ISSN 1440-771X

MONASH University

Australia

Department of Econometrics and Business Statistics

http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/

Non- and Semi-Parametric Panel Data Models: A Selective Review

Jia Chen, Degui Li and Jiti Gao

August 2013

Working Paper 18/13

Non– and Semi–Parametric Panel Data Models: A Selective Review

By Jia Chen[†], Degui Li[‡,*,1] and Jiti Gao[*]

[†] *University of Queensland,* [‡] *University of York and* [*] *Monash University*

Abstract

This article provides a selective review on the recent developments of some nonlinear nonparametric and semiparametric panel data models. In particular, we focus on two types of modelling frameworks: nonparametric and semiparametric panel data models with deterministic trends, and semiparametric single-index panel data models with individual effects. We also review various estimation methodologies which can consistently estimate both the parametric and nonparametric components in these models. The time series length and cross-sectional size in this article are allowed to be very large, under which the panel data are called "large dimensional panels".

JEL classification: C13, C14, C23.

Keywords: Deterministic trends, local linear fitting, panel data, semiparametric estimation, single-index models.

1. Introduction

Analysis of panel data has received a lot of attention in last three decades due to the prevalence of panel data in many disciplines such as economics, finance and biology. Panel data possess various advantages over purely time series or cross-sectional data. For instance, the double-index panel data enable researchers to estimate complex models and extract dynamic information which may be difficult to obtain by using purely cross-sectional data. Hence, there exists a rich literature on parametric linear and nonlinear panel data models as well as their statistical inference and econometric applications, see, for example, Baltagi (1995), Arellano (2003) and Hsiao (2003).

However, it is well known that the pre-specified parametric assumption on the panel data models might be too restrictive, and the parametric models might be misspecified in some practical applications. Such misspecification may lead to inconsistent estimates and hence incorrect conclusions being drawn. To address such issues, some more flexible nonlinear modelling frameworks, and nonparametric and semiparametric estimation methodologies have been introduced in recent years by allowing the panel data "speak for themselves".

[1]Corresponding Author: Degui Li, Department of Mathematics, University of York, Heslington, YO10 5DD, UK. Email: degui.li@york.ac.uk.

Recent studies in this field include Ullah and Roy (1998), Hjellvik *et al* (2004), Su and Ullah (2006, 2010), Wu and Zhang (2006), Li and Racine (2007), Cai and Li (2008), Henderson *et al* (2008), Mammen *et al* (2009), Li *et al* (2011a), Chen *et al* (2012a, 2012b, 2013a, 2013b), Su and Lu (2013).

In this paper, we will provide a selective review on the recent developments of nonparametric and semiparametric panel data models. In particular, we focus on two modelling frameworks: nonparametric and semiparametric panel data models where deterministic trends are involved, and semiparametric single-index panel data models with individual effects, and also review related nonparametric and semiparametric estimation methodologies. Unlike the model assumptions in statistics literature (e.g. Wu and Zhang, 2006) where the panel data are collected for many units over a short time span, we assume in this paper that both the time series length and cross-sectional size are very large, and the panel data under this case are called "large dimensional panels" (Bai *et al*, 2009). In recent years, many researchers have come up with new ideas to exploit the rich information contained in large panels. The assumption of large dimensional panel is extremely useful in exploring stochastic trends (Phillips and Moon, 1999; Bai *et al*, 2009) or deterministic trends (Robinson, 2011; Chen *et al*, 2012b) in econometric analysis of panel data.

Trends are the dominant characteristic in much of economic, financial and climatic data and have been extensively studied in recent years. Phillips (2001, 2010) provide a review on the developments and challenges on the modelling of trends. Mainly due to the limitation and implausibility of parametric models, recent literature mainly focuses on time-varying coefficient trending models and uses nonparametric and semiparametric estimation methods. Such studies include Robinson (1989), Cai (2007) and Li *et al* (2009, 2011b). Gao and Hawthorne (2006) propose using a semiparametric time series model to model the trend in global and hemispheric temperature series while at the same time allow for the inclusion of some explanatory variables in a parametric component. In Section 2, we will review the extension of such models to the panel data case: semiparametric trending panel data models (Chen *et al*, 2012b) and nonparametric time-varying coefficient panel data models (Li *et al*, 2011a). As fixed effects are involved in these models, to obtain unbiased estimates of the model parameters, the developed nonparametric and semiparametric procedures eliminate the influence of these fixed effects by treating them as nuisance parameters. Furthermore, we will review the asymptotic theory for these nonparametric and semiparametric estimators.

When there are more than three covariates, it is well known that the underlying nonparametric regression function cannot be estimated with reasonable accuracy due to the so-called "curse of dimensionality". The circumvention of the curse of dimensionality is an important issue in nonparametric regression analysis. Many approaches have been developed to address this issue in nonlinear time series (or cross-sectional) analysis (e.g., Fan and Yao, 2003; Gao, 2007; and Li and Racine, 2007). In Section 3, we will review some recent developments on semiparametric single-index or partially linear single-index panel data models. These modelling frameworks avoid the curse of dimensionality problem; allow for nonlinear relationship between the response variable and explanatory variables; and retain much of the ease of interpretation of the fully linear panel data models. Meanwhile, any existing information concerning possible linearity of some of the components can be taken into account in such models.

The rest of the paper is organized as follows. Section 2 reviews the nonparametric and semiparametric trending panel data models as well as their estimation procedures. Section 3 reviews semiparametric single-index and partially linear single-index panel data models as well as the associated estimation procedures. Section 4 discusses some possible future research topics and concludes this paper.

2. Nonlinear trending panel data models

As mentioned in the introductory section, there has been a rich literature on parametric and nonparametric time series models with various trending characteristics. There has also been increasing interest in identifying and estimating the trend in a panel data set in recent years. For example, Atak *et al* (2011) propose a semiparametric panel data model to deal with the modeling of climate change in the United Kingdom, and consider using a model with a dummy variable in the parametric component while allow for the time trend function to be nonparametrically estimated. In this section, we will review the semiparametric trending panel data model studied by Chen *et al* (2012b) and the nonparametric time-varying coefficient panel data model studied by Li *et al* (2011a). The panel data under investigation can be cross-sectionally dependent.

2.1. Semiparametric trending panel data models

Consider a panel data model of the form (Chen *et al*, 2012b)

$$Y_{it} = X_{it}^{\top}\beta + f_t + \alpha_i + e_{it}, \tag{2.1}$$

$$X_{it} = g_t + x_i + v_{it}, \quad i = 1, \cdots, N, \quad t = 1, \cdots, T, \tag{2.2}$$

where β is a d-dimensional vector of unknown parameters, $f_t = f\left(\frac{t}{T}\right)$ and $g_t = g\left(\frac{t}{T}\right)$ are two unknown time trend functions, α_i and x_i are individual specific effects, and $\{e_{it}\}$ and $\{v_{it}\}$ are serially stationary and dependent both over time and cross-sectionally. Note that $\{\alpha_i\}$ is allowed to be correlated with $\{X_{it}\}$ through some unknown structure, and hence is a sequence of fixed effects. For the purpose of identifiability, we impose the following restrictions on α_i and x_i:

$$\sum_{i=1}^{N} \alpha_i = 0 \quad \text{and} \quad \sum_{i=1}^{N} x_i = \mathbf{0}_d, \tag{2.3}$$

where $\mathbf{0}_d$ is a d-dimensional null vector. When $\beta = \mathbf{0}_d$, model (2.1) reduces to the nonparametric model discussed in Robinson (2012). When $N = 1$, model (2.1) reduces to the model discussed in Gao and Hawthorne (2006). Meanwhile, (2.2) allows $\{X_{it}\}$ to have a trend and thus be nonstationary. The practical applications of model (2.1) include the modelling of the dependence between the share consumption, $\{Y_{it}\}$, and the total consumption, $\{X_{it}\}$.

In the case of $\beta = \mathbf{0}_d$, Robinson (2012) introduces an averaged kernel estimation method to estimate the function $f(\cdot)$. Such a method can be extended to estimating β and $f(\cdot)$ in model (2.1). However the convergence rates are not satisfactory as shown in Section 4.1 of Chen *et al* (2012b). Hence, Chen *et al* (2012b) propose a semiparametric profile likelihood estimation method. Before presenting this method, we introduce the following notation:

$$\widetilde{Y} = (Y_{11}, \cdots, Y_{1T}, Y_{21}, \cdots, Y_{2T}, \cdots, Y_{N1}, \cdots, Y_{NT})^\top,$$
$$\widetilde{X} = (X_{11}, \cdots, X_{1T}, X_{21}, \cdots, X_{2T}, \cdots, X_{N1}, \cdots, X_{NT})^\top,$$
$$\alpha = (\alpha_2, \cdots, \alpha_N)^\top, \quad D = (-i_{N-1}, I_{N-1})^\top \otimes i_T,$$
$$\widetilde{f} = i_N \otimes (f_1, \cdots, f_T)^\top, \quad \widetilde{e} = (e_{11}, \cdots, e_{1T}, e_{21}, \cdots, e_{2T}, \cdots, e_{N1}, \cdots, e_{NT})^\top,$$

where $\otimes$ denotes the Kronecker product, i_k is the $k \times 1$ vector of ones and I_k is the $k \times k$ identity matrix. As $\sum_{i=1}^{N} \alpha_i = 0$ as per the identification condition, model (2.1) can be rewritten as

$$\widetilde{Y} = \widetilde{X}\beta + \widetilde{f} + D\alpha + \widetilde{e}. \tag{2.4}$$

Let $W(\tau) = \text{diag}\left(K(\frac{1-\tau T}{Th}), \cdots, K\left(\frac{T-\tau T}{Th}\right)\right)$ and $\widetilde{W}(\tau) = I_N \otimes W(\tau)$, where $K(\cdot)$ is a kernel function and h is a bandwidth, and

$$Z(\tau) = \begin{pmatrix} 1 & \frac{1-\tau T}{Th} \\ \vdots & \vdots \\ 1 & \frac{T-\tau T}{Th} \end{pmatrix}, \quad \widetilde{Z}(\tau) = i_N \otimes Z(\tau).$$

The pooled semiparametric profile likelihood estimation procedure contains the following three steps.

Step (i) Define the loss function by:

$$L_1(a, b) = \left[\widetilde{Y} - \widetilde{X}\beta - D\alpha - \widetilde{Z}(\tau)(a, b)^\top\right]^\top \widetilde{W}(\tau) \left[\widetilde{Y} - \widetilde{X}\beta - D\alpha - \widetilde{Z}(\tau)(a, b)^\top\right].$$

For given α and β, estimate $f(\tau)$ and $hf'(\tau)$ by

$$\begin{pmatrix} \widehat{f}_{\alpha,\beta}(\tau) \\ h\widehat{f}'_{\alpha,\beta}(\tau) \end{pmatrix} = \arg\min_{(a,b)^\top} L_1(a,b).$$

By simple calculation, we have

$$\widehat{f}_{\alpha,\beta}(\tau) = (1,0)S(\tau)(\widetilde{Y} - \widetilde{X}\beta - D\alpha) = s(\tau)(\widetilde{Y} - \widetilde{X}\beta - D\alpha), \tag{2.5}$$

where $s(\tau) = (1,0)S(\tau)$ and $S(\tau) = \left[\widetilde{Z}^\top(\tau)\widetilde{W}(\tau)\widetilde{Z}(\tau)\right]^{-1} \widetilde{Z}^\top(\tau)\widetilde{W}(\tau)$.

Step (ii) Denote

$$\widetilde{f}_{\alpha,\beta} = i_N \otimes \left(\widehat{f}_{\alpha,\beta}(1/T), \cdots, \widehat{f}_{\alpha,\beta}(T/T)\right)^\top = \widetilde{S}(\widetilde{Y} - \widetilde{X}\beta - D\alpha),$$

where $\widetilde{S} = i_N \otimes \left(s^\top(1/T), \cdots, s^\top(T/T)\right)^\top$. Then, estimate α and β by

$$\begin{aligned}
(\widehat{\alpha}^\top, \widehat{\beta}^\top)^\top &= \arg\min_{(\alpha^\top,\beta^\top)^\top} \sum_{i=1}^{N}\sum_{t=1}^{T}\left[Y_{it} - X_{it}\beta - \alpha_i - \widehat{f}_{\alpha,\beta}\left(\frac{t}{T}\right)\right]^2 \\
&= \arg\min_{(\alpha^\top,\beta^\top)^\top} \left(\widetilde{Y} - \widetilde{X}\beta - D\alpha - \widetilde{f}_{\alpha,\beta}\right)^\top \left(\widetilde{Y} - \widetilde{X}\beta - D\alpha - \widetilde{f}_{\alpha,\beta}\right) \\
&= \arg\min_{(\alpha^\top,\beta^\top)^\top} \left(\widetilde{Y}^* - \widetilde{X}^*\beta - D^*\alpha\right)^\top \left(\widetilde{Y}^* - \widetilde{X}^*\beta - D^*\alpha\right),
\end{aligned} \tag{2.6}$$

where $\widetilde{Y}^* = \left(I_{NT} - \widetilde{S}\right)\widetilde{Y}$, $\widetilde{X}^* = \left(I_{NT} - \widetilde{S}\right)\widetilde{X}$ and $D^* = \left(I_{NT} - \widetilde{S}\right)D$. Letting $M^* = I_{NT} - D^*\left(D^{*\top}D^*\right)^{-1}D^{*\top}$, we can obtain the solution to the minimization problem (2.6) as follows:

$$\widehat{\beta} = \left(\widetilde{X}^{*\top}M^*\widetilde{X}^*\right)^{-1}\widetilde{X}^{*\top}M^*\widetilde{Y}^*, \tag{2.7}$$

$$\widehat{\alpha} = \left(D^{*\top}D^*\right)^{-1}D^{*\top}\left(\widetilde{Y}^* - \widetilde{X}^*\widehat{\beta}\right). \tag{2.8}$$

Step (iii) Obtain the final nonparametric estimate of $f(\tau)$ as

$$\widehat{f}(\tau) = s(\tau)\left(\widetilde{Y} - \widetilde{X}\widehat{\beta} - D\widehat{\alpha}\right). \tag{2.9}$$

Under some technical assumptions, Chen *et al* (2012b) derive some nice asymptotic properties for $\widehat{\beta}$ and $\widehat{f}(\tau)$ by letting $T \to \infty$ and $N \to \infty$ simultaneously, an approach which is called the "joint limit approach" in Phillips and Moon (1999). Specifically they show that

$$\sqrt{NT}\left(\widehat{\beta} - \beta\right) \xrightarrow{d} N\left(\mathbf{0}_d,\ \Sigma_v^{-1}\Sigma_{v,e}\Sigma_v^{-1}\right) \tag{2.10}$$

and

$$\sqrt{NTh}\left[\widehat{f}(\tau) - f(\tau) - b_f(\tau)h^2 + o_P(h^2)\right] \xrightarrow{d} N\left(0,\ \nu_0\sigma_e^2\right), \tag{2.11}$$

where $b_f(\tau) = \frac{1}{2}\mu_2 f''(\tau)$, $f''(\cdot)$ is the second-order derivative of the nonlinear trend function $f(\cdot)$, $\mu_j = \int u^j K(u)du$ and $\nu_j = \int u^j K^2(u)du$. The definitions of Σ_v, Σ_{ve} and σ_e^2 are given in Section 3.1 of Chen *et al* (2012b). The finite sample performance of the developed model and estimation method as well as their applications in analyses of Australian CPI data and input-output data can also be found in Chen *et al* (2012b).

2.2. Nonparametric time-varying coefficient panel data models

We next generalise model (2.1) by allowing the relation between $\{Y_{it}\}$ and $\{X_{it}\}$ to vary over time. Such an extension is useful for analysing panel data observed over long time horizons over which economic mechanisms are likely to evolve and subject to changing institutional or regulatory conditions. The evolving relation between $\{Y_{it}\}$ and $\{X_{it}\}$ in response to such changes may be captured by replacing β by $\beta_t = \beta(\frac{t}{T})$. This leads to the following model (Li *et al*, 2011a):

$$Y_{it} = X_{it}^\top \beta_t + f_t + \alpha_i + e_{it}, \quad i = 1, \cdots, N, \quad t = 1, \cdots, T, \tag{2.12}$$

where $f_t = f(\frac{t}{T})$ and α_i are defined as in Section 2.1. It is easy to see that model (2.12) includes model (2.1) and nonparametric time-varying coefficient time series model (Robinson, 1989; Cai *et al*, 2007) as special cases.

Our main interest lies in the estimation of the functions $f(\cdot)$ and $\beta(\cdot)$. Li *et al* (2011a) develop two types of local linear estimation approaches: averaged local linear estimation method and pooled local linear estimation method, and show that the latter has better performance in numerical studies. Hence, we next only introduce the pooled local linear estimation method, which is also called the local linear dummy variable approach (e.g. Li *et al*, 2011a). Note that (2.12) can be rewritten in the matrix form

$$\widetilde{Y} = \widetilde{f} + \widetilde{B}(X, \widetilde{\beta}) + D\alpha + \widetilde{e}, \tag{2.13}$$

where $\widetilde{Y}$, $\widetilde{f}$, D, α and $\widetilde{e}$ are defined as in Section 2.1, and

$$\widetilde{B}(X, \widetilde{\beta}) = \left(X_{11}^\top \beta_1, \cdots, X_{1T}^\top \beta_T, X_{21}^\top \beta_1, \cdots, X_{NT}^\top \beta_T\right)^\top.$$

Let $\beta_*(\cdot) = \left(f(\cdot), \beta^\top(\cdot)\right)^\top$. The two-step algorithm in the pooled local linear method is described as follows.

Step (i) For given $\widetilde{\beta}$ and $0 < \tau < 1$, first solve the optimization problem:

$$\min_{\alpha} \left(\widetilde{Y} - \widetilde{f} - \widetilde{B}(X, \widetilde{\beta}) - D\alpha\right)^\top \widetilde{W}(\tau) \left(\widetilde{Y} - \widetilde{f} - \widetilde{B}(X, \widetilde{\beta}) - D\alpha\right), \tag{2.14}$$

where $\widetilde{W}(\tau)$ is defined as in Section 2.1. Taking derivative of (2.14) with respect to α and setting it to zero, we obtain

$$\widehat{\alpha} \equiv \widehat{\alpha}(\tau) = \left(D^\top \widetilde{W}(\tau) D\right)^{-1} D^\top \widetilde{W}(\tau) \left(\widetilde{Y} - \widetilde{f} - \widetilde{B}(X, \widetilde{\beta})\right).$$

Step (ii) Replace α in (2.14) with $\widehat{\alpha}$ and estimate $\beta_*(\tau)$ and $h\beta_*'(\tau)$ by solutions to the following optimization problem:

$$\min_{a,b} \left(\widetilde{Y} - \widetilde{X}(\tau)(a^\top, b^\top)^\top\right)^\top \widetilde{W}^*(\tau) \left(\widetilde{Y} - \widetilde{X}(\tau)(a^\top, b^\top)^\top\right)^\top, \tag{2.15}$$

where $\widetilde{X}^\top(\tau) = \left(X_1^\top(\tau), \cdots, X_N^\top(\tau)\right)$ with

$$X_i(\tau) = \begin{pmatrix} 1 & X_{i1}^\top & \frac{1-\tau T}{Th} & \frac{1-\tau T}{Th} X_{i1}^\top \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{iT}^\top & \frac{T-\tau T}{Th} & \frac{T-\tau T}{Th} X_{iT}^\top \end{pmatrix},$$

$\widetilde{W}^*(\tau) = W_*^\top(\tau)\widetilde{W}(\tau)W_*(\tau)$ and $W_*(\tau) = I_{NT} - D\left(D^\top \widetilde{W}(\tau) D\right)^{-1} D^\top \widetilde{W}(\tau)$.

As $\widetilde{W}(\tau)D\alpha = \mathbf{0}_{NT}$, the fixed effects term $D\alpha$ contained in $\widetilde{Y}$ can be eliminated from the minimization objective function in (2.15). By simple calculation, we obtain the local linear estimate of $\beta_*(\tau)$, which is defined as the solution a to (2.15), as follows

$$\widehat{\beta}_*(\tau) = [I_{d+1},\ \mathbf{0}_{d+1}] \left[\widetilde{X}^\top(\tau)\widetilde{W}^*(\tau)\widetilde{X}(\tau)\right]^{-1} \widetilde{X}^\top(\tau)\widetilde{W}^*(\tau)\widetilde{Y}. \tag{2.16}$$

We call $\widehat{\beta}_*(\tau)$ the pooled local linear estimate of $\beta_*(\tau)$. Under some technical conditions, Li *et al* (2011a) establish the asymptotic distribution theory for $\widehat{\beta}_*(\tau)$, where the point-wise root-(NTh) convergence rate is derived. Li *et al* (2011a) also study the finite sample behavior of the pooled local linear estimation, and apply the developed model and methodology in the analysis of a UK climate data set.

3. Semiparametric single-index panel data models

Single-index modelling is one of the most commonly used semiparametric modelling techniques. It searches for a linear combination of multi-dimensional covariates which can capture most information about the relationship between the response variable and covariates, while still allowing for curvature between them by superimposing a nonlinear link function between the linear combination and the response. The single-index models for cross-sectional or time series data and their estimation procedures have been extensively studied in the existing literature (e.g. Härdle *et al*, 1993; Carroll *et al*, 1997; Xia *et al*, 2002; Xia and Härdle, 2006). In this section, we will review extensions of these models to the panel data case: single-index panel data models with heterogeneous link functions (Chen *et al*, 2013a) and partially linear single-index panel data models with fixed effects (Chen *et al*, 2013b).

3.1. Single-index panel data models with heterogeneous link functions

Consider the following panel data model (Chen *et al*, 2013a):

$$Y_{it} = g_i(\theta_0^\top X_{it}) + e_{it}, \quad i = 1, \cdots, N, \quad t = 1, \cdots T, \tag{3.1}$$

where $g_i(\cdot)$ are unknown link functions that vary across individuals, X_{it} are p-dimensional covariate vectors and $\theta_0 = (\theta_{01}, \cdots, \theta_{0p})^\top$ is a p-dimensional unknown parameter vector. For the purpose of identifiability, we assume $\|\theta_0\| = 1$ and that the first nonzero entity of θ_0 is positive. We next use a so-called refined minimum average variance estimation (RMAVE) method (e.g. Xia *et al*, 2002) to estimate the parameter θ_0.

Given $\theta^\top X_{it} = u$, define $\sigma^2_{\theta,i}(u) = \mathsf{E}\big[(Y_{it} - g_i(\theta^\top X_{it}))^2 | \theta^\top X_{it} = u\big]$ for $i = 1, \cdots, N$. Then, the estimate of the parameter θ_0 can be obtained by minimizing

$$\sum_{i=1}^{N} \mathsf{E}\left(Y_{it} - g_i(\theta^\top X_{it})\right)^2 = \sum_{i=1}^{N} \mathsf{E}_u\left[\sigma^2_{\theta,i}(u)\right].$$

As the link functions $g_i(\cdot)$ are unknown, we estimate them by the local linear method by assuming that $g_i(\cdot)$ are differentiable up to the third order for each $1 \leq i \leq N$. Following a Taylor's expansion of the link functions, we have

$$Y_{it} - g_i(\theta^\top X_{it}) \approx Y_{it} - g_i(\theta^\top x) - g_i'(\theta^\top x)\theta^\top (X_{it} - x),$$

when X_{it} are close to x. The parameter θ_0 can be estimated by minimizing

$$\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} K\Big(\frac{\theta^\top X_{its}}{h}\Big)\left[Y_{it} - g_i(\theta^\top X_{is}) - g_i'(\theta^\top X_{is})\theta^\top X_{its}\right]^2, \tag{3.2}$$

where $X_{its} = X_{it} - X_{is}$, $K(\cdot)$ is a kernel function and h is a bandwidth.

We use the following algorithm to estimate the parameter θ_0.

Step (i) For given θ, calculate

$$\begin{aligned}\begin{pmatrix} a_{is} \\ b_{is} \end{pmatrix} &= \Big[\frac{1}{T}\sum_{t=1}^{T} K_h(\theta^\top X_{its})\begin{pmatrix} 1 \\ \theta^\top X_{its}\end{pmatrix}\begin{pmatrix} 1 \\ \theta^\top X_{its}\end{pmatrix}^\top\Big]^{-1} \\ &\quad \times\Big[\frac{1}{T}\sum_{t=1}^{T} K_h(\theta^\top X_{its})\begin{pmatrix} 1 \\ \theta^\top X_{its}\end{pmatrix} Y_{it}\Big],\end{aligned} \tag{3.3}$$

for each $i = 1, \cdots, N$ and $s = 1, \cdots, T$, where $K_h(\cdot) = \frac{1}{h}K(\cdot/h)$.

Step (ii) In (3.2), replace $g_i(\theta^\top X_{is})$ and $g_i'(\theta^\top X_{is})$ with a_{is} and b_{is} from (3.3) and minimize the resulting weighted least squares with respect to θ to obtain

$$\begin{aligned}\widetilde{\theta} &= \Big[\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} K_h(\theta^\top X_{its}) b_{is}^2 X_{its} X_{its}^\top / \sum_{t=1}^{T} K_h(\theta^\top X_{its})\Big]^+ \\ &\quad \times \Big[\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} K_h(\theta^\top X_{its}) b_{is} X_{its} (Y_{it} - a_{is}) / \sum_{t=1}^{T} K_h(\theta^\top X_{its})\Big], \end{aligned} \tag{3.4}$$

where A^+ denotes the pseudo-inverse of matrix A.

Step (iii) Update θ with $\theta = \widetilde{\theta}/\|\widetilde{\theta}\|$, and then repeat *Steps (i)* and *(ii)* until θ converges.

Denote the final estimate of θ_0 by $\widehat{\theta}$. Chen *et al* (2013a) derive the asymptotic normal distribution theory for $\widehat{\theta}$ with root-NT convergence rate when both N and T tend to infinity simultaneously. They also investigate the finite sample performance of the above iterative procedure.

3.2. Partially linear single-index panel data models with fixed effects

It is well known that the nonparametric components in partially linear models may only accommodate covariates with low dimensions as they are also subject to the curse of dimensionality when the dimension of the covariates is larger than three. The partially linear single-index models have thus been introduced to address this issue in time series or cross-sectional data case (e.g. Carroll *et al*, 1997; Xia and Härdle, 2006; Liang *et al*, 2010). We next review the extension of such models to the panel data case, where the fixed effects are involved. Consider a partially linear single-index panel data model of the form (Chen *et al*, 2013b)

$$Y_{it} = Z_{it}^\top \beta_0 + \eta(X_{it}^\top \theta_0) + \alpha_i + e_{it}, \quad i = 1, \cdots, N, \quad t = 1, \cdots, T, \tag{3.5}$$

where $Z_{it} = (Z_{it,1}, \cdots, Z_{it,d})^\top$ and $X_{it} = (X_{it,1}, \cdots, X_{it,p})^\top$ are the respective d-dimensional and p-dimensional covariate vectors, β_0 and θ_0 are unknown parameters with dimensions d and p, respectively, $\eta(\cdot)$ is an unknown link function, α_i are the fixed effects which satisfy the first identification condition in (2.3), and $\{e_{it}\}$ are the random errors. The parameter θ_0 satisfies the identification condition in Section 3.1.

Model (3.5) differs from model (3.1) in the following aspects. (i) Model (3.5) has a linear component which is not accommodated in the single-index model (3.1). This linear term allows us to take into account prior information, if there is any, that some of the covariates have linear relationship with the response. (ii) The two models take into account of cross-sectional heterogeneity in different ways. The fixed effects in model (3.5) are used to capture the heterogeneity, while in model (3.1), heterogeneous link functions are used. (iii) Chen *et al* (2013a) need to assume that the time series dimension T is larger than the cross-sectional dimension N in order to achieve the root-NT convergence rate for the single-index parametric estimator. However, this assumption is not required for model (3.5).

In order to construct consistent estimators for β_0 and θ_0, we have to eliminate the fixed effects. Due to the single-index structure in model (3.5), the traditional differencing method will complicate the estimation of the link function. Hence, we introduce a local linear dummy variable estimation approach, which is motivated by the least squares dummy variable method used for parametric panel data analysis (Hsiao, 2003), and which has also been used in Section 2.

Define $\widetilde{Z} = (Z_{11}, \cdots, Z_{1T}, Z_{21}, \cdots, Z_{NT})^\top$ and

$$\eta(X, \theta) = \left(\eta(X_{11}^\top\theta), \cdots, \eta(X_{1T}^\top\theta), \eta(X_{21}^\top\theta), \cdots, \eta(X_{NT}^\top\theta)\right)^\top,$$

and let $\widetilde{Y}$, D, α and $\widetilde{e}$ be as defined in Section 2.1. Rewrite model (3.5) as

$$\widetilde{Y} = \widetilde{Z}\beta_0 + \eta(X, \theta_0) + D\alpha + \widetilde{e}. \tag{3.6}$$

For X_{it} close to x, we have the following local linear approximation:

$$\eta(X_{it}^\top\theta_0) \approx \eta(x^\top\theta_0) + \eta'(x^\top\theta_0)(X_{it} - x)^\top\theta_0,$$

where $\eta'(u)$ is the derivative of $\eta(u)$ at u. With a rationale similar to that in Section 3.1, we aim to minimize

$$\sum_{i=1}^{N}\sum_{t=1}^{T}\left[\widetilde{Y} - \widetilde{Z}\beta - D\alpha - (i_{NT}, X_{it}(\theta))\,(a_{it}, b_{it})^\top\right]^\top W_{it}\left[\widetilde{Y} - \widetilde{Z}\beta - D\alpha - (i_{NT}, X_{it}(\theta))\,(a_{it}, b_{it})^\top\right] \tag{3.7}$$

with respect to β, θ and $(a_{it}, b_{it})^\top$, where

$$X_{it}(\theta) = \left((X_{11} - X_{it})^\top\theta, \cdots, (X_{1T} - X_{it})^\top\theta, (X_{21} - X_{it})^\top\theta, \cdots, (X_{NT} - X_{it})^\top\theta\right)^\top,$$

$W_{it} = \mathsf{diag}(w_{11,it}, \cdots, w_{1T,it}, w_{21,it}, \cdots, w_{nT,it})$ is a diagonal matrix with its elements satisfying $\sum_{j=1}^{n}\sum_{s=1}^{T} w_{js,it} = 1$ for each pair (i,t), and i_{NT} is an NT-dimensional vector with all elements being 1.

To solve the minimization problem (3.7), Chen *et al* (2013b) propose the following iterative procedure:

Step (i) For given θ and β, minimize

$$\left[\widetilde{Y} - \widetilde{Z}\beta - D\alpha - (i_{NT}, X_{it}(\theta))\,(a_{it}, b_{it})^\top\right]^\top W_{it}\left[\widetilde{Y} - \widetilde{Z}\beta - D\alpha - (i_{NT}, X_{it}(\theta))\,(a_{it}, b_{it})^\top\right] \tag{3.8}$$

with respect to α to get

$$\alpha_{it} = \left(D^\top W_{it} D\right)^{-1} D^\top W_{it} \left[\widetilde{Y} - \widetilde{Z}\beta - (i_{NT}, X_{it}(\theta))\,(a_{it}, b_{it})^\top\right]. \tag{3.9}$$

Then, replace α in (3.8) by the right hand side of (3.9) and minimize the resulting weighted least squares with respect to $(a_{it}, b_{it})^\top$ to obtain the local linear estimator of $\left(\eta(X_{it}^\top\theta), \eta\prime(X_{it}^\top\theta)\right)^\top$:

$$(a_{it}, b_{it})^\top = \left(\overline{X}_{it,*}^\top(\theta) W_{it} \overline{X}_{it,*}(\theta)\right)^{-1} \overline{X}_{it,*}^\top(\theta) W_{it} \left(Y_{it,*} - Z_{it,*}\beta\right), \tag{3.10}$$

where

$$\begin{aligned}
\overline{X}_{it,*}(\theta) &= \left[I_{NT} - D\left(D^\top W_{it} D\right)^{-1} D^\top W_{it}\right](i_{NT}, X_{it}(\theta)),\\
Y_{it,*} &= \widetilde{Y} - D\left(D^\top W_{it} D\right)^{-1} D^\top W_{it}\widetilde{Y},\\
Z_{it,*} &= \widetilde{Z} - D\left(D^\top W_{it} D\right)^{-1} D^\top W_{it}\widetilde{Z}.
\end{aligned}$$

Step (ii) Substitute α and $(a_{it}, b_{it})^\top$ in (3.7) with the right hand sides of (3.9) and (3.10) for each pair of (i, t) and solve the minimization problem with respect to β and θ to obtain the updated estimates

$$(\beta^\top, \theta^\top)^\top = \begin{pmatrix} Z_*^\top W Z_* & Z_*^\top W X_* \\ X_*^\top W Z_* & X_*^\top W X_* \end{pmatrix}^{-1} \begin{pmatrix} Z_*^\top \\ X_*^\top \end{pmatrix} W\left(Y_* - A_*\right), \tag{3.11}$$

where $W = \mathsf{diag}\left(W_{11}, \cdots, W_{1T}, W_{21}, \cdots, W_{NT}\right)$,

$$\begin{aligned}
Y_* &= \left(Y_{11,*}^\top, \cdots, Y_{1T,*}^\top, Y_{21,*}^\top, \cdots, Y_{NT,*}^\top\right)^\top,\\
Z_* &= \left(Z_{11,*}^\top, \cdots, Z_{1T,*}^\top, Z_{21,*}^\top, \cdots, Z_{NT,*}^\top\right)^\top,\\
X_* &= \left(b_{11}X_{11,*}^\top, \cdots, b_{1T}X_{1T,*}^\top, b_{21}X_{21,*}^\top, \cdots, b_{NT}X_{NT,*}^\top\right)^\top,\\
X_{it,*} &= \left[I_{NT} - D\left(D^\top W_{it} D\right)^{-1} D^\top W_{it}\right]\widetilde{X}_{it},\\
\widetilde{X}_{it} &= \left((X_{11} - X_{it}), \cdots, (X_{1T} - X_{it}), (X_{21} - X_{it}), \cdots, (X_{NT} - X_{it})\right)^\top,\\
A_* &= \left(a_{11}i_{11,*}^\top, \cdots, a_{1T}i_{1T,*}^\top, a_{21}i_{21,*}^\top, \cdots, a_{nT}i_{NT,*}^\top\right)^\top,\\
i_{it,*} &= \left[I_{nT} - D\left(D^\top W_{it} D\right)^{-1} D^\top W_{it}\right] i_{NT}.
\end{aligned}$$

Step (iii) With the updated values of β and θ, repeat the above two steps until convergence.

Following the argument in Xia *et al* (2002), we use two sets of weights in the above iterative procedure. The first is a set of multidimensional kernel weights defined as

$$w_{js,it} = \frac{H\left((X_{js} - X_{it})/h_1\right)}{\sum_{j=1}^{N}\sum_{s=1}^{T} H\left((X_{js} - X_{it})/h_1\right)}, \tag{3.12}$$

where $H(\cdot)$ is a p-variate symmetric kernel function and h_1 is a bandwidth. Choosing any d-dimensional vector β and p-dimensional vector θ with $\|\theta\| = 1$ and first nonzero element being positive and following the above iterations, we can obtain initial estimators of β_0 and θ_0, which is shown to be weakly consistent by Chen *et al* (2013b). Such initial estimators of β_0 and θ_0 are denoted $\widetilde{\beta}$ and $\widetilde{\theta}$, respectively. Although being consistent, the estimators based on the p-variate kernel $H(\cdot)$ are not efficient due to the curse of dimensionality. To improve the efficiency, we use the second set of single-index weights which are defined as

$$w^{\theta}_{js,it} = \frac{K\left((X_{js} - X_{it})^{\top}\theta/h_2\right)}{\sum_{j=1}^{N}\sum_{s=1}^{T} K\left((X_{js} - X_{it})^{\top}\theta/h_2\right)}, \tag{3.13}$$

where $K(\cdot)$ is a univariate symmetric kernel function and h_2 is a bandwidth. Using the initial estimates $\widetilde{\beta}$ and $\widetilde{\theta}$ and following steps (i)–(iii) with the weights in (3.13), we then obtain the final estimators $\widehat{\beta}$ and $\widehat{\theta}$. By substituting β, θ and $X_{it}^{\top}\theta$ in (3.10) with $\widehat{\beta}$, $\widehat{\theta}$ and u, we obtain the local linear estimate of $\eta(u)$, which is denoted by $\widehat{\eta}(u)$.

Under some technical conditions, and by letting both N and T tend to infinity simultaneously, Chen *et al* (2013b) show that $\widehat{\beta}$ and $\widehat{\theta}$ have the asymptotic normal distribution with a convergence rate of root-NT, and $\widehat{\eta}(u)$ is also asymptotically normal with a convergence rate of root-NTh. They also conduct extensive simulation studies, and apply the developed model and methodology to analyze the US cigarette consumption data in 46 states and the effect of foreign direct investment on the economic growth in 22 OECD countries.

4. An empirical example

As an illustration, we analyse a cigarette demand data set using three of the four models surveyed here. The data set is from Baltagi *et al* (2000) and contains annual data from 46 states in the USA over the period 1963–1992. Baltagi *et al* (2000) used a linear dynamic panel data model

$$\ln C_{it} = \beta_0 + \beta_1 \ln C_{i,t-1} + \theta_1 \ln DI_{it} + \theta_2 \ln P_{it} + \theta_3 \ln PN_{it} + u_{it} \tag{4.1}$$

to analyze the demand for cigarettes, where $i = 1, \cdots, 46$, denotes the i–th state, $t = 1, \cdots, 29$ denotes the t–th year, C_{it} is the real per capita sales of cigarettes (measured in packs), DI_{it} is the real per capita disposable income, P_{it} is the average retail price of a pack of cigarettes measured in real terms, PN_{it} is the minimum real price of cigarettes in any neighbouring state, and the disturbance term u_{it} in (4.1) is specified as

$$u_{it} = \mu_i + \lambda_t + v_{it}, \tag{4.2}$$

where μ_i denotes a state-specific effect, and λ_t denotes a year-specific effect, which can also be interpreted as a trend in time t.

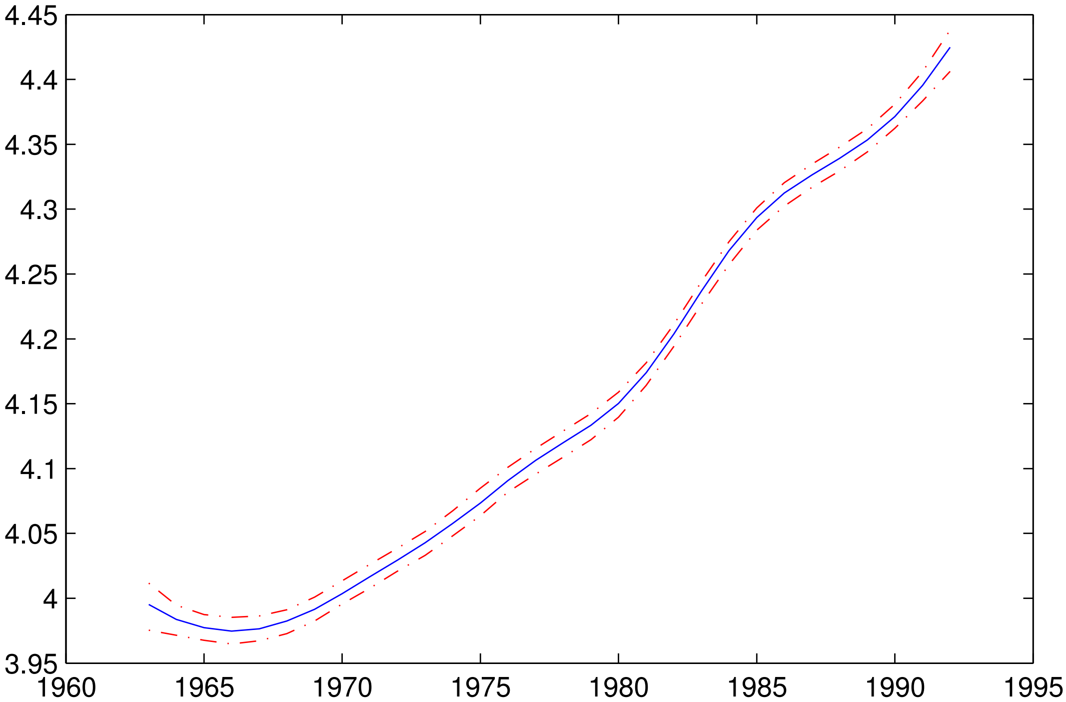

Figure 4.1. Solid line: estimated common trend in log per capita cigarette sales in 46 states of the USA from 1963–1992 using the semiparametric trending model; Dash-dotted lines: a 95% confidence band for the trend function.

Table 4.1. Parameter estimates in the semiparametric trending model for the cigar data

parameter	β_1 log disposable per-capita income	β_2 log price per pack	β_3 log min price in neighboring states
estimate (SD)	0.5133 (0.0033)	-0.9532 (0.0222)	0.0120 (0.0228)

Table 4.2. Parameter estimates in partially linear single-index model for the cigarette data

parameter	β detr. log sales in previous year	θ_1 detr. log disposable per-capita income	θ_2 detr. log price per pack	θ_3 detr. log min price in neighboring states
estimate (SD)	0.8480 (0.0073)	0.2594 (0.0217)	-0.8735 (0.0099)	0.4119 (0.0260)

We first fit the data with the semiparametric trending model (2.1)–(2.2) with $Y_{it} = \ln C_{it}$ and $X_{it} = (\ln DI_{it},\ \ln P_{it},\ \ln PN_{it})^\top$. The semiparametric profile likelihood method

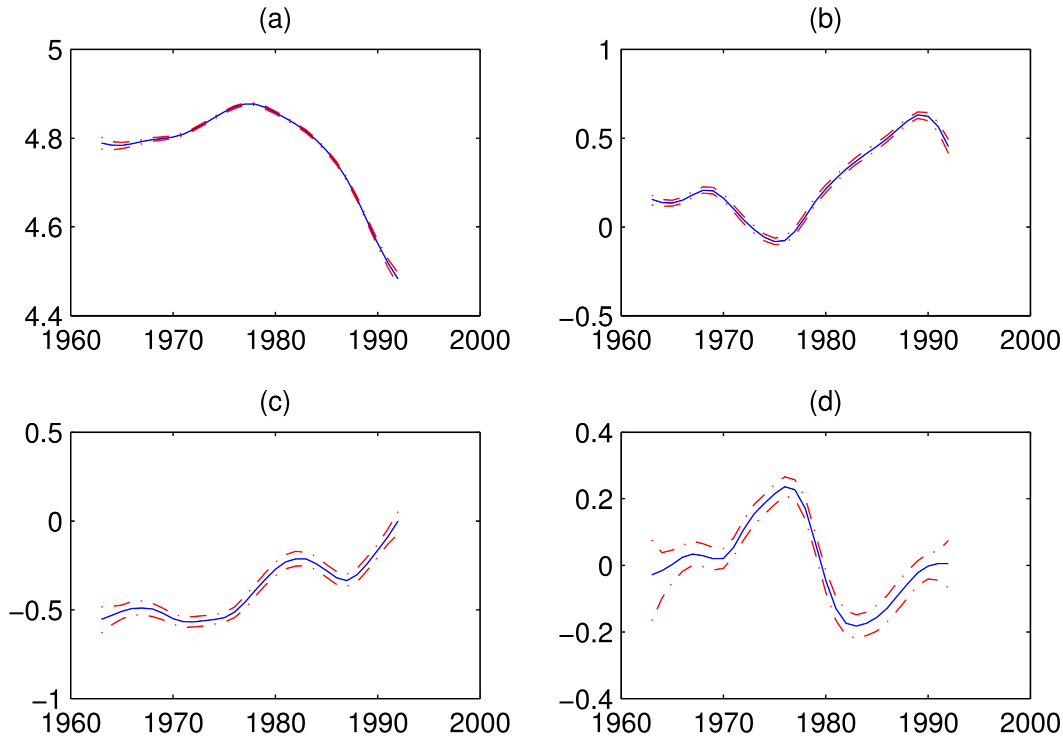

Figure 4.2. (a) Estimate of $f()$; (b) Estimate of $\beta_1()$; (c) Estimate of $\beta_2()$; (d) Estimate of $\beta_3()$. Solid lines: estimated coefficient functions in the nonparametric time-varying coefficient model for the cigar data; Dash-dotted lines: 95% confidence bands for the coefficient functions.

surveyed in Section 2.1 is used to estimate the model parameters and the trend, and the results are given in Table 4.1 and figure 4.1, respectively. With an estimate of 0.0120 and a standard error of 0.0222 for its coefficient, the variable $\ln PN_{it}$ (i.e., log minimum price in any neighbouring state) is not significant. The goodness-of-fit statistic R^2 is 0.8905 for the model.

The second model we fit to the data is the nonparametric time-varying coefficent model (2.12) with $\ln C_{it}$ as the response Y_{it} and $(\ln DI_{it},\ \ln P_{it},\ \ln PN_{it})^\top$ as the covariates. The covariate vectors X_{it} in this model are assumed to be stationary over time t, and since trends are present in each component of $(\ln DI_{it},\ \ln P_{it},\ \ln PN_{it})^\top$, we detrend each component series and let $X_{it} = (\ln DI_{it} - s_{DI}(t),\ \ln P_{it} - s_P(t),\ \ln PN_{it} - s_{PN}(t))^\top$, where $s_{DI}(t)$, $s_P(t)$, and $s_{PN}(t)$ are the nonparametric estimates of the trends in $\ln DI_{it}$, $\ln P_{it}$ and $\ln PN_{it}$. The pooled local linear estimation method is employed, and the resulting estimates of the coefficient functions and their 95% confidence bands are given in Figure 4.2. The estimate of the coefficient function for $\ln DI_{it}$ is mostly positive, and that for $\ln P_{it}$ is negative. It is also discernible from the plot at the bottom right of Figure 4.2 that the estimate of the coefficient function for $\ln PN_{it}$ has a wider confidence band compared to those of other coefficient

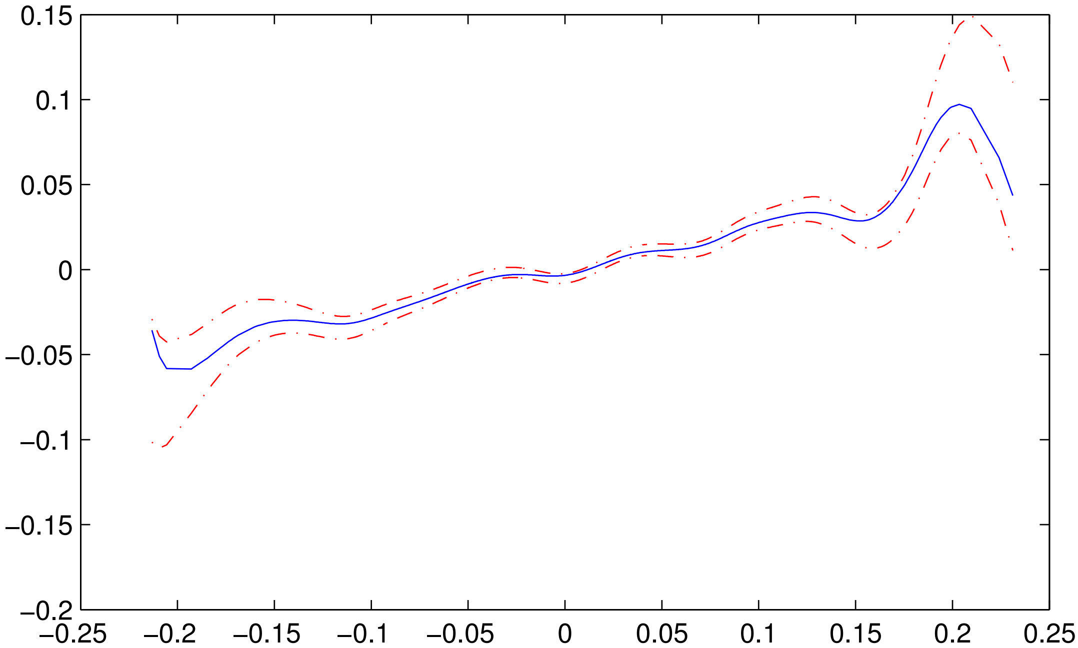

Figure 4.3. Solid line: estimated link function in the partially linear single-index model for the cigar data; Dash-dotted lines: a 95% confidence band for the link function.

functions. Such findings are in accordance with the results from the semiparametric trending model. The R^2 for the time-varying coefficient model is 0.9917.

The last model we consider for the same data set is the partially linear single-index model (3.5) surveyed in Section 3.2. Since all the variables (including the response) involved in this model are assumed to be stationary, we first detrend each data series and define $Y_{it} = \ln C_{it} - s_C(t)$, where $s_C(t)$ is the nonparametric estimate of the trend in $\ln C_{it}$, $Z_{it} = Y_{i,t-1}$, and $X_{it} = (\ln DI_{it} - s_{DI}(t),\ \ln P_{it} - s_P(t),\ \ln PN_{it} - s_{PN}(t))^\top$ as in the previous model. The reason for putting $Y_{i,t-1}$ in the linear component of model (3.5) is that it exhibits the strongest linearity with Y_{it}. The estimates of the parameters and the link function are presented in Table 4.2 and Figure 4.3, respectively. The signs of the parameter estimates echo those of the corresponding parameter estimates from the semiparametric trending model. The R^2 of the partially linear single-index model is 0.9698.

5. Conclusions and discussions

In this paper, we review the recent developments on two types of modelling frameworks: nonparametric and semiparametric panel data models with deterministic trends, and semiparametric single-index panel data models with individual effects. In Section 2, we review the semiparametric trending panel data models and nonparametric time-varying coefficient panel data models as well as their estimation procedures. Both the time series length and cross-sectional size in this article are allowed to be very large and the panel data could be cross-sectionally dependent. In Section 3, we review the single-index panel data models with heterogeneous link functions and the partially linear single-index panel data models with fixed effects, and the extension of the minimum average variance estimation method to the panel data case. We next discuss some possible future topics which are closely related to those introduced in Sections 2 and 3.

In Section 2, we mainly focus on nonlinear panel data models with deterministic trends by using nonparametric and semiparametric methods. A possible future topic is to consider stochastic trends in panel data (Phillips and Moon, 1999; Bai *et al*, 2009). For example, the regressors $\{X_{it}\}$ can be generated by a multivariate unit root process and we may expect some super-consistency asymptotic results for the developed estimation. Another possible extension is to replace the fixed effects in models (2.1) and (2.12) by a more complicated interactive fixed effects structure (Bai, 2009). However, the estimation procedures in Section 2 may have to be substantially extended to deal with such a complicated structure.

In Section 3, to establish nice asymptotic properties for the semiparametric estimators, we need to assume that both the time series length and cross-sectional size tend to infinity. A recent paper by Chen *et al* (2013b) relax this condition for the partially linear single-index panel data models and allow one of the two indices to be fixed and the other tend to infinity. Chen *et al* (2013b) also introduce a new semiparametric GEE method to improve the efficiency of the parametric estimation. Other possible extensions of the partially linear single-index panel data model (3.5) include models with discrete response variables or those with high-dimensional covariates, in which case a variable selection procedure is needed (Liang *et al*, 2010).

References

Arellano, M., 2003. *Panel data Econometics.* Oxford University Press: New York.

Atak, A., Linton, O., Xiao, Z., 2011. A semiparametric panel model for unbalanced data with application to climate change in the United Kingdom. *Journal of Econometrics* **164**, 92–115.

Bai, J., 2009. Panel data models with interactive fixed effects. *Econometrica* **77**,1229–1279.

Bai, J., Kao, C., Ng, S., 2009. Panel cointegration with global stochastic trends. *Journal of Econometrics* **149**, 82–99.

Baltagi, B. H., 1995. *Econometric Analysis of Panel Data.* New York: John Wiley & Sons.

Baltagi, B. H., Griffin, J. M. and Xiong, W. (2000). To pool or not to pool: Homogenous versus heterogenous estimators applied to cigarette demand. *Review of Economics and Statistics* **82**, 117-126.

Cai, Z., 2007. Trending time-varying coefficient time series models with serially correlated errors. *Journal of Econometrics* **136**, 163–188.

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

Carroll, R. J., Fan, J. Gijbels, I., Wand, M., 1997. Generalized partially linear single-index models. *Journal of the American Statistical Association* **92**, 477–489.

Chen, J., Gao, J., Li, D., 2012a. A new diagnostic test for cross-section uncorrelatedness in nonparametric panel data models. *Econometric Theory* **28**, 1144–1163.

Chen, J., Gao, J., Li, D., 2012b. Semiparametric trending panel data models with cross-sectional dependence. *Journal of Econometrics* **171**, 71–85.

Chen, J., Gao, J., Li, D., 2013a. Estimation in single-index panel data models with heterogeneous link functions. *Econometric Reviews* **32**, 928–955.

Chen, J., Gao, J., Li, D., 2013b. Estimation in partially linear single-index panel data models with fixed effects. *Journal of Business and Economic Statistics* **31**, 315–330.

Fan, J., Yao, Q., 2003. *Nonlinear Time Series: Nonparametric and Parametric Methods.* Springer, New York.

Gao, J., 2007. *Nonlinear Time Series: Semiparametric and Nonparametric Methods.* Chapman & Hall/CRC, London.

Gao, J., Hawthorne, K., 2006. Semiparametric estimation and testing of the trend of temperature series. *Econometrics Journal* **9**, 332–355.

Härdle, W., Hall, P., Ichimula, H., 1993. Optimal smoothing in single-index models. *Annals of Statistics* **21**, 157–178.

Henderson, D., Carroll, R., Li, Q., 2008. Nonparametric estimation and testing of fixed effects panel data models. *Journal of Econometrics* **144**, 257–275.

Hjellvik, V., Chen, R., Tjøstheim, D., 2004. Nonparametric estimation and testing in panels of intercorrelated time series. *Journal of Time Series Analysis* **25**, 831–872.

Hsiao, C., 2003. *Analysis of Panel Data.* Cambridge University Press, Cambridge.

Li, D., Chen, J., Gao, J., 2011a. Non-parametric time-varying coefficient panel data models with fixed effects. *Econometrics Journal* **14**, 387–408.

Li, D., Chen, J., Lin, Z., 2009. Variable selection in partially time-varying coefficient models. *Journal of Nonparametric Statistics* **21**, 553–566.

Li, D., Chen, J., Lin, Z., 2011b. Statistical inference in partially time-varying coefficient models. *Journal of Statistical Planning and Inference* **141**, 995–1013.

Li, Q., Racine, J., 2007. *Nonparametric Econometrics: Theory and Practice.* Princeton: Princeton University Press.

Liang, H., Liu, X., Li, R., Tsai, C., 2010. Estimation and testing for partially linear single-index models. *Annals of Statistics* **38**, 3811–3836.

Mammen, E., Støve, B., Tjøstheim, D., 2009. Nonparametric additive models for panels of time series. *Econometric Theory* **25**, 442–81.

Phillips, P. C. B., 2001. Trending time series and macroeconomic activity: some present and future challengers. *Journal of Econometrics* **100**, 21–27.

Phillips, P. C. B., 2010. The mysteries of trend. *Macroeconomic Review* **IX**, 82–89.

Phillips, P. C. B., Moon, H., 1999. Linear regression limit theory for nonstationary panel data. *Econometrica* **67**, 1057-1111.

Robinson, P. M., 1989. Nonparametric estimation of time-varying parameters. *Statistical Analysis and Forecasting of Economic Structural Change*. Hackl, P. (Ed.). Springer, Berlin, pp. 164-253.

Robinson, P. M., 2012. Nonparametric trending regression with cross-sectional dependence. *Journal of Econometrics* **169**, 4–14.

Su, L., Lu, X., 2013. Nonparametric dynamic panel data models: kernel estimation and specification testing. Forthcoming in *Journal of Econometrics*.

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

Su, L., Ullah, A., 2010. Nonparametric and semiparametric panel econometric models: estimation and testing. Working paper.

Ullah, A., Roy, N., 1998. Nonparametric and semiparametric econometrics of panel data. In *Handbook of Applied Economics and Statistics* edited by A. Ullah and D. E. A. Giles, 579–604. New York: Marcel Dekker.

Wu, H., Zhang, J., 2006. *Nonparametric Regression Methods for Longitudinal Data Analysis: Mixed-Effects Modeling Approaches*. John Wiley & Sons.

Xia, Y., Härdle, W., 2006. Semiparametric estimation of partially linear single-index models. *Journal of Multivariate Analysis* **97**, 1162–1184.

Xia, Y., Tong, H., Li, W. K., Zhu, L., 2002. An adaptive estimation of dimension reduction space. *Journal of the Royal Statistical Society Series B* **64**, 363–410.