

附录 1：Fong 等（2018）以及平均剂量响应函数估计量的补充说明

Fong 等（2018）采用如下方法估计平衡权重，其中 $\hat{w}$ 代表平衡权重的估计值：

$$\hat{w} = \arg \min_w \sum_{i=1}^n \log w_i$$

使得

$$\frac{1}{n} \sum_{i=1}^n w_i A_i X_i = \frac{1}{n} \sum_{i=1}^n A_i \frac{1}{n} \sum_{i=1}^n X_i, \quad \sum_{i=1}^n w_i X_i = \sum_{i=1}^n X_i, \quad \sum_{i=1}^n w_i A_i = \sum_{i=1}^n A_i$$

以及

$$\sum_{i=1}^n w_i = n, \quad w_i \geq 0, \quad \text{其中 } i=1, \dots, n$$

上述最优化为带等式和不等式约束的非线性规划。第一部分约束为等式约束，利用协变量和处理变量的一阶矩条件。第二部分约束包含等式约束和不等式约束，其中等式约束是对权重进行归一化。鉴于上述方法仅利用协变量和处理变量的部分信息且无法保证达到最优的协变量平衡，部分研究（Ai 等，2021；Huling 等，2023；Jiang 等，2022）优化协变量平衡方法，但除了 Ai 等（2021），最优化仍然为带等式和不等式约束的非线性规划，最优化求解相对复杂。

当平衡权重被估计完成，大部分研究基于 $EY(a) = E(w^S Y | A = a)$ 构造平均剂量响应函数的估计量，如 Fong 等（2018），Ai 等（2021），Vegetabile 等（2021），Huling 等（2023）等。令 $\hat{w}^S = \{\hat{w}_i^S\}_{i=1}^n$ 为 w^S 的估计量，采用一些参数或者非参数方法可得到 $\tau(a)$ 的估计量。例如，采用 Nadaraya-Watson 方法， $\tau(a)$ 可被估计为

$$\hat{\tau}_{NW}(a) = \frac{E_n[\hat{w}^S Y K_h(A-a)]}{E_n[K_h(A-a)]} = \frac{\sum_{i=1}^n \hat{w}_i^S Y_i K_h(A_i - a)}{\sum_{i=1}^n K_h(A_i - a)}$$

其中， $K_h(\cdot) = h^{-1}K(\cdot/h)$ ， $K(\cdot)$ 为核函数且 h 为带宽。

附录 2：引理 2 的证明

由于式（8）对任意可积函数 $g(X)$ 均成立，令 $g(X) = \exp\{a'Xi\}$ ，其中向量 $a \in \mathbb{R}^d$ ， i 为虚数单位，从而有

$$\begin{aligned} 0 &= E\left[\left(\frac{K_h(A-a)}{f(X)} - 1\right) \exp\{a'Xi\}\right] = E\left[\exp\{a'Xi\} E\left[\frac{K_h(A-a)}{f(X)} - 1 \middle| X\right]\right] \\ &= E\left[\exp\{a'Xi\} \frac{\pi^*(a, X) + O_p(h^2) - f(X)}{f(X)}\right] \end{aligned}$$

由级数转化的唯一性有 $\pi^*(a, X) + O_p(h^2) - f(X) = 0$ ，其中 $x \in X$ ，即 $f(X) = \pi^*(a, X) + O_p(h^2)$ 。

由于

$$E\left[\left(\frac{K_h(A-a)}{f(X)}-1\right)g(X)\right]=E\left[\frac{E[K_h(A-a)|X]-f(X)}{f(X)}g(X)\right]$$

推导可得 $E[K_h(A-a)|X]$ 等价于

$$\begin{aligned} E[K_h(A-a)|X] &= \int_A K_h(A-a)\pi^*(A,X)dA \\ &= \int_{t \in T} K(t)\pi^*(a+th,X)dt \\ &= \int_{t \in T} K(t)\left(\pi^*(a,X)+\frac{\partial \pi^*(a,X)}{\partial a}th+\frac{1}{2}\frac{\partial^2 \pi^*(\bar{a},X)}{\partial a^2}t^2h^2\right)dt \\ &= \pi^*(a,X)+\frac{h^2}{2}\frac{\partial^2 \pi^*(\bar{a},X)}{\partial a^2}\int_{t \in T} t^2K(t)dt \end{aligned}$$

其中 $\bar{a}$ 介于 a 和 $a+th$ 。第一个等式源于期望的定义，第二个等式源于积分变换，令 $t=(A-a)/h$ ，第三个等式源于泰勒展开，第四个等式源于假设 4 和假设 5。

$$\text{从而当 } f(X)=\pi^*(a,X)+\frac{h^2}{2}\frac{\partial^2 \pi^*(\bar{a},X)}{\partial a^2}\int_{t \in T} t^2K(t)dt \text{ 时有 } E\left[\left(\frac{K_h(A-a)}{f(X)}-1\right)g(X)\right]=0。 \text{同时，}$$

由假设 2 可知 $\pi^*(a,X)$ 远离 0，从而当 h 充分小时有 $f(X) \neq 0$ 。因此式 (8) 确实存在解。

Q.E.D

附录 3：定理 1 的证明

定理 1 的证明可拆解为 4 个部分，分别为 (1) 相关知识，(2) 引理 S1 及证明，(3) 引理 S2 及证明，(4) 定理 1 的证明。其中，引理 S1 和引理 S2 会在定理 1 的证明中被使用。

(1) 相关知识

对于任意矩阵 A ，令矩阵范数为 $\|A\|=\sqrt{\text{tr}(A'A)}$ ，其中 $\text{tr}(\cdot)$ 代表矩阵的迹， A' 代表矩阵 A 的转置。令 C 和 C_i 代表正常数。参考 Hirano 等 (2003)，为逼近广义倾向得分 $\pi^*(a,x)$ ，采用幂级数序列 $R_K(X)$ 并对序列进行正交化，从而有 $E[R_K(X)R_K(X)']=I_K$ 且 $\xi(K) \leq CK$ ，其中 $\xi(K)=\sup_{x \in X} \|R_K(x)\|$ 。相应地，有 $R_K(X)'R_K(X)=\|R_K(x)\|^2 \leq \xi(K)$ ，以及

$$\begin{aligned} E[R_K(X)'R_K(X)] &= E[\text{tr}(R_K(X)'R_K(X))] = E[\text{tr}(R_K(X)R_K(X)')] \\ &= \text{tr}[E(R_K(X)R_K(X)')] = \text{tr}(I_K) = K \end{aligned}$$

由于广义倾向得分 $\pi^*(a,x)$ 的取值范围为 $(0,+\infty)$ ，本文采用幂级数序列 $R_K(X)$ 逼近 $\pi^*(a,x)$ 的对数。假设 $\pi^*(a,x)$ 在定义域 X 上 s 阶连续可微，基于 Hirano 等 (2003)，有

$$\sup_{x \in X} |\ln \pi^*(a,x) - R_K(x)' \beta_K| < CK^{-s/d} \quad (\text{S1})$$

$\pi^*(a,x)$ 的估计量为 $\hat{\pi}_K(a,x) = \exp\{R_K(x)'\hat{\beta}_K\}$ ，其中

$$\hat{\beta}_K = \arg \min_{\beta} Q_n(\beta) \quad (\text{S2})$$

且

$$Q_n(\beta) = E_n \left[K_h(A-a) \exp \{-R_K(X)' \beta\} + R_K(X)' \beta \right] \quad (S3)$$

为后续证明方便，定义 β_K^* 为：

$$\beta_K^* = \arg \min_{\beta} Q^*(\beta) \quad (S4)$$

其中

$$Q^*(\beta) = E \left[\pi^*(a, X) \exp \{-R_K(X)' \beta\} + R_K(X)' \beta \right] \quad (S5)$$

相应地，定义 $\pi_K^*(a, x) = \exp \{R_K(x)' \beta_K^*\}$ 。

（2）引理 S1 及证明

引理 S1：假设

- ① 协变量 X 的定义域 $\mathcal{X}$ 为 $\mathbb{R}^d$ 上的紧致集；
- ② $\pi^*(a, x)$ 在定义域 $\mathcal{X}$ 上 s 阶连续可微，且有 $s/d \geq 4$ ；
- ③ $\pi^*(a, x)$ 在定义域 $\mathcal{X}$ 上有界，且取值下限远离 0；
- ④ 协变量 X 的密度函数在定义域 $\mathcal{X}$ 上远离 0。

从而有 $\|\beta_K - \beta_K^*\| = O_p(K^{-s/(2d)})$ ，以及 $\sup_{x \in \mathcal{X}} |\pi^*(a, x) - \pi_K^*(a, x)| = O(K^{-s/(2d)} \xi(K))$ 。

证明：令 $\eta = \inf \left\{ \inf_{x \in \mathcal{X}} \pi^*(a, x), \inf_{x \in \mathcal{X}} (1/\pi^*(a, x)) \right\}$ ，基于引理 S1 的假设③可得 $\eta > 0$ 。定义集合

$$B_K = \left\{ \beta \in \mathbb{R}^K : \inf_{x \in \mathcal{X}} \exp \{R_K(x)' \beta\}, \inf_{x \in \mathcal{X}} \exp \{-R_K(x)' \beta\} \geq \eta/2 \right\}$$

基于式 (S1) 以及指数函数的单调性，有

$$\exp \{R_K(x)' \beta_K - CK^{-s/d}\} - \exp \{R_K(x)' \beta_K\} \leq \pi^*(a, x) - \exp \{R_K(x)' \beta_K\} \leq \exp \{R_K(x)' \beta_K + CK^{-s/d}\} - \exp \{R_K(x)' \beta_K\}$$

基于中值定理，当 K 充分大时有

$$\begin{aligned} \exp \{R_K(x)' \beta_K - CK^{-s/d}\} - \exp \{R_K(x)' \beta_K\} &= -CK^{-s/d} \exp \{R_K(x)' \bar{\beta}\} \\ &\geq -CK^{-s/d} \exp \{R_K(x)' \beta_K\} \\ &\geq -CK^{-s/d} \exp \{\ln \pi^*(a, x) + CK^{-s/d}\} \\ &\geq -2CK^{-s/d} \pi^*(a, x) \\ &\geq \frac{-2C}{\eta} K^{-s/d} \end{aligned}$$

其中， $\bar{\beta}$ 使得 $R_K(x)' \bar{\beta}$ 介于 $R_K(x)' \beta_K - CK^{-s/d}$ 和 $R_K(x)' \beta_K$ 之间。第二个不等式源于式 (S1)，第

三个不等式源于 $\exp \{CK^{-s/d}\}$ 随 K 趋于零。同理有

$$\exp \{R_K(x)' \beta_K + CK^{-s/d}\} - \exp \{R_K(x)' \beta_K\} \leq \frac{2C}{\eta} K^{-s/d}$$

从而

$$\sup_{x \in \mathcal{X}} |\pi^*(a, x) - \exp\{R_K(x)' \beta_K\}| < CK^{-s/d} \quad (\text{S6})$$

当 K 足够大时有 $CK^{-s/d}$ 收敛到零，即 $\pi^*(a, x)$ 和 $\exp\{R_K(x)' \beta_K\}$ 足够接近，基于三角不等式有 $\beta_K \in B_K$ 。

定义函数 $Q_K(\beta)$ 为

$$Q_K(\beta) = E[\exp\{R_K(X)' \beta_K\} \exp\{-R_K(X)' \beta\} + R_K(X)' \beta] \quad (\text{S7})$$

将 $Q_K(\beta)$ 基于 β 求一阶导并令一阶导函数等于零，求解得出 $\beta = \beta_K$ 。从而有 $\beta_K = \arg \min_{\beta} Q_K(\beta)$ 。

由于当 $\beta \in B_K$ 时有 $\exp\{R_K(X)' \beta\}$ 有界，从而存在常数 C_1 使得

$$\sup_{\beta \in \tilde{B}_K} |Q^*(\beta) - Q_K(\beta)| < C_1 K^{-s/d} \quad (\text{S8})$$

围绕 β_K 定义如下紧致集 $\tilde{B}_K$ ，其中 C_2 为某个正常数，

$$\tilde{B}_K = \{\beta \in \mathbb{R}^K : \|\beta - \beta_K\| \leq C_2 K^{-s/(2d)}\}$$

由于

$$\begin{aligned} \sup_{x \in \mathcal{X}, \beta \in \tilde{B}_K} |\exp\{R_K(X)' \beta\} - \exp\{R_K(X)' \beta_K\}| &= \sup_{x \in \mathcal{X}, \beta \in \tilde{B}_K} |\exp\{R_K(x)' \bar{\beta}\} R_K(x)' (\beta - \beta_K)| \\ &\leq \sup_{x \in \mathcal{X}, \beta \in \tilde{B}_K} |\exp\{R_K(x)' \bar{\beta}\}| \sup_{x \in \mathcal{X}} \|R_K(x)\| \sup_{\beta \in \tilde{B}_K} \|\beta - \beta_K\| \\ &\leq \frac{4}{\eta} \xi(K) C_2 K^{-s/(2d)} \\ &= \frac{4C_2 C}{\eta} K^{1-s/(2d)} \end{aligned}$$

其中 $\bar{\beta}$ 介于 β 和 β_K 之间。上式第二个不等式源于 $\xi(K) \leq CK$ ， $s/d \geq 4$ 以及

$$\begin{aligned} \exp\{R_K(x)' \bar{\beta}\} &= \exp\{R_K(x)' (\bar{\beta} - \beta_K)\} \exp\{R_K(x)' \beta_K\} \\ &\leq \exp\{\|R_K(x)\| \|\bar{\beta} - \beta_K\|\} \exp\{R_K(x)' \beta_K\} \\ &\leq \frac{2}{\eta} \exp\{C_2 \xi(K) K^{-s/(2d)}\} \leq \frac{2}{\eta} \exp\{C_2 CK^{1-s/(2d)}\} \leq \frac{4}{\eta} \end{aligned}$$

相应地， $\sup_{x \in \mathcal{X}, \beta \in \tilde{B}_K} |\exp\{R_K(X)' \beta\} - \exp\{R_K(X)' \beta_K\}|$ 收敛到零，从而有 $\tilde{B}_K \in B_K$ 。

将 $Q_K(\beta)$ 基于 β 求二阶导，有

$$\frac{\partial^2 Q_K(\beta)}{\partial \beta \partial \beta'} = E[\exp\{R_K(X)' \beta_K\} \exp\{-R_K(X)' \beta\} R_K(X) R_K(X)']$$

令 $\lambda_{\min}(\cdot)$ 代表矩阵的最小特征值。当 $\beta \in \tilde{B}_K$ ，有

$$\lambda_{\min}\left(-\frac{\partial^2 Q_K(\beta)}{\partial \beta \partial \beta'}\right) = -\lambda_{\min}\left(E[\exp\{R_K(X)' \beta_K\} \exp\{-R_K(X)' \beta\} R_K(X) R_K(X)']\right) \geq -\frac{\eta^2}{4}$$

指定 $C_2 > \sqrt{\frac{16C_1}{\eta^2}}$ ，从而对任意位于 $\tilde{B}_K$ 边界上的 β ，即满足 $\|\beta - \beta_K\| = C_2 K^{-s/(2d)}$ ，有

$$\begin{aligned} Q^*(\beta_K) - Q^*(\beta) &\leq Q_K(\beta_K) - Q_K(\beta) + 2C_1 K^{-s/d} \\ &= -(\beta - \beta_K)' \frac{\partial Q_K(\beta_K)}{\partial \beta} - \frac{1}{2}(\beta - \beta_K)' \frac{\partial^2 Q_K(\tilde{\beta})}{\partial \beta \partial \beta'} (\beta - \beta_K) + 2C_1 K^{-s/d} \\ &\leq -\frac{\eta^2}{8} C_2^2 K^{-s/d} + 2C_1 K^{-s/d} < 0 \end{aligned}$$

其中 $\tilde{\beta}$ 介于 β 和 β_K 之间。基于 $Q^*(\cdot)$ 的连续性， $Q^*(\beta)$ 在紧致集 $\tilde{B}_K$ 上必然存在局部最小值点，且该局部最小值点为 $\tilde{B}_K$ 的内点。由于 $Q^*(\cdot)$ 为严格凸函数，全局最小值点为 β_K^* ，因此有 $\|\beta_K - \beta_K^*\| = O_p(K^{-s/(2d)})$ 。

由于

$$\begin{aligned} \left| \exp\{R_K(X)' \beta_K\} - \exp\{R_K(X)' \beta_K^*\} \right| &\leq \exp\{R_K(X)' \bar{\beta}\} R_K(X)' (\beta_K - \beta_K^*) \\ &\leq C \|R_K(X)\| \cdot \|\beta_K - \beta_K^*\| = O_p(\xi(K) K^{-s/(2d)}) \end{aligned}$$

基于式 (S6) 和三角不等式，有

$$\sup_{x \in X} |\pi^*(a, x) - \pi_K^*(a, x)| = O(K^{-s/(2d)} \xi(K))$$

从而得证。

Q.E.D

(3) 引理 S2 及证明

引理 S2：假设引理 S1 的假设①~④满足，并假设 $K = K(n) \rightarrow \infty$ ， $\xi(K)^2 K / (nh) \rightarrow 0$ ，

$\xi(K)^2 K h^4 \rightarrow 0$ 且 $\xi(K)^4 K^{1-s/d} \rightarrow 0$ 。从而有

$$\|\hat{\beta}_K - \beta_K^*\| = O_p\left(\sqrt{\frac{K}{nh}}\right) + O_p(h^2 \sqrt{K}) + O_p(\xi(K) K^{1/2-s/(2d)})$$

证明：经推导可得，

$$\begin{aligned} &E \left[\left\| \frac{\partial Q_n(\beta_K^*)}{\partial \beta} \right\|^2 \right] \\ &= E \left[\text{tr} \left(\frac{\partial Q_n(\beta_K^*)}{\partial \beta} \frac{\partial Q_n(\beta_K^*)}{\partial \beta'} \right) \right] \\ &= E \left[\frac{1}{n^2} \sum_{i=1}^n \left(1 - \frac{K_h(A_i - a)}{\pi_K^*(a, X_i)} \right)^2 R_K(X_i)' R_K(X_i) + \frac{1}{n^2} \sum_{i,j=1, i \neq j}^n \left(1 - \frac{K_h(A_i - a)}{\pi_K^*(a, X_i)} \right) R_K(X_i)' \left(1 - \frac{K_h(A_j - a)}{\pi_K^*(a, X_j)} \right) R_K(X_j) \right] \\ &= \frac{1}{n} E \left[\left(1 - \frac{K_h(A - a)}{\pi_K^*(a, X)} \right)^2 R_K(X)' R_K(X) \right] + \frac{n(n-1)}{n^2} E \left[\left(1 - \frac{K_h(A - a)}{\pi_K^*(a, X)} \right) R_K(X)' \right] E \left[\left(1 - \frac{K_h(A - a)}{\pi_K^*(a, X)} \right) R_K(X) \right] \end{aligned} \quad (\text{S9})$$

关于式 (S9) 等式右边的第一项，注意到

$$E\left[\left(\pi_K^*(a, X) - K_h(A-a)\right)^2 \middle| X\right] = \pi_K^*(a, X)^2 + E\left[K_h(A-a)^2 \middle| X\right] - 2\pi_K^*(a, X)E\left[K_h(A-a) \middle| X\right]$$

其中，上式等式右边第二项等价于

$$\begin{aligned} E\left[K_h(A-a)^2 \middle| X\right] &= \int_{A \in \mathcal{A}} K_h(A-a)^2 \pi^*(A, X) dA \\ &= \int_{t \in \mathcal{T}} \frac{1}{h} K(t)^2 \pi^*(a+ht, X) dt \\ &= \frac{1}{h} \pi^*(a, X) \int_{t \in \mathcal{T}} K(t)^2 dt + O_p(h) \end{aligned}$$

其中第二个等式源于积分变换，令 $t = (A-a)/h$ ，第三个等式源于泰勒展开。同理，有

$$E\left[K_h(A-a) \middle| X\right] = \pi^*(a, X) + O_p(h^2)$$

从而当 h 充分小时，存在常数 C_0 使得 $|E[K_h(A-a) \middle| X] - \pi^*(a, X)| \leq C_0 h^2$ ，同时存在常数 C_1 使得

$$E\left[K_h(A-a)^2 \middle| X\right] = \frac{C_1}{h} \text{ 以及 } E\left[\left(\pi_K^*(a, X) - K_h(A-a)\right)^2 \middle| X\right] \leq \frac{2C_1}{h}。 \text{ 由于 } \pi_K^*(a, x) = \exp\{R_K(x)' \beta_K^*\},$$

其中 $\beta_K^* \in \tilde{B}_K \subset B_K$ ， $E[R_K(X)' R_K(X)] = K$ ，从而式 (S9) 等式右边的第一项满足

$$\frac{1}{n} E\left[\left(1 - \frac{K_h(A-a)}{\pi_K^*(a, X)}\right)^2 R_K(X)' R_K(X)\right] \leq \frac{2C_1}{nh} \frac{4}{\eta^2} K = \frac{8C_1}{\eta^2} \frac{K}{nh}$$

注意此处的 η 与引理 S1 的证明中出现的 η 相同。

关于式 (S9) 等式右边的第二项，由于

$$\begin{aligned} &\frac{n(n-1)}{n^2} E\left[\left(1 - \frac{K_h(A-a)}{\pi_K^*(a, X)}\right) R_K(X)'\right] E\left[\left(1 - \frac{K_h(A-a)}{\pi_K^*(a, X)}\right) R_K(X)\right] \\ &= \frac{n-1}{n} \left\| E\left[\left(1 - \frac{K_h(A-a)}{\pi_K^*(a, X)}\right) R_K(X)\right] \right\|^2 \\ &= \frac{n-1}{n} \left\| \int_{x \in \mathcal{X}} \frac{\pi_K^*(a, x) - E[K_h(A-a) \middle| x]}{\pi_K^*(a, x)} R_K(x) dF(x) \right\|^2 \\ &\leq \left(\int_{x \in \mathcal{X}} \frac{|\pi_K^*(a, x) - E[K_h(A-a) \middle| x]|}{\pi_K^*(a, x)} \|R_K(x)\| dF(x) \right)^2 \\ &\leq \left(\int_{x \in \mathcal{X}} \frac{|\pi_K^*(a, x) - E[K_h(A-a) \middle| x]|}{\pi_K^*(a, x)} \|R_K(x)\| dF(x) + \int_{x \in \mathcal{X}} \frac{|\pi_K^*(a, x) - \pi^*(a, x)|}{\pi_K^*(a, x)} \|R_K(x)\| dF(x) \right)^2 \\ &\leq 2 \left(\frac{C_0 h^2}{\eta/2} \right)^2 (E\|R_K(X)\|)^2 + 2 \left(\frac{C_5 \xi(K) K^{-s/(2d)}}{\eta/2} \right)^2 (E\|R_K(X)\|)^2 \\ &\leq 2 \left(\frac{C_0 h^2}{\eta/2} \right)^2 E[\|R_K(X)\|^2] + 2 \left(\frac{C_5 \xi(K) K^{-s/(2d)}}{\eta/2} \right)^2 E[\|R_K(X)\|^2] \\ &= \frac{8C_0^2}{\eta^2} h^4 K + \frac{8C_5^2}{\eta^2} \xi(K)^2 K^{-s/d} K \end{aligned}$$

其中，第二个等式源于期望的定义，第二个不等式源于三角不等式，第三个不等式源于引理 S1 并令常数 C_5 满足 $\sup_{x \in \mathcal{X}} |\pi^*(a, x) - \pi_K^*(a, x)| = C_5 K^{-s/(2d)} \xi(K)$ ，以及 $(a+b)^2 \leq 2a^2 + 2b^2$ ，第四个不等式源于 Jensen 不等式，最后一个不等式源于 $E[\|R_K(X)\|^2] = E[R_K(X)'R_K(X)] = K$ 。

因此，基于上述分析有

$$E\left[\left\|\frac{\partial Q_n(\beta_K^*)}{\partial \beta}\right\|^2\right] = \frac{8C_1}{\eta^2} \frac{K}{nh} + \frac{8C_0^2}{\eta^2} h^4 K + \frac{8C_5^2}{\eta^2} \xi(K)^2 K^{1-s/d}$$

基于马尔科夫不等式，有

$$\left\|\frac{\partial Q_n(\beta_K^*)}{\partial \beta}\right\| = O_p\left(\sqrt{\frac{K}{nh}}\right) + O_p(h^2 \sqrt{K}) + O_p(\xi(K) K^{1/2-s/(2d)})$$

$$\text{令 } C^2 = \max\left\{\frac{8C_1}{\eta^2}, \frac{8C_0^2}{\eta^2}, \frac{8C_5^2}{\eta^2}\right\}, \text{ 定义集合}$$

$$B_K^* = \left\{\beta \in \mathbb{R}^K : \|\beta - \beta_K^*\| \leq C_2 C \left(\sqrt{\frac{K}{nh}} + h^2 \sqrt{K} + \xi(K) K^{1/2-s/(2d)}\right)\right\}$$

对于位于 B_K^* 边界上的任意 β ，即 $\|\beta - \beta_K^*\| = C_2 C \left(\sqrt{\frac{K}{nh}} + h^2 \sqrt{K} + \xi(K) K^{1/2-s/(2d)}\right)$ ，有

$$Q_n(\beta_K^*) - Q_n(\beta) = -(\beta - \beta_K^*)' \frac{\partial Q_n(\beta_K^*)}{\partial \beta} - \frac{1}{2}(\beta - \beta_K^*)' \frac{\partial^2 Q_n(\bar{\beta})}{\partial \beta \partial \beta'} (\beta - \beta_K^*) \quad (\text{S10})$$

其中 $\bar{\beta}$ 介于 β 和 β_K^* 之间，且有

$$-\frac{\partial^2 Q_n(\bar{\beta})}{\partial \beta \partial \beta'} = -E_n\left[K_h(A-a) \exp\{-R_K(X)' \bar{\beta}\} R_K(X) R_K(X)'\right]$$

为简化定义，令 $\eta_1 = \inf\left\{\inf_{x \in \mathcal{X}} \exp\{R_K(X)' \beta_K^*\}, \inf_{x \in \mathcal{X}} \exp\{-R_K(X)' \beta_K^*\}\right\}$ 。鉴于

$$\begin{aligned} \sup_{x \in \mathcal{X}, \beta \in B_K^*} \left| \exp\{R_K(X)' \beta_K^*\} - \exp\{R_K(X)' \beta\} \right| &= \sup_{x \in \mathcal{X}, \beta \in B_K^*} \left| \exp\{R_K(X)' \bar{\beta}\} R_K(X)' (\beta_K^* - \beta) \right| \\ &\leq C \xi(K) \left(\sqrt{\frac{K}{nh}} + h^2 \sqrt{K} + \xi(K) K^{1/2-s/(2d)} \right) \end{aligned}$$

其中第一个不等式源于 $\sup_{x \in \mathcal{X}, \beta \in B_K^*} \left| \exp\{R_K(X)' \bar{\beta}\} \right|$ 有界。由引理 S2 的假设可知，上述式子中不等式的右边趋于零，从而当 n 充分大时，有

$$\inf\left\{\inf_{x \in \mathcal{X}, \beta \in B_K^*} \exp\{R_K(X)' \beta\}, \inf_{x \in \mathcal{X}, \beta \in B_K^*} \exp\{-R_K(X)' \beta\}\right\} \geq \frac{\eta_1}{2}$$

同时，经推导可得

$$\begin{aligned}
& E \left[\left\| E_n \left[K_h(A-a) R_K(X) R_K(X)' \right] - E \left[K_h(A-a) R_K(X) R_K(X)' \right] \right\|^2 \right] \\
&= E \left[\text{tr} \left\{ E_n \left[K_h(A-a) R_K(X) R_K(X)' \right] E_n \left[K_h(A-a) R_K(X) R_K(X)' \right] \right\} \right. \\
&\quad \left. - 2 E \left[\text{tr} \left\{ E_n \left[K_h(A-a) R_K(X) R_K(X)' \right] E \left[K_h(A-a) R_K(X) R_K(X)' \right] \right\} \right] \right. \\
&\quad \left. + E \left[\text{tr} \left\{ E \left[K_h(A-a) R_K(X) R_K(X)' \right] E \left[K_h(A-a) R_K(X) R_K(X)' \right] \right\} \right] \right] \\
&= \text{tr} \left[E \left\{ E_n \left[K_h(A-a) R_K(X) R_K(X)' \right] E_n \left[K_h(A-a) R_K(X) R_K(X)' \right] \right\} \right. \\
&\quad \left. - 2 \text{tr} \left[E \left\{ E_n \left[K_h(A-a) R_K(X) R_K(X)' \right] E \left[K_h(A-a) R_K(X) R_K(X)' \right] \right\} \right] \right. \\
&\quad \left. + \text{tr} \left[E \left[K_h(A-a) R_K(X) R_K(X)' \right] E \left[K_h(A-a) R_K(X) R_K(X)' \right] \right] \right] \\
&= \text{tr} \left[E \left\{ \frac{1}{n^2} \sum_{i=1}^n K_h(A_i - a)^2 R_K(X_i) R_K(X_i)' R_K(X_i) R_K(X_i)' \right\} \right. \\
&\quad \left. + \text{tr} \left[E \left\{ \frac{1}{n^2} \sum_{i,j=1; i \neq j}^n K_h(A_i - a) K_h(A_j - a) R_K(X_i) R_K(X_i)' R_K(X_j) R_K(X_j)' \right\} \right] \right. \\
&\quad \left. - \text{tr} \left[E \left[K_h(A-a) R_K(X) R_K(X)' \right] E \left[K_h(A-a) R_K(X) R_K(X)' \right] \right] \right] \\
&= \frac{1}{n} \text{tr} \left\{ E \left[K_h(A-a)^2 R_K(X) R_K(X)' R_K(X) R_K(X)' \right] \right\} - \frac{1}{n} \text{tr} \left\{ E \left[K_h(A-a) R_K(X) R_K(X)' \right] E \left[K_h(A-a) R_K(X) R_K(X)' \right] \right\} \\
&\leq \frac{1}{n} \text{tr} \left\{ E \left[K_h(A-a)^2 R_K(X) R_K(X)' R_K(X) R_K(X)' \right] \right\} \\
&\leq \frac{C_1}{nh} \xi(K)^2 K
\end{aligned}$$

定义事件 E_n 如下所示,

$$E_n = \left\{ \beta : (\beta - \beta_K^*)' (E \left[K_h(A-a) R_K(X) R_K(X)' \right] - \frac{\eta}{4} I_K) (\beta - \beta_K^*) < (\beta - \beta_K^*)' E_n \left[K_h(A-a) R_K(X) R_K(X)' \right] (\beta - \beta_K^*) \right\}$$

由切比雪夫不等式可知

$$\begin{aligned}
& P(|(\beta - \beta_K^*)' E_n \left[K_h(A-a) R_K(X) R_K(X)' \right] (\beta - \beta_K^*) - (\beta - \beta_K^*)' E \left[K_h(A-a) R_K(X) R_K(X)' \right] (\beta - \beta_K^*)| \\
&\quad \geq \frac{\eta}{4} \|\beta - \beta_K^*\|^2, \beta \neq \beta_K^*) \\
&\leq \frac{16 \|\beta - \beta_K^*\|^4 E \left\| E_n \left[K_h(A-a) R_K(X) R_K(X)' \right] - E \left[K_h(A-a) R_K(X) R_K(X)' \right] \right\|^2}{\eta^2 \|\beta - \beta_K^*\|^4} \\
&= \frac{16}{\eta^2} \frac{C_1}{nh} \xi(K)^2 K
\end{aligned}$$

由于 $\frac{\xi(K)^2 K}{nh} \rightarrow 0$, 从而对于任意 $\varepsilon > 0$, 存在 $n_0(\varepsilon)$ 使得当 $n > n_0(\varepsilon)$ 时, 有

$$P((E_n)^c) < \frac{\varepsilon}{2} \quad (\text{S11})$$

选定 $C_2 > \sqrt{\frac{\varepsilon}{2}} \frac{16}{\eta_1 \eta}$ 。当 n 充分大时, 对于任意 $\beta \in E_n$, 如果 β 位于 B_K^* 的边界上, 即有

$$\|\beta - \beta_K^*\| = C_2 C \left(\sqrt{\frac{K}{nh}} + h^2 \sqrt{K} + \xi(K) K^{1/2-s(2d)} \right), \text{ 基于式 (S10) 以及 (S11), 有}$$

$$\begin{aligned}
Q_n(\beta_K^*) - Q_n(\beta) &= -(\beta - \beta_K^*)' \frac{\partial Q_n(\beta_K^*)}{\partial \beta} - \frac{1}{2}(\beta - \beta_K^*)' E_n \left[K_h(A - a) \exp\{-R_K(X)' \bar{\beta}\} R_K(X) R_K(X)' \right] (\beta - \beta_K^*) \\
&\leq -(\beta - \beta_K^*)' \frac{\partial Q_n(\beta_K^*)}{\partial \beta} - \frac{1}{2} \frac{\eta_1}{2} (\beta - \beta_K^*)' \left(E \left[K_h(A - a) R_K(X) R_K(X)' \right] - \frac{\eta}{4} I_K \right) (\beta - \beta_K^*) \\
&\leq \left\| \frac{\partial Q_n(\beta_K^*)}{\partial \beta} \right\| \left\| \beta - \beta_K^* \right\| - \frac{1}{2} \frac{\eta_1}{2} \frac{\eta}{4} \left\| \beta - \beta_K^* \right\|^2 \\
&\leq \left(\left\| \frac{\partial Q_n(\beta_K^*)}{\partial \beta} \right\| - \frac{\eta_1 \eta}{16} \left\| \beta - \beta_K^* \right\| \right) \left\| \beta - \beta_K^* \right\|
\end{aligned} \tag{S12}$$

其中第二个不等式源于，当 h 充分小时有

$$\frac{\eta}{2} I_K \leq E \left[K_h(A - a) R_K(X) R_K(X)' \right] \leq \frac{2}{\eta} I_K$$

同时，基于马尔科夫不等式，有

$$\begin{aligned}
P \left(\left\| \frac{\partial Q_n(\beta_K^*)}{\partial \beta} \right\| \geq \frac{\eta_1 \eta}{16} \left\| \beta - \beta_K^* \right\| \right) &\leq \frac{E \left[\left\| \frac{\partial Q_n(\beta_K^*)}{\partial \beta} \right\|^2 \right] 16^2}{\eta_1^2 \eta^2 \left\| \beta - \beta_K^* \right\|^2} \\
&\leq \frac{C^2 \left(\frac{K}{nh} + h^4 K + \xi(K)^2 K^{1-s/d} \right) 16^2}{\eta_1^2 \eta^2 \frac{2}{\varepsilon} \frac{16^2}{\eta_1^2 \eta^2} C^2 \left(\sqrt{\frac{K}{nh}} + h^2 \sqrt{K} + \xi(K) K^{1/2-s/(2d)} \right)^2} \\
&\leq \frac{\varepsilon}{2}
\end{aligned} \tag{S13}$$

综合式 (S11) 和 (S13)，有

$$P \left((E_n)^c \text{ 或者 } \left\| \frac{\partial Q_n(\beta_K^*)}{\partial \beta} \right\| \geq \frac{\eta_1 \eta}{8(2-\eta_1)} \left\| \beta - \beta_K^* \right\| \right) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

从而有

$$P \left(E_n \text{ 且 } \left\| \frac{\partial Q_n(\beta_K^*)}{\partial \beta} \right\| < \frac{\eta_1 \eta}{8(2-\eta_1)} \left\| \beta - \beta_K^* \right\| \right) > 1 - \varepsilon \tag{S14}$$

基于式 (S12) 和 (S14) 可得，当 β 位于 B_K^* 的边界上时，有

$$P(Q_n(\beta_K^*) - Q_n(\beta) < 0) > 1 - \varepsilon$$

基于 $Q_n(\beta)$ 的连续性， $Q_n(\beta)$ 在紧致集 B_K^* 上必然存在局部最小值点，且该局部最小值点为 B_K^* 的内点。由于 $Q_n(\beta)$ 为严格凸函数，全局最小值点为 $\hat{\beta}_K$ ，从而得证。 **Q.E.D**

(4) 定理 1 的证明

由引理 S2 以及 $\xi(K) = O_p(K)$ 可知，

$$\begin{aligned}
|\hat{\pi}_K(a, x) - \pi_K^*(a, x)| &= \left| \exp \{R_K(x)' \hat{\beta}_K\} - \exp \{R_K(x)' \beta_K^*\} \right| \\
&\leq C \|R_K(x)\| \cdot \|\hat{\beta}_K - \beta_K^*\| \\
&= O_p(\xi(K)) \cdot \left(O_p\left(\sqrt{\frac{K}{nh}}\right) + O_p(h^2 \sqrt{K}) + O_p(\xi(K) K^{1/2-s/(2d)}) \right) \\
&= O_p\left(\frac{K^{3/2}}{\sqrt{nh}}\right) + O_p(h^2 K^{3/2}) + O_p(K^{5/2-s/(2d)})
\end{aligned}$$

从而有

$$\begin{aligned}
|\hat{\pi}_K(a, x) - \pi^*(a, x)| &\leq |\hat{\pi}_K(a, x) - \pi_K^*(a, x)| + |\pi_K^*(a, x) - \pi^*(a, x)| \\
&\leq O_p\left(\frac{K^{3/2}}{\sqrt{nh}}\right) + O_p(h^2 K^{3/2}) + O_p(K^{5/2-s/(2d)}) + O_p(K^{-s/(2d)} \xi(K)) \\
&= O_p\left(\frac{K^{3/2}}{\sqrt{nh}}\right) + O_p(h^2 K^{3/2}) + O_p(K^{5/2-s/(2d)})
\end{aligned}$$

相应地，有

$$\begin{aligned}
|\hat{\tau}_K(a) - \tau(a)| &= \left| \frac{1}{n} \sum_{i=1}^n \frac{K_h(A_i - a) Y_i}{\hat{\pi}_K(a, X_i)} - \tau(a) \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^n \frac{K_h(A_i - a) Y_i}{\hat{\pi}_K(a, X_i)} - \frac{1}{n} \sum_{i=1}^n \frac{K_h(A_i - a) Y_i}{\pi^*(a, X_i)} \right| + \left| \frac{1}{n} \sum_{i=1}^n \frac{K_h(A_i - a) Y_i}{\pi^*(a, X_i)} - \tau(a) \right| \\
&:= M_1 + M_2
\end{aligned}$$

首先考察 M_2 。关于 M_2 ，有

$$\begin{aligned}
M_2 &= \left| \frac{1}{n} \sum_{i=1}^n \frac{K_h(A_i - a) Y_i}{\pi^*(a, X_i)} - \tau(a) \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^n \frac{K_h(A_i - a) Y_i}{\pi^*(a, X_i)} - E \left[\frac{K_h(A - a) Y}{\pi^*(a, X)} \right] \right| + \left| E \left[\frac{K_h(A - a) Y}{\pi^*(a, X)} \right] - \tau(a) \right| \\
&= O_p\left(\frac{1}{\sqrt{nh}}\right) + O_p(h^2)
\end{aligned}$$

其中，第三个等式可由非参数估计的收敛速度直接得到。

接着考察 M_1 。关于 M_1 ，有

$$\begin{aligned}
M_1 &= \left| \frac{1}{n} \sum_{i=1}^n \frac{K_h(A_i - a) Y_i}{\hat{\pi}_K(a, X_i)} - \frac{1}{n} \sum_{i=1}^n \frac{K_h(A_i - a) Y_i}{\pi^*(a, X_i)} \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^n \frac{K_h(A_i - a) Y_i (\hat{\pi}_K(a, X_i) - \pi^*(a, X_i))}{\hat{\pi}_K(a, X_i) \pi^*(a, X_i)} \right| \\
&\leq \frac{C}{n} \sum_{i=1}^n |K_h(A_i - a) Y_i| |\hat{\pi}_K(a, X_i) - \pi^*(a, X_i)|
\end{aligned}$$

其中第一个不等式源于 $\hat{\pi}_K(a, x)$ 和 $\pi^*(a, x)$ 在 $x \in X$ 均远离 0 和 1。接着考察上式中的

$\frac{1}{n} \sum_{i=1}^n |K_h(A_i - a)Y_i|$ ，由于

$$\begin{aligned} E \left[\frac{1}{n} \sum_{i=1}^n |K_h(A_i - a)Y_i| \right] &= E \left[|K_h(A - a)Y| \right] \\ &\leq \sqrt{E(K_h(A - a)Y)^2} \\ &= \sqrt{E \left\{ E[K_h(A - a)^2 Y^2 | A, X] \right\}} \\ &= \sqrt{E \left\{ K_h(A - a)^2 E[Y^2 | A, X] \right\}} \\ &= \sqrt{E \left\{ K_h(A - a)^2 \{ \mu(A, X)^2 + \sigma(A, X)^2 \} \right\}} \\ &= \sqrt{E \left[E \left\{ K_h(A - a)^2 \{ \mu(A, X)^2 + \sigma(A, X)^2 \} | X \right\} \right]} \\ &= \sqrt{E \left\{ \frac{1}{h} \int_{t \in T} K(t)^2 d\pi^*(a, X) \{ \mu(a, X)^2 + \sigma(a, X)^2 \} + O_p(h) \right\}} \\ &= O_p \left(\frac{1}{\sqrt{h}} \right) \end{aligned}$$

其中，第一个不等式源于 Jensen 不等式，第二和第五个等式源于迭代期望定理，第六个等式源于

泰勒展开。从而基于马尔科夫不等式有 $\frac{1}{n} \sum_{i=1}^n |K_h(A_i - a)Y_i| = O_p \left(\frac{1}{\sqrt{h}} \right)$ 。相应地，关于 M_1 ，有

$$\begin{aligned} M_1 &\leq \frac{C}{n} \sum_{i=1}^n |K_h(A_i - a)Y_i| |\hat{\pi}_K(a, X_i) - \pi^*(a, X_i)| \\ &= O_p \left(\frac{1}{\sqrt{h}} \right) \cdot \left\{ O_p \left(\frac{K^{3/2}}{\sqrt{nh}} \right) + O_p(h^2 K^{3/2}) + O_p(K^{5/2-s/(2d)}) \right\} \\ &= O_p \left(\frac{K^{3/2}}{\sqrt{nh^2}} \right) + O_p(h^{3/2} K^{3/2}) + O_p \left(\frac{K^{5/2-s/(2d)}}{\sqrt{h}} \right) \end{aligned}$$

为考察 $M_1 + M_2$ 的收敛性，分两种情况进行分析，分别为 $h \leq O(n^{-1/5})$ 和 $h > O(n^{-1/5})$ 。令 $K = O(n^{c_1})$ ， $h = O(n^{-c_2})$ ，其中 c_1 和 c_2 均大于零。

当 $h \leq O(n^{-1/5})$ 时，即 $c_2 \geq \frac{1}{5}$ ， $M_1 \leq O_p \left(\frac{K^{3/2}}{\sqrt{nh^2}} \right) + O_p \left(\frac{K^{5/2-s/(2d)}}{\sqrt{h}} \right)$ 。为使得 $M_1 + M_2$ 收敛，需满足

$$\left\{ \begin{array}{l} s/d \geq 4 \\ \xi(K)^2 K / (nh) \rightarrow 0, \xi(K)^2 K h^4 \rightarrow 0, \xi(K)^4 K^{1-s/d} \rightarrow 0 \\ \frac{1}{\sqrt{nh}} \rightarrow 0, h^2 \rightarrow 0 \\ \frac{K^{3/2}}{\sqrt{nh^2}} \rightarrow 0, \frac{K^{5/2-s/(2d)}}{\sqrt{h}} \rightarrow 0 \end{array} \right. \quad (S15)$$

其中，第一个式子源于引理 S1 的假设，第二个式子源于引理 S2 的假设，第三个式子源于基于 M_2 的不等式，第四个式子源于基于 M_1 的不等式。通过求解不等式可知，当满足如下不等式 (S16) 时，式 (S15) 能够被满足。注意，式 (S16) 是式 (S15) 的充分条件，非必要条件。

$$\left\{ \begin{array}{l} s/d \geq 7 \\ \frac{c_2}{s/d-5} < c_1 < \frac{1}{3} - \frac{2}{3}c_2 \\ \frac{1}{5} \leq c_2 < \frac{1}{3/(s/d-5)+2} \end{array} \right. \quad (\text{S16})$$

当 $h > O(n^{-1/5})$ 时，即 $0 < c_2 < \frac{1}{5}$ ， $M_1 \leq O_p(h^{3/2}K^{3/2}) + O_p\left(\frac{K^{5/2-s/(2d)}}{\sqrt{h}}\right)$ 。为使得 $M_1 + M_2$ 收敛，需满足

$$\left\{ \begin{array}{l} s/d \geq 4 \\ \xi(K)^2 K / (nh) \rightarrow 0, \xi(K)^2 K h^4 \rightarrow 0, \xi(K)^4 K^{1-s/d} \rightarrow 0 \\ \frac{1}{\sqrt{nh}} \rightarrow 0, h^2 \rightarrow 0 \\ h^{3/2} K^{3/2} \rightarrow 0, \frac{K^{5/2-s/(2d)}}{\sqrt{h}} \rightarrow 0 \end{array} \right. \quad (\text{S17})$$

其中，第一个式子源于引理 S1 的假设，第二个式子源于引理 S2 的假设，第三个式子源于基于 M_2 的不等式，第四个式子源于基于 M_1 的不等式。通过求解不等式可知，当满足如下不等式 (S18) 时，式 (S17) 能够被满足。注意，式 (S18) 是式 (S17) 的充分条件，非必要条件。

$$\left\{ \begin{array}{l} s/d \geq 7 \\ \frac{c_2}{s/d-5} < c_1 < c_2 \\ 0 < c_2 < \frac{1}{5} \end{array} \right. \quad (\text{S18})$$

从而得证。

Q.E.D

附录 4：变量的描述性统计结果

变量名称	样本组 1					样本组 2				
	均值	中位数	标准差	最小值	最大值	均值	中位数	标准差	最小值	最大值
家庭储蓄率	-0.051	0.156	0.727	-2.000	0.750	-0.036	0.168	0.722	-2.000	0.750
户主年龄	45.496	46.000	8.890	25.000	60.000	48.963	49.000	11.012	25.000	70.000
户主性别	0.564	1.000	0.496	0.000	1.000	0.581	1.000	0.493	0.000	1.000
户主受教育年限	7.626	9.000	4.418	0.000	22.000	7.274	9.000	4.495	0.000	22.000
户主婚姻状况	0.931	1.000	0.253	0.000	1.000	0.919	1.000	0.272	0.000	1.000
户主健康状况	0.847	1.000	0.360	0.000	1.000	0.827	1.000	0.378	0.000	1.000
户主工作状况	0.755	1.000	0.430	0.000	1.000	0.712	1.000	0.453	0.000	1.000

是否城镇家庭	0.452	0.000	0.498	0.000	1.000	0.454	0.000	0.498	0.000	1.000
是否自有住房	0.877	1.000	0.329	0.000	1.000	0.879	1.000	0.326	0.000	1.000
家庭老人数量	0.199	0.000	0.483	0.000	4.000	0.317	0.000	0.612	0.000	4.000
家庭少儿数量	0.996	1.000	1.070	0.000	5.000	0.959	1.000	1.076	0.000	5.000
家庭规模	4.379	4.000	1.762	1.000	10.000	4.321	4.000	1.850	1.000	10.000
家庭消费支出	10.313	10.317	0.800	5.193	13.397	10.283	10.292	0.813	5.193	13.397
地区经济发展	10.612	10.597	0.491	9.482	11.851	10.635	10.607	0.494	9.482	11.851

附录 5：户主年龄影响储蓄率的完整回归结果

变量	储蓄率	
	样本组 1	样本组 2
户主年龄	-0.023*** (-6.356)	-0.010*** (-3.792)
户主年龄平方/100	0.029*** (7.001)	0.013*** (5.002)
户主性别	-0.062*** (-8.501)	-0.059*** (-9.013)
受教育年限	0.024*** (27.679)	0.025*** (31.541)
婚姻状况	0.058*** (4.389)	0.042*** (3.815)
健康状况	0.094*** (9.481)	0.101*** (11.982)
工作状况	0.054*** (6.557)	0.017** (2.300)
是否城镇家庭	0.103*** (13.299)	0.118*** (16.688)
是否自有住房	0.041*** (4.038)	0.046*** (4.886)
家庭老人数量	-0.048*** (-6.412)	-0.056*** (-9.982)
家庭少儿数量	-0.076*** (-15.890)	-0.075*** (-17.015)
家庭规模	0.084*** (28.475)	0.091*** (35.903)
家庭消费支出	-0.482*** (-97.357)	-0.467*** (-106.315)
地区经济发展	0.000 (-0.013)	(0.015) (-0.448)
年份	控制	控制
省份	控制	控制
样本量	38818	47305
R ²	0.216	0.211

注：***、**和*分别表示 1%、5%、10%的显著性水平，括号内为稳健 t 值。由于户主年龄平方的系数估计值较小，此处将户主年龄平方除以 100。