具有截面相关的变系数异质面板数据模型的估计[①]

徐秋华　方颖

（厦门大学王亚南经济研究院）

【摘要】 本文建立了一类变系数异质面板数据模型。此模型假设自变量的系数是某一平滑变量的未知函数；允许自变量、平滑变量和误差项存在通过共同因子结构引入的截面相关。由于一些共同因子的不可观测性，本文采用局部线性共同相关效应估计方法对未知的函数进行估计并证明了估计量的渐近正态性质。蒙特卡洛模拟结果表明该估计方法具有良好的小样本性质。

关键词　面板数据模型　截面相关　变系数　局部线性估计　共同相关效应

中图分类号　　　　　　　　　　　　　**文献标识码**

Estimation of Varying-coefficient Heterogeneous Panel Data Models with Cross-sectional Dependence

Abstract: In this paper, we propose a varying-coefficient heterogeneous panel data model, allowing for cross-sectional dependence through a common factors structure. The local linear common correlated effect estimation technique is applied to estimating the heterogeneous varying coefficients and the asymptotic normality of these proposed estimators are established. Monte Carlo simulations demonstrate good finite sample performances for these estimators.

Key words: Panel Data Models; Cross-sectional Dependence; Varying Coefficient; Local Linear Estimation; Common Correlated Effects

一、引言

近年来面板数据模型在经济学和金融学等领域得到了广泛的应用。与横截面或时间序列数据相比，面板数据可以有效地控制个体异质性、提高自由度、提取横截面或时间序列数据无法获得的动态信息。然而，传统的面板数据模型假设各截面个体是独立同分布的（Arellano，2003、Hsiao，2003）。随着全球化进程的迅猛发展，如今不同国家或地区之间的联系日益加深。在使用面板数据研究经济问题时，我们

[①] 本文受到国家自然科学基金（70971113、71131008）与教育部新世纪优秀人才支持计划的资助。

不得不考虑截面相关问题。这使得对具有截面相关的面板数据模型进行估计和检验成为面板数据模型理论研究的热点之一。在文献中，处理截面相关的方法主要有两种：一种方法是使用空间权重矩阵外生设定截面个体之间的相关结构。空间权重矩阵由空间距离等地理变量或经济指标决定。例如，王火根和沈利生（2007）使用空间面板回归模型研究中国各省市区经济增长和能源消费的关系；王锐淇（2012）分析了我国区域技术创新能力空间相关性及扩散效应；张志强（2012）比较了 GMM、QML 等空间面板参数估计方法的小样本特性。空间相关是一种随“距离”增加而减弱的相关，通常被称为是截面弱相关；另一种方法假设回归方程的误差项包含影响所有截面个体的共同因子。在宏观经济学中，共同因子可以用来刻画影响所有个体的共同冲击。钱金保和钱彬彬（2012）使用含有共同因子的面板数据模型分析了共同冲击对广东省出口的影响。具有共同因子结构的截面相关通常被称为截面强相关，含有此类相关结构的面板数据模型引起了学者们的广泛兴趣。Bai（2009）考虑了具有交互固定效应（interactive fixed effects）的面板数据模型。所谓交互固定效应是指回归方程的误差项包含了有限多个不可观测的共同因子，而且不同的截面个体可以具有不同的因子载荷（factor loading）。因为经典面板数据模型中的个体效应和时间效应是交互固定效应的特例，所以这种模型设定更具有一般性。由于共同因子是不可观测的，Bai（2009）基于主成分分析的方法得到了共同因子张成的空间和自变量系数的一致估计。Moon 和 Weidner（2013a, b）将上述模型推广到动态面板模型，并使用准极大似然估计法（Quasi Maximum Likelihood, QMLE）得到了自变量系数的一致估计。上述文献都假设自变量的系数是同质的，即不同截面个体具有相同的斜率系数。Song（2013）将 Moon 和 Weidner 的研究推广到具有交互固定效应的动态异质面板数据模型。Pesaran（2006）考虑了具有交互固定效应的静态异质面板数据模型，他使用可观测变量的截面样本均值代替不可观测的共同因子，自变量系数的一致估计可通过普通最小二乘法得到。这种估计方法被称为共同相关效应（Common Correlated Effects, CCE）估计法。Chudik 和 Pesaran（2014）将 CCE 方法推广到动态异质面板数据模型。

上述模型都假设自变量的系数不随时间变化。但是，个体偏好、技术进步和政治制度等因素的变化会导致经济结构的改变。这种情况在时间维度较大的长面板数据中极有可能发生。这使得具有系数不变假定的线性面板数据模型可能存在模型误设，而基于误设模型得到的估计量通常是不一致的。为解决模型误设问题，一些学者提出了非参数或半参数的面板数据模型。Cai 和 Li（2008）研究了变系数动态面板数据模型。此模型假设自变量的系数是某一外生变量的未知平滑函数；Li、Chen 和 Gao（2011）研究了具有固定效应的非参数时变系数（time-varying coefficient）面板数据模型。他们假设自变量的系数是时间 t 的函数；Su 和 Lu（2013）讨论了具有固定效应的非参数动态面板数据模型的估计与检验；Lee（2013）考虑了具有固定效应的非参数自回归面板数据模型的估计问题。这四篇文献中的估计量的渐近性质成立的前提是各截面个体之间是独立同分布的，所以并不适用于分析具有截面相关的面板数据。Su 和 Jin（2012）将 Pesaran（2006）的模型推广到非参数静态异质面板数据模型并提出了基于 CCE 估计思想的 Sieve 估计；Huang（2013）研究了具有

固定效应和共同因子结构的面板数据模型并提出了基于 CCE 估计思想的局部线性估计方法；Su 和 Zhang（2013）研究了具有交互固定效应的非参数动态面板数据模型并结合 Sieve 估计和主成分分析方法对模型进行了估计和检验。Su 和 Zhang（2013）假设模型中非参数部分的未知函数是同质的，即不同截面个体具有相同的函数形式；并且其估计方法需要假设共同因子的个数是已知的。虽然 Su 和 Jin（2012）的模型假设了异质的未知函数并且不需要假设共同因子的个数是已知的，但是纯粹的非参数模型往往缺乏经济学解释。Huang（2013）除了使用可观测变量的截面样本均值代替不可观测的共同因子，还加入了时间维度的样本均值解决模型识别问题。但是其最后估计的模型实际上假设了不同截面个体具有相同的因子载荷。

基于上述分析，本文将建立具有共同因子结构的变系数异质面板数据模型。变系数模型是经典参数模型的自然推广。通过允许自变量的斜率系数是某些经济变量或时间的函数，变系数模型具有良好的经济学解释力和刻画非线性关系的能力。例如，方颖和郭萌萌（2009）使用时变系数模型检验中国主要宏观变量的稳定性；蔡宗武等（2012）运用函数系数模型来分析人民币对美元汇率收益率与波动率的非线性特征。Cai（2010）详细地综述了变系数模型在经济和金融领域的应用。本文的结构安排如下：第二部分建立变系数异质面板数据模型并给出此模型的估计方法；第三部分给出了变系数估计量的渐近分布；第四部分通过蒙特卡洛模拟观察估计量的小样本性质；最后对全文进行总结。所有数学证明在附录中给出。

二、模型设定和估计方法

1. 模型设定

令 y_{it} 表示第 i 个截面个体的被解释变量在时刻 t 的观测值，具有共同因子结构的变系数异质面板数据模型具有如下形式：

$$\begin{aligned} y_{it} &= \beta_i^{\mathrm{T}}(u_{it})x_{it} + \gamma_{1i}^{\mathrm{T}} f_{1t} + e_{it} \\ e_{it} &= \gamma_{2i}^{\mathrm{T}} f_{2t} + \varepsilon_{it} \end{aligned} \tag{1}$$

$i=1,\ldots,N$，$t=1,\ldots,T$。其中 x_{it} 是 $p\times 1$ 的自变量；f_{1t} 是 $m_1\times 1$ 的可观测共同因子；f_{2t} 是 $m_2\times 1$ 的不可观测共同因子；γ_{1i} 是 $m_1\times 1$ 的可观测共同因子的因子载荷；γ_{2t} 是 $m_2\times 1$ 的不可观测共同因子的因子载荷；ε_{it} 是误差项；$u_{it}\in\mathbb{R}$ 是平滑变量（smooth variable）[1]；$\beta_i(\cdot)$ 是定义在实数集 $\mathbb{R}$ 上的未知平滑函数，具有连续二阶

[1] 这里为了简化符号，我们假设 $u_{it}\in\mathbb{R}$。本文的研究可以容易地推广到 u_{it} 是多维向量的情况。由于非参数估计存在“维度诅咒”问题，在实际应用中 u_{it} 的维度通常很低。

导数，并且对不同的截面个体 i， $\beta_i(\cdot)$ 可以具有不同的函数形式。一般地，不可观测的共同因子除了影响被解释变量以外，还可能对自变量和平滑变量造成影响。为了允许这种情况，我们假设 x_{it} 和 u_{it} 满足如下形式：

$$z_{it}=\begin{pmatrix} x_{it} \\ u_{it} \end{pmatrix}=\Gamma_{1i}^{\mathrm{T}} f_{1t}+\Gamma_{2i}^{\mathrm{T}} f_{2t}+v_{it} \tag{2}$$

其中 Γ_{1i} 是 $m_1\times(p+1)$ 的可观测共同因子的因子载荷；Γ_{2i} 是 $m_2\times(p+1)$ 的不可观测共同因子的因子载荷。

模型（1）具有一般性，许多文献中的面板数据模型可以看作是上述模型的特例：① 如果对所有 i，$\beta_i(u_{it})=\beta$、 $f_{1t}=1$ 且 $\gamma_{2i}=0$，模型（1）就是传统的具有固定效应的面板数据模型；② Pesaran（2006）的模型对应于 $\beta_i(u_{it})=\beta_i$；③ 如果对所有 i，$\beta_i(u_{it})=\beta$ 且 $\gamma_{1i}=0$，模型（1）简化为 Bai（2009）中的模型；④ 如果 $x_{it}=1$，模型（1）成为 Su 和 Jin（2012）、Huang（2013）中的模型。

2. 模型估计

我们感兴趣的是如何估计模型（1）中的系数函数 $\beta_i(u_{it})$。基于 Pesaran（2006）中提出的 CCE 估计思想，我们可以使用可观测的共同因子以及自变量和平滑变量的截面样本均值代替不可观测的共同因子：令 $\bar{z}_t=N^{-1}\sum_{i=1}^{N} z_{it}$，类似地，定义 Γ_{1i}、Γ_{2i} 和 v_{it} 的截面样本均值 $\bar{\Gamma}_1$、$\bar{\Gamma}_2$ 和 $\bar{v}_t$。由模型（2），我们可以得到：

$$\bar{z}_t=\bar{\Gamma}_1^{\mathrm{T}} f_{1t}+\bar{\Gamma}_2^{\mathrm{T}} f_{2t}+\bar{v}_t \tag{3}$$

如果当 N 足够大时，$rank(\bar{\Gamma}_2)=m_2\le p+1$，则 f_{2t} 可表示为：

$$f_{2t}=\left(\bar{\Gamma}_2\bar{\Gamma}_2^{\mathrm{T}}\right)^{-1}\bar{\Gamma}_2\left(\bar{z}_t-\bar{\Gamma}_1^{\mathrm{T}} f_{1t}-\bar{v}_t\right) \tag{4}$$

给定 t，当 $N\to\infty$ 时，$\bar{v}_t\to_p 0$。所以，当 $N\to\infty$ 时，我们有：

$$f_{2t}-\left(\bar{\Gamma}_2\bar{\Gamma}_2^{\mathrm{T}}\right)^{-1}\bar{\Gamma}_2\left(\bar{z}_t-\bar{\Gamma}_1^{\mathrm{T}} f_{1t}\right)\to_p 0 \tag{5}$$

上式表明，当 N 足够大时，我们可以用 $\bar{z}_t$ 和 f_{1t} 的线性组合代替 f_{2t}。将（5）中的

线性组合代入模型（1），经过整理便得到如下我们要估计的模型：

$$y_{it}=\beta_i^{\mathrm{T}}(u_{it})x_{it}+\vartheta_i^{\mathrm{T}}q_t+\varepsilon_{it}^{*} \tag{6}$$

其中 $q_t=\begin{pmatrix} f_{1t} \\ \bar{z}_t \end{pmatrix}$；$\varepsilon_{it}^{*}$ 包括模型（1）中的误差项 ε_{it} 以及用 $\bar{z}_t$ 和 f_{1t} 的线性组合代替 f_{2t} 导致的近似误差。模型（6）具有半参数变系数部分线性模型的形式。在文献中，此类模型通常采用两步估计法，从而使线性参数部分的估计量具有更快的收敛速度。然而，在我们的模型中，ϑ_i 被看作是冗余参数（nuisance parameters），我们将借助于参数子集回归（partitioned regression）的思想对模型（6）进行估计。由于对不同的 i，$\beta_i(u_{it})$ 可以具有不同的函数形式，所以在估计模型（6）时，我们实际上考虑的是固定 i 后得到的时间序列。

令 $y_i=\left(y_{i1},\ldots,y_{iT}\right)^{\mathrm{T}}$、$Q=\left(q_1,\ldots,q_T\right)^{\mathrm{T}}$ 和 $\varepsilon_i^{*}=\left(\varepsilon_{i1}^{*},\ldots,\varepsilon_{iT}^{*}\right)^{\mathrm{T}}$，将（6）写成如下向量形式：

$$y_i=\begin{pmatrix} \beta_i^{\mathrm{T}}(u_{i1})x_{i1} \\ \vdots \\ \beta_i^{\mathrm{T}}(u_{iT})x_{iT} \end{pmatrix}+Q\vartheta_i+\varepsilon_i^{*} \tag{7}$$

为了估计（7）中的未知函数 $\beta_i(\cdot)$，我们采用局部线性估计方法。为此，令 h 表示窗宽（bandwidth）。给定 $u_0\in\mathbb{R}$，若 $\left|u_{it}-u_0\right|\le h$，$t=1,\ldots,T$，由一阶 Taylor 展开可以得到：

$$\begin{pmatrix} \beta_i^{\mathrm{T}}(u_{i1})x_{i1} \\ \vdots \\ \beta_i^{\mathrm{T}}(u_{iT})x_{iT} \end{pmatrix}=\tilde{X}_i\beta_i^{*}(u_0)+O_p(h^2) \tag{8}$$

其中 $\tilde{X}_i$ 是 $T\times 2p$ 的矩阵，其第 t 行是 $(x_{it}^{\mathrm{T}},x_{it}^{\mathrm{T}}(\frac{u_{it}-u_0}{h}))$，$t=1,\ldots,T$；$\beta_i^{*}(u_0)=\begin{pmatrix} \beta_i(u_0) \\ h\beta_i^{'}(u_0) \end{pmatrix}$。$\beta_i^{*}(u_0)$ 的估计量 $\hat{\beta}_i^{*}(u_0)$ 可以通过最小化如下加权平方和得到：

$$\sum_{t=1}^{T}[y_{it}-(x_{it}^{\mathrm{T}},x_{it}^{\mathrm{T}}(\frac{u_{it}-u_0}{h}))\beta_i^{*}(u_0)-\vartheta_i^{\mathrm{T}}q_t]^2k_h(u_{it}-u_0) \tag{9}$$

其中$k_h(u_{it}-u_0)=k(\dfrac{u_{it}-u_0}{h})/h$，$k(\cdot)$是核函数。容易看出，最小化（9）等价于对回归方程：

$$w_{it}y_{it}=w_{it}(x_{it}^{\mathrm{T}},x_{it}^{\mathrm{T}}(\frac{u_{it}-u_0}{h}))\beta_i^*(u_0)+w_{it}q_t^{\mathrm{T}}\vartheta_i+w_{it}\varepsilon_{it}^* \tag{10}$$

进行普通最小二乘估计，其中$w_{it}=\sqrt{k_h(u_{it}-u_0)}$。若采用（7）的向量形式，上式可以记为：

$$W_i^{1/2}y_i=W_i^{1/2}\tilde{X}_i\beta_i^*(u_0)+W_i^{1/2}Q\vartheta_i+W_i^{1/2}\varepsilon_i^* \tag{11}$$

其中$W_i=diag(w_{i1}^2,\ldots,w_{iT}^2)$。定义$M_i=I_T-W_i^{1/2}Q(Q^{\mathrm{T}}W_iQ)^{-1}Q^{\mathrm{T}}W_i^{1/2}$，$I_T$是$T\times T$单位阵。将（11）两端同时左乘以$M_i$，我们有：

$$M_iW_i^{1/2}y_i=M_iW_i^{1/2}\tilde{X}_i\beta_i^*(u_0)+M_iW_i^{1/2}\varepsilon_i^* \tag{12}$$

由（12）的普通最小二乘估计可得：

$$\hat{\beta}_i^*(u_0)=(\tilde{X}_i^{\mathrm{T}}W_i^{1/2}M_iW_i^{1/2}\tilde{X}_i)^{-1}\tilde{X}_i^{\mathrm{T}}W_i^{1/2}M_iW_i^{1/2}y_i \tag{13}$$

我们将（13）中得到的估计量称为共同相关效应局部线性（LCCE）估计量。

三、渐近性质

1. 符号和假设

为了下文叙述和推导的方便，我们先引入一些符号。令C表示一个一般的有限的常数，在不同处出现时可以取不同的值；对于实矩阵A，A^-表示A的广义逆；定义$\|A\|=[\mathrm{tr}(AA^{\mathrm{T}})]^{1/2}$，其中$\mathrm{tr}(\cdot)$表示矩阵的迹。IIDN 和 IIDU 分别表示正态独立同分布和均匀独立同分布。令$\mu_j=\int_{-\infty}^{\infty}u^jk(u)du$、$\lambda_j=\int_{-\infty}^{\infty}u^jk^2(u)du$、$\eta_t=\begin{pmatrix}f_{1t}\\ f_{2t}\end{pmatrix}$、$\Gamma=\begin{pmatrix}I_{m_1} & \bar{\Gamma}_1\\ 0 & \bar{\Gamma}_2\end{pmatrix}$和$b_t=\Gamma^{\mathrm{T}}\eta_t$。令$(x_i,u_i,v_i,b,q,f_2)$和$(x_{it},u_{it},v_{it},b_t,q_t,f_{2t})$具有相同的分布，$f(u_i,\xi_i)$表示$(u_i,\xi_i)$的联合概率密度函数，其中$\xi_i=x_i,q,b,v_i,f_2$和$(x_i,q)$。$f_u(u_i)$表示$u_i$的边际概率密度函数，$f_{1t}(u_{i1},u_{it})$是$u_{i1}$和$u_{it}$的联合概率密度函数。令$\theta_i=(x_i,u_i,q)$，再定义如下记号：

$$\sigma_{1t}\left(\theta_{i1},\theta_{it}\right)=E(\varepsilon_{i1}\varepsilon_{it}\mid\theta_{i1},\theta_{it})$$
$$\Lambda_{it}^{\omega_1\omega_2}(u_{i1},u_{it})=E(\omega_{11}\omega_{2t}^{\mathrm{T}}\sigma_{1t}\left(\theta_{i1},\theta_{it}\right)\mid u_{i1},u_{it})=E(\omega_{11}\omega_{2t}^{\mathrm{T}}\varepsilon_{i1}\varepsilon_{it}\mid u_{i1},u_{it})$$
$$\Omega_{i\omega_1\omega_2}(u_0)=E(\omega_1\omega_2^{\mathrm{T}}\mid u_i=u_0)$$
$$\Omega_{i\omega_1\omega_2}^{*}(u_0)=E(\omega_1\omega_2^{\mathrm{T}}\sigma_{11}\left(\theta_i,\theta_i\right)\mid u_i=u_0)$$

其中$\omega_1,\omega_2\in\{x_i,q\}$，令$\Omega(u_0)=\Omega_{ix_ix_i}(u_0)-\Omega_{ix_iq}(u_0)\Omega_{iqq}^{-1}(u_0)\Omega_{ix_iq}^{\mathrm{T}}(u_0)$。

为了推导 LCCE 估计量的渐近性质，我们需要对模型（1）和（2）施加如下假设：

假设 1：核函数$k(\cdot)$是支撑为[-1, 1]的有界概率密度函数且满足$\mu_1=0$。

假设 2：（a）给定 *i*，$\{(\varepsilon_{it},v_{it}):t\geq1\}$是严格平稳的$\alpha$-混合序列，其混合系数（mixing coefficient）$\alpha_i(j)$满足$\sum_{j=1}^{\infty}j^c[\alpha_i(j)]^{1-2/\delta}<\infty$，其中$\delta>2$，$c>1-2/\delta$。$E(\varepsilon_{it}\mid x_{it},u_{it},q_t)=0$。令$\varepsilon_i=\left(\varepsilon_{i1},\ldots,\varepsilon_{iT}\right)^{\mathrm{T}}$和$v_i=\left(v_{i1},\ldots,v_{iT}\right)^{\mathrm{T}}$，$(\varepsilon_i,v_i)$无截面相关且$E(v_i)=0$；（b）共同因子$\{(f_{1t},f_{2t}):t\geq1\}$是严格平稳的$\alpha$-混合序列，其混合系数$\alpha_0(j)$满足$\sum_{j=1}^{\infty}j^c[\alpha_0(j)]^{1-2/\delta}<\infty$。对于所有的 *i*，*t* 和 *s*，$(f_{1t},f_{2t})$与$v_{is}$独立；（c）$\sup_{N\geq1}\max_{1\leq i\leq N}E\left|\zeta_i\right|^{2\delta}<\infty$，其中$\zeta_i=f_{11},f_{21}$和$v_{i1}$。

假设 3：（a）因子载荷γ_{2i}，Γ_{1i}和Γ_{2i}满足独立同分布且具有有限2δ阶矩；（b）对于所有的 *j* 和 *t*，$(\gamma_{2i},\Gamma_{2i})$与$v_{jt}$和$(f_{1t},f_{2t})$独立；（c）对于所有的 *j* 和 *t*，$\Gamma_{1i}$与$\Gamma_{2j}$、$v_{jt}$和$(f_{1t},f_{2t})$独立；（d）$rank(\Gamma_2)=m_2\leq p+1$，其中$\Gamma_2=E(\Gamma_{2i})$。

假设 4： $f(u_i\mid\xi_i)\leq C<\infty$，其中$f(u_i\mid\xi_i)$表示给定$\xi_i$时$u_i$的条件概率密度函数且$\xi_i=x_i,q,b,v_i,f_2$和$(x_i,q)$；对任意的$\tau\geq1$，$f(u_{i0},u_{i\tau}\mid\xi_{i0},\xi_{i\tau})\leq C<\infty$，其中$f(u_{i0},u_{i\tau}\mid\xi_{i0},\xi_{i\tau})$表示给定$(\xi_{i0},\xi_{i\tau})$时$(u_{i0},u_{i\tau})$的条件概率密度函数；对任意 *t*，$\Lambda_{it}^{\omega_1\omega_2}(u_{i1},u_{it})$和$f_{1t}(u_{i1},u_{it})$在$(u_{i1}=u_0,u_{it}=u_0)$处连续。

结合假设 2 和假设 3（a），我们可知$\{(x_{it},u_{it}):t\geq1\}$和$\{q_t:t\geq1\}$都是$\alpha$-混合

序列，且它们的混合系数为$\bar{\alpha}_i(j) \equiv \max\{\alpha_i(j), \alpha_0(j)\}$的；对任意的 t，x_{it} 和 q_t 都具有有限的2δ阶矩。

假设 5：（a）$h \to 0$，$Th \to \infty$，$N \to \infty$。存在正整数s_T满足：当$T \to \infty$时，$s_T \to \infty$，$s_T = o(\sqrt{Th})$和$(T/h)^{1/2}\bar{\alpha}_i(s_T) \to 0$；

（b）存在$\delta^* > \delta$使得$E\left(\left\|x_i\varepsilon_{i1}\right\|^{\delta^*} \mid u_i\right)$和$E\left(\left\|q\varepsilon_{i1}\right\|^{\delta^*} \mid u_i\right)$在$u_i = u_0$处连续，并且$\bar{\alpha}_i(\tau) = O(\tau^{-\theta^*})$，其中$\theta^* \geq \delta\delta^* / \{2(\delta^* - \delta)\}$；

（c）$T^{1/2-\delta/4}h^{\delta/\delta^*-1/2-\delta/4} = O(1)$。

2. 渐近正态

对$\beta_i(u_{it})$进行二阶 Taylor 展开，可以得到：

$$\beta_i(u_{it}) = \beta_i(u_0) + h\beta_i^{'}(u_0)(\frac{u_{it}-u_0}{h}) + \frac{h^2}{2}\beta_i^{''}(u_0)(\frac{u_{it}-u_0}{h})^2 + R(u_{it}, u_0) \qquad (A.1)$$

其中$R(u_{it}, u_0)$是 Taylor 展开的余项。将（A.1）代入（1）并把截面个体 i 的所有观测值写成向量形式，我们有：

$$y_i = \breve{X}_i\beta_i^*(u_0) + \frac{h^2}{2}A_i\beta_i^{''}(u_0) + D_i + F_1\gamma_{1i} + F_2\gamma_{2i} + \varepsilon_i \qquad (A.2)$$

其中

$$A_i = \begin{pmatrix} x_{i1}^{\mathrm{T}}(\frac{u_{i1}-u_0}{h})^2 \\ \vdots \\ x_{iT}^{\mathrm{T}}(\frac{u_{iT}-u_0}{h})^2 \end{pmatrix}, \quad D_i = \begin{pmatrix} x_{i1}^{\mathrm{T}}R(u_{i1}, u_0) \\ \vdots \\ x_{iT}^{\mathrm{T}}R(u_{iT}, u_0) \end{pmatrix}$$

$$F_1 = \left(f_{11}, \ldots, f_{1T}\right)^{\mathrm{T}}, \quad F_2 = \left(f_{21}, \ldots, f_{2T}\right)^{\mathrm{T}}$$

将（A.2）代入（13），我们有：

$$\begin{aligned} \hat{\beta}_i^*(u_0) - \beta_i^*(u_0) = &\frac{h^2}{2}(\breve{X}_i^{\mathrm{T}}W_i^{1/2}M_iW_i^{1/2}\breve{X}_i)^{-1}\breve{X}_i^{\mathrm{T}}W_i^{1/2}M_iW_i^{1/2}A_i\beta_i^{''}(u_0) \\ &+(\breve{X}_i^{\mathrm{T}}W_i^{1/2}M_iW_i^{1/2}\breve{X}_i)^{-1}\breve{X}_i^{\mathrm{T}}W_i^{1/2}M_iW_i^{1/2}D_i \\ &+(\breve{X}_i^{\mathrm{T}}W_i^{1/2}M_iW_i^{1/2}\breve{X}_i)^{-1}\breve{X}_i^{\mathrm{T}}W_i^{1/2}M_iW_i^{1/2}F_2\gamma_{2i} \\ &+(\breve{X}_i^{\mathrm{T}}W_i^{1/2}M_iW_i^{1/2}\breve{X}_i)^{-1}\breve{X}_i^{\mathrm{T}}W_i^{1/2}M_iW_i^{1/2}\varepsilon_i \end{aligned} \qquad (A.3)$$

由（A.3）可以看出，我们将$\hat{\beta}_i^*(u_0)$和其真实值之差表示为四项的和。第一项是局

部线性回归的偏误项；第二项包含了 Taylor 展开的余项；第三项的出现是由于我们用 $\bar{z}_t$ 和 f_{1t} 的线性组合代替 f_{2t}；第四项包含了模型（1）的误差项 ε_i。下面的引理研究了这四项的渐近性质，通过这些引理我们可以证明本文提出的估计量渐近地服从正态分布。

引理 1：如果假设 1-4 成立，$f(u_i,x_i)$、$f(u_i,x_i,q)$ 和 $f(u_i,q)$ 在 u_0 处连续，并且当 $T\to\infty$ 时，$h\to 0$、$Th\to\infty$，则：

① $T^{-1}\mathring{X}_i^{\mathrm{T}}W_i\mathring{X}_i \to_p f_u(u_0)\begin{pmatrix}\Omega_{ix_ix_i}(u_0) & 0\\ 0 & \mu_2\Omega_{ix_ix_i}(u_0)\end{pmatrix}$；

② $T^{-1}\mathring{X}_i^{\mathrm{T}}W_iQ \to_p f_u(u_0)\begin{pmatrix}\Omega_{ix_iq}(u_0)\\ 0\end{pmatrix}$；

③ $T^{-1}Q^{\mathrm{T}}W_iQ \to_p f_u(u_0)\Omega_{iqq}(u_0)$；

④ $T^{-1}\mathring{X}_i^{\mathrm{T}}W_i^{1/2}M_iW_i^{1/2}\mathring{X}_i \to_p f_u(u_0)\begin{pmatrix}\Omega(u_0) & 0\\ 0 & \mu_2\Omega_{ix_ix_i}(u_0)\end{pmatrix}$。

引理 2：如果引理 1 的条件成立，则：

$$T^{-1}\mathring{X}_i^{\mathrm{T}}W_i^{1/2}M_iW_i^{1/2}A_i \to_p f_u(u_0)\begin{pmatrix}\mu_2\Omega(u_0)\\ \mu_3\Omega_{ix_ix_i}(u_0)\end{pmatrix}。$$

引理 3：如果引理 1 的条件成立，则：$T^{-1}\mathring{X}_i^{\mathrm{T}}W_i^{1/2}M_iW_i^{1/2}D_i = o_p(h^2)$。

引理 4：如果引理 1 的条件成立，并且 $f(u_i,v_i)$、$f(u_i,b)$ 和 $f(u_i,f_2)$ 在 u_0 处连续，则：

① $T^{-1}E\left\|W_i^{1/2}\bar{v}^*\right\|^2 = O(\frac{1}{N})$；

② $T^{-1}E\left\|W_i^{1/2}B\right\|^2 = O(1)$；

③ $T^{-1}E\left\|W_i^{1/2}F_2\right\|^2 = O(1)$；

④ 当 $N\to\infty$ 时，$M_i^BW_i^{1/2}F_2 = 0$ 的概率趋近于 1（*w.p.a.1*）；

⑤ $E\left\|T^{-1}\left(Q^{\mathrm{T}}W_iQ - B^{\mathrm{T}}W_iB\right)\right\| = O(\frac{1}{\sqrt{N}})$；

⑥ $T^{-1}(\tilde{X}_i^{\mathrm{T}}W_i^{1/2}M_iW_i^{1/2}F_2-\tilde{X}_i^{\mathrm{T}}W_i^{1/2}M_i^BW_i^{1/2}F_2)\gamma_{2i}=O_p(\frac{1}{\sqrt{N}})$。

引理 5：如果假设 1-6 成立，并且当 $T\to\infty$ 时，$h\to 0$、$Th\to\infty$，则：

① $h\,\mathrm{var}(G_{i1})\to\lambda_0 f_u(u_0)\Omega^*(u_0)$；

② $h\sum_{\tau=1}^{T-1}\left|\mathrm{cov}(G_{i1},G_{i,\tau+1})\right|=o(1)$；

③ $Th\,\mathrm{var}(\Psi_T)\to\lambda_0 f_u(u_0)\Omega^*(u_0)$。

（上述引理的详细证明我们将在本文的数学附录中给出。）

定理 1：如果假设 1-5 成立，$f(u_i,\xi_i)$ 在 u_0 处连续，则：

$$\sqrt{Th}\left(\hat{\beta}_i(u_0)-\beta_i(u_0)-\frac{h^2}{2}\mu_2\beta_i''(u_0)+O_p(\frac{1}{\sqrt{N}})\right)\to_d N(0,\Theta^2(u_0))$$

其中 $\Theta^2(u_0)=\dfrac{\lambda_0}{f_u(u_0)}\Omega^{-1}(u_0)\Omega^*(u_0)\Omega^{-1}(u_0)$，

$$\begin{aligned}\Omega^*(u_0)=&\Omega^*_{ix_ix_i}(u_0)-\Omega_{ix_iq}(u_0)\Omega^{-1}_{iqq}(u_0)\Omega^{*\mathrm{T}}_{ix_iq}(u_0)-\Omega^*_{ix_iq}(u_0)\Omega^{-1}_{iqq}(u_0)\Omega^{\mathrm{T}}_{ix_iq}(u_0)\\&+\Omega_{ix_iq}(u_0)\Omega^{-1}_{iqq}(u_0)\Omega^*_{iqq}(u_0)\Omega^{-1}_{iqq}(u_0)\Omega^{\mathrm{T}}_{ix_iq}(u_0)\end{aligned}$$
。

证明：我们使用 small-block 和 big-block 技术，即将 $\{1,\ldots,T\}$ 分割成 $2\rho_T+1$ 个子集，其中 big-block 的元素个数为 $r=r_T$，small-block 的元素个数为 $s=s_T$。令 $\rho=\rho_T=\left\lfloor\dfrac{T}{r_T+s_T}\right\rfloor$。根据 Cramer-Wold 定理，对任意的单位向量 $\varphi\in\mathbb{R}^p$，我们证明 $\sqrt{Th}\varphi^{\mathrm{T}}\Psi_T$ 服从渐近正态分布。令 $g_{T,it}=\sqrt{h}\varphi^{\mathrm{T}}G_{i,t+1}$，$t=0,\ldots,T-1$，则 $\sqrt{Th}\varphi^{\mathrm{T}}\Psi_T=\dfrac{1}{\sqrt{T}}\sum_{t=0}^{T-1}g_{T,it}$。由引理 5，

$$\begin{aligned}&\mathrm{var}(g_{T,i0})=\lambda_0 f_u(u_0)\varphi^{\mathrm{T}}\Omega^*(u_0)\varphi\left(1+o_p(1)\right)\equiv\psi^2(u_0)\left(1+o_p(1)\right)\\&\sum_{\tau=1}^{T-1}\left|\mathrm{cov}(g_{T,i0},g_{T,i\tau})\right|=o(1)\end{aligned}$$
。

对 $0\le j\le\rho-1$，定义如下随机变量：

$$\varpi_{lj}=\sum_{t=j(r+s)}^{j(r+s)+r-1}g_{T,it},\ \varpi_{sj}=\sum_{t=j(r+s)+r}^{(j+1)(r+s)-1}g_{T,it},\ \varpi_\rho=\sum_{t=\rho(r+s)}^{T-1}g_{T,it}。$$

则有：

$$\sqrt{Th}\varphi^{\mathrm{T}}\Psi_T = \frac{1}{\sqrt{T}}\left\{\sum_{j=0}^{\rho-1}\varpi_{lj} + \sum_{j=0}^{\rho-1}\varpi_{sj} + \varpi_{\rho}\right\}$$
$$\equiv \frac{1}{\sqrt{T}}\left\{\Xi_{T,1} + \Xi_{T,2} + \Xi_{T,3}\right\}$$

现证明如下事实：

$$\frac{1}{T}E\left(\Xi_{T,2}\right)^2 \to 0, \frac{1}{T}E\left(\Xi_{T,3}\right)^2 \to 0 \tag{A.4}$$

$$\left|E\left[\exp(it\Xi_{T,1})\right] - \prod_{j=0}^{\rho-1}E\left[\exp(it\varpi_{lj})\right]\right| \to 0 \tag{A.5}$$

$$\frac{1}{T}\sum_{j=0}^{\rho-1}E(\varpi_{lj}^2) \to \psi^2(u_0) \tag{A.6}$$

$$\frac{1}{T}\sum_{j=0}^{\rho-1}E\left[\varpi_{lj}^2 I\left\{\left|\varpi_{lj}\right| \geq \varepsilon\psi(u_0)\sqrt{T}\right\}\right] \to 0 \tag{A.7}$$

首先证明（A.4）。根据假设 6（a），存在正实数数列$l_T \to \infty$满足$l_T s_T = o(\sqrt{Th})$且$l_T (T/h)^{1/2}\overline{\alpha}_i(s_n) \to 0$，定义$r_T = \left\lfloor (Th)^{1/2}/l_T \right\rfloor$。于是，我们有：

$$s_T / r_T \to 0，\ r_T/T \to 0，\ r_T(Th)^{-1/2} \to 0，\ (T/r_T)\alpha(s_T) \to 0 \tag{A.8}$$

注意到

$$E\left(\Xi_{T,2}\right)^2 = \sum_{j=0}^{\rho-1}\mathrm{var}(\varpi_{sj}) + 2\sum_{0\leq k<j\leq\rho-1}\mathrm{cov}(\varpi_{sk},\varpi_{sj}) \equiv J_{10} + J_{11}$$

由严格平稳性和引理 5，可得：

$$J_{10} = \rho_T\ \mathrm{var}(\varpi_{s0}) = \rho_T\ \mathrm{var}(\sum_{j=1}^{s_T} g_{T,ij}) = \rho_T s_T\left[\psi^2(u_0) + o(1)\right]$$

对于J_{11}，令$r_j^* = j(r_T + s_T) + r_T$。注意到当$k \neq j$时，$\left|r_j^* + j_2 - r_k^* - j_1\right| \geq r_T$。于是，由引理 5②，我们有：

$$\begin{aligned}\left|J_{11}\right| &\leq 2\sum_{0\leq k<j\leq\rho-1}\sum_{j_1=0}^{s_T-1}\sum_{j_2=0}^{s_T-1}\left|\mathrm{cov}(g_{T,i,r_k^*+j_1}, g_{T,i,r_j^*+j_2})\right| \\ &\leq 2\sum_{j_1=0}^{T-r_T-1}\sum_{j_2=j_1+r_T}^{T-1}\left|\mathrm{cov}(g_{T,ij_1}, g_{T,ij_2})\right| \\ &\leq 2T\sum_{j=r_T}^{T-1}\left|\mathrm{cov}(g_{T,i0}, g_{T,ij})\right| \\ &= o(T)\end{aligned}$$

因此，$\dfrac{1}{T}E\left(\Xi_{T,2}\right)^2 = O(\rho_T s_T T^{-1}) + o(1) = o(1)$。对于$\dfrac{1}{T}E\left(\Xi_{T,3}\right)^2$，我们有：

$$\frac{1}{T}E\left(\Xi_{T,3}\right)^2 \le \frac{1}{T}\left[T-\rho_T(r_T+s_T)\right]\mathrm{var}(g_{T,i0})+2\sum_{j=1}^{T-1}\left|\mathrm{cov}(g_{T,i0},g_{T,ij})\right|$$
$$\le \frac{r_T+s_T}{T}\psi^2(u_0)+o(1)\to 0$$

（A.4）得证。

根据 Cai 和 Li（2008）中的引理 2，我们有：

$$\left|E\left[\exp(it\Xi_{T,1})\right]-\prod_{j=0}^{\rho-1}E\left[\exp(it\varpi_{lj})\right]\right|\le 16\rho_T\bar{\alpha}_i(s_T+1)$$
$$: \ 16\frac{T}{r_T}\bar{\alpha}_i(s_T)\to 0$$

于是，（A.5）得证。

对于（A.6），根据严格平稳性和引理 5，我们有：

$$E(\varpi_{l0}^2)=\mathrm{var}(\sum_{t=0}^{r-1}g_{T,it})=r_T\psi^2(u_0)(1+o(1))$$

因此，$\frac{1}{T}\sum_{j=0}^{\rho-1}E(\varpi_{lj}^2)=\frac{\rho_T r_T}{T}\psi^2(u_0)(1+o(1))\to\psi^2(u_0)$。

最后证明（A.7）：根据 Cai 和 Li（2008）中的引理 3，我们有：

$$E\left[\varpi_{l0}^2 I\left\{\left|\varpi_{l0}\right|\ge\varepsilon\psi(u_0)\sqrt{T}\right\}\right]\le CT^{1-\delta/2}E(\left|\varpi_{l0}\right|^{\delta})$$
$$\le CT^{1-\delta/2}r_T^{\delta/2}E(\left|g_{T,i0}\right|^{\delta^*})^{\delta/\delta^*}$$

与引理 5②的证明类似，我们有：$E(\left|g_{T,i0}\right|^{\delta^*})\le Ch^{1-\delta^*/2}$。因此，

$$E\left[\varpi_{l0}^2 I\left\{\left|\varpi_{l0}\right|\ge\varepsilon\psi(u_0)\sqrt{T}\right\}\right]\le CT^{1-\delta/2}r_T^{\delta/2}h^{(2-\delta^*)\delta/(2\delta^*)}$$

于是，由上式、r_T 的定义和假设 5（c），我们可以得到：

$$\frac{1}{T}\sum_{j=0}^{\rho-1}E\left[\varpi_{lj}^2 I\left\{\left|\varpi_{lj}\right|\ge\varepsilon\psi(u_0)\sqrt{T}\right\}\right]\le Cl_T^{1-\delta/2}T^{1/2-\delta/4}h^{\delta/\delta^*-1/2-\delta/4}\to 0$$

证毕。

从定理 1 可以看出，LCCE 估计量的偏误包括两部分。除了局部线性估计带来的偏误 $\frac{h^2}{2}\mu_2\beta_i^{''}(u_0)$ 之外，LCCE 估计量还含有 $O_p(\frac{1}{\sqrt{N}})$ 阶的偏误。这部分偏误的产生是由于我们使用了 $\bar{z}_t$ 和 f_{1t} 的线性组合代替不可观测的 f_{2t}。如果 $N\to\infty$，此项偏误依概率收敛于 0。在第四部分的蒙特卡洛模拟中，我们比较了 LCCE 估计量和直接使用 f_{2t} 数据的不可行估计量的小样本表现。

四、蒙特卡洛模拟

1. 数据生成过程

我们考虑如下数据生成过程：

$$
\begin{aligned}
&y_{it}=\beta_i(u_{it})x_{it}+\gamma_{1i}+\gamma_{2i,1}f_{2t,1}+e_{it}\\
&e_{it}=\gamma_{2i,2}f_{2t,2}+\varepsilon_{it}\\
&\beta_i(u_{it})=\frac{e^{u_{it}}}{e^{u_{it}}+1}+\delta_i(0.5u_{it}-0.25u_{it}^2)\\
&x_{it}=\Gamma_{1i,x}+\Gamma_{2i,x1}f_{2t,1}+\Gamma_{2i,x2}f_{2t,2}+v_{it,x}\\
&u_{it}=\Gamma_{1i,u}+\Gamma_{2i,u1}f_{2t,1}+\Gamma_{2i,u2}f_{2t,2}+v_{it,u}
\end{aligned}
$$

$i=1,\ldots,N$，$t=1,\ldots,T$。在上面的数据生成过程中，我们有一个可观测的共同因子 $f_{1t}=1$ 和两个不可观测的共同因子 $f_{2t,1}$、$f_{2t,2}$。δ_i 是否随着 i 变化决定了上述模型是同质面板还是异质面板变系数据模型。我们考虑 δ_i : IIDU(0, 1)情况下的估计。上述数据的具体生成过程如下：

⑴ ε_{it} 是 AR(1)过程：$\varepsilon_{it}=\rho_{i\varepsilon}\varepsilon_{i,t-1}+\sigma_i\sqrt{(1-\rho_{i\varepsilon}^2)}\zeta_{it}$，$t=-49,\ldots,1,\ldots,T$，其中 ζ_{it} : IIDN(0, 1)，$\rho_{i\varepsilon}$: IIDU(0.05, 0.95)，σ_i^2 : IIDU(0.05, 1.5)；

⑵ $v_{it,x}$ 是 AR(1)过程：$v_{it,x}=\rho_{vix}v_{i,t-1,x}+w_{it,x}$，$t=-49,\ldots,1,\ldots,T$，其中 $w_{it,x}$: IIDN(0, 1-ρ_{vix}^2)，ρ_{vix} : IIDU(0.05, 0.95)，$w_{i,-50,x}=0$，$v_{it,u}$ 的生成方式与 $v_{it,x}$ 的生成方式相同但彼此独立；

⑶ $f_{2t,j}$（j = 1，2）是彼此独立的 AR(1)过程，满足：$f_{2t,j}=\rho_{fj}f_{2,t-1,j}+v_{ft,j}$，$t=-49,\ldots,1,\ldots,T$，其中 $v_{ft,j}$: IIDN(0, 1-ρ_{fj}^2)，$\rho_{fj}=0.5$，$f_{2,-50,j}=0$；

⑷ $\Gamma_{1i}=\left(\Gamma_{1i,x},\Gamma_{1i,u}\right)^{\mathrm{T}}$: IIDN$\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1&0.5\\0.5&1\end{pmatrix}\right)$，$\gamma_{2i}=\left(\gamma_{2i,1},\gamma_{2i,2}\right)^{\mathrm{T}}$ 以同样方式生成；

⑸ $\gamma_{1i}=0.5\bar{x}_i$，$\bar{x}_i=T^{-1}\sum_{t=1}^{T}x_{it}$；

⑹ $vec(\Gamma_{2i})=\left(\Gamma_{2i,x1},\Gamma_{2i,x2},\Gamma_{2i,u1},\Gamma_{2i,u2}\right)^{\mathrm{T}}$: IIDN($\Gamma_{2,\tau}$, I_4)，τ = A, B，其中

$\Gamma_{2,\mathrm{A}} = (1,0,0,1)^{\mathrm{T}}$，$\Gamma_{2,\mathrm{B}} = (1,1,0,0)^{\mathrm{T}}$。

在假设 3（d）中，我们假设 $E(\Gamma_{2i})$ 是列满秩的。在（6）中，我们考虑两种生成 Γ_{2i} 的方式，情况 A 中 $E(\Gamma_{2i})$ 是列满秩的，而情况 B 中 $E(\Gamma_{2i})$ 不满足列满秩的假设。我们将通过蒙特卡洛模拟观察在不满足假设 3（d）的情况下 LCCE 估计量的小样本性质。

2. 模拟结果分析

我们选用 Epanechnikov 核函数。在估计异质模型时，我们取 $h = 2.34\hat{\sigma}T^{-1/5}$，其中 $\hat{\sigma}$ 是所有平滑变量观测值的样本标准差。令 $(u_1,\ldots,u_d)$ 是 d 个格点（grid points），为了比较异质 LCCE 估计量的小样本性质，我们定义如下均方根误差（RMSE）：

$$RMSE_{he} = \sqrt{\frac{1}{Nd}\sum_{i=1}^{N}\sum_{t=1}^{d}\left(\hat{\beta}_i(u_t) - \beta_i(u_t)\right)^2}$$

表 1 给出了 $RMSE_{he}$ 的 1000 次重复模拟的中位数。我们比较了假设 3（d）成立与否对 $RMSE_{he}$ 的影响，并且比较了 LCCE 估计量和直接使用 f_{2t} 数据的不可行估计量的小样本表现。

从表 1 可以看出：（1）固定 T，N 的增加没有减小异质系数函数估计量的 $RMSE_{he}$。这主要是因为 N 的增加导致了更严重的异质性，而且异质系数函数估计量的收敛速度依赖于 T。这与 Su 和 Jin（2012）的非参数异质估计量的小样本性质类似；（2）随着 T 的增加，异质系数函数估计量的 $RMSE_{he}$ 减小；（3）如果假设 3（d）的满秩条件不成立，LCCE 和不可行 LCCE 的 $RMSE_{he}$ 都比满秩条件成立时的结果有显著地增加；（4）通过与不可行估计量的结果比较可以发现，当满秩条件成立时，LCCE 的 $RMSE_{he}$ 和不可行 LCCE 的 $RMSE_{he}$ 相差很小；但当满秩条件不成立时，LCCE 有更大的效率损失（efficiency loss）。

五、结论

本文建立了一类变系数异质面板数据模型，其自变量、平滑变量和误差项存在通过共同因子结构引入的截面相关。由于同时允许自变量和误差项具有因子结构，

本文提出的模型适用于自变量是内生变量的情况。为了控制不可观测的共同因子，我们提出了基于 CCE 思想的局部线性变系数估计量。当 N 和 T 同时趋近于无穷时，在适当的正则假设条件下，本文提出的异质变系数估计量服从渐近正态分布。

本文构造的估计量及其渐近结果可以用于构造关于异质面板数据模型的检验统计量。例如，可以用于检验异质的系数函数是否为线性函数或是某一常数；也可以检验系数函数是否同质，等等。

本文的局限性在于使用 CCE 估计方法的前提是自变量和平滑变量满足（2）设定的形式。如果（2）不成立，我们可以考虑将 Song（2013）的方法用于估计变系数异质面板数据模型。如果我们想要估计同质的系数函数，可以考虑将 Bai（2009）的方法推广到变系数同质面板数据模型。

表 1　异质模型变系数估计量的 RMSE

	N/T	50	100	200
		满秩		
LCCE	50	0.5491	0.3323	0.2216
	100	0.6675	0.3975	0.2554
	200	0.8537	0.4900	0.2971
$LCCE_{inf}$	50	0.5631	0.3298	0.2228
	100	0.6413	0.3857	0.2413
	200	0.9267	0.4864	0.2941
		不满秩		
LCCE	50	0.8700	0.5250	0.3656
	100	1.1485	0.6672	0.4194
	200	1.5887	0.9120	0.5592
$LCCE_{inf}$	50	0.7591	0.4552	0.2886
	100	1.0247	0.5748	0.3502
	200	1.4124	0.7757	0.4651

参考文献

[1] Arellano, M., 2003, *Panel Data Econometrics* [M], Oxford University Press.

[2] Bai J., 2009, *Panel Data Models with Interactive Fixed Effects* [J], Econometrica, 77(4), 1229~1279.

[3] Bernstein, D. S., 2009, *Matrix Mathematics: Theory, Facts, and Formulas* [M], Princeton University Press.

[4] Cai Z., 2010, *Functional Coefficient Models for Economic and Financial Data* [A], Oxford Handbook on Statistics and Functional Data Analysis [M], Oxford University Press, 166~186.

[5] Cai Z., Fan, J., and Yao, Q., 2000, *Functional-coefficient Regression Models for Nonlinear time series* [J], Journal of the American Statistical Association, 95(451):941~956.

[6] Cai Z., Li, Q., 2008, *Nonparametric Estimation of Varying-coefficient Dynamic Panel Data Models* [J], Econometric Theory, 24(5):1321~1342.

[7] Chudik, A., Pesaran, M. H., 2014, *Common Correlated Effects Estimation of Heterogeneous Dynamic Panel Data Models with Weakly Exogenous Regressors* [R], CESifo Working Paper No. 4232.

[8] Fan, J., Yao, Q., 2005, *Nonlinear Time Series: Nonparametric and Parametric Methods* [M], Springer, New York.

[9] Hall, P., Heyde, C. C., 1980, *Martingale Limit Theory and Its Applications* [M], New York: Academic Press.

[10] Hsiao, C., 2003, *Analysis of Panel Data*, 2nd [M], Cambridge University Press.

[11] Huang, X., 2013, *Nonparametric Estimation in Large Panels with Cross-sectional Dependence* [J], Econometric Reviews, 32(5-6):754-777.

[12] Lee, Y., 2013, *Nonparametric Estimation of Dynamic Panel Models with Fixed Effects* [R], Working Paper, Michigan University.

[13] Li, D., Chen, J., and Gao, J., 2011, *Non-parametric Time-varying Coefficient Panel Data Models with Fixed Effect* [J], The Econometrics Journal, 14(3):387~408.

[14] Masry, E., 1996, *Multivariate regression estimation local polynomial fitting for time series* [J], Stochastic Processes and their Applications, 65(1):81~101.

[15] Moon, H. R., Weidner, M., 2013a, *Dynamic Linear Panel Regression Models with Interactive Fixed Effects* [R], Cemmap Working Paper No. CWP63/13.

[16] Moon, H. R., Weidner, M., 2013b, *Linear Regression for Panel with Unknown Number of Factors as Interactive Fixed Effects* [R], Cemmap Working Paper No. CWP49/13.

[17] Pesaran, M. H., 2006, *Estimation and inference in large heterogeneous panels with a multifactor error structure* [J], Econometrica, 74(4):967~1012.

[18] Shao, Q., Yu, H., 1996, *Weak Convergence of Weighted Empirical Processes of Dependent Sequences* [J], Annals of Probability, 24(4):2098-2127.

[19] Song, M., 2013, *Asymptotic Theory for Dynamic Heterogeneous Panels with Cross-sectional Dependence and Its Applications* [R], Mimeo Working Paper.

[20] Su, L., Jin, S., 2012, *Sieve Estimation of Panel Data Models with Cross Section Dependence*

[J], Journal of Econometrics, 169(1):34~47.

[21] Su, L., Lu, X., 2013, *Nonparametric Dynamic Panel Data Models: Kernel Estimation and Specification Testing* [J], Journal of Econometrics, 176(2):112~133.

[22] Su, L., Zhang, Y., 2013, *Nonparametric Dynamic Panel Data Models with Interactive Fixed Effects: Sieve Estimation and Specification Testing* [R], Working Paper, Singapore Management University.

[23] Wang, G., Wei, Y., and Qiao, S., 2004, *Generalized Inverse: Theory and Computations* [M], Science Press, New York.

[24] Wheeden, R. L., Zygmund, A., 1977, *Measure and Integral* [M], New York: Marcel Dekker.

[25] 蔡宗武、陈琳娜、方颖：《人民币汇率的半参数预测模型》[J]，《系统工程理论与实践》2012 年第 4 期。

[26] 方颖、郭萌萌：《中国主要宏观变量的稳定性检验：基于非参数估计与 Bootstrapping 的一个方法》[J]，《世界经济文汇》2009 年第 1 期。

[27] 钱金保、钱彬彬：《共同冲击对广东省出口的效应分析》[J]，《南方经济》2012 年第 10 期。

[28] 张志强：《空间面板参数估计的小样本特性探究》[J]，《数量经济技术经济研究》2012 年第 9 期。

[29] 王火根、沈利生：《中国经济增长与能源消费空间面板分析》[J]，《数量经济技术经济研究》2007 年第 12 期。

[30] 王锐淇：《我国区域技术创新能力空间相关性及扩散效应实证分析》[J]，《系统工程理论与实践》2012 年 11 月。