arXiv:2602.01631v1 [stat.ME] 2 Feb 2026

网络上的局部依赖下的双重差分*

Akihiro Sato[1] 和 Shonosuke Sugasawa[2]

[1]经济学研究生院，庆应义塾大学

[2]经济学部，庆应义塾大学

摘要

在干扰情形下估计因果效应十分重要，因为此时稳定个体处理值假设不成立，这一问题在区域经济学和公共经济学等领域尤为关键。现有关于干扰情形下因果推断的大部分研究都依赖于预先设定的“暴露映射”。本文聚焦于双重差分，针对观测网络上的局部干扰，提出了一种用于识别直接平均处理效应和间接平均处理效应的非参数识别策略。具体而言，我们提出了间接效应这一新概念，用于衡量处理的总体向外影响。基于给定邻域处理向量的平行趋势假设，我们构建了逆概率加权估计量和双重稳健估计量。我们建立了这些估计量的渐近性质，包括在满足一定正则条件时，即使干扰参数模型设定错误仍具有一致性。模拟研究和实证应用表明，所提出方法具有良好的有效性。

关键词：因果推断；暴露映射；网络效应；溢出效应；未知相互作用

*2026年2月3日

1 引言

估计干预的因果效应仍是实证研究中的核心挑战。双重差分（DID）框架长期以来一直是实现这一目标的基石，但其基本假设近来受到了广泛关注。近期文献表明，在处理采纳时间交错的情形下，由于先前接受处理的单位污染了对照组，传统的双重差分估计量可能会产生严重偏差（例如 Goodman-Bacon，2021年；Callaway 和 Sant'Anna，2021年；Sun 和 Abraham，2021年；de Chaisemartin 和 D'Haultfœuille，2020年）。网络上同时接受处理的单位产生的溢出效应，是对照组受到污染的另一种类似但不同的来源。在许多应用场景中，例如评估基于地点的政策或营销活动，处理效应并不局限于处理单元，还会溢出至其他单位。这种现象称为"干扰"（Cox，1958年），违反了稳定单元处理值假设（SUTVA；Rubin，1980年）——这一假设是因果推断的一项基本前提。

存在干扰时，对照组无法作为未接受处理这一情形的有效反事实。因此，简单比较处理组与对照组会导致对真实因果效应的有偏估计（例如，Perez-Heydrich et al., 2014; Butts, 2023）。这种偏差可能十分严重，甚至可能反转效应的符号，从而导致根本上有误的政策含义（Leung, 2024）。然而，由于必须考虑 2^N种可能的处理分配和 N 个体，干扰情形下的因果推断本质上颇具挑战性。若不施加结构性假设，几乎不可能识别有意义的因果效应（Manski, 2013）。探讨这一问题的早期重要研究包括有关社会互动（例如，Manski, 1993）和疫苗效力（例如，Hudgens and Halloran, 2008）的研究。

一种更灵活的基础是这样一个概念：个体的结果变量可能取决于整个处理分配向量，Manski（2013）将其形式化为有效处理。Aronow and Samii（2017）提出了一种系统性框架，称为暴

露映射，它将高维处理向量概括为低维的“暴露水平”。研究者可以据此定义干扰，并估计给定暴露水平下的效应（例如，Ugander 等人，2013；Basse 和 Feller，2018）。一项主要的实践局限在于必须预先指定干扰结构（例如，触发效应所需的受处理邻居数量），而且必须事先了解并正确设定该结构。然而，目前仍缺少一种灵活的双重差分框架，能够在不预先指定干扰结构的情况下实现识别。但社会和经济干扰往往复杂且具有非线性，因此研究者很少能预先知道正确的函数形式。依赖已知的暴露映射会带来重大风险；在复杂的现实世界网络中，映射设定错误可能导致因果效应估计出现严重偏差。近期理论研究强调了正确设定该映射的难度及设定错误的后果，例如 S¨avje（2024）探讨了设定错误情况下实现识别的条件，并凸显了这一问题的严重性。尽管网络因果推断领域近期取得了进展，例如 Liu 等人（2025）提出的自动双重稳健框架利用马尔可夫随机场和自回归模型成功处理了一般干扰，但这些方法通常依赖于对干扰函数的特定参数结构。相比之下，我们的方法利用面板数据的时间维度，以非参数方式识别因果效应，从而避免对溢出机制作出结构性假设；当该机制不为人所充分理解时，这提供了一种更稳健的替代方案（例如，Xu，2025）。

本文提出了一种新颖的识别策略，利用面板数据结构，在不施加已知暴露映射的情况下实现非参数识别。在条件平行趋势假设下，我们对邻域处理向量进行条件化；该方法无需对溢出效应结构作出函数形式假设，即可实现识别。估计依赖于灵活的方法（逆概率加权和双重稳健方法），即使在对该向量进行条件化时仍然可行。

近期研究将干扰纳入双重差分框架。早期空间 al

计量经济学研究提出了空间双重差分模型，用于考虑来自地理位置相近单位的溢出效应（Delgado和Florax，2015）。与我们的研究更为密切相关的是，近期一系列研究将这一整合拓展至一般网络结构。尽管Xu（2025）和Butts（2023）等研究作出了重要贡献，但它们通常依赖预先设定的暴露映射。我们的研究与近期放宽关于干扰函数形式的强假设这一趋势相一致（例如，Leung，2022）。我们通过将一种非参数识别策略——以邻域处理向量本身为条件——嵌入双重差分框架，为这一不断发展的文献作出贡献；在观察性情境下，双重差分框架仍是政策评估中使用最广泛的方法。

我们的一项独特贡献是对间接效应的定义。许多文献（包括 Xu（2025））将间接效应（或溢出效应）定义为个体自身的结果变量随暴露水平变化而产生的差异——内向溢出效应：“如果我所在邻域的处理状态不同，我的结果变量会发生怎样的变化？”（另见 Tchetgen Tchetgen 和 VanderWeele，2012）。而我们将处理组平均间接处理效应（AITT）定义为对某个个体实施处理对其邻居结果变量的影响，即外向溢出效应：“对个体 i进行干预后，其邻居 j 的结果变量平均会发生多大变化？”AITT 直接量化政策外部性，有别于中介效应（例如 Sobel，2006）。

我们提出了实现这一识别的逆概率加权（以及双重稳健）估计量，并建立了相应的渐近理论。在推断方面，我们采用 Kojevnikov等人（2021）的 ψ-依赖框架；该框架在观测值依赖于单一大型网络时，可在平衡网络密度与相依衰减的条件下得到渐近正态性；网络依赖下不适用标准的独立同分布渐近理论。与正确设定的暴露映射相比，对邻域处理向量进行条件化可能会降低效率（维数灾难），尤其是在稠密网络或小样本情形下。因此，我们的方法具有互补性：当干扰结构已知时，为提高效率，优先采用暴露映射；当干扰结构未知时，我们的非参数识别

结合灵活的参数估计，提供了一种稳健的替代方案。

本文结构安排如下。第 2 节介绍识别框架。第 3 节提出估计量和推断方法。第4节确立估计量的渐近性质。第 5 节报告模拟证据，第 6 节介绍一项实证应用，第 7 节给出总结性评论。展示所提方法的 Python 代码可在 GitHub 仓库中获取（ https://github. com/S10aki95/DIDUnderLocalDependenceOnNetworks）。

2 识别

我们定义设定、潜在结果和因果参数（ADTT 和 AITT），陈述识别假设——包括将平行趋势扩展至网络的假设——并确立主要识别结果。

2.1 设定与潜在结果

我们考虑由 n个个体 $i \in I_n$组成的总体，这些个体通过邻接矩阵A所表示的网络相连，其中 $l_A(i,\ j)$ 表示最短路径长度。数据为两期面板（$t = 1,\ 2$），包含处理指示变量 D_i 和结果变量 Y_{ti}；推断以固定属性z_i 为条件（有限总体视角）。我们考虑存在干扰（溢出效应）的情形，此时稳定单元处理值假设（SUTVA）不成立。具体而言，我们假设干扰是局部的，并沿网络结构发生。为明确起见，我们将个体 i的邻域记为 N_i，定义为该个体在网络中的 L 个最近邻，其中 L 是由研究者选定的固定数值；这样可以使估计所用条件集的维度保持固定。具体而言，对于每个个体 i，我们按照网络距离 $l_A(i,\ j)$ 从小到大对所有个体 $j \in I_n \setminus \{i\}$ 排序，并取前 L 个个体组成邻域 N_i。令 $\boldsymbol{D}_{N_i} \in \{0,1\}^L$ 表示这些邻居的处理向量，并按照其与个体 i的距离排序。也就是说，对于 $k = 1,\ \ldots,\ L,\ k$的第 $\boldsymbol{D}_{N_i}$ 个分量对应于N_i中第 k近邻的处理指示变量。

对于第4节中提出的渐近理论，我们假设干扰仅限于距离 K以内，其中 K 表示理论干扰范围（个体之间能够相互影响的最大网络距离）。我们将第 i个个体在距离 K 以内的邻域定义为

$$N(i;K) = \{j \in I_n \setminus \{i\} \mid l_A(i,j) \leq K\}.$$

对于任意个体 j，我们还将 $N(j;K)^+ = N(j;K) \cup \{j\}$定义为包含个体 j自身的扩展邻域。我们还将 $N^\partial(i;s) = \{j \in I_n \mid l_A(i,j) = s\}$定义为与个体 i距离恰好为 s 的单位集合。在考虑排除个体 i自身的个体j邻域的处理向量时，我们将其记为 $\boldsymbol{d}^{+,-i}_{N(j;K)} \in \{0,1\}^{|N(j;K)^+\setminus\{i\}|}$，它表示 j的扩展邻域中除个体 i之外所有单位的处理分配向量。

在这一设定下，我们将 $y_{ti}(d_i, \boldsymbol{d}_{N_i})$ 定义为个体 i 在其自身处理状态固定为 $d_i \in \{0,1\}$ 且其邻域处理向量固定为 $\boldsymbol{d}_{N_i} \in \{0,1\}^L$时的潜在结果。这一表述是一般性的，并涵盖了诸如 $y_{2i}(d_i, \boldsymbol{d}_{N_i}) = m(Y_{1i}, d_i, \boldsymbol{d}_{N_i}, \epsilon_i)$ 的结构，其中 ϵ_i 为误差项。该表述的一个关键特征是，它允许邻域处理的整个组合 $\boldsymbol{d}_{N_i}$影响结果变量，而不要求采用类似暴露映射的特定函数形式；文献中通常会作此假设（Hudgens and Halloran，2008年；Ugander 等人，2013年）。基于 L最近邻的固定长度向量 $\boldsymbol{d}_{N_i}$ 使条件集维度在不同个体之间保持不变。

2.2 因果估计量

在上一节所述的设定下，我们首先定义直接效应。**定义 1.** 受处理者的平均直接处理效应（*ADTT*）定义为

$$\tau_{\text{dir}} = \frac{1}{n}\sum_{i=1}^{n} E[y_{2i}(1, \boldsymbol{D}_{N_i}) - y_{2i}(0, \boldsymbol{D}_{N_i}) \mid D_i = 1, z_i],$$

该指标刻画了受处理子群体中处理的预期直接效应，并对观测到的邻域处理向量进行边际化。期望 $E[\cdot]$ 是针对个体 i的随机不可观测因素，条件是其处理状态和固定属性z_i。

这一界定确保了估计目标在存在干扰时定义明确；因此，τ_{dir} 被解释为在固定 $\boldsymbol{D}_{N_i}$ 的情况下，处理单元在 D_i =1与 $D_i = 0$之间的期望结果变量差异。该表述在文献中十分常见（例如，Aronow and Samii，2017；Xu，2025）。虽然估计目标的定义相同，但识别与估计策略存在显著差异。以往研究通常依赖暴露映射，将高维的处理分配向量缩减为低维的暴露水平（例如，Hudgens and Halloran，2008；Aronow and Samii，2017；Xu，2025）。我们在 ADTT 方面的主要贡献是，无须假设已知暴露映射即可实现非参数识别。为此，我们直接对观测到的邻域处理向量$\boldsymbol{D}_{N_i}$进行条件化。这避免了暴露映射中固有的设定错误风险，并使直接效应得以非参数识别。

对于间接效应，我们引入处理组平均间接处理效应（AITT），其不同于以往研究中常用的定义。

定义 2. 处理组平均间接处理效应（*AITT*），定义为

$$\tau_{\text{ind}} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{|N_i|}\sum_{j\in N_i} E\Big[\Delta y_{2j}|D_i=1,z_i,z_j\Big]\right),$$

用于量化对个体 i 进行处理所产生的邻居平均外部性，其中 $\Delta y_{2j} = y_{2j}(D_j, \boldsymbol{D}_{N_j}^{(1)}) - y_{2j}(D_j, \boldsymbol{D}_{N_j}^{(0)})$，而 $\boldsymbol{D}_{N_j}^{(d)}$ 是个体 j 的邻居处理向量，且 $D_i = d \in \{0,1\}$。期望 $E[\cdot]$ 是针对邻居 j 的随机不可观测因素计算的，并以 z_i 和 z_j 为条件。

内层求和捕捉对每个邻居的影响；外层汇总则是对个体 i实施处理所产生的“溢出效应范围”——干预产生的全部向外影响。

这一界定不同于文献中常见的间接效应（或溢出效应）定义（例如，Tchetgen Tchetgen和VanderWeele，2012年；Xu，2025年）。现有定义通常关注邻域暴露的变化对个体自身结果变量的影响（即："如果邻居的处理状态不同，我的结果变量会如何变化？"）。而我们的AITT捕捉的是对个体 i实施处理所产生的外向溢出效应。

2.3 识别结果

为识别这些参数，我们需要对未观测到的反事实作出假设。我们将双重差分中的标准平行趋势假设扩展，以适应网络干扰。

假设1（无预期效应）。 $E[y_{1i}|\boldsymbol{D}_{I_n}, z_i] = y_{1i}(z_i)$。

假设2（干扰下的条件平行趋势）。对于所有 $i \in I_n$ 和 $\boldsymbol{d}_{N_i} \in\{0,1\}^L$，均有

$$
\begin{aligned}
&E[y_{2i}(0, \boldsymbol{d}_{N_i}) - y_{1i}|D_i = 1, \boldsymbol{D}_{N_i} = \boldsymbol{d}_{N_i}, z_i] \\
&= E[y_{2i}(0, \boldsymbol{d}_{N_i}) - y_{1i}|D_i = 0, \boldsymbol{D}_{N_i} = \boldsymbol{d}_{N_i}, z_i]
\end{aligned}
$$

此外，对于所有 $j \in N_i$ 和 $\boldsymbol{d}_{N_j}^{-i} \in \{0,1\}^{L-1}$，均有

$$
\begin{aligned}
&E[y_{2j}(D_i = 0, D_j, \boldsymbol{D}_{N_j}^{-i} = \boldsymbol{d}_{N_j}^{-i}) - y_{1j} \mid D_i = 1, D_j, \boldsymbol{D}_{N_j}^{-i} = \boldsymbol{d}_{N_j}^{-i}, z_i, z_j] \\
&= E[y_{2j}(D_i = 0, D_j, \boldsymbol{D}_{N_j}^{-i} = \boldsymbol{d}_{N_j}^{-i}) - y_{1j} \mid D_i = 0, D_j, \boldsymbol{D}_{N_j}^{-i} = \boldsymbol{d}_{N_j}^{-i}, z_i, z_j]
\end{aligned}
$$

假设3（L 邻域的充分性与局部性）。 我们假设，所选邻居数量 L 足以捕捉相关干扰。此外，为确保基于网络距离推导出的渐近性质成立，我们假设，对于所有 $i \in I_n$ 和所有 $j \in N_i$，网络距离均满足 $l_A(i, j) \leq K$。也就是说， L-邻域 N_i 包含在理论干扰范围 K之内：

$N_i \subseteq \{j \in I_n \setminus \{i\} \mid l_A\ (i,\ \ j)\ \leq K\}$。

假设1称为“无预期假设”，要求处理前第一个时期的潜在结果不取决于处理。假设2指出，在给定相关邻域内所有其他个体的处理状态后，无处理情况下的预期结果趋势（$D_i = 0$）不受个体 i实际处理状态的影响。这是对双重差分分析中常用的标准条件平行趋势假设（含协变量）的自然扩展（Abadie，2005）。在标准条件平行趋势假设中，假定在给定一组观测协变量的条件下，平行趋势成立。在此，我们将这一思路扩展到网络环境，通过对邻域处理向量 $\boldsymbol{D}_{N_i}$进行条件化。通过控制邻域处理的确切配置，我们以关于反事实趋势的假设取代关于干扰函数的结构性假设，从而无需暴露映射即可实现识别。虽然这一假设强于标准无条件平行趋势假设，但它与近期其他关于网络干扰下非参数识别的研究所采用的假设类似。例如，Leung（2022）采用了类似的条件化策略（对邻域协变量和处理分配进行条件化），以在横截面环境中实现识别。

假设3弥合了理论干扰范围 K 与采用固定数量邻居 L的实际实施之间的差距。它确保所选邻域 N_i 涵盖所有相关干扰效应（充分性），并确保基于网络距离的渐近理论仍然成立（局部性）。具体而言，通过要求 L-邻域 N_i 中的所有个体与个体 i之间的距离均不超过 K，该假设保证基于 L-邻域的估计量满足在理论干扰范围 K下推导出的混合条件和弱依赖性质。在实践中，当 L 相对于网络结构和预期干扰范围足够大时，这一假设是合理的。

在假设 1、2 和 3 下，ADTT 和 AITT 均可按照以下

定理识别。证明见附录。

定理 1。 在假设 *1*、*2* 和 *3* 下，*ADTT* 可识别为

$$\tau_{\text{dir}} = \frac{1}{n}\sum_{i=1}^{n} E\left[\frac{D_i - e_i(\boldsymbol{D}_{N_i}, z_i)}{\pi_i(z_i)\{1 - e_i(\boldsymbol{D}_{N_i}, z_i)\}}(Y_{2i} - Y_{1i}) \mid z_i\right],$$

其中 $\pi_i(z_i) \equiv P(D_i = 1|z_i)$ 是给定z_i时 $D_i = 1$ 的边际概率，而 $e_i(\boldsymbol{D}_{N_i}, z_i) = P(D_i = 1|\boldsymbol{D}_{N_i}, z_i)$ 是给定邻域处理向量 $\boldsymbol{D}_{N_i}$ 和 z_i时的倾向得分。注意， e_i 同时是 $\boldsymbol{D}_{N_i}$ 和 z_i的函数。该加权方案不仅校正了个体处理选择，还校正了邻域处理的具体配置，从而确保比较具有相似局部干扰暴露程度的单位。权重 e_i/π_i 根据给定 $D_i = 1$时 $\boldsymbol{D}_{N_i}$ 的分布进行边际化，确保估计目标能够捕捉定义 *1* 中所定义的受处理子群体的平均直接效应。

定理 2。 在假设 *1*、*2* 和 *3* 下，*AITT* 可识别为

$$\tau_{\text{ind}} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{|N_i|}\sum_{j\in N_i} E\left[\frac{D_i - e'_{ij}(D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j)}{\pi_i(z_i)\{1 - e'_{ij}(D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j)\}}(Y_{2j} - Y_{1j}) \mid z_i, z_j\right]\right),$$

其中， $e'_{ij}(D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j) = P(D_i = 1|D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j)$ 表示个体 i 在给定其邻居 j的处理状态、 j的其他邻居以及固定属性 z_i 和 z_j的条件下接受处理的倾向。请注意， e'_{ij} 是 D_j、 $\boldsymbol{D}_{N_j}^{-i}$ 、 z_i和 z_j的函数。

3 估计与推断

第 2 节中的识别结果直接导出 ADTT 和 AITT 的逆概率加权（及双重稳健）估计量；本节将详细介绍其实现与推断。第 2 节中的识别结果是非参数的，因为它们不要求对干扰结构作任何函数形式假设，且仅以邻域处理向量 $\boldsymbol{D}_{N_i}$ （以及协变量）为条件。然而，在有限样本中，以高维向量为条件会导致维数灾难。为保持实际可行性

并确保统计性能，我们因此采用灵活的参数模型进行估计——具体而言，用逻辑回归估计倾向得分，用线性回归进行结果回归——并将完整的邻居处理向量（例如按距离排序的指标 $D_{(k)}$）作为协变量。这些设定可视为对非参数条件期望的参数近似；在适当情况下，研究者也可以使用机器学习方法来估计干扰函数。关键在于，识别策略仍然是非参数的，而估计 步骤则使用灵活的参数（或机器学习）近似，使该方法具有实践可行性。

3.1 逆概率加权估计量

定理1和定理2给出了以下用于估计 ADTT 和 AITT 的逆概率加权（IPW）估计量：

$$\widehat{\tau}_{\mathrm{dir}} = \frac{1}{n}\sum_{i=1}^{n} \frac{D_i - \widehat{e}_i(\boldsymbol{D}_{N_i}, z_i)}{\widehat{\pi}_i(z_i)\{1-\widehat{e}_i(\boldsymbol{D}_{N_i}, z_i)\}}(Y_{2i} - Y_{1i}), \tag{1}$$

$$\widehat{\tau}_{\mathrm{ind}} = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{1}{|N_i|}\sum_{j\in N_i}\left\{\frac{D_i - \widehat{e}'_{ij}(D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j)}{\widehat{\pi}_i(z_i)\{1-\widehat{e}'_{ij}(D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j)\}}(Y_{2j} - Y_{1j})\right\}\right\}. \tag{2}$$

此处，$\widehat{\pi}_i(z_i)$、$\widehat{e}_i(\boldsymbol{D}_{N_i}, z_i)$ 和 $\widehat{e}'_{ij}(D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j)$ 分别是 $\pi_i(z_i)$、$e_i(\boldsymbol{D}_{N_i}, z_i)$ 和 $e'_{ij}(D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j)$ 的估计量。

我们使用逻辑回归估计倾向得分。对于 ADTT，该模型以个体 i 的协变量 z_i 及其 L 个最近邻 $\boldsymbol{D}_{N_i}$（按距离排序）的处理状态为条件：

$$\mathrm{logit}(P(D_i = 1)) = \alpha + \beta^{\top} z_i + \sum_{k=1}^{L} \gamma_k D_{(k)}.$$

这一设定是对 $P(D_i = 1 \mid \boldsymbol{D}_{N_i}, z_i)$ 的灵活参数近似，既无需预先指定暴露映射，又能在有限样本中保持可行。通过逐一纳入处理指示变量 $D_{(k)}$，而非将它们汇总（例如，

汇总为一个总和），模型能够自然地捕捉异质性溢出效应：影响会随距离增加而减弱，或取决于受处理邻居的具体距离。同样，对于 AITT，我们在给定协变量 z_i，z_j、邻居处理 D_j以及 j 的其他邻居按距离排序的处理 $\boldsymbol{D}_{N_j}^{-i}$ 的条件下估计 e'_{ij}。在这两种情况下，特征向量的维度固定，由 z_i 中协变量的数量（AITT 中还包括 z_j ）加上 L 个邻居处理指示变量决定，从而确保计算可行，同时保留非参数识别 策略（估计时使用参数模型）。实际数据应用中使用的具体协变量详见第 6 节。

我们提出的 ADTT 估计量不同于干扰下现有的逆概率加权估计量，例如 Xu（2025）提出的估计量。尽管二者都采用逆概率加权结构，但关键区别在于倾向得分所使用的条件集。现有方法通常基于低维、预先设定的暴露映射 G_i，即 $P(D_i = 1|G_i)$，来估计倾向得分。相比之下，我们的估计量使用以完整邻域处理向量为条件的倾向得分，即 $e_i = P(D_i = 1|\boldsymbol{D}_{N_i})$。这种方法遵循我们的非参数识别策略，避免了暴露映射设定错误的风险，从而能够更稳健地估计直接效应。然而，这种稳健性需要付出代价。以高维向量 $\boldsymbol{D}_{N_i}$ 为条件估计倾向得分可能颇具挑战，尤其是在邻居数量 L 较多时（维数灾难）。与正确设定低维暴露映射的方法相比，这可能导致统计效率降低。如果研究者对干扰的函数形式具有较强的先验知识，采用暴露映射方法可能更有效率。当缺乏此类知识时，我们实施非参数识别的估计量提供了一种稳健的替代方案。

3.2 双重稳健估计量

尽管上文介绍的逆概率加权估计量基于我们的非参数识别策略，但它们依赖于对倾向得分模型 $\widehat{e}_i(\boldsymbol{D}_{N_i}, z_i)$ 和 $\widehat{e}'_{ij}(D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j)$ 的正确设定。为提高模型设定错误下的稳健性，我们提出双重稳健（DR）估计量，将逆概率加权方法与结果回归模型相结合。双重稳健估计量的主要优势在于：只要倾向得分模型或结果回归模型中的任意一个设定正确，该估计量就仍然具有一致性，从而能够防范其中一个组成部分的设定错误。

为构造双重稳健估计量，我们首先定义与识别假设相符的结果回归模型。对于 ADTT，我们定义条件期望函数：

$$\mu_{1i}(\boldsymbol{D}_{N_i}, z_i) = E[Y_{2i} - Y_{1i} | D_i = 1, \boldsymbol{D}_{N_i}, z_i],$$
$$\mu_{0i}(\boldsymbol{D}_{N_i}, z_i) = E[Y_{2i} - Y_{1i} | D_i = 0, \boldsymbol{D}_{N_i}, z_i].$$

这些函数以完整的邻域处理向量 $\boldsymbol{D}_{N_i}$为条件，从而确保与假设2一致。对于 AITT，我们为邻居 j的结果变量定义类似的结果回归模型，并以 D_i、 D_j和 $\boldsymbol{D}_{N_j}^{-i}$ 为条件：

$$\mu_{1ij}(D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j) = E[Y_{2j} - Y_{1j} | D_i = 1, D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j],$$
$$\mu_{0ij}(D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j) = E[Y_{2j} - Y_{1j} | D_i = 0, D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j].$$

基于这些结果回归模型，我们提出以下双重稳健估计量。在此，我们将观测结果差异记为 $\Delta Y_i = Y_{2i} - Y_{1i}$ （ΔY_j 也同理）。预测的结果差异定义为：

$$\Delta m_{1i} = \widehat{\mu}_{1i}(\boldsymbol{D}_{N_i}, z_i), \quad \Delta m_{0i} = \widehat{\mu}_{0i}(\boldsymbol{D}_{N_i}, z_i),$$
$$\Delta m'_{1,ij} = \widehat{\mu}_{1ij}(D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j), \quad \Delta m'_{0,ij} = \widehat{\mu}_{0ij}(D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j),$$

其中，$\widehat{\mu}_{1i}$、$\widehat{\mu}_{0i}$、$\widehat{\mu}_{1ij}$和 $\widehat{\mu}_{0ij}$ 是上述相应结果回归函数的估计量。

我们使用线性回归来估计结果回归模型。对于 ADTT，我们将ΔY_i 对 D_i、z_i 和按距离排序的邻居处理 $\boldsymbol{D}_{N_i}$进行回归。分别通过设定 $D_i = 1$ 和 $D_i = 0$，得到预测值 Δm_{1i} 和Δm_{0i}。对于 AITT，我们将ΔY_j 对 D_i、D_j、z_i、z_j和排序后的邻居处理 $\boldsymbol{D}_{N_j}^{-i}$ 进行回归。通过固定 D_i生成预测值。虽然我们使用普通最小二乘法，但也可以采用机器学习等灵活方法。

$$\widehat{\tau}_{\text{dir}}^{dr} = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{D_i}{\widehat{\pi}_i}\left(\Delta Y_i - \Delta m_{1i}\right) - \frac{(1-D_i)\widehat{e}_i}{\widehat{\pi}_i(1-\widehat{e}_i)}\left(\Delta Y_i - \Delta m_{0i}\right) + \frac{\widehat{e}_i}{\widehat{\pi}_i}\left(\Delta m_{1i} - \Delta m_{0i}\right)\right\}, \tag{3}$$

对于 AITT，

$$\begin{aligned}\widehat{\tau}_{\text{ind}}^{dr} = \frac{1}{n}\sum_{i=1}^{n}\Bigg\{\frac{1}{|N_i|}\sum_{j\in N_i}\Bigg(&\frac{D_i}{\widehat{\pi}_i}\left(\Delta Y_j - \Delta m_{1,ij}^{'}\right)\\ &- \frac{(1-D_i)\widehat{e'}_{ij}}{\widehat{\pi}_i(1-\widehat{e'}_{ij})}\left(\Delta Y_j - \Delta m_{0,ij}^{'}\right) + \frac{\widehat{e'}_{ij}}{\widehat{\pi}_i}(\Delta m_{1,ij}^{'} - \Delta m_{0,ij}^{'})\Bigg)\Bigg\}.\end{aligned} \tag{4}$$

从结构上看，这些估计量由标准逆概率加权部分构成，并加入了基于结果回归模型推导出的偏误校正项。如果倾向得分设定错误，这些回归项会校正剩余偏误；反之，如果结果模型设定错误，逆概率加权权重则能确保一致性。这里，$\widehat{e}_i$ 和 $\widehat{e'}_{ij}$ 分别是 $\widehat{e}_i(\boldsymbol{D}_{N_i}, z_i)$和 $\widehat{e'}_{ij}(D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j)$的简写。因此，双重稳健估计量具有双重稳健性：只要倾向得分模型（$\widehat{e}_i$、$\widehat{e'}_{ij}$）或结果回归模型（$\widehat{\mu}_{1i}$、$\widehat{\mu}_{0i}$、$\widehat{\mu}_{1ij}$、$\widehat{\mu}_{0ij}$）之一设定正确，它们对 ADTT 和 AITT 的估计就是一致的。这一性质能显著降低模型设定错误带来的影响，因为研究者只需正确设定两个模型组成部分中的一个，而不必两者都设定正确。当两个模型都设定正确时，与单独使用逆概率加权估计量相比，双重稳健估计量通常能提高效率。

3.3 方差估计

为了对 ADTT 和 AITT 进行统计推断，我们需要估计 $\widehat{\tau}_{\text{dir}}$ 和 $\widehat{\tau}_{\text{ind}}$的（渐近）方差。在网络依赖下，由于观测值之间存在相关性，标准方差估计量并不一致。这一问题在实际应用中尤为突出，例如 Conley（1999）分析的情形：同一行业内地理位置邻近的主体或企业会受到共同冲击，从而产生标准估计量未能考虑的相关扰动。因此，按照 Kojevnikov等人（2021）的做法，我们采用能够考虑网络依赖结构的异方差和自相关一致（HAC）估计量。具体而言，我们考虑以下基于核函数的 HAC 估计量：

$$\widehat{V}_n = \sum_{s \geq 0} \omega\left(\frac{s}{b_n}\right) \widehat{\Omega}_n(s), \tag{5}$$

其中，ω 是满足 $\omega(0) = 1$、具有有限支撑且对称的核函数。示例包括 Bartlett 核和 Parzen 核。术语 b_n 表示一个随 n增加的带宽参数。在实践中，带宽 b_n 通常根据网络结构和干扰范围 K来选择。遵循 Kojevnikov 等人（2021年）的框架，我们将带宽设为干扰范围 K的某个倍数（例如，对于某个常数 $c > 0$，取 $b_n = c \cdot K$），以考虑网络上的局部依赖结构。$\Omega_n(s)$是距离 s处的样本自协方差矩阵，定义如下：

$$\widehat{\Omega}_n(s) = n^{-1} \sum_{i \in I_n} \sum_{j \in N^{\partial}(i;s)} (\phi_i - \bar{\phi}_n)(\phi_j - \bar{\phi}_n)$$

其中，$N^{\partial}(i; s)$ 是与个体 i 的距离恰好为 s 的单位集合（定义见第 2 节），而 ϕ_i 是该估计目标的影响函数，定义如下：

$$\begin{aligned}
&\text{(ADTT)} \quad \phi_i^{\text{dir}} = \frac{D_i - e_i(\boldsymbol{D}_{N_i}, z_i)}{\pi_i(z_i)\{1 - e_i(\boldsymbol{D}_{N_i}, z_i)\}}(Y_{2i} - Y_{1i}), \\
&\text{(AITT)} \quad \phi_i^{\text{ind}} = \frac{1}{|N_i|} \sum_{j \in N_i} \left\{ \frac{D_i - e'_{ij}(D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j)}{\pi_i(z_i)\{1 - e'_{ij}(D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j)\}}(Y_{2j} - Y_{1j}) \right\}.
\end{aligned} \tag{6}$$

¯这里，ϕ_n 是其样本均值。

使用第3.2节介绍的双重稳健估计量时，影响函数 ϕ_i应替换为双重稳健估计量中相应的求和项。具体而言，对于双重稳健估计量，影响函数为：

$$
\begin{aligned}
(\mathrm{ADTT}) \quad \phi_i^{DR,\mathrm{dir}} &= \frac{D_i}{\widehat{\pi}_i}\left(\Delta Y_i - \Delta m_{1i}\right) - \frac{(1-D_i)\widehat{e}_i}{\widehat{\pi}_i(1-\widehat{e}_i)}\left(\Delta Y_i - \Delta m_{0i}\right) + \frac{\widehat{e}_i}{\widehat{\pi}_i}\left(\Delta m_{1i} - \Delta m_{0i}\right), \\
(\mathrm{AITT}) \quad \phi_i^{DR,\mathrm{ind}} &= \frac{1}{|N_i|}\sum_{j\in N_i}\left\{\frac{D_i}{\widehat{\pi}_i}\left(\Delta Y_j - \Delta m^{'}_{1,ij}\right)\right. \\
&\quad \left. - \frac{(1-D_i)\widehat{e'}_{ij}}{\widehat{\pi}_i(1-\widehat{e'}_{ij})}\left(\Delta Y_j - \Delta m^{'}_{0,ij}\right) + \frac{\widehat{e'}_{ij}}{\widehat{\pi}_i}(\Delta m^{'}_{1,ij} - \Delta m^{'}_{0,ij})\right\}.
\end{aligned}
\tag{7}
$$

该估计量通过考虑网络上相近单位之间的协方差，并随着单位对之间的距离增加而降低其权重，从而稳健地捕捉局部依赖结构。正如下一节所示，当 $n \to \infty$ 时，HAC$\widehat{\ }$ 估计量 V_n 是一致的。基于该方差估计量，可以对 ADTT 或 AITT 进行统计推断。例如，ADTT 的 Wald 型 100$(1-\alpha)$% 置信区间可构造为 $\widehat{\tau}_{\mathrm{dir}} \pm z_{\alpha/2}\sqrt{\widehat{V}_n/n}$。

---

4 渐近性质

本节确立所提出估计量的大样本性质（一致性与渐近正态性）。由于网络结构上的数据呈现出复杂的相依性，基于独立同分布假设的标准渐近理论并不适用。因此，我们采用弱依赖随机场的现代理论，具体而言是Kojevnikov等人（2021）提出的 ψ-相依框架，以给出严格的证明。该方法也与Leung（2022）对近似邻域干扰的分析相关。

4.1 网络上的局部依赖

首先，我们定义本分析所依赖的弱相依概念。ψ-相依是一种基于经非线性变换的变量之协方差来衡量相依强度的度量。

定义 3 (ψ-依赖)。若存在一组非负函数 $\{\psi_{h,h'}(\cdot,\cdot)\}$ 以及一个一致有界序列 $\{\theta_{n,s}\}$，满足 $\theta_{n,\,0}=1$ 使得对任意 $n, h, h', s>0$、任意有界利普希茨函数 $f\in L_h, f'\in L_{h'}$，以及任意一对集合 $(H,H')\in\mathcal{P}_n(h,h',s)$，只要它们之间的网络距离至少为 s，就有一组随机变量 $\{\phi_i\}_{i=1\ldots n}$ 称为ψ-依赖的：

$$|Cov(f(\phi_H),f'(Z_{H'}))|\le\psi_{h,h'}(f,f')\tilde{\theta}_{n,s}$$

其中，$\mathcal{P}_n(h,h',s)=\{(H,H')\mid l_A(H,H')\ge s,|H|=h,|H'|=h'\}$，$\mathcal{L}_a$ 表示 $\mathbb{R}^a$ 上有界利普希茨实函数的集合，即 $\mathcal{L}_a=\{f:\mathbb{R}^a\to\mathbb{R}:\|f\|_\infty<\infty, Lip(f)<\infty\}$，其中 $Lip(f)$ 表示 f 的利普希茨常数，[6]而 $\|\cdot\|_\infty$ 表示 f 的上确界范数，即 $\|f\|_\infty=sup_x|f(x)|$.

~该定义要求依赖系数 $\theta_{n,\,s}$ 随着距离 s的增加而衰减。

4.2 假设

为了推导估计量的大样本性质，我们引入一组正则条件，涉及干扰范围、随机性质以及网络的整体依赖结构。

假设4 (局部干扰)。干扰仅限于距离 K 内。也就是说，个体 i 的基本变量 (Y_{1i},D_i,ϵ_i) 不受距离超过 K 的单位的处理影响。对于所有 $i\in I_n$ 以及所有 $\boldsymbol{D}_{-i}\in\{0,1\}^{n-1}$:

$$P(Y_{1i},D_i,\epsilon_i|\boldsymbol{D}_{-i}=\boldsymbol{d}_{-i})=P(Y_{1i},D_i,\epsilon_i|\boldsymbol{D}_{N(i;K)}=\boldsymbol{d}_{N(i;K)}).$$

假设5（重叠）。所有倾向得分都一致地与0和1保持正距离。也就是说，存在 π, $\pi \in \in (0, 1)$，使得对于所有 $i, j \in I_n$, $\pi_i(z_i) = P(D_i = 1|z_i) \in [\pi, \pi]$, $e_i(\boldsymbol{D}_{N_i}, z_i) = P(D_i = 1|\boldsymbol{D}_{N_i}, z_i) \in [\pi, \pi]$ 以及 $e'_{ij}(D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j) = P(D_i = 1|D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j) \in [\pi, \pi]$。

假设6（有界结果）。结果变量是一致有界的。也就是说，存在一个常数 $Y < \infty$，使得对于所有 $i \in I_n$、$|Y_{1i}|$、$|Y_{2i}| < Y$，几乎处处成立。

接下来，我们介绍一个关于网络依赖结构的假设。令$x_i = (Y_{1i}, D_i, \epsilon_i)$。我们假设这些基本随机变量$\{x_i\}$ 的集合在空间上是弱依赖的。具体而言，我们假设满足 Jenish and Prucha (2009) 所考虑的一类空间 α-混合条件。相关依赖结构的极限定理已由同一作者在一系列论文中建立（例如，Jenish and Prucha, 2009, 2012）。

假设7（α-混合）。令 $\alpha_n(H, H') = \alpha(\sigma_n(\{x_i\}_{i\in H}), \sigma_n(\{x_i\}_{i\in H'}))$，其中 $\sigma(\cdot)$ 表示由随机变量集合生成的 σ-代数。存在一个有界函数 $\psi(\cdot, \cdot)$: $\mathbb{N} \times \mathbb{N} \to \mathbb{R}$ 以及一个序列 $\{\widehat{\alpha}(s)\}_{s\in\mathbb{R}}$，使得对于$\bar{\alpha}(h, h', s) = \sup_n \sup_{H,H'} \{\alpha_n(H, H') \mid |H| \leq h, |H'| \leq h', l_A(H, H') \geq s\}$，我们有：

$$\bar{\alpha}(h, h', s) \leq \psi(h, h')\widehat{\alpha}(s)$$

其中 $\widehat{\alpha}(s) \to 0$ 如 $s \to \infty$。这一假设决定了观测数据中潜在空间相依的强度。

以下技术假设规定了这种衰减率与网络密度之间的权衡，以确保大数定律（LLN）和中心极限定理（CLT）成立。

假设 8 (大数定律的弱相依)。 $\sum_{s=0}^{n} M_n^{\partial}(s)\tilde{\theta}_{n,\,s} = o(n)$，其中 $M_n^{\partial}(s) = \frac{1}{n}\sum_{i=1}^{n} |N^{\partial}(i;\ s)|$ 以及 $N^{\partial}(i;\ s)$ 的定义见第 2 节。

假设 9(中心极限定理的弱依赖**)**。 中心极限定理要求对网络结构和依赖衰减率施加更强的技术条件。存在一个正序列 $r_n \to \infty$，使得对于某些 $p > 4$ 和 $k = 1, 2$:

$$\frac{1}{\sigma_n^{2+k}} \sum_{i \in I_n} \sum_{s=0}^{\infty} s^{d-1} \max_{j \in I_n, s \le l_A(i,j) < s+1} |N(i; r_n)^+ \setminus N(j; s-1)^+|^k \tilde{\theta}_{n,s}^{1-\frac{2+k}{p}} \to 0$$

and

$$\frac{|I_n|^2 \tilde{\theta}_{n,r_n}^{1-\frac{1}{p}}}{\sigma_n} \to 0 \quad as\ n \to \infty$$

其中， σ_n^2 是随机变量之和的方差。

最后，我们根据 Kojevnikov 等人（2021年）的命题 4.1，列出 HAC方差估计量一致性所需的假设。

假设 10 （HAC一致性条件）。 存在某个 $p > 4$，使得核函数 $\omega(\cdot)$、带宽 b_n 以及弱依赖~系数$\theta_{n,\ s}$ 满足以下条件:

- ω（0） $= 1, \omega$（s/b_n） $= 0$ 对任意 $s > b_n$ 以及 $|\omega$（s/b_n）$| < \infty$ 对所有 n均成立。
- 极限$_{n\to\infty} \sum_{s\ge1} |\omega$ （s/b_n） $- 1| \, s^{d-1} \, \theta_n^{1-\frac{2}{p}}, \ \underline{s}$ $= 0$。
- 极限$_{n\to\infty} n^{-1} \sum_{s=0}^{\infty} s^{d-1}$最大值$_{j\in I_n,\ s\le l_A\ (i,\ j)\ <s+1} \left|N\ (i;\ \ b_n)\ ^+ \setminus N\ (j;\ \ s-1)\ ^+\right|^2 \ \theta_n^{1-\frac{4}{p}}, \ s$ $= 0$。
- $\delta_n(b_n) = o(n)$ 几乎处处，其中 $\delta_n(b_n) = \sum_{s=0}^{b_n} M_n^{\partial}(s)$ 以及 $M_n^{\partial}(s) = \frac{1}{n} \sum_{i=1}^{n} |N^{\partial}(i;\ \ s)|$ 表示每个个体恰好位于距离 s 处的单位平均数量。该条件确保带宽 b_n 范围内的网络密度相对于样本量不会增长得过快，这是 *HAC* 估计量具有一致性的必要条件。
- 序列 $\left\{ \tilde{\theta}_{n,\ s} / s^{p/\ (p-4)} \right\}$ 关于 $s \ge 1$是非递增的。

4.3 主要定理

首先，我们给出若干引理，说明所提出的估计量的各个组成部分满足 ψ-依赖的定义。这源于底层变量的 α-混合性质（假设7）和局部干扰（假设4）。证明见附录。在本小节中， $\mathbf{1}\{\cdot\}$ 表示指示函数：当条件 A 成立时取 $\mathbf{1}\{A\}=1$，否则取 $\mathbf{1}\{A\}=0$。

引理 1 （ψ-依赖的 ADTT 和 AITT 求和项）。 在假设 *4-7* 下，式（*6*）中定义的随机变量集合 $\{\phi_i^{\mathrm{dir}}\}_{i\in I_n}$ 和 $\{\phi_i^{\mathrm{ind}}\}_{i\in I_n}$ 分别是带有 ψ-依赖的$\theta_{n,\,s}=\mathbf{1}\{s<2K\}+\mathbf{1}\{s\geq 2K\}\widehat{\alpha}\,(s-2K)$ 以及 $\theta_{n,\,s}=\mathbf{1}\{s<4K\}+\mathbf{1}\{s\geq 4K\}\widehat{\alpha}\,(s-4K)$。

该引理确保我们可以在 ψ-依赖框架内分析估计量。通过应用 Kojevnikov等人（2021）的极限定理，我们得到一致性和渐近正态性。

定理 3 （一致性）。 在假设 *2* 和 *4-8* 下， $\widehat{\tau}_{\mathrm{dir}}$ 和 $\widehat{\tau}_{\mathrm{ind}}$ 在 $n\to\infty$下是一致的。

定理 4 （渐近正态性）。 在假设 *2*、*4–7* 和 *9* 下，有$\sqrt{n}\,(\widehat{\tau}_{\mathrm{dir}}-\tau_{\mathrm{dir}})\,/\sigma_{n,\,D}\to N(0,1)$ 且 $\sqrt{n}\,(\widehat{\tau}_{\mathrm{ind}}-\tau_{\mathrm{ind}})\,/\sigma_{n,\,I}\Rightarrow N(0,1)$ 当 $n\to\infty$ 时，其中$\sigma^2_{n,\,D}=\mathrm{Var}(n^{-1}\sum_{i\in I_n}\phi_i^{\mathrm{dir}})$ 以及 $\sigma^2_{n,\,I}=\mathrm{Var}(n^{-1}\sum_{i\in I_n}\phi_i^{\mathrm{ind}})$。

接下来，我们证明第 3.3 节中引入的 HAC 估计量的一致性。

定理 5 （方差估计量的一致性）。 在假设 *2*、*4–7*、*9* 和 *10* 下，式（*5*）中给出的 *HAC* 估计量 V_n 是渐近方差的一致估计量。具体而言，对于 *ADTT* 估计量，$\widehat{V}_n\to\sigma^2_{n,\,D}$ 当 $n\to\infty$时；对于 *AITT* 估计量，$\widehat{V}_n\to\sigma^2_{n,\,I}$ 当$n\to\infty$时，其中$\sigma^2_{n,\,D}$ 和 $\sigma^2_{n,\,I}$ 在定理 *4* 中定义。

现在，我们为双重稳健估计量提供理论保证。第一个定理确立了双重稳健性，这是双重稳健估计量相较于逆概率加权估计量的关键优势。

定理 6（双重稳健估计量的双重稳健性）。 在假设*1*和假设*2*下，如果倾向得分模型（e_i、 e'_{ij}）设定正确，或者结果变量回归模型（μ_{1i}、μ_{0i}、 μ_{1ij}、 μ_{0ij}）设定正确，则双重稳健估计量 $\widehat{\tau}^{DR}_{\text{dir}}$和 $\widehat{\tau}^{DR}_{\text{ind}}$分别是 τ_{dir}和 τ_{ind}的一致估计量。

定理 *7*（双重稳健估计量的渐近正态性）。在假设 *2*、*4–7* 和 *9* 下，如果倾向得分模型和结果回归模型均设定正确（或使用具有足够收敛速率的机器学习方法进行估计），那么当 $n \to \infty$ 时，$\sqrt{n}(\widehat{\tau}^{DR}_{dir} - \tau_{\text{dir}})/\sigma^{DR}_{n,D} \to N(0,1)$ 且 $\sqrt{n}(\widehat{\tau}^{DR}_{ind} - \tau_{\text{ind}})/\sigma^{DR}_{n,I} \to N(0,1)$，其中 $\sigma^{DR}_{n,D}{}^2 = Var(n^{-1}\sum_{i\in I_n}\phi^{DR}_{i,\text{dir}})$ 且 $\sigma^{DR}_{n,I}{}^2 = Var(n^{-1}\sum_{i\in I_n}\phi^{DR}_{i,\text{ind}})$，其中$\phi^{DR}_{i,\text{dir}}$和 $\phi^{DR}_{i,\text{ind}}$由公式（7）定义。

上述定理的所有证明均见附录。上述定理为通过$\widehat{}$使用 HAC估计量 V_n对 ADTT 和 AITT 进行统计推断提供了理论保证。

5 模拟研究

本小节通过一系列蒙特卡洛模拟评估我们所提出的估计量（ADTT 和 AITT）的有限样本表现。主要目标是在真实暴露映射未知且可能较为复杂的情形下，展示我们基于识别的方法的稳健性，并将其与依赖预先设定结构的现有方法进行比较。

5.1 数据生成过程

我们模拟一个包含 n 个随机分布于 20×20 二维空间中的个体的总体。网络结构由切比雪夫距离定义，干扰范围设为 $K = 1$。我们生成两类混杂因素：一个可观测的个体属性$z_i \sim N(0,1)$，以及一个不可观测的相关因子 $(z_{u1},\ldots,z_{un}) \sim N(0,\Sigma)$，其中协方差矩阵 $\Sigma_{ij} = 0.5^{l_A(i,j)}$ 刻画了网络依赖，其随网络距离增大而衰减

。网络距离为 $l_A(i, j)$。随后，处理根据这两个因素进行分配，具体如下：

$$P(D_i = 1 \mid z_i, z_{u,i}) = \text{logit}^{-1}(0.3z_i + 0.8z_{u,i}),$$

这会产生一种依赖于网络的聚类处理模式。结果变量根据以下两期模型生成：

$$Y_{1i} = 1.2z_i + 0.5z_{u,i} + \epsilon_{1i}$$

$$Y_{2i} = 1 + Y_{1i} + \tau D_i + f(S_i) + 0.1z_{u,i} + 0.2z_i + \epsilon_{2i}$$

其中 $\epsilon_{1i},\ \epsilon_{2i} \sim N(0,\ 1)$。真实直接处理效应设为 $\tau = 0.8$。本次模拟的核心特征是溢出函数 $f(S_i)$，其中 S_i 表示个体 i 的受处理邻居数量。我们假设其具有复杂的阶梯式递增结构：

$$f(S_i) = 0.8 \cdot I(S_i = 1) + 1.6 \cdot I(S_i = 2) + 2.4 \cdot I(S_i \geq 3).$$

这代表一种非线性干扰结构，其中溢出效应会达到饱和。我们假设研究者并不知道这一函数形式 $f(\cdot)$。真实 ADTT 等于 $\tau = 0.8$，而真实 AITT 则根据生成的数据计算条件平均值。具体而言，对于每个受处理个体 i，我们通过比较个体 i 受处理与未受处理时的结果变量，在观测网络结构和处理分配条件下，计算其邻居 j 所受到的平均溢出效应。根据 DGP 计算得出的真实 AITT 在各次模拟运行中的平均值约为 0.4961。图 1 展示了某次代表性模拟运行中 S_i 的分布，表明不同暴露水平的出现频率。

5.2 比较方法

我们首先评估所提出的估计量（依赖于非参数识别）与其他几种现有方法在估计处理的直接效应方面的表现，

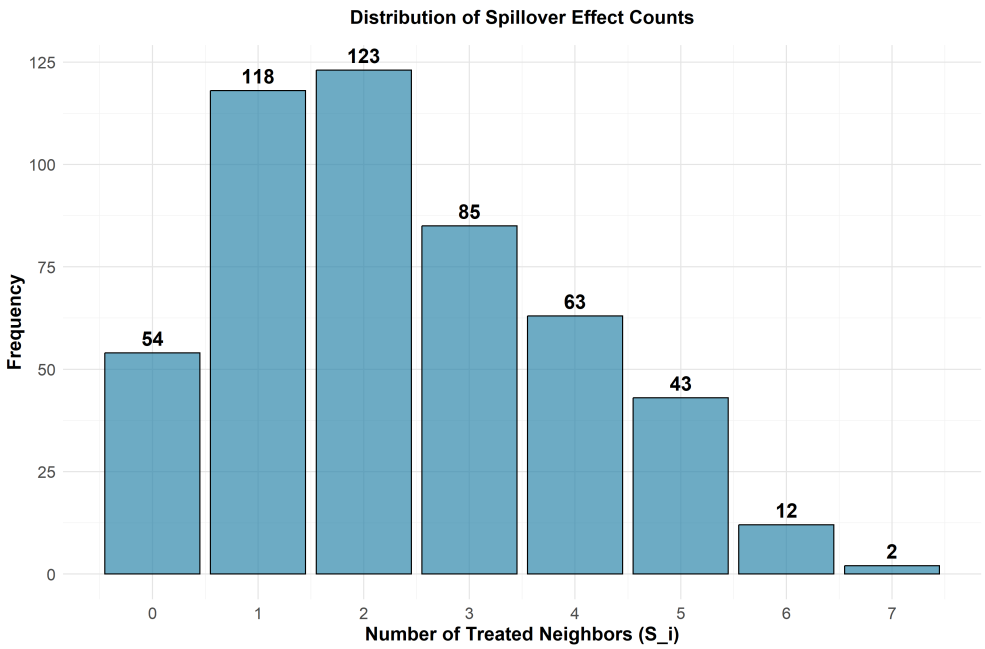

图1：代表性模拟运行中各个体所受处理邻居数量（S_i）的分布。该直方图展示了不同暴露水平的普遍程度，这对于理解我们数据生成过程中溢出效应的分布至关重要。

并对这些方法进行比较。对于所提出的估计量，我们根据观测协变量 z_i 和邻域处理向量$\boldsymbol{D}_{N_i}$（针对 ADTT）或 $\boldsymbol{D}_{N_j}^{-i}$（针对 AITT）进行条件化，以估计倾向得分，具体如定理1和定理2所述。按照第 3.1 节的方法，我们使用逻辑回归估计倾向得分。在模拟实验中，我们设定 $L = 10$。对于 AITT 估计，当个体 j的邻居数量超过 L时，我们会从距离在 K 以内的邻居中随机抽取 L 个，以保持计算效率，同时避免任意选择。

所有其他估计量仅使用观测协变量 z_i。在本次模拟中，我们报告了所提出估计量的逆概率加权和双重稳健版本的结果。对于双重稳健估计量，我们按照第3.2节所述，使用线性回归（普通最小二乘法）估计结果回归模型。

- **所提逆概率加权**：我们提出的用于 ADTT 和 AITT 的非参数逆概率加权估计量，按照第3.1节所述方法实现，并结合 $L = 10$。

• **所提双重稳健**：我们提出的用于 ADTT 和 AITT 的非参数双重稳健估计量，按照第3.2节所述方法实现。

• **Xu（2025年）方法**：Xu（2025年）提出的估计量需要指定一个暴露映射 G_i。我们实现了这些估计量的逆概率加权和双重稳健两个版本。使用以下三种暴露映射：

– 真实映射（不可行）：研究者知道真实的四级结构。 $G_i^{\text{Oracle}} = \min(S_i, 3)$。– MO（错误设定排序）：结构设定有误；根据真实暴露映射 G_i^{Oracle}，随机将 30% 的单位分配到不同的暴露组。– FM（完全错误设定）：研究者假设了一个过度简化的二元结构。 $G_i^{FM} = \mathbb{I}(S_i > 1)$。

请注意，在表5.3中，为与我们提出的双重稳健估计量进行比较，我们仅报告双重稳健版本（分别记为 Xu（Oracle）、Xu（MO）和 Xu（FM））。图2同时包含 Xu 方法的逆概率加权和双重稳健版本。

• **标准DID方法**：忽略干扰的标准DID估计量（典型逆概率加权、典型双向固定效应、双重稳健双重差分），以及一种扩展的双向固定效应模型（Xu（2025年）中的修正双向固定效应），该模型考虑了简单的二元暴露（$I(S_i \geq 1)$）。

5.3 结果

表5.3和图2总结了 ADTT（$\tau_{\text{dir}} = 0.8$）的结果。结果显示，与依赖设定错误的暴露映射或忽略干扰的方法相比，所提出的逆概率加权和双重稳健估计量通常能显著降低偏差和均方根误差。具体而言，所提出的 ADTT 估计量的偏差（分别为 0.062 和 0.033）和均方根误差与不可行的真实映射估计量相当。这可能是因为对完整的邻域处理向量 $\boldsymbol{D}_{N_i}$ 进行条件化，避免了暴露映射设定错误

同时条件平行趋势假设仍然成立。忽略干扰的基准估计量（典型逆概率加权、双向固定效应、双重稳健双重差分）因聚类处理带来的正向溢出效应而表现出显著的向上偏差（偏差 ≈ 0.21）；Xu（2025）提出的采用设定错误暴露映射的方法（MO、FM）也表现出显著偏差。

估计量	Bias	RMSE	CP
所提方法			
所提逆概率加权 (ADTT)	**0.0619**	**0.1171**	**1.0000**
所提双重稳健 (ADTT)	**0.0331**	**0.1040**	**0.9500**
Xu（2025）方法			
Xu（Oracle）	0.0567	0.1064	0.9900
Xu（MO）	0.1358	0.1830	0.9400
Xu（FM）	0.1032	0.1411	0.9900
标准*DID*方法			
典型逆概率加权	0.2115	0.2621	0.5200
典型双向固定效应	0.2077	0.2591	0.5100
双重稳健双重差分	0.2100	0.2613	0.5200
修正双向固定效应	0.1484	0.1936	0.7000

表1： $n = 500$ 下 ADTT 的模拟结果（100 次重复）。真实 ADTT =0.8。

估计量	Bias	RMSE	CP
所提方法			
所提逆概率加权（AITT）	**0.0473**	**0.1237**	**1.0000**
所提双重稳健（AITT）	**0.0139**	**0.0575**	**0.9200**

表2：基于 $n = 500$ 的AITT模拟结果：基于100次蒙特卡洛重复实验的性能比较。报告了95%置信水平下的偏差、均方根误差和覆盖概率（CP）。真实AITT为0.4961。各估计量的详情见第5.2节。

所提出的 AITT 估计量表现出色。具体而言，在可行的估计量中，双重稳健平均间接处理效应估计量的偏差（0.0139）和均方根误差（0.0575）最小；逆概率加权平均间接处理效应估计量的偏差为 0.0473，均方根误差为 0.1237，与相应的 ADTT 估计量相当或更优。这可能是因为，对个体 i及其邻居 j 的邻域处理配置进行条件化，无需暴露映射也能保留识别，而在中等邻域大小下，维数灾难仍在可控范围内。双重稳健平均间接处理效应估计量的优异表现，体现了双重稳健性对间接效应估计的益处。

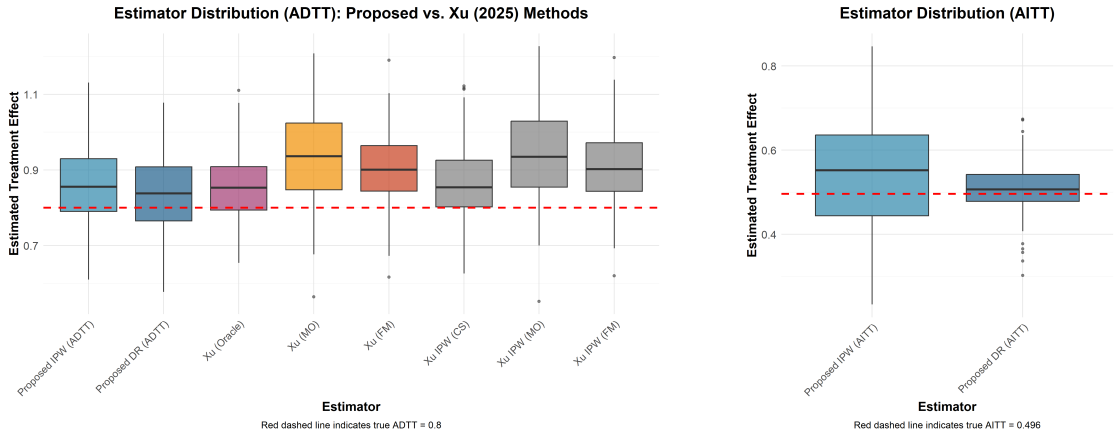

图2：蒙特卡洛重复实验中估计处理效应的分布。左：ADTT估计值。右：AITT估计值。虚线表示真实值（ADTT = 0.8，AITT = 0.4961）。方法分别标记为所提出的IPW/DR方法（我们的方法）、Xu（Oracle）、Xu（MO）和Xu（FM），对应于第5.2节所述的特定暴露映射设定。

我们还通过改变样本量 $n \in \{300, 500, 700\}$来考察估计量的渐近性质。图3展示了偏差随 n变化的情况。正如预期，随着样本量增加，所提出的逆概率加权和双重稳健 ADTT 估计量表现有所改善，偏差和均方根误差均逐渐趋近于零。这为我们估计量的一致性提供了经验证据（定理 3）。相比之下，忽略干扰或依赖设定错误的暴露映射的估计量，即使在样本量较大时仍存在显著偏差，这表明它们在该数据生成过程下不具一致性。

5.4 稳健性检验

我们进行了额外的稳健性检验，以考察所提出的方法在不同网络结构和依赖模式下的表现。这些实验进一步证明了我们非参数方法的稳健性。

为检验网络依赖的影响，我们通过改变基准相关参数 $\rho_0 \in \{0.2, 0.5, 0.8\}$来调整未观测混杂因素中空间相关性的强度。更强的空间相关性会增强相邻个体之间的依赖，使因果效应更难识别。图4表明，所提出的方法在不同空间相关性水平下仍具有稳健性，而

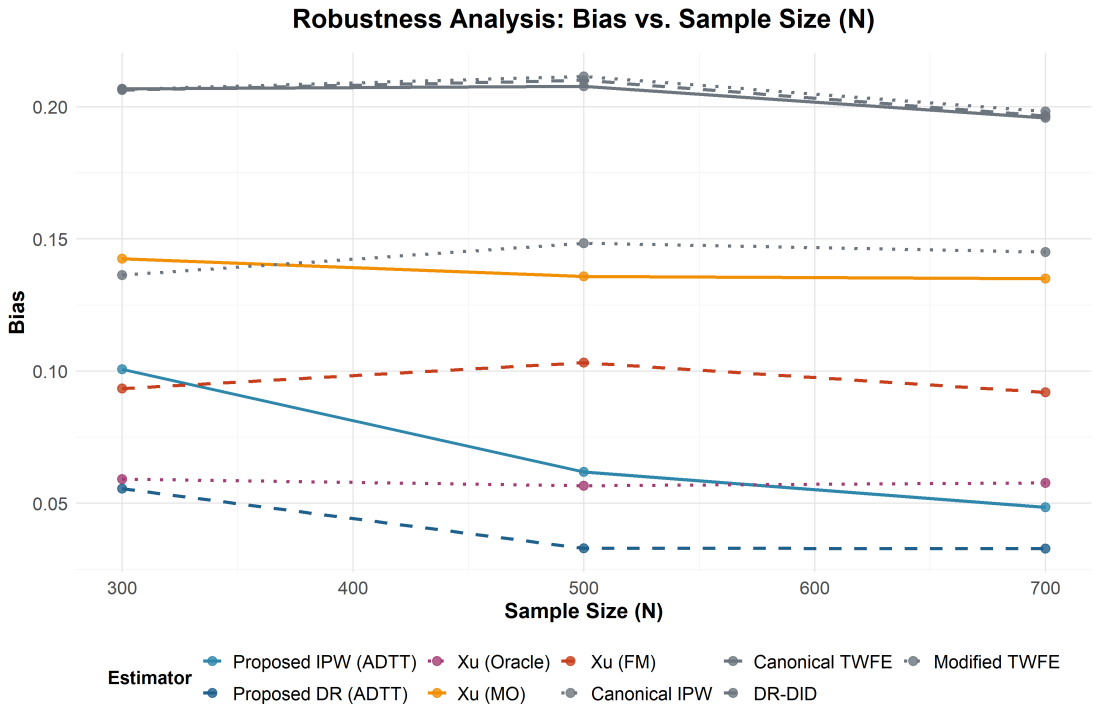

图3：所提出的 IPW 和双重稳健 ADTT 估计量及基准方法的偏差随样本量 $N \in \{300, 500, 700\}$ 的变化。随着样本量增加，所提出的 IPW 和双重稳健方法的偏差逐渐减小，为一致性提供了经验证据。忽略干扰或依赖设定错误的暴露映射的基准方法，即使在样本量更大时仍存在显著偏差。

忽略干扰的基准方法随着相关性增强，表现逐渐恶化。

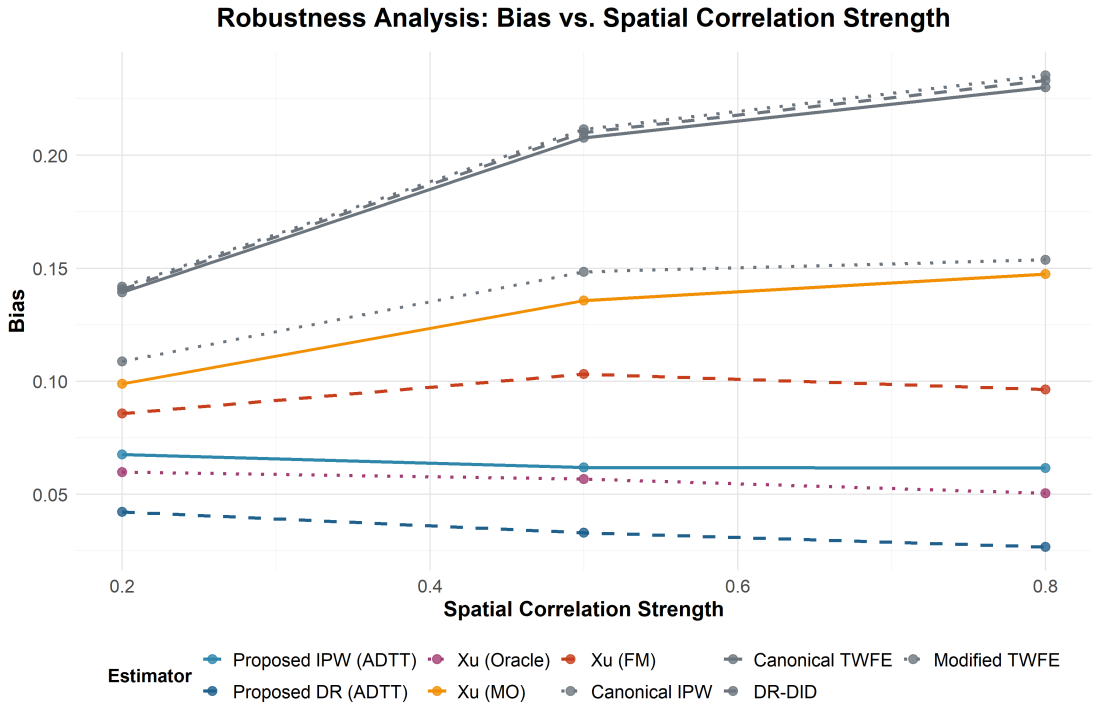

图4：偏差随空间相关性强度（$\rho_0 \in \{0.2, 0.5, 0.8\}$）的变化。

接下来，我们考察所提方法对邻域大小的敏感性。我们的方法可能面临的一个问题是，当邻域较大时会遭遇维数灾难。为此，我们通过改变倾向得分估计中使用的邻域特征数量进行敏感性分析。图5显示，随着有效维数增加，逆概率加权估计量的表现略有下降，而双重稳健估计量则保持稳定甚至有所改善，这表明双重稳健方法能够抵御维数灾难。即使邻域较大，两种方法仍能保持良好表现，说明我们的方法在实践中具有可行性。

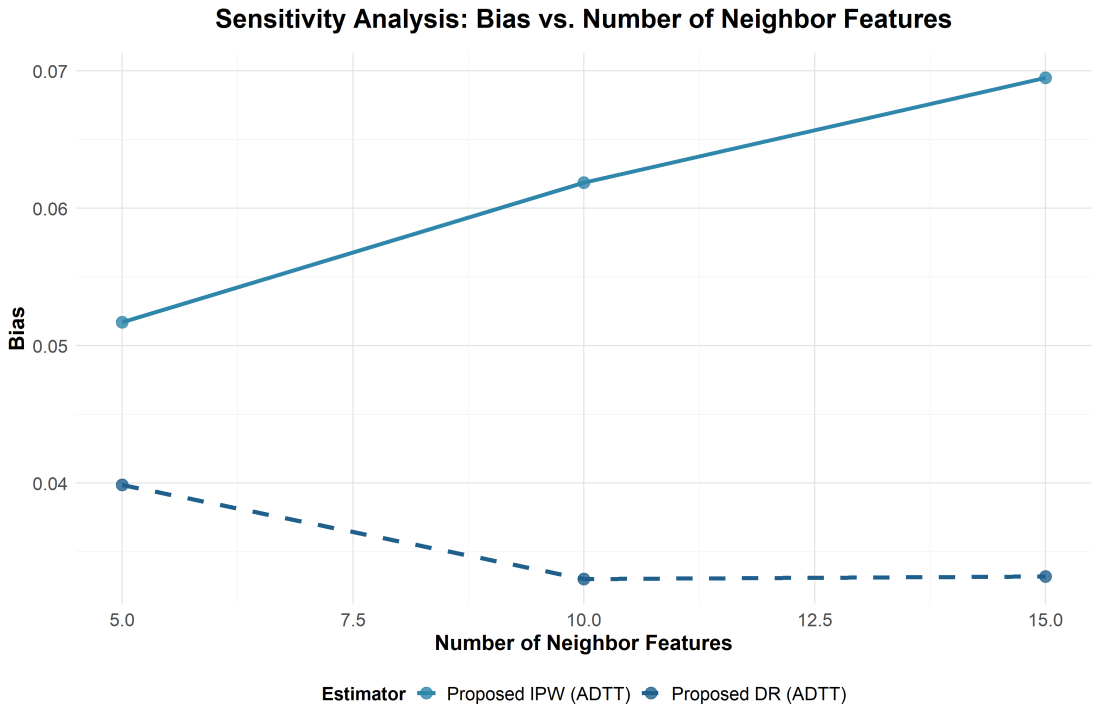

图5：偏差随倾向得分估计中使用的邻域特征数量的变化。

6 真实数据应用：中国经济特区政策

我们应用所提方法评估中国经济特区（SEZ）政策的效果。这是一项具有代表性的基于地点的政策干预。本应用展示了在政策评估中考虑溢出效应的重要性，并说明了我们的基于识别的方法如何应用于干扰结构未知的现实场景。

我们使用 Xu（2025）提供的中国村庄层面数据（2004年为处理前，2008年为处理后）。处理变量 D_i 表示村庄 i是否被指定为经济特区。由于无法获取地理坐标，我们根据协变量相似性在县级定义邻域。我们使用与第3.1节相同的模型设定，并采用 $L = 10$。作为基准，我们构建基于各县留一法经济特区比例的暴露映射。我们将暴露定义为每个村庄邻域内受处理邻居（经济特区村庄）的数量。村庄分为

“低暴露”（暴露低于中位数）和“高暴露”（暴露达到或超过中位数）。我们分析三个取对数的结果变量：资本、就业和产出。我们按照第3.1节所述实施所提出的逆概率加权估计量，并利用协变量相似性确定邻居排序。我们将所提出的估计量与 Xu（2025）的方法（需要暴露映射）以及忽略干扰的标准双重差分估计量进行比较。标准误采用HAC估计量计算，该估计量基于协变量相似性对县内空间相关性进行校正。由于县内无法获取地理距离，我们将协变量相似性作为一种实用的替代方案。

表3展示了结果。经济特区政策呈现出正向直接效应，对低暴露村庄的影响更大。我们发现，处理村庄之间存在负向溢出效应，这表明存在竞争效应。经典双重差分估计值显著大于考虑溢出效应后的估计值，说明存在正向选择偏差。ADTT的逆概率加权估计值与双重稳健估计值之间的差异（0.2318 对 0.0283）可能反映了模型设定或有限样本性质的差异；在各自的识别假设下，两种估计量都以处理组直接处理效应目标参数为估计目标。

表3：经济特区政策的直接效应和溢出效应估计值。括号内为标准误（HAC估计量）。所有结果变量均为对数尺度。

	结果变量		
效应	资本（对数）	就业（对数）	产出（对数）
不同暴露水平下的直接效应			
ADTT（低暴露）	0.0356	0.0309	0.0362
ADTT（高暴露）	0.0237	0.0197	0.0245
平均直接效应（*ADTT*）			
ADTT（双重稳健）	0.0283	0.0240	0.0290
提议的逆概率加权（ADTT）	**0.2318**	**0.2629**	**0.2066**
	(0.0269)	**(0.0181)**	**(0.0300)**
标准双重差分	0.3552	0.3464	0.3582
间接效应			
提出的逆概率加权（AITT）	**-0.2174**	**-0.1015**	**-0.2445**
	(0.0165)	**(0.0131)**	**(0.0209)**
溢出效应			
溢出效应（处理组）	-0.1930	-0.1438	-0.1955
溢出效应（对照组）	-0.0027	0.0328	-0.0023

7 结论

我们提出了一个局部网络干扰下的双重差分框架，该框架不依赖于已知的暴露映射。在给定邻域处理向量的条件平行趋势假设下，识别是非参数的；AITT估计目标刻画了干预对邻居的外部性。逆概率加权估计量的渐近理论建立在弱依赖随机场之上，我们还给出了一致的方差估计量用于推断。我们的模拟研究表明，在具有复杂且未知干扰结构的情形下，我们基于识别的方法（在识别阶段采用非参数方法）具有稳健性，其表现优于依赖设定错误的暴露映射或完全忽略干扰的方法。对中国经济特区政策的真实数据应用进一步说明了考察溢出效应的实际重要性，揭示出标准方法可能遗漏的显著政策效应和重要异质性模式。

我们强调，仍需进一步研究，以更清楚地了解在有限样本中，尤其是网络密度或邻域大小发生变化时，何种情况下我们的方法优于暴露映射方法。将这一框架扩展到交错处理采纳的情形，可以把本文的设定与双重差分文献近期指出的挑战结合起来。尽管这一扩展显然值得研究，但它会使识别和推断更加复杂，因此我们将其留待未来工作。另一个重要方向是，在邻域大小较大时，开发非参数地估计倾向得分的方法，并有望借助机器学习技术应对维数灾难。

致谢

本研究得到日本学术振兴会（JSPS KAK-ENHI）资助，项目编号为 24K00244 和 25H00546。

参考文献

Abadie, A. (2005). 半参数双重差分估计量。经济研究评论 *72*(1), 1–19。

Aronow, P. M. 与 C. Samii（2017）。在一般干扰下估计平均因果效应，并应用于社会网络实验。《应用统计年鉴》*11*（4），1912–1947。

Basse, G. 与 A. Feller（2018年）。在存在干扰的情况下分析两阶段实验。《美国统计协会杂志》*113*（521），41–55。

Butts, K.（2023）。存在空间溢出的双重差分。arXiv预印本 arXiv:2105.03737。arXiv:2105.03737。

Callaway, B. 与 P. H. C. Sant'Anna（2021年）。多时期双重差分。《计量经济学杂志》*225*（2），200–230。

Conley, T. G.（1999）。存在横截面相依的GMM估计。《计量经济学杂志》*92*（1），1–45。

Cox, D. R.（1958）。《实验设计》。Wiley。

de Chaisemartin, C.与X. D'Haultfœuille（2020）。存在异质性处理效应的双向固定效应估计量。《美国经济评论》*110*（9），2964–2996。

Delgado, M. S.与R. J. G. M. Florax（2015）。面向空间数据的双重差分方法：局部自相关与空间相互作用。《经济学快报》*137*，123–126。

Goodman-Bacon, A.（2021年）。存在处理时点差异的双重差分。计量经济学杂志*225*（2），254–277。

Hudgens, M. G. 和 M. E. Halloran（2008）。迈向存在干扰情况下的因果推断。《美国统计学会会刊》*103*(482)，832–842。

Jenish, N. 和 I. R. Prucha（2009）。随机场数组的中心极限定理与大数一致律。计量经济学杂志 *150*(1)，86–98。

Jenish, N. 和 I. R. Prucha（2012年）。近历元依赖下的空间过程与渐近推断。计量经济学杂志 *170*(1)，178–190。

Kojevnikov, D.、V. Marmer 和 K. Song（2021年）。网络依赖随机变量的极限定理。计量经济学杂志 *222*(2)，882–908。

Leung, M. P.（2022年）。近似邻域干扰下的因果推断。《计量经济学》*90*(1), 267–293。

Leung, M. P.（2024年）。利用暴露对比识别处理效应和溢出效应。*arXiv* 预印本 *arXiv:2403.08183*。

Liu, J.、D. Zhang 和 E. J. T. Tchetgen（2025年）。网络上因果效应的自动双重稳健估计。*arXiv* 预印本 *arXiv:2506.23332*。

Manski, C. F.（1993年）。内生社会效应的识别：反射问题。《经济研究评论》*60*(3), 531–542。

Manski, C. F.（2013年）。存在社会互动时处理反应的识别。《计量经济学杂志》*16*(1), S1–S23。

Perez-Heydrich, C., M. G. Hudgens, M. E. Halloran, J. D. Clemens, M. Ali 和 M. E. Emch（2014年）。存在干扰时评估霍乱疫苗接种的效果。《生物计量学》*70*(3), 731–741。

Rubin, D. B.（1980年）。实验数据的随机化分析：费希尔随机化检验：评论。《美国统计学会会刊》*75*(371), 591–593。

S¨avje, F.（2024年）。设定错误的暴露映射下的因果推断：区分定义与假设。《生物统计学》*111*(1)，1–16。

Sobel, M. E.（2006年）。住房流动随机研究证明了什么？面对干扰的因果推断。美国统计学会杂志 *101*(476)，1398–1407。

Sun, L. and S. Abraham（2021年）。在存在异质性处理效应的事件研究中估计动态处理效应。计量经济学杂志 *225*(2), 175–199。

Tchetgen Tchetgen, E. J. and T. J. VanderWeele（2012年）。论存在干扰时的因果推断。医学研究统计方法 *21*(1), 55–75。

Ugander, J., B. Karrer, L. Backstrom, and J. Kleinberg（2013年）。图聚类随机化：面向多个总体的网络暴露。载于第*19*届*ACM SIGKDD*知识发现与数据挖掘国际会议论文集，第3–12页。

Xu, R. (2025年). 存在干扰时的双重差分。arXiv 预印本 arXiv:2306.12003。arXiv:2306.12003。

《网络上的局部依赖下的双重差分》的补充材料

本补充材料提供了正文中各定理的证明。

S1 定理1的证明

根据 ADTT 的定义，我们对每个个体 i的邻域处理向量 $\boldsymbol{d} = \boldsymbol{d}_{N_i}$ 的所有可能组合进行边际化。

$$\begin{aligned}\tau_{\text{dir}} &= \frac{1}{n}\sum_{i=1}^{n} E[y_{2i}(1, \boldsymbol{D}_{N_i}) - y_{2i}(0, \boldsymbol{D}_{N_i})|D_i = 1, z_i] \\ &= \frac{1}{n}\sum_{i=1}^{n}\sum_{\boldsymbol{d}} E[y_{2i}(1, \boldsymbol{d}) - y_{2i}(0, \boldsymbol{d})|D_i = 1, \boldsymbol{D}_{N_i} = \boldsymbol{d}, z_i] \\ &\times P(\boldsymbol{D}_{N_i} = \boldsymbol{d}|D_i = 1, z_i)\end{aligned}$$

条件期望项可以改写为趋势差,

$$\begin{aligned}&E[y_{2i}(1, \boldsymbol{d}) - y_{2i}(0, \boldsymbol{d})|D_i = 1, \boldsymbol{D}_{N_i} = \boldsymbol{d}, z_i] \\ &= E[y_{2i}(1, \boldsymbol{d}) - y_{1i}|D_i = 1, \boldsymbol{D}_{N_i} = \boldsymbol{d}, z_i] \\ &- E[y_{2i}(0, \boldsymbol{d}) - y_{1i}|D_i = 1, \boldsymbol{D}_{N_i} = \boldsymbol{d}, z_i]\end{aligned}$$

根据假设2，第二项（反事实趋势）由对照组中的观测趋势（$D_i = 0$）识别:

$$E[y_{2i}(0, \boldsymbol{d}) - y_{1i}|D_i = 1, \boldsymbol{D}_{N_i} = \boldsymbol{d}, z_i] = E[Y_{2i} - Y_{1i}|D_i = 0, \boldsymbol{D}_{N_i} = \boldsymbol{d}, z_i]$$

设 $\Delta_{i,\ d} = E[Y_{2i} - Y_{1i}|D_i = 1, \boldsymbol{D}_{N_i} = \boldsymbol{d}, z_i] - E[Y_{2i} - Y_{1i}|D_i = 0, \boldsymbol{D}_{N_i} = \boldsymbol{d}, z_i]$。此外，根据贝叶斯定理，

$$\begin{aligned} P(\boldsymbol{D}_{N_i} = \boldsymbol{d}|D_i = 1, z_i) &= \frac{P(D_i = 1|\boldsymbol{D}_{N_i} = \boldsymbol{d}, z_i)P(\boldsymbol{D}_{N_i} = \boldsymbol{d}|z_i)}{P(D_i = 1|z_i)} \\ &= \frac{e_i(\boldsymbol{d}, z_i)P(\boldsymbol{D}_{N_i} = \boldsymbol{d}|z_i)}{\pi_i(z_i)} \end{aligned}$$

因此，ADTT 可以改写为类似逆概率加权的形式。令 $e_i = e_i(\boldsymbol{D}_{N_i},\ z_i)$。

$$\begin{aligned} \tau_{\text{dir}} &= \frac{1}{n}\sum_{i=1}^{n}\sum_{\boldsymbol{d}}\Delta_{i,d}\frac{e_i(\boldsymbol{d}, z_i)P(\boldsymbol{D}_{N_i} = \boldsymbol{d}|z_i)}{\pi_i(z_i)} \\ &= \frac{1}{n}\sum_{i=1}^{n} E\left[\left(E\left[\left(\frac{D_i}{e_i} - \frac{1 - D_i}{1 - e_i}\right)(Y_{2i} - Y_{1i})|\boldsymbol{D}_{N_i}, z_i\right]\right)\frac{e_i}{\pi_i}\middle| z_i\right] \\ &= \frac{1}{n}\sum_{i=1}^{n} E\left[E\left[\frac{D_i - e_i}{e_i(1 - e_i)}(Y_{2i} - Y_{1i})\middle|\boldsymbol{D}_{N_i}, z_i\right]\frac{e_i}{\pi_i}\middle| z_i\right] \\ &= \frac{1}{n}\sum_{i=1}^{n} E\left[\frac{D_i - e_i}{\pi_i(1 - e_i)}(Y_{2i} - Y_{1i})\middle| z_i\right] \end{aligned}$$

证明至此完成。

S2 定理2的证明

对于特定邻居 $j \in N_i$，我们考虑 AITT 定义中的期望项。令 $\boldsymbol{D}^d = \boldsymbol{D}_{N_j}^{(d)}$，$\boldsymbol{D}^{-i} = \boldsymbol{D}_{N_j}^{-i}$。我们的目标是确定：

$$\tau_{ij} = E[y_{2j}(D_j, \boldsymbol{D}^{(1)}) - y_{2j}(D_j, \boldsymbol{D}^{(0)})|D_i = 1, z_i, z_j],$$

其中 $\boldsymbol{D}^{(d)}$ 表示 j 的邻居的处理向量，并带有 $D_i = d \in \{0, 1\}$。我们对 D_j 和 $\boldsymbol{D}^{-i}$ 进行边缘化。

$$\begin{aligned} \tau_{ij} = \sum_{d_j, \boldsymbol{d}^{-i}} & E[y_{2j}(d_j, D_i = 1, \boldsymbol{d}^{-i}) - y_{2j}(d_j, D_i = 0, \boldsymbol{d}^{-i})|D_i = 1, D_j = d_j, \boldsymbol{D}_{N_j}^{-i} = \boldsymbol{d}^{-i}, z_i, z_j] \\ & \times P(D_j = d_j, \boldsymbol{D}_{N_j}^{-i} = \boldsymbol{d}^{-i}|D_i = 1, z_i, z_j). \end{aligned}$$

令CATT$_{ij}(d_j, \boldsymbol{d}^{-i})$表示条件期望项。我们将其改写为趋势差:

$$\begin{aligned}\text{CATT}_{ij}(d_j, \boldsymbol{d}^{-i}) = E[y_{2j}(d_j, D_i = 1, \boldsymbol{d}^{-i}) - y_{1j} | D_i = 1, D_j = d_j, \boldsymbol{D}_{N_j}^{-i} = \boldsymbol{d}^{-i}, z_i, z_j] \\ - E[y_{2j}(d_j, D_i = 0, \boldsymbol{d}^{-i}) - y_{1j} | D_i = 1, D_j = d_j, \boldsymbol{D}_{N_j}^{-i} = \boldsymbol{d}^{-i}, z_i, z_j]\end{aligned}$$

我们对个体 j 应用假设2，以识别反事实趋势:

$$\begin{aligned}&E[y_{2j}(d_j, D_i = 0, \boldsymbol{d}^{-i}) - y_{1j} | D_i = 1, D_j = d_j, \boldsymbol{D}_{N_j}^{-i} = \boldsymbol{d}^{-i}, z_i, z_j] \\ &= E[y_{2j}(d_j, D_i = 0, \boldsymbol{d}^{-i}) - y_{1j} | D_i = 0, D_j = d_j, \boldsymbol{D}_{N_j}^{-i} = \boldsymbol{d}^{-i}, z_i, z_j] \\ &= E[Y_{2j} - Y_{1j} | D_i = 0, D_j = d_j, \boldsymbol{D}_{N_j}^{-i} = \boldsymbol{d}^{-i}, z_i, z_j].\end{aligned}$$

因此，CATT$_{ij}(d_j,\ \boldsymbol{d}^{-i})$ 可通过关于个体 i的处理的个体 j结果变量的条件双重差分来识别。与定理1类似，可使用以下内容将其表示为逆概率加权形式

$$e'_{ij}(d_j, \boldsymbol{d}^{-i}, z_i, z_j) = P(D_i = 1 | D_j = d_j, \boldsymbol{D}_{N_j}^{-i} = \boldsymbol{d}^{-i}, z_i, z_j)$$

.

令 $e'_{ij} = e'_{ij}(d_j,\ \boldsymbol{d}^{-i},\ z_i,\ z_j)$，则有

$$\text{CATT}_{ij}(d_j, \boldsymbol{d}^{-i}) = E\left[\frac{D_i - e'_{ij}}{e'_{ij}(1 - e'_{ij})}(Y_{2j} - Y_{1j}) \middle| D_j = d_j, \boldsymbol{D}_{N_j}^{-i} = \boldsymbol{d}^{-i}, z_i, z_j\right].$$

现在我们将其代回 τ_{ij}。我们使用贝叶斯定理:

$$P(D_j = d_j, \boldsymbol{D}_{N_j}^{-i} = \boldsymbol{d}^{-i} | D_i = 1, z_i, z_j) = \frac{e'_{ij} P(D_j = d_j, \boldsymbol{D}_{N_j}^{-i} = \boldsymbol{d}^{-i} | z_i, z_j)}{P(D_i = 1 | z_i)}.$$

$$\tau_{ij} = E\left[E\left[\frac{D_i - e'_{ij}}{e'_{ij}(1-e'_{ij})}(Y_{2j} - Y_{1j})\middle| D_j, \boldsymbol{D}_{N_j}^{-i}, z_i, z_j\right] \times \frac{e'_{ij}}{\pi_i(z_i)}\middle| z_i, z_j\right]$$
$$= E\left[\frac{D_i - e'_{ij}}{\pi_i(z_i)(1-e'_{ij})}(Y_{2j} - Y_{1j})\middle| z_i, z_j\right].$$

最后，对 $j \in N_i$ 和 $i \in I_n$ 取平均，得到 AITT 的识别结果。

S3 引理1的证明

S3.1 ADTT 的证明

令 $\phi_i = \phi_i^{\text{dir}} = \frac{D_i - e_i}{e_i(z_i)\,\pi_i(z_i)(1-e_i(z_i))}(Y_{2i} - Y_{1i})$。我们验证 ψ-依赖的定义。考虑任意的 $h, h' \in \mathbb{N}$、有界利普希茨函数 $f \in \mathcal{L}_h, f' \in \mathcal{L}_{h'}$，以及任意一对集合 $(H, H') \in \mathcal{P}_n(h, h';\, s)$，使得 $l_A(H, H') \geq s > 0$。

情形1：$s \leq 2K$。根据假设5（重叠）和假设6（有界结果），$\{\phi_i\}$一致有界。由于 f 和 f' 有界，协方差有界：

$$|\operatorname{Cov}(f(\phi_H), f'(\phi_{H'}))| \leq 4\|f\|_\infty \|f'\|_\infty.$$

情形2：$s > 2K$。令 $x_i = (Y_{1i}, D_i, \epsilon_i)$ 为基本变量。根据假设4（局部干扰），ϕ_i 是 i的 K-邻域内变量的可测函数：

$$\phi_i = g(\{x_k\}_{k \in N(i;K)^+}).$$

根据假设5和假设6，g 是利普希茨连续的。

令 $H^K = \bigcup_{i \in H} N(i;\, K)^+$。那么，$\sigma(\{\phi_i\}_{i \in H}) \subset \sigma(\{x_i\}_{i \in H^K})$。当 $s > 2K$ 时，H^K 与 $(H')^K$ 之间的距离至少为 $s - 2K$。由于 $\{x_i\}$ 是 α-混合的（假设7），我们应用 Kojevnikov等人（2021）中的命题2.2来界定协方差：

$$|\operatorname{Cov}(f(\phi_H), f'(\phi_{H'}))| \leq 4\|f\|_\infty \|f'\|_\infty \alpha(\sigma(\{x_i\}_{i \in H^K}), \sigma(\{x_i\}_{i \in (H')^K}))$$
$$\leq 4\|f\|_\infty \|f'\|_\infty \bar{\alpha}(|H^K|, |(H')^K|, s - 2K).$$

根据假设7，$\bar{\alpha}(h,\ h',\ s) \leq \psi(h,\ h')\widehat{\alpha}(s)$。假设邻域大小 $|N(i;\ K)^{+}|$ 有界于 C_K，则有 $|H^K| \leq hC_K$。

$$|\operatorname{Cov}(f(\phi_H), f'(\phi_{H'}))| \leq 4\|f\|_\infty \|f'\|_\infty \psi(hC_K, h'C_K)\widehat{\alpha}(s-2K).$$

结合这两种情况，我们定义系数。令 $\Psi(h,\ h') = \max\{1, \psi(hC_K,\ h'C_K)\}$。我们设定：

$$\psi_{h,h'}(f,f') = 4\|f\|_\infty \|f'\|_\infty \Psi(h,h'),$$

$$\tilde{\theta}_{n,s} = \mathbf{1}\{s \leq 2K\} + \mathbf{1}\{s > 2K\}\widehat{\alpha}(s-2K).$$

这满足 ψ-依赖的定义。

S3.2 AITT 的证明

定义

$$\phi_i = \phi_i^{\text{ind}} = \frac{1}{|N_i|}\sum_{j\in N_i}\left\{\frac{D_i - e'_{ij}(z_i,z_j)}{\pi_i(z_i)(1-e'_{ij}(z_i,z_j))}(Y_{2j}-Y_{1j})\right\},$$

且 $x_i = (Y_{1i}, D_i, \epsilon_i)$。项 $(Y_{2j}-Y_{1j})$ 和 e'_{ij} 取决于 $\{x_k\}_{k\in N_j^+}$，其中 $N_j^+ = N_j \cup \{j\}$。根据假设3，$j \in N_i$ 意味着 $l_A(i,j) \leq K$；同理，对于 $k \in N_j^+$，我们有 $l_A(j,k) \leq K$（因为 N_j 由 j的最近邻组成）。因此，i 与任意 $k \in N_j^+$ 之间的距离至多为 $l_A(i,j) + l_A(j,k) \leq K + K = 2K$。因此，$\phi_i$ 是 i的 $2K$-邻域内变量的函数，该邻域定义为 $\phi_i = g(\{x_k\}_{k\in N(i;2K)^+})$。根据假设5和假设6，$g$ 是利普希茨连续且有界的。

考虑集合 H, H'，使得 $l_A(H,H') \geq s$。若 $s \leq 4K$，则协方差有界于 $4\|f\|_\infty\|f'\|_\infty$。若 $s > 4K$，令 $H^{2K} = \bigcup_{i\in H} N(i;\ 2K)^{+}$。$H^{2K}$与 $(H')^{2K}$ 之间的距离至少为 $s-4K$。应用 Kojevnikov等人（2021）中的命题2.2和假设7，我们得到

$$\begin{aligned}|\operatorname{Cov}(f(\phi_H), f'(\phi_{H'}))| &\leq 4\|f\|_\infty\|f'\|_\infty \bar{\alpha}(|H^{2K}|, |(H')^{2K}|, s-4K)\\ &\leq 4\|f\|_\infty\|f'\|_\infty \psi(|H^{2K}|, |(H')^{2K}|)\widehat{\alpha}(s-4K).\end{aligned}$$

假设邻域大小有界，我们可以类似于引理 1定义 ψ-依赖系数，其衰减率如下：

$$\tilde{\theta}_{n,s} = \mathbf{1}\{s \leq 4K\} + \mathbf{1}\{s > 4K\}\widehat{\alpha}(s - 4K),$$

由此完成证明。

S4 主要定理的证明

S4.1 定理 *3* 的证明

鉴于求和项是 ψ-依赖的（引理 1），并满足大数定律的弱依赖条件（假设 8），该结果可直接由 Kojevnikov等人（2021）的定理 3.1 得出。

S4.2 定理 *4* 的证明

根据引理 1 和假设 9，Kojevnikov等人（2021）的定理 3.2（中心极限定理）的条件均得到满足，从而确立了渐近正态性。

S4.3 定理 *5* 的证明

由于引理 1 表明 ϕ_i 是 ψ-依赖的，因此在假设 10 下应用 Kojevnikov等人（2021）的命题 4.1，即可建立一致性。

S4.4 定理 *6* 的证明

我们仅针对 ADTT 估计量证明双重稳健性，因为 AITT 的证明可以用同样的方法完成。ADTT 的双重稳健估计量可写为：

$$\begin{aligned}\widehat{\tau}_{\text{dir}}^{DR} = \frac{1}{n}\sum_{i=1}^{n}\Bigg\{ & \frac{D_i}{\widehat{\pi}_i(z_i)}\left(\Delta Y_i - \Delta m_{1i}(\boldsymbol{D}_{N_i}, z_i)\right) - \frac{(1-D_i)\widehat{e}_i(z_i)}{\widehat{\pi}_i(z_i)(1-\widehat{e}_i(z_i))}\left(\Delta Y_i - \Delta m_{0i}(\boldsymbol{D}_{N_i}, z_i)\right) \\ & + \frac{\widehat{e}_i(z_i)}{\widehat{\pi}_i(z_i)}\left(\Delta m_{1i}(\boldsymbol{D}_{N_i}, z_i) - \Delta m_{0i}(\boldsymbol{D}_{N_i}, z_i)\right)\Bigg\}.\end{aligned}$$

情形 1：倾向得分模型设定正确。当 $e_i = \widehat{e}_i$ 和 $\pi_i = \widehat{\pi}_i$ 设定正确时，我们可以证明回归项相互抵消。具体而言，根据迭代期望定律以及 $E[D_i|\boldsymbol{D}_{N_i}, z_i] = e_i(\boldsymbol{D}_{N_i}, z_i)$这一事实，我们有：

$$\begin{aligned} E\left[\frac{D_i}{\widehat{\pi}_i(z_i)}\Delta m_{1i}(\boldsymbol{D}_{N_i}, z_i)\right] &= E\left[\frac{1}{\pi_i(z_i)}E[D_i\Delta m_{1i}(\boldsymbol{D}_{N_i}, z_i)|\boldsymbol{D}_{N_i}, z_i]\right] \\ &= E\left[\frac{e_i(\boldsymbol{D}_{N_i}, z_i)}{\pi_i(z_i)}\Delta m_{1i}(\boldsymbol{D}_{N_i}, z_i)\right] \\ &= E\left[\frac{D_i}{\pi_i(z_i)}(Y_{2i} - Y_{1i})\right], \end{aligned}$$

对 Δm_{0i} 项也同理。第三项 $\widehat{e}_i(z_i)\left(\Delta m_{1i} - \Delta m_{0i}\right)/\widehat{\pi}_i(z_i)$ 的期望为

$$\begin{aligned} &E\left[\frac{\widehat{e}_i(z_i)}{\widehat{\pi}_i(z_i)}\left(\Delta m_{1i}(\boldsymbol{D}_{N_i}, z_i) - \Delta m_{0i}(\boldsymbol{D}_{N_i}, z_i)\right)\right] \\ &\quad = E\left[\frac{e_i(\boldsymbol{D}_{N_i}, z_i)}{\pi_i(z_i)}\left(\Delta m_{1i}(\boldsymbol{D}_{N_i}, z_i) - \Delta m_{0i}(\boldsymbol{D}_{N_i}, z_i)\right)\right], \end{aligned}$$

这使得模型预测项与前两项相加时相互抵消。因此，双重稳健估计量简化为逆概率加权估计量。根据定理1，在假设1和假设2成立时，该估计量是一致的。

情形2：结果回归模型设定正确。当 $\mu_{1i} = \Delta m_{1i}$ 和 $\mu_{0i} = \Delta m_{0i}$ 的设定正确时，我们可以证明逆概率加权残差项的期望为零。具体而言，根据迭代期望定律：

$$\begin{aligned} &E\left[\frac{D_i}{\pi_i(z_i)}\left(\Delta Y_i - \Delta m_{1i}(\boldsymbol{D}_{N_i}, z_i)\right)\middle|\boldsymbol{D}_{N_i}, z_i\right] \\ &\quad = \frac{e_i(\boldsymbol{D}_{N_i}, z_i)}{\pi_i(z_i)}\left(E[\Delta Y_i|D_i = 1, \boldsymbol{D}_{N_i}, z_i] - \mu_{1i}(\boldsymbol{D}_{N_i}, z_i)\right) = 0, \end{aligned}$$

第二项也同理。第三项 $\widehat{e}_i(z_i)(\Delta m_{1i} - \Delta m_{0i})/\widehat{\pi}_i(z_i)$ 的期望为——

望

$$E\left[\frac{\widehat{e}_i(z_i)}{\widehat{\pi}_i(z_i)}\left(\Delta m_{1i}(\boldsymbol{D}_{N_i}, z_i) - \Delta m_{0i}(\boldsymbol{D}_{N_i}, z_i)\right)\right] = E\left[\frac{\widehat{e}_i(z_i)}{\widehat{\pi}_i(z_i)}\left(\mu_{1i}(\boldsymbol{D}_{N_i}, z_i) - \mu_{0i}(\boldsymbol{D}_{N_i}, z_i)\right)\right].$$

当结果回归模型设定正确时，由于期望是对给定 $D_i = 1$条件下 $\boldsymbol{D}_{N_i}$ 的分布取的（由权重 $\widehat{e}_i(z_i)/\widehat{\pi}_i(z_i)$表示），该期望等于 ADTT。此时，双重稳健估计量退化为这一基于回归的估计量，该估计量在假设1和假设2下是一致的。这就确立了双重稳健性：只要倾向得分模型或结果回归模型中任意一个设定正确，双重稳健估计量就是一致的。

S4.5 定理7的证明

当倾向得分模型和结果回归模型均设定正确时，采用与定理 4 类似的论证，可以证明双重稳健估计量渐近正态。关键在于证明公式（7）中定义的双重稳健影响函数 ϕ_i^{DR}满足 ψ-依赖性质（类似于引理 1）以及弱依赖条件（假设9）。

由于 ϕ_i^{DR} 是相同基本变量（Y_{1i}， D_i， ϵ_i）及其邻域的函数，且结果变量回归函数 μ_{1i}、 μ_{0i} 等在假设 6 下有界，因此根据引理 1 中相同的论证可得 ψ-依赖结构。在假设 9 下，渐近正态性随即由 Kojevnikov等人（2021年）的定理3.2得出。