簇之旅：在网络干扰下利用内部节点进行 A/B 测试

Qianyi Chen[1], **Anpeng Wu**[2], **Bo Li**[1]*,**Lu Deng**[3] **& Yong Wang**[3]†
清华大学, [2]浙江大学, [3]腾讯{cqy22@mails,libo@sem}.
tsinghua.edu.cn anpwu@zju.edu.cn{
adamdeng,darwinwang}@tencent.com

arXiv:2602.04457v1 [stat.ME] 4 Feb 2026

摘要

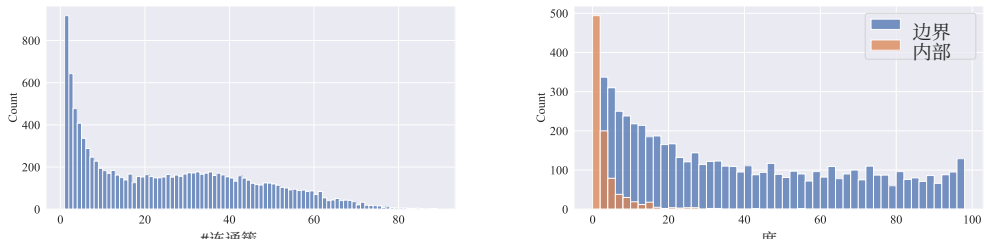
平台上的 A/B 测试常面临网络干扰带来的挑战，即实验单位的结果不仅取决于自身接受的处理，还取决于其网络邻居接受的处理。为应对这一问题，簇级随机化已成为标准做法，并使网络感知估计量得以应用。这些估计量通常会对数据进行截取，只保留信息性单元的一个子集；在适当条件下，这样做可以获得较低的偏差，但往往会导致较高的方差。本文首先证明，内部节点——即其所有邻居都位于同一簇内的单元——构成截取后子总体中的绝大多数。鉴于此，我们提出直接对内部节点取平均，构造内部均值（MII）估计量。该方法避免了现有网络感知估计量所需的复杂重新加权，并在经典设定下显著降低方差。然而，我们发现，内部节点往往无法代表总体，尤其是在网络依赖协变量方面，因此会产生显著偏差。随后，我们使用在整个网络上训练的反事实预测器来增广 MII 估计量，从而校正内部节点与总体之间的协变量分布偏移。通过重新整理表达式，我们揭示，增广 MII 估计量在解析形式上体现了预测增强推断框架中的点估计量（Angelopoulos 等人，2023a；b）。这一发现启发我们采用半监督视角，将内部节点视为受到选择偏差影响的带标签数据。大量且具有挑战性的模拟研究[1] 表明，我们的增广 MII 估计量在各种设定下均表现出色。

1 引言

长期以来，A/B 测试一直是现代平台决定是否推出新产品功能的黄金标准。然而，存在干扰时，基本的测试流程很容易失效并导致误导性结论——具体而言，当经典的稳定单元处理值假设（SUTVA）被违反，且某个实验单位的潜在结果受到相邻邻居所接受的处理影响时，便会出现这种情况。由于这些影响通常会通过网络拓扑结构传播，例如社交网络中的友谊关系，我们将这一现象称为网络干扰。

在大多数工业 A/B 测试场景中，关注的目标量是全局平均处理效应（GATE），定义为全局处理与全局对照下平均结果之差。在存在网络干扰的情况下，我们将 GATE 的估计视为一项并非易事的外推任务。我们认为，与其依赖大幅将维度降至一维或二维的、与图结构无关的方法（例如 Yu 等人（2022b）；Cortez-Rodriguez 等人（2024）；Bayati 等人（2024），利用已知的图结构更具吸引力。

*通讯作者。[1]代码可在
https://github.com/Cqyiiii/AMII-Harnessing-Interior-Nodes-for-Network-Experiments



(a) (b)

图 1：**(a)**各实验单位所连接的不同簇数量直方图。**(b)** 边界节点和内部节点度的直方图。网络拓扑结构源自 Facebook 网络（Gui 等人，2015），簇通过 Louvain 算法生成（Blondel 等人，2008）。

作为最基本的方法，均值差估计量与图结构无关，因此在存在干扰时会产生严重偏差（Karwa 和 Airoldi，2018）。这一问题在早期实验中尤为突出，此时处理比例通常低于10%。在这种情况下，处理组单位的邻居中往往有很大比例属于对照组，这意味着其一跳自我网络中的处理向量可能与全局处理向量显著不同。

为了利用图拓扑结构，核心方法论是图簇随机化（Hud-gens & Halloran，2008；Ugander 等人，2013），它确保簇内所有实验单位接受相同的处理。这些簇通常通过图聚类算法预先生成。与经典的单位级随机化不同，图簇随机化会使连接紧密的单元之间产生强相关性。这使内部节点所处的环境能够近似全局处理，从而便于进行 GATE 估计。

此外，广泛考虑的一种情形是邻域干扰假设（NIA），其非正式地假定干扰仅限于某个实验单位的一跳邻居。在 NIA 和簇级随机化下，采用网络暴露指标的 Horvitz–Thompson（HT）估计量（Ugander 等人，2013；Ugander & Yin，2023）是无偏的。该估计量仅纳入满足特定暴露条件的单元的结果，例如所有邻居都接受相同处理水平的单元。在不存在歧义时，我们简称其为 HT估计量，并将其正式定义留待第 2 节给出。

在图簇随机化中，位于簇内部的节点始终是干净节点，因此会纳入HT估计量。然而，这些内部节点仅占总体人口的一小部分。相比之下，簇边界上有更多节点，但它们被HT估计量选中的可能性要低得多。为说明这种差异，考虑一个与 c个不同簇相连的实验单位，在簇级进行处理比例为 p 的独立伯努利随机化。该实验单位成为无干扰处理节点的概率为 p^c，这个概率可能小到趋近于零。为帮助理解，图1a展示了我们的社交网络中 c 的分布。我们观察到，内部节点仅占总人口的8%。相比之下，边界节点的 c往往很大，导致估计量中出现形式为$(1/p)^c$ 的极端逆概率权重，从而使方差大幅膨胀。

为解决HT估计量方差爆炸的问题，人们已进行了大量探索。一方面，有研究提出了改进的随机化方案，以控制HT估计量的理论方差界（Ugander& Yin，2023；Kandiros等人，2025）。然而，鉴于实际社交网络固有的高密度，即使达到了紧的理论界，实际实现方差仍不尽如人意。另一方面，出现了一些替代估计量，它们以引入轻微偏差为代价换取显著的方差缩减，例如簇自适应估计量（CAE；Liu等人，2024年）。CAE同样使用暴露指示变量来筛选干净节点，但它以结果的双层平均取代逆概率加权——先在簇内平均，再在簇间平均。尽管如此，我们认为，实际的聚类算法很少满足CAE所要求的强结构假设，而且其双层平均步骤还可以进一步简化为单次平均，从而大幅缩减方差。

在本文中，我们首先提出内部均值估计量。该估计量通过为单元结果赋予温和得多的权重，模拟均值差估计量；这与指数级

权重形成对比，这些权重用于 HT 估计量，也是其方差极高的主要来源。我们的估计量基于簇级随机化的广泛实践，仅依赖簇的内部节点来计算平均结果。直观而言，这些内部节点往往处于更为纯净的环境中，更接近全局处理或对照条件。基于与 Liu等人(2024年相似的假设，具体而言，假设每个簇的内部节点足以代表整个簇，我们证明了 MII 估计量的一致性。

此外，我们观察到，HT、CAE 和 MII 估计量都排除了大多数边界节点；当内部单元与边界单元之间存在系统性差异时，这种排除不可避免地会引入选择偏差。为应对这一挑战，我们纳入了一个基于整个网络训练的反事实预测器。借助该预测器，我们为 MII 估计量构造了一个调整项，以减轻内部单元与边界单元之间的协变量不匹配。此外，通过重新排列估计量中的各项，我们从半监督学习的视角给出另一种解释：内部节点充当带标签数据的角色——尽管可能存在选择偏差——而边界节点则提供具有部分标签信息的代表性协变量。

相关工作 干扰下的因果推断已在早期研究中得到广泛探讨（Sobel，2006；Rosenbaum，2007）。该领域的一项奠基性方法是图簇随机化（Hudgens 和 Halloran，2008；Ugander 等人，2013），它在簇级分配处理，并确保簇内所有实验单位接受相同的处理。这种方法提升了若干经典估计量的统计性能，因此成为常见做法。此外，由于一般干扰通常难以处理，研究者提出了许多结构假设，以简化干扰模式并便于分析。

大量研究聚焦于部分干扰假设（ 2008；Bhattacharya 等，2020；Forastiere 等，2021；Candogan 等，2023），该假设认为网络可划分为互不相交的簇，且干扰仅发生在各簇内部。尽管这一假设大大简化了干扰结构，使簇级随机化能够产生无偏估计，但它具有一定限制，尤其不适用于社交网络和在线市场等场景。

除了这一特殊情形外，研究者还提出了三种主要的结构假设来约束干扰的形式。文献中研究最广泛的第一种干扰模式是邻域干扰（Forastiere 等，2022；Cortez 等，2022；Ugander & Yin，2023；Liu等人，2024），该模式假设干扰仅限于个体的直接邻居。在这一假设下，若某实验单位周围的邻居均接受相同的处理水平，则可认为该实验单位处于全局处理或对照环境中，因此特别适用于 GATE 估计。在更强的参数结构假设下，研究者提出暴露映射（Aronow & Samii，2017；Eckles 等，2016；Ugander 等人，2013；Baird 等，2018；Vazquez-Bare，2023），通过预先定义的函数或表征，总结整个处理向量对特定实验单位的影响。此外，研究者还利用潜在结果模型来设计更复杂的实验设计（Basse & Airoldi，2018；Yu 等人，2022a；Harshaw 等，2023；Chen 等，2023）。

此外，已有一系列研究将图神经网络（图神经网络）作为强大的回归模型，用于对观测数据进行反事实预测（Leung& Loupos，2024年；Cai 等人，2024年；Wu 等人，2025）。具体而言，GATE 可视为边界点处两个反事实之差，即全局处理与全局对照。

2 基本设定

2.1 预备知识

我们考虑一个由 n个实验单位构成的有限总体，这些实验单位通过具有已知拓扑结构的干扰网络相互连接。该网络表示为无向图 $\mathcal{G} = (\mathcal{V}, \mathcal{E})$，其中节点集 $\mathcal{V}$对应实验单位集 $[n] = \{1, 2, \ldots, n\}$，边由邻接矩阵 A描述。我们将节点 i 的度记为 $\deg_i$，并将实验单位 i 的 r 跳邻域定义为$\mathcal{N}(i, r) = \{j : \ell(i, j) = r\}$，其中 $\ell(i, j)$ 是实验单位 i与 j之间的最短路径距离。

我们采用 Neyman–Rubin 潜在结果框架（Rubin，1974）。处理分配由向量 $\mathbf{z} = (z_1, z_2, \ldots\ldots, z_n) \in \{0,1\}^n$表示。我们将实验单位 i的潜在结果定义为 $Y_i = Y_i(\mathbf{z})$，其中 Y_i 可能取决于其他实验单位的处理分配。

我们关注的参数是全局平均处理效应（GATE），其定义如下：

$$\tau := \frac{1}{n} \sum_{i \in [n]} \left(Y_i(\mathbf{1}) - Y_i(\mathbf{0}) \right), \tag{1}$$

其中**1**和**0**分别表示 n 维的全 1 向量和全 0 向量，分别代表全局处理条件和全局对照条件。

现在介绍我们的随机化方案——簇级独立伯努利随机化。首先，我们应用社区检测算法生成簇，并在本文中将这些簇视为已知。我们将得到的簇记为 $C_1, C_2, \ldots\ldots, C_K$，它们构成节点集 $\mathcal{V}$ 的一个划分。

其次，在每次实验中，处理在簇级进行分配，这意味着同一簇中的所有实验单位都接受相同的处理。这些簇的处理分配独立地从伯努利分布中抽取。

此外，对于给定的簇 C_k，内部集合定义如下：

$$\mathrm{Int}_k = \{ i : i \in C_k, \mathcal{N}(i,1) \subseteq C_k \}. \tag{2}$$

也就是说，如果一个节点的所有一跳邻居也都包含在同一个簇中，那么该节点就属于内部。边界集定义为簇内的补集：$\mathrm{Bnd}_k = C_k \setminus \mathrm{Int}_k$。内部集与边界集共同构成该簇的划分。直观而言，内部节点对于估计 GATE 尤为重要。在簇级随机化下，它们的一跳邻居具有相同的处理水平，因此其局部环境与全局处理或对照的情形十分相似。

接下来，我们正式引入邻域干扰假设（NIA），该假设将干扰效应限制在每个实验单位的一跳邻域内。

假设 2.1（NIA） 若 $\mathbf{z}^j = \mathbf{z}'_j$ 对所有 $j \in \mathcal{N}(i,1)$ 均成立，则有 $Y_i(\mathbf{z}) = Y_i(\mathbf{z}')$。

这一假设在文献中十分常见，因为它对干扰施加了高层次的结构约束，同时又不限定暴露函数的具体形式——也就是说，它不要求使用参数模型来描述邻居的处理如何汇总并影响结果。在本文余下部分，我们主要在这一假设下开展分析。

2.2 替代估计量

现在介绍三种常用的替代估计量，并将它们与我们的方法进行比较。第一种是均值差估计量（DIM），它不考虑干扰，只计算处理组单位与对照组单位之间的平均结果差异。

$$\hat{\tau}_{DIM} = \frac{z_i Y_i}{\sum_{i \in [n]} z_i} - \frac{(1 - z_i) Y_i}{\sum_{i \in [n]} (1 - z_i)}. \tag{3}$$

第二种是具有网络暴露的 HT估计量（Ugander & Yin，2023）。

$$\hat{\tau}_{HT} = \frac{1}{n} \sum_{i \in [n]} \left(\frac{\delta_i(1)}{\mathbb{E}\left[\delta_i(1)\right]} - \frac{\delta_i(0)}{\mathbb{E}\left[\delta_i(0)\right]} \right) Y_i, \tag{4}$$

其中，δ_i 是暴露指标：$\delta_i(z_0) = \mathbb{I}\{\sum_{j \in \mathcal{N}(i,1)} z_j / \text{度}_i = z_0,\ z_i = z_0\}$。这一定义意味着，HT估计量仅依赖干净节点的结果——这些节点的所有邻居都接受与该节点自身相同的处理水平。这种暴露条件通常称为全邻域暴露（Ugander 等人，2013）。在邻域干扰假设下，HT估计量可保证无偏，因此在理论上颇具吸引力。

最后，我们介绍一种更精细的估计量——CAE（Liu等人，2024年）。与前面讨论的HT估计量类似，CAE使用相同的指示变量来筛选干净节点。不过，它不使用逆概率权重，而是对观测结果计算双层平均。因此，它

可以解释为均值差估计量和HT估计量的混合体。为定义该估计量，设 t_k 表示分配给簇 C_k的处理。我们首先计算簇内平均值：

$$\hat{Y}_{k,z_0} = \sum_{i \in C_k} \frac{\delta_i(z_0) Y_i}{\sum_{j \in C_k} \delta_j(z_0)}. \tag{5}$$

随后，对各簇取外层平均值，得到 CAE：

$$\hat{\tau}_{CAE} = \frac{\sum_{k \in [K]} t_k Y_{k,1}}{\sum_{l \in [K]} t_l} - \frac{\sum_{k \in [K]} (1 - t_k) Y_{k,0}}{\sum_{l \in [K]} (1 - t_l)}. \tag{6}$$

3 利用内部节点

3.1 内部均值估计量

现在我们介绍 MII 估计量，它是处理组与对照组内部节点结果之差。这里，$\text{Int} = \cup_{k \in [K]} \text{Int}_k$ 表示各簇内部集合的并集。

$$\hat{\tau}_{MII} = \frac{\sum_{i \in \text{Int}} z_i Y_i}{\sum_{j \in \text{Int}} z_j} - \frac{\sum_{i \in \text{Int}} (1 - z_i) Y_i}{\sum_{j \in \text{Int}} (1 - z_j)}. \tag{7}$$

作为一种基于均值差法的估计量，MII 估计量对观测结果赋予适中的权重。实证上，我们发现，CAE 估计量的偏差与 MII 估计量相近，而 MII 估计量的方差则显著更低。此外，在与Liu等人（2024年）类似的假设下，我们证明了 MII 估计量具有一致性，证明见附录 B.1。

假设 3.1 *MII* 估计量一致性的技术假设。

1. 所有簇中内部节点的比例渐近趋于均匀：

$$\max_{i \in [K]} \left| \frac{|\,\text{Int}_i\,|}{\sum_{k \in [K]} |\,\text{Int}_k\,|} - \frac{n_i}{\sum_{k \in [K]} n_k} \right| = o_p\left(\frac{1}{K}\right). \tag{8}$$

2. 内部节点能够为每个簇提供具有充分代表性的样本：

$$\begin{aligned} \max_{k \in [K]} |\bar{Y}_k(\mathbf{1}) - \bar{Y}_{k,Int}(\mathbf{1})| &= o_p(1), \\ \max_{k \in [K]} |\bar{Y}_k(\mathbf{0}) - \bar{Y}_{k,Int}(\mathbf{0})| &= o_p(1). \end{aligned} \tag{9}$$

这里，$\bar{Y}_k(\mathbf{1})$ 和 $\bar{Y}_{k,\text{Int}}(\mathbf{1})$ 分别表示簇 k 内的平均结果和簇 k 内部节点的平均结果。渐近分析针对的是簇的数量 K。

定理 3.1 假设 *2.1* 和 *3.1* 成立；那么 *MII* 估计量是一致的，即：

$$\hat{\tau}_{MII} - \tau = o_p(1). \tag{10}$$

我们首先指出，假设 3.1 的两个部分分别规定了簇间的均匀性和内部节点的代表性。值得注意的是，我们的要求明显弱于Liu等人 (2024年)——例如，他们假设每个簇内的平均结果均为相同的常数。这种差异部分源于我们目前仅关注一致性，而他们的研究旨在建立中心极限定理（中心极限定理）。

在实践中，针对 GATE 等全局估计目标，在网络干扰下推导中心极限定理并进行后续统计推断本质上十分困难（Kandiros等人，2025）。即便采用特定的实验设计和估计量，并且接受各种无法验证的技术假设，对网络密度仍存在严格限制；例如，两跳连通图的最大度 $\mathcal{G}^2$ 应为 $o\left(n^{1/9}\right)$ （Kandiros等人，2025）。

总之，与偏差分析相比，刻画方差通常要复杂得多，因为网络会引致错综复杂的相互依赖。为提供直观理解，我们使用如下潜在结果模型来说明 MII 估计量的方差：

$$Y_i(\mathbf{z}) = \beta z_i + h\left(\frac{\sum_{j\in\mathcal{N}(i,1)} z_j}{\deg_i}, v_i\right). \tag{11}$$

这里，β 是直接处理效应，协变量 v_i 在所有实验单位间独立同分布，而 h 是干扰函数。该模型将经典的均值线性模型（Aronow &Samii，2017）和多项式干扰模型（Cortez 等，2022）纳为特例。

这类潜在结果模型的一个重要性质是，无论 h的复杂程度如何，每个内部节点都能为全局处理或对照下的平均结果提供无偏估计。这表明，只需对所有内部节点进行等权平均。根据柯西-施瓦茨不等式，当仅有内部节点可用时，等权重可得到方差最小的无偏估计量。原则上，加入边界节点可以进一步降低方差。然而，在网络干扰下，这样做需要使用精细设计的权重——例如，HT估计量中针对网络暴露的逆倾向权重，或 CAE 中的双层聚类相关权重——而这些权重实际上可能导致效率进一步降低。

4 调整协变量分布差异

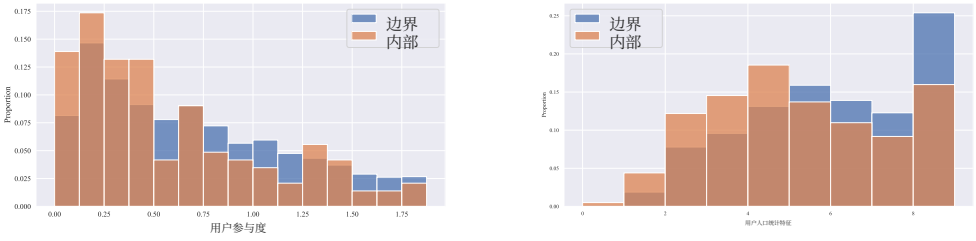
4.1 动机

首先，我们认为，对簇施加的结构假设——例如内部节点具有代表性，或簇内平均结果的期望为常数（Liu等人，2024年）——往往较为脆弱，恰恰是因为实践中使用的聚类算法通常较为粗糙。

在实践中，用于生成簇的可扩展算法通常是一种启发式算法。旨在最小化特定目标的图聚类通常是一项具有挑战性的任务，往往涉及 NP-hard 组合优化问题，例如相关聚类（Pouget-Abadie 等，2019）和平衡聚类（Brennan 等，2022）。在实践中，拥有数亿用户的社交平台通常会构建数百万个簇，以满足对 A/B 测试的巨大需求，而每周进行的 A/B 测试数量可能达到数千次。这一大规模问题要求采用启发式方案或快速但具有启发性的算法，从而使更为复杂的算法（例如因果聚类（

Viviano 等人，2025），这涉及求解半定规划，因此不适用。

事实上，在拥有十亿级用户的社交平台上，我们观察到，某些用户参与度指标——几乎所有实验都会记录的主要结果变量之一——在内部和边界子群体之间存在明显差异。图2展示了这一现象经脱敏处理后的实证证据。然而，正如引言中所述，大多数不含回归成分的估计量往往高度依赖内部节点，例如参见公式（4）和（6）。



(a) (b)

图 2：**(a)** 结果变量（用户参与度）以及**(b)** 人口统计属性在内部单元和边界单元中的分布。数据收集自一个拥有十亿级用户的社交平台。

4.2 借助反事实预测器弥合差距

由于我们的 MII 估计量也采用适度加权，而非逆倾向加权，因此同样容易受到上述选择偏差的影响。为了解决这一问题，我们提出在整个图上训练一个反事实预测器。具体而言，我们引入反事实预测的概念：给定处理向量 $\mathbf{z}$、协变量矩阵 X 和网络拓扑 A，我们定义一个函数 $f(\mathbf{z}\ X,\ A)$，用于预测在处理分配为 $\mathbf{z}$ 时所有实验单位的结果。函数f 可以灵活地参数化以适配数据；例如，它可以实例化为图神经网络。

在随机化下给定观测结果$\mathbf{Y(z)}$，我们通过回归训练函数 f。随后我们将处理机制切换为 $\mathbf{z}\ =\mathbf{1},\mathbf{0}$，并将增广 MII 估计量定义为：

$$\begin{aligned}\hat{\tau}_{AMII} = \hat{\tau}_{MII} + &\left(\frac{1}{n}\sum_{j\in[n]} f(\mathbf{1}, X, A)_j - \frac{1}{s_1}\sum_{i\in\text{Int}} z_i f(\mathbf{1}, X, A)_i\right) \\ &- \left(\frac{1}{n}\sum_{j\in[n]} f(\mathbf{0}, X, A)_j - \frac{1}{s_0}\sum_{i\in\text{Int}} (1-z_i) f(\mathbf{0}, X, A)_i\right).\end{aligned} \tag{12}$$

此处，s_1（s_0）表示内部集合中接受处理（对照）的单元数量。调整项刻画了总体的平均预测结果与接受处理（对照）的内部节点的平均预测结果之间的差异。为说明 AMII估计量背后的思想，我们考虑如下潜在结果模型，其中包含处理与一个可以依赖于网络的协变量之间的额外交互项。

$$Y_i(\mathbf{z}) = (\beta + \alpha u_i) z_i + h\left(\frac{\sum_{j\in\mathcal{N}(i,1)} z_j}{\deg_i}, v_i\right). \tag{13}$$

此处，v_i 是可能独立同分布的协变量，但我们允许 u_i 依赖于网络，因而为非独立同分布；例如，u_i 可以表示归一化度 $\deg_i/\deg$，其分布在边界与内部子群体之间有所不同。此外，β和 α 是未知标量参数，h是未知干扰函数，可能是非线性的。基于该模型以及 AMII估计量的形式，我们给出如下直观理解：输入选择得当的回归模型通常能够有效捕捉交互项。然而，在早期实验（较小的 p）中，它们可能难以外推干扰函数 h——而这正是 MII 估计量展现比较优势的部分。

更广泛地说，我们发现，在网络干扰下估计 GATE 的核心挑战，是对干扰项进行外推。实践中，我们通常只能观察到处理比例较低时的结果（例如 $p = 0\%$、5%、10%），而我们的目标是将其外推至完全处理的情形（$p = 100\%$）。一个关键观察是，对于边界单元 i，网络暴露项$\sum_{j\in\mathcal{N}(i,1)} z_j/\deg_i$往往集中在 p附近，这反映出一种远离全局处理的状态。这表明，外推的主要挑战在于恢复干扰函数 h。正如Chen &Li (2024) 所示，当 h 为非线性（例如平方根或二次函数）时，朴素的基于回归的估计量性能可能会显著下降；线性情形则不同。

接下来，我们将对 AMII 估计量的直觉进行形式化刻画，并着重说明调整项所实现的偏差降低。为处理网络依赖以及回归函数形式所带来的复杂性，我们对 f 施加参数假设。

假设 4.1（回归模型）给定 n 样本 $\{(u_i, v_i, z_i, Y_i)\}_{i=1}^n$，我们假设训练得到的回归函数 f 具有以下形式：

$$f(\mathbf{z}, X, A)_i = (\hat{\beta}_n + \hat{\alpha}_n u_i) z_i + \text{MEAN}(\{g_\theta(z_j, v_j) \mid j \in \mathcal{N}(i,1)\}). \tag{14}$$

ˆ这里，β_n,ˆα_n 表示线性部分的估计系数，g_θ 是一个可学习变换函数（例如多层感知机），而均值 表示对一组实数进行均值池化。

这种函数形式仿照部分线性模型。我们假设线性部分设定正确，而非线性部分——干扰函数——可能设定错误。对于这个非线性部分，我们采用一种单层图卷积形式。

定理 4.1 给定假设*4.1*、方程（*13*）中的潜在结果模型以及网络拓扑 $\mathcal{G}$，*MII* 估计量和 *AMII*估计量的偏差分别为：

$$\begin{aligned}\text{Bias}(\hat{\tau}_{MII}) &= \alpha(\mu_{\text{Int}} - \mu)\\ \text{Bias}(\hat{\tau}_{AMII}) &= (\mathbb{E}[\hat{\alpha}_n] - \alpha)(\mu - \mu_{\text{Int}}),\end{aligned} \tag{15}$$

其中 $\mu = \mathbb{E}[u_i]$、 $\mu_{\text{Int}} = \mathbb{E}[u_i \mid i \in \text{Int}]$。

定理4.1 的证明见附录 B.2。总的来说，我们观察到，MII 中的偏差源于内部与总体之间的差异。此外，当线性部分被正确指定时，量 $|\mathbb{E}[\hat{\alpha}_n] - \alpha|$ 通常远小于 $|\alpha|$，这表明偏差得到了显著降低。我们还发现，在内部与总体之间不存在分布差异时，AMII估计量是一种无害的调整；也就是说，它不会引入任何额外偏差。

由于这些结果依赖于参数假设，我们通过系统性模拟研究评估 AMII 估计量的性能，在该研究中，结构假设以各种方式遭到违反。尽管如此，AMII估计量仍表现出稳健性和出色的性能。

4.3 半监督的故事

我们随后给出我们方法的另一种解释。为简便起见，我们聚焦于估计全局处理下的平均结果。通过重新整理AMII估计量的表达式，我们得到：

$$\hat{\tau}_{AMII,1} = \frac{1}{n}\sum_{j\in[n]} f(\mathbf{1}, X, A)_j + \frac{1}{s_1}\sum_{i\in\text{Int}} z_i \left(Y_i - f(\mathbf{1}, X, A)_i\right). \tag{16}$$

这体现了PPI中点估计量的一种形式（Angelopoulos 等人，2023a；b），我们在这一严谨框架内解释我们的估计量。具体而言，我们将受处理的内部节点的结果视为可靠的有标签样本，而边界节点的结果仅用于训练预测器 f。目标是估计总体平均结果（即标签均值），即使用一小部分带标签数据（内部单元），以及大量仅观测到协变量的无标签样本。PPI背后的核心思想之一，是利用少量真实标签来校正纯基于模型的预测的偏差，从而减轻估计误差。这一视角与我们估计量的这一重组形式相一致。

然而，由于网络干扰的复杂性，“真实标签”的均值仍会产生偏差。如前所述，并如图2所示，该偏差来源于簇的形成，即内部单元与边界单元之间存在显著的分布差异，这与PPI框架中的完全随机缺失设定形成对照。因此，我们的AMII估计量应首先被定位为**利用预测去偏**，借此我们应对“真实标签”均值的这种选择偏差。相比之下，PPI点估计量与经典双重稳健估计量的核心思想，都是利用少量有标签样本对预测均值进行去偏。

5 模拟研究

5.1 基本设置

首先，由于我们的方法论旨在降低均方误差（MSE），而非单纯缩减方差，因此必须展示偏差降低的效果，尤其是针对所提出的 AMII 估计量。然而，由于真实平台实验中的效应十分复杂，例如存在许多并行实验、时间效应、推出其他新特征等，我们无法获取 GATE 的真实值，这妨碍了对偏差的评估。因此，我们遵循现有文献的惯例，进行 1,000 次重复的蒙特卡洛模拟，其中随机性来自潜在结果中的处理分配（随机化）和外生噪声。随后，我们根据这些重复实验评估偏差、方差和均方误差。

数据。我们使用的 Facebook 社交网络[2] 由 11,586 个节点和 568,309 条边组成。为生成簇，我们采用 Louvain 算法，并报告分辨率参数

[2]网络拓扑结构可在 https://networkrepository.com/socfb-Stanford3.php 获取。

$\gamma = 5$，由此得到 95 个簇。分辨率主要影响簇的数量，对内部节点的比例影响较小（约为 8%）。基于这一聚类结果，我们在簇级实施独立的伯努利随机化，并评估三种处理比例 $p \in \{0.1, 0.3, 0.5\}$，对应于上线前开展的三阶段实验。

我们重点关注全局处理下平均结果的估计，并为简单起见，设 $Y_i(\mathbf{0}) = 0$。若该假设不成立，只需进行简单调整——减去基线水平（Yu 等人，2022a）——即可将问题转化为这种情形。

潜在结果模型。我们采用 Chen 和 Li (2024年提出的二跳干扰框架，并进一步纳入两个协变量：度，以及每个实验单位所连接的不同簇的数量，分别记为 $\deg_i$ 和 c_i。这些协变量在内部节点和边界节点之间的分布有所不同，因此导致代表性假设不成立。

$$\mathbf{Y}(\mathbf{z}) = \beta\mathbf{z} + B\mathbf{z} + \frac{1}{2}\left(\frac{\deg}{\overline{\deg}} + \frac{\mathbf{c}}{\bar{c}}\right) \odot \mathbf{z} + \sigma\epsilon. \tag{17}$$

这里，B 是干扰矩阵，定义如下：

$$B = (\mathbf{1}\mathbf{1}^\top - I_n) \odot \left(\sum_{l=1}^{2} r_l (D^{-1}A)^l\right). \tag{18}$$

在此公式中，D 是对角度矩阵，$\odot$表示逐元素乘积。该模型自然地纳入了线性两跳干扰，其强度由 r_2刻画。为避免与直接处理效应混淆，B 的对角元素被移除。

此外，两个标准化协变量与处理向量 $\mathbf{z}$ 相互作用。最后，$\epsilon \sim \mathcal{N}(\mathbf{0}, I_n)$ 表示外生高斯噪声，而 σ 控制该噪声的强度。

参数设置。我们设定 $(\beta, r_1) = (1, 1)$，并考虑两种水平的两跳干扰，$r_2 \in \{0, 1\}$，分别对应邻域干扰假设成立以及存在显著两跳干扰的情况。此外，遵循Liu等人(2024年) 的做法，我们通过设定 $\sigma = 2$来考察信噪比较低的情形。这种情形在工业实践中尤为常见，因为大多数新产品特征对关注的结果往往只有较小的影响。

基线方法。我们总共评估五种方法：H´ajek 估计量、簇自适应估计量（CAE）、内部均值（MII）估计量、使用图神经网络（GNN）的反事实预测方法，以及增广 MII 估计量（AMII）。

我们现在补充一些说明。由于HT估计量的方差在复杂情形下往往急剧膨胀，我们采用其变体——H´ajek估计量（Ugander & Yin，2023）——它对权重进行自归一化，从而以偏差的小幅增加为代价大幅降低方差。鉴于潜在结果模型本质上是线性的，我们使用三个切比雪夫卷积层（Defferrard etal.，2016）来构建图神经网络，且不引入任何非线性激活函数。总体而言，这并非一种刻意追求精细的架构。

5.2 结果与讨论

由于篇幅限制，我们在正文中仅展示使用聚类分辨率 $\gamma = 5$ 的结果，如图3所示。完整实验——包括其他**分辨率水平**、不同的**协变量设定**、不同的**网络拓扑**，以及一系列消融研究——见附录A。在此，我们总结结果中的主要结论。

首先，我们观察到 AMII估计量的出色表现：与三个不含回归成分的估计量相比，它在保持方差相近的同时，显著降低了偏差。即使在 $p = 0.1$ 和图神经网络表现不佳时，这也意味着均方误差大幅降低。值得注意的是，存在两跳干扰（即 $r_2 = 1$）时，偏差降低的效果更加显著。此外，随着交互项系数从 0.5 增至 1，AMII 的优势也愈发明显。

其次，我们考察图神经网络估计量的表现。一个显著的模式是，它的表现高度依赖于较高的处理比例；在这种情况下，更多实验单位对于估计全局处理下的结果具有信息价值。

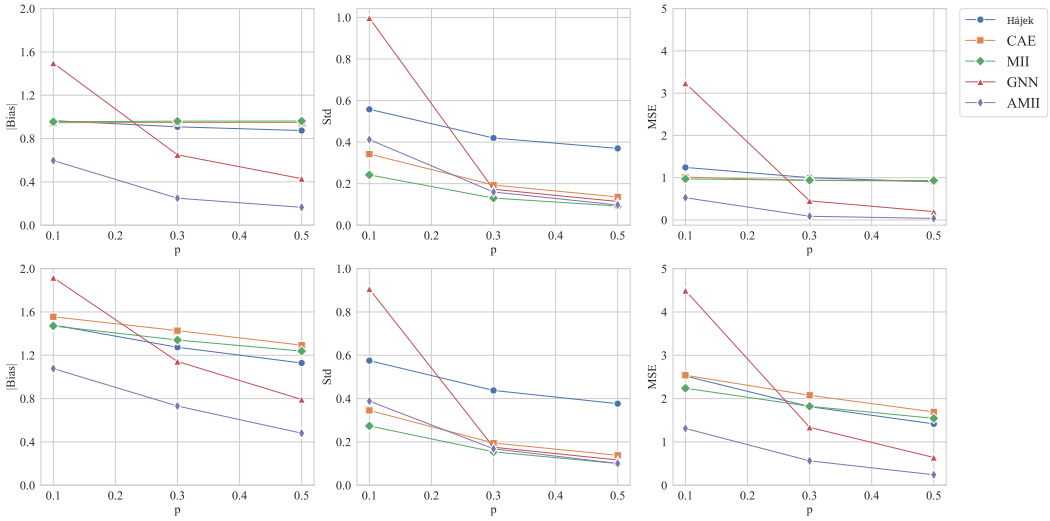


图3： $r_2 = 0$ （第一行）和 $r_2 = 1$ （第二行）下的统计性能。

接下来，我们考察 MII 估计量的表现。在所有设置中，它的方差始终最低，而偏差水平与 H´ajek 估计量和 CAE 相近。这种偏差是协变量$\deg_i$ 和 c_i的影响所致，属于预期之中。不存在交互项和两跳干扰时（见表 10），我们发现 MII 几乎无偏，与 CAE 类似，同时仍保持最低方差，因此在该设置下表现最佳。尽管如此，AMII 估计量的表现也很好。这体现了调整项的一种无害性性质。

此外，我们讨论了聚类分辨率的影响。相应的统计数据见表 1，结果见表 2 至 7。由于簇间边的模式会在聚类形成过程中发生变化，偏差的行为很难分析。关于方差，我们注意到，随着 γ 从 2 增加到 5，估计量的标准差通常会降低（Há´jek 估计量除外），这是因为“簇级样本”的数量增加了。尽管如此，仍然存在不可约方差分量，这一点从以下事实可以看出：当簇的数量进一步翻倍时（γ 从 5 增加到 10），估计量方差几乎保持不变。这与经典渐近框架存在显著差异，也再次说明了网络干扰下复杂的方差结构。

最后，我们讨论一项额外的消融研究，考察在 f 整个图上训练还是仅在边界节点上训练能够带来更好的表现。后者是一种自然的样本拆分思路，但我们注意到，网络数据中的复杂依赖关系使得很难从这一选择中得出明确结论。相关结果见表12 和13。从实证结果来看，当处理比例较低（例如 $p = 0.1$）时，在总体上训练的效果要好得多。当 p增加时，这两种方法的表现趋于相似。其背后的直觉是，随着 p 增加，边界节点会变得更具信息量；例如， $\mathbb{E}[\delta_i$期望暴露$] = p^{c_i}$ 对于水平较高的节点会迅速 c_i增长。这意味着，随着 p 上升，估计量能从局部邻域结构中捕获更多信号，从而有效利用边界信息，而这类信息在低暴露情形下通常较为稀疏。

6 结论

本文基于簇级随机化的大量实践，系统地指出了现有估计量在 GATE 估计中的局限性，并提出内部均值估计量，以消除不必要的重新加权，进一步缩减方差。鉴于该方法高度依赖内部节点，且聚类可能带来选择偏差，我们进一步提出一个调整项，并从一种新颖的半监督视角提升其地位。我们将这一调整概括为“利用预测进行去偏”，这对于纠正选择偏差本身也具有独立的研究价值。通过一系列具有挑战性的模拟研究，我们展示了该方法论的卓越表现，尤其是在早期实验中。

致谢

本研究得到国家自然科学基金委员会（编号：72171131）资助。

伦理声明

本研究遵循ICLR 道德准则。本研究不涉及人类受试者、可识别个人身份的信息或敏感数据。所有使用的数据集均为公开数据集，且已在研究社区中广泛采用；我们遵循推荐实践，以确保公平性、隐私和可复现性。我们提出的方法仅用于学术研究，不以造成伤害或助长滥用为目的。我们不了解与本研究相关的任何利益冲突、资助方面的顾虑或伦理风险。

可复现性声明

我们已尽力确保结果的可复现性。复现实验所需的全部代码和脚本均已在补充材料中提供。理论结果的完整证明也收录于附录中。这些资源旨在支持独立验证，并促进未来研究在我们工作的基础上继续开展。

大语言模型（LLM）的使用

作者仅使用大语言模型润色最终稿的措辞并纠正语法错误。所有观点、分析和结论均为作者原创。

参考文献

Anastasios N Angelopoulos、Stephen Bates、Clara Fannjiang、Michael I Jordan 和 Tijana Zrnic。预测驱动推断。《科学》，382(6671):669–674，2023a。

Anastasios N Angelopoulos、John C Duchi 和 Tijana Zrnic。Ppi++：高效的预测驱动推断。*arXiv* 预印本 *arXiv:2311.01453*，2023b。

Peter M Aronow 和 Cyrus Samii。在一般干扰下估计平均因果效应，并将其应用于一项社交网络实验。《应用统计年鉴》，11(4):1912–1947，2017。

¨Sarah Baird、J Aislinn Bohren、Craig McIntosh 和 Berk Ozler。《干扰存在时实验的最优设计》。《经济学与统计学评论》，100(5):844–860，2018年。

Guillaume W Basse 和 Edoardo M Airoldi。《网络相关结果下的模型辅助实验设计》。《生物计量学》，105(4):849–858，2018年。

Mohsen Bayati、Yuwei Luo、William Overman、Sadegh Shirani 和 Ruoxuan Xiong。《复杂干扰下实验的高阶因果消息传递》。载于 第三十八届神经信息处理系统年会，2024年。网址：https://openreview.net/forum?id=3vJbgcjgvd。

Rohit Bhattacharya、Daniel Malinsky 和 Ilya Shpitser。《干扰与网络不确定性下的因果推断》。载于 人工智能中的不确定性，第1028–1038页。PMLR，2020年。

Vincent D Blondel、Jean-Loup Guillaume、Renaud Lambiotte 和Etienne Lefebvre。在大型网络中快速发现社群。统计力学：理论与实验期刊，2008 (10)：P10008，2008。

Jennifer Brennan、Vahab Mirrokni 和 Jean Pouget-Abadie。《单侧二部实验的整群随机设计》。《神经信息处理系统进展》，35:37962– 37974，2022年。

Chencheng Cai、Xu Zhang 和 Edoardo Airoldi。《网络干扰下用于估计处理效应和溢出效应的独立集实验设计》。载于第十二届国际学习表征会议，2024年。Ozan Candogan、Chen Chen 和 Rad Niazadeh。《相关簇随机实验：稳健方差最小化》。《管理科学》，2023年。Qianyi Chen 和 Bo Li。《只需逐步增加：释放基于回归的估计量在网络干扰下 A/B 测试中的潜力》。*arXiv* 预印本 *arXiv:2410.12740*，2024年。Qianyi Chen、Bo Li、Lu Deng 和 Yong Wang。《干扰下社交网络 A/B 测试的优化协方差设计》。《神经信息处理系统进展》，36:37448– 37471，2023年。Mayleen Cortez、Matthew Eichhorn 和 Christina Yu。《阶梯式推出设计可在缺乏网络知识的情况下实现干扰下的因果推断》。载于《神经信息处理系统进展》，2022年。Mayleen Cortez-Rodriguez、Matthew Eichhorn 和 Christina Lee Yu。《结合推出设计与聚类，开展低阶干扰下的因果推断》。*arXiv* 预印本 *arXiv:2405.05119*，2024年。Micha¨el Defferrard、Xavier Bresson 和 Pierre Vandergheynst。《图上的快速局部谱滤波卷积神经网络》。《神经信息处理系统进展》，29，2016年。Dean Eckles、Brian Karrer 和 Johan Ugander。《网络中的实验设计与分析：减少干扰造成的偏差》。《因果推断期刊》，5(1):20150021，2016年。Laura Forastiere、Edoardo M Airoldi 和 Fabrizia Mealli。《网络观察性研究中处理效应和干扰效应的识别与估计》。《美国统计学会会刊》，116(534):901–918，2021年。Laura Forastiere、Fabrizia Mealli、Albert Wu 和 Edoardo M Airoldi。《利用贝叶斯广义倾向评分估计网络干扰下的因果效应》。《机器学习研究期刊》，23(289):1–61，2022年。Huan Gui、Ya Xu、Anmol Bhasin 和 Jiawei Han。《网络 A/B 测试：从抽样到估计》。载于第*24*届万维网国际会议，*WWW 2015*，第399–409页。计算机协会，2015年。Christopher Harshaw、Fredrik S¨avje、David Eisenstat、Vahab Mirrokni 和 Jean Pouget-Abadie。《线性暴露-响应模型下二部实验的设计与分析》。《电子统计学期刊》，17(1):464–518，2023年。Michael G Hudgens 和 M Elizabeth Halloran。《迈向干扰下的因果推断》。《美国统计学会会刊》，103(482):832–842，2008年。Vardis Kandiros、Charilaos Pipis、Constantinos Daskalakis 和 Christopher Harshaw。《冲突图设计：任意邻域干扰下因果效应的估计》。*arXiv* 预印本 *arXiv:2411.10908*，2025年。Vishesh Karwa 和 Edoardo M Airoldi。《系统考察因网络干扰误设而产生的经典因果推断策略》。*arXiv* 预印本 *arXiv:1810.08259*，2018年。Michael P Leung 和 Pantelis Loupos。《网络混杂下因果推断的图神经网络》。*arXiv* 预印本 *arXiv:2211.07823*，2024年。Yang Liu、Yifan Zhou、Ping Li 和 Feifang Hu。《簇自适应网络 A/B 测试：从随机化到估计》。《机器学习研究期刊》，25(170):1–48，2024年。Jean Pouget-Abadie、Kevin Aydin、Warren Schudy、Kay Brodersen 和 Vahab Mirrokni。《通过相关聚类降低二部实验中的方差》。《神经信息处理系统进展》，32，2019年。

Paul R Rosenbaum. 随机实验中单元间的干扰。《美国统计学会会刊》, 102(477):191–200, 2007. Donald B Rubin. 估计随机研究和非随机研究中的处理因果效应。教育心理学杂志, 66(5):688, 1974. Michael E Sobel. 住房流动的随机研究揭示了什么？面对干扰的因果推断。《美国统计学会会刊》, 101(476):1398–1407, 2006. Johan Ugander and Hao Yin. 随机图聚类随机化。《因果推断期刊》, 11(1):20220014, 2023. Johan Ugander, Brian Karrer, Lars Backstrom, and Jon Kleinberg. 图簇随机化：面向多个宇宙的网络暴露。载于第*19*届 *ACM SIGKDD* 知识发现与数据挖掘国际会议论文集, 第329–337页, 2013. Gonzalo Vazquez-Bare. 随机实验中溢出效应的识别与估计。计量经济学杂志, 237(1):105237, 2023. Davide Viviano, Lihua Lei, Guido Imbens, Brian Karrer, Okke Schrijvers, and Liang Shi. 因果聚类：网络干扰下聚类实验的设计。*arXiv* 预印本 *arXiv:2310.14983*, 2025. Anpeng Wu, Haiyi Qiu, Zhengming Chen, Zijian Li, Ruoxuan Xiong, Fei Wu, and Kun Zhang. 未知干扰下用于处理效应估计的因果图变换器。载于第十三届国际学习表征会议, 2025. Christina Lee Yu, Edoardo M Airoldi, Christian Borgs, and Jennifer T Chayes. 具有未知网络结构的随机实验中总处理效应的估计。美国国家科学院院刊, 119(44):e2208975119, 2022a. Christina Lee Yu, Edoardo M Airoldi, Christian Borgs, and Jennifer T Chayes. 图无关随机实验设计。*arXiv* 预印本 *arXiv:2205.12803*, 2022b.

A 详细模拟结果

本节将提供详细的数值结果。默认设置已在正文中介绍；具体而言，聚类分辨率设为 $\gamma = 5$，潜在结果模型遵循公式（17）。

A.1 聚类分辨率的影响

我们首先考察聚类分辨率 γ 的影响，因为这五种方法都依赖簇级随机化。总体而言，在层次聚类过程中，γ 控制合并所需的模块度增益。γ 的值越高，模块度增益越低，因此往往会保留更多较小的簇。

我们报告了三个不同 γ 水平下的基本统计数据，包括簇的数量、内部节点的比例以及簇内边的比例。

表 1：不同聚类方案的统计数据。

γ	#簇	%内部	%簇内
2	30	8.4	41.6
5	95	7.9	31.7
10	192	7.2	26.6

然后，我们报告 $\gamma \in \{2, 5, 10\}$ 表 2 至表 7中的结果。提醒一下，主论文中报告的默认设定为 $\gamma = 5$，这代表中等程度。这里的主要结论是，聚类分辨率对我们的方法论性能的影响很小，这表明 γ 的选择可以根据其他考量灵活决定。

A.2 协变量设定的影响

在本小节中，我们考察不同协变量设定以及交互项权重的影响。在介绍详细结果之前，我们明确说明，图神经网络（GNN）的输入特征仅由节点度 $\deg_i$ 和处理分配 z_i组成。这意味着，在默认设定下，GNN 显式地将协变量中的一个——度——作为其输入的一部分。因此，我们首先研究以下潜在结果模型，该模型从潜在结果模型中移除了归一化度：

$$\mathbf{Y}(\mathbf{z}) = \beta\mathbf{z} + B\mathbf{z} + \frac{\mathbf{c}}{\bar{c}} \odot \mathbf{z} + \sigma\epsilon. \tag{19}$$

在该模型中，图神经网络无法再直接获取与处理分配存在交互作用的协变量。参数设置与论文正文相同，结果见表 8 和表 9。与表 3 和表6相比，我们发现这种协变量设定的影响也很小。这表明，即使图神经网络没有明确利用仅由网络拓扑结构推导出的协变量，其性能仍基本不受影响。尽管如此，我们注意到，无法从网络结构推断出的人口统计特征应作为回归的输入协变量。

接下来，我们考察交互项权重的变化；在默认配置中，该权重设为 0.5 。从表 3和表6的结果可以看出，即使处理比例增加，CAE、H´ajek 和 MII 估计量仍难以有效处理协变量分布差异。这表明，随着交互项权重增大，AMII估计量的优势会更加明显。因此，我们不再讨论更高权重下的实验，而是转向相反的情形——权重为零，此时交互项不存在。

此情境中使用的潜在结果模型如下，相应结果见表10和11。

$$\mathbf{Y}(\mathbf{z}) = \beta\mathbf{z} + B\mathbf{z} + \sigma\epsilon. \tag{20}$$

当邻域干扰假设成立时（即 $r_2 = 0$），我们观察到 CAE、H´ajek 与 MII估计量均实现了近似无偏性，与理论预期一致。在此设定下，MII估计量取得

显著更低的方差，使其成为该情形下表现最佳的方法。相比之下，GNN估计量并未显著受益于该设定的简单性——它仍需要更大量的高质量数据（或更高的处理比例）才能表现良好。有趣的是，尽管 GNN估计量表现不佳，AMII估计量仍取得了与 MII 相当的结果。这一现象体现了一种无害性性质，进一步验证了我们方法论的稳健性。

在 $r_2 = 1$的情形下，我们观察到所有五种估计量都存在显著偏差。尽管如此，MII 估计量的偏差最低，且方差也最低，进一步巩固了它作为该情形下最有效方法的地位。这一观察也进一步验证了我们为结果组成部分赋予适中且均衡权重的选择。此外，即使搭配的图神经网络架构设计得不够精细，AMII 的表现仍与 MII 相当。

总之，我们发现，在不存在交互项、因而也没有相关差异问题时，MII 的表现最佳。此外，即使使用简单的图神经网络，AMII 也展现出有竞争力的表现。由于现实场景几乎总是涉及复杂的交互效应，而且式20的简洁结构过于理想化，因此只要能够训练反事实预测器，我们就建议在实践中使用 AMII。

A.3 训练样本选择的影响

在本小节中，我们探讨在完整总体上训练反事实预测器，还是仅在边界单元上训练，能带来更好的表现。这个问题自然源于样本拆分的想法，尽管这种方法无法完全消除数据中的复杂相互依赖关系。我们在表12和13中报告结果，并将其与表3 和表6进行比较。我们观察到，仅在边界节点上训练反事实预测器，在 $p = 0.1$时表现较差；但在 $p = 0.3$、0.5时表现略好。其背后的直觉是，随着 p 增加，边界节点会变得更具信息量，因为暴露（例如，接受处理的邻居所占的比例）会逐渐接近全局处理机制。尽管如此，这两种方法之间的差异仍然很小。

A.4 网络拓扑的影响

最后，我们针对另一种网络拓扑开展了一项额外实验[3]，其规模相近，但连接更稀疏——具体而言，其平均度约为早前实验所用网络的一半。尽管潜在结果模型仍是本模拟研究的主要关注点，我们仍纳入这项实验以求完整。详细结果见表 14 和 15，并与表 3 和 6进行比较。

我们发现，另外三个不含回归成分的估计量，其性能基本不受网络拓扑变化的影响。此外，基于GNN的反事实预测估计量在这张更稀疏的图上表现更佳，这种提升也始终转化为AMII估计量性能的改善。

表 2：统计性能（$r_2 = 0, \gamma = 2$）

p		0.1			0.3			0.5	
指标 / **估计量**	Bias	Std	MSE	Bias	Std	MSE	Bias	Std	MSE
H´ajek	0.925	0.405	1.020	0.871	0.289	0.842	0.828	0.197	0.725
CAE	0.942	0.455	1.094	0.940	0.237	0.940	0.923	0.159	0.878
MII	0.947	0.302	0.987	0.944	0.130	0.908	0.940	0.092	0.892
GNN	1.796	1.070	4.369	0.632	0.216	0.446	0.429	0.126	0.199
AMII	0.694	0.465	0.699	0.265	0.169	0.099	0.187	0.100	0.045

[3]网络数据可在以下网址获取：https://networkrepository.com/socfb-USF51.php。

表3：统计性能（$r_2 = 0\,, \gamma = 5$）

p		0.1			0.3			0.5	
指标 **估计量**	Bias	Std	MSE	Bias	Std	MSE	Bias	Std	MSE
哈*耶克	0.965	0.558	1.243	0.908	0.420	1.001	0.874	0.370	0.901
CAE	0.947	0.342	1.014	0.948	0.193	0.937	0.948	0.136	0.917
MII	0.955	0.242	0.971	0.961	0.130	0.940	0.961	0.092	0.933
GNN	1.497	0.997	3.235	0.650	0.173	0.452	0.430	0.114	0.198
AMII	0.597	0.412	0.526	0.249	0.160	0.088	0.165	0.097	0.037

表4：统计性能（$r_2 = 0, \gamma = 10$）

p		0.1			0.3			0.5	
指标 **估计量**	Bias	Std	MSE	Bias	Std	MSE	Bias	Std	MSE
H´ajek	0.954	0.593	1.261	0.929	0.466	1.081	0.890	0.400	0.952
CAE	0.982	0.349	1.086	0.981	0.192	0.999	0.968	0.140	0.956
MII	0.975	0.241	1.010	0.974	0.129	0.965	0.971	0.098	0.952
GNN	1.608	1.009	3.603	0.678	0.169	0.489	0.448	0.110	0.213
AMII	0.642	0.401	0.572	0.242	0.147	0.080	0.153	0.101	0.034

表5：统计性能（$r_2 = 1, \gamma = 2$）

p		0.1			0.3			0.5	
指标 **估计量**	Bias	Std	MSE	Bias	Std	MSE	Bias	Std	MSE
H´ajek	1.396	0.428	2.133	1.202	0.310	1.541	1.061	0.217	1.173
CAE	1.466	0.459	2.360	1.352	0.243	1.888	1.227	0.168	1.533
MII	1.422	0.329	2.131	1.294	0.149	1.697	1.200	0.103	1.451
GNN	2.247	1.170	6.418	1.101	0.228	1.263	0.770	0.157	0.618
AMII	1.150	0.492	1.563	0.713	0.192	0.545	0.486	0.124	0.251

表6：统计性能（$r_2 = 1\,, \quad \gamma = 5$）

p		0.1			0.3			0.5	
指标 **估计量**	Bias	Std	MSE	Bias	Std	MSE	Bias	Std	MSE
H´ajek	1.480	0.575	2.520	1.274	0.438	1.814	1.129	0.377	1.416
CAE	1.555	0.345	2.538	1.428	0.195	2.076	1.293	0.138	1.692
MII	1.471	0.274	2.239	1.341	0.154	1.822	1.239	0.100	1.544
GNN	1.917	0.906	4.494	1.142	0.175	1.336	0.792	0.117	0.641
AMII	1.077	0.388	1.311	0.731	0.168	0.563	0.480	0.100	0.241

表7：统计性能（$r_2 = 1, \quad \gamma = 10$）

p		0.1			0.3			0.5	
指标 **估计量**	Bias	Std	MSE	Bias	Std	MSE	Bias	Std	MSE
H´ajek	1.493	0.605	2.597	1.319	0.476	1.965	1.164	0.408	1.522
CAE	1.628	0.353	2.775	1.486	0.193	2.246	1.334	0.141	1.798
MII	1.514	0.264	2.362	1.379	0.141	1.921	1.271	0.103	1.625
GNN	1.979	0.897	4.722	1.184	0.176	1.433	0.824	0.111	0.691
AMII	1.131	0.395	1.434	0.753	0.150	0.590	0.489	0.100	0.249

表 8：**含一个协变量**时的统计性能（$r_2 = 0, \gamma = 5$）

p		0.1			0.3			0.5	
指标 **估计量**	Bias	Std	MSE	Bias	Std	MSE	Bias	Std	MSE
H´ajek	0.958	0.557	1.228	0.905	0.416	0.992	0.870	0.366	0.890
CAE	0.935	0.342	0.992	0.936	0.193	0.913	0.933	0.136	0.889
MII	0.949	0.242	0.959	0.957	0.130	0.933	0.958	0.092	0.927
GNN	1.428	0.991	3.023	0.639	0.173	0.439	0.425	0.116	0.195
AMII	0.578	0.394	0.490	0.274	0.160	0.101	0.197	0.100	0.049

表9：**含一个协变量**时的统计性能（$r_2 = 1, \gamma = 5$）

p		0.1			0.3			0.5	
指标 **估计量**	Bias	Std	MSE	Bias	Std	MSE	Bias	Std	MSE
H´ajek	1.472	0.572	2.493	1.271	0.431	1.800	1.124	0.371	1.401
CAE	1.544	0.345	2.502	1.415	0.195	2.040	1.279	0.138	1.654
MII	1.465	0.270	2.218	1.337	0.151	1.811	1.236	0.099	1.536
GNN	1.857	0.876	4.215	1.139	0.160	1.322	0.785	0.108	0.627
AMII	1.076	0.377	1.301	0.754	0.161	0.595	0.507	0.097	0.267

表 10：**未使用协变量**时的统计性能（$r_2 = 0, \gamma = 5$）

p		0.1			0.3			0.5	
指标 **估计量**	Bias	Std	MSE	Bias	Std	MSE	Bias	Std	MSE
H´ajek	0.019	0.558	0.312	0.001	0.415	0.172	0.013	0.369	0.136
CAE	0.023	0.342	0.118	0.016	0.193	0.038	0.008	0.136	0.019
MII	0.011	0.242	0.059	0.003	0.130	0.017	0.002	0.092	0.009
GNN	1.708	0.555	3.224	0.561	0.148	0.337	0.389	0.089	0.159
AMII	0.011	0.264	0.070	0.058	0.177	0.035	0.034	0.118	0.015

表11：**无协变量时的统计性能**($r_2 = 1, \gamma = 5$)

p		0.1			0.3			0.5	
指标 **估计量**	Bias	Std	MSE	Bias	Std	MSE	Bias	Std	MSE
H´ajek	0.533	0.572	0.611	0.366	0.427	0.316	0.268	0.371	0.209
CAE	0.586	0.345	0.463	0.463	0.194	0.252	0.338	0.138	0.133
MII	0.504	0.270	0.327	0.377	0.151	0.165	0.275	0.099	0.086
GNN	2.397	0.792	6.373	1.098	0.150	1.227	0.771	0.109	0.607
AMII	0.552	0.284	0.385	0.464	0.176	0.247	0.309	0.120	0.110

表12：**边界回归**($r_2 = 0, \gamma = 5$)下的统计性能

p		0.1			0.3			0.5	
指标 **估计量**	Bias	Std	MSE	Bias	Std	MSE	Bias	Std	MSE
H´ajek	0.965	0.558	1.243	0.908	0.420	1.001	0.874	0.370	0.901
CAE	0.947	0.342	1.014	0.948	0.193	0.937	0.948	0.136	0.917
MII	0.955	0.242	0.971	0.961	0.130	0.940	0.961	0.092	0.933
GNN	2.281	1.030	6.265	0.662	0.167	0.466	0.446	0.118	0.213
AMII	0.740	0.421	0.725	0.144	0.204	0.062	0.092	0.131	0.026

表 13：**边界回归** ($r_2 = 1, \gamma = 5$) 的统计性能

p		0.1			0.3			0.5	
指标 / **估计量**	Bias	Std	MSE	Bias	Std	MSE	Bias	Std	MSE
H´ajek	1.480	0.575	2.520	1.274	0.438	1.814	1.129	0.377	1.416
CAE	1.555	0.345	2.538	1.428	0.195	2.076	1.293	0.138	1.692
MII	1.471	0.274	2.239	1.341	0.154	1.822	1.239	0.100	1.544
GNN	2.463	1.211	7.535	1.167	0.190	1.398	0.812	0.117	0.673
AMII	1.080	0.492	1.408	0.625	0.222	0.440	0.386	0.136	0.167

表14：FB-USF ($r_2 = 0, \gamma = 5$) 上的统计性能

p		0.1			0.3			0.5	
指标 / **估计量**	Bias	Std	MSE	Bias	Std	MSE	Bias	Std	MSE
H´ajek	0.906	0.646	1.238	0.794	0.508	0.889	0.710	0.435	0.694
CAE	0.954	0.311	1.007	0.925	0.165	0.883	0.896	0.109	0.815
MII	0.944	0.241	0.950	0.943	0.131	0.906	0.943	0.101	0.899
GNN	0.821	0.381	0.819	0.559	0.187	0.347	0.346	0.094	0.128
AMII	0.302	0.256	0.157	0.158	0.138	0.044	0.144	0.101	0.031

表 15：FB-USF（$r_2 = 1, \quad \gamma = 5$）上的统计性能

p		0.1			0.3			0.5	
指标 / **估计量**	Bias	Std	MSE	Bias	Std	MSE	Bias	Std	MSE
哈耶克	1.432	0.651	2.475	1.188	0.511	1.672	0.989	0.442	1.173
CAE	1.549	0.314	2.498	1.400	0.171	1.988	1.253	0.117	1.583
MII	1.483	0.250	2.261	1.364	0.138	1.881	1.259	0.106	1.597
GNN	1.457	0.366	2.258	1.001	0.142	1.021	0.714	0.094	0.518
AMII	0.879	0.281	0.851	0.633	0.137	0.419	0.504	0.105	0.265

B 证明

B.1 证明定理3.1

首先，介绍基本记号：$m_k = |\text{Int}_k|$ 表示第 k 个簇的内部集合大小，$n_k = |C_k|$ 表示第 k 个簇的大小。此外，我们用 $Y_{k,\ Int}(\mathbf{1})Y_{k,\ Int}(\mathbf{0})$分别表示全局处理和对照条件下第 k 个簇内部节点的平均结果。相应地，$Y_k(\mathbf{1})$ $Y_k(\mathbf{0})$分别表示全局处理和对照条件下第 k 个簇中节点的平均结果。最后，我们用 t_k 表示第 k 个簇的处理分配指标。

由于我们采用簇级随机化，同一簇内实验单位的处理分配高度相关。在邻域干扰假设下，我们首先将估计量重新排列如下：

$$\begin{aligned}
\hat{\tau}_{MII} &= \frac{\sum_{i\in \text{Int}} z_i Y_i}{\sum_{j\in \text{Int}} z_j} - \frac{\sum_{i\in \text{Int}} (1-z_i) Y_i}{\sum_{j\in \text{Int}} (1-z_j)} \\
&= \frac{\sum_{i\in \text{Int}} z_i Y_i(\mathbf{1})}{\sum_{j\in \text{Int}} z_j} - \frac{\sum_{i\in \text{Int}} (1-z_i) Y_i(\mathbf{0})}{\sum_{j\in \text{Int}} (1-z_j)} \\
&= \frac{\sum_{k\in[K]} \sum_{i\in \text{Int}_k} z_i Y_i(\mathbf{1})}{\sum_{j\in \text{Int}} z_j} - \frac{\sum_{k\in[K]} \sum_{i\in \text{Int}_k} (1-z_i) Y_i(\mathbf{0})}{\sum_{j\in \text{Int}} (1-z_j)} \\
&= \frac{\sum_{k\in[K]} t_k m_k \bar{Y}_{k,Int}(\mathbf{1})}{\sum_{k\in[K]} t_k m_k} - \frac{\sum_{k\in[K]} (1-t_k) m_k \bar{Y}_{k,Int}(\mathbf{0})}{\sum_{k\in[K]} (1-t_k) m_k}.
\end{aligned} \tag{21}$$

第二个等式成立是由于邻域干扰假设（假设 2.1）和簇级随机化。第三个等式将对所有内部单元的求和转换为先对第 k 个簇的内部求和，再对各个簇求和。第四个等式是定义。

通过应用斯卢茨基引理，我们消除了处理分配中的随机性，从而得到：

$$\hat{\tau}_{MII} = \frac{\sum_{k\in[K]} m_k \bar{Y}_k(\mathbf{1})}{\sum_{k\in[K]} m_k} - \frac{\sum_{k\in[K]} m_k \bar{Y}_k(\mathbf{0})}{\sum_{k\in[K]} m_k} + o_p(1). \tag{22}$$

假设3.1 改编自 CAE 原始论文中的假设 4.1.1 和 4.1.2（Liu等人，2024年），而我们的假设更弱，部分原因是我们只追求一致性。给定假设 3.1，我们可以将内部节点的分量替换为整个簇的分量，从而得到：

$$\begin{aligned}
\hat{\tau}_{MII} &= \frac{\sum_{k\in[K]} n_k \bar{Y}_{k,Int}(\mathbf{1})}{\sum_{k\in[K]} n_k} - \frac{\sum_{k\in[K]} n_k \bar{Y}_{k,Int}(\mathbf{0})}{\sum_{k\in[K]} n_k} + o_p(1) \\
&= \frac{\sum_{k\in[K]} n_k \bar{Y}_k(\mathbf{1})}{\sum_{k\in[K]} n_k} - \frac{\sum_{k\in[K]} n_k \bar{Y}_k(\mathbf{0})}{\sum_{k\in[K]} n_k} + o_p(1) \\
&= \tau + o_p(1).
\end{aligned} \tag{23}$$

B.2 定理 4.1 的证明

首先，我们定义 $h(e) = \mathbb{E}_v[h(e,\ v)]$。回顾潜在结果模型为：

$$Y_i(\mathbf{z}) = (\beta + \alpha u_i) z_i + h\Big(\sum_{j\in\mathcal{N}(i,1)} z_j/\deg_i, v_i\Big). \tag{24}$$

因此，真实的 GATE 由下式给出：

$$\tau = \beta + \alpha\mu + h(1) - h(0). \tag{25}$$

接下来，基于该模型，MII 估计量的期望为：

$$\hat{\tau}_{MII} = \beta + \alpha\mu_{Int} + h(1) - h(0). \tag{26}$$

综合以上结果可得：

$$\mathrm{Bias}(\hat{\tau}_{MII}) = \alpha(\mu - \mu_{Int}). \tag{27}$$

接下来，我们考察 AMII 估计量。定义如下：

$$\hat{h}(p) = \mathbb{E}_{\mathbf{z}\sim\mathrm{Ber}(p),v}[\mathrm{MEAN}(\{g(z_j, v_j) \mid j \in \mathcal{N}(i,1)\})]. \tag{28}$$

这表示在随机化的处理比例等于 p的条件下，对干扰部分估计值取期望。请注意，随机性同时存在于处理分配和协变量 v 中。尽管如此，当 $p = 0$ 或 $p = 1$ 时，前者会消失，因为在相应情况下$\mathbf{z} = \mathbf{0}$,$\mathbf{1}$ 成立。

回归函数的函数形式有一个重要特征：

$$\hat{h}(1) = \mathbb{E}_v \left[\frac{1}{|\,\mathrm{Int}\,|} \sum_{i\in\mathrm{Int}} \mathrm{MEAN}(\{g(1, v_j) \mid j \in \mathcal{N}(i,1)\}) \middle| \mathcal{G} \right]. \tag{29}$$

这是因为在度 deg_i 固定时，可以交换均值和期望 $\mathbb{E}_v$。另一方面，很容易验证：

$$\hat{h}(1) = \mathbb{E}_v \left[\frac{1}{n} \sum_{i\in[n]} \mathrm{MEAN}(\{g(1, v_j) \mid j \in \mathcal{N}(i,1)\}) \middle| \mathcal{G} \right]. \tag{30}$$

因此，我们有：

$$\begin{aligned} \mathbb{E}[\hat{\tau}_{AMII}] &= \beta - \mathbb{E}[\hat{\beta}_n] + (\alpha - \mathbb{E}[\hat{\alpha}_n])\mu_{Int} + h(1) - \hat{h}(1) + \mathbb{E}[\hat{\beta}_n] + \mathbb{E}[\hat{\alpha}_n]\mu + \hat{h}(1) \\ &\quad - \Big(h(0) - \hat{h}(0) + \hat{h}(0)\Big) \\ &= \beta + (\alpha - \mathbb{E}[\hat{\alpha}_n])\mu_{Int} + \mathbb{E}[\hat{\alpha}_n]\mu + h(1) - h(0). \end{aligned} \tag{31}$$

这给出了 AMII估计量的偏差：

$$\mathrm{Bias}(\hat{\tau}_{AMII}) = (\mathbb{E}[\hat{\alpha}] - \alpha)(\mu - \mu_{Int}). \tag{32}$$