9-2007

More efficient estimation of nonparametric panel data models with random effects

Liangjun SU
Singapore Management University, ljsu@smu.edu.sg

Aman ULLAH
University of California, Riverside

Published in Economics Letters, 2007 September, Volume 96, Issue 3, Pages 375-380
http://doi.org/10.1016/j.econlet.2007.02.018

More efficient estimation of nonparametric panel data models with random effects ☆

Liangjun Su [a,*], Aman Ullah [b,1]

[a] *Guanghua School of Management, Peking University, Beijing, 100871, PR China*
[b] *Department of Economics, University of California, Riverside, CA 92521-0427, United States*

Abstract

We propose a class of two-step estimators for nonparametric panel data models with random effects that are more efficient than the conventional least squares estimators. We establish asymptotic normality for the proposed estimators and derive the most efficient estimator in the class.

Keywords: Nonparametrics; Panel data models; Random effects; Efficiency

JEL classification: C1; C14; C33

1. Introduction

Recently nonparametric panel data models have become a research topic of much interest. This is especially true for random effects nonparametric panel data models. See Li and Stengos (1996), Li and Ullah (1998), Lin and Carroll (2000), Henderson and Ullah (2005), among others. In such models the regression mean is generally estimated by the local linear least squares (LLLS) estimator. But this estimator ignores the covariance structure in the disturbance term and thus is inefficient. To take into account the covariance structure, Henderson and Ullah (2005) propose local linear weighted least squares

☆ The first author gratefully acknowledges the financial support from the NSFC (Project 70501001). The second author gratefully acknowledges the financial support from the Academic Senate, UCR.
* Corresponding author. Tel.: +86 10 62767444.
E-mail addresses: lsu@gsm.pku.edu.cn (L. Su), aman.ullah@ucr.edu (A. Ullah).
[1] Tel.: +1 951 8271591.

estimation. Unfortunately, as Lin and Carroll (2000) demonstrate, one cannot achieve asymptotic improvement over the LLLS estimator by this approach.

In this paper, we follow the idea of Ruckstuhl et al. (2000) and propose a class of two-step estimators that employ the covariance structure to achieve asymptotic improvement over the LLLS estimator. In comparison with their approach, our estimators are different in at least three aspects. First, we allow the regressor to be a random vector. Second, we consider more efficient estimation of both the nonparametric regression mean and its first order derivatives. Third, we use two different bandwidth sequences in the two steps that will greatly facilitate the comparison of our estimator with the conventional LLLS estimator.

The paper is structured as follows. We introduce a class of two-step estimators in Section 2 and study their asymptotic property in Section 3. We conduct Monte Carlo simulations in Section 4.

2. The estimator

We consider a two-step estimation of nonparametric panel data models with random effects:

$$y_{it} = m(z_{it}) + u_i + v_{it},\ \ i = 1, 2, \ldots, n, t = 1, 2, \ldots T, \tag{2.1}$$

where z_{it} is a $q\times 1$ vector of exogenous variables, u_i is i.i.d. (0, σ_u^2), v_{it} is i.i.d. (0, σ_v^2), u_i and v_{jt} are uncorrelated for all i, j=1, 2,…n, and m (·) is an unknown smooth function. We are interested in consistent estimation of $m(z)$ and its first order derivatives $\dot{m}(z)$ at an interior point z. We establish the asymptotic theory by letting n approach infinity and holding T fixed.

Let $\varepsilon_{it}=u_i+v_{it}$ and $\varepsilon_i=(\varepsilon_{i1}, \ldots, \varepsilon_{iT})'$. Then $\Sigma\equiv E(\varepsilon_i\varepsilon_i')$ takes the form $\Sigma=\sigma_v^2 I_T+\sigma_u^2 i_T i_T'$, where I_T is a $T\times T$ identity matrix and i_T is a $T\times 1$ vector of ones. Let K denote a kernel function on $\mathbb{R}^q$ and H=diag (h_1, …, h_q), a matrix of bandwidth sequences. Set $K_H(z)=|H|^{-1}K(H^{-1}z)$ and $\vec{Z}_{it}(z) = [1, \{H^{-1}(z_{it}-z)\}']'$, where $|H|$ is the determinant of H. Further denote $\mathbf{K}_H(z)=\text{diag}(K_H(z_{11}-z), \ldots, K_H(z_{1T}-z), \ldots, K_H(z_{n1}-z), \ldots, K_H(z_{nT}-z))$ and $\vec{Z}(z) = (\vec{Z}_{11}(z), \ldots, \vec{Z}_{1T}(z), \ldots, \vec{Z}_{n1}(z), \ldots, \vec{Z}_{nT}(z))'$. The LLLS estimator of $M(z)\equiv(m(z), (H\dot{m}(z))')'$ is obtained by

$$\widetilde{M}_H(z) = \underset{M\in\mathbb{R}^{q+1}}{\text{argmin}}\left(Y-\vec{Z}(z)M\right)'\mathbf{K}_H(z)\left(Y-\vec{Z}(z)M\right), \tag{2.2}$$

where $Y=(y_{11}, \ldots, y_{1T}, \ldots, y_{n1}, \ldots, y_{nT})'$. That is, $\widetilde{M}(z)=S(z)Y$, where $S(z) = [\vec{Z}(z)'\mathbf{K}_H(z)\vec{Z}(z)]^{-1}\vec{Z}(z)'\mathbf{K}_H(z)$ is a smoothing operator. In particular, the estimator for $m(z)$ is given by $\widetilde{m}_H(z)=s(z)'Y$, where $s(z)'=e'S(z)$, and $e=(1, 0, \ldots, 0)'$ is a $(q+1)\times 1$ vector.

The LLLS estimator $\widetilde{M}_H(z)$ ignores the information contained in the variance matrix Σ and it is inefficient. For this reason, Henderson and Ullah (2005) propose local linear weighted least squares estimation. However, Lin and Carroll (2000) demonstrate that one cannot achieve asymptotic improvement over the LLLS estimator by such an approach. Here, we follow the idea of Ruckstuhl et al. (2000) and propose a class of two-step estimators that employ the covariance structure to achieve asymptotic improvement over the LLLS estimator.

For the moment, we assume that Σ is known. Let $P=\Sigma^{-1}$ and let $P^{1/2}$ be its symmetric square root. The procedure is as follows:

1. Choose bandwidth $H_0=\text{diag}(h_{01}, \ldots, h_{0q})$ and obtain the usual LLLS estimator of $M(\tilde{z})$ as in Eq. (2.2) and denote it by $\widetilde{M}_{H_0}(\tilde{z})$ for $\tilde{z}=z_{11}, \ldots, z_{nT}$.

2. Define $Y_i^* = \tau P^{1/2} Y_i - (\tau P^{1/2} - I_T)\widetilde{m}_{H_0}(z_i)$ and denote its typical element as y_{it}^*, where $Y_i = (y_{i1}, \ldots, y_{iT})'$ and $\widetilde{m}_{H_0}(z_i) = (\widetilde{m}_{H_0}(z_{i1}) \ldots, \widetilde{m}_{H_0}(z_{iT}))'$. Obtain the estimator $\widetilde{M}_H(z)$ by regressing $y_{it}^*{}'s$ on $z_{it}'\ s$:

$$\widetilde{M}_H(z) \equiv \arg\min_{M \in \mathbb{R}^{q+1}} \left(Y^* - \overrightarrow{Z}(z)M\right)' \mathbf{K}_H(z) \left(Y^* - \overrightarrow{Z}(z)M\right), \tag{2.3}$$

where $Y^* = (Y_1^*{}', \ldots, Y_n^*{}')'$. In particular, $m(z)$ is estimated by $\widetilde{m}_H(z) = e'\widehat{M}_H(z)$.

The intuition for the above estimator is simple. Write $Y_i = m(Z_i) + \varepsilon_i$, multiply both sides by $\tau P^{1/2}$ and rearrange terms so that we have $\tau P^{1/2} Y_i - (\tau P^{1/2} - I_T)m(Z_i) = m(Z_i) + \tau P^{1/2}\varepsilon_i$. Y_i^* is a consistent estimator of the left hand side expression but the disturbance terms on the right hand side are now i.i.d. We will show that $\widehat{M}_H(z)$ is more efficient than $\widetilde{M}_H(z)$ under weak conditions. Also, replacing Σ by its consistent estimator does not affect the asymptotic relative efficiency of $\widehat{M}_H(z)$.

3. Asymptotic property

First, we make the following assumptions.

A1. (u_i, v_i, z_i), $i = 1, \ldots, n$, are i.i.d., where $v_i = (v_{i1}, \ldots, v_{iT})'$ and z_i is similarly defined. A2. z_{it} has a continuous density function $f_t(\cdot)$ with compact support C_f on $\mathbb{R}^q$. $f_t(\cdot)$ is bounded away from zero and infinity on C_f for $t = 1, \ldots, T$. A3. $m(\cdot)$ has bounded second partial derivatives on C_f. A4. $K(\cdot)$ is a product kernel of $k(\cdot)$ that is a continuous density with compact support on $\mathbb{R}$. All odd order moments of k vanish. A5. As $n \to \infty$, $\|H_0\| \to 0$, $\|H\| \to 0$, $n|H_0| \to \infty$, $n|H|^2 \to \infty$, $\|H\|^4 |H|^{-1} \to 0$, and $n|H|\ \|H\|^4 \to c \in [0, \infty)$, where $\|H\| = \{\mathrm{tr}(H'H)\}^{1/2}$.

To study the asymptotic property of $\widehat{M}_H(z)$, let $\bar{f}(z) = \sum_{t=1}^T f_t(z)$. Let $\tilde{p}_d$ and $\tilde{p}_o$ be the diagonal and the off-diagonal elements of $P^{1/2}$, respectively. Define $c_1 = \tilde{p}d \sum_{j=1}^q h_{0j}^2 \partial^2 m(z)/\partial z_j^2$, and $c_2 = \tilde{p}_0 \bar{f}(z)^{-1} \sum_{t=1}^T f_t(z) \sum_{s \neq t}^T \sum_{j=1}^q h_{0j}^2 E[\partial^2 m(z_{1s})/\partial z_j^2 | z_{1t} = z]$. Let $\gamma_{ij} = \int_{\mathbb{R}} u^i k(u)^j du$. Then by Assumption A4, $\int_{\mathbb{R}^q} uu'K(u)du = \gamma_{21} I_q$, $\int_{\mathbb{R}^q} K(u)^2 du = \gamma_{02}^q$, and $\int_{\mathbb{R}^q} uu'K(u)^2 du = \gamma_{02}^{q-1}\gamma_{22} I_q$.

Theorem 3.1. *(i) Under Assumptions A1–A5,*

$$\sqrt{n|H|}\left(\widehat{M}_H(z) - M(z) - \left(\frac{\gamma_{21}}{2}\sum_{j=1}^q \frac{\partial^2 m(z)}{\partial z_j^2} h_j^2 0\right) - B_1 - B_2\right) \xrightarrow{d} N(0, \tau^2 Q^{-1}\Gamma Q^{-1}), \tag{3.1}$$

where $B_1 = -\frac{\gamma_{21}(\tau\tilde{p}_d - 1)}{2}\left(\sum_{j=1}^q h_{0j}^2 \partial^2 m(z)/\partial z_j^2, 0'\right)'$, $B_2 = -(\gamma_{21} c_2 \tau/2, 0')'$,

$$Q = Q(z) = \bar{f}(z)\begin{pmatrix} 1 & 0' \\ 0 & \gamma_{21} I_q \end{pmatrix}, \text{ and } \Gamma = \bar{f}(z)\begin{pmatrix} \gamma_{02}^q & 0' \\ 0 & \gamma_{02}^{q-1}\gamma_{22} I_q \end{pmatrix}.$$

(ii) The optimal τ to minimize the mean squared error of $\widehat{m}_H(z)$ is given by

$$\tau^* = (c_1 + c_2)\sum_{j=1}^q \frac{\partial^2 m(z)}{\partial z_j^2}\left(h_{0j}^2 + h_j^2\right)\left\{(c_1 + c_2)^2 + 4n^{-1}|H|^{-1}\gamma_{02}^q \bar{f}^{-1}(z)\gamma_{21}^{-2}\right\}^{-1}. \tag{3.2}$$

(iii) If $\|H_0\| \ll \|H\|$, then B_1 and B_2 are asymptotically negligible in Eq. (3.1).

The proof is given in the Appendix. We make two remarks. (a) Theorem 3.1(i) says that the estimator for $m(z)$ is asymptotically independent of the estimator for $\dot{m}(z)$, and they have different rates of convergence. (b) As Theorem 3.1(iii) says, if we use undersmoothing in the first step, the asymptotic bias of $\widehat{M}_H(z)$ is equal to that of the conventional LLLS estimator $\widetilde{M}_H(z)$. Nevertheless, noticing that $\text{var}\{\sqrt{n|H|}\widetilde{M}_H(z)\} = (\sigma_u^2 + \sigma_\upsilon^2)Q^{-1}\Gamma Q^{-1}$, the asymptotic variance of $\widetilde{M}_H(z)$ is smaller than that of $\widetilde{M}_H(z)$ as long as $\tau^2 < \sigma_u^2 + \sigma_\upsilon^2$. This means that the 2-step estimator is asymptotically more efficient than the LLLS estimator. For example, if one sets $\tau = \tilde{p}_d^{-1}$ as in Martins-Filho and Yao (2005), then $\tau^2 = \sigma_\upsilon^2[1 - \{1-(1-d_T)^{1/2}\}/T]^{-2} < \sigma_u^2 + \sigma_\upsilon^2$, where $d_T = T\sigma_u^2/(\sigma_\upsilon^2 + T\sigma_u^2)$, and B_1 vanishes.

Since $\widehat{M}_H(z)$ depends on the unknown parameters $(\sigma_u^2, \sigma_\upsilon^2)$, it is infeasible. To provide a feasible estimator of $M(z)$, we need to estimate $(\sigma_u^2, \sigma_\upsilon^2)$ consistently. This is done in Ruckstuhl et al. (2000) and Martins-Filho and Yao (2005). Denote the resultant feasible estimator as $M_H^+(z) = (m^+(z), (H\dot{m}^+(z))')'$. It is not difficult to show that $M_H^+(z)$ has the asymptotic property as $\widehat{M}_H(z)$.

4. Monte Carlo simulations

To illustrate the finite sample performance of $M_H^+(z)$ in comparison with the conventional LLLS estimator $\widehat{M}_H(z)$, we consider the following data generating process:

$$y_{it} = m(z_{it}) + u_i + v_{it}, i = 1, \ldots .n, t = 1\ldots, T, \tag{4.1}$$

where $z_{it} = z_{it}^+ 1(|z_{it}^+| \le 3)$, z_{it}^+, u_i and v_{it} are independent i.i.d. $N(0, 1)$ sequences, $m(z) = z - 0.5z^2$. We choose H by the least-squares cross validation and set $H_0 = H^{5/4}$. We use $K(u) = 0.75(1-u^2)1(|u| \le 1)$. The consistent estimators for σ_u^2 and σ_υ^2 are obtained by using the formula in Ruckstuhl et al. (2000 p. 61).

For each estimator, we use $(n, T) = (100, 4)$ and consider 5000 repetitions. To save time, we only consider the estimation of $m(z)$ and $\dot{m}(z)$ at $z = 0.5$. Fig. 1 plots the estimated bias, standard error and root

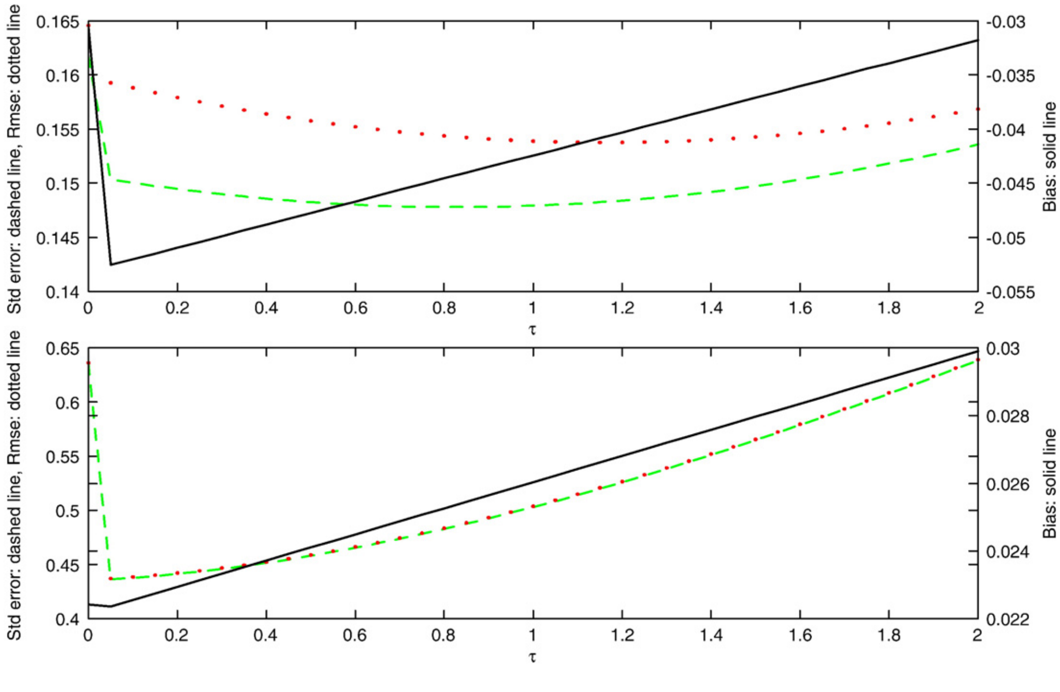

Fig. 1. Bias, standard deviation and root mean squared error (upper plot: $m(0.5)$, lower plot: $m?(0.5)$).

mean square error (Rmse) (averaged over repetitions) for our two-step estimators along with $\tau \in \{0, 0.05, 0.1, \ldots, 2\}$. $\tau=0$ corresponds to the conventional LLLS case. We find that: (a) As τ increases from 0.05, the bias also decreases in absolute value for $m_H^+(z)$ and increases for $\dot{m}_H^+(z)$. (b) The standard deviation of $m_H^+(z)$ only increases proportionally to τ^2 as $\tau>0.8$ whereas that of $\dot{m}_H^+(z)$ increases as τ increases. One reasonable explanation is that for small value of τ, the asymptotically negligible term in the variance expansion of $m_H^+(z)$ also plays a significant role in finite samples. (c) When $\tau=0.05$, the Rmse for our estimator of $\dot{m}(z)$ is smallest whereas for intermediate value of τ (around 1), the Rmse for our estimator of $m(z)$ is smallest. (d) For both estimators, we can achieve reduction in Rmse for a large range of $\tau \in (0, 2)$.

Appendix A

Proof of Theorem 3.1. By construction, $Y_i^* = m(Z_i) - (\tau P^{1/2} - I_T)(\tilde{m}_{H_0}(Z_i) - m(Z_i)) + \tau P^{1/2}\varepsilon_i$. It is standard to show that $\tilde{m}_{H_0}(z) \approx m(z) + \frac{\gamma_{21}}{2}\sum_{j=1}^{q} h_{0j}^2 \partial^2 m(z_{it})/\partial z_j^2 + n^{-1} e' Q^{-1}(z) \sum_{j=1}^{n}\sum_{s=1}^{T} \vec{Z}_{js}(z)\mathbf{K}_{H_0}(z_{js}-z)\varepsilon_{js}$ uniformly in z, where $a \approx b$ denote $a=b(1+o_p(1))$. Thus $Y_I^* \approx m(Z_i) - \tau P^{1/2} - I_T)(A_{1i}+A_{2i}) + \tau P^{1/2}\varepsilon_i$, where for $l=1,\ 2, A_{li} \equiv (a_l\ i1, \ldots, a_{liT})'$, $a_{1it} = \frac{\gamma_{21}}{2}\sum_{j=1}^{q} h_{0j}^2 \partial^2 m(z_{it})/\partial z_j^2$, and $a_{2it} = n^{-1} e' Q^{-1}(z_{it}) \sum_{j=1}^{n}\sum_{s=1}^{T} \vec{Z}_{js}(z) K_{H_0}(z_{js}-z_{it})\varepsilon_{js}$. By Taylor expression, $m(z_{it}) = \vec{Z}_{it}(z)' M(z) + \frac{1}{2}(z_{it}-z)\ddot{m}(z)(z_{it}-z) + o_P(\|H\|^2)$. Noticing $[n^{-1}\vec{Z}(z)'\mathbf{K}_H(z)\vec{Z}(z)]^{-1} \approx Q^{-1}$, we have $\sqrt{n|H|}(\widehat{M}_H(z) - M(z)) \approx B_{11} - B_{12} - B_{13} + B_{14}$, where

$$B_{11} = \frac{\sqrt{n|H|}}{2} n^{-1} Q^{-1} \sum_{i=1}^{n}\sum_{t=1}^{T} \vec{Z}_{it}(z) K_H(z_{it}-z)(z_{it}-z)\ddot{m}(z)(z_{it}-z),$$

$$B_{12} = \sqrt{n|H|} n^{-1} Q^{-1} \sum_{i=1}^{n} \left(\vec{Z}_{i1}(z)K_H(z_{i1}-z), \ldots, \vec{Z}_{iT}(z)K_H(z_{iT}-z)\right)\left(\tau P^{1/2} - I_T\right)A_{1i},$$

$$B_{13} = \sqrt{n|H|} n^{-1} Q^{-1} \sum_{i=1}^{n} \left(\vec{Z}_{i1}(z)K_H(z_{i1}-z), \ldots, \vec{Z}_{iT}(z)K_H(z_{iT}-z)\right)\left(\tau P^{1/2} - I_T\right)A_{2i},$$

$$B_{14} = \sqrt{n|H|} n^{-1} Q^{-1} \sum_{i=1}^{n} \left(\vec{Z}_{i1}(z)K_H(z_{i1}-z), \ldots, \vec{Z}_{iT}(z)K_H(z_{iT}-z)\right)\tau P^{1/2}\varepsilon_i$$

It is easy to show $B_{11} \approx (\sqrt{n|H|}/2)(\gamma_{21}\sum_{j=1}^{q} h_j^2 \partial^2 m(z_{it})/\partial z_j^2, 0')'$ and $B_{14} \xrightarrow{d} N(0, \tau^2 Q^{-1}\Gamma Q^{-1})$. Noting that B_{13} can be written as a second order U statistic which can easily be symmetrized. By direct calculation, $B_{13} = O_p(\sqrt{n|H|} n^{-1}|H_0|^{-1/2}|H|^{-1/2}) = o_p(1)$.

$$\begin{aligned} B_{12} &= \frac{\sqrt{n|H|}\gamma_{21}(\tau\tilde{p}_d - 1)}{2} n^{-1} Q^{-1} \sum_{i=1}^{n}\sum_{t=1}^{T} \vec{Z}_{it}(z) K_H(z_{it}-z) \sum_{j=1}^{q} \frac{\partial^2 m(z_{it})}{\partial z_j^2} h_{0j}^2 \\ &\quad + \frac{\sqrt{n|H|}\gamma_{21}\tau\tilde{p}_o}{2} n^{-1} Q^{-1} \sum_{i=1}^{n}\sum_{t=1}^{T} \vec{Z}_{it}(z) K_H(z_{it}-z) \sum_{s\neq t}^{T}\sum_{j=1}^{q} \frac{\partial^2 m(z_{is})}{\partial z_j^2} h_{0j}^2 \\ &\approx -\sqrt{n|H|}(B_1 + B_2). \end{aligned}$$

This completes the proof of part (i). The proofs of (ii) and (iii) are obvious and thus omitted. □

References

Henderson, D.J., Ullah, A., 2005. A nonparametric random effects estimator. Economics Letters 88, 403–407.

Li, Q., Stengos, T., 1996. Semiparametric estimation of partially linear panel data models. Journal of Econometrics 71, 389–397.

Li, Q., Ullah, A., 1998. Estimating partially linear panel data models with one-way error components. Econometric Reviews 17, 145–166.

Lin, X., 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.

Martins-Filho, C., Yao, F., 2005. Nonparametric Regression Estimation with General Parametric Error Covariance, Mimeo, Dept. of Economics, Oregon State University.

Ruckstuhl, A.F., Welsh, A.H., Carroll, R.J., 2000. Nonparametric function estimation of the relationship between two repeatedly measured variables. Statistica Sinica 10, 51–71.