arXiv:2602.01595v1 [stat.ME] 2 Feb 2026

通过最小方差加权对一般连续处理模型进行数据驱动的均匀推断*

艾春荣 ††，黄伟 ‡‡，张征 §⋆

†管理与经济学院，香港中文大学（深圳）

‡ 数学与统计学院，墨尔本大学⋆应用统计中心，统计学与大数据研究院，中国人民大学

摘要

Ai 等人（2021年）研究了连续处理的一般剂量反应函数（GDRF）的估计，该函数包括平均剂量反应函数、分位数剂量反应函数，以及剂量反应分布的其他期望分位数。他们将 GDRF 设定为仅依赖处理状态的参数化函数，并提出了一种加权回归，其加权函数通过最大熵方法估计。本文将 GDRF 设定为处理状态的非参数函数，提出了用于估计 GDRF 的加权局部线性回归，并开发了一种用于构造一致置信带的自助法程序。我们提出了方差最小的稳定权重，同时消除处理与混杂变量之间的样本关联。所提出的权重具有闭式表达式，因此能够在自助抽样中高效计算。在一定条件下，我们推导出所提出的 GDRF 估计量的一致 Bahadur 表示，并证明了相应一致置信带的有效性。文中还包括一种完全由数据驱动的欠平滑调优参数选择方法，以及一种由数据驱动的偏差控制置信带。模拟研究和一项应用展示了所提出方法的实用性。

*作者按姓氏字母顺序排列。

†电子邮箱：chunrongai@cuhk.edu.cn

‡电子邮箱：wei.huang@unimelb.edu.au§电子邮箱：zhengzhang@ruc.edu.cn

关键词：因果推断；连续处理；协变量平衡；分位数处理效应；一致置信带。

1 引言

Ai 等人(2021年) 研究了一般剂量反应函数（GDRF）的估计，该函数涵盖平均剂量反应函数（ADRF）、分位数剂量反应函数（QDRF）以及剂量反应分布的其他期望分位数（百分位数）。他们将 GDRF 设定为仅关于处理状态的参数化函数，并提出了一种加权回归估计量，其权重函数通过最大熵方法估计。在一些充分条件下，他们建立了估计模型参数的大样本性质，例如，根n 渐近正态性和效率界。本文将 GDRF 设定为处理状态的非参数函数，并提出加权局部线性回归。在一些充分条件下，我们建立了 GDRF 非参数估计量及其一致置信带的渐近性质。这些结果很有帮助，因为对 GDRF 进行错误的参数化可能会导致错误结论，而一致置信带则能对处理效应参数进行更有力的推断。例如，我们可能希望检验母亲年龄在某一年龄范围内变化（即连续处理）是否会显著影响婴儿出生体重，或者分析商业广告投放量（即连续处理）对销售额的影响。

我们的论文还对一致非参数推断文献作出了贡献。不同于现有关于概率密度函数（Chernozhukov 等人，2014a；Kato 和 Sasaki，2018）、累积分布函数（Adusumilli 等人，2020）以及条件均值回归函数（Horowitz 和 Lee，2017；Kato 和 Sasaki，2019；Li 和 Liao，2020）的一致非参数推断研究——这些研究中的感兴趣参数具有闭式表达式，且不依赖额外的干扰参数——我们的广义剂量反应函数（GDRF）被定义为一般加权优化问题的一个隐式解，其中加权函数未知。准则函数可能是非光滑的（例如分位数损失函数）。这些显著差异带来了新的挑战，使以往的方法无法适用。为解决上述问题，我们提出了一种通过最小化样本方差来估计权重的新方法，同时施加数量不断增加的矩约束。最小方差权重具有闭式表达式，可以稳定且高效地计算。随后，我们基于估计量的准则函数，通过加权自助法构建了一致推断程序，同时还纳入了完全由数据驱动的推断程序。我们推导了所提出的 GDRF 估计量的一致 Bahadur 表示，并验证了一致置信带的构造。蒙特卡洛模拟研究和真实数据应用表明，所提出方法具有实用价值。

本文结构安排如下。第2节介绍 GDRF 参数及本文的研究目标。第3 节介绍我们的方法，包括 GDRF 参数的非参数估计与推断。第4 节阐述估计得到的 GHTE 的理论性质，以及一致推断的理论有效性。第5节给出实用建议，包括数据驱动的欠平滑调参选择器，以及数据驱动的偏差

校正置信带。第6 节报告模拟研究，第7 节介绍一项实证应用。所有技术证明及其他模拟结果均收录于补充材料中。

2 基本框架

令 T 表示观测到的连续处理状态变量，其支撑为 $\mathcal{T} \subset \mathbb{R}$，概率密度函数为 $f_T(t)$。令 $Y^*(t)$ 表示分配处理 t 时的潜在响应。$Y^*(t)$ 的概率密度存在，记为 $f_{Y^*(t)}$，且连续可微。令 $Y = Y^*(T)$ 表示观测到的响应。令 $\mathcal{L}(\cdot)$ 表示一个非负、严格凸且可能非光滑的损失函数，并满足 $\mathcal{L}(0) = 0$，且对于所有 $v \in \mathbb{R}$，$\mathcal{L}(v) \geq 0$。$\mathcal{L}(\cdot)$ 的导数记为 $\mathcal{L}'(\cdot)$，其几乎处处存在且非常数。令 $\boldsymbol{X} \in \mathbb{R}^{d_X}$ 表示一个可观测协变量向量，其支撑为 $\mathcal{X}$，其中 $d_X \geq 1$ 为某个整数。

我们考虑一般剂量反应函数（GDRF）$g(\cdot)$，其定义为以下凸优化问题的唯一解：

$$g(t) := \underset{a \in \mathbb{R}}{\operatorname{argmin}} \, \mathbb{E}\left[\mathcal{L}\left\{Y^*(t) - a\right\}\right], \tag{2.1}$$

对每一个 $t \in \mathcal{T}$。一般连续处理效应（广义因果处理效应）定义为 $\tau(t_1, t_0) := g(t_1) - g(t_0)$，即当处理状态从 $T = t_0$ 离散变化到 $T = t_1$ 时；或一般边际处理效应（GMTE）$g'(t)$，即 $g(t)$的导数，当处理状态发生边际变化时。为了根据可观测变量识别（2.1）中的 $g(\cdot)$，本文始终施加以下被广泛接受的无混杂条件（Hirano和Imbens，2004）。

假设1 (无混杂)。 对所有 $t \in \mathcal{T}$，给定 $\boldsymbol{X}$，T 独立于 $Y^*(t)$，即$Y^*(t) \perp T|\boldsymbol{X}$，对所有 $t \in \mathcal{T}$。

令 $f_{T|X}$ 表示 T 在给定观测协变量 $\boldsymbol{X}$时的条件概率密度函数，在Hirano和Imbens（2004）中称为广义倾向得分（GPS）。在假设1下，广义剂量反应函数 $g(t)$，求解

$$g(t) = \underset{a \in \mathbb{R}}{\operatorname{argmin}} \, \mathbb{E}\left[\pi(T, \boldsymbol{X})\mathcal{L}\left\{Y - a\right\} | T = t\right]. \tag{2.2}$$

其中

$$\pi(T, \boldsymbol{X}) := \frac{f_T(T)}{f_{T|X}(T|\boldsymbol{X})} \tag{2.3}$$

是 稳定化权重函数 (Robins 等人，2000；Ai 等人，2021年）。式（2.2）的证明见附录A。

模型（2.1）涵盖了多种感兴趣的连续处理效应参数。例如，取 $\mathcal{L}(v) = v^2$ (2.1)，则得到 $g(t) = \mathbb{E}\{Y^*(t)\}$平均剂量反应函数（平均剂量反应函数），该函数在现有文献中得到广泛研究，参见Kennedy等人（2017）；Colangelo和Lee（2020）；Huang和Zhang（2023）。取 $\mathcal{L}(v) = v\{q - \mathbb{1}(v \leq 0)\}$，对某个$q \in (0, 1)$，模型（2.1）给出分位数剂量反应函数（QDRF）$g(t) = F^{-1}_{Y^*}(t)(q) = \inf\{\rho : \mathbb{P}\{Y^*(t) \geq \rho\} \leq q\}$. 值得注意的是，上述所有论文仅关注逐点

平均剂量反应函数的非参数推断，我们尚未发现任何文献研究其一致渐近性质，更不用说关于广义剂量反应函数一致推断的更一般结果。本文通过为一般因果效应函数$g(\cdot)$、 $\tau(\cdot,\cdot)$和 $g'(\cdot)$建立非参数估计量、给出其一致渐近结果并构造一致置信带，填补了这一空白。一致置信带具有重要作用，因为我们可以用它们来检验原假设：

$$H_0 : g(t) = g_0(t) \text{ for all } t \in \mathcal{T}; \tag{2.4}$$
$$H_0 : \tau(t_1, t_0) = \tau_0(t_1, t_0) \text{ for all } (t_0, t_1) \in \mathcal{T} \times \mathcal{T}; \tag{2.5}$$

其中 $g_0(t)$ 与 $\tau_0(t_1, t_0)$ 被假定为已知函数，以及如下 nu ll

$$H_0 : g'(t) = 0 \text{ for all } t \in \mathcal{T}. \tag{2.6}$$

我们首先简要说明构造 g 的非参数一致带的思路。针对g' 和 τ的带可以用同样的方式建立。基于（2.2），一旦我们有了估计量 $\widehat{\pi}$ 用于 π，我们考虑估计 $g(t)$ 和 $g'(t)$，同时采用局部线性回归：[1]

$$(\widehat{g}_h(t), \widehat{g'}_h(t)) = \underset{(\theta_1,\theta_2)\in\mathbb{R}^2}{\operatorname{argmin}} \sum_{i=1}^{N} \widehat{\pi}(T_i, \boldsymbol{X}_i)\mathcal{L}\left\{Y_i - \theta_1 - \theta_2(T_i - t)\right\}\mathcal{K}_h\left(T_i - t\right), \tag{2.7}$$

其中 $\mathcal{K}_h(T_i - t) := \mathcal{K}((T_i - t)/h)$， $\mathcal{K}(\cdot)$ 为核函数， h 为带宽。因此，我们可以将广义因果处理效应 $\tau(t_1, t_0)$ 估计为 $\widehat{\tau}(t_1, t_0) := \widehat{g}_h(t_1) - \widehat{g}_h(t_0)$。令

$$Z_{g,N}(t) := \frac{\widehat{g}_h(t) - g(t)}{\sigma_{g,N}(t, h)},$$

and

$$I_{g,N}(t) := [\widehat{g}_h(t) - c_N(\alpha) \cdot \sigma_{g,N}(t, h), \widehat{g}_h(t) + c_N(\alpha) \cdot \sigma_{g,N}(t, h)],$$

其中 $\sigma_{g,N}(t, h)$ 是 $\widehat{g}_h(t)$ 的极限分布的标准差， $c_N(\alpha)$ 是 $\sup_{t\in\mathcal{T}} |Z_{g,N}(t)|$ 的极限分布的 $(1 - \alpha)$-分位数。则 $\{I_{g,N}(t)\ t \in \mathcal{T}\}$ 是 g 的 $(1 - \alpha)$ 非参数一致带，其意义在于

$$\lim_{N\to\infty} \mathbb{P}\{g(t) \in I_{g,N}(t), \forall t \in \mathcal{T}\} = 1 - \alpha.$$

为确定临界值 $c_N(\alpha)$，我们考虑估计 $c_N(\alpha)$通过加权自助法算法（Zhang andZheng, 2020）。具体而言，令 B 为自助法的次数，第 b^{th} 次自助法 $\widehat{}$版本 $(\widehat{g}_h,\ g'_h)$ 为

$$(\widehat{g}_h^{(b)}(t), \widehat{g'}_h^{(b)}(t)) = \underset{(\theta_1,\theta_2)}{\operatorname{argmin}} \sum_{i=1}^{N} \xi_i^{(b)}\widehat{\pi}^{(b)}(T_i, \boldsymbol{X}_i)\mathcal{L}\left\{Y_i - \theta_1 - (T_i - t)\theta_2\right\}\mathcal{K}_h\left(T_i - t\right), \tag{2.8}$$

[1]一般而言， $\widehat{g}$和 $\widehat{\tau}$缺乏显式表达式，因此需要通过迭代数值方法进行估计。例如，要计算分位数处理效应，可以使用Lejeune 和 Sarda (1988) 提出的迭代重加权最小二乘算法。

其中，$\xi_i^{(b)}$ 为相互独立的正随机变量，均值为 1、方差为 1，且$b = 1, \ldots, B$。例如，我们可以取一组独立同分布的 $2 \times \text{Bernoulli}(1/2)$ 随机变量。权重 $\widehat{\pi}^{(b)}$ 是 $\widehat{\pi}$ 的 b^{th} 自助法版本。随后，我们用 $\{\widehat{Z}_{g,N}^{(b)}\}_{b=1}^{B}$ 的 $(1-\alpha)$ 经验分位数估计 $c_N(\alpha)$，其中

$$\widehat{Z}_{g,N}^{(b)} := \frac{\widehat{g}_h^{(b)}(t) - \widehat{g}_h(t)}{\widehat{\sigma}_{g,N}(t,h)},$$

其中，$\widehat{\sigma}_{g,N}(t,h)$ 是 $\sigma_{g,N}(t,h)$ 的一个估计量。

为从实践和理论两方面确立我们的一致置信带的有效性，我们需要完成以下三个部分：

第一部分：提出一个能够轻松、高效计算的 $\widehat{\pi}$ 估计量及其自助法版本 $\{\widehat{\pi}^{(b)}\}_{b=1}^{B}$，用于估计 π。正如（2.8）中所强调的，构造一致置信带需要计算自助法权重 $\{\widehat{\pi}^{(b)}\}_{b=1}^{B}$，重复 B 次，这可能会带来较大的计算负担。[2] 为提高计算效率和可行性，我们在第 3 节提出一种新的 π 估计量及其自助法对应版本，它们具有闭式解，从而大大简化了推断过程。

第二部分：建立 $(\widehat{g}_h, \widehat{g'}_h)$ 和 $(\widehat{g}_h^{(b)}, \widehat{g'}_h^{(b)})$ 对 $b = 1, \ldots, B$ 的一致巴哈杜尔表示。借助一致巴哈杜尔表示，通过应用经验过程上确界的高斯逼近理论（Chernozhukov 等人(2014b)）以及紧高斯过程的反集中性质（Chernozhukov 等人 (2014a)），我们将在第 4 节确立自助法一致置信带的有效性。

第三部分：提供实践细节，包括数据驱动的欠平滑调优参数，以及在第5节中构造具有理论有效性的数据驱动偏差控制置信带。

注1. *Ai 等人（2021年）*研究了参数响应函数 $g(t, \boldsymbol{\beta}_0)$，其中有限维参数 $\boldsymbol{\beta}_0 \in \mathbb{R}^p$ 求解如下优化问题：

$$\boldsymbol{\beta}_0 = \underset{\boldsymbol{\beta}}{\operatorname{argmin}}\, \mathbb{E}\left[\pi(T, \boldsymbol{X})\mathcal{L}\{Y - g(T;\boldsymbol{\beta})\}\right],$$

其中 $g(\cdot, \cdot)$ 是一个已知函数。这种参数响应函数可能被错误设定，从而导致错误的结论。

注 2. 估计广义剂量反应函数的局部线性回归 (2.7) 的另一种方法是 *Chen* 等人 提出的筛选广义经验似然（*GEL*）*(2019)* 。注意，$g(T)$满足以下条件矩约束，

$$\mathbb{E}\left[\pi(T, \boldsymbol{X})\mathcal{L}'(Y - g(T))|T\right] = 0,$$

[2] Ai 等人（2021年）为 π 通过最大熵方法提出了一个估计量。然而，他们的估计量没有闭式表达式，需要使用数值方法计算，这在实践中不适合进行自助法重复。

这意味着

$$\mathbb{E}\left[\pi(T,\boldsymbol{X})\mathcal{L}'(Y-g(T))v_{K_1}(T)\right]=0,$$

其中 $v_{K_1}(T)$ 是一个 K_1维逼近筛。令 $s(\cdot)$ 为严格凹的二次连续可微函数，具有*Lipschitz*连续二阶导数，并满足 $s'(0)=s''(0)=-1$。遵循 *Chen* 等人 *(2019)*， $g(t)$的序贯*GEL*估计量由下式给出

$$\widehat{g}_{GEL}(t)=\widehat{\alpha}_{K_0}^{\top}v_{K_0}(t),$$

其中 $\widehat{\alpha}_{K_0}\in\mathbb{R}^{K_1}$ 求解如下鞍点优化问题：

$$\widehat{\alpha}_{K_0}=\arg\min_{\alpha\in\mathbb{R}^{K_0}}\max_{\gamma\in\mathbb{R}^{K_1}}\frac{1}{N}\sum_{i=1}^{N}s\left\{\widehat{\pi}(T_i,\boldsymbol{X}_i)\mathcal{L}'(Y_i-\alpha^{\top}v_{K_0}(T_i))\cdot\gamma^{\top}v_{K_1}(T_i)\right\}-s(0).$$

然而，对于一般的 $\mathcal{L}'(\cdot)$ （其可能不可微）以及估计得到的 $\widehat{\pi}$， $g(\cdot)$ 与 $g'(\cdot)$ 的一致推断更难以建立。我们将在未来的研究中推进这一扩展。

注3. 双重稳健处理效应估计已日益流行。对于连续处理，*Kennedy*等人 *(2017)*基于双重稳健表示提出了平均剂量反应函数的非参数估计量，即 $\mathbb{E}[Y^*(t)]=\mathbb{E}[\xi(T,\boldsymbol{X},Y)|T=t]$，其中

$$\xi(T,\boldsymbol{X},Y):=\pi(T,\boldsymbol{X})\{Y-m(T,\boldsymbol{X})\}+\int m(T,\boldsymbol{x})f_X(\boldsymbol{x})d\boldsymbol{x}$$

且 $m(T,\boldsymbol{X}):=\mathbb{E}[Y|T,\boldsymbol{X}]$ 是结果回归函数。随后，他们通过局部线性^多项式回归，将估计得到的伪结果 $\xi(T,\boldsymbol{X},Y)$ 对 $T=t$ 进行回归，从而估计 $\mathbb{E}[Y^*(t)]$，即 $\theta_1(t)$，其中

$$(\widehat{\theta}_1(t),\widehat{\theta}_2(t)):=\arg\min_{(\theta_1,\theta_2)\in\mathbb{R}^2}\sum_{i=1}^{N}\mathcal{K}_h\left(T_i-t\right)\left\{\widehat{\xi}(T_i,\boldsymbol{X}_i,Y_i)-\theta_1-(T_i-t)\theta_2\right\}^2,$$

$$\widehat{\xi}(T,\boldsymbol{X},Y):=\widehat{\pi}(T,\boldsymbol{X})\{Y-\widehat{m}(T,\boldsymbol{X})\}+\frac{1}{N}\sum_{i=1}^{N}\widehat{m}(T,\boldsymbol{X}_i),$$

其中， $\widehat{\pi}(\cdot)$ 和 $\widehat{m}(\cdot)$ 是 $\pi(\cdot)$ 和 $m(\cdot)$ 的相合非参数估计量。*Kennedy*^等人（*2017*，定理*2*）针对每个固定的$t\in\mathcal{T}$，建立了 $\theta_1(t)$ 的逐点渐近结果，并证明了其双重稳健性。然而，他们的方法依赖于平均剂量反应函数 *ADRF* 双重稳健表示的闭式形式，因此难以将其推广到估计我们的广义剂量反应函数 *GDRF*$g(\cdot)$；后者是一个隐式定义的因果函数。此外，关于 $\mathbb{E}[Y^*(t)]$ 在 $t\in\mathcal{T}$上的一致推断，目前尚无相关研究。未来工作中，我们将研究 $g(\cdot)$ 的双重稳健估计及其一致推断。

3 最小方差权重与一致推断算法

可以采用多种方法来估计 π。例如，可以分别估计（2.3）中的分子和分母，然后构造比率估计量。然而，这种估计量

已知会因对分母估计的敏感性而不稳定（Kang等人，2007年）。为解决这一不稳定问题，Ai 等人（2021年）提出直接估计 π，方法是最小化 $\{N^{-1}\pi(T_i, \boldsymbol{X}_i)\}_{i=1}^N$ 与均匀设计权重 $\{N^{-1}\}_{i=1}^N$ 之间的 Kullback-Leibler（KL）散度，并施加一组不断扩展的乘积矩约束。该方法虽解决了不稳定问题，但无法提供闭式解，且计算时需要采用迭代数值方法。为克服上述问题，我们对Ai 等人（2021年）的方法进行扩展，以平方损失替代 KL 散度，使估计得到的权重具有闭式表达式，从而能够稳定、高效地计算。

我们注意到，对于任意适当的函数 $u(t, \boldsymbol{x})$，权重函数 $\pi(t, \boldsymbol{x})$ 满足矩约束，

$$\mathbb{E}\{\pi(T, \boldsymbol{X})u(T, \boldsymbol{X})\} = \int u(t, \boldsymbol{x})f_T(t)f_X(\boldsymbol{x})dtd\boldsymbol{x}. \tag{3.1}$$

方程 (3.1) 确定了 $\pi(t, \boldsymbol{x})$，因此我们可以通过求解 (3.1) 的样本类比来估计 $\pi(T_i, \boldsymbol{X}_i)$。挑战在于 (3.1) 包含无限多个约束。我们不可能对有限个样本观测施加无限多个约束。为克服这一困难，我们用一列有限维筛空间逼近一个函数空间。具体而言，令 $u_K(T, \boldsymbol{X}) = (u_{K,1}(T, \boldsymbol{X}) \ldots, u_{K,K}(T, \boldsymbol{X}))^\top$ 表示逼近筛，例如 B样条和幂级数（关于筛逼近的更多讨论，参见 Chen，2007年）。$\pi(T, \boldsymbol{X})$ 也满足

$$\mathbb{E}\{\pi(T, \boldsymbol{X})u_K(T, \boldsymbol{X})\} = \int u_K(t, \boldsymbol{x})f_T(t)f_X(\boldsymbol{x})dtd\boldsymbol{x}\,. \tag{3.2}$$

注意到 $\mathbb{E}\{\pi(T, \boldsymbol{X})\} = 1$ 且 $\text{var}\{\pi(T, \boldsymbol{X})\} = \mathbb{E}[\{\pi(T, \boldsymbol{X}) - 1\}^2]$，结合矩约束（3.2），我们提出非参数最小方差权重 $\widehat{\pi}_K(T_i, \boldsymbol{X}_i)$:

$$\begin{cases} \{\widehat{\pi}_K(T_i, \boldsymbol{X}_i)\}_{i=1}^N = \text{argmin}_{\pi_i} \sum_{i=1}^N (\pi_i - 1)^2/2 \\ \text{subject to } \frac{1}{N}\sum_{i=1}^N \pi_i u_K(T_i, \boldsymbol{X}_i) = \frac{1}{N(N-1)}\sum_{j=1, j\neq i}^N \sum_{i=1}^N u_K(T_i, \boldsymbol{X}_j). \end{cases} \tag{3.3}$$

注意，当 $u_{K,\,1}(T, \boldsymbol{X}) = 1$ 包含在筛基中时，(3.3) 保证$\sum_{i=1}^N \widehat{\pi}_K(T_i, \boldsymbol{X}_i)/N = 1$。在附录 B中，我们证明原始问题 (3.3) 存在闭式解，

$$\begin{aligned} \widehat{\pi}_K(t, \boldsymbol{x}) = &- \left(\frac{1}{N}\sum_{i=1}^N u_K(T_i, \boldsymbol{X}_i) - \frac{1}{N(N-1)}\sum_{i=1}^N \sum_{j=1, i\neq j}^N u_K(T_i, \boldsymbol{X}_j)\right)^\top \\ &\times \left(\frac{1}{N}\sum_{i=1}^N u_K(T_i, \boldsymbol{X}_i)\, u_K^\top(T_i, \boldsymbol{X}_i)\right)^{-1} u_K(t, \boldsymbol{x}) + 1. \end{aligned} \tag{3.4}$$

相应地，自助法权重为

$$\widehat{\pi}_K^{(b)}(t, \boldsymbol{x}) = - \left(\frac{1}{N}\sum_{i=1}^N \xi_i^{(b)} u_K(T_i, \boldsymbol{X}_i) - \frac{1}{N(N-1)}\sum_{i=1}^N \sum_{j=1, i\neq j}^N u_K(T_i, \boldsymbol{X}_j)\right)^\top$$

$$\times \left(\frac{1}{N} \sum_{i=1}^{N} \xi_i^{(b)} u_K\left(T_i, \boldsymbol{X}_i\right) u_K^{\top}\left(T_i, \boldsymbol{X}_i\right) \right)^{-1} u_K\left(t, \boldsymbol{x}\right) + 1. \tag{3.5}$$

现在，我们介绍为 g 构建一致置信带的过程。对于 g' 和 τ，我们可以使用相同的算法构建一致置信带：

1. 在 $\widehat{g}_h(t)$ 中，根据3.4) 中的 $\widehat{\pi}$ 替换为 $\widehat{\pi}_K$，计算（2.2）中的估计量，并在其支撑上取足够精细的网格，$\mathcal{T}$。

2. 取一个足够大的整数 B。在相同网格上，对 $\widehat{g}_h^{(b)}(t)$，根据3.5) 中的 $\widehat{\pi}^{(b)}$ 替换为 $\widehat{\pi}_K^{(b)}$，计算（2.8）中的自助估计量，令 $b = 1, \ldots\ldots, B$。

3. 计算渐近方差的一致估计，$\widehat{\sigma}_{g,N}(t)$。例如，我们可以用自助标准差估计 $\sigma_{g,N}(t)$，即 $\{\widehat{g}_h^{(b)}(t) - \widehat{g}_h(t)\}_{b=1}^{B}$的标准差。或者，我们可以采用自助标准化四分位距，即 $\widehat{\sigma}_{g,N}(t) = \{q_{0.75}(t) - q_{0.25}(t)\}/\{z_{0.75}(t) - z_{0.25}(t)\}$，其中 $q_p(t)$ 和 $z_p(t)$ 分别是 $\{\widehat{g}_h^{(b)}(t) - \widehat{g}_h(t)\}_{b=1}^{B}$和 $N(0,1)$ 的第 p 分位数。

4. 计算 $M_{g,b} := \sup_{t\in\mathcal{T}} |\widehat{Z}_{g,N}^{(b)}(t)|$，其中 $b = 1, \ldots, B$，所选网格点上的最大值近似于上确界。

5. 给定置信水平 $1-\alpha$，求数集 $\{M_{g,b}: b = 1, \ldots, B\}$ 的经验 $1-\alpha$ 分位数，记为 $C_{g,\alpha}$。

6. g 的双侧一致置信带构造如下

$$\widehat{I}_{g,N} := \left\{ \left(\widehat{g}_h(t) - \widehat{C}_{g,\alpha}\widehat{\sigma}_{g,N}(t,h), \widehat{g}_h(t) + \widehat{C}_{g,\alpha}\widehat{\sigma}_{g,N}(t,h) \right) : t \in \mathcal{T} \right\}.$$

$g'(\cdot)$ 和 $\tau(\cdot, \cdot)$ 的置信带分别记为 $I_{g',N}$ 和 $I_{\tau,N}$，可通过在上述算法中分别将 $\{\widehat{g}_h, \widehat{g}_h^{(b)}\}_{b=1}^{B}$ 替换为 $\{\widehat{g'}_h, \widehat{g'}_h^{(b)}\}_{b=1}^{B}$ 和 $\{\widehat{\tau}(t_1, t_0) := \widehat{g}_h(t_1) - \widehat{g}_h(t_0)\ \widehat{\tau}^{(b)} := \widehat{g}_h^{(b)}(t_1) - \widehat{g}_h^{(b)}(t_0)\}_{b=1}^{B}$ 来构造。

注 *4. Yiu* 和 *Su*（*2018*）中的最小方差权重是参数化的。请注意，T与 $\boldsymbol{X}$之间由 $f_{T|X}(T|\boldsymbol{X})$刻画的关联，在经过 $\pi(T, \boldsymbol{X})$加权后可以消除，即 $\pi(T, \boldsymbol{X}) f_{T|X}(T|\boldsymbol{X}) = f_T(T)$。作者考虑参数模型 $T \sim f_T(T; \alpha)$和 $T|\boldsymbol{X} \sim f_0(T|\boldsymbol{X}; \beta = (\beta_b, \beta_d))$，其中 β_d指描述 T 与 $\boldsymbol{X}$ 之间依赖关系的参数，而 β_b 是基线分布的参数。他们假设 $f_0(T|\boldsymbol{X}; \beta_b, \beta_d = 0) = f_T(T; \alpha)$。然后，给定一组权重 $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_N)$，可以通过求解加权得分方程来估计 β_d

$$\sum_{i=1}^{N} \pi_i \frac{\partial}{\partial \beta(\boldsymbol{\pi})} \log f_0(T_i|\boldsymbol{X}_i; \beta(\boldsymbol{\pi})) \Big|_{\beta_b(\boldsymbol{\pi})=\alpha} = 0.$$

他们的想法是，如果权重 $\boldsymbol{\pi}$ 满足协变量关联消除条件，就像$\pi(T, \boldsymbol{X})$ 那样，那么 $\beta_d = 0$ 就应该是加权得分方程的解。因此，他们将最小方差权重估计量定义如下：

$$\begin{cases} \{\widehat{\pi}_i\}_{i=1}^N = \arg\min \sum_{i=1}^N \{\pi_i - 1\}^2 /2 \\ \textit{subject to } \sum_{i=1}^N \pi_i \frac{\partial}{\partial \beta(\boldsymbol{\pi})} \log f_0(T_i | \beta(\boldsymbol{\pi})^\top \boldsymbol{X}_i) \Big|_{\beta_b(\boldsymbol{\pi})=\widehat{\alpha}, \beta_d(\boldsymbol{\pi})=0} = 0, \\ \frac{1}{N}\sum_{i=1}^N \pi_i = 1,\ \pi_i > 0\ (i = 1, ..., N), \end{cases}$$

其中，$\widehat{\alpha}$通过拟合模型 $f_T(T_i;\ \alpha)$得到，该模型基于观测数据。然而，*Yiu* 和 *Su(2018)* 并未研究权重 $\{\widehat{\pi}_i\}_{i=1}^N$ 的理论性质，也未研究由此得到的处理效应估计量。相比之下，我们提出非参数最小方差权重，并推导了权重及其所产生的处理效应估计量的理论结果。

4 理论性质

本节将通过完成第2节中的第二部分，建立我们估计量的一致渐近性质，并验证我们的一致置信带。为便于阐述，我们引入一些记号。对于任意矩阵 ,:tr()表示的弗罗贝尼乌斯范数。对于任意可测函数()和任意概率测度，令，$_2 = \left(\int |f(\boldsymbol{o})|^2 dQ(\boldsymbol{o})\right)^{1/2}$表示 $L^2(Q)$范数。对于任意序列 a_N，$b_N \in \mathbb{R}^+$,$a_N \prec b_N$ 表示 $a_N/b_N \to 0$ 当 $N \to \infty$时成立，而 $a_N \asymp b_N$ 表示对所有 N均有界且远离零。对于任意 a，$b \in \mathbb{R}$，记 $a \vee b =$ max(a，b)。

4.1 一致巴哈杜尔表示

^本节将建立我们的估计量 $\widehat{g}_h, g'_h$和 $\widehat{\tau}$的一致线性（Bahadur）表示，它们具有影响函数展开。

本文余下部分均要求并维持以下条件。

假设2。 函数 $g(t)$ 关于 $t \in \mathcal{T}$二次连续可微。

假设3。 $\{T_i, \boldsymbol{X}_i, Y_i\}_{i=1}^N$ 是从$(T, \boldsymbol{X}, Y)$的分布中独立抽取的观测值。$(T, \boldsymbol{X})$ 的支撑 $\mathcal{T} \times \mathcal{X}$是紧集。

假设4。 存在两个正常数 η和 η，使得对于所有$(t, \boldsymbol{x}) \in \mathcal{T} \times \mathcal{X}$，$0 < \eta \leq \pi(t, \boldsymbol{x}) - \leq \eta < \infty$。

假设5。存在 $\boldsymbol{\lambda}_K \in \mathbb{R}^K$ 和一个正数 $\omega > 0$，使得上确界$_{(t,\boldsymbol{x})\in\mathcal{T}\times\mathcal{X}} |\pi(t, \boldsymbol{x}) - \boldsymbol{\lambda}_K^\top u_K(t, \boldsymbol{x})| = O(K^{-\omega})$。

假设6. $\mathbb{E}[u_K(T, \boldsymbol{X}) u_K^\top(T, \boldsymbol{X})]$的特征值在 K上统一有上界，且与零保持正距离。

假设7。

~
(i) $\phi(a|t) := \mathbb{E}\{\mathcal{L}(Y-a)|T=t\}$ 和 $\phi(a|t) := \mathbb{E}\{\pi(T,\boldsymbol{X})\mathcal{L}(Y-a)|T=t\}$ 关于 a $t \in \mathcal{T}$二次连续可微，并且对于每个固定的 a，所有
~
导数均连续。此外，$\inf_{t\in\mathcal{T}} \partial_a^2 \phi(g(t)|t) > 0$。

（*ii*）核函数 $\mathcal{K}(v) \geq 0$ 连续且有界，并且 $\kappa_{j1} := \int_{-\infty}^{\infty} v^j \mathcal{K}(v)\, dv$满足 $\kappa_{0,\,1} = 1$, $\kappa_{1,\,1} = 0$和 $\kappa_{j,\,1} < \infty$ 对于 $j = 2$、3 和 4成立。

（*iii*）密度函数 $f_T(t)$ 连续，并满足 $\inf_{t\in\mathcal{T}} f_T(t) > 0$。

（*iv*）对于每个 $y \in \mathbb{R}$，条件密度函数 $f_{Y|T}(y|t)$ 关于 t 连续。存在正常数 $\varepsilon > 0$ 和正函数 $G(y|t)$，使得 $\sup_{\|t'-t\|\leq\epsilon} f_{Y|T}(y|t') \leq G(y|t)$ 对几乎所有 y成立，并且

$$\int \{\mathcal{L}(y-\delta) - \mathcal{L}(y) + \mathcal{L}'(y)\cdot\delta\}^2 G(y|t)dy = O(\delta^4) \text{ as } \delta \to 0$$

在 $t \in \mathcal{T}$ 上一致成立。

~
（*v*）函数 $\partial_a^2\phi(g(t)|t)$ 和 $e(t)$：$(g(t)|t)\, f_T(t)$ 关于 t 连续可微，其导数在 $t \in \mathcal{T}$ 上一致有界。

假设8。

(i) 存在一个充分大的常数 M，使得

$$\sup_{t\in\mathcal{T}} \{|g(t) - t\cdot g'(t)| \vee |g'(t)|\} \leq M.$$

函数类 $\mathcal{F}_1 = \{(y,t) \mapsto \mathcal{L}(y-\alpha-\beta\cdot t) \colon \alpha,\beta \in [-M,M]\}$ 和 $\mathcal{F}_2 = \{(y,t) \mapsto \mathcal{L}'(y-\alpha-\beta\cdot t) \colon \alpha,\beta \in [-M,M]\}$ 是 *Vapnik-Chervonenkis*（*VC*）型的，即对于某些常数 A 和 v,

$$\sup_Q \mathcal{N}(\mathcal{F}_j, \|\cdot\|_{Q,2}, \varepsilon\|F_j\|_{Q,2}) \leq \left(\frac{A}{\varepsilon}\right)^v,\ j = 1,2,$$

其中 $\mathcal{N}(\mathcal{F}_j, \|\cdot\|_{Q,2}, \varepsilon)$ 是 $\mathcal{F}_j$ 在尺度 ε 下关于范数$\|\cdot\|_{Q,2}$的覆盖数，F_j 是 $\mathcal{F}_j$的包络函数，且 Q 的上确界取遍 $\mathbb{R}^2$上的所有有限离散分布。

(ii) $\mathbb{E}\left[\sup_{\alpha,\beta\in[-M,M]} |\mathcal{L}'(Y-\alpha-\beta\cdot T)|^q\right] < \infty$ 对于某个 $q > 2$。

假设9。

(i) 假设对于某些常数 c_1、$c_2 > 0$，$h \asymp N^{-c_1}$ 和 $K \asymp N^{c_2}$ 成立。

（*ii*）当 $\zeta(K) = \text{上确界}_{(t,\boldsymbol{x})\in\mathcal{T}\times\mathcal{X}} \|u_K(t,\boldsymbol{x})\|$，$\zeta(K)^2 K \prec N^{1/2} \prec K^\omega$以及 $\{\zeta(K)^2 K^{1-2\omega}\} \vee (K^2/N) \prec N^{1-2/q}h^3$时，其中 ω定义见假设 5；

(*iii*) $h^2 \prec (Nh)^{-1/2}$。

假设3要求随机变量有界。这一条件虽然限制较强，但便于进行大样本推导。假设4要求权重函数有界且远离零。这一条件在协变量平衡文献中很常见。假设5要求筛逼近误差以多项式速率趋于零。假设6排除了逼近基函数之间的近似多重共线性。这一条件在筛回归文献中广为人知（Chen，2007年）。假设7与Fan等人（1994年）的条件A类似。利用Chernozhukov 等（2014b）的推论5.1建立一致推断结果时，需要假设8；常用的损失函数满足该假设。关于在$\mathcal{L}(v) = v^2$ 和 $\mathcal{L}(v) = v \cdot \{\tau - \mathbb{1}(v \leq 0)\}$下验证假设8(i)，请参见附录C。假设9是一个欠平滑条件。对于幂级数，$\zeta(K) = O(K)$；对于B样条，$\zeta(K) = O(\sqrt{K})$。

—

在这些条件下，我们为 $\widehat{g}$，建立如下渐近表示。$\widehat{g}$ 及其自助法，$\widehat{g}^{(b)}$，$\widehat{g'}^{(b)}$，对于 $b = 1, \ldots, B$。为了便于表述这些定理，我们定义

$$\psi_{0,h}(T_i, \boldsymbol{X}_i, Y_i; t) = \frac{1}{e(t)} \cdot \Big[\pi(T_i, \boldsymbol{X}_i)\mathcal{L}'(Y_i - g(t) - g'(t)(T_i - t)) - \mathbb{E}\{\pi(T_i, \boldsymbol{X}_i)\mathcal{L}'(Y_i - g(t) - g'(t)(T_i - t))|T_i, \boldsymbol{X}_i\}\Big]\mathcal{K}_h(T_i - t)$$

$$\text{and } \psi_{1,h}(T_i, \boldsymbol{X}_i, Y_i; t) = \frac{1}{e(t) \cdot \kappa_{21}} \cdot \Big[\pi(T_i, \boldsymbol{X}_i)\mathcal{L}'(Y_i - g(t) - g'(t)(T_i - t)) - \mathbb{E}\{\pi(T_i, \boldsymbol{X}_i)\mathcal{L}'(Y_i - g(t) - g'(t)(T_i - t))|T_i, \boldsymbol{X}_i\}\Big](T_i - t)\mathcal{K}_h(T_i - t)\,.$$

定理1. 在假设 *1-9*下，我们得到 $\widehat{g}_h(t)$ 和 $\widehat{g'}_h(t)$ 在 上一致成立的线性（*Bahadur*）表示：

$$\sqrt{Nh}\,\{\widehat{g}_h(t) - g(t)\} = \frac{1}{\sqrt{Nh}} \sum_{i=1}^{N} \psi_{0,h}(T_i, \boldsymbol{X}_i, Y_i; t) + R_{0,N}(t)$$

and

$$\sqrt{Nh^3}\left\{\widehat{g'}_h(t) - g'(t)\right\} = \frac{1}{\sqrt{Nh^3}} \sum_{i=1}^{N} \psi_{1,h}(T_i, \boldsymbol{X}_i, Y_i; t) + R_{1,N}(t)\,.$$

对于 $\widehat{g}^{(b)}$ 和 $\widehat{g'}^{(b)}$，对于 $b = 1$，……，B，我们有：

$$\sqrt{Nh}\left\{\widehat{g}_h^{(b)}(t) - \widehat{g}_h(t)\right\} = \frac{1}{\sqrt{Nh}} \sum_{i=1}^{N} \{\xi^{(b)} - 1\}\psi_{0,h}(T_i, \boldsymbol{X}_i, Y_i; t) + R_{0,N}^{(b)}(t)$$

and

$$\sqrt{Nh^3}\left\{\widehat{g'}_h^{(b)}(t) - \widehat{g'}_h(t)\right\} = \frac{1}{\sqrt{Nh^3}} \sum_{i=1}^{N} \{\xi^{(b)} - 1\}\psi_{1,h}(T_i, \boldsymbol{X}_i, Y_i; t) + R_{1,N}^{(b)}(t)\,.$$

此外，我们表明

$$\sup_{t\in\mathcal{T}}|R_{0,N}(t)| = o_P\left(\log^{-1} N\right), \quad \sup_{t\in\mathcal{T}}|R_{1,N}(t)| = o_P\left(\log^{-1} N\right);$$

以及对于 $b = 1$、.....、 B,

$$\sup_{t\in\mathcal{T}}|R^{(b)}_{0,N}(t)| = o_P\left(\log^{-1} N\right), \quad \sup_{t\in\mathcal{T}}|R^{(b)}_{1,N}(t)| = o_P\left(\log^{-1} N\right).$$

备注 *5.* 在定理*1*的证明中，我们表明，估计量 $\widehat{g}_h(t)$ 和 $\widehat{g'}_h(t)$ 的偏差速率为 h^2，且在 $t \in \mathcal{T}$ 上一致。方差 $Var\{\psi_{0,h}(T_i, \boldsymbol{X}_i, Y_i; t)\} \asymp h$ 和$Var\{\psi_{1,h}(T_i, \boldsymbol{X}_i, Y_i; t)\} \asymp h^3$。因此，对于估计量 $\widehat{g}_h(t)$ 和 $\widehat{g'}_h(t)$，在欠平滑条件假设*9 (ii)* 下，偏差相对于方差可以忽略不计。

$\widehat{\tau}(t_1, t_0)$，当处理水平从 t_0 变为 t_1时，估计一般连续处理效应函数的估计量，其渐近分布可通过应用定理 1得出。

推论1. 在假设*1-9*下，我们得到

$$\begin{aligned}&\sqrt{Nh}\{\widehat{\tau}(t_1,t_0) - \tau(t_1,t_0)\}\\ =&\frac{1}{\sqrt{Nh}}\sum_{i=1}^{N}\{\psi_{0,h}(T_i,\boldsymbol{X}_i,Y_i;t_1) - \psi_{0,h}(T_i,\boldsymbol{X}_i,Y_i;t_0)\} + R_{1,N}(t_1,t_0),\end{aligned}$$

且对于 $b = 1, \ldots, B$,

$$\begin{aligned}&\sqrt{Nh}\left\{\widehat{\tau}^{(b)}(t_1,t_0) - \widehat{\tau}(t_1,t_0)\right\}\\ =&\frac{1}{\sqrt{Nh}}\sum_{i=1}^{N}\{\xi^{(b)} - 1\}\{\psi_{0,h}(T_i,\boldsymbol{X}_i,Y_i;t_1) - \psi_{0,h}(T_i,\boldsymbol{X}_i,Y_i;t_0)\} + R^{(b)}_{1,N}(t_1,t_0),\end{aligned}$$

其中

$$\sup_{(t_0,t_1)\in\mathcal{T}\times\mathcal{T}}|R_N(t_0,t_1)| = o_P\left(\log^{-1} N\right) \textit{ and } \sup_{(t_0,t_1)\in\mathcal{T}\times\mathcal{T}}|R^{(b)}_N(t_0,t_1)| = o_P\left(\log^{-1} N\right).$$

4.2 置信带的有效性

根据定理1 和推论1，渐近主导项 $\widehat{g}_h(t) - g(t)$、 $g'_h(t) - g'(t)$ 以及 $\widehat{\tau}(t_1,t_0) - \tau(t_1,t_0)$ 的方差为

$$\sigma^2_{g,N}(t,h) = \frac{Var\{\psi_{0,h}(T_i,\boldsymbol{X}_i,Y_i;t)\}}{Nh^2}, \quad \sigma^2_{g',N}(t) = \frac{Var\{\psi_{1,h}(T_i,\boldsymbol{X}_i,Y_i;t)\}}{Nh^6},$$

and

$$\sigma^2_{\tau,N}(t_1,t_0,h) := \frac{Var\{\psi_{0,h}(T,\boldsymbol{X},Y;t_1) - \psi_{0,h}(T,\boldsymbol{X},Y;t_0)\}}{Nh^2},$$

分别。

应用李雅普诺夫中心极限定理（CLT）、定理 1 和推论 1，我们得到依分布的逐点收敛，

$$\lim_{N\to\infty} \frac{\widehat{\theta}^{(b)}(\boldsymbol{w}) - \widehat{\theta}(\boldsymbol{w})}{\sigma_{\theta,N}(\boldsymbol{w},h)} \overset{d}{=} \lim_{N\to\infty} \frac{\widehat{\theta}(\boldsymbol{w}) - \theta(\boldsymbol{w})}{\sigma_{\theta,N}(\boldsymbol{w},h)} \overset{d}{=} \mathcal{N}(0,1)\,, \tag{4.1}$$

对于 通用参数记号 $\theta(\boldsymbol{w}) = g(t)\ g'(t)$ 或 $\tau(t_1,\ t_0)$，对于每个 $t,\ t_0,\ t_1 \in \mathcal{T}$.

请注意，(4.1) 是基于自助法程序对 $g(t)$、$g'(t)$ 和 $\tau(t_1,t_0)$ 进行逐点推断有效性的基础。然而，上述逐点依分布收敛的结果无法加强为随机过程的（统一）弱收敛，因为 ^随机过程 $\{\widehat{g}_h(t) :\ t \in \mathcal{T}\}, \{\widehat{g'}_h(t) :\ t \in \mathcal{T}\}$ 和 $\{\widehat{\tau}(t_1,t_0) : (t_1,t_0)\ \in \mathcal{T}\times\mathcal{T}\}$ 并非渐近紧，即

$$\overline{\lim_{N\to\infty}} \sup_{\boldsymbol{w}\in\mathcal{W}} \frac{\widehat{\theta}(\boldsymbol{w}) - \theta(\boldsymbol{w})}{\sigma_{\theta,N}(\boldsymbol{w},h)} \to \infty,$$

对于 $\theta(\boldsymbol{w}) = g(t)\ g'(t)$ 或 $\tau(t_1,\ t_0)$，该情况以趋近于 1 的概率发生。

尽管标准化过程的上确界不存在极限分布，但秉承 Chernozhukov 等人 (2014a) 和 Chernozhukov 等人 (2014b) 的思路，我们将证明存在一个紧高斯过程，其上确界能够很好地逼近标准化过程的上确界。由此，我们可以验证一致推断。为完成这一步，我们需要以下附加条件：

假设10。 对于每个 $\widehat{\sigma}_{g,\ N}, \widehat{\sigma}_{g',\ N}$和 $\widehat{\sigma}_{\tau,\ N}$，对于任意 $\epsilon > 0$，存在 N_ϵ，使得对于任意 $N \geq N_\epsilon$，

$$\mathbb{P}\left(\sup_t \left|\frac{\widehat{\sigma}_{g,N}(t,h)}{\sigma_{g,N}(t,h)} - 1\right| \geq \epsilon/\log N\right) \leq \epsilon\,, \quad \mathbb{P}\left(\sup_t \left|\frac{\widehat{\sigma}_{g',N}(t)}{\sigma_{g',N}(t)} - 1\right| \geq \epsilon/\log N\right) \leq \epsilon\,,$$

and

$$\mathbb{P}\left(\sup_{t_1,t_0} \left|\frac{\widehat{\sigma}_{\tau,N}(t_1,t_0)}{\sigma_{\tau,N}(t_1,t_0)} - 1\right| \geq \epsilon/\log N\right) \leq \epsilon\,.$$

假设11。 $\mathbb{E}\left[\sup_{\alpha,\ \beta\in[-M,M]} |\mathcal{L}'(Y-\alpha-\beta\cdot T)|^q\right] < \infty$ 对于某个 $q \geq 4$ 和 $\sup_{t\in\mathcal{T}} \mathbb{E}\left[\sup_{\alpha,\ \beta\in[-M,M]} |\mathcal{L}'(Y-\alpha-\beta\cdot T)|^4 | T=t\right] < \infty$

只要渐近方差的估计量满足假设10，就可以确立所提出的置信带的渐近有效性。例如，正如 Chernozhukov 等人（2013，注3.2）所讨论的，提出于第3节的自助法标准化四分位距估计量和自助标准差在满足一些附加条件时符合该假设。具体而言，对于自助法标准化四分位距估计量，如果相应的目标方差 $\sigma^2_{g,N}(t)$（分别为 $\sigma^2_{g',N}(t)$ 和 $\sigma^2_{\tau,N}(t_1,t_0)$）对 $t \in \mathcal{T}$（分别为 $t \in \mathcal{T}$ 和 $(t_1,t_0) \in \mathcal{T}\times\mathcal{T}$）有正的下界，则可以验证 $\widehat{\sigma}^2_{g,N}(t)$（分别为 $\widehat{\sigma}^2_{g',N}(t)$ 和 $\widehat{\sigma}^2_{\tau,N}(t_1,t_0)$）满足该条件。当 $\{Nh(\widehat{g}^{(b)}_h(t) - \widehat{g}_h(t))^2 :\ t \in \mathcal{T}\}$（分别为 $\{Nh^3(\widehat{g'}^{(b)}_h(t) - \widehat{g'}_h(t))^2 :\ t \in \mathcal{T}\}$）一致可积时，自助标准差满足该假设（见 Kato，2011）。我们的模拟研究采用了自助标准差，结果表明其表现良好。此外，Chernozhukov 等（2014b）所要求的假设11，对于估计量渐近过程的紧高斯过程逼近至关重要。

定理 2. 假设*1-11* 成立。我们有

$$\lim_{N\to\infty}\mathbb{P}\left(g\in\widehat{I}_{g,N}\right)=\lim_{N\to\infty}\mathbb{P}\left(\tau_{g'}\in\widehat{I}_{g',N}\right)=\lim_{N\to\infty}\mathbb{P}\left(\tau\in\widehat{I}_{\tau,N}\right)=1-\alpha.$$

5 实践细节

5.1 基函数

在数值研究中，我们将筛基 u_K 设为乘积幂级数。具体而言，我们定义 $u_K(T,\ \boldsymbol{X})=w_{K_1}(T)\otimes w_{K_2}(\boldsymbol{X})$ 和 $K=(K_1+1)\cdot(K_2+1)$，其中，对于任意正整数 p 和 r，$w_p(\cdot):\mathbb{R}^r\mapsto\mathbb{R}^{pr+1}$:

$$w_p(\boldsymbol{v})=(1,\boldsymbol{v}_1^{1:p},\ldots,\boldsymbol{v}_r^{1:p})^\top,$$

对于 $j=1,\ \ldots\ldots,\ r,\ \ \boldsymbol{v}_j^{1:p}=(v_j,\ v_j^2,\ \ldots\ldots,\ v_j^p)$，且 v_j是任意变量 $\boldsymbol{v}\in\mathbb{R}^r$的 j^{th}元素。此外，核函数 $\mathcal{K}$ 选为标准正态密度函数。

5.2 数据驱动的调优参数和置信带

调优参数 K 和 h必须满足一些充分条件，即假设9，其中 h在渐近推导中必须以更快的速度减小（即欠平滑；见注释5）。然而，这些条件在实践中并不能给出唯一的选择。本节介绍两种用于实践推断的数据驱动方法。我们首先采用交叉验证（CV）选择筛函数数量和初始带宽，以获得最优的均方误差估计结果。然后，我们借鉴Bissantz 等人(2007年)的思路，介绍一种选择更小带宽、以满足欠平滑条件的方法；同时介绍一种基于 Lepski 方法的替代方案，通过控制偏差来确保置信带有效（例如，参见 Chernozhukov 等人，2014a；Chen 等人，2025）。我们还证明了 Lepski 方法的有效性。

5.2.1 初始参数

具体而言，在第一步中，我们将数据集随机划分为 F 个集合，其中 $2\le F\le N$ 为整数。$\widehat{\ }\quad\widehat{\ }\quad\widehat{\ }\quad\widetilde{\ }$原则上，我们选择 $g(\cdot)$、$K=(K_1+1)$、$\cdot(K_2+1)$ 和 h 的初始调优参数，以使调整后的 F 折交叉验证准则最小化：

$$G(K_1,K_2,h)=\sum_{j=1}^{F}\frac{1}{|S_j|}\sum_{k\in S_j}\widehat{\pi}_K(T_k,\mathbf{X}_k)\mathcal{L}\{Y_k-\widehat{g}_h^{(-j)}(T_k)\}\Big/(1-K/N)^2,$$

对于某个正整数 P 和集合 $\mathcal{H}$，在 $(K_1,\ K_2,\ h)\in\{1,\ldots,P\}^2\times\mathcal{H}$ 上，其中，对于 $j=1,\ \ldots\ldots,\ F,\ S_j$表示数据集的 j^{th} 集合，$|S_j|$是第 j 个集合的样本量，而 $\widehat{g}_h^{(-j)}$的计算方式与 $\widehat{g}_h$ 相同，但不包括 S_j 中的观测值。通过这种方式，我们进行交叉验证以选择 h。为避免因 K过大而导致过拟合，我们将 F 折交叉验证准则乘以因子 $1/(1-K/N)^2$ 进行调整。虽然可以对 K 和 h进行完整的交叉验证，但这种方法计算量大，在某些情况下可能并不实用。在数值研究中，我们取 $F=5$。

5.2.2 方法1（欠平滑）：数据驱动欠平滑带宽

为确保在第二步中获得合适的欠平滑带宽，Bissantz 等人，2007年建议使用略微过平滑的带宽作为试验带宽，即试验$\tilde{}$带宽 $h_0 = \gamma h$，其中 $\gamma > 1$。对于 g 和 τ的推断，我们设定 $\gamma = 1.1$。

在第二步中，我们旨在选择一个正常数 $c < 1$，使得 $h = c \cdot h_0$ 满足欠平滑要求。令 $d_{L_\infty}(g_1, g_2)$ 表示函数 g_1 和 g_2 之间的 L_∞ 距离。根据备注 5 中估计量的偏差和方差的收敛速率，对于足够精细的网格 $\mathbb{G} \in \mathcal{T}$，上确界距离 $d_{L_\infty}(\widehat{g}, g) := \sup_{t\in\mathbb{G}} |\widehat{g}_h(t) - g(t)|$ 在接近最优偏差-方差权衡带宽的带宽下相对稳定。然而，对于偏差相较于方差可忽略不计的欠平滑带宽，这些距离会随着带宽减小而急剧增大。因此，当带宽从略微过度平滑的尺度逐渐减小时，我们应观察到 $d_{L_\infty}(\widehat{g}, g)$ 急剧增大。尽管无法计算 $d_{L_\infty}(\widehat{g}, g)$，Bissantz 等人（2007年）提出，可计算带宽逐渐减小时 $\widehat{g}$ 之间的 L_∞ 距离，以找到这种急剧增大的现象。

具体而言，考虑一个整数 J；对于 $j = 0, \ldots, J-1$，我们使用带宽 $h_j = (J-j)/J \cdot h_0$计算估计量，记为 $\widehat{g}_{h_j}$。接着，对于 $j = 1, \ldots, J-1$，我们计算 $d_{L_\infty}(\widehat{g}_{h_j}, \widehat{g}_{h_{j-1}})$。$\widehat{}$我们选择欠平滑带宽为 $h_u := (J-j)/J \cdot h_0$，取能够在 $d_{L_\infty}(\widehat{g}_{h_j}, \widehat{g}_{h_{j-1}})$ 中观察到突然陡增（转折点）的最小 j；如我们在第7节的实证研究图 2 所示。对于导数 g'，备注 5 指出，其带宽的最优速率大于 f 的带宽速率。我们令 K 与 g 的取值相同，但初始带宽 $h_0 = 4\widetilde{h} \cdot N^{1/5} \cdot N^{-1/7}$。然后，我们对 g 的推断$\widehat{}$执行与上述相同的欠平滑步骤。最终得到的欠平滑带宽记为 h'_u。

正如 Bissantz 等人（2007年）所建议的，我们取 $J = 20$。在模拟研究中，我们考察条件平均值和分位数剂量反应函数，即取 $\mathcal{L}(v) = v^2$ 和$\mathcal{L}(v) = v\{q - \mathbb{1}(v \le 0)\}$，其中 $q \in (0, 1)$。在所有情形下，我们发现显著的陡增大多对应于 $j \in [15, 18]$。为简单起见，我们在整个模拟研究中设定$j = 18$。

5.2.3 方法2（Lepski）：数据驱动置信带

我们将 Lepski 方法（例如，参见 Chernozhukov 等人，2014a；Chen 等人，2025）应用于我们的情境。设 $J_{\max,N}$ 和 $J_{\min,N}$ 为两个整数序列，相应的候选带宽集为 $\mathcal{H}_N := \{h = 2^{-\ell} : \ell \in [J_{\min,N}, J_{\max,N}] \cap \mathbb{Z}\}$。对于任意带宽 $h \in \mathcal{H}_N$，令$D_h(t) = \widehat{g}_h(t) - g(t)$。定义 $\mathcal{V}_N := \{(h, h_2) \in \mathcal{H}_N : h_2 < h\}$，对于任意 h，$h_2 \in \mathcal{V}_N$，有$D_h(t) - D_{h_2}(t) = \widehat{g}_h(t) - \widehat{g}_{h_2}(t)$。此外，$D_h(t) - D_{h_2}(t)$ 的加权自助法版本为

$$D_h^{(b)}(t) - D_{h_2}^{(b)}(t) := \{\widehat{g}_h^{(b)}(t) - \widehat{g}_h(t)\} - \{\widehat{g}_{h_2}^{(b)}(t) - \widehat{g}_{h_2}(t)\},$$

and

$$Z(t, h, h_2) := \frac{D_h^{(b)}(t) - D_{h_2}^{(b)}(t)}{\widehat{\sigma}_{g,N}(t, h, h_2)}$$

是一个随机过程，其中 $\widehat{\sigma}_{g,N}(t, h, h_2)$ 是 $\sigma_{g,N}(t, h, h_2)$ 的相合估计量，即 $D_h^{(b)}(t) - D_{h_2}^{(b)}(t)$ 的截断渐近标准差。具体而言，

$$\sigma_{g,N}^2(t, h, h_2) := \widetilde{\sigma}_{g,N}^2(t, h, h_2) \vee \left\{c_\sigma^2 \sigma_{g,N}^2(t, h_2)\right\},$$

其中

$$\widetilde{\sigma}^2_{g,N}(t,h,h_2) := Var\left\{\frac{1}{Nh}\sum_{i=1}^{N}\varphi_{0,h}(T_i,\boldsymbol{X}_i,Y_i;t) - \frac{1}{Nh_2}\sum_{i=1}^{N}\varphi_{0,h_2}(T_i,\boldsymbol{X}_i,Y_i;t)\right\},$$

小常数 c_σ $\widetilde{\sigma}^2_{g\,N}\ t\ h\ h_2\ >$的 $c^2_\sigma\sigma^2_{g\,N}\ t\ h_2$ $h\ >\ \varpi h_2$ ϖ 选取使得以趋于1的概率(,,) ,(,)对成立，且是一个常数（见补充材料中的引理**??**）。在我们的数值研究中，我们取$\widehat{\sigma}_{g,N}(t,h,h_2)$为$\{D^{(b)}_h(t) - D^{(b)}_{h_2}(t)\}_{b=1}^{B}$ 和 $c_\sigma = 0.25$的样本标准差。

设 γ_N为正数序列且收敛于零。我们定义 $\widetilde{c}_N(\gamma_N)$为在给定观测数据 $\{(T_i,\boldsymbol{X}_i,Y_i):$ $1 \le i \le N\}$的条件下，$\sup_{t\in\mathcal{T},(h,h_2)\in\mathcal{V}_N}|Z(t,h,h_2)|$ 的条件$(1-\gamma_N)$-分位数。令$\mathcal{H}_{N,h}:$ $= \{h_2 \in \mathcal{H}_N:\ h_2 < h\}$，对某个与 N无关的常数 $v > 1$，带宽的Lepski型估计量为

$$\widehat{h} := \sup\left\{h \in \mathcal{H}_N : \sup_{h_2\in\mathcal{H}_{N,h}}\sup_{t\in\mathcal{T}}\left|\frac{\widehat{g}_h(t)-\widehat{g}_{h_2}(t)}{\widehat{\sigma}_{g,N}(t,h,h_2)}\right| \le v\widetilde{c}_N(\gamma_N)\right\}. \tag{5.1}$$

Lepski 型数据驱动置信区间进一步校正置信区间 $I_{g,\ N}(t)\widehat{\ }$中由估计带宽 h引入的偏差。具体而言，最终的 Lepski 型置信区间定义为

$$\tilde{I}_{g,N} := [\widehat{g}_{\widehat{h}}(t) - \{\widehat{c}_N(\alpha) + u_N\widetilde{c}_N(\gamma_N)\}\cdot\widehat{\sigma}_{g,N}(t,\widehat{h}), \widehat{g}_{\widehat{h}}(t) + \{\widehat{c}_N(\alpha) + u_N\widetilde{c}_N(\gamma_N)\}\cdot\widehat{\sigma}_{g,N}(t,\widehat{h})],$$

其中，u_N为某个调整参数，$\widehat{c}_N(\alpha)$ 是在给定观测数据 $\{(T_i,\ \boldsymbol{X}_i,\ Y_i): 1 \le i \le N\}$ 的条件下，$\sup_{t\in\mathcal{T},\ h\in\mathcal{H}_N}\left|\left(\widehat{g}^{(b)}_h(t) - \widehat{g}_h(t)\right)/\widehat{\sigma}_{g,\ N}(t,\ h)\right|$ 的 $(1-\alpha)$ 分位数。

参照 Chernozhukov 等（2014a）和 Chen 等人（2025），我们设定 $J_{\max,N} = \lceil\max(\tilde{J}_N, -\log_2(\tilde{h}_0/10))\rceil$，以及 $J_{\min,N} = \lceil\max(\log(J_{\max,N}) - \log_2(\tilde{h}_0\cdot\log^{1/5}(N)/3))\rceil$，其中$\tilde{J}_N = \log_2(10\cdot N\cdot\tilde{h}_0/\log^4(N))$，$\tilde{h}_0 = \widetilde{h}$ 是 g 的交叉验证带宽。此外，$\gamma_N = \min(0.5, \{\log(J_{\max,N})/J_{\max,N}\}^{1/2})$，$v = 1.1$, $u_N = 0.5$。

可通过类似的方法构造导数 g' 的数据驱动型 Lepski 方法带宽和置信区间：在上述算法中，将 $\widehat{g}_h, \widehat{g}^{(b)}_h$ 分别替换为 $\widehat{g}'_h$ 和 $\widehat{g'}^{(b)}_h$，并对所有 h 充分$\widehat{\ }$执行。我们分别将所得带宽和置信区间记为 h' 和 $\tilde{I}_{g',N}$。此外，在数值研究中，将 $J_{\max,N}$ 中的 $\tilde{h}_0$ 替换为 $3\widetilde{h}\cdot N^{1/5}\cdot N^{-1/7}$。所得候选带宽集记为 $\mathcal{H}'_N$。

5.3 Lepski 方法的有效性

我们定义在带宽 h 下 $g(h)$ 和 $g'(t)$ 的近似为

$$(g_h(t), g'_h(t)) := \underset{(\theta_1,\theta_2)\in\mathbb{R}^2}{\operatorname{argmin}}\ \mathbb{E}\left[\pi_0(T,\boldsymbol{X})\mathcal{L}\left\{Y - \theta_1 - \theta_2(T-t)\right\}\mathcal{K}_h(T-t)\right].$$

假设12。假设 g 属于函数类 $\mathcal{G}$。存在常数 $0 < h_0$、c_3、C_3、s、$\bar{s} < \infty$，使得对每个 $g \in \mathcal{G}$，都有

$$c_3h^2 \le \sup_{t\in\mathcal{T}}|g_h(t) - g(t)| \le C_3h^2 \text{ and } \sup_{t\in\mathcal{T}}|g'_h(t) - g'(t)| \le C_3h^2$$

对所有 $h \le h_0$。

假设13。 存在常数 c_4、 $C_4 > 0$ ， 使得对于每个 $g \in \mathcal{G}$ ， 都存在 $h \in \mathcal{H}_N$和$h' \in \mathcal{H}'_N$ ， 满足

$$\left(c_4 \frac{\log N}{N}\right)^{1/5} \leq h \leq \left(C_4 \frac{\log N}{N}\right)^{1/5},$$

and

$$\left(c_4 \frac{\log N}{N}\right)^{1/7} \leq h' \leq \left(C_4 \frac{\log N}{N}\right)^{1/7}.$$

假设14. 对每一个 $b = 0$，……， B，均有

$$\sqrt{Nh}\left\{\widehat{g}_h^{(b)}(t) - g_h(t)\right\} = \frac{1}{\sqrt{Nh}} \sum_{i=1}^{N} \xi_i^{(b)} \psi_{0,h}(T_i, \boldsymbol{X}_i, Y_i; t) + \widetilde{R}_{0,N}(t, h, g)$$

and

$$\sqrt{Nh^3}\left\{\widehat{g}'^{(b)}_{h'}(t) - g'_{h'}(t)\right\} = \frac{1}{\sqrt{N(h')^3}} \sum_{i=1}^{N} \xi_i^{(b)} \psi_{1,h'}(T_i, \boldsymbol{X}_i, Y_i; t) + \widetilde{R}_{1,N}(t, h', g)$$

其中 $\xi_i^{(0)} \equiv 1$与$\widetilde{R}_{0,N}(t, h, g)$ 以及$\widetilde{R}_{1,N}(t, h', g)$ 满足

$$\sup_{g \in \mathcal{G}} \mathbb{P}\left(\sup_{t \in \mathcal{T}, h \in \mathcal{H}_N, h' \in \mathcal{H}'_N} \left\{\left|\widetilde{R}_{0,N}(t, h, g)\right| \vee \left|\widetilde{R}_{1,N}(t, h', g)\right|\right\} \geq o((\log N)^{-1/2})\right) = o(1)$$

当 $N \to \infty$时。

假设15. 设$\mathcal{V}_N := \{(t, h, h_2) : t \in \mathcal{T}, h, h_2 \in \mathcal{H}_N, h_2 < h\}$, $\mathcal{V}'_N := \{(t, h, h_2) : t \in \mathcal{T}, h, h_2 \in \mathcal{H}'_N, h_2 < h\}$，且$\delta_{\sigma,N}$为一列满足$\delta_{\sigma,N} = o((\log N)^{-1})$的非负数。则有

$$\sup_{g \in \mathcal{G}} \mathbb{P}\left(\sup_{t \in \mathcal{T}, h \in \mathcal{H}_N} \left|\frac{\widehat{\sigma}_{g,N}(t, h)}{\sigma_{g,N}(t, h)} - 1\right| \geq \delta_{\sigma,N}\right) = 1 - o(1),$$

$$\sup_{g \in \mathcal{G}} \mathbb{P}\left(\sup_{t \in \mathcal{T}, h \in \mathcal{H}'_N} \left|\frac{\widehat{\sigma}'_{g,N}(t, h)}{\sigma'_{g,N}(t, h)} - 1\right| \geq \delta_{\sigma,N}\right) = 1 - o(1),$$

$$\sup_{g \in \mathcal{G}} \mathbb{P}\left(\sup_{(t,h,h_2) \in \overline{\mathcal{V}}_N} \left|\frac{\widehat{\sigma}_{g,N}(t, h, h_2)}{\sigma_{g,N}(t, h, h_2)} - 1\right| \geq o((\log N)^{-1})\right) = 1 - o(1),$$

and

$$\sup_{g \in \mathcal{G}} \mathbb{P}\left(\sup_{(t,h,h_2) \in \overline{\mathcal{V}}'_N} \left|\frac{\widehat{\sigma}_{g',N}(t, h, h_2)}{\sigma_{g',N}(t, h, h_2)} - 1\right| \geq o((\log N)^{-1})\right) = 1 - o(1)$$

当 $N \to \infty$时。

假设16。 存在某个常数 $c > 0$ ， 使得

$$c^2 \leq \sigma^2_{Y|T,\boldsymbol{X}} := Var\left(\mathcal{L}'(Y_i - g(t) - g'(t)(T_i - t)) \mid T, \boldsymbol{X}\right).$$

假设17。 函数类 $\{t \mapsto \mathcal{K}((t-t')/h) : t' \in \mathcal{T}, h \in \mathcal{H}_N\}$ 属于 $VC(b, a, v)$ 类型，其中常数 b, a 和 v 与 N 无关；*VC* 类型类的定义见 *Chernozhukov 等人 (2014a, 第1801页)*。

假设12对逼近误差施加了一些限制，这是一个较弱的假设，类似于 Chernozhukov 等人(2014a)中的条件L2。假设13 可由 $\mathcal{H}_N$ 和 $\mathcal{H}'_N$ 的构造满足。假设14 可用与定理1类似的论证来证明。假设15 与假设10类似。假设17是一个温和的假设，参见Gin´e 和 Guillou(2002，第911页）。我们有以下定理。

定理3。假设假设1-4、7-8、11-17成立。此外，我们假设 $(\log N)^4 / \{N(h_{\min,N} \wedge (h'_{\min,N})^3)\} = o(N^{-c_5})$, $h_{\max,N}, h'_{\max,N} \to 0$, $h_{\min N}/h_{\max N}$ $h'_{\min N}/h'_{\max N} \to 0$ $\gamma_N \to 0$ | $\gamma_N| = O$ N $\delta_{\sigma N} u_N = o$ N^{-1} , , , , , , , log (log) and , ((log)), where $h_{\min,N} := \inf \mathcal{H}_N$, $h_{\max,N} := \sup \mathcal{H}_N$, $h'_{\min,N} := \inf \mathcal{H}'_N$, $h'_{\max,N} := \sup \mathcal{H}'_N$ and $c_5 > 0$ is a constant. Then it holds that (i)

$$\liminf_{N\to\infty} \inf_{g\in\mathcal{G}} \mathbb{P}\left(g(t) \in \tilde{I}_{g,N}(t),\ \forall t \in \mathcal{T}\right) \geq 1-\alpha$$

以及（*ii*）

$$\liminf_{N\to\infty} \inf_{g\in\mathcal{G}} \mathbb{P}\left(g'(t) \in \tilde{I}_{g',N}(t),\ \forall t \in \mathcal{T}\right) \geq 1-\alpha.$$

6 数值研究

6.1 模拟研究

我们针对连续处理模型开展了一项模拟研究，以评估⌢提出的估计量 $\hat{g}$ 和 $\partial_t g$的有限样本性能。我们考虑三个响应函数。具体而言，令U_w、 U_t 和 U_y 为三个相互独立的标准正态随机变量。这些模型表示为$\boldsymbol{X}$ =（Z， W）$^\top$，其中

$$Z \sim \text{均匀分布}(-0.65, 0.65) \quad W = 0.5Z + 0.5U_w, \quad T = 0.1W + 0.1Z + 0.4U_t.$$
$$\textbf{DGP0}: Y^*(t) = \exp(0.2Z - 0.5W) + 0.1U_y, \quad \textbf{DGP1L}: Y^*(t) = 0.6t + 0.7Z - 2W + 0.1U_y \text{（关于 } t \text{呈线性）}, \quad \textbf{DGP1NL}: Y^*(t) = 0.75 \exp(t) + 0.6Z^3 - 3W + 0.1U_y \text{（关于 } t \text{呈非线性）}。$$

我们生成 $J = 500$ 个样本，样本量分别为 $N = 400$、800 和 1200。对于 $\mathcal{L}(v) = v^2$ 和$\mathcal{L}(v) = v\{q - \mathbb{1}(v \leq 0)\}$，分别取 $q = 0.25$、0.35 和 0.45 时，我们在包含 25 个等距点的网格上计算条件平均响应 $g(t) = \mathbb{E}\{Y^*(t)\}$ 和条件分位数响应 $g(t) = F^{-1}_{Y^*(t)}(q)$的估计量 $\widehat{g}$ 和 $\partial_t g$，其中 $\mathcal{V}$的范围为 T的第 5 至第 95 百分位数。请注意，对于条件平均剂量-反应，我们的估计量 $\hat{g}$和 g' 有闭式解，但对于条件分位数剂量-反应函数则没有。对于任意分位数水平 q，估计 $g(t) = F^{-1}_{Y^*(t)}(q)$所需的数值最小化耗时较长。为加快计算，我们采用迭代加权最小二乘

表1：使用方法1（欠平滑）通过500次蒙特卡洛模拟计算得到的 g 置信带的经验覆盖概率（平均宽度）

	N	方法	DGP0 99%	DGP0 95%	DGP0 90%	DGP1L 99%	DGP1L 95%	DGP1L 90%	DGP1NL 99%	DGP1NL 95%	DGP1NL 90%
平均值	400	提出的	0.970 (0.7)	0.898 (0.5)	0.842 (0.4)	0.978 (4.1)	0.926 (2.8)	0.868 (2.4)	0.968 (4.2)	0.904 (2.9)	0.860 (2.5)
		朴素	0.944 (0.7)	0.830 (0.5)	0.734 (0.4)	0.958 (3.9)	0.836 (2.7)	0.732 (2.4)	0.944 (3.9)	0.782 (2.7)	0.684 (2.4)
	800	提出的	0.978 (0.4)	0.940 (0.3)	0.868 (0.3)	0.978 (1.9)	0.938 (1.6)	0.872 (1.4)	0.978 (1.9)	0.920 (1.6)	0.846 (1.5)
		朴素	0.922 (0.4)	0.776 (0.3)	0.648 (0.3)	0.896 (2.0)	0.748 (1.6)	0.624 (1.5)	0.864 (2.0)	0.670 (1.6)	0.502 (1.5)
	1200	提出的	0.978 (0.2)	0.926 (0.2)	0.854 (0.1)	0.958 (0.6)	0.888 (0.5)	0.818 (0.5)	0.968 (1.1)	0.900 (0.9)	0.820 (0.8)
		朴素	0.720 (0.2)	0.460 (0.2)	0.308 (0.1)	0.656 (0.7)	0.358 (0.6)	0.222 (0.5)	0.540 (1.1)	0.278 (0.9)	0.174 (0.9)
$q = 0.45$	400	提出的	1.000 (1.7)	0.976 (1.1)	0.934 (0.9)	0.992 (5.1)	0.950 (3.2)	0.906 (2.7)	0.992 (8.4)	0.968 (5.4)	0.940 (4.4)
		朴素	0.994 (1.5)	0.962 (1.0)	0.914 (0.8)	0.984 (5.3)	0.926 (3.3)	0.870 (2.7)	0.986 (7.7)	0.928 (5.0)	0.842 (4.1)
	800	提出的	0.996 (0.6)	0.968 (0.4)	0.940 (0.4)	0.996 (2.2)	0.972 (1.7)	0.930 (1.5)	0.998 (3.4)	0.964 (2.5)	0.916 (2.2)
		朴素	0.978 (0.6)	0.900 (0.4)	0.842 (0.4)	0.974 (2.0)	0.892 (1.6)	0.784 (1.4)	0.964 (3.2)	0.826 (2.4)	0.704 (2.2)
	1200	提出的	0.990 (0.3)	0.932 (0.2)	0.856（0.2）	0.974（1.0）	0.894（0.8）	0.806（0.7）	0.988（1.6）	0.906（1.3）	0.838 (1.2)
		朴素	0.908 (0.3)	0.684 (0.2)	0.566 (0.2)	0.826 (1.0)	0.560 (0.8)	0.410 (0.7)	0.748 (1.7)	0.472 (1.3)	0.346 (1.2)
$q = 0.35$	400	提出的	0.996 (1.5)	0.982 (0.9)	0.954 (0.8)	0.992 (6.1)	0.958 (3.8)	0.902 (3.1)	1.000 (8.6)	0.976 (5.5)	0.910 (4.5)
		朴素	0.994 (1.3)	0.962 (0.9)	0.914 (0.7)	0.980 (5.1)	0.930 (3.3)	0.866 (2.8)	0.992 (8.9)	0.924 (5.6)	0.840 (4.6)
	800	提出的	0.994 (0.6)	0.964 (0.5)	0.918 (0.4)	0.986 (2.3)	0.932 (1.7)	0.878 (1.5)	0.978 (3.4)	0.944 (2.6)	0.876 (2.3)
		朴素	0.978 (0.6)	0.908 (0.5)	0.836 (0.4)	0.970 (2.2)	0.876 (1.7)	0.772 (1.5)	0.958 (3.7)	0.842 (2.8)	0.714 (2.4)
	1200	提出的	0.988 (0.3)	0.936 (0.2)	0.860 (0.2)	0.968 (1.0)	0.888 (0.8)	0.828 (0.8)	0.970 (1.9)	0.898 (1.5)	0.804 (1.3)
		朴素	0.894 (0.3)	0.676 (0.2)	0.528 (0.2)	0.812 (1.1)	0.606 (0.9)	0.482 (0.8)	0.786 (1.8)	0.552 (1.4)	0.390 (1.3)
$q = 0.25$	400	提出的	0.994 (1.5)	0.960 (1.0)	0.906 (0.8)	0.988 (5.7)	0.920 (3.6)	0.868 (3.0)	0.988 (8.1)	0.932 (5.3)	0.862 (4.4)
		朴素	0.988 (1.4)	0.944 (0.9)	0.906 (0.7)	0.976 (5.3)	0.898 (3.4)	0.816 (2.9)	0.980 (8.3)	0.884 (5.4)	0.790 (4.5)
	800	提出的	0.978 (0.6)	0.904 (0.4)	0.830 (0.4)	0.976 (2.4)	0.904 (1.8)	0.824 (1.6)	0.968 (3.7)	0.886 (2.8)	0.812 (2.5)
		朴素	0.966 (0.6)	0.866 (0.4)	0.760 (0.4)	0.936 (2.4)	0.800 (1.8)	0.710 (1.6)	0.936 (3.6)	0.768 (2.8)	0.664 (2.5)
	1200	提出的	0.982 (0.3)	0.912 (0.2)	0.830 (0.2)	0.948 (1.2)	0.860 (0.9)	0.766 (0.8)	0.950 (1.9)	0.842 (1.5)	0.752 (1.4)
		朴素	0.902 (0.3)	0.730 (0.2)	0.610 (0.2)	0.806 (1.1)	0.578 (0.9)	0.438 (0.8)	0.782 (1.9)	0.536 (1.5)	0.406 (1.3)

算法，由Lejeune and Sarda (1988提出。在自助法程序中，我们将自助法重复次数设为 $B = 500$。

我们计算朴素估计量，$\widehat{g}_{\mathrm{N}}$ 和 $\widehat{\partial_t g}_{\mathrm{N}}$，通过设定 $\widehat{\pi}_K \equiv 1$。表 1 和表 2 报告了第 5 节中为 g 从 DGP0、DGP1L 和 DGP1NL 引入的两种自适应置信带的经验覆盖概率及其平均宽度，置信水平分别为 0.99、0.95 和 0.90。此外，表 3 和表 4 展示了在显著性水平 0.01、0.05 和 0.1 下，式 (2.6) 中原假设 H_0 的经验拒绝概率。需要注意的是，尽管原假设 H_0 在 DGP0 下成立，但在 DGP1L 和 DGP1NL 下并不成立。这些表格共同证实了我们方法中混杂调整的有效性，这体现在一致推断具有良好的检验水平和检验功效，同时置信带的宽度也相当。总体而言，与欠平滑方法相比，Lepski 方法得到的置信带更保守，但更窄。

为了更深入地了解一致推断结果，我们在表格中报告了网格上的平均平方偏差和平均方差 $\mathcal{V}$ 以及 500 个模拟样本中欠平滑估计量的相应结果。5和6。这些模拟结果表明，与朴素估计量相比，我们提出的估计量始终表现出显著更小的偏差，且方差相当。具体而言，与朴素估计量不同，我们的估计量的偏差相对于其方差可以忽略不计。此外，随着分位数趋近中位数，方差有所降低；而随着样本量或分位数水平增加，所提出方法的偏差和方差均持续减小。

为直观展示模拟结果，图1 展示了从模型 g 中估计 **DGP1L** 和 g' 的 **DGP1NL**时，采用提出的方法和朴素方法得到的一致90%置信带的若干示例。这些图表明，提出的方法所估计的 g 和 g' 与真实曲线高度吻合，并且

表2：使用方法2（Lepski）通过500次蒙特卡洛模拟计算得到的 g 置信带的经验覆盖概率（平均宽度）

	N	方法	DGP0 99%	DGP0 95%	DGP0 90%	DGP1L 99%	DGP1L 95%	DGP1L 90%	DGP1NL 99%	DGP1NL 95%	DGP1NL 90%
平均值	400	提出的	0.998（0.2）	0.986（0.2）	0.980（0.2）	0.986（0.9）	0.956（0.8）	0.922（0.8）	0.984 (0.9)	0.944 (0.8)	0.914 (0.8)
		朴素	0.992 (0.3)	0.960 (0.2)	0.914 (0.2)	0.984 (1.3)	0.954 (1.2)	0.914 (1.1)	0.970 (1.3)	0.916 (1.2)	0.838 (1.1)
	800	提出的	0.996 (0.1)	0.982 (0.1)	0.976 (0.1)	0.976 (0.7)	0.954 (0.6)	0.922 (0.6)	0.976 (0.7)	0.940 (0.6)	0.912 (0.6)
		朴素	0.958 (0.2)	0.892 (0.2)	0.834 (0.1)	0.934 (1.0)	0.820 (0.9)	0.730 (0.8)	0.864 (1.0)	0.690 (0.9)	0.568 (0.8)
	1200	提出的	0.992 (0.1)	0.974 (0.1)	0.952 (0.1)	0.978 (0.3)	0.936 (0.3)	0.910 (0.3)	0.960 (0.5)	0.924 (0.5)	0.884 (0.4)
		朴素	0.846 (0.1)	0.656 (0.1)	0.538 (0.1)	0.738 (0.5)	0.532 (0.4)	0.406 (0.4)	0.596 (0.8)	0.368 (0.7)	0.236 (0.6)
$q = 0.45$	400	提出的	0.996 (0.3)	0.982 (0.2)	0.974 (0.2)	0.988 (0.9)	0.972 (0.8)	0.964 (0.7)	0.976 (1.3)	0.960 (1.1)	0.938 (1.1)
		朴素	0.998 (0.3)	0.984 (0.3)	0.974 (0.3)	0.996 (1.1)	0.970 (1.0)	0.942 (0.9)	0.992 (1.7)	0.964 (1.5)	0.922 (1.4)
	800	提出的	1.000 (0.2)	0.998 (0.2)	0.994 (0.2)	0.998 (0.7)	0.988 (0.6)	0.984 (0.5)	0.994 (1.0)	0.990 (0.9)	0.978 (0.8)
		朴素	0.996 (0.2)	0.968 (0.2)	0.950 (0.2)	0.972 (0.8)	0.940 (0.7)	0.890 (0.7)	0.958 (1.2)	0.868 (1.1)	0.800 (1.0)
	1200	提出的	1.000 (0.1)	0.998 (0.1)	0.992 (0.1)	0.994 (0.5)	0.978 (0.4)	0.948 (0.4)	0.986 (0.8)	0.954 (0.7)	0.934 (0.6)
		朴素	0.958 (0.2)	0.870 (0.1)	0.796 (0.1)	0.894 (0.6)	0.786 (0.5)	0.686 (0.5)	0.840 (0.9)	0.658 (0.8)	0.504 (0.8)
$q = 0.35$	400	提出的	1.000 (0.3)	0.994 (0.2)	0.994 (0.2)	0.998 (1.1)	0.994 (0.9)	0.990 (0.9)	0.996 (1.6)	0.984 (1.4)	0.980 (1.3)
		朴素	0.994 (0.3)	0.978 (0.3)	0.970 (0.3)	0.994 (1.3)	0.970 (1.1)	0.956 (1.0)	0.986 (1.9)	0.964 (1.7)	0.936 (1.5)
	800	提出的	1.000 (0.2)	0.994 (0.2)	0.992 (0.2)	1.000 (0.7)	0.988 (0.6)	0.980 (0.6)	1.000 (1.1)	0.980 (1.0)	0.964 (0.9)
		朴素	0.986 (0.2)	0.960 (0.2)	0.938 (0.2)	0.970 (0.9)	0.940 (0.8)	0.908 (0.7)	0.946 (1.3)	0.880 (1.2)	0.816 (1.1)
	1200	提出的	1.000 (0.1)	0.992 (0.1)	0.984 (0.1)	0.998 (0.5)	0.980 (0.5)	0.964 (0.4)	0.990 (0.8)	0.966 (0.7)	0.936 (0.7)
		朴素	0.948 (0.2)	0.888 (0.1)	0.818 (0.1)	0.920 (0.7)	0.790 (0.6)	0.726 (0.5)	0.850 (1.0)	0.700 (0.9)	0.586 (0.8)
$q = 0.25$	400	提出的	0.998 (0.3)	0.992 (0.3)	0.988 (0.2)	0.996 (1.2)	0.988 (1.1)	0.980 (1.0)	1.000 (1.9)	0.992 (1.6)	0.978 (1.5)
		朴素	0.994 (0.3)	0.980 (0.3)	0.966 (0.3)	0.996 (1.4)	0.974 (1.2)	0.970 (1.1)	0.980 (2.1)	0.950 (1.8)	0.926 (1.7)
	800	提出的	0.998 (0.2)	0.992 (0.2)	0.986 (0.2)	1.000 (0.9)	0.988 (0.7)	0.974 (0.7)	0.996 (1.3)	0.976 (1.1)	0.964 (1.1)
		朴素	0.990（0.2）	0.966（0.2）	0.942（0.2）	0.972（1.0）	0.930（0.8）	0.900（0.8）	0.960 (1.5)	0.904 (1.3)	0.834 (1.2)
	1200	提出的	1.000 (0.2)	0.990 (0.1)	0.980 (0.1)	0.994 (0.6)	0.976 (0.5)	0.966 (0.5)	0.986 (1.0)	0.946 (0.9)	0.920 (0.8)
		朴素	0.960 (0.2)	0.900 (0.1)	0.854 (0.1)	0.912 (0.7)	0.800 (0.6)	0.728 (0.6)	0.862 (1.1)	0.712 (0.9)	0.626 (0.9)

一致置信带能够有效覆盖目标。相反，朴素方法会因混杂问题引入明显偏差，导致置信带未能覆盖目标。

7 实证研究

在本节中，我们应用提出的估计方法，分析一组来自美国总统竞选的数据集，以了解播放的政治广告数量如何对非竞争州的竞选捐款产生因果影响。该数据集常用于连续处理效应文献，涵盖非竞争州邮政编码中从 0 到 22379 的范围（参见例如 $N = 16265$ 邮政编码Urban 和 Niebler，2014年；Fong 等人，2018；Ai 等人，2021年；Huang 等人，2022年）。

所考虑的协变量包括 $\boldsymbol{X}$ log(人口)、年龄超过 65 岁的人口占比、log(家庭收入中位数 + 1)、西班牙裔人口占比、黑人口占比、log(人口密度 + 1)、大学毕业生占比，以及该地区是否可以通勤至竞争州的指标。更多˜信息可参见Fong 等人（2018）。我们将协变量标准化为 $\boldsymbol{X} = \{\boldsymbol{X} - \min(\boldsymbol{X})\}/\{\max(\boldsymbol{X}) - \min(\boldsymbol{X})\}$。

Urban 和 Niebler (2014年) 使用二元模型分析了该数据中广告与竞选捐款之间的因果关系。他们比较了收到超过 1000 条广告的 5230 个邮政编码地区的竞选捐款，以及其余收到少于 1000 条广告的 11035 个邮政编码地区的竞选捐款。他们的分析显示，在非竞争州，广告对竞选捐款金额有显著的因果效应

表3：检验H_0的经验拒绝概率（所提出的一致置信带的平均宽度）： $g'=0$，使用方法1（欠平滑）通过500次蒙特卡洛模拟计算得出

	N	方法	DGP0 99%	DGP0 95%	DGP0 90%	DGP1L 99%	DGP1L 95%	DGP1L 90%	DGP1NL 99%	DGP1NL 95%	DGP1NL 90%
平均值	400	所提出的	0.002 (1.2)	0.050 (1.0)	0.110 (0.9)	0.088 (6.7)	0.228 (5.6)	0.396 (5.0)	0.172 (7.1)	0.420 (5.9)	0.608 (5.3)
		朴素	0.038 (1.2)	0.160 (1.0)	0.288 (0.9)	0.018 (6.5)	0.062 (5.4)	0.116 (4.9)	0.032 (6.6)	0.108 (5.5)	0.194 (5.0)
	800	提出的	0.010 (0.8)	0.032 (0.7)	0.086 (0.6)	0.154 (4.5)	0.398 (3.8)	0.538 (3.5)	0.358 (4.6)	0.676 (3.9)	0.808 (3.6)
		朴素	0.066 (0.8)	0.212 (0.7)	0.384 (0.6)	0.012 (4.6)	0.054 (3.9)	0.114 (3.6)	0.040 (4.7)	0.124 (4.0)	0.208 (3.6)
	1200	提出的	0.006 (0.4)	0.046 (0.3)	0.114（0.3）	0.930（1.2）	0.956（1.0）	0.962（0.9）	0.888（2.2）	0.924（1.9）	0.956 (1.7)
		朴素	0.696 (0.4)	0.850 (0.3)	0.910 (0.3)	0.524 (1.2)	0.734 (1.0)	0.814 (0.9)	0.198 (2.1)	0.412 (1.8)	0.544 (1.6)
$q=0.45$	400	提出的	0.002 (2.1)	0.016 (1.6)	0.068 (1.5)	0.090 (7.4)	0.288 (5.9)	0.442 (5.3)	0.046（11.6）	0.220（9.2）	0.366（8.3）
		朴素	0.010（2.0）	0.068（1.6）	0.124 (1.4)	0.022 (7.0)	0.080 (5.5)	0.150 (4.9)	0.016 (10.0)	0.046 (8.0)	0.090 (7.2)
	800	提出的	0.002 (1.2)	0.030 (1.0)	0.056 (0.9)	0.156 (4.9)	0.440 (4.0)	0.620 (3.6)	0.122 (7.5)	0.376 (6.2)	0.510 (5.6)
		朴素	0.024 (1.2)	0.104 (1.0)	0.190 (0.9)	0.022 (4.2)	0.096 (3.5)	0.168 (3.1)	0.016 (6.9)	0.054 (5.8)	0.122 (5.2)
	1200	提出的	0.010 (0.5)	0.062 (0.4)	0.112 (0.4)	0.848 (1.7)	0.900 (1.4)	0.918 (1.3)	0.768 (3.2)	0.862 (2.6)	0.894 (2.4)
		朴素	0.400 (0.6)	0.632 (0.5)	0.750 (0.4)	0.286 (1.9)	0.514 (1.6)	0.654 (1.4)	0.090 (3.2)	0.216 (2.6)	0.336 (2.4)
$q=0.35$	400	提出的	0.002 (2.0)	0.022 (1.6)	0.042 (1.4)	0.064 (8.5)	0.172 (6.7)	0.330 (6.0)	0.050 (12.1)	0.178 (9.6)	0.282 (8.6)
		朴素	0.020 (1.8)	0.070 (1.4)	0.106 (1.3)	0.018 (7.2)	0.058 (5.8)	0.118 (5.2)	0.006 (11.0)	0.036 (8.9)	0.084 (7.9)
	800	提出的	0.010 (1.3)	0.028 (1.1)	0.064 (1.0)	0.130 (4.7)	0.388 (3.9)	0.572 (3.5)	0.124 (7.5)	0.340 (6.2)	0.506 (5.6)
		朴素	0.032 (1.1)	0.112 (0.9)	0.188 (0.8)	0.014 (4.7)	0.102 (3.9)	0.172 (3.5)	0.016 (7.6)	0.064 (6.3)	0.112 (5.7)
	1200	提出的	0.012 (0.5)	0.062 (0.4)	0.098 (0.4)	0.834 (1.8)	0.900 (1.5)	0.936 (1.4)	0.746 (3.5)	0.862 (2.9)	0.900 (2.6)
		朴素	0.380 (0.5)	0.644 (0.4)	0.758 (0.3)	0.234 (1.9)	0.456 (1.6)	0.600 (1.4)	0.080 (3.3)	0.226 (2.7)	0.332 (2.5
$q=0.25$	400	提出的	0.002 (2.2)	0.018 (1.7)	0.044 (1.5)	0.036 (8.7)	0.142 (6.9)	0.268 (6.1)	0.042 (12.5)	0.138 (10.1)	0.244 (9.0)
		朴素	0.010 (1.9)	0.054 (1.5)	0.112 (1.4)	0.004 (7.8)	0.050 (6.3)	0.098 (5.6)	0.000 (12.1)	0.052 (9.6)	0.084 (8.6)
	800	提出的	0.004 (1.2)	0.034 (1.0)	0.082 (0.9)	0.104 (5.2)	0.306 (4.3)	0.482 (3.8)	0.076 (8.4)	0.280 (6.9)	0.436 (6.2)
		朴素	0.024 (1.2)	0.112 (1.0)	0.176 (0.9)	0.018 (5.1)	0.078 (4.2)	0.138 (3.8)	0.008 (7.9)	0.044 (6.6)	0.106 (5.9)
	1200	提出的	0.010 (0.6)	0.044 (0.5)	0.088 (0.4)	0.822 (2.4)	0.878 (2.0)	0.928 (1.8)	0.746 (3.7)	0.850 (3.0)	0.896 (2.7)
		朴素	0.276 (0.5)	0.568 (0.4)	0.696 (0.3)	0.192 (1.9)	0.432 (1.5)	0.582 (1.4)	0.080 (3.3)	0.192 (2.7)	0.302 (2.5)

。

相比之下，Ai 等人 (2021年) 将政治广告的数量视为连续变量，并假设平均响应函数服从线性模型。他们对结果变量和处理变量分别进行了对数变换：$\log(\text{竞选捐款}+1)$ 和 $\log(\#\text{广告}+1)$，其中 # 广告表示广告数量。他们的分析发现，广告对竞选捐款没有显著的因果效应，这与Fong 等人 (2018) 的发现一致。

Huang 等人 (2022年) 提出了一套用于连续处理效应模型设定检验的统一框架，并拒绝了Ai 等人 (2021年) 所假设的线性模型。他们建议采用 Tobit模型，并结合观测捐款的 Box-Cox变换和 #ads 的复合对数变换，以更好地拟合数据。具体而言，他们的 Box-Cox 变换定义为 $\text{BoxCox}(\text{捐款}\ \lambda_1,\ \lambda_2)=\{(\text{捐款}+\lambda_2)^{\lambda_1}-1\}/\tilde{\lambda_1}$ 其中 $(\lambda_1,\ \lambda_2)=(0.1397,\ 0.0176)$。随后，他们考察了观测结果变量，

$$Y^*(T)=Y=\text{BoxCox}(\text{Contribution},\tilde{\lambda}_1,\tilde{\lambda}_2)-\min\left\{\text{BoxCox}(\text{Contribution},\tilde{\lambda}_1,\tilde{\lambda}_2)\right\},$$

以及处理变量，$T=\log(\log(\log(\#\text{广告}+1)\ +1)\ +2)$。他们估计的 Tobit模型显示，当 $\#$ 广告 $\in[0,20]$ 时，竞选捐款会迅速增加，此后增幅则变得有限。然而，Huang 等人 (2022年) 并未给出该研究的任何推断结果。

这些先前研究中相互矛盾的发现表明，广告与竞选捐款之间存在复杂的因果关系，而参数模型可能无法捕捉这种关系。对于 Y，我们采用与 Huang 等人 (2022年)相同的 Box-Cox变换，但取 $\tilde{T}=\{T-\min(T)\}/\{\max(T)-\min(T)\}$，其中 $T=\log(\#\text{ads}+1)$。随后，应用提出的非参数方法，我们估计分位数剂量反应函数（QDRF）$g(t)=F^{-1}_{Y^*(t)}(q)$和偏导数 $\partial_t g$，并构造 95% 一致置信带，其中 $q=0.5$ and0.75。

表4：使用方法2（Lepski）基于500次蒙特卡洛模拟计算的检验H_0： $g' = 0$ 的经验拒绝概率（一致置信带的平均宽度）

			DGP0			DGP1L			DGP1NL		
	N	方法	99%	95%	90%	99%	95%	90%	99%	95%	90%
平均值	400	提出的	0.000 (0.5)	0.000 (0.5)	0.006 (0.4)	0.622 (2.8)	0.712 (2.5)	0.746 (2.3)	0.692 (3.1)	0.766 (2.7)	0.814 (2.5)
		朴素	0.048 (0.6)	0.112 (0.5)	0.202 (0.5)	0.000 (3.8)	0.006 (3.3)	0.012 (3.0)	0.004 (3.8)	0.024 (3.3)	0.046 (3.0)
	800	提出的	0.000 (0.4)	0.002 (0.4)	0.004 (0.4)	0.698 (2.3)	0.748（2.0）	0.778（2.0）	0.756（2.6）	0.806（2.3）	0.836（2.1）
		朴素	0.160（0.5）	0.296（0.4）	0.378（0.4）	0.000（2.9）	0.006（2.5）	0.010（2.4）	0.010 (3.1)	0.032 (2.7)	0.058 (2.6)
	1200	提出的	0.000 (0.3)	0.004 (0.3)	0.016 (0.3)	0.826 (1.1)	0.872 (1.0)	0.906 (0.9)	0.804 (2.0)	0.852 (1.8)	0.880 (1.7)
		朴素	0.470 (0.4)	0.626 (0.3)	0.706 (0.3)	0.238 (1.4)	0.428 (1.2)	0.524 (1.1)	0.032 (2.2)	0.086 (2.0)	0.158 (1.8)
$q = 0.45$	400	提出的	0.002 (0.6)	0.002 (0.5)	0.006 (0.4)	0.620 (2.0)	0.712 (1.7)	0.772 (1.6)	0.508 (3.2)	0.644 (2.8)	0.734 (2.5)
		朴素	0.012 (0.6)	0.058 (0.5)	0.128 (0.5)	0.014 (2.3)	0.052 (2.0)	0.078 (1.8)	0.002 (3.7)	0.016 (3.2)	0.026 (3.0)
	800	提出的	0.000 (0.5)	0.000 (0.4)	0.002 (0.4)	0.766 (1.7)	0.838 (1.5)	0.894 (1.3)	0.692 (2.7)	0.776 (2.4)	0.836 (2.2)
		朴素	0.060 (0.5)	0.138 (0.4)	0.218 (0.4)	0.024 (1.8)	0.080 (1.6)	0.150 (1.4)	0.004 (3.3)	0.010 (2.9)	0.034 (2.7)
	1200	提出的	0.000 (0.4)	0.002 (0.3)	0.006 (0.3)	0.858 (1.2)	0.894 (1.0)	0.910 (0.9)	0.818 (1.9)	0.872 (1.6)	0.896 (1.5)
		朴素	0.252 (0.4)	0.472 (0.3)	0.570 (0.3)	0.124 (1.4)	0.306 (1.2)	0.414 (1.1)	0.018 (2.1)	0.062 (1.9)	0.086 (1.7)
$q = 0.35$	400	提出的	0.000 (0.6)	0.002 (0.5)	0.004 (0.5)	0.458 (2.5)	0.618 (2.1)	0.688 (1.9)	0.296 (3.7)	0.516 (3.1)	0.608 (2.9)
		朴素	0.010 (0.6)	0.060 (0.5)	0.108 (0.5)	0.008 (2.5)	0.032 (2.1)	0.052 (2.0)	0.000 (4.0)	0.012 (3.4)	0.022 (3.2)
	800	提出的	0.000 (0.6)	0.000 (0.5)	0.004 (0.4)	0.692 (1.8)	0.782 (1.6)	0.820 (1.4)	0.632 (2.9)	0.738 (2.5)	0.794 (2.3)
		朴素	0.052 (0.5)	0.146 (0.5)	0.232 (0.4)	0.016 (2.0)	0.076 (1.7)	0.120 (1.6)	0.002 (3.2)	0.016 (2.8)	0.032 (2.6)
	1200	提出的	0.000 (0.3)	0.002 (0.3)	0.006 (0.3)	0.852 (1.2)	0.898 (1.0)	0.912 (1.0)	0.820 (2.1)	0.856 (1.8)	0.888 (1.7)
		朴素	0.194 (0.3)	0.396 (0.3)	0.516 (0.3)	0.096 (1.3)	0.246 (1.1)	0.342 (1.1)	0.022 (2.4)	0.060 (2.1)	0.086 (2.0)
$q = 0.25$	400	提出的	0.002 (0.6)	0.006 (0.5)	0.006 (0.5)	0.308 (2.5)	0.508 (2.1)	0.610 (1.9)	0.192 (3.8)	0.388 (3.2)	0.494 (3.0)
		朴素	0.004 (0.6)	0.040 (0.5)	0.076 (0.5)	0.000 (2.7)	0.022 (2.3)	0.046 (2.1)	0.002 (4.3)	0.010 (3.7)	0.024 (3.4)
	800	提出的	0.000 (0.5)	0.006 (0.4)	0.008 (0.4)	0.612 (1.9)	0.740 (1.6)	0.808 (1.5)	0.512 (3.4)	0.676 (2.9)	0.752 (2.7)
		朴素	0.036 (0.5)	0.140 (0.4)	0.208 (0.4)	0.014 (2.2)	0.056 (1.9)	0.088 (1.7)	0.008 (3.4)	0.024 (2.9)	0.040 (2.7)
	1200	提出的	0.000 (0.4)	0.010 (0.3)	0.012 (0.3)	0.830 (1.5)	0.890 (1.3)	0.916 (1.2)	0.778 (2.7)	0.830 (2.3)	0.868 (2.2)
		朴素	0.170 (0.3)	0.330 (0.3)	0.444 (0.3)	0.066 (1.5)	0.194 (1.3)	0.312 (1.2)	0.012 (2.5)	0.050 (2.2)	0.070 (2.0)

关于调节参数，我们采用与 Ai 等人（2021）和 Huang 等人（2022）相同的 $K_1 = 2$ 和 $K_2 = 1$。带宽通过第 5.2 节所述的方法选取。具体而言，$\tilde{}$ $\tilde{}$初始带宽 $h_0 = 1.1h$，其中 h 通过最小化第 5.2 节定义的 $G(2, 1, h)$（相对于 h）选出。图 2右侧的图分别展示了估计的中位数 QDRF及其导数在 j 上的上确界距离。我们选择图中首次出现显著上升转折点的位置 j，由此得到中位数剂量-反应 g 的 $j = 16$，以及导数的 $j = 17$。

图 2 展示了估计得到的分位数剂量-反应函数（QDRF）及其一阶导数，以及使用第 5 节介绍的方法 1 和方法 2 得到的 95% 一致置信带。这些导数及其置信带表明，广告对竞选捐款分布的因果影响不存在异质性。与 Ai 等人 (2021年) 和 Huang 等人 (2022年) 的平均剂量反应函数（ADRF）结果一致，广告对竞选捐款分布的因果影响在广告投放量相对较小时显著（#广告 < 400），但超过这一阈值后便变得微乎其微。

表 5： 1000 ×欠平滑估计量（方法 1）的平均平方偏差（平均方差），基于网格 $\mathcal{V}$ 上的 500 次蒙特卡洛模拟

			DGP0		DGP1L		DGP1NL	
	N	方法	$\widehat{g}_{\widehat{h}_{\mathrm{u}}}$	$\widehat{g'}_{\widehat{h}'_{\mathrm{u}}}$	$\widehat{g}_{\widehat{h}_{\mathrm{u}}}$	$\widehat{g'}_{\widehat{h}'_{\mathrm{u}}}$	$\widehat{g}_{\widehat{h}_{\mathrm{u}}}$	$\widehat{g'}_{\widehat{h}'_{\mathrm{u}}}$
平均值	400	提出的	0.03 (3.95)	0.14 (75.9)	1.13 (119)	5.65 (2509)	1.26 (130)	8.90 (2628)
		朴素	1.23 (4.19)	7.73 (62.1)	47.2 (120)	286 (2242)	62.9 (125)	416 (2373)
	800	提出的	0.02 (2.05)	0.07 (26.8)	0.42 (61.6)	3.33 (790)	0.57 (64.4)	2.72 (783)
		朴素	1.23 (2.14)	7.33 (27.6)	44.7 (63.8)	273 (858)	62.0 (66.1)	407 (807)
	1200	提出的	0.00 (0.84)	0.05 (36.0)	0.04 (9.77)	0.54 (303)	0.12 (28.6)	10.7 (590)
		朴素	1.19（0.82）	7.44（53.8）	19.4（9.82）	120（141）	60.9（27.4）	416 (671)
$q = 0.45$	400	提出的	0.19 (5.90)	0.45 (140)	1.84 (81.5)	3.83 (1938)	5.65 (195)	10.9 (4662)
		朴素	1.37 (6.59)	7.37 (117)	22.0 (78.1)	128 (1715)	69.8 (175)	427 (3293)
	800	提出的	0.08 (2.93)	0.13 (50.9)	0.68 (42.7)	4.61 (939)	1.77 (102)	6.01 (2071)
		朴素	1.18 (2.84)	6.83 (56.1)	20.1 (39.7)	116 (493)	64.4 (96.5)	408 (1616)
	1200	提出的	0.06 (1.20)	0.06 (28.0)	0.77 (15.8)	0.52 (236)	2.04 (43.1)	8.80 (626)
		朴素	1.12 (1.22)	6.42 (27.1)	20.3 (16.5)	118 (347)	64.0 (41.0)	423 (729)
$q = 0.35$	400	提出的	0.28 (6.37)	0.38 (122)	3.09 (112)	6.38 (2208)	8.12 (211)	8.40 (4012)
		朴素	1.25 (5.54)	6.22 (88.8)	22.5 (91.3)	129 (1654)	70.0 (200)	415 (3633)
	800	提出的	0.14 (2.93)	0.31 (55.0)	1.75 (45.4)	4.10 (804)	4.94 (107)	9.91 (1981)
		朴素	1.11 (2.80)	6.18 (33.7)	21.1 (45.2)	120 (701)	68.8 (110)	418 (1990)
	1200	提出的	0.09 (1.11)	0.07 (28.3)	1.13 (17.0)	1.02 (296)	3.74 (47.8)	7.11 (1235)
		朴素	1.01 (1.12)	5.80 (16.5)	20.2 (17.9)	120 (282)	65.7 (46.2)	419 (697)
$q = 0.25$	400	提出的	0.46 (6.88)	0.44 (133)	5.31 (120)	4.40 (2111)	14.1 (245)	12.7 (4657)
		朴素	1.23 (5.85)	6.07 (114)	23.5 (99.4)	125 (1711)	71.7 (247)	451 (4235)
	800	提出的	0.21 (2.85)	0.23 (40.8)	3.24 (51.3)	6.68 (769)	8.21 (122)	12.5 (2858)
		朴素	1.05 (2.79)	5.54 (38.6)	22.6 (51.8)	127 (678)	70.8 (119)	410 (1906)
	1200	提出的	0.14 (1.23)	0.07 (27.9)	2.22 (21.1)	2.14 (628)	5.61 (53.1)	2.93 (793)
		朴素	0.93 (1.14)	4.97 (17.9)	20.8 (18.5)	117 (257)	66.7 (49.5)	412 (711)

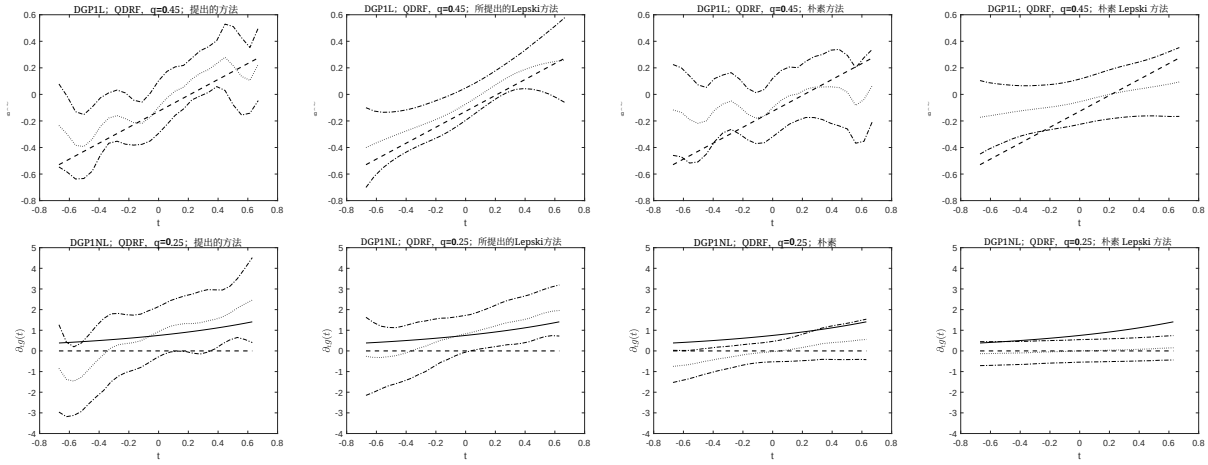


图 1：根据满足 $N = 1200$ 的样本绘制的 g or g' 估计量（点线图），以及针对 g 的 90% 一致置信带（点划线）和用于检验 H_0 的 $g' = 0$（虚线）。真实的 g' 曲线以实线表示（第 2 行）。

表6： 1000 ×基于网格进行500次蒙特卡洛模拟所得的 Lepski 估计量（方法2）的平均平方偏差（平均方差） $\mathcal{V}$

	N	方法	DGP0 $\widehat{g}_{\widehat{h}}$	DGP0 $\widehat{g'}_{\widehat{h}'}$	DGP1L $\widehat{g}_{\widehat{h}}$	DGP1L $\widehat{g'}_{\widehat{h}'}$	DGP1NL $\widehat{g}_{\widehat{h}}$	DGP1NL $\widehat{g'}_{\widehat{h}'}$
平均值	400	提出的	0.01（0.72）	0.01（11.6）	0.06（19.1）	0.62（391）	0.17（19.3）	18.6（540）
		朴素	1.24 (0.86)	7.56 (15.3)	45.7 (23.8)	279 (854)	60.5 (24.7)	422 (549)
	800	提出的	0.01 (0.39)	0.02 (10.9)	0.07 (10.5)	0.94 (253)	0.12 (11.6)	11.2 (295)
		朴素	1.21 (0.42)	7.48 (8.58)	44.8 (12.5)	276 (333)	60.4 (13.1)	414 (336)
	1200	提出的	0.00 (0.21)	0.02 (8.47)	0.01 (2.74)	0.19 (90.4)	0.17 (7.15)	20.1 (293)
		朴素	1.18 (0.28)	7.42 (5.81)	19.3 (3.43)	120 (84.5)	59.8 (9.23)	420 (237)
$q = 0.45$	400	提出的	0.10 (0.79)	0.03 (7.65)	2.19 (9.73)	0.37 (100)	7.39 (23.0)	18.6 (273)
		朴素	1.25 (0.88)	6.86 (7.42)	21.8 (11.9)	119 (123)	69.5 (27.4)	425 (317)
	800	提出的	0.05 (0.37)	0.03 (6.64)	0.69 (5.26)	0.46 (82.3)	2.68 (12.8)	14.1 (240)
		朴素	1.14 (0.40)	6.76 (5.41)	19.5 (5.95)	118 (86.2)	63.0 (14.2)	422 (364)
	1200	提出的	0.06 (0.25)	0.02 (5.17)	0.88 (3.56)	0.28 (59.5)	3.54 (7.96)	26.9 (155)
		朴素	1.12 (0.29)	6.47 (4.98)	20.3 (4.09)	118 (77.2)	64.2 (9.20)	435 (207)
$q = 0.35$	400	提出的	0.09 (0.77)	0.04 (6.87)	1.58 (11.8)	0.60 (128)	6.36 (28.4)	15.2 (286)
		朴素	1.06 (0.88)	6.18 (7.46)	21.4 (14.1)	124 (141)	67.7 (33.1)	431 (330)
	800	提出的	0.08 (0.37)	0.06 (8.38)	1.21 (5.90)	0.59 (71.6)	4.49 (14.0)	14.3 (216)
		朴素	1.05 (0.40)	6.10 (7.49)	20.4 (6.46)	121 (81.8)	66.3 (15.5)	438 (240)
	1200	提出的	0.08 (0.26)	0.02 (8.39)	1.32 (4.04)	0.16 (57.0)	5.31 (9.68)	24.3 (181)
		朴素	1.00 (0.27)	5.70（3.63）	20.3（4.60）	119（58.8）	65.3（10.8）	433（223）
$q = 0.25$	400	提出的	0.10 (0.89)	0.04 (8.54)	1.89 (14.5)	0.95 (105)	7.42 (32.4)	10.5 (227)
		朴素	0.94 (0.93)	5.49 (8.73)	21.3 (16.2)	121 (120)	67.8 (35.8)	430 (328)
	800	提出的	0.10 (0.43)	0.04 (6.02)	1.62 (7.06)	0.57 (73.1)	5.93 (17.8)	11.7 (488)
		朴素	0.96 (0.44)	5.48 (6.42)	21.1 (8.18)	124 (104)	66.1 (18.6)	428 (478)
	1200	提出的	0.11 (0.27)	0.02 (8.39)	1.93 (4.82)	0.39 (94.3)	7.12 (12.5)	16.8 (355)
		朴素	0.93 (0.29)	5.02 (3.79)	20.5 (5.11)	117 (76.3)	66.0 (12.1)	427 (195)

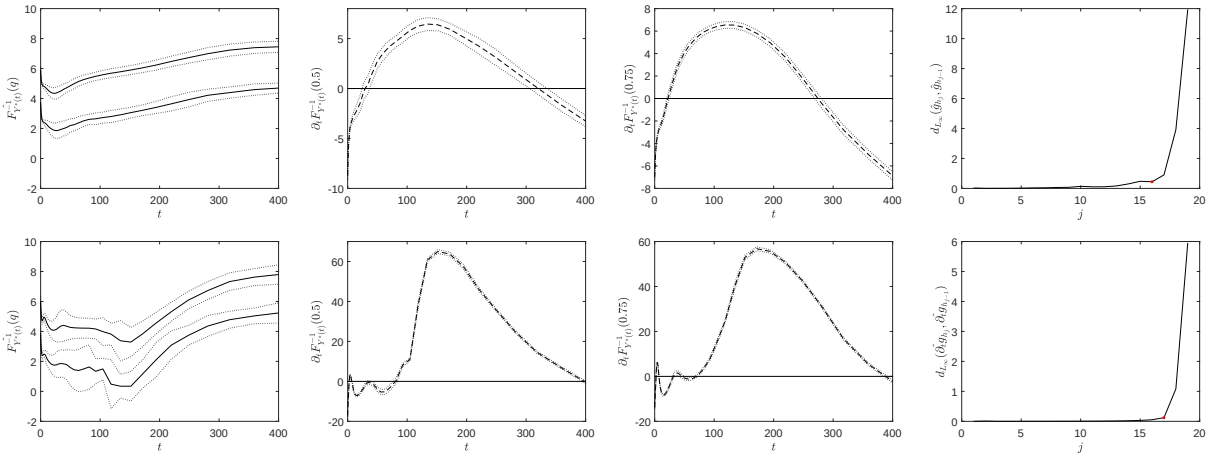


图 2：使用方法1（左上）和方法2（左下）得到的 QDRF 估计值及其 95% 一致置信带（从下到上分别为 $q = 0.5$ 和 0.75），使用方法1（上）和方法2$\widehat{\ }$ $\widehat{\ }$（下）得到的 QDRF 一阶导数估计值及其 95% 置信带（第 2、3 列），以及方法1得到的 $d_{L_\infty}(\widehat{g}_{h_j}, \widehat{g}_{h_{j-1}})$（右上）和 $d_{L_\infty}(\partial_t g_{h_j}, \partial_t g_{h_{j-1}})$（右下）的图，其中 g 为中位剂量-反应函数。

附录

A 的证明(2.2)

根据假设 1，即 $T \perp Y^*(t)|\boldsymbol{X}$，并利用条件期望的塔式性质，对于每个 $t \in \mathcal{T}$，广义剂量反应函数 $g(t)$ 可按点识别如下：

$$
\begin{aligned}
g(t) &= \arg\min_{a\in\mathbb{R}} \mathbb{E}\left[\mathcal{L}\left(Y^*(t)-a\right)\right] \\
&= \arg\min_{a\in\mathbb{R}} \int \mathbb{E}\left[\mathcal{L}\left(Y^*(t)-a\right)|\boldsymbol{X}=\boldsymbol{x}\right] f_X(\boldsymbol{x})d\boldsymbol{x} \\
&= \arg\min_{a\in\mathbb{R}} \int \mathbb{E}\left[\mathcal{L}\left(Y^*(t)-a\right)|\boldsymbol{X}=\boldsymbol{x}, T=t\right] f_X(\boldsymbol{x})d\boldsymbol{x} \quad (Assumption\ 1) \\
&= \arg\min_{a\in\mathbb{R}} \int \mathbb{E}\left[\mathcal{L}\left(Y-a\right)|\boldsymbol{X}=\boldsymbol{x}, T=t\right] f_X(\boldsymbol{x})d\boldsymbol{x} \\
&= \arg\min_{a\in\mathbb{R}} \int \mathbb{E}\left[\pi(T,\boldsymbol{X})\mathcal{L}\left(Y-a\right)|\boldsymbol{X}=\boldsymbol{x}, T=t\right] \cdot f_{X|T}(\boldsymbol{x}|t)d\boldsymbol{x} \\
&= \arg\min_{a\in\mathbb{R}} \mathbb{E}\left[\pi(T,\boldsymbol{X})\mathcal{L}\left(Y-a\right)|T=t\right].
\end{aligned}
$$

B 的推导(3.4)

Let

$$
b_K := \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1, i\neq j}^{N} u_K\left(T_i, \boldsymbol{X}_i\right).
$$

此外，令 $U_{K\times N} := (u_K(T_1, \boldsymbol{X}_1) \ldots, u_K(T_N, \boldsymbol{X}_N)) \in \mathbb{R}^{K\times N}$，$\boldsymbol{\pi} = (\pi_1, \ldots, \pi_N)$，以及 $F(\boldsymbol{\pi})$: $= \sum_{i=1}^{N}(\pi_i - 1)^2/2$。那么 (3.3) 可写为

$$
\begin{cases}
\min_{\boldsymbol{\pi}} F\left(\boldsymbol{\pi}\right) \\
\text{subject to } U_{K\times N} \cdot \boldsymbol{\pi} = N \cdot b_K .
\end{cases}
\tag{B.1}
$$

共轭凸函数（Tseng 和 Bertsekas，1991）$F(\cdot)$ 定义为

$$
F^*(\boldsymbol{z}) = \sup_{\boldsymbol{\pi}} \sum_{i=1}^{N} \left\{z_i\pi_i - (\pi_i - 1)^2/2\right\} = \sum_{i=1}^{N} \left\{z_i\pi_i^* - (\pi_i^* - 1)^2/2\right\},
$$

其中 π_i^* 满足一阶条件：

$$
z_i = \pi_i^* - 1 \Rightarrow \pi_i^* = z_i + 1, \quad i = 1, \ldots, N.
$$

将 $\pi_i^* = z_i + 1$ 代入 $F^*(\boldsymbol{z})$，得到

$$
F^*(\boldsymbol{z}) = \sum_{i=1}^{N} \left\{z_i(z_i+1) - z_i^2/2\right\} = \sum_{i=1}^{N} z_i^2/2 + z_i = \sum_{i=1}^{N} \frac{1}{2}(z_i+1)^2 - \frac{1}{2}.
$$

根据 Tseng 和 Bertsekas (1991)， （B.1）的对偶问题是求解 $z = (-\gamma^*)^\top U_{K\times N}$，其中 γ^*为待求解变量

$$\begin{aligned}
&\max_{\gamma\in\mathbb{R}^K}\left\{-F^*\left(-\gamma^\top U_{K\times N}\right)+(-\gamma)^\top(N\cdot b_K)\right\}\\
&=\max_{\gamma\in\mathbb{R}^K}\left[\sum_{i=1}^N\left\{-\frac{1}{2}(\gamma^\top u_K(T_i,\boldsymbol{X}_i)-1)^2+\frac{1}{2}\right\}-\gamma^\top(N\cdot b_K)\right]\\
&=\max_{\gamma\in\mathbb{R}^K}N\{\widehat{G}(\gamma)+1/2\},
\end{aligned}$$

其中

$$\widehat{G}(\gamma)=-\frac{1}{2N}\sum_{i=1}^N\left\{\gamma^\top u_K\left(T_i,\boldsymbol{X}_i\right)-1\right\}^2-\frac{1}{N(N-1)}\sum_{i=1}^N\sum_{j=1,i\neq j}^N\gamma^\top u_K\left(T_i,\boldsymbol{X}_i\right).$$

因此,

$$\widehat{\pi}_K(T_i,\boldsymbol{X}_i)=-\widehat{\gamma}^\top u_K(T_i,\boldsymbol{X}_i)+1,$$

其中

$$\begin{aligned}
&\widehat{\gamma}=\arg\max\widehat{G}(\gamma)\\
&=\left(\frac{1}{N}\sum_{i=1}^N u_K\left(T_i,\boldsymbol{X}_i\right)u_K^\top\left(T_i,\boldsymbol{X}_i\right)\right)^{-1}\left(\frac{1}{N}\sum_{i=1}^N u_K\left(T_i,\boldsymbol{X}_i\right)-\frac{1}{N(N-1)}\sum_{i=1}^N\sum_{j=1,i\neq j}^N u_K\left(T_i,\boldsymbol{X}_i\right)\right).
\end{aligned}$$

C 假设验证 8(i)

当 $\mathcal{L}(v) = v^2$ 时，我们验证假设 8(i) 中对 $\mathcal{F}_1$ 和 $\mathcal{F}_2$ 所施加的 Vapnik–Chervonenkis (VC) 型条件。假设3结合 $g(t)$ 和 $g'(t)$ 的连续性，保证了 M 的存在。对于固定的 a 和 b，函数 $(y,t)\mapsto\mathcal{L}(y-a-bt)$ 的次图满足

$$\begin{aligned}
&\{(y,t,u):u\leq\mathcal{L}(y-a-bt)\}=\{(y,t,u):u\leq(y-a-bt)^2\}\\
=&\{(y,t,u):y-a-bt\geq\sqrt{u\vee 0}\}\cup\{(y,t,u):y-a-bt\leq-\sqrt{u\vee 0}\}\in\mathcal{G}\cup\mathcal{G},
\end{aligned}$$

其中

$$\mathcal{G}:=\{\{(y,t,u):x_1\cdot y+x_2\cdot t+x_3\cdot\sqrt{u\vee 0}+x_4\leq 0\}:x_1,x_2,x_3,x_4\in\mathbb{R}\},$$

是由四个可测函数生成的集合：$(y,t,u)\mapsto y$、$(y,t,u)\mapsto t$、$(y,t,u)\mapsto\sqrt{u\vee 0}$ 以及$(y,t,u)\mapsto 1$。根据 Van Der Vaart 和 Wellner（1996，引理 2.6.15、引理 2.6.18 (iii)），$\mathcal{G}$ 是一个 VC 集类。随后，根据 Van Der Vaart 和 Wellner（1996，引理 2.6.17 (iii)），$\mathcal{G}\cup\mathcal{G}$也是一个 VC 集类。因此，假设 8(i)中的 $\mathcal{F}_1$ 是一个 VC 次图类。根据 Van Der Vaart 和 Wellner（1996，定理 2.6.7），假设 8(i)中对 $\mathcal{F}_1$ 施加的 VC 型条件成立。同理，假设 8(i)中对 $\mathcal{F}_2$ 施加的 VC 型条件也成立。

当 $\mathcal{L}(v) = v\{\tau-\mathbb{1}(v\leq 0)\}$ 时，可以类似地验证 $\mathcal{F}_1$ 和 $\mathcal{F}_2$ 的 VC 型条件。

参考文献

Adusumilli, K., D. Kurisu, T. Otsu 和 Y.-J. Whang（2020）。测量误差下分布函数的推断。《计量经济学杂志》*215*(1)，131–164。Ai, C., O. Linton, K. Motegi 和 Z. Zhang（2021）。一般处理模型高效估计的统一框架。《数量经济学》*12*(3)，779–816。Bissantz, N., L. D¨umbgen, H. Holzmann 和 A. Munk（2007）。反卷积密度估计中的非参数置信带。《英国皇家统计学会杂志*B*辑：统计方法学》*69*(3)，483–506。Chen, X.（2007）。半非参数模型的大样本筛估计。《计量经济学手册》*6*(B)，5549–5632。Chen, X., T. Christensen 和 S. Kankanala（2025）。非参数结构函数和弹性的自适应估计与一致置信带。经济研究评论 *92*(1)，162–196。Chen, X., D. Pouzo 和 J. L. Powell（2019）。用于非参数分位数工具变量回归加权平均导数的惩罚筛广义经验似然。《计量经济学杂志》*213*(1)，30–53。Chernozhukov, V., D. Chetverikov 和 K. Kato（2014a）。反集中与诚实、自适应置信带。统计年鉴 *42*(5)，1787–1818。Chernozhukov, V., D. Chetverikov 和 K. Kato（2014b）。经验过程上确界的高斯近似。统计年鉴 *42*(4)，1564–1597。Chernozhukov, V., I. Fern´andez-Val 和 B. Melly（2013）。反事实分布的推断。计量经济学 *81*(6)，2205–2268。Colangelo, K. 和 Y.-Y. Lee（2020）。连续处理下双重去偏机器学习非参数推断。*arXiv* 预印本 *arXiv:2004.03036*。Fan, J., T.-C. Hu 和 Y. K. Truong（1994）。稳健的非参数函数估计。斯堪的纳维亚统计学杂志，433–446。Fong, C., C. Hazlett 和 K. Imai（2018）。连续处理的协变量平衡倾向得分：政治广告效果的应用。应用统计年鉴 *12*(1)，156–177。Gin´e, E. 和 A. Guillou（2002）。多元核密度估计量强一致收敛的速率。亨利·庞加莱研究所年鉴（*B*辑）：概率与统计 *38*(6)，907–921。Hirano, K. 和 G. W. I mbens（2004）。连续处理的倾向得分。载于 A. Gelman 和 X.-L. Meng（编），不完整数据视角下的贝叶斯建模与因果推断，第7章，第73–84页。约翰·威立父子有限公司。

Horowitz, J. L. 和 S. Lee（2017）。形状约束下的非参数估计与推断。《计量经济学杂志》*201*(1), 108–126。Huang, W., O. Linton 和 Z. Zhang（2022）。连续处理效应模型设定检验的统一框架。《商业与经济统计学杂志》*40*(4), 1817–1830。Huang, W. 和 Z. Zhang（2023）。存在测量误差时连续处理效应的非参数估计。《皇家统计学会会刊*B*辑：统计方法学》*85*(2), 474–496。Kang, J. D., J. L. Schafer 等（2007年）。揭开双重稳健性的神秘面纱：从不完整数据估计总体均值的替代策略比较。《统计科学》*22*(4), 523–539。Kato, K.（2011）。关于自助法M估计量矩收敛的一则说明。《统计与决策》*28*(1), 51–61。Kato, K. 和 Y. Sasaki（2018）。误差分布未知时反卷积的一致置信带。《计量经济学杂志》*207*(1), 129–161。Kato, K. 和 Y. Sasaki（2019）。非参数变量误差回归的一致置信带。《计量经济学杂志》*213*(2), 516–555。Kennedy, E. H., Z. Ma, M. D. McHugh 和 D. S. Small（2017）。连续处理效应双重稳健估计的非参数方法。《皇家统计学会会刊*B*辑（统计方法学）》*79*(4), 1229–1245。Lejeune, M. G. 和 P. Sarda（1988）。分位数回归：一种非参数方法。《计算统计与数据分析》*6*(3), 229–239。Li, J. 和 Z. Liao（2020）。时间序列的一致非参数推断。《计量经济学杂志》*219*(1), 38–51。Robins, J. M., M. A. Hern´an 和 B. Brumback（2000）。边际结构模型与流行病学中的因果推断。《流行病学》*11*(5), 550–560。Tseng, P. 和 D. P. Bertsekas（1991）。具有严格凸成本和线性约束问题的松弛方法。《运筹学数学》*16*(3), 462–481。Urban, C. 和 S. Niebler（2014年）。人行道上的钱：美国总统候选人是否应该在无竞争州投放广告？《美国政治科学杂志》*58*(2), 322–336。Van Der Vaart, A. W. 和 J. A. Wellner（1996）。弱收敛与经验过程及其在统计学中的应用。施普林格。Yiu, S. 和 L. Su（2018）。协变量关联消除权重：因果效应估计的统一加权框架。《生物统计学》*105*(3), 709–722。Zhang, Y. 和 X. Zheng（2020）。协变量自适应随机化下的分位数处理效应与自助法推断。《数量经济学》*11*(3), 957–982。