

具有自相关误差结构的面板数据部分线性单指标模型的统计推断

徐群芳¹, 刘高生², 柏杨^{2*}

1. 宁波大学商学院, 宁波 315211;
2. 上海财经大学统计与管理学院, 上海 200433
E-mail: tsunfangxu@163.com, gaosheng@163.sufe.edu.cn, statbyang@mail.shufe.edu.cn

收稿日期: 2018-09-21; 接受日期: 2019-01-28; 网络出版日期: 2019-05-20; * 通信作者
教育部人文社会科学研究规划基金(批准号: 15YJA910004)、宁波大学王宽诚幸福基金和国家自然科学基金(批准号: 11771268)
资助项目

摘要 部分线性单指标模型是在科学的研究中具有广泛应用的经典半参数模型之一。本文主要研究具有自相关误差结构的面板数据的部分线性单指标模型的统计推断问题。通过结合局部多项式和纠偏广义估计方程方法, 本文提出模型参数的可行加权广义估计(feasible weighted generalized estimating equation estimation, GEE-FW), 证明该估计具有相合性和渐近正态性, 并且在渐近方差意义下阐明该估计比工作独立的广义估计(generalized estimating equation estimation based on working independence, GEE-WI) 更加有效。此外, 本文对模型中未知连接函数提出两阶段局部线性估计(two step local linear generalized estimating equation estimation, GEE-TS), 建立该估计的渐近性质。数值模拟研究和实际数据分析都表明了本文所提出的方法是有效的, 在理论和应用方面均具有良好的表现。

关键词 面板数据 部分线性单指标模型 序列自相关

MSC (2010) 主题分类 62G08, 62J02

1 引言

面板数据是指对给定的多个不同研究个体在一段时间的多个时间点进行多次观测所得到的数据, 比横截面数据和时间序列数据可以提供更多的信息。面板数据分析已经广泛应用于经济、金融、生物和医学等众多领域。面板数据参数模型的统计推断已经被广泛深入研究, 可参见文献[1–4]。然而简单易解释的参数模型却有模型误设的风险。为降低参数模型的限制, Ruckstuhl 等^[5]提出了非参数面板数据回归模型, 探求潜在的结构, 减少建模偏差。但是非参数模型也会存在如维数灾难难以解释和不易外推等缺陷。为此, 近十几年来, 很多文献提出用部分线性半参数模型对面板数据建模并进行统计

英文引用格式: Xu Q F, Liu G S, Bai Y. Statistical inference for partially linear single-index model of panel data with serially correlated error structure (in Chinese). Sci Sin Math, 2019, 49: 899–930, doi: 10.1360/N012018-00230

推断研究(参见文献[6–8]). 然而, 当非参数协变量的个数很多时, 部分线性模型仍然存在维数灾难问题. 因此考虑如下起到降维、避免维数灾难问题的部分线性单指标模型:

$$Y_{it} = \mathbf{X}_{it}^T \boldsymbol{\beta} + g(\mathbf{Z}_{it}^T \boldsymbol{\theta}) + \varepsilon_{it}, \quad i = 1, \dots, n, \quad t = 1, \dots, T, \quad (1.1)$$

其中 Y_{it} 是响应变量; $(\mathbf{X}_{it}^T, \mathbf{Z}_{it}^T)^T = (X_{it1}, \dots, X_{itp}, Z_{it1}, \dots, Z_{itq})^T$ 是解释变量, 它与模型误差 ε_{it} 独立; $\boldsymbol{\beta}$ 和 $\boldsymbol{\theta}$ 分别为 p 和 q 维未知的回归系数; $g(\cdot)$ 是未知的连接函数.

显见, 模型(1.1)包含很多常见的参数、半参数和非参数模型. 例如, 当 $q = 1$ 时, 模型简化为常见的面板数据部分线性模型. 已有很多学者把该模型运用到实际问题分析中, 如 Zeger 和 Diggle^[9]. 当 $g(\cdot) \equiv 0$ 时, 模型变成了面板数据线性回归模型(参见文献[1, 3]). 当 $T \equiv 1$ 时, 模型退化为经典的截面数据部分线性单指标模型(参见文献[10–17]).

个体内部相关性和横截面异质性是面板数据很重要的特性. 如何合理描述其误差结构是值得研究的重要问题之一. 单因素自相关误差结构是实际中比较常见的情形. 为此, 进一步假定模型(1.1)中的随机误差项 ε_{it} 服从如下单因素 $AR(s)$ 自相关的误差结构:

$$\varepsilon_{it} = \mu_i + \nu_{it}, \quad \nu_{it} = \rho_1 \nu_{i,t-1} + \dots + \rho_s \nu_{i,t-s} + e_{it}, \quad (1.2)$$

其中 μ_i 和 e_{it} 为独立同分布均值为 0、方差分别为 σ_μ^2 和 σ_e^2 的随机变量; $\boldsymbol{\rho} = (\rho_1, \dots, \rho_s)^T$ 为满足平稳条件 $1 - \rho_1 z - \dots - \rho_s z^s \neq 0$ ($|z| \leq 1$) 的未知自相关系数; 个体随机效应 μ_i 与服从 $AR(s)$ 过程的剩余扰动 ν_{it} 相互独立. 根据文献[18]可知, $AR(s)$ 过程是平稳 s -阶相依过程($E(\nu_{it} \nu_{i,t-j}) = \gamma_j$, 其中 $\gamma_{-j} = \gamma_j$)的一个特例.

误差项(1.2)比传统的假定(ε_{it} 和 $\varepsilon_{it'}$ 具有等相关的误差结构, 如文献[3])更具有一般性. 在模型(1.2)下, 同一个个体的多个观测间 Y_{i1}, \dots, Y_{iT} 允许相关. 误差结构(1.2)与 Chen 等^[19]提出的主要不同之处在于: 他们考虑用固定效应来刻画横截面异质性, 而我们将它视为随机效应. 另外, 他们提出的估计渐近理论是在时间 T 和样本量 n 都趋于无穷条件下得到的, 而本文研究假定 T 固定不趋于无穷并且 $T \geq s$, 这在很多领域的面板数据中是较常见的情形.

令 $\mathbf{Y} = (\mathbf{Y}_1^T, \dots, \mathbf{Y}_n^T)^T$, 其中 $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{iT})^T$, $\mathbf{X} = (\mathbf{X}_1^T, \dots, \mathbf{X}_n^T)^T$, $\mathbf{X}_i = (\mathbf{X}_{i1}, \dots, \mathbf{X}_{iT})^T$, $\mathbf{Z} = (\mathbf{Z}_1^T, \dots, \mathbf{Z}_n^T)^T$, $\mathbf{Z}_i = (\mathbf{Z}_{i1}, \dots, \mathbf{Z}_{iT})^T$, $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1^T, \dots, \boldsymbol{\varepsilon}_n^T)^T$, $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \dots, \varepsilon_{iT})^T$, 则模型(1.1)可以写成如下向量矩阵形式:

$$\begin{aligned} \mathbf{Y}_i &= \mathbf{X}_i \boldsymbol{\beta} + g(\mathbf{Z}_i \boldsymbol{\theta}) + \boldsymbol{\varepsilon}_i, \quad i = 1, \dots, n, \\ \mathbf{Y} &= \mathbf{X} \boldsymbol{\beta} + g(\mathbf{Z} \boldsymbol{\theta}) + \boldsymbol{\varepsilon}. \end{aligned}$$

误差项模型(1.2)写成如下向量形式:

$$\begin{aligned} \boldsymbol{\varepsilon}_i &= \mathbf{1}_T \boldsymbol{\mu}_i + \boldsymbol{\nu}_i, \quad i = 1, \dots, n, \\ \boldsymbol{\varepsilon} &= (\mathbf{I}_n \otimes \mathbf{1}_T) \boldsymbol{\mu} + \boldsymbol{\nu}, \end{aligned}$$

其中 $\boldsymbol{\nu}_i = (\nu_{i1}, \dots, \nu_{iT})^T$, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^T$, $\boldsymbol{\nu}^T = (\nu_{11}, \dots, \nu_{1T}, \dots, \nu_{n1}, \dots, \nu_{nT})$, $\mathbf{1}_T$ 为所有元素为 1 的 T 维向量, \mathbf{I}_n 是维数为 n 的单位阵, \otimes 表示 Kronecker 乘积. $\boldsymbol{\varepsilon}_i$ 和 $\boldsymbol{\varepsilon}$ 的方差和协方差矩阵可分别表示为

$$\boldsymbol{\Sigma}_0 = E(\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^T) = \sigma_\mu^2 \mathbf{J}_T + \boldsymbol{\Psi}, \quad \boldsymbol{\Omega} = E(\boldsymbol{\varepsilon} \boldsymbol{\varepsilon}^T) = \sigma_\mu^2 (\mathbf{I}_n \otimes \mathbf{J}_T) + \mathbf{I}_n \otimes \boldsymbol{\Psi} = \mathbf{I}_n \otimes \boldsymbol{\Sigma}_0,$$

其中 $\mathbf{J}_T = \mathbf{1}_T \mathbf{1}_T^T$ 是所有元素为 1 的 $T \times T$ 矩阵; $\boldsymbol{\Psi} = \sigma_v^2 \mathbf{V} = \mathbf{E}(\boldsymbol{\nu}_i \boldsymbol{\nu}_i^T)$ 为剩余扰动项 $\boldsymbol{\nu}_i$ (假定每个个体都相同) 的方差协方差矩阵; \mathbf{V} 假定为对称正定矩阵, 并且 $\sigma_v^2 = \gamma_0 = \mathbf{E}(\nu_{it}^2)$. 特别地, 记 $s = 1$ 和 $\nu_{i0} \sim N(0, \sigma_e^2/(1 - \rho^2))$, 则 $\boldsymbol{\Psi} = \sigma_v^2 \mathbf{V}$, $\sigma_v^2 = \sigma_e^2/(1 - \rho^2)$,

$$\mathbf{V} = \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{pmatrix}.$$

对于模型 (1.1) 和 (1.2), 未知参数分量和非参数连接函数的有效估计是值得重点关注的问题. 其特殊情形之一, 即针对截面数据部分线性单指标模型, 已有不少文献进行了研究. 例如, Ma 和 Zhu^[17] 提出了一类稳健的纠偏广义估计方程方法, 但这些研究方法无法直接简单应用到个体内具有相关性和横截面具有异质性的面板数据的部分线性单指标模型的统计推断中. 针对面板数据部分线性单指标模型, Li 等^[20] 提出了纠偏广义估计方程和纠偏二次推断函数方法. Xu 等^[21] 考虑了面板数据边际广义部分线性单指标模型, 提出了剖面广义估计方程估计参数, 通过调整核加权估计方程来估计未知函数, 并证明了其估计具有相合性和渐近正态性. 但文献 [20, 21] 都没有考虑相关误差结构的设定, 只基于工作相关阵的假定, 若假定错误则会导致估计有效性的损失, 而且没有考虑在更加有效的参数估计基础上, 进行未知连接函数的两阶段估计以提高估计效果.

因此, 本文基于具有自相关误差结构的面板数据部分线性单指标模型, 在考虑误差自相关结构条件下提高未知参数分量和非参数连接函数的估计有效性. 具体而言, 首先不考虑组内相关, 类似于 Ma 和 Zhu^[17] 提出的双稳健的纠偏广义估计方程方法, 通过剖面最小二乘和局部线性平滑方法, 先将未知连接函数表示为模型中参数部分的局部多项式函数, 再利用纠偏广义估计方程来实现参数和非参数初始估计; 然后, 利用上述参数和非参数的初始估计来实现模型中单因素自相关误差结构的拟合与刻画, 得到未知参数 β 和 θ 的一种可行加权广义估计 (GEE-FW); 之后证明该估计具有相合性和渐近正态性, 并且该估计比不考虑组内相关的未加权 GEE (generalized estimating equation) 估计 (GEE-WI) 更渐近有效. 在更加有效的参数估计基础上, 结合自相关误差对连接函数估计的影响, 本文提出了考虑相关结构的未知连接函数的两阶段局部线性估计 (GEE-TS), 并且考察了 GEE-TS 的渐近性质, 从而提高估计效率.

本文余下的内容安排如下: 第 2 节给出在工作独立条件下基于局部多项式和广义估计方程的未知参数 β 、 θ 及连接函数 $g(\cdot)$ 的初始估计. 第 3 节介绍时间序列自相关误差结构的拟合估计. 第 4 节主要研究合理考虑相关误差结构的参数分量的有效估计, 提出可行加权广义矩估计并推导其大样本性质. 基于广义估计方程的非参数连接函数的两阶段局部线性估计放在第 5 节. 第 6 节是模拟研究. 第 7 节通过一个关于工资的面板数据的案例分析来说明本文所研究的模型及其估计方法在实际问题中的应用. 第 8 节是总结. 理论结果的证明放在附录 A 中.

2 初始估计

为保证模型可识别, 假定参数 θ 的第一个分量为 1, 用 \mathbf{Z}_{-1} 表示 \mathbf{Z} 向量去掉第一个分量构成的 $q - 1$ 维向量, 即 $\theta = (1, \theta_1, \dots, \theta_{q-1})^T$, $\mathbf{Z}_{-1,it} = (\mathbf{Z}_{it2}, \dots, \mathbf{Z}_{itq})^T$. 该单指标分量和参数的识别条件比较简单清晰. 对于模型 (1.1) 参数分量的初始估计, 先不考虑组内相关即在工作独立条件下, 类似 Ma

和 Zhu^[17] 所研究的方法, 利用 Xue 和 Zhu^[13] 的纠偏思想, 构造如下纠偏广义估计方程:

$$\begin{aligned} 0 &= \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \{Y_{it} - \mathbf{X}_{it}^T \boldsymbol{\beta} - g(\mathbf{Z}_{it}^T \boldsymbol{\theta})\} \{\mathbf{X}_{it} - \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})\}, \\ 0 &= \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \{Y_{it} - \mathbf{X}_{it}^T \boldsymbol{\beta} - g(\mathbf{Z}_{it}^T \boldsymbol{\theta})\} g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})\}, \end{aligned} \quad (2.1)$$

其中 $g'(\cdot)$ 是 $g(\cdot)$ 的一阶导数.

Ma 和 Zhu^[17] 已经证明上述估计方程具有稳健性, 即当连接函数 $g(\cdot)$ 不是相合估计或被误设时所得到的参数分量 $\boldsymbol{\beta}$ 和 $\boldsymbol{\theta}$ 的估计仍然是相合估计. 根据文献 [22] 可知, 局部多项式平滑具有可以减少传统 Nadaraya-Watson 估计的偏差和传统 Gasser-Müller 估计的方差, 以及自动适应边界点等优点. 故对估计方程 (2.1) 中所包含的需要估计的 4 个未知函数 $g(\cdot)$ 、 $g'(\cdot)$ 、 $\mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})$ 和 $\mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})$, 利用 Fan 和 Gijbels^[23] 提出的剖面最小二乘和局部线性平滑方法估计.

对给定的参数分量 $\boldsymbol{\beta}$ 和 $\boldsymbol{\theta}$, 局部线性平滑估计是 $\mathbf{Z}_{it}^T \boldsymbol{\theta}$ 在给定 u 的局部邻域内, 关于参数 a_ν ($\nu = 0, 1$), 最小化如下损失函数:

$$\sum_{i=1}^n \sum_{t=1}^T \{(Y_{it} - \mathbf{X}_{it}^T \boldsymbol{\beta}) - a_0 - a_1(\mathbf{Z}_{it}^T \boldsymbol{\theta} - u)\}^2 K_h(\mathbf{Z}_{it}^T \boldsymbol{\theta} - u),$$

其中 $K_h(\cdot) = h^{-1}K(\cdot/h)$ 是定义在 \mathcal{R}^1 上的窗宽为 h 的核 $K(\cdot)$ 的一个正则化核函数. 记 h 和 h_1 分别表示估计 $g(\cdot)$ 和 $g'(\cdot)$ 的窗宽. 通过简单计算, 得到未知函数 $g(\cdot)$ 和 $g'(\cdot)$ 的局部线性平滑估计可表示为

$$\hat{g}(u; \boldsymbol{\beta}, \boldsymbol{\theta}) = \sum_{i=1}^n \sum_{t=1}^T W_{nit}(u, \boldsymbol{\theta})(Y_{it} - \mathbf{X}_{it}^T \boldsymbol{\beta})$$

和

$$\hat{g}'(u; \boldsymbol{\beta}, \boldsymbol{\theta}) = \sum_{i=1}^n \sum_{t=1}^T \widetilde{W}_{nit}(u, \boldsymbol{\theta})(Y_{it} - \mathbf{X}_{it}^T \boldsymbol{\beta}),$$

其中

$$\begin{aligned} W_{nit}(u; \boldsymbol{\theta}) &= \frac{K_h(\mathbf{Z}_{it}^T \boldsymbol{\theta} - u)\{S_{n,2}(u; \boldsymbol{\theta}, h) - (\mathbf{Z}_{it}^T \boldsymbol{\theta} - u)S_{n,1}(u; \boldsymbol{\theta}, h)\}}{S_{n,0}(u; \boldsymbol{\theta}, h)S_{n,2}(u; \boldsymbol{\theta}, h) - S_{n,1}^2(u; \boldsymbol{\theta}, h)}, \\ \widetilde{W}_{nit}(u; \boldsymbol{\theta}) &= \frac{K_{h_1}(\mathbf{Z}_{it}^T \boldsymbol{\theta} - u)\{(\mathbf{Z}_{it}^T \boldsymbol{\theta} - u)S_{n,0}(u; \boldsymbol{\theta}, h_1) - S_{n,1}(u; \boldsymbol{\theta}, h_1)\}}{S_{n,0}(u; \boldsymbol{\theta}, h_1)S_{n,2}(u; \boldsymbol{\theta}, h_1) - S_{n,1}^2(u; \boldsymbol{\theta}, h_1)}, \\ S_{n,l}(u; \boldsymbol{\theta}, h) &= \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T (\mathbf{Z}_{it}^T \boldsymbol{\theta} - u)^l K_h(\mathbf{Z}_{it}^T \boldsymbol{\theta} - u), \quad l = 0, 1, 2. \end{aligned}$$

类似得到未知函数 $\mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})$ 和 $\mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})$ 的如下局部线性平滑估计:

$$\hat{\mathbb{E}}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) = \sum_{i=1}^n \sum_{t=1}^T W_{nit}(u; \boldsymbol{\theta}) \mathbf{X}_{it}, \quad \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) = \sum_{i=1}^n \sum_{t=1}^T W_{nit}(u; \boldsymbol{\theta}) \mathbf{Z}_{-1,it}.$$

因此, 参数分量 β 和 θ 的初始可行纠偏广义估计方程为

$$\begin{aligned} 0 &= \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \{Y_{it} - \mathbf{X}_{it}^T \beta - \hat{g}(\mathbf{Z}_{it}^T \theta)\} \{\mathbf{X}_{it} - \hat{\mathbb{E}}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \theta)\}, \\ 0 &= \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \{Y_{it} - \mathbf{X}_{it}^T \beta - \hat{g}(\mathbf{Z}_{it}^T \theta)\} \hat{g}'(\mathbf{Z}_{it}^T \theta) \{\mathbf{Z}_{-1,it} - \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \theta)\}. \end{aligned} \quad (2.2)$$

记 $(\hat{\beta}_n^T, \hat{\theta}_n^T)$ 是估计方程 (2.2) 的解, 则未知函数分量 $(g(u), g'(u))$ 的初始可行局部线性平滑估计为

$$\begin{aligned} \hat{g}(u) &= \hat{g}(u; \hat{\beta}_n, \hat{\theta}_n) = \sum_{i=1}^n \sum_{t=1}^T W_{nit}(u; \hat{\theta}_n)(Y_{it} - \mathbf{X}_{it}^T \hat{\beta}_n), \\ \hat{g}'(u) &= \hat{g}'(u; \hat{\beta}_n, \hat{\theta}_n) = \sum_{i=1}^n \sum_{t=1}^T \widetilde{W}_{nit}(u; \hat{\theta}_n)(Y_{it} - \mathbf{X}_{it}^T \hat{\beta}_n). \end{aligned}$$

为得到初始估计 $\hat{\beta}_n$ 、 $\hat{\theta}_n$ 、 $\hat{g}_n(u)$ 及之后的最终估计的渐近性质, 给出如下假设条件.

假设 2.1 对于 $i = 1, \dots, n$, 向量 $((\mathbf{X}_{i1}^T, \mathbf{Z}_{i1}^T)^T, \dots, (\mathbf{X}_{iT}^T, \mathbf{Z}_{iT}^T)^T)^T$ 是独立同分布的.

假设 2.2 (i) \mathbf{Z}_{it} 的分布具有紧支集 \mathcal{A} ; (ii) 对个体 $i = 1, \dots, n$, $(\mathbf{Z}_{i1}^T \theta, \dots, \mathbf{Z}_{iT}^T \theta)$ 的联合密度函数存在, $\mathbf{Z}_{it}^T \theta$ 的密度函数为正并且对于 θ 在 θ_0 的邻域内满足 1 阶 Lipschitz 条件. 此外, $\mathbf{Z}_{it}^T \theta_0$ 在 \mathcal{U} 上具有正的有界的密度函数 $p_t(u)$, 其中 $\mathcal{U} = \{u = \mathbf{Z}_{it}^T \theta_0 : z \in \mathcal{A}\}$.

假设 2.3 (i) 函数 $g(\cdot)$ 和 $g_{2k}(\cdot)$ 具有二阶连续有界偏导数, 其中 $g_{2k}(\cdot)$ 是函数向量 $g_2(u)$ 的第 k 个分量, $1 \leq k \leq p$, $g_2(u) = \mathbb{E}(\mathbf{X} | \mathbf{Z}_{it}^T \theta_0 = u)$;

(ii) $g_{3j}(\cdot)$ 满足 1 阶 Lipschitz 条件, 其中 $g_{3j}(\cdot)$ 为 $g_3(u)$ 的第 j 个分量, $1 \leq j \leq q-1$, $g_3(u) = \mathbb{E}(\mathbf{Z}_{-1} | \mathbf{Z}_{it}^T \theta_0 = u)$.

假设 2.4 (i) 核函数 $K(\cdot)$ 是有界连续的概率密度函数, 满足

$$\int_{-\infty}^{\infty} u^2 K(u) du \neq 0, \quad \int_{-\infty}^{\infty} |u|^2 K(u) du < \infty;$$

(ii) 核函数 $K(\cdot)$ 在 \mathcal{R}^1 满足 Lipschitz 条件.

假设 2.5 带宽序列 h 和 h_1 满足

(i) $nh^2 / \log^2 n \rightarrow \infty$, $\limsup_{n \rightarrow \infty} nh^5 < \infty$;

(ii) $n h h_1^3 / \log^2 n \rightarrow \infty$, $nh^4 \log n \rightarrow 0$, $\limsup_{n \rightarrow \infty} nh_1^5 < \infty$.

假设 2.6 在如下定理 2.1 中定义的矩阵 Ω_1 是正定矩阵.

注 2.1 假设 2.1 是对个体独立性的假设, 详见文献 [24, 假设 2(ii)]. 假设 2.2 和 2.3 中的 Lipschitz 条件和二阶导数条件是标准的平滑性条件. 假设 2.4 为二阶核密度的普通假设. 假设 2.5(i) 估计 $\hat{g}(\cdot)$ 时用到的窗宽 h 包含最优窗宽 $n^{-1/5}$, 当分析 $\hat{\theta}_n$ 的渐近性质时, 需要估计 $g(\cdot)$ 的导数 $g'(\cdot)$. 在同样窗宽时, $\hat{g}'(\cdot)$ 收敛速度比 $\hat{g}(\cdot)$ 收敛速度慢, 这会导致 $\hat{\theta}_n$ 的收敛速度低于 \sqrt{nT} , 因此, 采用窗宽 h_1 控制 $\hat{g}'(\cdot)$ 的变化, 假设 2.5(ii) 给出了窗宽 h 与 h_1 的关系, 假设 2.5 详见文献 [16, 25]. 假设 2.6 保证了参数估计的方差极限存在.

未知参数分量的初始估计量 $\hat{\beta}_n$ 和 $\hat{\theta}_n$ 有如下的渐近结论.

定理 2.1 令假设 2.1–2.6 成立, 当 $n \rightarrow \infty$ 时,

$$\sqrt{nT} \{(\hat{\beta}_n^T, \hat{\theta}_n^T)^T - (\beta^T, \theta^T)^T\} \rightarrow_D N(0, \Omega_1^{-1} \Omega_2 \Omega_1^{-1}),$$

其中

$$\begin{aligned}\Omega_1 &= \begin{pmatrix} \Omega_{1(11)} & \Omega_{1(12)} \\ \Omega_{1(21)} & \Omega_{1(22)} \end{pmatrix}, \quad \Omega_2 = \begin{pmatrix} \Omega_{2(11)} & \Omega_{2(12)} \\ \Omega_{2(21)} & \Omega_{2(22)} \end{pmatrix}, \\ \Pi_X &= (\mathbf{X}_{i1}, \dots, \mathbf{X}_{iT}) - (\mathbb{E}(\mathbf{X}_{i1} | \mathbf{Z}_{i1}^T \boldsymbol{\theta}), \dots, \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})), \\ \Pi_Z &= (g'(\mathbf{Z}_{i1}^T \boldsymbol{\theta}) \mathbf{Z}_{-1,i1}, \dots, g'(\mathbf{Z}_{iT}^T \boldsymbol{\theta}) \mathbf{Z}_{-1,it}) - (g'(\mathbf{Z}_{i1}^T \boldsymbol{\theta}) \mathbb{E}(\mathbf{Z}_{-1,i1} | \mathbf{Z}_{i1}^T \boldsymbol{\theta}), \dots, g'(\mathbf{Z}_{iT}^T \boldsymbol{\theta}) \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})), \\ \Omega_{1(11)} &= T^{-1} \mathbb{E}(\Pi_X \Pi_X^T r), \quad \Omega_{1(12)} = T^{-1} \mathbb{E}(\Pi_X \Pi_Z^T), \quad \Omega_{1(21)} = T^{-1} \mathbb{E}(\Pi_Z \Pi_X^T), \\ \Omega_{1(22)} &= T^{-1} \mathbb{E}(\Pi_Z \Pi_Z^T), \quad \Omega_{2(11)} = T^{-1} \mathbb{E}(\Pi_X \Sigma_0 \Pi_X^T), \quad \Omega_{2(12)} = T^{-1} \mathbb{E}(\Pi_X \Sigma_0 \Pi_Z^T), \\ \Omega_{2(21)} &= T^{-1} \mathbb{E}(\Pi_Z \Sigma_0 \Pi_X^T), \quad \Omega_{2(22)} = T^{-1} \mathbb{E}\{\Pi_Z \Sigma_0 \Pi_Z^T\}.\end{aligned}$$

记

$$\varrho_j = \int_{-\infty}^{\infty} u^j K(u) du, \quad \varsigma_j = \int_{-\infty}^{\infty} u^j K^2(u) du, \quad j = 0, 1, 2, 3.$$

未知函数的初始估计量 $\hat{g}(u; \hat{\beta}_n, \hat{\theta}_n)$ 和 $\hat{g}'(u; \hat{\beta}_n, \hat{\theta}_n)$ 有如下的渐近结论.

定理 2.2 令假设 2.1–2.6 成立, 当 $n \rightarrow \infty$ 时,

$$\sqrt{nTh} \left[\mathbf{H} \left\{ \begin{pmatrix} \hat{g}(u; \hat{\beta}_n, \hat{\theta}_n) \\ \hat{g}'(u; \hat{\beta}_n, \hat{\theta}_n) \end{pmatrix} - \begin{pmatrix} g(u) \\ g'(u) \end{pmatrix} \right\} - \frac{h^2}{2} \frac{1}{\varrho_2 - \varrho_1^2} \begin{pmatrix} (\varrho_2^2 - \varrho_1 \varrho_3) g''(u) \\ (\varrho_3 - \varrho_1 \varrho_2) g''(u) \end{pmatrix} \right] \xrightarrow{D} N(0, \Phi),$$

其中 $\mathbf{H} = \text{diag}(1, h)$, $g''(u) = \partial^2 g(u) / \partial u^2$,

$$\Phi = \frac{\sum_{t=1}^T \sigma_{tt}^2 p_t(u)}{(\sum_{t=1}^T p_t(u))^2 (\varrho_2 - \varrho_1^2)^2} \begin{pmatrix} \varrho_2^2 \varsigma_0 - 2\varrho_1 \varrho_2 \varsigma_1 + \varrho_1^2 \varsigma_2 & (\varrho_1^2 + \varrho_2) \varsigma_1 - \varrho_1 \varrho_2 \varsigma_0 - \varrho_1 \varsigma_2 \\ (\varrho_1^2 + \varrho_2) \varsigma_1 - \varrho_1 \varrho_2 \varsigma_0 - \varrho_1 \varsigma_2 & \varsigma_2 - \varrho_1 (2\varsigma_1 + \varrho_1 \varsigma_0) \end{pmatrix},$$

这里 σ_{tt}^2 为 $\text{diag}(\Sigma_0)$ 的第 t 个元素.

以上初始估计 $(\hat{\beta}_n^T, \hat{\theta}_n^T)^T$ 、 $\hat{g}(u; \hat{\beta}_n, \hat{\theta}_n)$ 和 $\hat{g}'(u; \hat{\beta}_n, \hat{\theta}_n)$ 没有考虑序列相关误差结构, 因此不是渐近有效估计. 下面通过引入估计的误差方差及相关误差构造更加有效的估计.

3 序列自相关误差结构的拟合

基于初始估计 $(\hat{\beta}_n^T, \hat{\theta}_n^T)^T$ 和 $\hat{g}(u; \hat{\beta}_n, \hat{\theta}_n)$ 可得到如下残差的估计:

$$\hat{\varepsilon}_{it} = Y_{it} - \mathbf{X}_{it}^T \hat{\beta}_n - \hat{g}(\mathbf{Z}_{it}^T \hat{\theta}_n; \hat{\beta}_n, \hat{\theta}_n), \quad i = 1, \dots, n, \quad t = 1, \dots, T.$$

定义

$$\begin{aligned}\hat{Q}_{0,n} &= \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} (\hat{\varepsilon}_{i,t+(s-1)}, \dots, \hat{\varepsilon}_{it})^T (\hat{\varepsilon}_{i,t+(s-1)}, \dots, \hat{\varepsilon}_{it}), \\ \hat{Q}_{1,n} &= \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} (\hat{\varepsilon}_{i,t+(s-1)}, \dots, \hat{\varepsilon}_{it})^T (\hat{\varepsilon}_{i,t+s}, \dots, \hat{\varepsilon}_{i,t+1}),\end{aligned}$$

$$\hat{\mathbf{Q}}_{2,n} = \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} (\hat{\varepsilon}_{i,t+(s-1)}, \dots, \hat{\varepsilon}_{it})^T \hat{\varepsilon}_{i,t+s},$$

$$\hat{\mathbf{Q}}_{3,n} = \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} (\hat{\varepsilon}_{i,t+(s-1)}, \dots, \hat{\varepsilon}_{it})^T \hat{\varepsilon}_{i,t+s+1},$$

则用

$$\hat{\boldsymbol{\rho}}_n = (\hat{\rho}_{1,n}, \dots, \hat{\rho}_{s,n})^T = (\hat{\mathbf{Q}}_{0,n} - \hat{\mathbf{Q}}_{1,n})^{-1} (\hat{\mathbf{Q}}_{2,n} - \hat{\mathbf{Q}}_{3,n})$$

估计 $\boldsymbol{\rho} = (\rho_1, \dots, \rho_s)^T$.

此外, 注意到

$$\varepsilon_{i,t+s} = \mu_i + \nu_{i,t+s} = \mu_i + \rho_1 \nu_{i,t+(s-1)} + \dots + \rho_s \nu_{it} + e_{i,t+s}.$$

因此,

$$\begin{aligned} \varepsilon_{i,t+s} - \rho_1 \varepsilon_{i,t+(s-1)} - \dots - \rho_s \varepsilon_{it} \\ = (\mu_i - \rho_1 \mu_i - \dots - \rho_s \mu_i) + \nu_{i,t+s} - \rho_1 \nu_{i,t+(s-1)} - \dots - \rho_s \nu_{it} \\ = (\mu_i - \rho_1 \mu_i - \dots - \rho_s \mu_i) + e_{i,t+s}. \end{aligned}$$

记 $\ell_{it} = \varepsilon_{i,t+s} - \rho_1 \varepsilon_{i,t+(s-1)} - \dots - \rho_s \varepsilon_{it}$, 对于 $t = 1, \dots, T-s$. 基于以下事实:

$$\sigma_e^2 = E(\ell_{it}^2) - E(\ell_{it} \ell_{i,t+1}), \quad \sigma_\mu^2 = \rho_c^{-2} E(\ell_{it} \ell_{i,t+1}), \quad \rho_c = 1 - \rho_1 - \dots - \rho_s,$$

对于 $t = 1, \dots, T-(s+1)$ 和 $\hat{\rho}_c = 1 - \hat{\rho}_{1,n} - \dots - \hat{\rho}_{s,n}$, 定义 σ_e^2 和 σ_μ^2 的矩估计量分别为

$$\hat{\sigma}_{e,n}^2 = \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} \hat{\ell}_{it}^2 - \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} \hat{\ell}_{it} \hat{\ell}_{i,t+1},$$

$$\hat{\sigma}_{\mu,n}^2 = \frac{\hat{\rho}_c^{-2}}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} \hat{\ell}_{it} \hat{\ell}_{i,t+1},$$

其中 $\hat{\ell}_{it} = \hat{\varepsilon}_{i,t+s} - \hat{\rho}_{1,n} \hat{\varepsilon}_{i,t+(s-1)} - \dots - \hat{\rho}_{s,n} \hat{\varepsilon}_{it}$.

下面的定理给出估计量 $(\hat{\rho}_{1,n}, \dots, \hat{\rho}_{s,n})^T$ 、 $\hat{\sigma}_{e,n}^2$ 和 $\hat{\sigma}_{\mu,n}^2$ 的渐近性质.

定理 3.1 令假设 2.1-2.6 成立, 当 $n \rightarrow \infty$ 时,

(i) $\sqrt{nT}(\hat{\boldsymbol{\rho}}_n - \boldsymbol{\rho}) \rightarrow_D N(0, \frac{2T}{T-(s+1)} \sigma_e^2 \boldsymbol{\Sigma}_{\rho 1} \boldsymbol{\Sigma}_{\rho 2} \boldsymbol{\Sigma}_{\rho 1}^{-1})$, 其中

$$\boldsymbol{\Sigma}_{\rho 1} = \begin{pmatrix} \gamma_\nu(0) & \gamma_\nu(1) & \cdots & \gamma_\nu(s-1) \\ \gamma_\nu(1) & \gamma_\nu(0) & \cdots & \gamma_\nu(s-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_\nu(s-1) & \gamma_\nu(s-2) & \cdots & \gamma_\nu(0) \end{pmatrix} - \begin{pmatrix} \gamma_\nu(1) & \gamma_\nu(2) & \cdots & \gamma_\nu(s) \\ \gamma_\nu(0) & \gamma_\nu(1) & \cdots & \gamma_\nu(s-1) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_\nu(s) & \gamma_\nu(s-1) & \cdots & \gamma_\nu(1) \end{pmatrix},$$

$$\boldsymbol{\Sigma}_{\rho 2} = \sigma_\mu^2 \mathbf{1}_s \mathbf{1}_s^T + \begin{pmatrix} \gamma_\nu(0) & \gamma_\nu(1) & \cdots & \gamma_\nu(s-1) \\ \gamma_\nu(1) & \gamma_\nu(0) & \cdots & \gamma_\nu(s-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_\nu(s-1) & \gamma_\nu(s-2) & \cdots & \gamma_\nu(0) \end{pmatrix},$$

及 $\gamma_\nu(j) = \text{E}(\nu_{it}\nu_{i,t+j})$, $j = 0, 1, \dots, s-1$, $\mathbf{1}_s$ 是元素都为 1 的 s 维列向量.

(ii) $\sqrt{nT}(\hat{\sigma}_{e,n}^2 - \sigma_e^2) \rightarrow_D N(0, \omega_e)$, 其中

$$\omega_e = \frac{T}{(T-(s+1))^2} \{ \text{Var}(e_{it}^2) + 2(1-\rho_1-\dots-\rho_s)^2 \sigma_\mu^2 \sigma_e^2 + \sigma_e^4 \}.$$

(iii) $\sqrt{nT}(\hat{\sigma}_{\mu,n}^2 - \sigma_\mu^2) \rightarrow_D N(0, \omega_\mu)$, 特别地, 对于 $s=1$ 有如下表示:

$$\begin{aligned} \omega_\mu &= \frac{T}{\{\rho(2-T) + T(1-\rho)\}^2} \text{Var} \left\{ (\sqrt{1-\rho^2}(\mu_i + \nu_{i1}), (1-\rho)\mu_i + e_{i2}, \dots, (1-\rho)\mu_i + e_{iT}) \right. \\ &\quad \times \left. \left(\bar{\mathbf{J}}_T^\rho - \frac{1}{T-1} \mathbf{E}_T^\rho \right) (\sqrt{1-\rho^2}(\mu_i + \nu_{i1}), (1-\rho)\mu_i + e_{i2}, \dots, (1-\rho)\mu_i + e_{iT})^\top \right\}, \end{aligned}$$

其中 $\mathbf{E}_T^\rho = \mathbf{I}_T - \bar{\mathbf{J}}_T^\rho$, 这里 $\bar{\mathbf{J}}_T^\rho = 1/[(T-1)+(1+\rho)/(1-\rho)]\mathbf{1}_T\mathbf{1}_T^\top$, $\mathbf{1}_T = (\sqrt{(1+\rho)/(1-\rho)}, \mathbf{1}_{T-1}^\top)^\top$, \mathbf{I}_T 表示 T 维单位向量, $\mathbf{1}_T$ 表示元素都为 1 的 T 维列向量.

接下来, 将研究如何利用本节得到的残差和自相关误差结构的估计来提高模型 (1.1) 中线性部分和单指标中的未知参数分量 β 和 θ 以及未知连接函数 $g(\cdot)$ 的估计有效性.

4 可行的加权广义估计方程

记 $\hat{\Sigma}_{0n}^{-1} \triangleq (\hat{\sigma}_{\mu,n}^2 \mathbf{J}_T + \hat{\Psi}_n)^{-1}$, 其中 $\mathbf{J}_T = \mathbf{1}_T \mathbf{1}_T^\top$ 表示 Σ_0^{-1} 的估计, 则考虑协方差结构得到模型 (1.1) 中参数分量的可行的加权纠偏广义估计方程如下:

$$\begin{aligned} 0 &= \frac{1}{nT} \sum_{i=1}^n \begin{pmatrix} \hat{g}'(\mathbf{Z}_{i1}^\top \theta) \{ \mathbf{Z}_{-1,i1}^\top - \hat{\text{E}}(\mathbf{Z}_{-1,i1}^\top | \mathbf{Z}_{i1}^\top \theta) \} \\ \vdots \\ \hat{g}'(\mathbf{Z}_{iT}^\top \theta) \{ \mathbf{Z}_{-1,iT}^\top - \hat{\text{E}}(\mathbf{Z}_{-1,iT}^\top | \mathbf{Z}_{iT}^\top \theta) \} \end{pmatrix}^\top \hat{\Sigma}_{0n}^{-1} \begin{pmatrix} Y_{i1} - \mathbf{X}_{i1}^\top \beta - \hat{g}(\mathbf{Z}_{i1}^\top \theta) \\ \vdots \\ Y_{iT} - \mathbf{X}_{iT}^\top \beta - \hat{g}(\mathbf{Z}_{iT}^\top \theta) \end{pmatrix}, \\ 0 &= \frac{1}{nT} \sum_{i=1}^n \begin{pmatrix} \{ \mathbf{X}_{i1}^\top - \hat{\text{E}}(\mathbf{X}_{i1}^\top | \mathbf{Z}_{i1}^\top \theta) \} \\ \vdots \\ \{ \mathbf{X}_{iT}^\top - \hat{\text{E}}(\mathbf{X}_{iT}^\top | \mathbf{Z}_{iT}^\top \theta) \} \end{pmatrix}^\top \hat{\Sigma}_{0n}^{-1} \begin{pmatrix} Y_{i1} - \mathbf{X}_{i1}^\top \beta - \hat{g}(\mathbf{Z}_{i1}^\top \theta) \\ \vdots \\ Y_{iT} - \mathbf{X}_{iT}^\top \beta - \hat{g}(\mathbf{Z}_{iT}^\top \theta) \end{pmatrix}. \end{aligned} \quad (4.1)$$

假定 $(\hat{\beta}_n^{w\tau}, \hat{\theta}_n^{w\tau})$ 是以上估计方程 (4.1) 的解, 则称 $(\hat{\beta}_n^{w\tau}, \hat{\theta}_n^{w\tau})$ 是未知参数分量 β 和 θ 的可行加权广义估计 (GEE-FW), 并给出如下渐近性质.

定理 4.1 令假设 2.1–2.6 成立, 当 $n \rightarrow \infty$ 时,

$$\sqrt{nT} \{ (\hat{\beta}_n^{w\tau}, \hat{\theta}_n^{w\tau})^\top - (\beta^\top, \theta^\top)^\top \} \rightarrow_D N(0, \Omega_3^{-1}),$$

其中

$$\begin{aligned} \Omega_3 &= \begin{pmatrix} \Omega_{3(11)} & \Omega_{3(12)} \\ \Omega_{3(21)} & \Omega_{3(22)} \end{pmatrix}, \\ \Omega_{3(11)} &= T^{-1} \text{E}(\Pi_X \Sigma_0^{-1} \Pi_X^\top), \quad \Omega_{3(12)} = T^{-1} \text{E}(\Pi_X \Sigma_0^{-1} \Pi_Z^\top), \\ \Omega_{3(21)} &= T^{-1} \text{E}(\Pi_Z \Sigma_0^{-1} \Pi_X^\top), \quad \Omega_{3(22)} = T^{-1} \text{E}(\Pi_Z \Sigma_0^{-1} \Pi_Z^\top). \end{aligned}$$

注 4.1 记 $\mathbf{O} = \boldsymbol{\Omega}_1^{-1} \mathbf{\Pi}_i^T \boldsymbol{\Sigma}_{b0}^{1/2} - \boldsymbol{\Omega}_3^{-1} \mathbf{\Pi}_i^T \boldsymbol{\Sigma}_{b0}^{-1/2}$, $\mathbf{\Pi}_i = (\mathbf{\Pi}_X^T \mathbf{\Pi}_Z^T)$, $\boldsymbol{\Sigma}_{b0} = \text{blkdiag}(\boldsymbol{\Sigma}_0, \boldsymbol{\Sigma}_0)$, 其中 blkdiag 表示分块对角阵, 则 $\boldsymbol{\Omega}_1 = E(\mathbf{\Pi}_i^T \mathbf{\Pi}_i)$, $\boldsymbol{\Omega}_2 = E(\mathbf{\Pi}_i^T \boldsymbol{\Sigma}_{b0} \mathbf{\Pi}_i)$, $\boldsymbol{\Omega}_3 = E(\mathbf{\Pi}_i^T \boldsymbol{\Sigma}_{b0}^{-1} \mathbf{\Pi}_i)$ 且

$$\mathbf{O}\mathbf{O}^T = \boldsymbol{\Omega}_1^{-1} \mathbf{\Pi}_i^T \boldsymbol{\Sigma}_{b0} \mathbf{\Pi}_i \boldsymbol{\Omega}_1^{-1} - \boldsymbol{\Omega}_3^{-1} \mathbf{\Pi}_i^T \mathbf{\Pi}_i \boldsymbol{\Omega}_1^{-1} - \boldsymbol{\Omega}_1^{-1} \mathbf{\Pi}_i^T \mathbf{\Pi}_i \boldsymbol{\Omega}_3^{-1} + \boldsymbol{\Omega}_3^{-1} \mathbf{\Pi}_i^T \boldsymbol{\Sigma}_{b0}^{-1} \mathbf{\Pi}_i \boldsymbol{\Omega}_3^{-1}.$$

由于 $\mathbf{O}\mathbf{O}^T$ 为非正定矩阵, 则

$$0 \leq T^{-1} E(\mathbf{O}\mathbf{O}^T) = \boldsymbol{\Omega}_1^{-1} \boldsymbol{\Omega}_2 \boldsymbol{\Omega}_1^{-1} - \boldsymbol{\Omega}_3^{-1} \boldsymbol{\Omega}_1 \boldsymbol{\Omega}_1^{-1} - \boldsymbol{\Omega}_1^{-1} \boldsymbol{\Omega}_1 \boldsymbol{\Omega}_3^{-1} + \boldsymbol{\Omega}_3^{-1} \boldsymbol{\Omega}_3 \boldsymbol{\Omega}_3^{-1} = \boldsymbol{\Omega}_1^{-1} \boldsymbol{\Omega}_2 \boldsymbol{\Omega}_1^{-1} - \boldsymbol{\Omega}_3^{-1},$$

因此, $\boldsymbol{\Omega}_3^{-1} \leq \boldsymbol{\Omega}_1^{-1} \boldsymbol{\Omega}_2 \boldsymbol{\Omega}_1^{-1}$. 这表明加权广义估计 $\hat{\beta}_n^w$ 具有更小的渐近协方差矩阵, 因此比未加权的初始广义估计 $\hat{\beta}_n$ 更加渐近有效.

5 非参数函数的两阶段估计

除了模型中的未知连接函数 $g(u)$ 外, 在一些情形下, 其导函数 $g'(u)$ 也是感兴趣的函数, 可参见文献 [26]. 本节将研究未知连接函数及其导函数 $(g(u), g'(u))^T$ 的有效估计. 在上一节得到的更加有效的参数估计基础上, 结合序列相关误差对连接函数估计的影响, 本节提出考虑自相关结构的未知连接函数的两阶段局部线性估计 (GEE-TS), 并且建立其渐近分布. You 和 Zhou^[27] 基于相关结构完全已知情形, 针对不含任何参数分量的非参数可加模型, 提出了函数的两阶段估计方法, 但只建立了 $g(u)$ 的估计量的渐近性质. 本文所提出的两阶段局部线性估计拓展了 You 和 Zhou^[27] 的结果. 类似 Fan 和 Zhang^[28], 我们所建立的函数向量 $(g(u), g'(u))^T$ 的估计量的渐近分布可用来构建单指标函数的同时置信带, 也可用来检查估计的函数是否显著或者检查可加函数是否为常数等.

设 $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{iT})^T$, $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \dots, \varepsilon_{iT})^T$, $\boldsymbol{\Sigma}_0 = E(\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^T) = \sigma_\mu^2 \mathbf{1}_T \mathbf{1}_T^T + \sigma_e^2 \mathbf{V} = (\sigma_{t_1 t_2}^2)_{t_1, t_2=1}^T$, $\boldsymbol{\Sigma}_0^{-1} = (\sigma^{t_1 t_2})_{t_1, t_2=1}^T$, 则 $\boldsymbol{\Sigma}_0^{-1} \mathbf{Y}_i$ 的第 t 个元素有如下形式:

$$\tilde{Y}_{it} = \sum_{t_1=1}^T \sigma^{tt_1} Y_{it_1} = \sigma^{tt} Y_{it} + \sum_{t_1 \neq t} \sigma^{tt_1} Y_{it_1}.$$

记 $a_{it} = \mathbf{Z}_{it}^T \boldsymbol{\beta} + g(\mathbf{Z}_{it}^T \boldsymbol{\theta})$, 则有

$$(\sigma^{tt})^{-1} \left(\tilde{Y}_{it} - \sigma^{tt} \mathbf{X}_{it}^T \boldsymbol{\beta} - \sum_{t_1 \neq t} \sigma^{tt_1} a_{it_1} \right) = g(\mathbf{Z}_{it}^T \boldsymbol{\theta}) + (\sigma^{tt})^{-1} \sum_{t_1=1}^T \sigma^{tt_1} \varepsilon_{it_1}.$$

设 $\tilde{Y}_{it}^* = (\sigma^{tt})^{-1} (\tilde{Y}_{it} - \sigma^{tt} \mathbf{X}_{it}^T \boldsymbol{\beta} - \sum_{t_1 \neq t} \sigma^{tt_1} a_{it_1})$. 由于

$$E(\tilde{Y}_{it}^* | \mathbf{Z}_{it}^T \boldsymbol{\theta}) = g(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \quad \text{和} \quad \text{Var} \left\{ (\sigma^{tt})^{-1} \sum_{t_1=1}^T \sigma^{tt_1} \varepsilon_{it_1} \right\} = (\sigma^{tt})^{-1} \leq \sigma_{tt}^2$$

(上述不等式成立可参见文献 [29, 推论 1(iii)]), 故对 \tilde{Y}_{it}^* 应用局部多项式估计, 可得到具有自相关误差结构的面板数据部分线性单指标模型 (1.1) 和 (1.2) 中未知连接函数 $g(\cdot)$ 的更加有效估计.

记 $U_{it} = \mathbf{Z}_{it}^T \boldsymbol{\theta}$, 则 U_{it} 在 u 的小邻域内 $g(U_{it})$ 可被近似为

$$g(U_{it}) \approx g(u) + g'(u)(U_{it} - u) \equiv a + b(U_{it} - u).$$

求解如下局部最小二乘问题, 即最小化如下目标函数得到 $\{(a, b)\}$ 的估计:

$$\sum_{i=1}^n \sum_{t=1}^T [\tilde{Y}_{it}^* - \{a + b(U_{it} - u)\}]^2 K_h(U_{it} - u), \quad (5.1)$$

其中 $K_h(\cdot)$ 为核函数, $K_h(\cdot) = h^{-1}K(\cdot/h)$, h 为窗宽. 经过简单计算得到 (5.1) 的解:

$$(\tilde{a}, \tilde{b})^T = (\mathbf{D}_u^T \mathbf{W}_u \mathbf{D}_u)^{-1} \mathbf{D}_u \mathbf{W}_u \tilde{\mathbf{Y}}^*,$$

其中

$$\begin{aligned} \tilde{\mathbf{Y}}^* &= (\tilde{Y}_{11}^*, \dots, \tilde{Y}_{1T}^*, \dots, \tilde{Y}_{nT}^*)^T, \\ \mathbf{W}_u &= \text{diag}(K_h(U_{11} - u), \dots, K_h(U_{1T} - u), \dots, K_h(U_{nT} - u)), \\ \mathbf{D}_u &= \begin{pmatrix} 1 & \cdots & 1 & \cdots & 1 \\ (U_{11} - u) & \cdots & (U_{1T} - u) & \cdots & (U_{nT} - u) \end{pmatrix}^\tau. \end{aligned}$$

在实际应用中, a_{it} 、 β 、 θ 、 $g(\cdot)$ 和 $\sigma^{t_1 t_2}$ 往往都是未知的, 故需要用它们的估计来代替近似. 对于 \tilde{Y}_{it}^* , 可采用 $\hat{a}_{it} = \mathbf{X}_{it}^T \hat{\beta}_n^w + \hat{g}_n(\mathbf{Z}_{it}^T \hat{\theta}_n^w)$ 、 $\hat{\beta}_n$ 、 $\hat{\theta}_n$ 、 $\hat{g}_n(\cdot)$ 和 $\hat{\sigma}^{t_1 t_2}$ 来代替未知量, 从而得到其估计量 $\hat{\mathbf{Y}}_{it}^*$, 即

$$\hat{Y}_{it}^* = (\hat{\sigma}^{tt})^{-1} \left(\hat{Y}_{it} - \hat{\sigma}^{tt} \mathbf{X}_{it}^T \hat{\beta} - \sum_{t_1 \neq t}^T \hat{\sigma}^{tt_1} \hat{a}_{it_1} \right),$$

其中 $\hat{Y}_{it} = \sum_{t_1=1}^T \hat{\sigma}^{tt_1} Y_{it_1}$. 因此得到 $(g(u), g'(u))$ 的一个可行的两阶段估计

$$(g_n^{TS}(u), \hat{g}_n'^{TS}(u))^T = (\mathbf{D}_u^T \mathbf{W}_u \mathbf{D}_u)^{-1} \mathbf{D}_u \mathbf{W}_u \hat{\mathbf{Y}}^*,$$

其中 $\hat{\mathbf{Y}}^* = (\hat{Y}_{11}^*, \dots, \hat{Y}_{1T}^*, \dots, \hat{Y}_{nT}^*)^T$.

下面给出未知函数向量的两阶段估计量 $(\hat{g}_n^{TS}(u), \hat{g}_n'^{TS}(u))^T$ 的渐近性质.

定理 5.1 令假设 2.1–2.6 成立, 则当 $n \rightarrow \infty$ 时,

$$\sqrt{nTh} \left[\mathbf{H} \left\{ \begin{pmatrix} \hat{g}_n^{TS}(u; \hat{\beta}_n, \hat{\theta}_n) \\ \hat{g}_n'^{TS}(u; \hat{\beta}_n, \hat{\theta}_n) \end{pmatrix} - \begin{pmatrix} g(u) \\ g'(u) \end{pmatrix} \right\} - \frac{h^2}{2} \frac{1}{\varrho_2 - \varrho_1^2} \begin{pmatrix} (\varrho_2^2 - \varrho_1 \varrho_3) g''(u) \\ (\varrho_3 - \varrho_1 \varrho_2) g''(u) \end{pmatrix} \right] \xrightarrow{D} N(0, \Phi^{TS}),$$

其中 $\mathbf{H} = \text{diag}(1, h)$, $g''(u) = \partial^2 g(u) / \partial u^2$,

$$\Phi^{TS} = \frac{\sum_{t=1}^T (\sigma^{tt})^{-1} p_t(u)}{(\sum_{t=1}^T p_t(u))^2 (\varrho_2 - \varrho_1^2)^2} \begin{pmatrix} \varrho_2^2 \varsigma_0 - 2\varrho_1 \varrho_2 \varsigma_1 + \varrho_1^2 \varsigma_2 & (\varrho_1^2 + \varrho_2) \varsigma_1 - \varrho_1 \varrho_2 \varsigma_0 - \varrho_1 \varsigma_2 \\ (\varrho_1^2 + \varrho_2) \varsigma_1 - \varrho_1 \varrho_2 \varsigma_0 - \varrho_1 \varsigma_2 & \varsigma_2 - \varrho_1 (2\varsigma_1 + \varrho_1 \varsigma_0) \end{pmatrix}.$$

推论 5.1 令假设 2.1–2.6 成立, 当 $n \rightarrow \infty$ 时,

$$\sqrt{nTh} \left\{ \hat{g}_n^{TS}(u) - g(u) - \frac{h^2}{2} \frac{\varrho_2^2 - \varrho_1 \varrho_3}{\varrho_2 - \varrho_1^2} g''(u) \right\} \xrightarrow{D} N(0, \sigma^{TS}),$$

其中

$$\sigma^{TS} = (\varrho_2^2 \varsigma_0 - 2\varrho_1 \varrho_2 \varsigma_1 + \varrho_1^2 \varsigma_2) \left\{ \left(\sum_{t=1}^T p_t(u) \right)^2 (\varrho_2 - \varrho_1^2)^2 \right\}^{-1} \sum_{t=1}^T (\sigma^{tt})^{-1} p_t(u).$$

注 5.1 为应用推论 5.1 对函数 $g(\cdot)$ 做统计推断, 需 σ^{ts} 的相合估计. 由于核函数给定时 $\varsigma_0, \varsigma_1, \varsigma_2, \varrho_1, \varrho_2$ 和 ϱ_3 都是已知常数, 故只需估计 σ^{tt} 和 $p_t(\cdot)$ ($t = 1, \dots, T$). 注意到 $\Sigma_{0n}^{-1} = (\sigma^{t_1 t_2})_{t_1, t_2=1}^T$, 其相应估计

$$\hat{\Sigma}_{0n}^{-1} = (\hat{\sigma}_{\mu, n}^{t_1 t_2})_{t_1, t_2=1}^T = (\hat{\sigma}_{\mu, n}^2 \mathbf{J}_T + \hat{\Psi}_n)^{-1},$$

其中 $\mathbf{J}_T = \mathbf{1}_T \mathbf{1}_T^T$. 根据定理 3.1 可知, $\hat{\sigma}^{tt}$ 是 σ^{tt} 的相合估计. 对于 $p_t(\cdot)$, 可用普通的核密度估计方法来估计, 即 $\hat{p}_t(\cdot) = \frac{1}{h} \sum_{i=1}^n K_h(U_{it} - \cdot)$.

6 模拟研究

本节通过随机模拟研究, 给出本文所提出的估计方法在有限样本量下的表现. 考虑从如下具有有序相关误差结构的面板数据部分线性单指标模型中产生数据:

$$Y_{it} = X_{it,1}\beta_1 + X_{it,2}\beta_2 + X_{it,3}\beta_3 + g(Z_{it,1} + Z_{it,2}\theta_1 + Z_{it,3}\theta_2 + Z_{it,4}\theta_3) + \varepsilon_{it},$$

误差

$$\varepsilon_{it} = \mu_i + \nu_{it}, \quad \nu_{it} = \rho\nu_{i,t-1} + e_{it}, \quad i = 1, \dots, n, \quad t = 1, \dots, T,$$

其中感兴趣的参数 $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)^T = (1.5, 0.5, -0.5)^T$ 和 $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)^T = (2, -1, 1.5)^T$; 协变量 $X_{it,1}, X_{it,2}$ 和 $X_{it,3}$ 是 i 和 t 独立且 $X_{it,1}$ 和 $X_{it,2}$ 来自均匀分布 $U(1, 2)$ 以及 $X_{it,3}$ 来自成功概率为 0.5 的 Bernoulli 分布; 协变量 $\mathbf{Z}_{it} = (Z_{it,1}, Z_{it,2}, Z_{it,3}, Z_{it,4})^T$ 独立产生于 q ($q = 4$) 维正态分布, 其均值为 0, 协方差矩阵为 $(\sigma_{ij})_{q \times q}$, 其中, 对于 $i = j$, 有 $\sigma_{ii} = 1$, 并且对于 $i \neq j$, 有 $\sigma_{ij} = 0.5$; $\mu_i \sim N(0, \sigma_\mu^2)$, $\nu_{it} \sim N(0, \sigma_e^2)$, $\nu_{i0} \sim N(0, \sigma_e^2)$, $\sigma_\mu^2 = 1$, $\sigma_e^2 = 1$, $\sigma^2 = \sigma_e^2/(1 - \rho^2)$. 为了比较和节省空间, 选取 n 分别为 150、200 和 300, T 分别为 5 和 10, ρ 分别为 0.3 和 0.6. 窗宽采用广义交叉验证 (generalized cross-validation, GCV) 方法选择:

$$GCV(h) = \frac{1}{h} \sum_{i=1}^n \sum_{t=1}^T \frac{\{Y_{it} - \mathbf{X}_{it}^T \hat{\boldsymbol{\beta}}_n - \hat{g}_h(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n)\}^2}{\{n^{-1} \text{tr}(\mathbf{I} - \mathbf{S}_h)\}},$$

其中 \hat{g}_h 是窗宽为 h 的 $g(\cdot)$ 的估计, \mathbf{S}_h 是窗宽为 h 的平滑矩阵. 最小化 $GCV(h)$ 可得到估计 $g(\cdot)$ 的窗宽, $h_{\text{opt}} = cn^{-1/5}$, 其中 $c > 0$. 因此用 h_{opt} 作为估计 $\hat{g}(\cdot)$ 的最优窗宽. 当计算 $\hat{\boldsymbol{\theta}}$ 时, 选择如下的窗宽: $h = h_{\text{opt}} n^{-1/20} (\log n)^{-1/2}$, $h_1 = h_{\text{opt}}$. 由于 $h = cn^{-1/4} (\log n)^{1/2}$, $h_1 = cn^{-1/5}$, 其中 $c > 0$, 因此, 窗宽满足假设 2.5(ii). 模拟以及之后的实证分析均采用 Gauss 核函数: $K_h(\cdot) = \frac{1}{h\sqrt{2\pi}} \exp\left\{-\frac{(\cdot)^2}{2h^2}\right\}$, 其中 h 表示窗宽. 所有情形的模拟结果都基于 1,000 次重复实验. 采用如下指标评价参数估计的精度: 估计偏差 (bias)、估计标准差 (std)、估计的标准误差的均值 (ste) 和均方误差 (mse). 表 1 和 2 分别汇总了 T 为 5 和 10、参数分量基于工作独立的 GEE 估计 (GEE-WI)、可行的加权 GEE 估计 (GEE-FW) 和对应的假定真实协方差模型的基准估计 (generalized estimating equation benchmark estimation based on true covariance, GEE-BW) 的结果. 连接函数的估计精确度可由根号均方误差 (the square-root of averaged squared errors, RASE) 评估: $RASE(\hat{g}) = \left\{ \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T [\hat{g}(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}) - g(\mathbf{Z}_{it}^T \boldsymbol{\theta})]^2 g \right\}^{\frac{1}{2}}$.

表 3 分别给出了连接函数的基于工作独立的不加权 GEE 估计 (GEE-WI)、可行加权 GEE 估计 (GEE-FW) 和两阶段局部线性估计 (GEE-TS) 的 RASE 的样本中位数. 误差序列自相关自回归系数和误差方差的参数估计量 $\hat{\rho}, \hat{\sigma}_e^2$ 和 $\hat{\sigma}_\mu^2$ 的有限样本表现在表 4 中给出.

表 1 参数估计的模拟结果, $T = 5$

n			$\rho = 0.3$						$\rho = 0.6$					
			β_1	β_2	β_3	θ_1	θ_2	θ_3	β_1	β_2	β_3	θ_1	θ_2	θ_3
150	GEE-WI	bias	-0.0064	-0.0008	0.0003	0.0681	-0.0283	0.0553	-0.0083	-0.0004	0.0013	0.0938	-0.0369	0.0743
		std	0.1732	0.1890	0.1052	0.4178	0.2003	0.3314	0.1910	0.2063	0.1159	0.4931	0.2374	0.3833
		ste	0.1714	0.1829	0.1022	0.3654	0.1767	0.2841	0.1879	0.2015	0.1123	0.4114	0.1971	0.3181
		mse	0.0300	0.0357	0.0111	0.1792	0.0409	0.1129	0.0366	0.0426	0.0134	0.2519	0.0577	0.1524
	GEE-FW	bias	-0.0092	0.0008	0.0052	0.0431	-0.0207	0.0413	-0.0113	0.0008	0.0076	0.0436	-0.0216	0.0416
		std	0.1282	0.1298	0.0763	0.2662	0.1310	0.2159	0.1164	0.1170	0.0701	0.2487	0.1203	0.2023
		ste	0.1281	0.1325	0.0746	0.2514	0.1220	0.2000	0.1190	0.1227	0.0688	0.2309	0.1107	0.1837
		mse	0.0165	0.0168	0.0058	0.0727	0.0176	0.0483	0.0137	0.0137	0.0050	0.0637	0.0149	0.0426
	GEE-BW	bias	-0.0092	0.0003	0.0053	0.0432	-0.0209	0.0414	-0.0114	0.0004	0.0081	0.0437	-0.0221	0.0418
		std	0.1274	0.1301	0.0759	0.2644	0.1304	0.2144	0.1158	0.1174	0.0695	0.2479	0.1196	0.2005
		ste	0.1278	0.1322	0.0744	0.2510	0.1217	0.1996	0.1161	0.1196	0.0671	0.2253	0.1077	0.1793
		mse	0.0163	0.0169	0.0058	0.0718	0.0174	0.0477	0.0135	0.0138	0.0049	0.0633	0.0148	0.0419
200	GEE-WI	bias	-0.0021	-0.0028	0.0043	0.0592	-0.0290	0.0495	-0.0019	-0.0044	0.0046	0.0772	-0.0387	0.0675
		std	0.1573	0.1586	0.0911	0.3610	0.1715	0.2954	0.1776	0.1743	0.1011	0.4299	0.2002	0.3538
		ste	0.1549	0.1575	0.0885	0.3122	0.1536	0.2502	0.1704	0.1735	0.0973	0.3492	0.1712	0.2794
		mse	0.0247	0.0252	0.0083	0.1338	0.0302	0.0897	0.0315	0.0304	0.0102	0.1908	0.0416	0.1297
	GEE-FW	bias	-0.0049	-0.0005	0.0052	0.0469	-0.0197	0.0417	-0.0068	-0.0002	0.0061	0.0489	-0.0213	0.0434
		std	0.1124	0.1092	0.0663	0.2387	0.1108	0.1959	0.1034	0.0997	0.0597	0.2210	0.1008	0.1826
		ste	0.1134	0.1138	0.0642	0.2170	0.1057	0.1768	0.1049	0.1048	0.0589	0.1986	0.0960	0.1622
		mse	0.0127	0.0119	0.0044	0.0592	0.0127	0.0401	0.0107	0.0099	0.0036	0.0512	0.0106	0.0352
	GEE-BW	bias	-0.0054	-0.0001	0.0054	0.0466	-0.0193	0.0413	-0.0075	0.0003	0.0065	0.0491	-0.0211	0.0432
		std	0.1119	0.1086	0.0661	0.2372	0.1102	0.1956	0.1027	0.0994	0.0594	0.2196	0.1002	0.1825
		ste	0.1131	0.1134	0.0639	0.2163	0.1053	0.1764	0.1027	0.1026	0.0577	0.1945	0.0939	0.1589
		mse	0.0125	0.0118	0.0044	0.0584	0.0125	0.0400	0.0106	0.0099	0.0036	0.0506	0.0105	0.0352
300	GEE-WI	bias	-0.0073	-0.0006	0.0029	0.0461	-0.0135	0.0355	-0.0067	-0.0011	0.0043	0.0521	-0.0152	0.0416
		std	0.1259	0.1314	0.0751	0.2756	0.1327	0.2368	0.1386	0.1447	0.0829	0.3152	0.1523	0.2691
		ste	0.1280	0.1286	0.0734	0.2528	0.1237	0.2049	0.1408	0.1414	0.0808	0.2790	0.1363	0.2258
		mse	0.0159	0.0173	0.0057	0.0781	0.0178	0.0574	0.0193	0.0209	0.0069	0.1021	0.0234	0.0741
	GEE-FW	bias	-0.0067	0.0031	0.0027	0.0311	-0.0126	0.0263	-0.0060	0.0047	0.0037	0.0307	-0.0135	0.0255
		std	0.0895	0.0928	0.0506	0.1922	0.0901	0.1620	0.0822	0.0840	0.0463	0.1764	0.0817	0.1488
		ste	0.0924	0.0925	0.0526	0.1750	0.0856	0.1445	0.0848	0.0848	0.0482	0.1594	0.0776	0.1321
		mse	0.0081	0.0086	0.0026	0.0379	0.0083	0.0269	0.0068	0.0071	0.0022	0.0320	0.0069	0.0228
	GEE-BW	bias	-0.0069	0.0036	0.0027	0.0311	-0.0125	0.0264	-0.0060	0.0052	0.0036	0.0311	-0.0138	0.0259
		std	0.0894	0.0924	0.0505	0.1918	0.0898	0.1611	0.0821	0.0837	0.0463	0.1761	0.0815	0.1478
		ste	0.0921	0.0923	0.0525	0.1746	0.0854	0.1443	0.0835	0.0833	0.0474	0.1569	0.0763	0.1302
		mse	0.0080	0.0086	0.0026	0.0377	0.0082	0.0267	0.0068	0.0070	0.0022	0.0320	0.0068	0.0225

表2 参数估计的模拟结果, $T = 10$

n	GEE-WI	$\rho = 0.3$						$\rho = 0.6$						
		β_1	β_2	β_3	θ_1	θ_2	θ_3	β_1	β_2	β_3	θ_1	θ_2	θ_3	
150	bias	0.0039	0.0056	0.0029	0.0184	-0.0094	0.0165	0.0052	0.0065	0.0046	0.0222	-0.0117	0.0202	
	std	0.1336	0.1238	0.0687	0.2728	0.1242	0.2134	0.1459	0.1356	0.0745	0.3094	0.1408	0.2434	
	ste	0.1312	0.1283	0.0705	0.2505	0.1176	0.1909	0.1446	0.1411	0.0771	0.2736	0.1285	0.2088	
	mse	0.0179	0.0154	0.0047	0.0748	0.0155	0.0458	0.0213	0.0184	0.0056	0.0962	0.0200	0.0596	
	GEE-FW	bias	0.0025	0.0039	0.0004	0.0119	-0.0060	0.0070	0.0032	0.0026	0.0006	0.0092	-0.0057	0.0042
	std	0.0913	0.0872	0.0499	0.1826	0.0846	0.1476	0.0832	0.0786	0.0447	0.1681	0.0781	0.1371	
200	ste	0.0913	0.0894	0.0507	0.1700	0.0814	0.1341	0.0837	0.0814	0.0460	0.1532	0.0736	0.1212	
	mse	0.0083	0.0076	0.0025	0.0335	0.0072	0.0218	0.0069	0.0062	0.0020	0.0283	0.0061	0.0188	
	GEE-BW	bias	0.0024	0.0037	0.0003	0.0121	-0.0061	0.0071	0.0031	0.0024	0.0005	0.0095	-0.0058	0.0042
	std	0.0912	0.0873	0.0499	0.1828	0.0849	0.1474	0.0830	0.0787	0.0448	0.1681	0.0782	0.1366	
	ste	0.0912	0.0892	0.0506	0.1698	0.0813	0.1340	0.0823	0.0800	0.0452	0.1506	0.0724	0.1192	
	mse	0.0083	0.0076	0.0025	0.0336	0.0072	0.0218	0.0069	0.0062	0.0020	0.0283	0.0062	0.0187	
	GEE-WI	bias	0.0029	-0.0037	0.0017	0.0055	-0.0060	0.0026	0.0011	-0.0045	0.0016	0.0102	-0.0077	0.0069
	std	0.1111	0.1127	0.0656	0.2173	0.1058	0.1767	0.1230	0.1245	0.0705	0.2418	0.1173	0.1995	
	ste	0.1130	0.1128	0.0631	0.2073	0.0992	0.1617	0.1248	0.1242	0.0691	0.2270	0.1085	0.1774	
	mse	0.0123	0.0127	0.0043	0.0473	0.0112	0.0312	0.0151	0.0155	0.0050	0.0586	0.0138	0.0398	
	GEE-FW	bias	0.0008	-0.0002	-0.0001	0.0078	-0.0075	0.0024	0.0004	-0.0002	-0.0003	0.0056	-0.0072	0.0012
	std	0.0773	0.0781	0.0438	0.1483	0.0722	0.1207	0.0705	0.0701	0.0393	0.1350	0.0654	0.1100	
	ste	0.0781	0.0773	0.0441	0.1413	0.0686	0.1122	0.0713	0.0703	0.0399	0.1269	0.0617	0.1010	
	mse	0.0060	0.0061	0.0019	0.0221	0.0053	0.0146	0.0050	0.0049	0.0015	0.0182	0.0043	0.0121	

(续表)

n		$\rho = 0.3$						$\rho = 0.6$					
		β_1	β_2	β_3	θ_1	θ_2	θ_3	β_1	β_2	β_3	θ_1	θ_2	θ_3
GEE-FW	bias	0.0008	-0.0002	-0.0001	0.0078	-0.0075	0.0024	0.0004	-0.0002	-0.0003	0.0056	-0.0072	0.0012
	std	0.0773	0.0781	0.0438	0.1483	0.0722	0.1207	0.0705	0.0701	0.0393	0.1350	0.0654	0.1100
	ste	0.0781	0.0773	0.0441	0.1413	0.0686	0.1122	0.0713	0.0703	0.0399	0.1269	0.0617	0.1010
	mse	0.0060	0.0061	0.0019	0.0221	0.0053	0.0146	0.0050	0.0049	0.0015	0.0182	0.0043	0.0121
GEE-BW	bias	0.0008	-0.0002	-0.0002	0.0079	-0.0074	0.0024	0.0005	-0.0002	-0.0003	0.0055	-0.0072	0.0011
	std	0.0773	0.0780	0.0437	0.1485	0.0723	0.1209	0.0705	0.0700	0.0393	0.1351	0.0655	0.1101
	ste	0.0780	0.0772	0.0440	0.1412	0.0685	0.1121	0.0703	0.0693	0.0394	0.1252	0.0608	0.0996
	mse	0.0060	0.0061	0.0019	0.0221	0.0053	0.0146	0.0050	0.0049	0.0015	0.0183	0.0043	0.0121
300 GEE-WI	bias	-0.0006	-0.0025	0.0005	0.0121	-0.0040	0.0028	-0.0001	-0.0027	0.0003	0.0152	-0.0051	0.0046
	std	0.0920	0.0905	0.0522	0.1870	0.0879	0.1483	0.1010	0.0997	0.0576	0.2137	0.0994	0.1712
	ste	0.0915	0.0901	0.0516	0.1744	0.0840	0.1377	0.1009	0.0992	0.0566	0.1910	0.0918	0.1513
	mse	0.0085	0.0082	0.0027	0.0351	0.0077	0.0220	0.0102	0.0099	0.0033	0.0459	0.0099	0.0293
GEE-FW	bias	-0.0000	-0.0029	-0.0018	0.0040	-0.0033	-0.0027	-0.0003	-0.0020	-0.0022	0.0019	-0.0023	-0.0052
	std	0.0626	0.0629	0.0360	0.1262	0.0603	0.0979	0.0565	0.0572	0.0320	0.1147	0.0552	0.0888
	ste	0.0633	0.0626	0.0359	0.1184	0.0581	0.0945	0.0574	0.0568	0.0324	0.1061	0.0520	0.0849
	mse	0.0039	0.0040	0.0013	0.0159	0.0036	0.0096	0.0032	0.0033	0.0010	0.0132	0.0030	0.0079
GEE-BW	bias	-0.0001	-0.0029	-0.0018	0.0039	-0.0032	-0.0029	-0.0004	-0.0020	-0.0022	0.0017	-0.0022	-0.0054
	std	0.0625	0.0628	0.0360	0.1259	0.0602	0.0975	0.0565	0.0572	0.0320	0.1144	0.0551	0.0885
	ste	0.0633	0.0626	0.0359	0.1183	0.0580	0.0944	0.0568	0.0562	0.0321	0.1050	0.0515	0.0840
	mse	0.0039	0.0040	0.0013	0.0159	0.0036	0.0095	0.0032	0.0033	0.0010	0.0131	0.0030	0.0079

表 3 连接函数估计的模拟结果

			$\rho = 0.3$			$\rho = 0.6$		
			$n = 150$	$n = 200$	$n = 300$	$n = 150$	$n = 200$	$n = 300$
$T = 5$	GEE-WI	sm(RASE)	0.3455	0.2983	0.2531	0.3792	0.3314	0.2777
	GEE-FW	sm(RASE)	0.2741	0.2502	0.2083	0.2756	0.2583	0.2126
	GEE-TS	sm(RASE)	0.2463	0.2248	0.1856	0.2255	0.2151	0.1744
	GEE-FW	ratio(RASE)	0.7934	0.8387	0.8231	0.7268	0.7793	0.7656
	GEE-TS	ratio(RASE)	0.7130	0.7535	0.7332	0.5946	0.6492	0.6279
$T = 10$	GEE-WI	sm(RASE)	0.2547	0.2161	0.1825	0.2808	0.2383	0.1999
	GEE-FW	sm(RASE)	0.2049	0.1861	0.1530	0.2100	0.1875	0.1593
	GEE-TS	sm(RASE)	0.1798	0.1595	0.1315	0.1698	0.1464	0.1258
	GEE-FW	ratio(RASE)	0.8046	0.8858	0.8382	0.7478	0.7868	0.7969
	GEE-TS	ratio(RASE)	0.7061	0.7592	0.7206	0.6046	0.6144	0.6293

注: 其中 sm(RASE) 表示 1,000 次重复 RASE 的样本中位数, ratio(RASE) 为 GEE-FW/GEE-TS 估计的 RASE 样本中位数除以 GEE-WI 估计的 RAESE 样本中位数

表 4 $\hat{\rho}$ 、 $\hat{\sigma}_e^2$ 和 $\hat{\sigma}_\mu^2$ 的样本实现结果

			$T = 5$			$T = 10$		
			$n = 150$	$n = 200$	$n = 300$	$n = 150$	$n = 200$	$n = 300$
$\rho = 0.3$	$\hat{\rho}$	bias	-0.0137	-0.0129	-0.0136	-0.0096	-0.0075	-0.0064
		std	0.0943	0.0814	0.0678	0.0520	0.0452	0.0355
	$\hat{\sigma}_e^2$	bias	0.0023	0.0029	-0.0004	-0.0006	-0.0010	-0.0009
		std	0.0657	0.0600	0.0475	0.0397	0.0363	0.0274
$\rho = 0.6$	$\hat{\rho}$	bias	-0.0035	-0.0007	-0.0047	-0.0007	-0.0028	-0.0038
		std	0.1225	0.1054	0.0873	0.0611	0.0557	0.0424
	$\hat{\sigma}_e^2$	bias	-0.0523	-0.0424	-0.0388	-0.0302	-0.0258	-0.0193
		std	0.1178	0.1000	0.0850	0.0613	0.0534	0.0426
	$\hat{\sigma}_\mu^2$	bias	0.0173	0.0157	0.0072	0.0114	0.0079	0.0053
		std	0.0735	0.0653	0.0521	0.0414	0.0381	0.0284
	$\hat{\sigma}_\mu^2$	bias	0.0048	0.0085	-0.0020	0.0058	0.0009	-0.0013
		std	0.1342	0.1144	0.0957	0.0659	0.0589	0.0450

从表 1–4, 我们可以得到以下结论:

- (1) 参数分量的所有估计都是渐近无偏的并且标准误差估计 (ste) 几乎在所有情形下与真实的标准误差 (std) 很接近.
- (2) 参数分量的可行的加权广义估计 (GEE-FW), 对比忽略误差结构的不加权的估计 (GEE-WI), 有更小的标准差 (std) 和均方误差 (mse). 对连接函数, 加权估计 (GEE-FW) 的样本 RASE 的中位数几乎一致地小于不加权的估计 (GEE-WI). 与加权估计 (GEE-FW) 比较, 两阶段估计 (GEE-TS) 的样本 RASE 的中位数也是几乎一致地小. 数值模拟结果表明, 忽略个体内部相关的不加权估计 (GEE-WI) 会损失一定的估计有效性. 这与理论结果相符合.
- (3) 当样本量增加时, 加权估计 (GEE-FW) 表现和假定真实协方差模型的基准估计 (GEE-BW) 几乎一致.

乎很接近. 此模拟结果符合自回归系数和误差方差估计是相合的并且是有效的理论.

7 实际数据分析

通过一个关于工资的面板数据的案例分析来说明本文所研究的模型及其估计方法在实际问题中的应用. 此数据来源于美国劳工部的国家纵向调查数据 (national longitudinal surveys, NLS). 该分析基于对年龄在 14 到 24 周岁之间的青年女性在时间 1983、1985、1987 和 1989 年 (分别用 $t = 1, 2, 3, 4$ 表示) 进行的调查数据的一个子集, 共包括 716 名妇女. 详细的数据说明参见文献 [30, 第 15.4.3 小节]. 我们的分析模型选取该子集的以下 7 个观测变量: 对数工资 Y_{it} 为响应变量; 虚拟变量 $X_{it,1}$, 取值为 1 表示该妇女来自南部, 否则取值为 0; 虚拟变量 $X_{it,2}$ 取值为 1 表示该妇女为黑人, 否则取值为 0; 虚拟变量 $X_{it,3}$ 取值为 1 表示是工会的一员, 否则取值为 0; 总的工作经验 (年) $Z_{it,1}$; 工作年限 $Z_{it,2}$; 以及受教育年限 $Z_{it,3}$. 基于 Hill 等^[30] 已经分析得到总的工作经验和工作年限对工资的影响为非线性二次效应, 因此, 为了放松二次效应的假定, 考虑如下具有序列自相关误差结构的面板数据部分线性单指标模型:

$$Y_{it} = X_{it,1}\beta_1 + X_{it,2}\beta_2 + X_{it,3}\beta_3 + g(Z_{it,1} + Z_{it,2}\theta_1 + Z_{it,3}\theta_2) + \varepsilon_{it},$$

误差 $\varepsilon_{it} = \mu_i + \nu_{it}$, $\nu_{it} = \rho\nu_{i,t-1} + e_{it}$, 其中 $i = 1, \dots, 716$; $t = 1, \dots, 4$.

参数 $(\rho, \sigma_\mu^2, \sigma_e^2)^T$ 的估计为 $(0.1757, 0.1017, 0.0376)^T$. 表 5 给出了基于独立工作阵的 GEE 估计 (GEE-WI) 和所提出的可行加权 GEE 估计 (GEE-FW) 这两种不同情形下参数分量的计算结果, 其中标准误差估计来自渐近协方差矩阵. 此外, 两阶段局部线性估计 (GEE-TS) 曲线及经验的 95% 置信区间在图 1 中.

从表 5 中可看到, 所提出的加权估计 (GEE-FW) 比工作独立的 GEE 估计 (GEE-WI) 具有更小的估计误差和更短的置信区间. 此外, β_1 和 β_2 的估计为负, 其他系数估计为正. 这表明来自南部以及是黑人对工资有负效应. 从图 1 中可看到连接函数与单指标间具有递增趋势关系. 这反映了累计工作经验、工作年限和教育的线性组合对增加工资具有更大的潜力. 单指标估计系数全部为正, 表明工作经验、工作年限和教育对工资均有正效应. 这些结论与传统的劳动经济学理论相符合, 可以为相关的经济理论提供统计支持.

表 5 基于 NLS 数据的参数估计结果

参数	估计方法	估计值	标准误差	95% 的置信区间
β_1	GEE-WI	-0.2135	0.0277	[-0.2677, -0.1593]
β_2	GEE-WI	-0.1808	0.0290	[-0.2376, -0.1240]
β_3	GEE-WI	0.1438	0.0257	[0.0934, 0.1942]
θ_1	GEE-WI	0.2648	0.0881	[0.0922, 0.4374]
θ_2	GEE-WI	2.9102	0.3111	[2.3005, 3.5198]
β_1	GEE-FW	-0.1573	0.0236	[-0.2034, -0.1111]
β_2	GEE-FW	-0.1746	0.0280	[-0.2294, -0.1198]
β_3	GEE-FW	0.0847	0.0145	[0.0563, 0.1131]
θ_1	GEE-FW	0.2743	0.0460	[0.1842, 0.3645]
θ_2	GEE-FW	2.9277	0.1622	[2.6098, 3.2457]

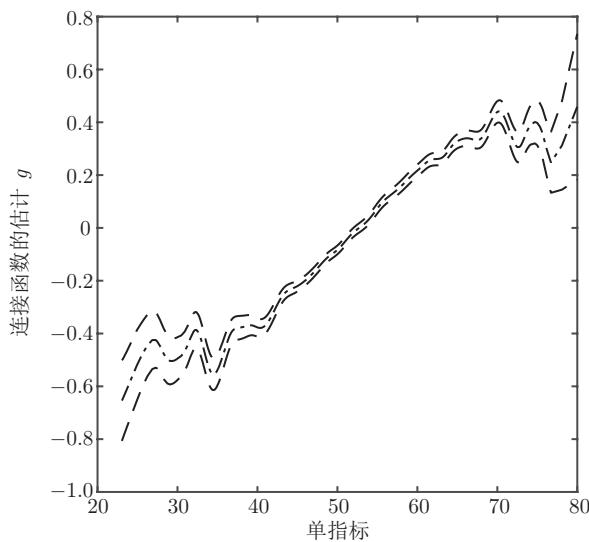


图 1 NLS 数据分析结果: 连接函数 $g(\cdot)$ 的 GEE-TS 估计. 其中点虚线和上下两条虚线分别表示曲线估计与相应的逐点置信区间 (95% 置信水平)

8 总结

本文主要讨论了具有自相关误差结构的面板数据部分线性单指标模型的统计推断问题, 通过结合局部多项式方法和纠偏的广义估计方程, 对回归模型的线性部分参数和单指标部分参数分量提出了可行加权广义估计 (GEE-FW), 证明了 GEE-FW 的相合性和渐近正态性, 并且在渐近方差意义下说明了该估计比工作独立性条件下的未加权 GEE 估计更有效. 此外, 本文在 GEE 估计和 GEE-FW 的基础上, 还提出了考虑自相关误差结构的未知连接函数的两阶段局部线性估计 (GEE-TS), 建立了该估计的渐近性质并说明其渐近有效性. 数值模拟研究和实际数据分析都表明了本文所提出的方法是有效的, 并且在理论和应用方面均具有良好表现. 本文所提出的参数以及非参数估计方法同样适用于个体观察次数不相同的情形, 非参数部分的结果可参见文献 [31].

众所周知, 在正确设定的情形下, 对比非参数模型和半参数模型, 参数模型更加简约. 因此需要检验非线性函数 $g(\cdot)$ 是否具有参数结构. Fan 等^[32] 在经典的非参数回归中提出了广义似然比统计量检验某个未知函数是否具有参数结构. 将这种检验统计量推广到我们的模型中需要进一步研究.

在模型中, 我们并没有考虑滞后响应变量 $Y_{i,t-d}$ 项. 如果此滞后响应变量加到回归模型中, 模型具有内生性, 需要考虑其他新的估计方法 (如工具变量方法). 此外, 本文仅考虑了大 n 小 T 情形. 在一些情形下会出现观察的时间点要比样本个数多或是与样本量一起增加, 这些都需要进一步研究.

致谢 衷心感谢匿名审稿人及编委对本文提出的许多宝贵意见.

参考文献

- 1 Ahn S C, Schmidt P. Estimation of linear panel data models using GMM. In: Generalized Method of Moments Estimation. Cambridge: Cambridge University Press, 1999, 211–247
- 2 Hsiao C. Analysis of Panel Data, 2nd ed. Cambridge: Cambridge University Press, 2003
- 3 Baltagi B. Econometric Analysis of Panel Data. Chichester: John Wiley & Sons, 2008
- 4 Verbeke G, Molenberghs G. Linear Mixed Models for Longitudinal Data. New York: Springer, 2009

- 5 Ruckstuhl A F, Welsh A H, Carroll R J. Nonparametric function estimation of the relationship between two repeatedly measured variables. *Statist Sinica*, 2000, 10: 51–71
- 6 Horowitz J L, Markatou M. Semiparametric estimation of regression models for panel data. *Rev Econom Stud*, 1996, 63: 145–168
- 7 Li Q, Ullha A. Estimating partially linear panel data models with one-way error components. *Econometric Rev*, 1998, 17: 145–166
- 8 You J, Zhou X, Zhou Y. Statistical inference for panel data semiparametric partially linear regression models with heteroscedastic errors. *J Multivariate Anal*, 2010, 101: 1079–1101
- 9 Zeger S L, Diggle P J. Semiparametric models for longitudinal data with application to CD4 cell numbers in HIV seroconverters. *Biometrics*, 1994, 50: 689–699
- 10 Ichimura H. Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. *J Econometrics*, 1993, 58: 71–120
- 11 Yu Y, Ruppert D. Penalized spline estimation for partially linear single-index models. *J Amer Statist Assoc*, 2002, 97: 1042–1054
- 12 Xia Y, Härdle W. Semi-parametric estimation of partially linear single-index models. *J Multivariate Anal*, 2006, 97: 1162–1184
- 13 Xue L G, Zhu L. Empirical likelihood for single-index models. *J Multivariate Anal*, 2006, 97: 1295–1312
- 14 Liang H, Liu X, Li R, et al. Estimation and testing for partially linear single-index models. *Ann Statist*, 2010, 38: 3811–3836
- 15 Cui X, Härdle W K, Zhu L. The EFM approach for single-index models. *Ann Statist*, 2011, 39: 1658–1688
- 16 Wang J, Xue L, Zhu L, et al. Estimation for a partial-linear single-index model. *Ann Statist*, 2011, 38: 246–274
- 17 Ma Y, Zhu L. Doubly robust and efficient estimators for heteroscedastic partially linear single-index models allowing high dimensional covariates. *J R Stat Soc Ser B Stat Methodol*, 2013, 75: 305–322
- 18 Brockwell P J, Davis R A. *Introduction to Time Series and Forecasting*. New York: Taylor & Francis, 2002
- 19 Chen J, Gao J, Li D. Estimation in partially linear single-index panel data models with fixed effects. *J Bus Econom Statist*, 2013, 31: 315–330
- 20 Li G, Lai P, Lian H. Variable selection and estimation for partially linear single-index models with longitudinal data. *Stat Comput*, 2015, 25: 579–593
- 21 Xu P, Zhang J, Huang X, et al. Efficient estimation for marginal generalized partially linear single-index models with longitudinal data. *TEST*, 2016, 25: 413–431
- 22 Fan J. Local linear regression smoothers and their minimax efficiencies. *Ann Statist*, 1993, 21: 196–216
- 23 Fan J, Gijbels I. *Local Polynomial Modelling and Its Applications: Monographs on Statistics and Applied Probability*. London: Chapman & Hall/CRC, 1996
- 24 Chen J, Li D, Liang H, et al. Semiparametric GEE analysis in partially linear single-index models for longitudinal data. *Ann Statist*, 2015, 43: 1682–1715
- 25 Pang Z, Xue L. Estimation for the single-index models with random effects. *Comput Statist Data Anal*, 2012, 56: 1837–1853
- 26 Mundra K. Immigration and international trade: A semiparametric empirical investigation. *J Int Trade Economic Dev*, 2005, 14: 65–91
- 27 You J, Zhou H. Two-stage efficient estimation of longitudinal nonparametric additive models. *Statist Probab Lett*, 2007, 77: 1666–1675
- 28 Fan J, Zhang J T. Two-step estimation of functional linear models with applications to longitudinal data. *J R Stat Soc Ser B Stat Methodol*, 2000, 62: 303–322
- 29 Wang N. Marginal nonparametric kernel regression accounting for within-subject correlation. *Biometrika*, 2003, 90: 43–52
- 30 Hill R C, Griffiths W E, Lim G C. *Principles of Econometrics*. Hoboken: John Wiley & Sons, 2008
- 31 Yao W, Li R. New local estimation procedure for nonparametric regression function of longitudinal data. *J R Stat Soc Ser B Stat Methodol*, 2013, 75: 123–138
- 32 Fan J, Zhang C, Zhang J. Generalized likelihood ratio statistics and Wilks phenomenon. *Ann Statist*, 2001, 29: 153–193
- 33 Mack Y P, Silverman B W. Weak and strong uniform consistency of kernel regression estimates. *Z Wahrscheinlichkeitstheorie Verw Gebiete*, 1982, 61: 405–415
- 34 Carroll R J, Härdle W, Liang H. Errors-in-variables model. *Ann Statist*, 1999, 27: 1519–1535
- 35 Horowitz J L, Mammen E. Nonparametric estimation of an additive model with a link function. *Ann Statist*, 2004, 32: 2412–2443

附录 A 主要理论结果的证明

引理 A.1 令假设 2.1、2.4 和 2.5(i) 成立并且 $nh^8 \rightarrow 0$, 则当 $n \rightarrow \infty$ 时,

$$\sup_{u \in \mathcal{U}} \left| \frac{1}{nTh} \sum_{i=1}^n \sum_{t=1}^T K\left(\frac{U_{it} - u}{h}\right) \left(\frac{U_{it} - u}{h}\right)^k - \frac{1}{T} \sum_{t=1}^T p_t(u) \varrho_k \right| = O_p \left\{ h^2 + \left(\frac{\log n}{nh}\right)^{1/2} \right\}$$

和

$$\sup_{u \in \mathcal{U}} \frac{1}{nTh} \sum_{i=1}^n \sum_{t=1}^T K\left(\frac{U_{it} - u}{h}\right) \left(\frac{U_{it} - u}{h}\right)^k \varepsilon_{it} = O_p \left\{ \left(\frac{\log n}{nh}\right)^{1/2} \right\},$$

其中 $k = 0, 1, 2, 4$.

引理 A.1 可直接从文献 [33] 中的结果得到.

引理 A.2 令假设 2.1、2.4 和 2.5 成立并且 $\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \leq c_1 n^{-1/2}$, c_1 为固定常数, 则当 $n \rightarrow \infty$ 时,

$$\max_{1 \leq i \leq n, 1 \leq t \leq T} |\hat{g}_n(\mathbf{Z}_{it}\boldsymbol{\theta}) - g(\mathbf{Z}_{it}\boldsymbol{\theta})| = O_p \left\{ h^2 + \left(\frac{\log n}{nh}\right)^{1/2} \right\}$$

和

$$\max_{1 \leq i \leq n, 1 \leq t \leq T} |\hat{g}'_n(\mathbf{Z}_{it}\boldsymbol{\theta}) - g'(\mathbf{Z}_{it}\boldsymbol{\theta})| = O_p \left\{ h + \left(\frac{\log n}{nh^3}\right)^{1/2} \right\}.$$

引理 A.2 可从文献 [25, 定理 1] 的结果得到.

引理 A.3 令假设 2.1、2.4 和 2.5 成立, 则当 $n \rightarrow \infty$ 时,

$$\frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} \left\{ \sum_{i_1=1}^n \sum_{t_1=1}^T W_{nit}(\mathbf{Z}_{i_1 t_1}^\top \boldsymbol{\theta}) \right\} = o_p(n^{-\frac{1}{2}}).$$

引理 A.4 令假设 2.1、2.4 和 2.5 成立, 则当 $n \rightarrow \infty$ 时,

$$\frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} \{ \hat{g}_n(\mathbf{Z}_{it}\boldsymbol{\theta}) - g(\mathbf{Z}_{it}\boldsymbol{\theta}) \} = o_p(n^{-\frac{1}{2}}).$$

由引理 A.2, 结合文献 [34, 引理 A.5 和 A.6] 可得到引理 A.3 和 A.4 成立. 这里略去细节.

定理 2.1 的证明 由 (2.1) 得

$$\begin{aligned} 0 &= \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \{Y_{it} - \mathbf{X}_{it}^\top \hat{\boldsymbol{\beta}}_n - \hat{g}(\mathbf{Z}_{it}^\top \hat{\boldsymbol{\theta}}_n)\} \hat{g}'(\mathbf{Z}_{it}^\top \hat{\boldsymbol{\theta}}_n) \{\mathbf{Z}_{-1,it} - \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^\top \hat{\boldsymbol{\theta}}_n)\} \\ &= \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \{ \mathbf{X}_{it}^\top (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n) + (g(\mathbf{Z}_{it}^\top \boldsymbol{\theta}) - \hat{g}(\mathbf{Z}_{it}^\top \boldsymbol{\theta})) + (\hat{g}(\mathbf{Z}_{it}^\top \boldsymbol{\theta}) - g(\mathbf{Z}_{it}^\top \hat{\boldsymbol{\theta}}_n)) + \varepsilon_{it} \} \\ &\quad \times \hat{g}'(\mathbf{Z}_{it}^\top \hat{\boldsymbol{\theta}}_n) \{(\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^\top \boldsymbol{\theta})) - (\mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^\top \boldsymbol{\theta}) - \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^\top \hat{\boldsymbol{\theta}}_n)) \}. \end{aligned}$$

上式方程中的右边可化为

$$\frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \hat{g}'(\mathbf{Z}_{it}^\top \hat{\boldsymbol{\theta}}_n) \{\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^\top \boldsymbol{\theta})\} \mathbf{X}_{it}^\top (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n)$$

$$\begin{aligned}
& + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \hat{g}'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) \{ \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) - \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) \} \mathbf{X}_{it}^T (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n) \\
& + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \hat{g}'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) \{ \mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} (g(\mathbf{Z}_{it}^T \boldsymbol{\theta}) - \hat{g}(\mathbf{Z}_{it}^T \boldsymbol{\theta})) \\
& + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \hat{g}'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) \{ \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) - \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) \} (g(\mathbf{Z}_{it}^T \boldsymbol{\theta}) - \hat{g}(\mathbf{Z}_{it}^T \boldsymbol{\theta})) \\
& + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \hat{g}'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) \{ \mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} (\hat{g}(\mathbf{Z}_{it}^T \boldsymbol{\theta}) - g(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n)) \\
& + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \hat{g}'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) \{ \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) - \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) \} (\hat{g}(\mathbf{Z}_{it}^T \boldsymbol{\theta}) - g(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n)) \\
& + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} g'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) \{ (\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})) - (\mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) - \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n)) \} \\
& =: J_1 + J_2 + \cdots + J_7.
\end{aligned}$$

第一部分 J_1 可以写为

$$\begin{aligned}
J_1 & = \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \hat{g}'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) \{ \mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \{ \mathbf{X}_{it}^T - \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n) \\
& + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \hat{g}'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) \{ \mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n).
\end{aligned}$$

进一步可化简为

$$\begin{aligned}
J_1 & = \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{ \mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \{ \mathbf{X}_{it}^T - \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n) \\
& + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{ \mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n) \\
& + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \{ \hat{g}'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) - g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \{ \mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \\
& \times \{ \mathbf{X}_{it}^T - \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n) \\
& + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \{ \hat{g}'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) - g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \{ \mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n) \\
& =: J_{11} + J_{12} + J_{13} + J_{14}.
\end{aligned}$$

由大数定律可得

$$J_{11} = \sum_{t=1}^T \mathbb{E}[g'(\mathbf{Z}_{it}^T \boldsymbol{\theta})(\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}))(\mathbf{X}_{it}^T - \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}))] \sqrt{nT} (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n) + o_p(1).$$

由于

$$\sum_{t=1}^T \mathbb{E}[g'(\mathbf{Z}_{it}^T \boldsymbol{\theta})(\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})) \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})]$$

$$\begin{aligned}
&= \sum_{t=1}^T \mathbb{E}[\mathbb{E}[g'(\mathbf{Z}_{it}^T \boldsymbol{\theta})(\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})) \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) | \mathbf{Z}_{it}^T \boldsymbol{\theta}]] \\
&= 0,
\end{aligned}$$

以及假定 $\sum_{t=1}^T g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})\} \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})$ 是独立同分布的随机矩阵并且每个方差元素都是有限的, 因此, $J_{12} = O_p(1) \cdot \|\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n\| = o_p(1)$.

此外,

$$\begin{aligned}
\|J_{13}\| &\leq \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \|\hat{g}'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) - g'(\mathbf{Z}_{it}^T \boldsymbol{\theta})\| \\
&\quad \times \|\{\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})\} \{\mathbf{X}_{it}^T - \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})\}\| \|\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n\| \\
&= O_p(n^{\frac{1}{2}}) \cdot O_p(n^{\frac{1}{2}}) \cdot O_p\{h + (nh^3)^{-\frac{1}{2}} \log n\} \\
&\quad \times \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \|\{\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})\} \{\mathbf{X}_{it}^T - \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})\}\| \\
&= O_p\{h + (nh^3)^{-\frac{1}{2}} \log n\} \\
&= o_p(1).
\end{aligned}$$

类似可得到 $J_{14} = o_p(1)$. 因此,

$$J_1 = \sum_{t=1}^T \mathbb{E}[g'(\mathbf{Z}_{it}^T \boldsymbol{\theta})(\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}))(\mathbf{X}_{it}^T - \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}))] \sqrt{nT} (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n) + o_p(1).$$

由于 $g'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n)$ 的相合性以及 $\hat{\boldsymbol{\beta}}_n$ 的 \sqrt{n} 相合性, 易得 $J_2 = o_p(1)$.

对 J_3 , 有如下式子成立:

$$\begin{aligned}
J_3 &= \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})\} (g(\mathbf{Z}_{it}^T \boldsymbol{\theta}) - \hat{g}(\mathbf{Z}_{it}^T \boldsymbol{\theta})) \\
&\quad + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \{\hat{g}'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) - g'(\mathbf{Z}_{it}^T \boldsymbol{\theta})\} \{\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})\} (g(\mathbf{Z}_{it}^T \boldsymbol{\theta}) - \hat{g}(\mathbf{Z}_{it}^T \boldsymbol{\theta})) \\
&=: J_{31} + J_{32}.
\end{aligned}$$

由于

$$\mathbb{E}[g'(\mathbf{Z}_{it}^T \boldsymbol{\theta})(\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})) | \mathbf{Z}_{it}^T \boldsymbol{\theta}] = 0,$$

通过引理 A.3 可得到 $J_{31} = o_p(1)$. 另一方面, 对于 J_{32} , 可得

$$J_{32} = O_p\{h + (nh^3)^{-\frac{1}{2}} \log n\} \cdot O_p\{h^2 + (nh)^{-\frac{1}{2}} \log n\} \cdot O_p(n^{\frac{1}{2}}) = o_p(1).$$

因此, $J_3 = o_p(1)$.

对于 J_4 , 应用 $\hat{\boldsymbol{\theta}}_n$ 的相合性以及引理 A.2 可得

$$J_4 = O_p\{h + (nh^3)^{-\frac{1}{2}} \log n\} \cdot O_p\{h^2 + (nh)^{-\frac{1}{2}} \log n\} \cdot O_p(n^{\frac{1}{2}}) = o_p(1).$$

对于 J_5 , 易得

$$\begin{aligned} J_5 &= \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \{g'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) - g'(\mathbf{Z}_{it}^T \boldsymbol{\theta})\} \{\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})\} (\hat{g}(\mathbf{Z}_{it}^T \boldsymbol{\theta}) - g(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n)) \\ &\quad + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})\} (\hat{g}(\mathbf{Z}_{it}^T \boldsymbol{\theta}) - \hat{g}(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n)) \\ &=: J_{51} + J_{52}. \end{aligned}$$

类似 J_4 , 以下式子成立:

$$J_{51} = O_p\{h^2 + (nh)^{-\frac{1}{2}} \log n\} \cdot O_p\{h^2 + (nh)^{-\frac{1}{2}} \log n\} \cdot O_p(n^{\frac{1}{2}}) = o_p(1).$$

易得

$$\begin{aligned} &\{\hat{g}(\mathbf{Z}_{it}^T \boldsymbol{\theta}) - \hat{g}(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n)\} - \{g(\mathbf{Z}_{it}^T \boldsymbol{\theta}) - g(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n)\} \\ &= \{\hat{g}'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \mathbf{Z}_{it}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) + o_p(n^{-\frac{1}{2}})\} - \{g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \mathbf{Z}_{it}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) + o_p(n^{-\frac{1}{2}})\} \\ &= (\hat{g}'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) - g'(\mathbf{Z}_{it}^T \boldsymbol{\theta})) \mathbf{Z}_{it}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) + o_p(n^{-\frac{1}{2}}). \end{aligned}$$

因此,

$$\begin{aligned} J_{52} &= \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})\} \\ &\quad \times \{\{g(\mathbf{Z}_{it}^T \boldsymbol{\theta}) - g(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n)\} + (\hat{g}'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) - g'(\mathbf{Z}_{it}^T \boldsymbol{\theta})) \mathbf{Z}_{it}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) + o_p(n^{-\frac{1}{2}})\} \\ &= \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})\} \{g(\mathbf{Z}_{it}^T \boldsymbol{\theta}) - g(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n)\} + o_p(n^{-\frac{1}{2}}) \\ &= \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \{g'(\mathbf{Z}_{it}^T \boldsymbol{\theta})\}^2 \{\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})\} \\ &\quad \times \{\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})\}^T (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_n) + o_p(n^{-\frac{1}{2}}) \\ &= \sum_{t=1}^T \mathbb{E}[\{g'(\mathbf{Z}_{it}^T \boldsymbol{\theta})\}^2 \{\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})\} \{\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})\}^T] \\ &\quad \times \sqrt{nT}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_n) + o_p(n^{-\frac{1}{2}}). \end{aligned}$$

类似 J_4 可得

$$J_6 = O_p\{h^2 + (nh)^{-\frac{1}{2}} \log n\} \cdot O_p\{h^2 + (nh)^{-\frac{1}{2}} \log n\} \cdot O_p(n^{\frac{1}{2}}) = o_p(1).$$

对于 J_7 , 可得

$$\begin{aligned} J_7 &= \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{(\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})) - (\mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) - \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n))\} \\ &\quad + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} \{g'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) - g'(\mathbf{Z}_{it}^T \boldsymbol{\theta})\} \end{aligned}$$

$$\begin{aligned}
& \times \{(\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})) - (\mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) - \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n))\} \\
& = \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{ \mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \\
& \quad + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{ \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \\
& \quad + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{ \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) - \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \\
& \quad + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} \{ g'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) - g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \{ \mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \\
& \quad + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} \{ g'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) - g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \{ \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \\
& \quad + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} \{ g'(\mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) - g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \{ \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) - \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \\
& =: J_{71} + J_{72} + \cdots + J_{76}.
\end{aligned}$$

由于

$$\mathbb{E}(\varepsilon_{it} g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) | \mathbf{Z}_{it}^T \boldsymbol{\theta}) = 0 \quad \text{和} \quad \mathbb{E}\{\varepsilon_{it} g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{ \mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \mid \mathbf{Z}_{it}^T \boldsymbol{\theta} \}\} = 0,$$

类似引理 A.3 的证明可得

$$J_{72} = o_p(1) \quad \text{和} \quad J_{74} = o_p(1).$$

此外, 以下式子成立:

$$\begin{aligned}
J_{73} & = \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{ \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \\
& \quad \times \{ \{ \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) - \hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} - \{ \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \hat{\boldsymbol{\theta}}_n) - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \} \\
& = \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{ \mathbb{E}'(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \mathbf{Z}_{it}^T (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) + o_p(n^{-\frac{1}{2}}) \} \\
& \quad + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{ \{ \hat{\mathbb{E}}'(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) - \mathbb{E}'(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \mathbf{Z}_{it}^T (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) + o_p(n^{-\frac{1}{2}}) \} \\
& = o_p(n^{-\frac{1}{2}}).
\end{aligned}$$

应用 $\hat{\boldsymbol{\theta}}_n$ 和 $\hat{\mathbb{E}}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})$ 的相合性可得

$$\begin{aligned}
J_{75} & = O_p\{h + (nh^3)^{-\frac{1}{2}} \log n\} \cdot O_p\{h^2 + (nh)^{-\frac{1}{2}} \log n\} \cdot O_p(n^{\frac{1}{2}}) = o_p(1), \\
J_{76} & = O_p\{h + (nh^3)^{-\frac{1}{2}} \log n\} \cdot O_p\{h^2 + (nh)^{-\frac{1}{2}} \log n\} \cdot O_p(n^{\frac{1}{2}}) = o_p(1).
\end{aligned}$$

因此,

$$J_7 = \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{ \mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} + o_p(1).$$

则由以上式子可得

$$\begin{aligned}
& \sum_{t=1}^T E[(g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}))^2 (\mathbf{Z}_{-1,it} - E(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})) (\mathbf{Z}_{-1,it} - E(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}))^T] \\
& \times \sqrt{nT}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_n) + \sum_{t=1}^T E[g'(\mathbf{Z}_{it}^T \boldsymbol{\theta})(\mathbf{Z}_{-1,it} - E(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}))(\mathbf{X}_{it}^T - E(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}))]^T \sqrt{nT}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n) \\
& = \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) \{ \mathbf{Z}_{-1,it} - E(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} + o_p(1).
\end{aligned}$$

类似得到

$$\begin{aligned}
& \sum_{t=1}^T E[\{ \mathbf{Z}_{-1,it} - E(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} \{ \mathbf{Z}_{-1,it} - E(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \}^T] \sqrt{nT}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n) \\
& + \sum_{t=1}^T E[(\mathbf{X}_{it}^T - E(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta})) g'(\mathbf{Z}_{it}^T \boldsymbol{\theta}) (\mathbf{Z}_{-1,it} - E(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}))]^T \sqrt{nT}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_n) \\
& = \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} \{ \mathbf{X}_{it} - E(\mathbf{X}_{it} | \mathbf{Z}_{it}^T \boldsymbol{\theta}) \} + o_p(1).
\end{aligned}$$

结合 Slutsky 定理及中心极限定理可得定理 2.1 成立. \square

定理 2.2 的证明 由 $\hat{\boldsymbol{\theta}}_n$ 的 \sqrt{n} 相合性以及标准的非参数回归证明过程可得定理 2.2 成立. \square

定理 3.1 的证明 (i) 设

$$\begin{aligned}
\mathbf{Q}_{0,n} &= \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} (\varepsilon_{i,t+(s-1)}, \dots, \varepsilon_{it})^T (\varepsilon_{i,t+(s-1)}, \dots, \varepsilon_{it}), \\
\Delta_{it} &= \mathbf{X}_{it}^T (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n) + g(U_{it,n}) - \hat{g}_n(\hat{U}_{it,n}).
\end{aligned}$$

由于

$$\hat{\varepsilon}_{it} = Y_{it} - \mathbf{X}_{it}^T \hat{\boldsymbol{\beta}}_n - \hat{g}_n(\hat{U}_{it,n}) = \varepsilon_{it} + \Delta_{it},$$

因此下式成立:

$$\begin{aligned}
\hat{\mathbf{Q}}_{0,n} &= \mathbf{Q}_{0,n} + \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} (\Delta_{i,t+(s-1)}, \dots, \Delta_{it})^T (\Delta_{i,t+(s-1)}, \dots, \Delta_{it}) \\
&+ \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} (\varepsilon_{i,t+(s-1)}, \dots, \varepsilon_{it})^T (\Delta_{i,t+(s-1)}, \dots, \Delta_{it}) \\
&+ \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} (\Delta_{i,t+(s-1)}, \dots, \Delta_{it})^T (\varepsilon_{i,t+(s-1)}, \dots, \varepsilon_{it}) \\
&=: \mathbf{Q}_{0,n} + J_{1,n} + J_{2,n} + J_{3,n}.
\end{aligned}$$

根据定理 2.1、2.2 和假设 2.3, 可得 $\Delta_{it} = O_p(n^{-1/2}) + O_p(h^2 + (nh)^{-1/2})$, 则

$$J_{1,n} = O_p(n^{-1}) + O_p\left(\frac{1}{nh}\right) = o_p((nh)^{-1/2}).$$

对于 $J_{2,n}$, 可得

$$\begin{aligned} n(T - (s + 1))J_{2,n} &= \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} (\varepsilon_{i,t+(s-1)}, \dots, \varepsilon_{it})^T (\Delta_{i,t+(s-1)}, \dots, \Delta_{it}) \\ &=: J_{2,n}^{(1)} + J_{2,n}^{(2)} + J_{2,n}^{(3)} + J_{3,n}^{(4)}. \end{aligned}$$

根据 Δ_{it} 的定义得到

$$\begin{aligned} J_{2,n}^{(1)} &= \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} (\varepsilon_{i,t+(s-1)}, \dots, \varepsilon_{it})^T (\mathbf{X}_{i,t+(s-1),n}^T(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n), \dots, \mathbf{X}_{it}^T(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n)) \\ &\quad + \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} (\varepsilon_{i,t+(s-1)}, \dots, \varepsilon_{it})^T (g(U_{i,t+(s-1),n}) - \hat{g}_n(\hat{U}_{i,t+(s-1),n})) \\ &=: J_{2,n}^{(1)*} + J_{2,n}^{(1)**}. \end{aligned}$$

类似定理 2.1 的证明, 对于 $0 \leq m \leq s - 1$, 有

$$\frac{1}{n(T - (s + 1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} \mathbf{X}_{i,t+m}(\varepsilon_{i,t+(s-1)}, \dots, \varepsilon_{it}) = O_p(n^{-\frac{1}{2}}).$$

因此, 由定理 2.1 得 $J_{2,n}^{(1)*} = O_p(1) = o(n^{1/2})$. 由引理 A.4, 以及 $\hat{\boldsymbol{\beta}}$ 和 $\hat{\boldsymbol{\theta}}$ 的相合性可得 $J_{2,n}^{(1)**} = o_p(n^{1/2})$.

于是, $J_{2,n}^{(1)} = o_p(n^{1/2})$. 类似易得 $J_{2,n}^{(2)} = J_{2,n}^{(3)} = o_p(n^{1/2})$ 和 $J_{2,n}^{(4)} = o_p(n^{1/2})$. 这表明

$$\hat{\mathbf{Q}}_{0,n} = \mathbf{Q}_{0,n} + o_p((nh)^{-\frac{1}{2}}).$$

类似可得

$$\hat{\mathbf{Q}}_{1,n} = \mathbf{Q}_{1,n} + o_p((nh)^{-\frac{1}{2}}).$$

对于 $\hat{\boldsymbol{\rho}}_n$, 可得

$$\begin{aligned} \sqrt{nT}(\hat{\boldsymbol{\rho}}_n - \boldsymbol{\rho}) &= \sqrt{nT}\{(\hat{\mathbf{Q}}_{0,n} - \hat{\mathbf{Q}}_{1,n})^{-1}(\hat{\mathbf{Q}}_{2,n} - \hat{\mathbf{Q}}_{3,n}) - \boldsymbol{\rho}\} \\ &= \sqrt{nT}[(\hat{\mathbf{Q}}_{0,n} - \hat{\mathbf{Q}}_{1,n})^{-1}((\hat{\mathbf{Q}}_{2,n} - \hat{\mathbf{Q}}_{3,n}) - (\hat{\mathbf{Q}}_{0,n} - \hat{\mathbf{Q}}_{1,n})\boldsymbol{\rho})] \\ &= \sqrt{nT}(\mathbf{Q}_{0,n} - \mathbf{Q}_{1,n})^{-1} \cdot \{(\mathbf{Q}_{2,n} - \mathbf{Q}_{3,n}) - (\mathbf{Q}_{0,n} - \mathbf{Q}_{1,n})\boldsymbol{\rho}\} + o_p(1). \end{aligned}$$

根据 $\mathbf{Q}_{0,n}$ 和 $\mathbf{Q}_{1,n}$ 的定义, 当 $n \rightarrow \infty$ 时,

$$\begin{aligned} \mathbf{Q}_{0,n} - \mathbf{Q}_{1,n} &= \frac{1}{n(T - (s + 1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} (\varepsilon_{i,t+(s-1)}, \dots, \varepsilon_{it})^T (\varepsilon_{i,t+(s-1)}, \dots, \varepsilon_{it}) \\ &\quad - \frac{1}{n(T - (s + 1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} (\varepsilon_{i,t+(s-1)}, \dots, \varepsilon_{it})^T (\varepsilon_{i,t+s}, \dots, \varepsilon_{i,t+1}) \\ &\xrightarrow{p} \mathbb{E}\{(\varepsilon_{i,t+(s-1)}, \dots, \varepsilon_{it})^T (\varepsilon_{i,t+(s-1)}, \dots, \varepsilon_{it})\} \\ &\quad - \mathbb{E}\{(\varepsilon_{i,t+(s-1)}, \dots, \varepsilon_{it+1})^T (\varepsilon_{i,t+s}, \dots, \varepsilon_{i,t+1})\} \end{aligned}$$

$$\begin{aligned}
&= E\{(\nu_{i,t+(s-1)}, \dots, \nu_{it})^T (\nu_{i,t+(s-1)}, \dots, \nu_{it})\} \\
&\quad - E\{(\nu_{i,t+(s-1)}, \dots, \nu_{it})^T (\nu_{i,t+s}, \dots, \nu_{i,t+1})\} \\
&= \begin{pmatrix} \gamma_\nu(0) & \gamma_\nu(1) & \cdots & \gamma_{\nu,n}(s-1) \\ \gamma_\nu(1) & \gamma_\nu(0) & \cdots & \gamma_\nu(s-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_\nu(s-1) & \gamma_\nu(s-2) & \cdots & \gamma_\nu(0) \end{pmatrix} - \begin{pmatrix} \gamma_\nu(1) & \gamma_\nu(2) & \cdots & \gamma_\nu(s) \\ \gamma_\nu(0) & \gamma_\nu(1) & \cdots & \gamma_\nu(s-1) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_\nu(s) & \gamma_\nu(s-1) & \cdots & \gamma_\nu(1) \end{pmatrix}.
\end{aligned}$$

根据 (1.2) 可得

$$\begin{aligned}
\varepsilon_{i,t+s} &= \mu_i + \nu_{i,t+s} = \mu_i + \rho_1 \nu_{i,t+(s-1)} + \cdots + \rho_s \nu_{it} + e_{i,t+s} \\
&= \rho_1 \varepsilon_{i,t+(s-1)} + \cdots + \rho_s \varepsilon_{it} + (1 - \rho_1 - \cdots - \rho_s) \mu_i + e_{i,t+s}
\end{aligned}$$

和

$$\begin{aligned}
\varepsilon_{i,t+(s+1)} &= \mu_i + \nu_{i,t+(s+1)} = \mu_i + \rho_1 \nu_{i,t+s} + \cdots + \rho_s \nu_{i,t+1} + e_{i,t+(s+1)} \\
&= \rho_1 \varepsilon_{i,t+s} + \cdots + \rho_s \varepsilon_{i,t+1} + (1 - \rho_1 - \cdots - \rho_s) \mu_i + e_{i,t+(s+1)}.
\end{aligned}$$

因此,

$$\begin{aligned}
&\sqrt{n(T-(s+1))} \{ \mathbf{Q}_{2,n} - \mathbf{Q}_{3,n} - (\mathbf{Q}_{0,n} - \mathbf{Q}_{1,n}) \boldsymbol{\rho} \} \\
&= \frac{1}{\sqrt{n(T-(s+1))}} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} \{ (\varepsilon_{i,t+(s-1)}, \dots, \varepsilon_{it})^T e_{i,t+s} - (\varepsilon_{i,t+(s-1)}, \dots, \varepsilon_{it})^T e_{i,t+(s+1)} \} \\
&= \frac{1}{\sqrt{n(T-(s+1))}} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} \{ \{ (\nu_{i,t+(s-1)} + \mu_i, \dots, \nu_{it} + \mu_i)^T e_{i,t+s} \\
&\quad - (\nu_{i,t+(s-1)} + \mu_i, \dots, \nu_{it} + \mu_i)^T e_{i,t+(s+1)} \} \} \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\frac{1}{\sqrt{(T-(s+1))}} \sum_{t=1}^{T-(s+1)} [(\nu_{i,t+(s-1)} + \mu_i, \dots, \nu_{it} + \mu_i)^T e_{i,t+s} \right. \\
&\quad \left. - (\nu_{i,t+(s-1)} + \mu_i, \dots, \nu_{it} + \mu_i)^T e_{i,t+(s+1)}] \right] \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^n \chi_i,
\end{aligned}$$

其中 $\{\chi_i\}_{i=1}^n$ 是独立同分布的随机向量序列, 其均值为 0, 协方差为

$$\begin{aligned}
\text{Cov}(\chi_i) &= \text{Cov}\left(\frac{1}{\sqrt{T-(s+1)}} \sum_{t=1}^{T-(s+1)} [(\nu_{i,t+(s-1)} + \mu_i, \dots, \nu_{it} + \mu_i)^T (e_{i,t+s} - e_{i,t+(s+1)})] \right) \\
&= 2\sigma_e^2 \left\{ \sigma_\mu^2 \mathbf{1}_s \mathbf{1}_s^T + \begin{pmatrix} \gamma_\nu(0) & \gamma_\nu(1) & \cdots & \gamma_{\nu,n}(s-1) \\ \gamma_\nu(1) & \gamma_\nu(0) & \cdots & \gamma_\nu(s-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_\nu(s-1) & \gamma_\nu(s-2) & \cdots & \gamma_\nu(0) \end{pmatrix} \right\},
\end{aligned}$$

其中 $\gamma_\nu(j) = \text{E}(\nu_{it}\nu_{i,t+j})$.

(ii) 由于 $\hat{\ell}_{it,n} = \hat{\varepsilon}_{i,t+s} - \hat{\rho}_{1,n}\hat{\varepsilon}_{i,t+(s-1)} - \cdots - \hat{\rho}_{s,n}\hat{\varepsilon}_{it}$,

$$\begin{aligned}\hat{\ell}_{it,n} &= \{\varepsilon_{i,t+s} - \hat{\rho}_{1,n}\varepsilon_{i,t+(s-1)} - \cdots - \hat{\rho}_{s,n}\varepsilon_{it}\} + \{\Delta_{i,t+s} - \hat{\rho}_{1,n}\Delta_{i,t+(s-1)} - \cdots - \hat{\rho}_{s,n}\Delta_{it}\} \\ &= \{\varepsilon_{i,t+s} - \rho_1\varepsilon_{i,t+(s-1)} - \cdots - \rho_s\varepsilon_{it}\} + l\{(\rho_1 - \hat{\rho}_{1,n})\varepsilon_{i,t+(s-1)} - \cdots - (\rho_s - \hat{\rho}_{s,n})\varepsilon_{it}\} \\ &\quad + \{\Delta_{i,t+s} - \hat{\rho}_{1,n}\Delta_{i,t+(s-1)} - \cdots - \hat{\rho}_{s,n}\Delta_{it}\} \\ &= e_{i,t+s} + \{(1 - \rho_1 - \cdots - \rho_s)\mu_i\} + \{(\rho_1 - \hat{\rho}_{1,n})\varepsilon_{i,t+(s-1)} - \cdots - (\rho_s - \hat{\rho}_{s,n})\varepsilon_{it}\} \\ &\quad + \{\Delta_{i,t+s} - \hat{\rho}_{1,n}\Delta_{i,t+(s-1)} - \cdots - \hat{\rho}_{s,n}\Delta_{it}\} \\ &= e_{i,t+s} + J_{4,n} + J_{5,n} + J_{6,n},\end{aligned}$$

因此,

$$\begin{aligned}&\frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} \hat{\ell}_{it,n}^2 \\ &= \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} e_{i,t+s}^2 + \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} J_{4,n}^2 \\ &\quad + \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} J_{5,n}^2 + \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} J_{6,n}^2 \\ &\quad + \frac{2}{n(T-s)} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} e_{i,t+s} J_{4,n} + \frac{2}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} e_{i,t+s} J_{5,n} \\ &\quad + \frac{2}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} e_{i,t+s} J_{6,n} + \frac{2}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} J_{4,n} J_{5,n} \\ &\quad + \frac{2}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} J_{4,n} J_{6,n} + \frac{2}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} J_{5,n} J_{6,n}.\end{aligned}$$

由定理 3.1(i) 可得

$$\frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} J_{5,n}^2 = O_p(n^{-1}) \quad \text{和} \quad \frac{2}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} e_{i,t+s} J_{5,n} = O_p(n^{-1}).$$

由定理 2.1 和引理 A.2 可得

$$\frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} J_{6,n}^2 = O_p(n^{-1}) = o_p(n^{-\frac{1}{2}}).$$

根据 Cauchy-Schwarz 不等式得

$$\frac{2}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} J_{5,n} J_{6,n} = O_p(n^{-1}) = o_p(n^{-\frac{1}{2}}).$$

此外,

$$\frac{2}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} J_{4,n} J_{5,n}$$

$$= -\frac{2}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} J_{4,n}\{(\rho_1 - \hat{\rho}_{1,n}) - \cdots - (\rho_s - \hat{\rho}_{s,n})\} \mu_i.$$

因此,

$$\begin{aligned} \hat{E}(l_{it}^2) &= \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} \hat{\ell}_{it,n}^2 \\ &= \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} e_{i,t+s}^2 + \frac{1}{n} \sum_{i=1}^n (1 - \rho_1 - \cdots - \rho_s)^2 \mu_i^2 \\ &\quad + \frac{2}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} e_{i,t+s} (1 - \rho_1 - \cdots - \rho_s) \mu_i \\ &\quad - \frac{2}{n} \sum_{i=1}^n (1 - \rho_1 - \cdots - \rho_s) \{(\rho_1 - \hat{\rho}_{1,n}) - \cdots - (\rho_s - \hat{\rho}_{s,n})\} \mu_i^2. \end{aligned}$$

类似得到

$$\begin{aligned} \hat{E}(l_{it} l_{i,t+1}) &= \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} \hat{\ell}_{it,n} \hat{\ell}_{i,t+1,n} \\ &= \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} e_{i,t+s} e_{i,t+s+1} + \frac{1}{n} \sum_{i=1}^n (1 - \rho_1 - \cdots - \rho_s)^2 \mu_{i,n}^2 \\ &\quad + \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} e_{i,t+s} (1 - \rho_1 - \cdots - \rho_s) \mu_i \\ &\quad + \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} e_{i,t+s+1} (1 - \rho_1 - \cdots - \rho_s) \mu_i \\ &\quad - \frac{2}{n} \sum_{i=1}^n (1 - \rho_1 - \cdots - \rho_s) \{(\rho_1 - \hat{\rho}_{1,n}) - \cdots - (\rho_s - \hat{\rho}_{s,n})\} \mu_i^2. \end{aligned}$$

因此,

$$\begin{aligned} \hat{\sigma}_{e,n}^2 &= \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} \hat{\ell}_{it,n}^2 - \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} \hat{\ell}_{it,n} \hat{\ell}_{i,t+1,n} \\ &= \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} (e_{i,t+s}^2 - e_{i,t+s} e_{i,t+s+1}) \\ &\quad + \frac{1}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} (e_{i,t+s} - e_{i,t+s+1})(1 - \rho_{1,n} - \cdots - \rho_{s,n}) \mu_i. \end{aligned}$$

则

$$\begin{aligned} \sqrt{nT}(\hat{\sigma}_{e,n}^2 - \sigma_e^2) &= \frac{\sqrt{nT}}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} (e_{i,t+s}^2 - \sigma_e^2 - e_{i,t+s} e_{i,t+s+1}) \\ &\quad + \frac{\sqrt{nT}}{n(T-(s+1))} \sum_{i=1}^n \sum_{t=1}^{T-(s+1)} (e_{i,t+s} - e_{i,t+s+1})(1 - \rho_1 - \cdots - \rho_s) \mu_i + o_p(1) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\frac{\sqrt{T}}{T - (s+1)} \sum_{t=1}^{T-(s+1)} (e_{i,t+s}^2 - \sigma_e^2 - e_{i,t+s} e_{i,t+s+1} \right. \\
&\quad \left. - (e_{i,t+s} - e_{i,t+s+1})(1 - \rho_1 - \cdots - \rho_s) \mu_i) \right] + o_p(1) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^n \varsigma_i,
\end{aligned}$$

其中 $\varsigma_{i,n}$ 是独立同分布的随机变量向量序列, 其均值为 0, 方差为

$$\text{Var}(\varsigma_i) = \frac{T}{(T - (s+1))^2} \{ \text{Var}(e_{it}^2) + 2(1 - \rho_1 - \cdots - \rho_s)^2 \sigma_\mu^2 \sigma_e^2 + \sigma_e^4 \}.$$

(iii) 对于 $\hat{\sigma}_{en}^2$, 类似以上过程, 可得定理 3.1(iii) 成立. 这里省略其详细证明过程. \square

定理 4.1 的证明 类似定理 2.1 的证明可得

$$\begin{aligned}
&\sum_{t=1}^T \mathbb{E}[(g'(\mathbf{Z}_{it}^\top \boldsymbol{\theta}))^2 \hat{\Sigma}_{0n}^{-1} (\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^\top \boldsymbol{\theta})) (\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^\top \boldsymbol{\theta}))^\top] \cdot \sqrt{nT}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_n) \\
&+ \sum_{t=1}^T \mathbb{E}[g'(\mathbf{Z}_{it}^\top \boldsymbol{\theta}) \hat{\Sigma}_{0n}^{-1} (\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^\top \boldsymbol{\theta})) (\mathbf{X}_{it}^\top - \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^\top \boldsymbol{\theta}))] \sqrt{nT}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n) \\
&= \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} g'(\mathbf{Z}_{it}^\top \boldsymbol{\theta}) \hat{\Sigma}_{0n}^{-1} \{ \mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^\top \boldsymbol{\theta}) \} + o_p(1).
\end{aligned}$$

则得到

$$\begin{aligned}
&\sum_{t=1}^T \mathbb{E}[(\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^\top \boldsymbol{\theta})) \hat{\Sigma}_{0n}^{-1} (\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^\top \boldsymbol{\theta}))^\top] \sqrt{nT}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_n) \\
&+ \sum_{t=1}^T \mathbb{E}[(\mathbf{X}_{it}^\top - \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^\top \boldsymbol{\theta})) g'(\mathbf{Z}_{it}^\top \boldsymbol{\theta}) \hat{\Sigma}_{0n}^{-1} (\mathbf{Z}_{-1,it} - \mathbb{E}(\mathbf{Z}_{-1,it} | \mathbf{Z}_{it}^\top \boldsymbol{\theta}))] \sqrt{nT}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_n) \\
&= \frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^T \varepsilon_{it} \hat{\Sigma}_{0n}^{-1} \{ \mathbf{X}_{it} - \mathbb{E}(\mathbf{X}_{it} | \mathbf{Z}_{it}^\top \boldsymbol{\theta}) \} + o_p(1).
\end{aligned}$$

结合 Slutsky 定理及中心极限定理可得定理 4.1 成立. \square

定理 5.1 的证明 设

$$\begin{aligned}
\hat{Y}_{it}^{**} &= (\sigma^{tt})^{-1} \left(\hat{Y}_{it} - \sum_{t_1 \neq t}^T \sigma^{tt_1} \hat{a}_{it_1} \right) = (\sigma^{tt})^{-1} \left(\sum_{t_1=1}^T \sigma^{tt_1} Y_{it_1} - \sum_{t_1 \neq t}^T \sigma^{tt_1} \hat{a}_{it_1} \right), \\
(\hat{g}_n^{TS*}(u), \hat{g}_n'^{TS*}(u))^\top &= (\mathbf{D}_u^\top \mathbf{W}_u \mathbf{D}_u)^{-1} \mathbf{D}_u \mathbf{W}_u \hat{\mathbf{Y}}^{**},
\end{aligned}$$

其中 $\hat{\mathbf{Y}}^{**} = (\hat{Y}_{11}^{**}, \dots, \hat{Y}_{1T}^{**}, \dots, \hat{Y}_{nT}^{**})^\top$, 则

$$(\hat{g}_n^{TS*}(u), \hat{g}_n'^{TS*}(u))^\top - (g(u), g'(u))^\top =: J_1 + J_2 + J_3 - J_4,$$

其中

$$J_1 = (\mathbf{D}_u^\top \mathbf{W}_u \mathbf{D}_u)^{-1} \sum_{i=1}^n \sum_{t=1}^T \binom{1}{U_{it}} \frac{1}{h} K\left(\frac{U_{it} - u}{h}\right) g(U_{it}) - (g(u), g'(u))^\top,$$

$$\begin{aligned}
J_2 &= (\mathbf{D}_u^T \mathbf{W}_u \mathbf{D}_u)^{-1} \sum_{i=1}^n \sum_{t=1}^T \binom{1}{U_{it}} \frac{1}{h} K\left(\frac{U_{it}-u}{h}\right) \sum_{t_1=1}^T (\sigma^{tt})^{-1} \sigma^{tt_1} \varepsilon_{it_1}, \\
J_3 &= (\mathbf{D}_u^T \mathbf{W}_u \mathbf{D}_u)^{-1} \sum_{i=1}^n \sum_{t=1}^T \binom{1}{U_{it}} \frac{1}{h} K\left(\frac{U_{it}-u}{h}\right) (\sigma^{tt})^{-1} \sum_{t_1 \neq t}^T \sigma^{tt_1} (g(U_{it_1}) - \hat{g}_n(\hat{U}_{it_1})), \\
J_4 &= (\mathbf{D}_u^T \mathbf{W}_u \mathbf{D}_u)^{-1} \sum_{i=1}^n \sum_{t=1}^T \binom{1}{U_{it}} \frac{1}{h} K\left(\frac{U_{it}-u}{h}\right) \sum_{t_1=1}^T \sigma^{tt_1} (g(U_{it_1}) - \hat{g}_n(\hat{U}_{it_1})).
\end{aligned}$$

由引理 A.1、假设 2.1–2.6 成立以及 $\hat{\beta}$ 和 $\hat{\theta}$ 的相合性, 类似文献 [35, 引理 3 和 4] 的证明, 可以得到 $J_3 = J_4 = o_p(1)$.

注意到

$$\mathbf{D}_u^T \mathbf{W}_u \mathbf{D}_u = \begin{pmatrix} \sum_{i=1}^n \sum_{t=1}^T K_h(U_{it}-u) & \sum_{i=1}^n \sum_{t=1}^T \left(\frac{U_{it}-u}{h}\right) K_h(U_{it}-u) \\ \sum_{i=1}^n \sum_{t=1}^T \left(\frac{U_{it}-u}{h}\right) K_h(U_{it}-u) & \sum_{i=1}^n \sum_{t=1}^T \left(\frac{U_{it}-u}{h}\right)^2 K_h(U_{it}-u) \end{pmatrix}.$$

上式矩阵每一个元素都是核光滑估计的形式, 由引理 A.1 可得

$$\frac{1}{n} \mathbf{D}_u^T \mathbf{W}_u \mathbf{D}_u = \sum_{t=1}^T p_t(u) \otimes \mathbf{H} \begin{pmatrix} 1 & \varsigma_1 \\ \varsigma_1 & \varsigma_2 \end{pmatrix} \mathbf{H} \cdot O_p \left(1 + \left\{ \frac{\log n}{nh} \right\}^{1/2} \right).$$

因此, 通过非参数回归的结果可得

$$\sqrt{nTh} \left[\mathbf{H}^{-1} \left\{ J_1 - \begin{pmatrix} g(u) \\ g'(u) \end{pmatrix} \right\} - \frac{h^2}{2} \frac{1}{\varrho_2 - \varrho_1^2} \begin{pmatrix} (\varrho_2^2 - \varrho_1 \varrho_3) g''(u) \\ (\varrho_3 - \varrho_1 \varrho_2) g''(u) \end{pmatrix} + o(h^2) \right] = o_p(1).$$

接下来证明

$$\sqrt{nTh} \mathbf{H}^{-1} J_2 \xrightarrow{D} n(0, \Phi^{TS}), \quad \text{当 } n \rightarrow \infty.$$

设

$$Q = \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \left\{ d_1 + d_2 \left(\frac{U_{it}-u}{h} \right) \right\} K_h(U_{it}-u) \sum_{t_1=1}^T (\sigma^{tt})^{-1} \sigma^{tt_1} \varepsilon_{it_1},$$

其中 d_1 和 d_2 是非零的常数. 易得 $E(Q) = 0$ 以及

$$\begin{aligned}
\text{Var}(\sqrt{nTh} Q) &= \frac{h}{nT} \sum_{i=1}^n \sum_{t_1=1}^T \sum_{t_2=1}^T E \left\{ d_1 + d_2 \left(\frac{U_{it_1}-u}{h} \right) \right\} \left\{ d_1 + d_2 \left(\frac{U_{it_2}-u}{h} \right) \right\} \\
&\quad \times K_h(U_{it_1}-u) K_h(U_{it_2}-u) \left\{ \sum_{t_3=1}^T (\sigma^{t_1 t_1})^{-1} \sigma^{t_1 t_3} \varepsilon_{it_3} \right\} \left\{ \sum_{t_3=1}^T (\sigma^{t_2 t_2})^{-1} \sigma^{t_2 t_3} \varepsilon_{it_3} \right\} \\
&= \frac{h}{nT} \sum_{i=1}^n \sum_{t_1=1}^T \sum_{t_2=1}^T E \left[\left(d_1 + d_2 \left(\frac{U_{it_1}-u}{h} \right) \right) \left(d_1 + d_2 \left(\frac{U_{it_2}-u}{h} \right) \right) \right. \\
&\quad \times K_h(U_{it_1}-u) K_h(U_{it_2}-u) \left. \right] E \left\{ \left(\sum_{t_3=1}^T (\sigma^{t_1 t_1})^{-1} \sigma^{t_1 t_3} \varepsilon_{it_3} \right) \left(\sum_{t_3=1}^T (\sigma^{t_2 t_2})^{-1} \sigma^{t_2 t_3} \varepsilon_{it_3} \right) \right\}.
\end{aligned}$$

注意到对于 $t_1 \neq t_2$, 有

$$\begin{aligned} \mathbb{E}\{K_h(U_{it_1} - u)K_h(U_{it_2} - u)\} &= O(1), \\ \mathbb{E}\left\{\left(\frac{U_{it_2} - u}{h}\right)K_h(U_{it_1} - u)K_h(U_{it_2} - u)\right\} &= O(1), \\ \mathbb{E}\left\{\left(\frac{U_{it_1} - u}{h}\right)K_h(U_{it_1} - u)K_h(U_{it_2} - u)\right\} &= O(1), \\ \mathbb{E}\left(\frac{U_{it_1} - u}{h}\right)\left(\frac{U_{it_2} - u}{h}\right)K_h(U_{it_1} - u)K_h(U_{it_2} - u) &= O(1). \end{aligned}$$

此外,

$$\mathbb{E}\left\{\sum_{t_2=1}^T (\sigma^{t_1 t_1})^{-1} \sigma^{t_1 t_2} \varepsilon_{it_2}\right\}^2 = (\sigma^{t_1 t_1})^{-2} \mathbf{e}_{t_1 T} \boldsymbol{\Sigma}_0^{-1} \mathbf{e}_{t_1 T} = (\sigma^{t_1 t_1})^{-1},$$

其中 $\mathbf{e}_{t_1 T}$ 表示第 t_1 个位置元素为 1、其余元素为 0 的 T 维向量. 因此,

$$\begin{aligned} \text{Var}(\sqrt{nTh}Q) &= o(1) + \frac{h}{nT} \sum_{i=1}^n \sum_{t_1=1}^T \mathbb{E}\left[\left(d_1 + d_2\left(\frac{U_{it_1} - u}{h}\right)\right)^2 K_{h_x}^2(U_{it_1} - u)\right] \mathbb{E}\left(\sum_{t_2=1}^T (\sigma^{t_1 t_1})^{-1} \sigma^{t_1 t_2} \varepsilon_{it_2}\right)^2 \\ &= o(1) + \frac{h}{nT} \sum_{i=1}^n \sum_{t=1}^T (\sigma^{tt})^{-1} \mathbb{E}\left[\left\{d_1 + d_2\left(\frac{U_{it_1} - u}{h}\right)\right\}^2 K_h^2(U_{it} - u)\right] \\ &= o(1) + \frac{hd_1 d_2}{nT} \sum_{i=1}^n \sum_{t=1}^T (\sigma^{tt})^{-1} \mathbb{E}\{K_h^2(U_{it} - u)\} \\ &\quad + \frac{hd_1 d_2}{nT} \sum_{i=1}^n \sum_{t=1}^T (\sigma^{tt})^{-1} \mathbb{E}\left\{\left(\frac{U_{it} - u}{h}\right)^2 K_h^2(U_{it} - u)\right\} \\ &\quad + \frac{2hd_1 d_2}{nT} \sum_{i=1}^n \sum_{t=1}^T (\sigma^{tt})^{-1} \mathbb{E}\left\{\left(\frac{U_{it} - u}{h}\right) K_h^2(U_{it} - u)\right\} \\ &=: o(1) + J_5 + J_6 + J_7. \end{aligned}$$

通过计算可得

$$J_5 \rightarrow_p d_1 d_2 \sum_{t=1}^T (\sigma^{tt})^{-1} p_t(u) \varrho_0, \quad J_6 \rightarrow_p d_1 d_2 \sum_{t=1}^T (\sigma^{tt})^{-1} p_t(u) \varrho_2$$

和

$$J_7 \rightarrow_p 2d_1 d_2 \sum_{t=1}^T p_t(u) \varrho_1 \quad (n \rightarrow \infty).$$

因此,

$$\text{Var}(\sqrt{nTh}Q) = d_1 d_2 \sum_{t=1}^T (\sigma^{tt})^{-1} p_t(u) (\varrho_0 + 2\varrho_1 + \varrho_2) + o(1).$$

设 $b_i = \sqrt{h} \sum_{t=1}^T \{d_1 + d_2(\frac{U_{it}-u}{h})\} K_h(U_{it} - u) \{\sum_{t_1=1}^T (\sigma^{tt})^{-1} \sigma^{tt_1} \varepsilon_{it_1}\}$ 和 $B_n^2 = \sum_{i=1}^n \mathbb{E}(b_i^2)$, 则

$$B_n^2 = nd_1 d_2 \sum_{t=1}^T (\sigma^{tt})^{-1} p_t(u) (\varrho_0 + 2\varrho_1 + \varrho_2) + o(n).$$

简单计算可得

$$\sum_{i=1}^n \mathbb{E}|b_i|^3 \leq O(1) \cdot \sum_{i=1}^n \sum_{t=1}^T h^{3/2} \mathbb{E} \left\{ |d_1| + |d_2| \cdot \left| \frac{U_{it} - u}{h} \right| \right\}^3 K_h^3 (U_{it} - u) = O(h^{-1/2}).$$

因此, Linderberg 条件 $\lim_{n \rightarrow \infty} B_n^{-3} \sum_{i=1}^n \mathbb{E}|b_i^3| = 0$ 满足. 应用中心极限定理以及 $\hat{\sigma}^{tt_1}$ 的相合性可得定理 5.1 成立. \square

Statistical inference for partially linear single-index model of panel data with serially correlated error structure

Qunfang Xu, Gaosheng Liu & Yang Bai

Abstract Due to its flexibility, partially linear single-index model arises in many contemporary scientific endeavor. In this paper, we set foot on its inference under settings of panel data and a serially correlated error component structure. By combining local polynomial technique with the bias-corrected generalized estimating equations, we propose a feasible weighted generalized estimating equation estimation (GEE-FW) for unknown regression parameters. The GEE-FW is shown to be asymptotically normal and more efficient than the unweighted generalized estimating equation estimation based on working independence (GEE-WI). Moreover, a two stage local linear generalized estimating equation estimation (GEE-TS) of the unknown link function is also proposed, which takes the contemporaneous correlation into account and achieves an increase in its estimated efficiency. The asymptotic property of the GEE-TS is established as well and we show that it is asymptotically more efficient than the one which ignores the contemporaneous correlation. Conducted numerical simulation studies and actual data analysis show that the proposed estimate methods are effective and have good performance in both theory and application.

Keywords panel data, partially linear single-index model, serial correlation

MSC(2010) 62G08, 62J02

doi: 10.1360/N012018-00230