 2 are consistent with the findings in the nonparametric tests literature in that a smaller bandwidth is good for size approximation and a larger bandwidth is good for power approximation.

6 Conclusion

In this paper we proposed using a local least squares method to estimate a semiparametric panel data model with fixed effects and smooth coefficients. In addition, we suggested using a bootstrap procedure to test for random effects against fixed effects. A data-driven method has been introduced to automatically find the optimal bandwidth for the proposed FE estimator. Monte Carlo simulations indicate that the proposed estimator and test statistic have good finite sample performance. Finally, the choice of bandwidths for the test is important when investigating the size and power of the test.

Acknowledgments

Sun's research was supported from the Social Sciences and Humanities Research Council of Canada (SSHRC). Carroll's research was supported by a grant from the National Cancer Institute (CA-57030), and by the Texas A&M Center for Environmental and Rural Health via a grant from the National Institute of Environmental Health Sciences (P30-ES09106). Li's research was partially supported by the Private Enterprise Research Center, Texas A&M University.

Appendix A: Technical Sketch–Proof of Theorem 3

To make our mathematical formula short, we introduce some simplified notations first: for each i, t, and s, $\lambda_{it} = K_H(Z_{it}, z)$, $c_H(Z_i, z)^{-1} = \sum_{t=1}^{m} \lambda_{it}$, and for any positive integers i, j, t, s

$$
\begin{aligned}
[\cdot]_{it,js} &= G_{it}(z,H)\,G_{js}^{T}(z,H) = \begin{bmatrix} 1 & G_{js1} & \cdots & G_{jsq} \\ G_{it1} & G_{it1}G_{js1} & \cdots & G_{it1}G_{jsq} \\ \vdots & \vdots & \ddots & \vdots \\ G_{itq} & G_{itq}G_{js1} & \cdots & G_{itq}G_{jsq} \end{bmatrix} \\
&= \begin{bmatrix} 1 & \left(H^{-1}(Z_{js}-z)\right)^{T} \\ H^{-1}(Z_{it}-z) & H^{-1}(Z_{it}-z)\left(H^{-1}(Z_{js}-z)\right)^{T} \end{bmatrix} \qquad (1)
\end{aligned}
$$

where the $(l+1)^{th}$ element of $G_{js}(z,H)$ is $G_{jsl} = (Z_{jsl} - z_l)/h_l$, $l = 1, \cdots, q$. Simple calculations show that

$$
\begin{aligned}
[\cdot]_{i_1t_1,i_2t_2}[\cdot]_{j_1s_1,j_2s_2} &= \left(1 + \sum_{j=1}^{q} G_{j_1s_1j}G_{i_2t_2j}\right)[\cdot]_{i_1t_1,j_2s_2}, \qquad (2) \\
R_i(z,H)^{T} K_H(Z_i,z)\, e_m e_m^{T} K_H(Z_j,z)\, R_j(z,H) &= \sum_{s=1}^{m}\sum_{t=1}^{m} \lambda_{it}\lambda_{js}[\cdot]_{it,js} \otimes \left(X_{it}X_{js}^{T}\right) \qquad (3)
\end{aligned}
$$

In addition, we have for a finite positive integer j

$$
\begin{aligned}
|H|^{-1}\sum_{t=1}^{m} E\left[\lambda_{it}^{j}[\cdot]_{it,it}|X_{it}\right] &= \sum_{t=1}^{m} E\left(S_{t,j,1}|X_{it}\right) + O_p\left(\|H\|^{2}\right), \qquad (4) \\
|H|^{-1}\sum_{t=1}^{m} E\left[\lambda_{it}^{2j}\sum_{j'=1}^{q} G_{itj'}^{2}[\cdot]_{it,it}|X_{it}\right] &= \sum_{t=1}^{m} E\left(S_{t,j,2}|X_{it}\right) + O_p\left(\|H\|^{2}\right), \qquad (5)
\end{aligned}
$$

where

$$
\begin{aligned}
S_{t,j,1} &= \begin{bmatrix} f_t(z|X_{it})\int K^{j}(u)\,du & \frac{\partial f_t(z|X_{it})}{\partial z^{T}} H R_{K,j} \\ R_{K,j} H \frac{\partial f_t(z|X_{it})}{\partial z} & f_t(z|X_{it})\, R_{K,j} \end{bmatrix} \qquad (6) \\
S_{t,j,2} &= \begin{bmatrix} f_t(z|X_{it})\int K^{2j}(u)\,u^{T}u\,du & \frac{\partial f_t(z|X_{1t})}{\partial z^{T}} H \Gamma_{K,2j} \\ \Gamma_{K,2j} H \frac{\partial f_t(z|X_{it})}{\partial z} & f_t(z|X_{it})\,\Gamma_{K,2j} \end{bmatrix} \qquad (7)
\end{aligned}
$$

where $R_{K,j} = \int K^{j}(u)\,uu^{T}du$ and $\Gamma_{K,2j} = \int K^{2j}(u)\left(u^{T}u\right)\left(uu^{T}\right)du$.

Moreover, for any finite positive integer j_1 and j_2, we have

$$
\begin{aligned}
&|H|^{-2}\sum_{t=1}^{m}\sum_{s\neq t} E\left[\lambda_{it}^{j_1}\lambda_{is}^{j_2}[\cdot]_{it,is}|X_{it},X_{is}\right] \qquad (8) \\
&= \sum_{t=1}^{m}\sum_{s\neq t} E\left(T_{j_1,j_2,1}^{(t,s)}|X_{it},X_{is}\right) + O_p\left(\|H\|^{2}\right)
\end{aligned}
$$

$$|H|^{-2}\sum_{t=1}^{m}\sum_{s\neq t}E\left[\lambda_{it}^{j_1}\lambda_{is}^{j_2}\left(\sum_{j'=1}^{q}G_{itj'}G_{isj'}\right)[\cdot]_{it,is}\,|X_{it},X_{is}\right] \qquad (9)$$
$$= \sum_{t=1}^{m}\sum_{s\neq t}E\left(T_{j_1,j_2,2}^{(t,s)}|X_{it},X_{is}\right)+O_p\left(\|H\|^2\right)$$

where we define $b_{j1,j_2,i_1,i_2=}\int K^{j_1}(u)\,u_1^{2i_1}du\int K^{j_2}(u)\,u_1^{2i_2}du$

$$T_{j_1,j_2,1}^{(t,s)}=\begin{bmatrix} f_{t,s}(z,z|X_{it},X_{is})\,b_{j_1,j_2,0,0} & \bigtriangledown_s^T f_{t,s}(z,z|X_{it},X_{is})\,Hb_{j_1,j_2,0,1} \\ H\bigtriangledown_t f_{t,s}(z,z|X_{it},X_{is})\,b_{j_1,j_2,1,0} & H\bigtriangledown_{t,s}^2 f_{t,s}(z,z|X_{it},X_{is})\,Hb_{j_1,j_2,1,1}\end{bmatrix}$$

and

$$T_{j_1,j_2,2}^{(t,s)}=\begin{bmatrix} tr\left(H\bigtriangledown_{t,s}^2 f_{t,s}(z,z|X_{it},X_{is})\,H\right) & \bigtriangledown_t^T f_{t,s}(z,z|X_{it},X_{is})\,H \\ H\bigtriangledown_s f_{t,s}(z,z|X_{it},X_{is}) & f_{t,s}(z,z|X_{it},X_{is})\,I_{q\times q}\end{bmatrix}b_{j_1,j_2,1,1},$$

with $\bigtriangledown_s f_{t,s}(z,z|X_{it},X_{is})=\partial f_{t,s}(z,z|X_{it},X_{is})/\partial z_s$ and $\bigtriangledown_{t,s}^2 f_{t,s}(z,z|X_{it},X_{is})=$ $\partial^2 f_{t,s}(z,z|X_{it},X_{is})/\left(\partial z_t\partial z_s^T\right)$.

The conditional bias and variance of $vec\left(\widehat{\beta}(z)\right)$ are given as follows:

$Bias\left[vec\left(\widehat{\beta}(z)\right)|\{X_{it},Z_{it}\}\right]=\frac{1}{2}\left[R(z,H)^T S_H(z)\,R(z,H)\right]^{-1}R(z,H)^T S_H(z)\,\Pi(z,H)$
and $Var\left[vec\left(\widehat{\beta}(z)\right)|\{X_{it},Z_{it}\}\right]=\sigma_v^2\left[R(z,H)^T S_H(z)\,R(z,H)\right]^{-1}\left[R(z,H)^T S_H^2(z)\,R(z,H)\right]$
$\times\left[R(z,H)^T S_H(z)\,R(z,H)\right]^{-1}$.

Lemma 6 *Under Assumptions 1-3, we have*

$$\left[n^{-1}|H|^{-1}R(z,H)^T S_H(z)\,R(z,H)\right]^{-1}\approx\begin{bmatrix}\Phi^{-1} & O_p(\|H\|) \\ O_p(\|H\|) & \left[\int K(u)\,uu^Tdu\right]^{-1}\otimes\Phi^{-1}\end{bmatrix},$$

where $\Phi=\sum_{t=1}^{m}\left(1-\sqrt{\frac{f_t(z)}{f(z)}}\right)E\left[f_t(z|X_{1t})\,X_{1t}X_{1t}^T\right]$.

Proof: First, simple calculation gives

$$
\begin{aligned}
A_n &= R(z,H)^T S_H(z) R(z,H) = R(z,H)^T W_H(z) M_H(z) R(z,H) \\
&= \sum_{i=1}^{n} R_i(z,H)^T K_H(Z_i,z) R_i(z,H) \\
&\quad - \sum_{j=1}^{n}\sum_{i=1}^{n} q_{ij} R_i(z,H)^T K_H(Z_i,z) e_m e_m^T K_H(Z_j,z) R_j(z,H) \\
&= \sum_{i=1}^{n}\sum_{t=1}^{m} \lambda_{it} [\cdot]_{it,it} \otimes \left(X_{it}X_{it}^T\right) - \sum_{i=1}^{n} q_{ii} \sum_{t=1}^{m} \lambda_{it}^2 [\cdot]_{it,it} \otimes \left(X_{it}X_{it}^T\right) \\
&\quad - \sum_{i=1}^{n} q_{ii} \sum_{s=1}^{m}\sum_{t\neq s} \lambda_{it}\lambda_{is} [\cdot]_{it,is} \otimes \left(X_{it}X_{is}^T\right) - \sum_{j=1}^{n}\sum_{i\neq j} q_{ij} \sum_{s=1}^{m}\sum_{t=1}^{m} \lambda_{it}\lambda_{js} [\cdot]_{it,js} \otimes \left(X_{it}X_{js}^T\right) \\
&= A_{n1} - A_{n2} - A_{n3} - A_{n4},
\end{aligned}
$$

where $M_H(z) = I_{(nm)\times(nm)} - \left(Q \otimes e_m e_m^T\right) W_H(z)$, and the typical elements of Q are $q_{ii} = c_H(Z_i,z) - c_H(Z_i,z)^2 / \sum_{i=1}^n c_H(Z_i,z)$ and $q_{ij} = -c_H(Z_i,z)\, c_H(Z_j,z) / \sum_{i=1}^n c_H(Z_i,z)$ for $i \neq j$. Simple calculations give

$$
\begin{aligned}
\sum_{t=1}^{m} K_H(Z_{it},z) &= |H| f(z) + |H|^{1/2} \left(f(z) \int K^2(u)\, du \right)^{1/2} + O_p\left(|H|^{\frac{1}{2}} \|H\| \right), \\
K_H(Z_{it},z) &= |H| f_t(z) + |H|^{1/2} \left(f_t(z) \int K^2(u)\, du \right)^{1/2} + O_p\left(|H|^{\frac{1}{2}} \|H\| \right),
\end{aligned}
$$

and it follows that $c_H(Z_i,z) = |H|^{-1/2} \left(f(z) \int K^2(u)\, du\right)^{-1/2} + o_p\left(|H|^{-\frac{1}{2}}\right)$ is the leading term of q_{ii}, and that $q_{ij} = O_p\left(n^{-1}\right)$. As a result, it is easy to show that A_{n4} is dominated by the other three terms.

Applying (4), (5), (8), and (9) to A_{n1}, we have $n^{-1} |H|^{-1} A_{n1} \approx \sum_{t=1}^m E\left[S_{t,1,1} \otimes \left(X_{it}X_{it}^T\right)\right] + O_p\left(\|H\|^2\right) + O_p\left(n^{-\frac{1}{2}} |H|^{-\frac{1}{2}}\right)$ if $\|H\| \to 0$ and $n|H| \to \infty$ as $n \to \infty$.

To cope with the difficulty caused by the random denominator $c_H(Z_i,z)$ when taking expectations and variations of A_{n2} and A_{n3}, we apply the technique used for kernel curve estimation. Specifically, define $\hat{\omega}_{it}(z) = \lambda_{it}^{-1} c_H(Z_i,z)^{-1}$. It then follows that $\hat{\omega}_{it}(z) = (f(z)/f_t(z))^{1/2} + o_p(1)$. Define $\omega_t(z) = (f(z)/f_t(z))^{1/2}$

and $\hat{g}_{it}(z) = \lambda_{it} [\cdot]_{it,it} \otimes \left(X_{it} X_{it}^T\right)$. Then

$$A_{n2} \approx \sum_{i=1}^{n} \sum_{t=1}^{m} \frac{\hat{g}_{it}(z)}{\hat{\omega}_{it}(z)} = \sum_{i=1}^{n} \sum_{t=1}^{m} \frac{\hat{g}_{it}(z)}{\hat{\omega}_{it}(z)} \left[\frac{\hat{\omega}_{it}(z)}{\omega_t(z)} + \left(1 - \frac{\hat{\omega}_{it}(z)}{\omega_t(z)}\right) \right].$$

Since $\hat{\omega}_{it}(z) = \omega_t(z) + o_p(1)$, it can be shown that $\sum_{i=1}^{n} \sum_{t=1}^{m} \hat{g}_{it}(z) / \omega_t(z)$ is the leading term of A_{n2}. Again applying (4), (5), (8), and (9), we have $n^{-1} |H|^{-1} \sum_{i=1}^{n} \sum_{t=1}^{m} \hat{g}_{it}(z) / \omega_t(z) \approx \sum_{t=1}^{m} \omega_t(z)^{-1} E\left[S_{t,1,1} \otimes \left(X_{it} X_{it}^T\right)\right] + O_p\left(\|H\|^2\right) + O_p\left(n^{-\frac{1}{2}} |H|^{-\frac{1}{2}}\right)$.

Similarly, for A_{n3}, we define $\hat{\omega}_i(z) = |H|^{-1/2} c_H(Z_i, z)^{-1}$ and $\omega_i(z) = \left(f(z) \int K^2(u)\, du\right)^{1/2}$, and we can show that $n^{-1} |H|^{-1} A_{n3} \approx O_p\left(|H|^{1/2}\right) + O_p\left(n^{-1/2} |H|^{-1/2}\right)$ if $\|H\| \to 0$ and $n|H| \to \infty$ as $n \to \infty$. Taking all the results together and will complete the proof of this lemma, where the inverse matrix is derived by using Theorem A.4.4. in Poirier (1995, p.627) .

Lemma 7 *Under Assumptions 1-3, we have* $n^{-1} |H|^{-1} C_n \approx \begin{bmatrix} \Phi\left(\int K(u)\, uu^T du\right) \Theta_H(z) \\ O_p\left(\|H\|^3\right) \end{bmatrix}$, *where* $\Theta_H(z) = \left[tr\left(H \frac{\partial^2 \theta_1(z)}{\partial z \partial z^T} H\right), \cdots, tr\left(H \frac{\partial^2 \theta_k(z)}{\partial z \partial z^T} H\right)\right]^T$.

Proof: Simple calculations give

$$\begin{aligned}
C_n &= R(z,H)^T S_H(z) \Pi(z,H) \\
&= \sum_{i=1}^{n} \sum_{t=1}^{m} \lambda_{it} (G_{it} \otimes X_{it}) X_{it}^T r_H\left(\tilde{Z}_{it}, z\right) - \sum_{j=1}^{n} \sum_{i=1}^{n} q_{ij} \sum_{s=1}^{m} \sum_{t=1}^{m} \lambda_{js} \lambda_{it} (G_{it} \otimes X_{it}) X_{js}^T r_H\left(\tilde{Z}_{js}, z\right) \\
&= \sum_{i=1}^{n} \sum_{t=1}^{m} \lambda_{it} (G_{it} \otimes X_{it}) X_{it}^T r_H\left(\tilde{Z}_{it}, z\right) - \sum_{i=1}^{n} q_{ii} \sum_{t=1}^{m} \lambda_{it}^2 (G_{it} \otimes X_{it}) X_{it}^T r_H\left(\tilde{Z}_{it}, z\right) \\
&\quad - \sum_{i=1}^{n} q_{ii} \sum_{s=1}^{m} \sum_{t \neq s} \lambda_{is} \lambda_{it} (G_{it} \otimes X_{it}) X_{is}^T r_H\left(\tilde{Z}_{is}, z\right) \\
&\quad - \sum_{j=1}^{n} \sum_{i \neq j} q_{ij} \sum_{s=1}^{m} \sum_{t=1}^{m} \lambda_{js} \lambda_{it} (G_{it} \otimes X_{it}) X_{js}^T r_H\left(\tilde{Z}_{js}, z\right) \\
&= C_{n1} - C_{n2} - C_{n3} - C_{n4},
\end{aligned}$$

where $\Pi(z,H)$ is defined in Section 3. Similar to the proof of Lemma 6, it is easy to show that $C_{n,3}$ and $C_{n,4}$ are dominated by the other two terms of C_n.

For $l = 1, \cdots, k$ we have

$$|H|^{-1} E\left[\lambda_{it} r_{H,l}(Z_{it}, z) | X_{it}\right] = f_t(z|X_{it})\, tr\left(\left(\int K(u)\, uu^T du\right) H \frac{\partial^2 \theta_l(z)}{\partial z \partial z^T} H\right) + O_p\left(\|H\|^4\right)$$
$$|H|^{-1} E\left[\lambda_{it} r_{H,l}(Z_{it}, z) H^{-1}(Z_{it} - z) | X_{it}\right] = O_p\left(\|H\|^3\right),$$

and $E\left(n^{-1}|H|^{-1} C_{n1}\right) \approx \left\{\left[\Phi\left(\int K(u)\, uu^T du\right) \Theta_H(z)\right]^T, O\left(\|H\|^3\right)\right\}^T$. Similarly we can show that $Var\left(n^{-1}|H|^{-1} C_{n1}\right) = O\left(n^{-1}|H|^{-1}\|H\|^4\right)$ if $E\left(\left\|X_{it} X_{is}^T X_{it} X_{is}^T\right\|\right) < M < \infty$ for all t and s.

For C_{n2}, we use the same method to cope with the random denominator problem as in the proof of Lemma 6. Then we can show that the dominant term of C_{n2} is $\sum_{i=1}^n \sum_{t=1}^m \sqrt{\frac{f_t(z)}{f(z)}} \lambda_{it} (G_{it} \otimes X_{it}) X_{it}^T r_H(Z_{it}, z)$.

This will complete the proof of this lemma.

Lemma 8 *Under Assumptions 1-3, we have*

$$n^{-1}|H|^{-1} B_n \approx 2\begin{bmatrix} \left(\int K^2(u)\, du\right)\Phi & O_p(\|H\|) \\ O_p(\|H\|) & \left[\int K^2(u)\, uu^T du\right] \otimes \Phi \end{bmatrix}.$$

Proof: Simple calculations give

$$\begin{aligned}
B_n &= R(z,H)^T S_H^2(z) R(z,H) = R(z,H)^T W_H(z) M_H(z) M_H(z)^T W_H(z) R(z,H) \\
&= \sum_{i=1}^n R_i(z,H)^T K_H^2(Z_i,z) R_i(z,H) - \sum_{j=1}^n \sum_{i=1}^n q_{ji} R_j(z,H)^T K_H^2(Z_j,z) e_m e_m^T K_H(Z_i,z) R_i(z,H) \\
&\quad - \sum_{j=1}^n \sum_{i=1}^n q_{ij} R_i(z,H)^T K_H(Z_i,z) e_m e_m^T K_H^2(Z_j,z) R_j(z,H) \\
&\quad + \sum_{j=1}^n \sum_{i=1}^n \sum_{i'=1}^n q_{ij} q_{ji'} R_i(z,H)^T K_H(Z_i,z) e_m e_m^T K_H^2(Z_j,z) e_m e_m^T K_H(Z_{i'},z) R_{i'}(z,H) \\
&= B_{n1} - B_{n2} - B_{n2}^T + B_{n3}.
\end{aligned}$$

Similar to the proof of Lemma 6, we can show that $n^{-1}|H|^{-1} B_n \approx 2n^{-1}|H|^{-1}(B_{n1} - B_{n2})$ and that the leading term of B_{n1} is $\sum_{i=1}^n \sum_{t=1}^m \lambda_{it}^2 [\cdot]_{it,it} \otimes \left(X_{it} X_{it}^T\right)$ and the leading term of B_{n2} is $\sum_{i=1}^{n} q_{ii} \sum_{t=1}^{m} \lambda_{it}^{3} [\cdot]_{it,it} \otimes \left(X_{it} X_{it}^{T}\right)$. Then we obtain $n^{-1} |H|^{-1} B_n \approx \sum_{t=1}^{m} \left(1 - \sqrt{\frac{f_t(z)}{f(z)}}\right) E\left[S_{t,2,1} \otimes \left(X_{it} X_{it}^{T}\right)\right] + O_p\left(\|H\|^2\right) + O_p\left(n^{-1/2} |H|^{-1/2}\right)$.

The three lemmas above are enough to give the result of Theorem 3. Moreover, applying Liaponuov's CLT will give the result of Theorem 4. Since the proof is a rather standard procedure, we drop the details for compactness of the paper.

Appendix B: Technical Sketch–Random Effects Estimator

The RE estimator, $\hat{\theta}_{RE}(\cdot)$, is the solution to the following optimization problem:

$$\min_{\beta(z)} \left[Y - R(z,H)\, vec(\beta(z))\right]^T W_H(z) \left[Y - R(z,H)\, vec(\beta(z))\right];$$

that is, we have

$$\begin{aligned} vec\left(\hat{\beta}_{RE}(z)\right) &= \left[R(z,H)^T W_H(z) R(z,H)\right]^{-1} R(z,H)^T W_H(z) Y \\ &= vec(\beta(z)) + \left[R(z,H)^T W_H(z) R(z,H)\right]^{-1} R(z,H)^T W_H(z) D\mu \\ &\quad + \left[R(z,H)^T W_H(z) R(z,H)\right]^{-1} R(z,H)^T W_H(z) \left(\tilde{A}_n/2 + \tilde{B}_n\right) \end{aligned}$$

where $\tilde{A}_n = R(z,H)^T W_H(z) \Pi(z,H)$ and $\tilde{B}_n = R(z,H)^T W_H(z) V$. Its asymptotic results are as follows.

Lemma 9 *Under Assumptions 1-3, $\sqrt{n|H|}\,\|H\|^2 = o(1)$ as $n \to \infty$, and $E\left(|v_{it}|^{2+\delta}\right) < \infty, E\left(|\mu_i|^{2+\delta}\right) < M < \infty$ and $E\left(\|X_{it}\|^{2+\delta}\right) < M < \infty$ for all i and t and for some $\delta > 0$, we have under H_0*

$$\sqrt{n|H|}\left(\widehat{\theta}_{RE}(z) - \theta(z)\right) \xrightarrow{d} N\left(0, \Sigma_{\theta(z),RE}\right), \tag{10}$$

where $\Sigma_{\theta(z),RE} = \left(\frac{2}{n}\sum_{i=2}^{n} \sigma_i^2 + \sigma_v^2\right) \Psi^{-1} \int K^2(u)\, du$ and $\Psi = \sum_{t=1}^{m} E\left[f_t(z|X_{1,t})\, X_{1,t} X_{1,t}^T\right]$.

Under H_1, we have

$$\begin{aligned} Bias\left(\widehat{\theta}_{RE}(z)\right) &= \left(\int K(u)\, uu^T du\right)\Theta_H(z) \\ &+\frac{1}{2n}\sum_{i=1}^{n}\sum_{t=1}^{m} E\left[f_t\left(z|X_{i,t},\mu_i\right)\mu_i X_{i,t}\right] + o\left(\|H\|^2\right) \\ Var\left[\hat{\theta}_{RE}(z)\,|\,\{X_{it},Z_{it}\}\right] &= \sigma_v^2 n^{-1}|H|^{-1}\Psi^{-1}\int K^2(u)\,du. \end{aligned}$$

Proof of Lemma 9: First, we have the following decomposition

$$\sqrt{n|H|}\left[\hat{\theta}_{RE}(z)-\theta(z)\right] = \sqrt{n|H|}\left[\hat{\theta}_{RE}(z)-E\left(\hat{\theta}_{RE}(z)\right)\right]+\sqrt{n|H|}\left[E\left(\hat{\theta}_{RE}(z)\right)-\theta(z)\right],$$

where we are going to show that the first term converges to a normal distribution with mean zero and the second term contributes to the asymptotic bias. Without confusing readers, we drop the subscription 'RE'.

a. Under H_0, the conditional bias and variance of $\hat{\theta}(z)$ are as follows: $Bias_0\left[\hat{\theta}(z)\,|\,\{X_{it},Z_{it}\}\right] = S_k\left[R(z,H)^T W_H(z)\, R(z,H)\right]^{-1} R(z,H)^T W_H(z)\,\Pi(z,H)/2$ and $Var\left[\hat{\theta}(z)\,|\,\{X_{it},Z_{it}\}\right] = S_k\left[R(z,H)^T W_H(z)\, R(z,H)\right]^{-1}\times\left[R(z,H)^T W_H(z)\, Var(UU^T) W_H(z)\, R(z,H)\right.$ $\left.\left[R(z,H)^T W_H(z)\, R(z,H)\right]^{-1} S_k^T\right.$. It is simple to show that $Var(UU^T) = D\Sigma_\mu D^T + \sigma_v^2 I_{(nm)\times(nm)}$, where $\Sigma_\mu = diag\left(\sigma_2^2,\cdots,\sigma_n^2\right)$ is the covariance matrix of $(\mu_2,\cdots,\mu_n)$.

b. Under H_1, we notice that $Bias_1\left[\hat{\theta}(z)\,|\,\{X_{it},Z_{it}\}\right]$ is the sum of $Bias_0\left[\hat{\theta}(z)\,|\,\{X_{it},Z_{it}\}\right]$ plus an additional term $S_k\left[R(z,H)^T W_H(z)\, R(z,H)\right]^{-1} R(z,H)^T W_H(z)\, D\mu$. It is easy to show that under Assumptions 1-3, and $E\left(|\mu_i|^{2+\delta}\right) < M < \infty$ and $E\left(\|X_{it}\|^{2+\delta}\right) < M < \infty$ for all i and t and for some $\delta > 0$

$$\begin{aligned} & n^{-1}|H|^{-1} S_k R(z,H)^T W_H(z)\, D\mu \\ = \;& n^{-1}\sum_{i=1}^{n}\sum_{t=1}^{m} E\left[f_t\left(z|X_{it},\mu_i\right)\mu_i X_{it}\right] + O_p\left(\|H\|^2\right) + O_p\left(\frac{1}{\sqrt{n|H|}}\right) \end{aligned} \quad (11)$$

which is $o_p(1)$ under H_0 and is a non-zero constant plus a term of $o_p(1)$ under H_1. Based on the facts that $R(z,H)^T W_H(z)\, R(z,H)$ is A_{n1} in Lemma 6 and

$R(z,H)^T W_H(z)\Pi(z,H)$ is C_{n1} in Lemma 7 we have

$$Bias_0\left[\hat{\theta}(z)\,|\,\{X_{it},Z_{it}\}\right]=\left(\int K(u)\,uu^T du\right)\Theta_H(z)+o\left(\|H\|^2\right),$$

which is the same as the bias term of the FE estimator. Our results indicate that under H_1 the bias of the RE estimator will not vanish as $n\to\infty$ and this leads to the inconsistency of the RE estimator under H_1.

As for the conditional variance, if we denote $\Psi=\sum_{t=1}^m E\left[f_t(z|X_{1t})\,X_{1t}X_{1t}^T\right]$, we can easily show that under H_0

$$Var\left[\hat{\theta}(z)\,|\,\{X_{it},Z_{it}\}\right]=n^{-1}|H|^{-1}\left(2n^{-1}\sum_{i=2}^n\sigma_{u_i}^2+\sigma_v^2\right)\Psi^{-1}\int K^2(u)\,du, \tag{12}$$

and under H_1

$$Var\left[\hat{\theta}(z)\,|\,\{X_{it},Z_{it}\}\right]=\sigma_v^2 n^{-1}|H|^{-1}\Psi^{-1}\int K^2(u)\,du, \tag{13}$$

where we have recognized that $R(z,H)^T W_H(z)^2 R(z,H)$ is B_{n1} in Lemma 8.

A Appendix C: Technical Sketch–Proof of Theorem 5

Define $\triangle_i=(\triangle_{i,1},\cdots,\triangle_{i,m})^T$ with $\triangle_{i,t}=X_{i,t}^T\left(\theta(Z_{i,t})-\hat{\theta}_{RE}(Z_{i,t})\right)$. Since $M_D D=0$, we can decompose the proposed statistic into four terms

$$\begin{aligned}\hat{T}_n &= \frac{1}{n^2|H|}\sum_{i=1}^n\sum_{j\neq i}\hat{U}_i^T Q_m A_{i,j} Q_m \hat{U}_j\\ &= \frac{1}{n^2|H|}\sum_{i=1}^n\sum_{j\neq i}\triangle_i^T Q_m A_{i,j} Q_m\triangle_j+\frac{2}{n^2|H|}\sum_{i=1}^n\sum_{j\neq i}\triangle_i^T Q_m A_{i,j} Q_m V_j\\ &\quad+\frac{1}{n^2|H|}\sum_{i=1}^n\sum_{j\neq i}V_i^T Q_m A_{i,j} Q_m V_j\\ &= T_{n1}+2T_{n2}+T_{n3}\end{aligned}$$

where $V_i=(v_{i,1},\cdots,v_{i,m})^T$ is the $m\times 1$ error vector. Since $\hat{\theta}_{RE}(Z_{i,t})$ is the leave-one-unit-out estimator for a pair of (i,j), it is easy to see that $E(T_{n2})=0$.

The proofs fall into the standard procedures seen in the literature of nonparametric tests. We therefore give a very brief proof below.

Firstly, applying Hall's (1984) CLT, we can show that under both H_0 and H_1

$$n\sqrt{|H|}T_{n3} \xrightarrow{d} N\left(0, \sigma_0^2\right) \tag{14}$$

by defining $H_n\left(\chi_i, \chi_j\right) = V_i^T Q_m A_{i,j} Q_m V_j$ with $\chi_i = (X_i, Z_i, V_i)$, which a symmetric, centred and degenerate variable. We are able to show that

$$\frac{E\left[G_n^2\left(\chi_1, \chi_2\right)\right] + n^{-1}E\left[H_n^4\left(\left(\chi_1, \chi_2\right)\right)\right]}{\left\{E\left[H_n^2\left(\left(\chi_1, \chi_2\right)\right)\right]\right\}^2} = \frac{O\left(|H|^3\right) + O\left(n^{-1}|H|\right)}{O\left(|H|^2\right)} \to 0$$

if $|H| \to 0$ and $n|H| \to \infty$ as $n \to \infty$, where $G_n\left(\chi_1, \chi_2\right) = E_{\chi_i}\left[H_n\left(\left(\chi_1, \chi_i\right)\right) H_n\left(\left(\chi_2, \chi_i\right)\right)\right].$ In addition, $|H|\sigma_0^2 = 2|H|E\left(H_n^2\left(\chi_1, \chi_2\right)\right) \approx 2\left(1 - m^{-1}\right)^2 \sigma_v^2 |H| \sum_{t=1}^m \sum_{s=1}^m E\left[K_H^2\left(Z_{1s}, Z_{2t}\right)\left(X_{1s}^T X_{2t}\right)^2\right].$

Secondly, we can show that $n\sqrt{|H|}T_{n2} = O_p\left(\|H\|^2\right) + O_p\left(n^{-1/2}|H|^{-1/2}\right)$ under H_0 and $n\sqrt{|H|}T_{nj} = O_p(1)$ under H_1. Moreover, we have, under H_0, $n\sqrt{|H|}T_{n1} = O_p\left(n\sqrt{|H|}\,\|H\|^4\right)$; under H_1, $n\sqrt{|H|}T_{n1} = O_p\left(n\sqrt{|H|}\right).$

Finally, to estimate σ_0^2 consistently under both H_0 and H_1, we replace the unknown V_i and V_j in T_{n3} by the estimated residual vectors from FE estimator. Simple calculations show that the typical element of $\hat{V}_i Q_m$ is $\widetilde{\tilde{v}}_{it} = y_{it} - X_{it}^T \hat{\theta}_{FE}\left(Z_{it}\right) - v_{it} - \left(\bar{y}_i - m^{-1}\sum_{t=1}^m X_{it}^T \hat{\theta}_{FE}\left(Z_{it}\right) - \bar{v}_i\right) = \triangle_{it} - \left(v_{it} - \bar{v}_i\right)$, where $\triangle_{it} = X_{it}^T\left(\theta\left(Z_{it}\right) - \hat{\theta}_{FE}\left(Z_{it}\right)\right) - m^{-1}\sum_{t=1}^m X_{it}^T\left(\theta\left(Z_{it}\right) - \hat{\theta}_{FE}\left(Z_{it}\right)\right) = \sum_{l=1}^m q_{lt} X_{il}^T\left(\theta\left(Z_{il}\right) - \hat{\theta}_{FE}\left(Z_{il}\right)\right)$ with $q_{tt} = 1 - 1/m$ and $q_{lt} = -1/m$ for $l \neq m$. The leave-two-unit-out FE estimator does not use the observations from the ith and jth units, and this leads to $E\left(\hat{V}_i^T Q_m A_{i,j} Q_m \hat{V}_j\right)^2 \approx \sum_{t=1}^m \sum_{s=1}^m E\left[K_H^2\left(Z_{is}, Z_{jt}\right)\left(X_{is}^T X_{jt}\right)^2 \left(\triangle_{it}^2 \triangle_{js}^2 + \triangle_i^2\right.\right.$ where $\tilde{v}_{it} = v_{it} - \bar{v}_i$ and $\bar{v}_i = m^{-1}\sum_{t=1}^m v_{it}$.

References

Hall, P. (1984). Central limit theorem for integrated square error of multivariate nonparametric density estimators. *Annals of Statistics*, 14, 1-16.

Härdle, W. and Mammen, E. (1991). Bootstrap simultaneous error bars for nonparametric regression. *Annals of Statistics*, 19, 778-796.

Henderson, D. J. and Ullah, A. (2005). A Nonparametric Random Effects Estimator. *Economics Letters*, 88, 403-407.

Ke, C. and Y. Wang (2001). Semiparametric Nonlinear Mixed-Effects Models and their Applications. *Journal of the American Statistical Association*, 96, 1272-1281.

Li, Q. and Huang, C.J., Li, D. and Fu, T. (2002). Semiparametric Smooth Coefficient Models. *Journal of Business & Economic Statistics*, 20, 412-422.

Li, Q. and Stengos, T. (1996). Semiparametric Estimation of Partially Linear Panel Data Models. *Journal of Econometrics*, 71, 389-397.

Li, Q. and Wang, S. (1998). A Simple Consistent Bootstrap Test for a Parametric Regression Function. *Journal of Econometrics*, 87, 145-165.

Lin, D. Y. and Ying, Z. (2001). Semiparametric and Nonparametric Regression Analysis of Longitudinal Data (with Discussion). *Journal of the American Statistical Association*, 96, 103-126.

Lin, X., and Carroll, R. J. (2000). Nonparametric Function Estimation for Clustered Data when the Predictor is Measured without/with Error. *Journal of the American Statistical Association*, 95, 520-534.

Lin, X. and Carroll, R. J. (2001). Semiparametric Regression for Clustered Data using Generalized Estimation Equations. *Journal of the American Statistical Association*, 96, 1045-1056.

Lin, X. and Carroll, R. J. (2006). Semiparametric Estimation in General Repeated Measures Problems. *Journal of the Royal Statistical Society, Series*

B, 68, 68-88.

Lin, X., Wang, N., Welsh, A. H. and Carroll, R. J. (2004). Equivalent Kernels of Smoothing Splines in Nonparametric Regression for Longitudinal/Clustered Data. *Biometrika*, 91, 177-194.

Poirier, D.J. (1995). Intermediate Statistics and Econometrics: a Comparative Approach. The MIT Press.

Ruckstuhl, A. F., Welsh, A. H. and Carroll, R. J. (2000). Nonparametric Function Estimation of the Relationship Between Two Repeatedly Measured Variables. *Statistica Sinica*, 10, 51-71.

Su, L. and Ullah, A. (2006). Profile likelihood estimation of partially linear panel data models with fixed effects. *Economics Letters 92, 75-81.*

Ullah, A. and Roy, N. (1998). Nonparametric and Semiparametric Econometrics of Panel Data, in *Handbook of Applied Economics Statistics*, A. Ullah and D. E. A. Giles, eds., Marcel Dekker, New York, 579-604.

Wang, N. (2003). Marginal Nonparametric Kernel Regression Accounting for Within-Subject Correlation. *Biometrika*, 90, 43-52.

Wu, H. and Zhang, J. Y. (2002). Local Polynomial Mixed-Effects Models for Longitudinal Data. *Journal of the American Statistical Association*, 97, 883-897.

Zheng, J. (1996). A Consistent Test of Functional Form via Nonparametric Estimation Techniques. *Journal of Econometrics 75, 263-289.*

Table 1: Average mean squared errors (AMSE) of the fixed and random effects estimators when the data generation process is a random effects model and when it is a fixed effects model.

	Random Effects Estimator			Fixed Effects Estimator		
Data Process	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
Estimating $\theta_1(\cdot)$:						
$c_0 = 0$	.0951	.0533	.0277			
$c_0 = 0.5$	.6552	.5830	.5544	.1381	.1163	.1021
$c_0 = 1.0$	2.2010	2.1239	2.2310			
Estimating $\theta_2(\cdot)$:						
$c_0 = 0$	.1562	.0753	.0409			
$c_0 = 0.5$	.8629	.7511	.7200	.1984	.1379	.0967
$c_0 = 1.0$	2.8707	2.4302	2.5538			

Table 2: Percentage Rejection Rate from 1000 Monte Carlo Simulations with c=1.0

c_0	$n = 50$			$n = 100$		
	1%	5%	10%	1%	5%	10%
0	.011	.023	.041	.025	.04	.062
0.5	.682	.780	.819	.935	.943	.951
1.0	.908	.913	.921	.962	.966	.967