Econometric Reviews

ISSN: 0747-4938 (Print) 1532-4168 (Online) Journal homepage: http://www.tandfonline.com/loi/lecr20

Local GMM Estimation of Semiparametric Panel Data with Smooth Coefficient Models

Kien C. Tran & Efthymios G. Tsionas

To cite this article: Kien C. Tran & Efthymios G. Tsionas (2009) Local GMM Estimation of Semiparametric Panel Data with Smooth Coefficient Models, Econometric Reviews, 29:1, 39-61, DOI: 10.1080/07474930903327856

To link to this article: https://doi.org/10.1080/07474930903327856

Published online: 10 Nov 2009.

Submit your article to this journal

Article views: 213

View related articles

Citing articles: 12 View citing articles

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=lecr20

Econometric Reviews, 29(1):39–61, 2010
Copyright © Taylor & Francis Group, LLC
ISSN: 0747-4938 print/1532-4168 online
DOI: 10.1080/07474930903327856

LOCAL GMM ESTIMATION OF SEMIPARAMETRIC PANEL DATA WITH SMOOTH COEFFICIENT MODELS

Kien C. Tran[1] and Efthymios G. Tsionas[2]

[1]*Department of Economics, University of Lethbridge, Lethbridge, Alberta, Canada*
[2]*Department of Economics, Athens University of Economics and Business, Athens, Greece*

□ *In this article, we consider the estimation of semiparametric panel data smooth coefficient models. We propose a class of local generalized method of moments (LGMM) estimators that are simple and easy to implement in practice. We show that the proposed LGMM estimators are consistent and asymptotically normal. Monte Carlo simulations suggest that our proposed estimator performs quite well in finite samples. An empirical application using a large panel of U.K. firms is also presented.*

Keywords Local Generalized Method of Moments; Monte Carlo simulation; Semiparametric panel data model; Smooth coefficient.

JEL Classification C13; C14; C33.

1. INTRODUCTION

The choice of a regression functional form is very important for economic analysis. Economic theory rarely provides a specific functional form for the regression relationship. Thus, unless the functional form is correctly specified, the performance of a model will generally be poor. Accordingly, it seems more desirable to work with the nonparametric/semiparametric regression models to avoid the misspecification of the regression functional form. One of the advantages of the nonparametric/semiparametric method is that little prior restriction is imposed on the model's structure, and it may provide useful guidance for the construction of parametric models.

Address correspondence to Kien C. Tran, Department of Economics, University of Lethbridge, 4401 University Drive, Lethbridge, Alberta T1K 3M4, Canada; E-mail: kien.tran@uleth.ca

In the context of panel data models, there has been a recent focus on semiparametric panel data models or "partial linear" panel data models. For example, Horowitz and Markatou (1996), Ullah and Roy (1998), Li and Hsiao (1998), Ullah and Mundra (1999), Knieser and Li (2002) considered semiparametric panel data models with exgoneous regressors, while Li and Stengos (1996), Li and Ullah (1998), Papalia (1999), Baltagi and Li (2002) considered semiparametric panel data models with endogenous regressors. However, all the above-mentioned works assumed that the regression coefficients on the parametric part are constant over time and across individuals. In practice, it is possible that, in one set of data, this assumption produces more reasonable empirical results, but in another data set, allowing for the coefficients to vary across individuals, or over time, or both on the parametric part, may lead to more plausible empirical results. Thus, an important extension to the semiparametric panel data models is to allow for the regression coefficients on the parametric part to vary according to the smooth coefficient model. The smooth coefficient model lets the marginal effect of a given variable be an unknown function of an observable covariate, and hence, introduces heterogeneity into the marginal effect. This specification also nests the traditional linear model as a special case when the marginal effect is found to be constant over the support of the observable covariate. Smooth coefficient models have received a lot of attention in the statistics/econometrics literature recently and have been used in various applications; see for example, Hastie and Tibshirani (1993), Carroll et al. (1998), Gozalo and Linton (2000), Fan and Zhang (1999), Cai et al. (2000), Das (2005), Cai et al. (2006), just to name a few. However, most of the works are focused on models with exogenous regressors, and very little attention has been paid to the case of semiparametric panel data with endogenous regressors.[1]

The purpose of this article is to extend the semiparametric panel data models with endogenous regressors to allow for the slope coefficient heterogeneity in the parametric part of the model by allowing it to have a smooth coefficient form. We propose a consistent two-step local generalized method of moments estimator to estimate the smooth coefficients, and establish its asymptotic properties. After the initial revised and resubmission of our article, an associate editor and a referee directed our attention to the recent work of Cai and Li (2008). In that

[1]Das (2005) and Cai et al. (2006) considered cross-section varying–coefficient models with endogenous regressors whereas Carroll et al. (1998) and Gozalo and Linton (2000) considered models with exogenous regressors. Also, it is worth to point out that Carroll et al. (1998) method for estimating varying coefficient is based on M-estimation procedure using the local first-order conditions; while Gozalo and Linton (2000) method is initially parameterizing an unknown regression function $g(x)$, and then using local polynomial fitting to recover the $g(x)$ and its derivatives.

article, they suggest a nonparametric generalized method of moments (NPGMM) approach to estimate a varying coefficient panel data model with endogenous regressors. Their approach is based on a local linear fitting, and they consider the case where both $N, T \to \infty$. Although their approach has some overlap with our proposed method, there are some important differences. First, they only consider a one-step NPGMM with a weighting matrix that is the identity matrix, while we allow for more general weighting matrix and consider a two-step (and/or iterative two-step) estimation approach. Consequently, the advantage of our method over Cai and Li's (2008) approach is a potential asymptotic efficiency gain from the use of such two-step estimator. Second, their article contains neither simulations nor an empirical application, while we provide some Monte Carlo simulations to examine the finite sample performance, and an empirical application to illustrate the usefulness of the proposed method.

The article is organized as follows. Section 2 introduces the semiparametric panel model with smooth coefficients. Section 3 derives the local generalized method of moments (GMM) estimators and establishes the asymptotic properties of the proposed estimators. Monte Carlo simulations are presented in Section 4. Section 5 provides an empirical illustration. Section 6 briefly discusses the fixed effects model. Concluding remarks are given in Section 7. Proof of the theorem is given in the Appendix.

2. THE MODEL

We consider the following semiparametric panel data with varying–coefficient model:

$$y_{it} = x_{it}'\delta(z_{it}) + u_{it}, \qquad (i = 1, \dots, N; t = 1, \dots, T), \tag{1}$$

where the prime denotes the transpose of a matrix or vector, x_{it} is of dimension $(p \times 1)$ with its first element $x_{it,1} = 1$, z_{it} is of dimension $(q \times 1)$ which does not contain a constant, $\delta(z_{it})$ is a $(p \times 1)$ vector of unknown and unspecified smooth functions, and u_{it} is the usual random disturbance. We allow some or all components of x_{it} to be correlated with the error u_{it}. We assume the data are independent across the i index but there is no restriction on the time index t, and $E(u_{it} \mid z_{it}) = 0$. We consider the common empirical case of N large and small T.

We also allow the possibility that the error term u_{it} is serially correlated. For example, when u_{it} follows a one-way error component specification, $u_{it} = \mu_i + \varepsilon_{it}$, where μ_i a random individual specific effect with $\mu_i \sim i.i.d.(0, \sigma_\mu^2)$ and $\varepsilon_{it} \sim i.i.d.(0, \sigma_\varepsilon^2)$, which render the errors serially correlated.

There are several interesting features of model (1) worth mentioning. First, (1) is an extension and generalization of the cross-section model considered by Li et al. (2002). Second, when $x_{1,it} = 1$ and $\delta_j(z_{it}) = \delta_0$, $j = 2, \ldots, p$, model (1) reduces to the semiparametric panel model of Baltagi and Li (2002) and Li and Stengos (1996). Third, model (1) covers semiparametric IV models considered by Das (2005) for discrete endogenous regressors and Cai et al. (2006) for both discrete and continuous endogenous regressors. Finally, if there are no endogenous variables and the coefficients $\delta_j(\cdot)$, $j = 2, \ldots, p$ are threshold functions such as

$$\delta_j(z) = \delta_{j1} I(z \leq \alpha_j) + \delta_{j2} I(z > \alpha_j),$$

then model (1) may describe a threshold nondynamic panel regression model of Hansen (1999). Thus, model (1) includes some interesting special cases that arise commonly in empirical research.

3. LOCAL GENERALIZED METHOD OF MOMENTS ESTIMATION

By stacking all T observations for the ith individual, (1) can be written as

$$y_i = X_i \delta(Z_i) + u_i, \qquad i = 1, \ldots, N, \tag{2}$$

where $y_i = (y_{i1}, \ldots, y_{iT})'$ is a $(T \times 1)$ vector, X_i is a $(T \times p)$ matrix having rows x_{it}', $t = 1, \ldots, T, \delta(Z_i) = (\delta(z_{i1}), \ldots, \delta(z_{iT}))'$, Z_i is a $(T \times q)$ matrix having rows $z_{it}', t = 1, \ldots, T$, and $u_i = (u_{i1}, \ldots, u_{iT})'$ is a $(T \times 1)$ vector. Assume that there exists a $(T \times l)$ (where $l \geq p$) matrix of instruments W_i having rows w_{it}' (where the first component of $w_{it}, w_{it,1} = 1$), $t = 1, \ldots, T$, such that (X_i, Z_i, W_i, u_i) are *iid* over $i = 1, \ldots, N$ and

$$E(w_{it} u_{it} \mid z_{it}) = 0, \qquad t = 1, \ldots, T,$$

which implies, for a given $z_{it} = z$

$$E(W_i' u_i \mid Z_i = \tau_T z') = \sum_{t=1}^{T} E(w_{it} u_{it} \mid z_{it} = z) = 0, \tag{3}$$

where $\tau_T = (1, \ldots, 1)'$ is a $(T \times 1)$ vector of ones. Thus, Eq. (3) provides the moment conditions that form the basis for identification and our estimation below. In the cross-section context, Cai et al. (2006) provided the conditions for which $\delta(\cdot)$ is identified (up to an additive constant) and proposed a two-stage estimation method. However, their identification conditions are not directly applicable in our case since our estimation method will be based on the conditional moment restrictions in (3). Furthermore, their proposed estimation method will require a two-step nonparametric estimation procedure which will complicate the asymptotic analysis of the resulting estimator. To avoid these shortcomings, we suggest a simple estimation which will require only one nonparametric estimation procedure.

To obtain identification conditions for our case, note that from (3) fixing $Z_i = \tau_T z'$ and for any $\delta_1(z)$, we have[2]

$$\begin{aligned} E\big[W_i'(Y_i - X_i\delta_1(z)) \mid Z_i = \tau_T z'\big] &= E(W_i' u_i \mid Z_i = \tau_T z') \\ &\quad + E\big[W_i' X_i(\delta(z) - \delta_1(z)) \mid Z_i = \tau_T z'\big] \\ &= E(W_i' X_i \mid Z_i = \tau_T z')(\delta(z) - \delta_1(z)). \end{aligned}$$

Thus, the necessary and sufficient condition for identify $\delta(z)$ is that $E(W_i' X_i \mid Z_i = \tau_T z')$ has full column rank since $E(W_i' X_i \mid Z_i = \tau_T z')(\delta(z) - \delta_1(z)) = 0$ if and only if $\delta(z) = \delta_1(z)$.

For the remaining part of the article, we assume that the vector $\delta(\cdot)$ is identified and twice continuously differentiable. Then for a given point $z \in \Re^q$ and for $\{z_{it}\}$ in the neighborhood of z, Eq. (3) provides the conditional moment restrictions that can be used to construct an estimator similar to the GMM of Hansen (1982) for parametric models. Thus the local-GMM (LGMM) criterion function is

$$J_N(\delta) = \left[\frac{1}{N}(Y - X\delta(z))' K W\right] R_N^{-1} \left[\frac{1}{N} W' K (Y - X\delta(z))\right], \tag{4}$$

where $Y = (Y_1', \ldots, Y_N')'$ is a $(NT \times 1)$ vector, $W = (W_1', \ldots, W_N')'$ is a $(NT \times l)$ matrix of instruments, $X = (X_1', \ldots, X_N')'$ is a $(NT \times p)$ matrix of regressors, R_N is some $(l \times l)$ positive definite weighting matrix, and K is an $(NT \times NT)$ matrix of kernel weights with $K = \operatorname{diag}\{K_1^T, \ldots, K_N^T\}$, $K_i^T = \operatorname{diag}(K_H(z_{it} - z)), t = 1, \ldots, T,$ is a $(T \times T)$ matrix, and $K_H(\xi) = \prod_{j=1}^q h_j^{-1} k(\xi_j / h_j)$, in which $k(\varphi) \geq 0$, is a bounded univariate symmetric function with $\int k(\varphi) d\varphi = 1$, $\int \varphi^2 k(\varphi) d\varphi = \omega > 0$, $\int k^2(\varphi) d\varphi = \upsilon > 0$, so that $K(\xi) = \prod_{j=1}^q k(\xi_j)$, and $H = \operatorname{diag}\{h_1, \ldots, h_q\}$ is a $(q \times q)$ matrix of bandwidths with $|H| = \prod_{j=1}^q h_j$, and $h_j > 0$.

For a given $z_{it} = z$, minimizing (5) with respect to δ, we obtain

$$\hat{\delta}(z) = \left[X' K W R_N^{-1} W' K X\right]^{-1} X' K W R_N^{-1} W' K Y. \tag{5}$$

The estimator given in (5) is termed a LGMM estimator. It is consistent because $E(w_{it} u_{it} \mid z_{it} = z) = 0$. However, to implement (5) one needs to

[2]We owe this observation to an anonymous referee.

specify the weighting matrix R_N. Different full-rank weighting matrices R_N lead to different local GMM estimators, except in the just-identified case where $l = p$. The two leading choices are given below.

(1) One-Step LGMM Estimator:
Under *i.i.d.* assumption, the one-step LGMM estimator uses weighting matrix.[3]

$$R_{1N} = R(z) \equiv E\left(W_1' K_1^T W_1\right) = f(z) E(W_1' W_1 \mid Z_1 = \tau_T z') = O(1),$$

leading to

$$\hat{\delta}_{OS}(z) = \left[X'KW(W'KW)^{-1}W'KX\right]^{-1} X'KW(W'KW)^{-1}W'KY, \tag{6}$$

where in (6) we have replaced R_{1N} by its consistent estimator, $\widehat{R}_{1N} = N^{-1}\sum_{i=1}^{N} W_i' K_i^T W_i$. The motivation for this estimator is that it can be shown to be the optimal LGMM estimator based on the conditional moment restrictions (3) if $u_i \mid z_i$ is $iid(0, \sigma^2 I_T)$. Also, it is interesting to note that the one-step local GMM estimator given in (6) is numerically equivalent to the two-stage smooth coefficient least squares estimator (Li et al., 2002) where in the first stage, a smooth coefficient least squares regression of X on W, yielding prediction $\widehat{X}$, and in the second stage, a smooth coefficient least squares regression of Y on $\widehat{X}$ using the same kernel K and bandwidth H.

(2) Two-Step Local GMM Estimator:
Under *i.i.d.* assumption over i and stationarity assumption over t, the two-step LGMM estimator uses the weighting matrix (see Appendix for derivation)

$$\begin{aligned}
R_{2N} = S(z) &\equiv \lim_{N\to\infty} \operatorname{var}\left\{\sqrt{N|H|}\left(W_1' K_1^T u_1\right)\right\} \\
&= f(z) \int K^2(z)\,dz' E\left(W_1' V_1 W_1 \mid Z_1 = \tau_T z'\right) \\
&= f(z) \int K^2(z)\,dz' E\left(\sum_{t=1}^{T} u_{1t}^2 w_{1t} w_{1t}' \mid z_{1t} = z\right) \\
&= \Omega_0(z) f(z) \int K^2(z)\,dz',
\end{aligned}$$

where $V_1 = \operatorname{diag}\left(u_{11}^2, \ldots, u_{1T}^2\right)$ and $\Omega_0 = E\left(W_1' V_1 W_1 \mid Z_1 = \tau_T z'\right) = E\left(\sum_{t=1}^{T} u_{1t}^2 w_{1t} w_{1t}' \mid z_{1t} = z\right)$. Let $A_i^* = \left(K_i^T\right)^{1/2} A_i$ and $A^* = \left(A_1^*, \ldots, A_N^*\right)'$, where

[3] We use the subscript 1 to signify "typical i."

$\left(K_i^T\right)^{1/2} = \mathrm{diag}\left(\sqrt{K_H(z_{it} - z)}\right)$, $t = 1, \ldots, T$, and $A_i = X_i, W_i$, or Y_i, then by standard arguments of kernel smoothing, $S(z)$ can be consistently estimated by

$$\widehat{S}_N(z) = f(z) \int K^2(z)\,dz' \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{u}_{1t}^2 w_{1t} w_{1t}' K_H(z_{it} - z)$$
$$= f(z) \int K^2(z)\,dz' \left\{ N^{-1} \sum_{i=1}^{N} W_i^{*'} \widehat{V}_i W_i^* \right\},$$

where $\hat{u}_{it} = y_{it} - x_{it}' \hat{\delta}_{OS}(z)$.

Thus, by replacing R_N in (5) with $\widehat{S}_N(z)$, we obtain the following two-step LGMM estimator

$$\hat{\delta}_{TS}(z) = \left[X^{*'} W^* \left(W^{*'} \widehat{V} W^*\right)^{-1} W^{*'} X^*\right]^{-1} X^{*'} W^* \left(W^{*'} \widehat{V} W^*\right)^{-1} W^{*'} Y^*, \qquad (7)$$

where $\widehat{V} = \mathrm{diag}\left(\widehat{V}_1, \ldots, \widehat{V}_N\right)$ and $\widehat{V}_i = \mathrm{diag}\left(\hat{u}_{i1}^2, \ldots, \hat{u}_{iT}^2\right)$ is a consistent estimator of V_i. We call this a two-step LGMM estimator because a first-step consistent estimator of $\delta(z)$ such as $\hat{\delta}_{OS}(z)$ is needed to form the residuals $\hat{u}$ used to compute $\widehat{V}$.

The two-step LGMM estimator may be iterated by recomputing the residuals after computing (7) and then reentering the computation. The asymptotic properties of the iterated estimator are the same as those of two-step estimator which will be discussed next.

Asymptotic Properties

First, we introduce some additional notation. Let $\kappa_2 = \int \varphi^2 K(\varphi) d\varphi$, $z_0 = \tau_T z'$, $\delta_j(z) = \partial\delta(z)/\partial z_j$, and $\delta_{jj}(z) = \partial^2\delta(z)/\partial z_j^2$ be $(q \times 1)$ vectors. In addition, $\mathscr{G}_v^\mu$ denote the class of functions such that if $f \in \mathscr{G}_v^\mu$, then f is v times continuously differentiable, and its derivatives up to order v are all bounded by some function that has μth order finite moments. We make the following assumptions:

(A.1): (i) For each fixed t, $\{(y_{it}, x_{it}, w_{it}, z_{it}, u_{it})\}$ are *iid* in the i subscript and strictly stationary over t for each fixed i. Let $f(z)$ be the marginal density function of z_{it}, and let $A(z) = f(z)E(W_i' X_i \mid Z_i = z_0)$. $f(z) \in \mathscr{G}_{v-1}^\infty$, $\delta(z) \in \mathscr{G}_v^{4+\alpha}$, and $A(z) \in \mathscr{G}_{v-1}^{4+\alpha}$, for some $\alpha > 0$ and some positive integer $v \geq 2$. $\delta(z), f(z)$, $A(z)$, and $f(z)\Omega_0(z)$, all satisfy some Lipschitz-type of conditions in z.

(ii) For each $t \geq 1$, let $\Omega_r(z_1, z_2) = E\left(u_{1t} w_{1t} u_{1t-r} w_{1t-r} \mid z_{1t} = z_1, z_{1t-r} = z_2\right)$ and $g_r(z_1, z_2)$, the joint density of (z_{1t}, z_{1t-r}) is continuous at $\big(z_{1t} = z_1,$

$z_{1t-r} = z_2\big)$. In addition, $\sup_{t\geq 1}\left|\Omega_r(z_1, z_2)g(z_1, z_2)\right| \leq C(z) < \infty$ for some function $C(z)$.

(A.2): There exists a $(T \times l)$ (where $l \geq p$) matrix of instruments W_i having rows $w'_{it}, t = 1, \ldots, T$ such that $E(W_i' u_i \mid Z_i = z_0) = 0$ and $E[W_i' X_i \mid Z_i = z_0]$ is of full column rank for all z.

(A.3): $E|u_i|^2 < \infty$, $E\|W_i' X_i\|^2 < \infty$ and $E\|W_i' W_i\|^2 < \infty$.

(A.4): $k(\cdot) \geq 0$ is a bounded symmetric function with $\int k(\varphi)d\varphi = 1$, $\int \varphi^2 k(\varphi)d\varphi = \omega > 0$, $\int k^2(\varphi)d\varphi = \upsilon > 0$, and $\int K(\xi)d\xi' = 1$. As $N \to \infty$, $\sqrt{N|H|} \to \infty$ and $h_j \to 0$.

Assumption (A.1) requires that observations are *iid* across i and stationary across t which is a standard assumption in the panel data literature. It also requires that $\{z_{it}\}$ has a common distribution over t, and gives some smoothness conditions on functionals involved. (A.2) provides the necessary and sufficient condition for model identification. (A.3) gives some standard moment conditions. (A.4) provides conditions on a kernel function and smoothing parameter.

The following theorem establishes the consistency and asymptotic normality of $\hat{\delta}_{TS}(z)$ given in (7).

Theorem 1. *Under the assumptions* (A.1)–(A.4), *we have*

$$\text{(a)}\quad \hat{\delta}_{TS}(z) - \delta(z) - \Psi(z)\sum_{j=1}^{q} h_j^2 B_j(z) = o_p\left(\sum_{j=1}^{q} h_j^2 + (N|H|)^{-1/2}\right)$$

$$\text{(b)}\quad \sqrt{N|H|}\left\{\hat{\delta}_{TS}(z) - \delta(z) - \Psi(z)\sum_{j=1}^{q} h_j^2 B_j(z)\right\}$$

$$\xrightarrow{d} N\big(0, \{A(z)'S(z)^{-1}A(z)\}^{-1}\big),$$

where

$$\Psi(z) = \left[A(z)'S(z)^{-1}A(z)\right]^{-1}A(z)'S(z)^{-1},$$

$$B_j(z) = (1/2)\kappa_2\left\{A(z)\delta_{jj}(z) + 2\frac{\partial A(z)}{\partial z_j}\delta_j(z)\right\},$$

$$A(z) = f(z)E\big(W_1'X_1 \mid Z_1 = z_0\big),$$

$$S(z) = T\Omega_0 f(z)\int K^2(z)dz = f(z)\int K^2(z)dz' E(W_1^{*'}V_1 W_1^{*} \mid Z_1 = z_0).$$

The proof of Theorem 1 is given in the Appendix. Note that the unknown quantities $A(z)$ and $S(z)$ can be consistently estimated by

$\widetilde{A}_N(z) = N^{-1}\sum_{i=1}^{N} W_i' X_i K_i^T(z)$ and $\widetilde{S}_N(z) = N^{-1}\int K^2(z)dz'\{\sum_{i=1}^{N} W_i^{*\prime}\widetilde{V}_i W_i^*\}$, respectively, where $\widetilde{V}_i = \mathrm{diag}(\tilde{u}_{i1}^2, \ldots, \tilde{u}_{iT}^2)$, and $\tilde{u}_i = Y_i - X_i\hat{\delta}_{TS}(z)$ is a $(T \times 1)$ estimated residual vector from the two-step LGMM estimator in (7).

Remark 1. It is interesting to note that the results of Theorem 1 also covers the results in the cross-sectional data case (e.g., $T = 1$). Furthermore, when there is no endogenous variable in the model (e.g., $w_{it} = x_{it}$), it also covers the results in Li et al. (2002).

Remark 2. To implement the estimator in (6) or (7), one needs to specify the choice of the kernel function, the smoothing parameters, and the set of instrument variables. In practice, the most commonly used kernel function is a Gaussian kernel, although any other function that satisfies the conditions in assumption (A.4) could be used. However, it is known that the choice of the kernel function is of less importance compared to the choice of the smoothing parameters. In practice, one could use $h_j = sd(z_j)N^{-1/(q+4)}$, where $sd(z_j)$ is the sample standard deviation of z_j, $j = 1, 2, \ldots, q$. Alternatively, one may use some data-driven method such as cross-validation to select h_j. As for the choice of instrument variables, by the exogenous assumption, we know that $E(u_{it} \mid z_{it}) = 0$, so that z_{it} is uncorrelated with u_{it}, and hence z_{it} or any function of z_{it} can be used as part of the set of instruments.[4] On the other hand, Newey (1990) and Baltagi and Li (2002) offer discussion on how to choose optimal instruments in the context of semiparametric panel data models, and if we restrict our attention to the case where the instruments are functions of z_{it}, we can use the results of Newey (1990) or Baltagi and Li (2002) to construct the optimal instruments. The readers are referred to these articles for more detailed.

4. MONTE CARLO SIMULATION

In this section, we report some simulation results to examine the finite sample performance of our proposed two-step LGMM estimator, and also compare it with the NPGMM estimator suggested by Cai and Li (2008). We consider the following data generating process (DGP):[5]

$$y_{it} = \gamma(z_{it})y_{i,t-1} + x_{it}\beta(z_{it}) + \mu_i + v_{it},$$

[4]Note that when x_{it} contains lagged of dependent variable, z_{it} is assumed to be weakly exogenous in the sense that $E(u_{it} \mid z_{is}) = 0$ for $s \leq t$.

[5]Part of the DGP are taken from Baltagi and Li (2002).

where

$$\gamma(z_{it}) = \exp[-(0.5z_{it} - 1.5)^2] \qquad \text{and} \qquad \beta(z_{it}) = z_{it} + \sin(z_{it}).$$

The error term v_{it} is $i.i.d.$ $N(0, \sigma_v^2)$, μ_i is $i.i.d.$ $N(0, \sigma_\mu^2)$, z_{it} is generated as by the $i.i.d.$ uniform[2,6] distribution, and $x_{it} = \zeta_{1it} + \zeta_{2it}$, where ζ_{jit}, $j = 1, 2$, are $i.i.d.$ uniform[0,2]. We fixed the total error variance $\sigma^2 = \sigma_\mu^2 + \sigma_v^2 = 1.0$, and define $\rho = \sigma_\mu^2/\sigma^2 = \{0.2, 0.5, 0.8\}$. The sample sizes are $N = \{100, 200\}$ and $T = 5$, and the number of replications is 500 for all cases.

Note that the above DPG is a special case, where the model is a dynamic one-way error component model with an exogenous regressor. The main reason why we chose this model for our simulation study is because it seems to be the most common case encountered in practice.

In our simulation, a Gaussian kernel function is used and the smoothing parameter is chosen as $h = \alpha sd(z)N^{-1/5}$, where $\alpha = 0.8, 1.0$ and 1.2. However, our results do not seem to be sensitive to the choice of α, and consequently, we set $\alpha = 1.0$ in all of our experiments. We compare the estimated mean average square error (*MASE*) of $\hat{\theta}_j(\cdot) = \left(\hat{\gamma}_j(\cdot), \hat{\beta}_j(\cdot)\right)'$ defined by

$$MASE(\hat{\theta}(\cdot)) = \frac{1}{1000}\sum_{j=1}^{1000}\left\{\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\hat{\theta}_j(z_{it}) - \theta_j(z_{it})\right)\right\}^2,$$

where $\hat{\theta}_j(\cdot)$ is the estimate of $\theta_j(\cdot) = \left(\gamma_j(\cdot),\ \beta_j(\cdot)\right)'$ from the jth replication based on one of the two methods: the two-step LGMM method or the NPGMM method.

We use two sets of instruments for each method. The first set of instruments consists of the $w_{it}^{(1)} = \left\{y_{i,t-2}, z_{it-1}, x_{it}, x_{it-1}\right\}$, and in the second set, we use the optimal instruments given by $w_{it}^{(o)} = \left\{E(y_{it-1} \mid z_{it-1}), E(y_{it-1} \mid z_{it-2}, E(y_{it-1} \mid z_{it-1}, z_{it-2}), x_{it}\right\}$ (see Baltagi and Li, 2002). Note that some of the optimal instruments in $w_{it}^{(o)}$ are not feasible because the conditional expectations involved are unknown. However, these conditional expectations can be consistently estimated using some nonparametric approach such as kernel method, series method, etc. In this article, we suggest using the density-weighted kernel smoothing approach (see Powell et al., 1989) to estimate these unknown conditional expectations.

The simulation results are presented in Table 1, where the instrument set $w_{it}^{(1)}$ is used in the estimation. From Table 1, we first see that both NPGMM and the two-step LGMM methods perform quite well in the finite sample. Second, by comparing the two methods, we observed that there are substantial efficiency gains from using the two-step LGMM method as

TABLE 1 $MASE(\hat{\theta}(\cdot))$ by NPGMM and two-step LGMM methods (*Regular*) Instrument set $= w_{it}^{(1)}$

	$\rho = 0.2$		$\rho = 0.5$		$\rho = 0.8$	
	$N = 100, T = 5$		$N = 100, T = 5$		$N = 100, T = 5$	
	$MASE(\hat{\gamma}(\cdot))$	$MASE(\hat{\beta}(\cdot))$	$MASE(\hat{\gamma}(\cdot))$	$MASE(\hat{\beta}(\cdot))$	$MASE(\hat{\gamma}(\cdot))$	$MASE(\hat{\beta}(\cdot))$
NPGMM	0.00389	0.01185	0.00651	0.01239	0.01167	0.01296
Two-step LGMM	0.00119	0.01026	0.00175	0.01108	0.00184	0.01130
	$N = 200, T = 5$		$N = 200, T = 5$		$N = 200, T = 5$	
	$MASE(\hat{\gamma}(\cdot))$	$MASE(\hat{\beta}(\cdot))$	$MASE(\hat{\gamma}(\cdot))$	$MASE(\hat{\beta}(\cdot))$	$MASE(\hat{\gamma}(\cdot))$	$MASE(\hat{\beta}(\cdot))$
NPGMM	0.00252	0.00716	0.00526	0.00745	0.01067	0.00771
Two-step LGMM	0.00092	0.00631	0.00096	0.00706	0.00107	0.00729

opposed to the NPGMM method. This is especially true for the estimated coefficient on the endogenous regressor. Also as the value of ρ increases, the efficiency gains become more pronounced. Finally, we observe that as the sample size N increases, the MASE for both methods decrease.

Table 2 reports the MASE results for the NPGMM and two-step LGMM methods when the optimal instrument set $w_{it}^{(o)}$ is used. From Table 2, we see that similar results are observed as in Table 1. Moreover, estimations using "optimal instruments" provide better performance than estimations that are based on "regular instruments" in term of MASE.

TABLE 2 $MASE(\hat{\theta}(\cdot))$ by NPGMM and two-step LGMM methods *(Optimal)* Instrument set $= w_{it}^{(o)}$

	$\rho = 0.2$		$\rho = 0.5$		$\rho = 0.8$	
	$N = 100, T = 6$		$N = 100, T = 6$		$N = 100, T = 6$	
	$MASE(\hat{\gamma}(\cdot))$	$MASE(\hat{\beta}(\cdot))$	$MASE(\hat{\gamma}(\cdot))$	$MASE(\hat{\beta}(\cdot))$	$MASE(\hat{\gamma}(\cdot))$	$MASE(\hat{\beta}(\cdot))$
NPGMM	0.00198	0.01147	0.00356	0.01204	0.00692	0.01384
Two-step LGMM	0.00101	0.01019	0.00129	0.01092	0.00162	0.01106
	$N = 200, T = 6$		$N = 200, T = 6$		$N = 200, T = 6$	
	$MASE(\hat{\gamma}(\cdot))$	$MASE(\hat{\beta}(\cdot))$	$MASE(\hat{\gamma}(\cdot))$	$MASE(\hat{\beta}(\cdot))$	$MASE(\hat{\gamma}(\cdot))$	$MASE(\hat{\beta}(\cdot))$
NPGMM	0.00108	0.00686	0.00261	0.00693	0.00289	0.00709
Two-step LGMM	0.00073	0.00527	0.00091	0.00556	0.00097	0.00585

5. EMPIRICAL APPLICATION

In this section, we apply the new techniques to data on a substantial number of U.K. manufacturing companies from 1982 to 1994 as used by Nickell (1996) and Nickell et al. (1997). These authors have investigated the role of competition in productivity and productivity growth since "[the] general belief in the efficacy of competition exists despite the fact that it is not supported either by any strong theoretical foundation or by a large corpus of hard empirical evidence in its favor" (Nickell, 1996, p. 725). The evidence provided in Nickell (1996) and Nickell et al. (1997) rests upon strong parametric assumptions, the assumption of constant returns to scale in a Cobb–Douglas production technology,[6] and the assumption that competition measures do not have a firm-specific or time-specific effect on productivity. All these assumptions are troublesome and could be responsible for biased results.

The basic model in Nickell (1996) is

$$y_{it} = \lambda y_{i.t-1} + (1-\lambda)\alpha_i n_{it} + (1-\lambda)(1-\alpha_i)k_{it} + \beta h_{it} + f_{it} + u_{it},$$

where y_{it} is log of output, n_{it} is log employment, k_{it} is capital stock, h_{it} is a business cycle component, f_{it} reflects all factors influencing the level of productivity, u_{it} is an error term, and we have omitted time-specific and firm-specific fixed effects. Nickell (1996) has estimated the model in first-differenced form using the Arellano and Bond (1991) GMM technique. The variables used in f_{it} that affect productivity or productivity growth are the following: market share, rents normalized on value-added, concentration ratio, and import penetration (imports divided by home sales). For detailed construction of these variables, see Nickell (1996) and Nickell et al. (1997).

In this article, we use an extended unbalanced panel data set of 582 companies taken from Nickell et al. (1997). We estimate the following dynamic panel data with smooth coefficients model

$$\begin{aligned} y_{it} = {} & \lambda(z_{it})y_{it-1} + (1-\lambda(z_{it}))\alpha_1(z_{it})n_{it} + (1-\lambda(z_{it}))(1-\alpha_1(z_{it}))k_{it} \\ & + \beta(z_{it})h_{it} + \delta'(z_{it})f_{it} + \mu_i + u_{it}, \end{aligned}$$

where the covariate z_{it} is taken to be logarithm of debt. Thus our empirical model suggests that the dynamic adjustment, labor, capital, and other inputs coefficients may vary directly with the firm's debt values. As a result, the returns to scale may also be a function of debt. Note that Nickell (1996)

[6]Nickell (1996) tried to deal with both assumptions in a parametric way. For example, he added a CES component to check whether the Cobb–Douglas assumption is responsible for serious differences in the results.

treats k_{it} and n_{it} as endogenous in the model, so we follow the same practice here.

We estimate the above model using the LGMM procedure given in (7). We use a standard normal kernel for $k(\cdot)$ and since z_{it} is a scalar, univariate crossvalidation bandwidth selection procedure is used to determine the optimal bandwidth. For the selection of the instruments, we use the optimal instrument discussed in Baltagi and Li (2002). Specifically, we use the density-weighted kernel estimates of $\{E(y_{it-1} \mid z_{it-1}), E(y_{it-2} \mid z_{it-2}), E(k_{it} \mid z_{it}), E(k_{it-1} \mid z_{it-1}), E(n_{it} \mid z_{it}), E(n_{it-1} \mid z_{it-1})\}$ as instrument set for $\{y_{it-1}, k_{it}, n_{it}\}$.

We present our empirical results in graphical form in Figs. 1 and 2. The descriptive statistics are reported in Table 3. In the figures, we report kernel densities of the various firm- and time-specific coefficients. The dynamic adjustment coefficient averages 0.288 (with standard deviation 0.15), and the labor and capital coefficients are 0.416 and 0.295 (with standard errors 0.16 and 0.10, approximately).

The average coefficient on $y_{i,t-1}$ is fairly close to the value reported by Nickell (1996). However, what is *not* uncovered by the results in Nickell (1996) or Nickell et al. (1997) is the fact that the distribution of this

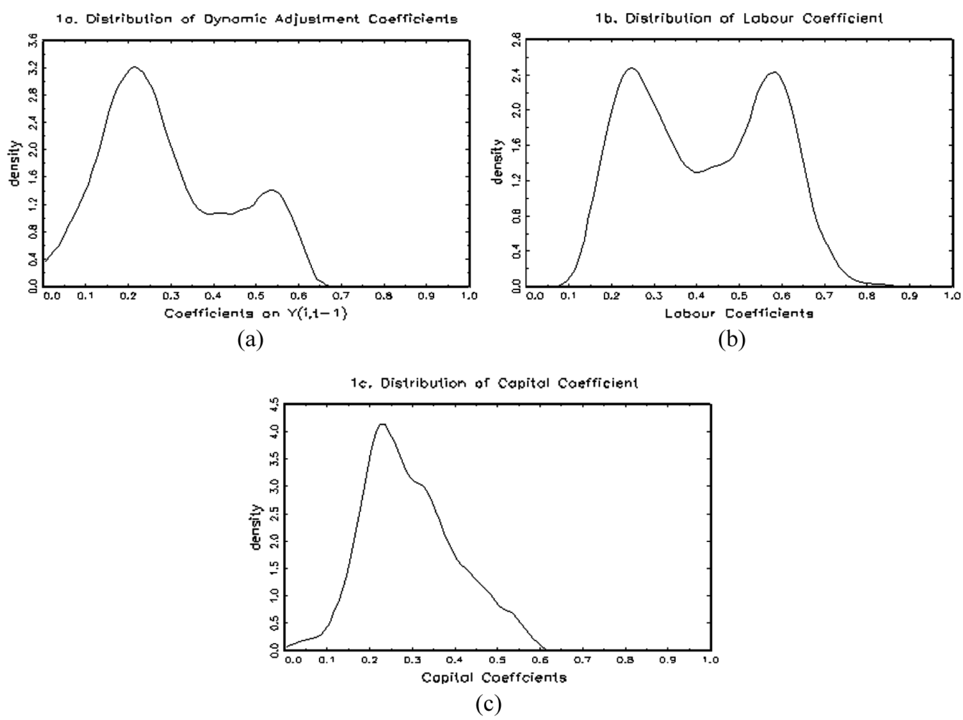

FIGURE 1 Distribution of coefficients on endogenous variables.

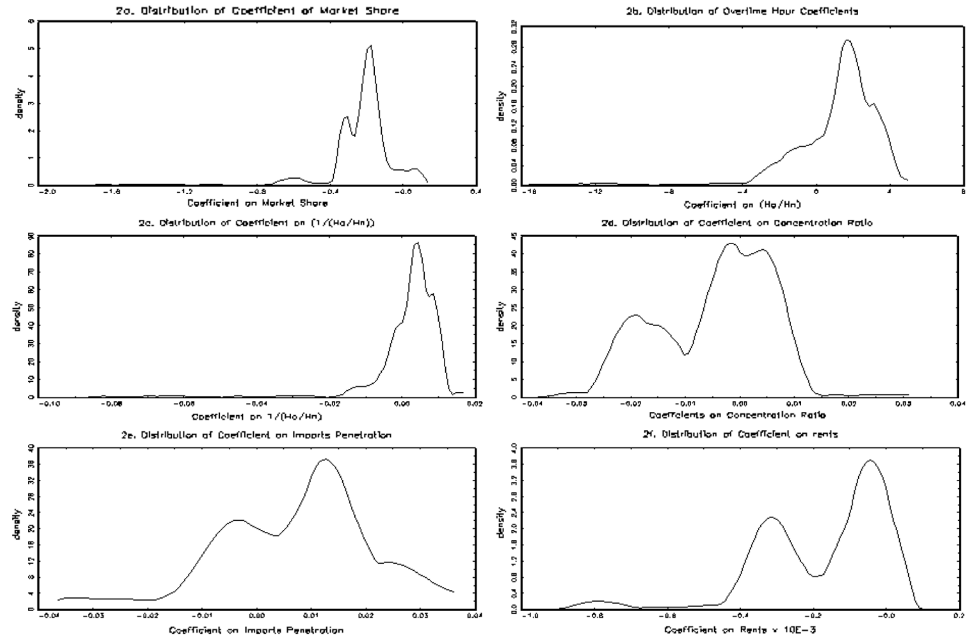

FIGURE 2 Distribution of coefficients on exogenous variables.

coefficient is bimodal with modes at about 0.2 and 0.5, the first mode being the dominant one. The maximum value of the coefficient is close to 0.6154, a value that could not have been predicted using an asymptotic normal approximation along with the parameter estimate and its standard

TABLE 3 Descriptive statistics of the estimated coefficients

Variable	Mean	Std Dev	Minimum	Maximum
Const	1.1495	0.7402	−0.4729	5.9648
$^*y_{i,t-1}$	0.2888	0.1512	0.0149	0.6154
$^*l_{i,t}$	0.4161	0.1604	0.1494	0.8055
$^*k_{i,t}$	0.2951	0.1075	0.0067	0.5676
$mktsh_{i,t-2}$	−0.2500	0.2521	−1.9611	0.1490
H_{oit}/H_{nit}	1.0187	2.6912	−15.8321	5.1019
$(H_{oit}/H_{nit})^{-1}$	0.0011	0.0122	−0.0868	0.0175
$(conc_i)t$	−0.0045	0.0105	−0.0373	0.0315
$(imp_i)t$	0.0070	0.0146	−0.0367	0.0367
$10^{-3} \times (rent_i)t$	−0.1792	0.1873	−1.4978	0.0085

Note: Total number of firms is 582, and the number of observations is 5273. The dependent variable is log(real sale). Instruments include $\{E(n_{it} \mid z_{it}), E(n_{it-1} \mid z_{it-1}), E(y_{it-1} \mid z_{it-1}), E(y_{it-2} \mid z_{it-2}), E(k_{it} \mid z_{it}), E(k_{it-1} \mid z_{it-1})\}$.
*Variables that are treated as endogenous.

error from Nickell (1996) or Nickell et al. (1997). Our average coefficients of labor and capital are also in agreement with Nickell (1996), although estimates in Nickell et al. (1997) tend to be somewhat higher.[7] Again, the distribution of the labor coefficients is bimodal (the modes are close to 0.2 and 0.6) whereas the distribution of the capital coefficient is skewed to the right. These results indicate that there is considerable heterogeneity in the sample that *cannot* be ignored and this heterogeneity is of complex form.

Market share and rents (see Figs. 2(a) and 2(f)) provide a clear conclusion: The effect of these variables on productivity is unambiguously negative suggesting that competition has a positive effect on productivity. The distributions are clearly non-normal, and the distribution of the effect of rents is even trimodal with modes at 0, −0.4, and −0.8, suggesting that the effect of competition on productivity may vary by industry and also by firm. The effect of concentration ratio appears to be close to zero on the average (−0.0045 with standard deviation 0.0105), but from Fig. 2(d) we see that the mode near −0.02 suggests a positive effect of competition on productivity *growth* for a non-negligible portion of the sample. Thus, once again we observe that competition improves performance, although not so clearly as in the case of rents and market share. For some firms and industries, this effect does not exist (in fact, for a large but not dominating part). For others, the effect of competition on productivity growth is clearly positive.

The effect of import penetration (Fig. 2(e)) ranges from −0.04–0.04, suggesting either that the effect is too heterogeneous or that the effect is really zero (this is, in fact the conclusion from the estimates reported in Nickell, 1996 and Nickell et al., 1997). Finally, the effects of overtime and its inverse (Figs. 2(b) and 2(c)) is as expected from the results in Nickell (1996) and Nickell et al. (1997) although *t*-statistics reported by these authors seem to severely understate the sampling variability, possibly due to their parametric assumption and the highly asymmetric and non-normal pattern of the heterogeneity in coefficients.

Such results strongly suggest that competition is good for productivity and productivity growth, but the pattern is complicated, and the effects are highly heterogeneous and cannot be described adequately by normal, symmetric, or unimodal distributions. In a sense, our results help to invigorate the sometimes weak results reported by in Nickell (1996) and Nickell et al. (1997). To summarize, our analysis suggests an unambiguous positive effect of market share and rents on the level of productivity and an ambiguous positive effect of concentration of productivity *growth*. The

[7]Our measure of "long run" returns to scale is close to 0.88. Focusing on averages it does not appear that constant returns to scale is at odds with this data set. Given the shapes of the distributions in Figs. 1(a)–(c), however, this would be a gross simplification of reality.

meaning of "ambiguous" here is that the effect exists for certain industries and firms, but not for all firms and all industries in the sample—so this is not in fact a weak result. Thus we feel that the techniques presented here can shed additional light on key debates of the industrial organization literature and can enrich our understanding of firm-level and industry-wide heterogeneity.

6. POSSIBLE EXTENSION

In this section, we will briefly discuss how to estimate the varying coefficient $\delta(\cdot)$ in a fixed effects model. The model is same as in (1) except that $u_{it} = \mu_i + \varepsilon_{it}$ with μ_i being individual fixed effects. Taking the first differences to eliminate the fixed effects, we obtain

$$y_{it}^* = x_{it}'\delta(z_{it}) - x_{i,t-1}'\delta(z_{i,t-1}) + \varepsilon_{it}^*, \tag{8}$$

where y_{it}^*, x_{it}^*, ε_{it}^* are first differences variables. Note that equation (8) has an additive form, and ε_{it}^* has at least an $MA(1)$ structure. In principle, the method discussed above can be modified coupled with integrating method to estimate $\delta(z_{it})$ and $\delta(z_{i,t-1})$. However, one drawback with this approach is that it does not impose the restriction that the two additive functions in (8) have the same functional form $\delta(\cdot)$. To overcome this shortcoming, an alternative approach is to use the series methods to estimate (8). Series methods can easily impose the restriction that two additive functions have the same functional form, see for example by Li (2000), Ahmad et al. (2005). We leave the estimation problem of (8) using series methods as a future research topic.

7. CONCLUSIONS

In this article, we propose semiparametric smooth coefficient models for panel data. The proposed models are useful and flexible specification for examining the varying coefficients in the general regression relationship. We suggest a local generalized method of moments with kernel weights to estimate the smooth coefficient functions. The consistency and asymptotic normality of the proposed estimator are established. Limited Monte Carlo simulations suggest that our estimator performs quite well in finite sample. We apply the proposed method to data on the U.K. manufacturing companies from 1982–1994 to examine the effects of competition and corporate performance on productivity growth. We found that the analysis not only reinforce the findings in Nickell (1996) but also uncovered some of the heterogeneous effect that are not captured in the parametric specification of the model.

We did not consider the hypothesis testing problems in this article. It would be interesting and useful to test (i) whether or not a homogenous effect exists and (ii) whether or not serial correlation exists. Li et al. (2002), Fan et al. (2001), and Li and Hsiao (1998) provide testing frameworks in the cross-sectional/time series context, and their methods can be extended to the semiparametric panel data case. We leave these topics for future research.

APPENDIX

Proof of Theorem 1. Let $z_0 = \tau_T z'$, $\|H\| = \sqrt{\sum_{j=1}^{q} h_j^2}$, and $\xi(z + h\varphi) = \xi(z_1 + h_1\varphi_1, \ldots, z_q + h_q\varphi_q)$. It is more convenient to express the two-step estimator in (7) as

$$\begin{aligned}
\hat{\delta}_{TS}(z) &= \left[X'KW\widehat{S}_N^{-1}W'KX\right]^{-1}X'KW\widehat{S}_N^{-1}W'KY \\
&= \left[X'KW\widehat{S}_N^{-1}W'KX\right]^{-1}X'KW\widehat{S}_N^{-1}W'K\left\{X\delta(z) + X[\delta(Z) - \delta(z)] + u\right\} \\
&= \delta(z) + \left[X'KW\widehat{S}_N^{-1}W'KX\right]^{-1}X'KW\widehat{S}_N^{-1}W'K\left\{X[\delta(Z) - \delta(z)] + u\right\} \\
&= \delta(z) + \left\{C_N'(z)\widehat{S}_N(z)^{-1}C_N(z)\right\}^{-1}C_N'(z)\widehat{S}_N(z)^{-1}\left\{D_{1,N}(z) + D_{2,N}(z)\right\},
\end{aligned} \tag{A1}$$

where

$$C_N(z) = W'KX = N^{-1}\sum_{i=1}^{N} W_i'K_i^T X_i,$$

$$D_{1,N}(z) = W'KX\theta(Z) = N^{-1}\sum_{i=1}^{N} W_i'K_i^T X_i\theta_i(Z) \qquad \text{where } \theta_i(Z) = \delta(Z_i) - \delta(z_0),$$

$$D_{2,N}(z) = W'Ku = N^{-1}\sum_{i=1}^{N} W_i'K_i^T u_i.$$

Theorem 1 will be proved if we can show the following:

(i) $C_N(z) = A(z) + o_p(1)$, where $A(z) = f(z)E(W_1'X_1 \mid Z_1 = z_0)$;
(ii) $\widehat{S}_N(z) = S(z) + o_p(1)$, where $S(z) = T\Omega_0 f(z)\int K^2(z)dz'$;
(iii) $D_{1,N}(z) = \sum_{j=1}^{q} h_j^2 B_j(z) + o_p\left(\sum_{j=1}^{q} h_j^2\right)$;
(iv) $D_{2,N}(z) = o_p(1)$;
(v) $(N|H|)^{1/2}D_{2,N}(z) \xrightarrow{d} N(0, S(z))$.

These results are proven next.

Proof of (i). Under *i.i.d.* assumption, and by the law of iterative expectation, we have

$$
\begin{aligned}
E(C_N(z)) &= N^{-1}\sum_{i=1}^{N} E(W_i' X_i K_i^T(z)) = E(W_1' X_1 K_1^T(z)) \\
&= \sum_{t=1}^{T} E\big\{E(w_{1t}x_{1t}' \mid z_{1t})K_1^T(z_{1t}-z)\big\} \\
&= \sum_{t=1}^{T} \int E(w_{1t}x_{1t}' \mid z_{1t}) f(z_{1t}) K_1^T(z_{1t}-z)\, dz_{1t}' \\
&= \sum_{t=1}^{T} E(w_{1t}x_{1t}' \mid z_{1t}=z) f(z)\bigg[\int K_1^T(\varphi)\, d\varphi' + O(\|H\|)\bigg] \\
&= f(z)E(W_1' X_1 \mid Z_1 = z_0) + o(1) = A(z) + o(1).
\end{aligned}
$$

Similarly, one can show that

$$\operatorname{var}(C_N(z)) = O((N|H|)^{-1}) = o(1).$$

Thus,

$$C_N(z) = A + o_p(1). \tag{A2}$$

Proof of (ii) (*Sketch of Proof*). Under *i.i.d.* assumption and recall that

$$
\begin{aligned}
\widehat{S}_N(z) &= N^{-1}\int K^2(z)\, dz' \bigg\{\sum_{i=1}^{N} W_i^{*\prime}\widehat{V}_i W_i^*\bigg\} \\
&= N^{-1}\int K^2(z)\, dz' \bigg\{\sum_{i=1}^{N}\sum_{t=1}^{T} \hat{u}_{it}^2 w_{it} w_{it}' K_H(z_{it}-z)\bigg\}
\end{aligned}
$$

and by standard arguments for uniform convergence (e.g., Marsy, 1996), $\hat{\delta}_{TS}(z) - \delta(z) = O_p((N|H|)^{-1/2}(\ln N)^{1/2} + \|H\|^2)$ uniformly in z, and hence it is easy to show that $\hat{u}_{it} = u_{it} + o_p(1)$ uniformly implying $\hat{u}_{it}^2 = u_{it}^2 + o_p(1)$ uniformly. Thus, by the law of large numbers coupled with the law of iterative expectations, we have

$$
\begin{aligned}
\widehat{S}_N &\xrightarrow{p} \int K^2(z)\, dz' E\bigg\{\sum_{t=1}^{T} u_{it}^2 w_{it} w_{it}' K_H(z_{it}-z)\bigg\} \\
&= f(z)\int K^2(z)\, dz' \bigg\{\sum_{t=1}^{T} E\big(u_{it}^2 w_{it} w_{it}' \mid z_{it} = z\big)\bigg\} = S(z)
\end{aligned}
\tag{A3}
$$

Proof of (iii).

$$\begin{aligned}
E(D_{1,N}(z)) &= N^{-1}\sum_{i=1}^{N} E\big(W_i' K_i^T X_i \theta_i(z)\big) = E\big[W_1' K_1^T X_1(\delta(Z_1)-\delta(z))\big] \\
&= \sum_{t=1}^{T} E\big\{E(w_{1t}x_{1t}' \mid z_{1t})(\delta(z_{1t})-\delta(z))K_1^T(z_{1t}-z)\big\} \\
&= \sum_{t=1}^{T}\int E(w_{1t}x_{1t}' \mid z_{1t})(\delta(z_{1t})-\delta(z))f(z_{1t})K_1^T(z_{1t}-z)\,dz_{1t}' \\
&= \sum_{t=1}^{T}\int E(w_{1t}x_{1t}' \mid z_{1t}=z)\Big[\sum_{j=1}^{q} h_j\varphi_j\delta_j(z) + (1/2)\sum_{j=1}^{q} h_j^2\varphi_j^2\delta_{jj}(z)\Big] \\
&\quad \times \Big[f(z)+\sum_{j=1}^{q} f_j(z)h_j\varphi_j\Big]K_1^T(\varphi)\,d\varphi' + O\Big(\sum_{j=1}^{q} h_j^3\Big) \\
&= \frac{1}{2}\kappa_2\sum_{j=1}^{N} h_j^2\Big\{f(z)\sum_{t=1}^{T} E(w_{1t}x_{1t}' \mid z_{1t}=z)\delta_{jj}(z) \\
&\qquad + 2f_j(z)\sum_{t=1}^{T} E(w_1 x_{1t}' \mid z_{1t}=z)\delta_j(z)\Big\} + O\Big(\sum_{j=1}^{q} h_j^3\Big) \\
&= \frac{1}{2}\kappa_2\sum_{j=1}^{N} h_j^2\Big[f(z)E(W_1'X_1 \mid Z_1=z_0)\delta_{jj}(\cdot) \\
&\qquad + 2\frac{\partial f(z)E(W_1'X_1 \mid Z_1=z_0)}{\partial z_j}\delta_j(\cdot)\Big] + O\Big(\sum_{j=1}^{q} h_j^3\Big) \\
&= \sum_{j=1}^{q} h_j^2 B_j + O\Big(\sum_{j=1}^{q} h_j^3\Big).
\end{aligned}$$

Similarly, one can show that

$$\mathrm{Var}[D_{1,N}(z)] = O\Big(\sum_{j=1}^{q} h_j^2 (N|H|)^{-1} + \sum_{j=1}^{q} h_j^5\Big).$$

Thus,

$$\Big\{D_{1,N}(z) - \sum_{j=1}^{q} h_j^2 B_j(z)\Big\} = o_p\Big(\sum_{j=1}^{q} h_j^2 + (N|H|)^{-1/2}\Big). \tag{A4}$$

Proof of (iv). By the law of iterative expectation, we have

$$\begin{aligned}
E(D_{2N}(z)) &= E\big(W_1' K_1^T u_1\big) = E\big(E(W_1' K_1^T u_1 \mid Z_1 = z_0)\big)\\
&= E\Bigg(\sum_{t=1}^{T} E(w_{1t}u_{1t} \mid z_{1t} = z)K_H(z_{1t} - z))\Bigg)\\
&= f(\varphi)E(w_{1t}u_{1t} \mid z_{1t} = \varphi) = 0,
\end{aligned}$$

and it can be shown that $\operatorname{var}(D_{2N}) = O((N|H|)^{-1})$; see the proof (v) below. Thus,

$$D_{2N}(z) \xrightarrow{p} 0. \tag{A5}$$

Proof of (v). $\sqrt{N|H|}\, D_{2,N}(z)$ has mean zero, and its variance is given by

$$\begin{aligned}
\operatorname{Var}[D_{2,N}(z)] &= N^{-1}E\big(W_1' K_1^T u_1 u_1' K_1^T W_1\big)\\
&= N^{-1}\sum_t \big\{E(u_{1t}^2 w_{1t} w_{1t}')K_H^2(z_{1t} - z)\big\}\\
&\quad + N^{-1}\sum_t\sum_r \big\{E(u_{1t}u_{1r}w_{1t}w_{1r}')K_H(z_{1t} - z)K_H(z_{1r} - z)\big\}\\
&= N^{-1}\Gamma_0 + N^{-1}\sum_{r=1}^{T-1}\left(\frac{T-r}{T}\right)\big(\Gamma_r + \Gamma_r'\big)\\
&= J_1 + J_2,
\end{aligned}$$

where $\Gamma_0 = \sum_t E\big\{u_{1t}^2 w_{1t} w_{1t}' K_H^2(z_{it} - z)\big\}$ and $\Gamma_r = E\big\{u_{1t}u_{1t-r}w_{1t}w_{1t-r}' K_H(z_{it} - z)\times K_H(z_{it-r} - z)\big\}$.

Now consider the first term J_1. By strict stationarity and the law of iterative expectations, we have

$$\begin{aligned}
\Gamma_0 &= \sum_t E\big\{E\big(u_{1t}^2 w_{1t} w_{1t}' \mid z_{1t} = c\big)K_H^2(c - z)\big\}\\
&= \int \Omega_0(c)K_H^2(c - z)f(c)\,dc'\\
&= |H|^{-1}\int \Omega_0(z + H\xi)K^2(\xi)f(z + H\xi)\,d\xi',
\end{aligned}$$

where the third equality follows by making the substitution $\xi = H^{-1}(c - z)$. Thus by assumptions (A.1) and (A.3), it follows that

$$N|H|\left\{\frac{1}{N}\sum_t \Gamma_0(z)\right\} \to \Omega_0(z)f(z)\int K^2(z)\,dz'.$$

Next we show that the second term J_2 above is of order $O(N^{-1})$. By the law of iterative expectations, and making substitutions $s_1 = H^{-1}(c_1 - z)$ and $s_2 = H^{-1}(c_2 - z)$, we obtain

$$\begin{aligned}\Gamma_r &= E\big\{E\big(u_{1t}u_{1t-r}w_{1t}w'_{1t-r} \mid z_{1t} = c_1, z_{1t-r} = c_2\big)K_H(z_{it} - z)K_H(z_{it-r} - z)\big\} \\ &= \int \Omega_r(c_1, c_2)K_H(c_1 - z)K_H(c_2 - z)g_r(c_1, c_2)\,dc'_1\,dc'_2 \\ &= \int \Omega_r(z + Hs_1, z + Hs_2)K(s_1)K(s_2)g_r(z + Hs_1,\ z + Hs_2)\,ds'_1\,ds'_2.\end{aligned}$$

By assumptions (A.1) and (A.4), it follows that $\Gamma_r \to \Omega_r(z, z)g_r(z, z)$, implying $J_2 = O(N^{-1})$. Thus,

$$\lim_{N\to\infty} \mathrm{Var}\Big[\sqrt{N|H|}D_{2,N}(z)\Big] = \Omega_0(z)f(z)\int K^2(z)\,dz' = S(z).$$

It is straightforward to check that the conditions of Lyapounov's central limit theorem hold. Thus,

$$\sqrt{N|H|}D_{2,N}(z) \xrightarrow{d} N(0, S(z)). \tag{A6}$$

Combining (A2)–(A6), we have shown that,

(a)
$$\begin{aligned}\{\hat{\delta}_{TS}(z_0) - \delta(z_0)\} &= \{C_N(z)'\widehat{S}(z)^{-1}C_N(z)\}^{-1}C_N(z)'\widehat{S}(z)^{-1} \\ &\quad \times \big[D_{1,N}(z) + D_{2,N}(z)\big] \\ &= \Psi(z)\sum_{j=1}^{q} h_j^2 B_j(z) + o_p\left(\sum_{j=1}^{q} h_j^2 + (N|H|^{-1/2}\right)\end{aligned}$$

Hence, $\hat{\delta}_{TS}(z) - \delta(z) - \Psi(z)\sum_{j=1}^{q} h_j^2 B_j(z) = o_p\big(\sum_{j=1}^{q} h_j^2 + (N|H|)^{-1/2}\big)$;

(b)
$$\begin{aligned}&\sqrt{N|H|}\{\hat{\delta}_{TS}(z_0) - \delta(z_0)\} \\ &= \{C_N(z)'\widehat{S}(z)^{-1}C_N(z)\}^{-1}C_N(z)'\widehat{S}(z)^{-1}\sqrt{N|H|}\big[D_{1,N}(z) + D_{2,N}(z)\big] \\ &= \big\{A(z)'S(z)^{-1}A(z)\big\}^{-1}A(z)'S(z)^{-1} \\ &\quad \times \left[\sqrt{N|H|}\sum_{j=1}^{q} h_j^2 B_j + \sqrt{N|H|}D_{2,N}(z) + o_p(1)\right] \\ &\xrightarrow{d} N\left(\Psi(z)\sum_{j=1}^{q} h_j^2 B_j, \big\{A(z)'S(z)^{-1}A(z)\big\}^{-1}\right)\end{aligned}$$

$\sqrt{N|H|}\big\{\hat{\delta}_{TS}(z_0) - \delta(z_0) - \Psi(z)\sum_{j=1}^{q} h_j^2 B_j\big\} \xrightarrow{d} N\big(0, \{A(z)'S(z)^{-1}A(z)\}^{-1}\big)$. □

ACKNOWLEDGMENTS

We would like to thank an associate editor and two anonymous referees for detailed insightful comments and suggestions that led to a substantial improvement of the article. We would also like to thank the participants of the Hellenic Workshop on Efficiency and Productivity Measurement (Patras, Greece, 2006) and especially Robin Sickles and Dawit Zerom for many useful comments on an earlier version of this article.

REFERENCES

Ahmad, I., Leelahanon, S., Li, Q. (2005). Efficient estimation of a semiparametric partially linear varying coefficient model. *The Annals of Statistics* 33:258–283.

Arellano, M., Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58:277–297.

Baltagi, B. H., Li, Q. (2002). On instrumental variable estimation of semiparametric dynamic panel data models. *Economics Letters* 76:1–9.

Cai, Z., Li, Q. (2008). Nonparametric estimation of varying coefficient dynamic panel data models. *Econometric Theory* 24:1321–1342.

Cai, Z., Fan, J., Zhao, Q. (2000). Functional coefficient regression models for nonlinear time series. *Journal of the American Statistical Association* 95:941–956.

Cai, Z., Das, M., Xiong, H., Wu, X. (2006). Functional coefficient instrumental variables models. *Journal of Econometrics* 133:207–241.

Carroll, R., Ruppert, J., Welsh, A. (1998). Local estimating equations. *Journal of the American Statistical Association* 93:214–227.

Das, M. (2005). Instrumental variables estimators for nonparametric models with discrete endogenous regressors. *Journal of Econometrics* 124:335–361.

Gozalo, P., Linton, O. (2000). Local nonlinear least squares: using parametric information in non-parametric regression. *Journal of Econometrics* 99:63–106.

Fan, J., Zhang, W. (1999). Statistical estimation in varying coefficient models. *The Annals of Statistics* 27:1491–1518.

Fan, J., Zhang, C., Zhang, J. (2001). Generalized likelihood ratio statistics and Wilks phenomenon. *The Annals of Statistics* 29:153–193.

Hansen, B. E. (1999). Threshold effects in non-dynamic panels: estimation, testing and inference. *Journal of Econometrics* 93:345–368.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica* 50:1029–1054.

Hastie, T., Tibshirani, R. (1993). Varying-coefficient models. *Journal of the Royal Statistical Society B* 55:757–796.

Horowitz, J. L., Markatou, M. (1996). Semiparametric estimation of regression models for panel data. *The Review of Economic Studies* 63:145–168.

Knieser, T. J., Li, Q. (2002). Nonlinearity in dynamic adjustment: semiparametric estimation of panel labor supply. *Empirical Economics* 27:131–148.

Li, Q. (2000). Efficient estimation of additive partially linear models. *International Economic Review* 41:1073–1092.

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

Li, Q., Hsiao, C. (1998). Testing for serial correlation in semiparametric panel data models. *Journal of Econometrics* 87:207–237.

Li, Q., Ullah, A. (1998). Estimating partially linear models with one-way error components. *Econometric Reviews* 17:145–166.

Li, Q., Huang, C., Li, D., Fu, T.-T. (2002). Semiparametric smooth coefficient models. *Journal of Business and Economic Statistics* 20:412–422.

Marsy, E. (1996). Multivariate local polynomial regression for time series: uniform strong consistency rates. *Journal of Time Series Analysis* 17:571–599.

Newey, W. (1990). Efficient instrumental variable estimation of nonlinear models. *Econometrica* 58:809–837.

Nickell, S. J. (1996). Competition and corporate performance. *Journal of Political Economy* 104:724–746.

Nickell, S. J., Nicolitsas, D., Dryden, N. (1997). What makes firms perform well? *European Economic Review* 41:783–796.

Papalia, R. (1999). Local generalized method of moments estimation based on kernel weights: An application to panel data. *Journal of Applied Statistics* 26(8):1005–1015.

Powell, J. L., Stock, J. H., Stoker, T. M. (1989). Semiparametric estimation of index coefficients. *Econometrica* 57:1043–1430.

Ullah, A., Roy, N. (1998). Nonparametric and semiparametric econometrics of panel data. In: Ullah, A., Giles, D. E. A., eds. *Handbook of Applied Economic Statistics*. Monticello, New York: Marcel Dekker, Chapter 17, pp. 579–604.

Ullah, A., Mundra, K. (1999). Semiparametric panel data estimation: An application to immigrates homelink effect on US producer trade flows. Department of Economics, University of California at Riverside, Working Paper 15.