arXiv:2305.05017v1 [stat.ME] 8 May 2023

一种贝叶斯非参数方法

C用于存在处理后混杂因素的因果中介分析

混杂因素

裴宇静
佛罗里达大学统计系，佛罗里达州盖恩斯维尔，邮编 32611，
w
oojung.bae
@ufl.edu

Michael J. Daniels
佛罗里达大学统计学系，佛罗里达州盖恩斯维尔，邮编 32611
daniels@ufl.edu

Michael G. Perri
佛罗里达大学临床与健康心理学系，佛罗里达州盖恩斯维尔，邮编 32611
mperri@ufl.edu

2023年5月10日

摘要

我们提出一种新的贝叶斯非参数（BNP）方法，用于估计存在处理后混杂因素时的中介因果效应。我们指定一个扩展狄利克雷过程混合模型（EDPM），以对观测数据（结局、中介变量、处理后混杂因素、处理和基线混杂因素）的联合分布进行建模。所提出的贝叶斯非参数模型允许基于混杂变量的簇数量多于结局和中介变量的簇数量。为确保可识别性，我们采用 Hong 等人 [2022] 提出的标准序贯可忽略性的扩展版本。观测数据模型和因果识别假设使我们能够估计和识别中介的因果效应、$i.e.$、自然直接效应（NDE）和间接效应（NIE）。我们开展模拟研究，以评估所提方法的表现。此外，我们将该方法应用于评估 Rural LITE 试验中的因果中介效应，展示其在现实场景中的实用价值。

关键词：因果推断；增强型狄利克雷过程混合模型。

1 引言

社会科学、行为科学和健康科学领域的研究人员通常致力于进行

关于因果效应的推断 [Wu 等人，2011年；Woo 等人，2015年；Zhao 等人，2018年；Sohn 和 Li，2019]。其中一些研究关注对由中间变量中介的处理对结局的因果影响进行推断 [Kraemer 等人， 2002]。几十年来，处理通过中间变量影响结局的过程一直是一个“黑箱” [Imai 等人，2011], ，研究人员已提出各种方法对其进行考察。关于中介分析，频率学派文献十分丰富 [Pearl，2013年；MacKinnon 等人，2002年；Robins，2003年；Preacher 和 Hayes，2004年；Petersen 等人，2006年；VanderWeele，2009年；Imai 等人，2010年；Albert 和 Nelson，2011年；Tchetgen Tchetgen 和 Shpitser，2012年；Valeri 和 VanderWeele， 2013]。贝叶斯方法包括 Yuan 和 MacKinnon [2009]；Elliott 等人， [2010]；Schwartz 等人， [2011]；Daniels 等人， [2012]；Mattei 等人， [2013]；Kim 等人， [2017, 2019]。在 Baron 和 Kenny 提出的传统结构方程模型框架下 [1986], Yuan 和 MacKinnon [2009] 采用贝叶斯方法计算系数乘积的后验分布，从而能够对所需的因果效应进行推断。然而，Elliott 等人 [2010]；Schwartz 等人，[2011]；Mattei 等人， [2013] 中的方法主要旨在估计主要因果效应，这不同于我们所关注的因果估计量，即 NDE 和 NIE。此外，贝叶斯非参数（BNP）模型尚未被广泛用于估计因果中介效应。Kim 等人 [2017] 建立了一个用于评估因果中介效应的 BNP 框架。具体而言，他们将多元正态分布的狄利克雷过程混合（DPMN）作为观测数据（结局、中介变量和混杂变量）联合分布的先验分布，以估计因果中介效应。他们考虑了两组不可检验假设，分别由 Imai 等人[2010] 和 Daniels 等人 [2012], 提出。Kim 等人 [2017] 中的 DPMN 方法并不

适用于包含多于少数几个混杂变量或非连续结局的情形，-

中介变量或混杂变量。此外，存在治疗后混杂因素时，上述假设并不成立。存在治疗后混杂因素时，通常认为中介变量与结局之间的标准跨世界反事实独立性不切实际 [Hong 等人，2022；Rudolph 等人， 2023]。因此，最常见的替代方法会对治疗后混杂因素提出额外假设，将其视为按顺序排列的多个中介变量中的第一个。这些假设通常伴随着广义线性结构方程模型框架中的一组基于模型的假设 [Albert 和 Nelson，2011；Albert 等人，2019；Daniel 等人，2015；Imai 和 Yamamoto，2013]。此外，也有研究提出使用基于加权的方法调整治疗后混杂 [Huber，2014；Hong，2015；Hong 等人， 2018], 这些方法同样对治疗后混杂因素提出了较强的假设。另一种做法是，许多研究者提出不同的因果估计量，不将平均处理效应（ATE）分解为自然间接效应（NIE）和自然直接效应（NDE） [Geneletti，2007；Didelez 等人，2012；VanderWeele 等人，2014；Lin 和 VanderWeele，2017；Miles 等人，2017；Vansteelandt 和 Daniel，2017；Rudolph 等人，2018a、b；Miles 等人，2020；Wodtke 和 Zhou，2020；Hejazi 等人，2020；Díaz 等人，2021；Nguyen 等人，2021；Miles，2022；Tai 和 Lin，2022]。其中一些研究应用干预方法，在不假设不存在治疗后混杂因素的情况下，推导出用于估计直接效应和间接效应的新因果估计量。具体而言，Hejazi 等人 [2020]；Díaz 等人 [2021] 近期的研究在非参数结构方程模型（NPSEM）框架下，给出了干预效应的理论性质及不依赖于跨世界反事实独立性的假设。

最近，Hong 等人 [2022] 和 Rudolph 等人 [2023] 提出了在存在治疗后混杂因素时估计 NIE 和 NDE 的假设。Rudolph 等人[2023] 提出的方法旨在估计 NDE 和 NIE，同时考虑中间混杂以及单调性约束，适用于治疗后混杂因素为二元变量的情况。然而，当治疗后混杂因素为连续变量时，他们的假设并不充分。Hong 等人[2022] 提出了一个放宽版的跨世界反事实独立性假设，并结合中介变量概率加权（RMPW），从而允许存在处理与中介变量的交互作用 [Hong，2010，2015；Hong 和 Nomi，2012；Hong 等人，2015；Lange 等人，2012；Tchetgen Tchetgen 和 Shpitser， 2012]。利用高斯 Copula 模型 [Nelsen， 2007], Hong 等人 [2022] 考察了基于实际治疗条件和反事实治疗条件的治疗后混杂因素之间的条件相关性，并提出了一种用于处理治疗后混杂的敏感性分析方法。

本文提出一种贝叶斯非参数方法，用于在存在治疗后混杂因素时估计中介的因果效应。我们指定一个扩展的狄利克雷过程 [EDPM；Wade 等，2011年、2014年；Roy 等人， 2018] ，以对观测数据的联合分布建模，其中包括结局、中介变量、处理、治疗后混杂因素和基线混杂因素。所提出的贝叶斯非参数模型允许混杂因素的簇数多于结局和中介变量的簇数。为保证可识别性，我们考虑针对治疗后混杂因素的序贯可忽略性扩展版本 [Hong 等人， 2022]。观测数据模型与因果假设相结合，使得中介的因果效应 $i.e.$、自然直接效应和自然间接效应得以估计和识别。与以往研究采用的方法类似，我们使用高斯Copula模型对治疗后混杂因素建模 [Daniels 等人，2012年；Kim 等人，2017年；Hong 等人， 2022]。这是因为治疗后混杂因素仅部分观

测到，这意味着若某个研究对象被分配至治疗组，其在反事实对照条件下的潜在治疗后混杂因素值仍不可观测。高斯Copula模型有助于处理这类部分“缺失”数据。Hong 等人 [2022] 采用基于中介变量概率加权（RMPW）比值的参数方法，而我们的方法通过基于EDPM从后验分布中抽样，能够更灵活地估计中介因果效应，并避免使用RMPW。此外，与Hong 等人 [2022] 的方法不同，我们的方法可以轻松计算给定部分混杂变量条件下的自然直接效应和自然间接效应；由于我们对联合分布进行指定，因此在可忽略缺失假设下，可以处理观测数据中的任何缺失。

我们将所提出的方法应用于 Rural LITE 试验的数据 [Perri 等人， 2014]。Rural LITE 试验旨在考察通过农村社区的合作推广办公室实施的不同强度行为减重治疗及维持治疗的效果和成本。受试者被随机分配接受低、中或高剂量的行为治疗，或接受对照条件。我们关注出勤率这一潜在中介变量，它可能影响研究结束时的体重变化。在第 0 个月（基线）、第 6 个月和第 24 个月测量体重，主要结局为基线至最终随访（第 24 个月）的体重变化。基线至第 6 个月的体重变化是一个潜在的治疗后混杂因素。我们收集了以下基线混杂因素：性别、年龄、种族以及基线时的体质指数。本文其余部分安排如下。第 2 节介绍我们关注的因果效应，并提出一个用于观测数据的贝叶斯非参数模型（BNP）。第 3 节提出一组足以实现识别的假设。第 4 节讨论所提出的 BNP 模型的后验计算。第 5 节介绍为评估 BNP 模型性能而开展的模拟研究。所提出的

方法将在第 6 节应用于 Rural LITE 数据，以估计因果中介效应。最后，第 7 节进行讨论。

2 因果效应的定义与观测数据模型的设定

模型

令 Z_i 表示 i^{th} 个体（$i = 1, \cdots, n$）的二元处理指标。用 $\boldsymbol{C}_i$ 表示一组基线混杂因素，并令 V_i 为一个处理后混杂因素。令 M_i 表示中介变量的取值， Y_i 表示结局的取值。为定义因果中介效应，我们采用 Rubin [1974] 的潜在结局框架。在稳定单元处理值假设（SUTVA）下， [Rubin， 1980], 处理后混杂因素的潜在取值不会随分配给其他个体的处理而变化。也就是说，对于随机分配处理的向量 $\boldsymbol{Z}^{(n)} \in \{0,1\}^{\otimes n}$, $V_{i,\boldsymbol{z}^{(n)}} = V_{i,Z_i}$ ，其中 $V_{i,\boldsymbol{z}^{(n)}}$ 表示 i^{th} 个体在处理向量 $\boldsymbol{Z}^{(n)}$ 设为 $\boldsymbol{z}^{(n)}$时本应观测到的处理后混杂因素 V_i 的取值。类似地，随机变量$M_{i,\boldsymbol{z}^{(n)},\boldsymbol{v}^{(n)}}$ 等于M_{i,Z_i,V_i}，其中$M_{i,\boldsymbol{z}^{(n)},\boldsymbol{v}^{(n)}}$ 表示 i^{th} 个体在处理向量 $\boldsymbol{Z}^{(n)}$ 设为 $\boldsymbol{z}^{(n)}$ 且处理后混杂因素$\boldsymbol{V}^{(n)}$ 的取值设为 $\boldsymbol{v}^{(n)}$时本应观测到的中介变量取值。最后，随机变量$Y_{i,\boldsymbol{z}^{(n)},\boldsymbol{v}^{(n)},\boldsymbol{m}^{(n)}}$ 等于Y_{i,Z_i,V_i,M_i}，其中$Y_{i,\boldsymbol{z}^{(n)},\boldsymbol{v}^{(n)},\boldsymbol{m}^{(n)}}$ 表示 i^{th} 个体在处理向量 $\boldsymbol{Z}^{(n)}$ 设为 $\boldsymbol{z}^{(n)}$、处理后混杂因素向量 $\boldsymbol{V}^{(n)}$ 设为 $\boldsymbol{v}^{(n)}$，且中介变量 $\boldsymbol{M}^{(n)}$ 的取值设为 $\boldsymbol{m}^{(n)}$时本应观测到的结局取值。为简化记号并专注于评估自然效应，本文沿用 [Robins 和 Greenland（1992；Pearl， 2013]）的做法，省略个体索引 i 。此外，对于 $z, z' \in \{0, 1\}$，我们不再使用M_{z,V_z} 和$Y_{z,V_z,M_{z'},V_{z'}}$，而改写为 M_z 和$Y_{z,M_{z'}}$。观测到的处理后混杂因素、中介变量和结局定义如下： $V = ZV_1 + (1-Z)\, V_0$, $M = ZM_1 + (1-Z)\, M_0$，以及 $Y = ZY_{1,M_1} + (1-Z)\, Y_{0,M_0}$。我们定义以基线混杂因素为条件的 NDE 和 NIE

$\boldsymbol{C}=\boldsymbol{c}$分别为 NIE $(\boldsymbol{c})=\mathrm{E}\left[Y_{1,M_1}-Y_{1,M_0}|\boldsymbol{c}\right]$和 NDE $(\boldsymbol{c})=\mathrm{E}\left[Y_{1,M_0}-Y_{0,M_0}|\boldsymbol{c}\right]$。自然间接效应 NIE $(\boldsymbol{c})$量化了在混杂因素 $\boldsymbol{C}=\boldsymbol{c}$取固定值时，处理经由中介变量产生的效应。自然直接效应 NDE $(\boldsymbol{c})$量化了在基线混杂因素 $\boldsymbol{C}$取固定值的条件下，将中介变量 M设为其自然值 M_0（即未接受处理时中介变量的取值）时，处理对结局的效应。总效应是这两种效应之和：ATE $(\boldsymbol{c})=$NIE $(\boldsymbol{c})+$NDE $(\boldsymbol{c})=\mathrm{E}\left[Y_{1,M_1}-Y_{0,M_0}|\boldsymbol{c}\right]$。对基线混杂因素积分后，可得到边际因果效应 NIE、NDE 和 ATE。我们的目标是根据观测数据 $(Y,M,V,Z,\boldsymbol{C})$识别因果效应。下文首先在第 2.1 节详细介绍观测数据的贝叶斯非参数模型。

2.1 用于观测数据的贝叶斯非参数模型

为了估计第 3 节中介绍的因果效应，我们首先需要估计联合分布 P $(Y,M,V,Z,\boldsymbol{C})$。令 $\boldsymbol{X}=\left(Z,\boldsymbol{C}^{\top}\right)^{\top}$，并考虑 P $(Y,M,V,\boldsymbol{X})$。我们提出使用 EDPM [Wade 等，2011 年，2014；Roy 等人，2018] 来建模 $(Y,M,V,\boldsymbol{X})$ 的联合分布，具体如下：

$$
\begin{aligned}
Y_i|M_i,V_i,\boldsymbol{X}_i;\theta_i^y &\sim f\left(y_i|m_i,v_i,\boldsymbol{x}_i;\theta_i^y\right)\\
M_i|V_i,\boldsymbol{X}_i;\theta_i^m &\sim f\left(m_i|v_i,\boldsymbol{x}_i;\theta_i^m\right)\\
V_i|\boldsymbol{X}_i;\omega_i^v &\sim f\left(v_i|\boldsymbol{x}_i;\omega_i^v\right)\\
\boldsymbol{X}_{i,q};\omega_{i,q}^x &\sim f\left(x_{i,q};\omega_{i,q}^x\right),\ \ q=1,\cdots,p_1+p_2+1\\
(\theta_i,\omega_i)\,|\,\mathrm{G} &\sim \mathrm{G}\\
\mathrm{G} &\sim \mathrm{EDP}\left(\alpha^{\theta},\alpha^{\omega},\mathrm{G}_0\right)
\end{aligned}
\tag{2.1}
$$

其中 $\theta_i=(\theta_i^y,\theta_i^m)$ 和 $\omega_i=(\omega_i^v,\omega_i^x)$。记号 G $\sim$EDP $\left(\alpha^{\theta},\alpha^{\omega},\ \mathrm{G}_0\right)$ 表示 $\mathrm{G}^{\theta}\sim$DP $\left(\alpha^{\theta},\mathrm{G}_0^{\theta}\right)$，且 $\mathrm{G}^{\omega|\theta}\sim$DP $\left(\alpha^{\theta},\ \mathrm{G}_0^{\omega|\theta}\right)$，其基测度为 $\mathrm{G}_0=\mathrm{G}_0^{\theta}\times\mathrm{G}_0^{\omega|\theta}$。

在 EDPM 设定中，每个受试者 i 都有自己的参数 (θ_i,ω_i)，但由于 G [Ferguson 的离散性，同一簇中的受试者共享相同的参数值，(θ_i,ω_i)。

1973]。假设基线协变量 $\boldsymbol{X}$ 在簇内相互独立，而治疗后混杂因素 V 允许存在局部依赖。EDPM 有两个浓度参数 α^{θ} 和 α^{ω}，而狄利克雷过程混合模型（DPM）只有一个浓度参数。$y\&m$-簇的数量取决于浓度参数 α^{θ}，而给定 $y\&m$-簇时，嵌套簇 $v\&\boldsymbol{x}$-簇的数量取决于浓度参数 α^{ω}。较低的 α^{θ} 和 α^{ω} 表示簇的数量较少，这与 DPM 中的浓度参数类似。这使得 $v\&\boldsymbol{x}$-簇的数量可以多于 $y\&m$-簇，这一点至关重要，因为 $v\&\boldsymbol{x}$ 的维度通常远大于 $y\&m$的维度。为简便起见，以下各节将分别使用 y-簇和 $\boldsymbol{x}$-簇，而不使用 $y\&m$-簇和 $v\&\boldsymbol{x}$-簇。

在每个 y-簇中，我们假设两个广义线性模型，一个用于$f\left(y|m,v,\boldsymbol{x};\ \theta^{y}\right)$另一个用于 $f\left(m|v,\boldsymbol{x};\ \theta^{y}\right)$。如果 Y是连续变量，我们设定 $Y|\mathbb{M};\ \beta^{y},\sigma^{y,2}\ \sim\mathrm{N}(\mathbb{M}\beta^{y},\sigma^{y,2})$，其中 $\theta^{y}=(\beta^{y},\sigma^{y,2})$ 且 $\mathbb{M}=\left(1,M,V,Z,\boldsymbol{C}^{\top}\right)^{\top}$（一个包含 M, V 和 $\boldsymbol{X}$ 的设计矩阵）。如果 Y是二元变量，则可以设定 $Y|\mathbb{M};\ \theta^{y}\sim$伯努利分布$\left(\mathrm{probit}^{-1}(\mathbb{M}\theta^{y})\right)$。类似地，我们假设在每个 y-簇（$y\&m$-簇）中，$f\left(m|v,\boldsymbol{x};\ \theta^{m}\right)$服从一个广义线性模型。对于治疗后混杂因素 V，给定基线混杂因素 $\boldsymbol{X}$和处理 Z，我们设定了一个局部广义线性模型。与 [Roy 等人 2018], 的做法类似，我们假设基线混杂因素 $\boldsymbol{X}$ 局部独立。值得注意的是，尽管我们假设每个簇内的 y, m, v和 $\boldsymbol{x}$分别服从参数模型，但所有变量在整体上相互依赖，且可能存在非线性关系。

EDPM模型 (2.1) 具有方块分解表示 [Wade 等人， 2011],

$$\mathrm{P}\left(y,m,v,\boldsymbol{x}|\ \mathrm{G}\right)=\sum_{k=1}^{\infty}\gamma_{k}f\left(y|m,v,\boldsymbol{x};\theta_{k}^{y}\right)f\left(m|\boldsymbol{x};\theta_{k}^{m}\right)\sum_{r=1}^{\infty}\gamma_{r|k}f\left(v|\boldsymbol{x};\omega_{r|k}^{v}\right)f\left(\boldsymbol{x};\omega_{r|k}^{x}\right),$$

其中， k 索引 y-簇，而 $f(\cdot)$ 是相应的分布。权重的先验分布为 $\gamma'_k \sim$ 贝塔分布$\left(1, \alpha^{\theta}\right)$ 和 $\gamma'_{r|k} \sim$ 贝塔分布$(1, \alpha^{\omega})$，其中 $\gamma_k = \gamma'_k \prod_{l<k}(1-\gamma'_l)$ 且 $\gamma_{r|k} = \gamma'_{r|k} \prod_{d<r}\left(1-\gamma'_{d|k}\right)$。

根据联合分布，我们可以推导出以下条件分布形式（我们将在第4.1节的 G 计算算法中使用）：

$$\mathrm{P}(y|m,v,\boldsymbol{x}) = \sum_{k=1}^{\infty} \Lambda_k^y(m,v,\boldsymbol{x}) f\left(y|m,v,\boldsymbol{x};\theta_k^y\right),$$

其中 $\Lambda_k^y(m,v,\boldsymbol{x}) = \dfrac{\gamma_k f\left(m|\boldsymbol{x}\right.}{;\quad \theta_k^m)\sum_{r=1}^{\infty}\gamma_{r|k} f\left(v|\boldsymbol{x}\ \omega_k^V\right) f\left(\boldsymbol{x}\ \omega_{r|k}^x\right);;;;;}{\sum_{d=1}^{\infty}\gamma_d f\left(m|\boldsymbol{x}\ \theta d^m\right)\sum_{r=1}^{\infty}\gamma_{r|d} f\left(v|\boldsymbol{x}\ \omega_k^V\right) f\left(\boldsymbol{x}\right.}$ $\omega_{r|d}\big) = \dfrac{\lambda_k^y(m}{,\quad v\ \boldsymbol{x}),,,\boldsymbol{x}),}{\sum_{d=1}^{\infty}\lambda_d^y(m\ v}$，权重 $\Lambda_k^y(m,v,\boldsymbol{x})$ 取决于 $m, v, \boldsymbol{x}$。注意，$\mathrm{P}(m|v,\boldsymbol{x})$ 和 $\mathrm{P}(v|\boldsymbol{x})$ 具有相似的形式；详情请参阅补充材料。因此，尽管局部上假设 $f(y|m,v,\boldsymbol{x};\ \theta^y)$, $f(m|v,\boldsymbol{x};\ \theta^m)$ 和 $f(v|\boldsymbol{x};\ \omega^v)$ 分别服从简单的广义线性模型，$\mathrm{P}(y|m,v,\boldsymbol{x})$、$\mathrm{P}(m|v,\boldsymbol{x})$ 和 $\mathrm{P}(v|\boldsymbol{x})$ 仍是计算上易于处理、灵活的非线性、非加性模型。

3 因果效应的识别假设与推断

本节定义我们关注的因果效应及识别假设。需要注意的是，估计观测数据的联合分布，与定义因果效应和指定识别假设是彼此独立的步骤。我们采用 Hong 等人 [2022], 提出的一组假设。这些假设是对 [Imai 等人 2010], 提出的标准序贯可忽略性的扩展，用于根据存在治疗后混杂因素时的观测数据分布，评估我们关注的因果效应。

下方的因果图 1 展示了存在治疗后混杂因素 V 且治疗为随机分配 Z 时的因果中介结构。

Hong 等人 [2022] 对 Imai 等人的序贯可忽略性假设 [2010] 作如下修改：对于 z、$z'=0$、1 和 $z \neq z'$:

(1) 给定观测到的基线混杂因素时，处理分配可忽略：

$$\{Y_{z,m}, M_z, M_{z'}, V_z, V_{z'}\} \perp\!\!\!\perp Z | \boldsymbol{C} = \boldsymbol{c}. \tag{1}$$

鉴于 V_z 和 $V_{z'}$ 均是在固定治疗条件下指定的，在给定潜在治疗后混杂因素值 $V_z = v$ 和$V_{z'} = v'$的情况下，治疗分配也是可忽略的，

$$\{Y_{z,m}, M_z, M_{z'}\} \perp\!\!\!\perp Z | V_z, V_{z'}, \boldsymbol{C} = \boldsymbol{c}. \tag{1'}$$

当治疗经过随机分配时，假设 (1) 和 (1') 必然成立。在准实验研究中，假设 (1) 可能过于强，因为在 $\boldsymbol{C}$的各个水平内，治疗选择仍可能与未观测到的基线混杂因素相关。

(2) 在每种治疗条件下，给定观测到的基线混杂因素和治疗后混杂因素，中介变量值分配可忽略：

$$Y_{z,m} \perp\!\!\!\perp \{M_z, M_{z'}\} | V_z = v, Z = z, \boldsymbol{C} = \boldsymbol{c}. \tag{2}$$

假设 (2) 排除了中介变量与结局之间的跨世界联系。与标准可忽略性假设 (1') 相比，该假设更为合理；后者仅以基线混杂因素为条件，因为治疗后混杂往往无法避免。根据假设 (2)，具有相同基线混杂因素值$\boldsymbol{C} = \boldsymbol{c}$ 和相同治疗后 $V_z = v$ 的受试者，其潜在结局$Y_{z,\ m}$ 在两种治疗条件 M_z 和 $M_{z'}$下分别相互独立。然而，显然，如果中介变量与结局的关系受到未测量的基线或治疗后混杂因素的混杂，假设 (2) 可能不成立。

(3) 治疗后混杂因素与中介变量之间的条件跨世界独立性：

$$M_{z'} \perp\!\!\!\perp V_z | V_{z'} = v', Z = z', \boldsymbol{C} = \boldsymbol{c}. \tag{3}$$

假设 (3) 表明，对于给定的治疗条件 z'，在已知 $V_{z'}$ 和 $\boldsymbol{C}$ 的情况下，V_z 不会包含任何关于 $M_{z'}$ 的额外信息。如果存在未测量的基线混杂因素或治疗后混杂因素，它会影响 $M_{z'}$ 和 V_z，但不会影响 $V_{z'}$，那么这一假设也可能不成立。

(4) 高斯Copula：假设给定基线混杂因素后，两个潜在治疗后混杂因素的联合分布遵循高斯Copula模型 [Nelsen,2007]:

$$F_{V_{z'},V_z}\left(v_{z'},v_z|\boldsymbol{C}=\boldsymbol{c}\right)=\Phi_2\left(\Phi_1^{-1}\left(F_{V_{z'}}\left(v_{z'}|\boldsymbol{C}=\boldsymbol{c}\right)\right),\Phi_1^{-1}\left(F_{V_z}\left(v_z|\boldsymbol{C}=\boldsymbol{c}\right)\right)\right), \quad (4)$$

其中，Φ_1 是一元标准正态累积分布函数（CDF），Φ_2是均值为 0、方差为 1、相关系数为 $\rho\in(-1,1)$ 的二元正态累积分布函数。重要的是，这种对潜在治疗后混杂因素联合分布的设定，并未对模型 $V_{z'}|\boldsymbol{C}$ 和 $V_z|\boldsymbol{C}$ 施加任何约束；我们使用 EDPM 对这两个模型进行估计。由于我们从未同时观测到 $V_{z'}$ 和 V_z，因此无法根据数据估计 ρ。因此，可以将 ρ 视为已知，并将其纳入敏感性分析 [Daniels 等人, 2012], ；或者指定一个先验分布，例如 Unif$(0,1)$，以假设潜在治疗后混杂因素之间仅存在正相关。

Hong 等人 [2022] 在假设（1）和（3）下证明了以下结论。

定理3.1

在假设（1）和（3）下，

$$\begin{aligned}&\mathrm{P}\left(M_{z'}=m'|V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c}\right)\\&=\int\mathrm{P}\left(M_{z'}=m'|V_{z'}=v',Z=z',\boldsymbol{C}=\boldsymbol{c}\right)\mathrm{P}\left(V_{z'}=v'|V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c}\right)dv'.\end{aligned}$$

该定理意味着，我们可以通过对条件分布进行边缘化，识别反事实条件 $Z=z$ 下某个中介变量取值$M_{z'}=m'$的条件概率。

分布、$V_{z'}|V_z=v$、$Z=z$、$\boldsymbol{C}=\boldsymbol{c}$。请注意，条件分布 $V_{z'}|V_z=v\,Z=z$、$\boldsymbol{C}=\boldsymbol{c}$ 由（4）部分识别。

以下定理给出了自然间接效应（NIE）和自然直接效应（NDE）的识别。

定理 3.2

在假设 (1)、(2)、(3) 和 (4) 下，可以识别自然间接效应（NIE）和自然直接效应（NDE）：对于 $z\neq z'$ 和 $v\neq v'$,

$$E\big[Y_{z,M_{z'}}|V_z=v,\boldsymbol{C}=\boldsymbol{c}\big]=\iint E[Y_{z,m'}|M_z=m',V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c}]$$
$$\mathrm{dF}_{M_{z'}|V_{z'}=v',Z=z',\boldsymbol{C}=\boldsymbol{c}}(m')\,\mathrm{dF}_{V_{z'}|V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c}}(v').$$

证明。

$$\begin{aligned}
&\mathrm{E}\big[Y_{z,M_{z'}}|V_z=v,\boldsymbol{C}=\boldsymbol{c}\big]\\
&=\int \mathrm{E}[Y_{z,m'}|M_{z'}=m',V_z=v,\boldsymbol{C}=\boldsymbol{c}]\,\mathrm{dF}_{M_{z'}|V_z=v,\boldsymbol{C}=\boldsymbol{c}}(m')\\
&\overset{(1)}{=}\int \mathrm{E}[Y_{z,m'}|M_{z'}=m',V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c}]\,\mathrm{dF}_{M_{z'}|V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c}}(m')\\
&\overset{(2)}{=}\int \mathrm{E}[Y_{z,m'}|M_z=m',V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c}]\,\mathrm{dF}_{M_{z'}|V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c}}(m')\\
&\overset{3.1}{=}\int \mathrm{E}[Y_{z,m'}|M_z=m',V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c}]\int \mathrm{dF}_{M_{z'}|V_{z'}=v',Z=z',\boldsymbol{C}=\boldsymbol{c}}(m')\,\mathrm{dF}_{V_{z'}|V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c}}(v')\\
&=\iint \mathrm{E}[Y_{z,m'}|M_z=m',V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c}]\,\mathrm{dF}_{M_{z'}|V_{z'}=v',Z=z',\boldsymbol{C}=\boldsymbol{c}}(m')\,\mathrm{dF}_{V_{z'}|V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c}}(v')
\end{aligned}$$

□

定理 3.2 的证明中，Hong 等人 [2022] 得到了以下结果：

$$\mathrm{E}\big[Y_{z,M_{z'}}|V_z=v,\boldsymbol{C}=\boldsymbol{c}\big]=\iiint \psi\left(m',v,v',\boldsymbol{c}\right)\mathrm{E}[Y_{z,m'}|M_z=m',V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c}]$$
$$\mathrm{dF}_{M_z|V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c}}(m')\,\mathrm{dF}_{V_{z'}|V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c}}(v').$$

其中 $\psi\left(m',v,v',\boldsymbol{c}\right)=\frac{\mathrm{P}(\cancel{M_{z'}=m'|V_{z'}=v',Z=z',\boldsymbol{C}=\boldsymbol{c}})\,\mathrm{P}(M_z=m'|V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c})}{\mathrm{P}(\cancel{M_z=m'|V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c}})\,\mathrm{P}(M_z=m'|V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c})}\overset{(=1)}{}$是通过将乘到定理 3.2 结果的被积函数上得到的。Hong 等人采用半参数方法并进行重新加权。由于在定理 3.2 中，我们对所有分布都采用了默认的灵活设定，因此这一步重新加权并无必要。此外，采用我们的方法，可以轻松计算 NDE

以及在给定（部分）混杂变量条件下的 NIE，这与 Hong 等人不同。 [2022] 而对于 $(Y, M, V, \boldsymbol{X})$ 的任何缺失，其中 $\boldsymbol{X} = \left(Z, \boldsymbol{C}^\top\right)^\top$，均可在可忽略缺失假设下通过数据增广处理。

4 后验计算

我们使用 Wade 等人提出的算法 [2014], , 该算法是对 [Neal 算法 82000] 的扩展，用于生成观测数据模型参数的后验样本；详情请参见补充材料。通过对观测数据模型参数的每个后验样本执行若干后处理步骤，计算 NIE、NDE 和平均处理效应的后验分布，以及相应的点估计和可信区间。

4.1 因果效应估计的后处理步骤

潜在结果分布的任意泛函，例如 $\mathrm{E}\left[Y_{z,\ M_{z'}}\right]$、$\mathrm{E}\left[Y_{z,\ M_{z'}}|V\right]$ 或 $\mathrm{E}\left[Y_{z,\ M_{z'}}|\boldsymbol{C}\right]$，均可利用后验分布中的参数以及第 3 节的识别假设进行计算。以下说明如何计算潜在结果的期望 $\mathrm{E}\left[Y_{z,\ M_{z'}}\right]$,

$$\mathrm{E}\big[Y_{z,M_{z'}}|V_z = v, \boldsymbol{C} = \boldsymbol{c}\big] = \iint \mathrm{E}[Y_{z,m'}|M_z = m', V_z = v, Z = z, \boldsymbol{C} = \boldsymbol{c}]$$
$$\mathrm{dF}_{M_{z'}|V_{z'}=v',Z=z',\boldsymbol{C}=\boldsymbol{c}}\left(m'\right)\mathrm{dF}_{V_{z'}|V_z=v,Z=z,\boldsymbol{C}=\boldsymbol{c}}\left(v'\right).$$

令 θ 表示结局回归和中介变量回归的参数，ω 表示处理和混杂变量（包括处理后混杂变量和基线混杂因素）的参数。此外，令 K 表示非空 y-簇（y 和 m-簇）的数量，$K_{r|k}$ 表示 k^{th} y-簇内非空 x-簇（z 和 x-簇）的数量。给定 $\{\theta^*, \omega^*, s\}$ 的一个后验样本时，条件期望的形式为

$$\begin{aligned}&\mathrm{E}[Y|M = m', V = v, Z = z, \boldsymbol{C} = \boldsymbol{c}; \theta^*, \omega^*, s] \\ &= \frac{\lambda^y_{K+1}\left(m', v, z, \boldsymbol{c}\right)\mathrm{E}_0\left[y|m', v, z, \boldsymbol{c}\right] + \sum_{l=1}^{K}\lambda^y_l\left(m', v, z, \boldsymbol{c}\right)\mathrm{E}\left[y|m', v, z, \boldsymbol{c}; \theta^{y*}_l\right]}{\lambda^y_{K+1}\left(m', v, z, \boldsymbol{c}\right) + \sum_{l=1}^{K}\lambda^y_l\left(m', v, z, \boldsymbol{c}\right)}\end{aligned} \tag{4.1}$$

其中 $\lambda_{K+1}^{y}\left(m'、\ v、\ z、\ \boldsymbol{c}\right)=\frac{\alpha^{\theta}}{\alpha^{\theta}+n}f_0\left(m'|v、\ z、\ \boldsymbol{c}\right)$，$\lambda_{l}^{y}\left(m'、\ v、\ z、\ \boldsymbol{c}\right)=\frac{n_l}{\alpha^{\theta}+n}f\left(m'|v、\ z、\ \boldsymbol{c};\ \theta_l^{m*}\right)$，$f_0\left(z、\ \boldsymbol{c}\right)$、$\left\{\frac{\alpha^{\omega}}{\alpha^{\omega}+n_l}f_0\left(v、\ z、\ \boldsymbol{c}\right)+\sum_{r=1}^{K_l}\frac{n_{r|l}}{\alpha^{\omega}+n_l}f\left(v、\ z、\ \boldsymbol{c};\ \omega_{r|l}^{*}\right)\right\}$.

对先验分布中的参数进行边缘化后，我们可以分别得到分布和均值，$f_0\left(z,\boldsymbol{c}\right)$，$f_0\left(v,z,\boldsymbol{c}\right)$，$f_0\left(m'|v,z,\boldsymbol{c}\right)$，以及 $\mathrm{E}_0\left[y|m',v,z,\boldsymbol{c}\right]$：$f_0\left(z,\boldsymbol{c}\right)=\int f\left(z,\boldsymbol{c};\ \omega^{x}\right)d\,\mathrm{G}_0\left(\omega^{x}\right)$，$f_0\left(v|z,\boldsymbol{c}\right)=\int f\left(v|z,\boldsymbol{c};\ \omega\right)d\,\mathrm{G}_0\left(\omega^{v}\right)$，$f_0\left(m'|v,z,\boldsymbol{c}\right)=\int f\left(m'|v,z,\boldsymbol{c};\ \theta^{m}\right)d\,\mathrm{G}_0\left(\theta^{m}\right)$，以及 $\mathrm{E}_0\left[y|m',v,z,\boldsymbol{c}\right]=\int \mathrm{E}\left[y|m',v,z,\boldsymbol{c};\ \theta^{y}\right]d\,\mathrm{G}_0\left(\theta^{y}\right)$。我们应用蒙特卡罗积分，在给定参数当前取值$\{\theta^{*},\omega^{*},s\}$的情况下，将 $(V,\boldsymbol{C},\boldsymbol{S})$ 边缘化。给定吉布斯采样器得到的参数 $\{\theta^{*},\omega^{*},s\}$，我们可以从 $\mathrm{P}\left(m|V=v,Z=z,\boldsymbol{C}=\boldsymbol{c};\ \theta^{*},\omega^{*},s\right)$ 中抽取 m；详情请参阅补充材料。我们还可以从定义为 $\mathrm{P}\left(v|Z=z,\boldsymbol{C}=\boldsymbol{c};\ \theta^{*},\omega^{*},s\right)$ 的分布中抽取 v

$$\begin{aligned}&\mathrm{P}\left(v|Z=z,\boldsymbol{C}=\boldsymbol{c};\theta^{*},\omega^{*},s\right)\\&=\frac{\lambda_{K_K+1}^{v}\left(z,\boldsymbol{c}\right)f_0\left(v|z,\boldsymbol{c}\right)+\sum_{l=1}^{K}\sum_{r=1}^{K_l}\lambda_{r|l}^{v}\left(z,\boldsymbol{c}\right)f\left(v|z,\boldsymbol{c};\omega_{r|l}^{v*}\right)}{\lambda_{K_K+1}^{v}\left(z,\boldsymbol{c}\right)+\sum_{l=1}^{K}\sum_{r=1}^{K_l}\lambda_{r|l}^{v}\left(z,\boldsymbol{c}\right)},\end{aligned}\tag{4.2}$$

其中，$\lambda_{K_K+1}^{v}\left(z、\ \boldsymbol{c}\right)=\left\{\frac{\alpha^{\theta}}{\alpha^{\theta}+n}+\sum_{l=1}^{K}\frac{n_l}{\alpha^{\theta}+n}\frac{\alpha^{\omega}}{\alpha^{\omega}+n_l}\right\}f_0\left(z、\ \boldsymbol{c}\right)$和$\lambda_{r|l}^{v}\left(z、\ \boldsymbol{c}\right)=\frac{n_l}{\alpha^{\theta}+n}\frac{n_{r|l}}{\alpha^{\omega}+n_l}f\left(z、\ \boldsymbol{c};\ \omega_{r|l}^{x*}\right)$。为了从$F_{V'|V=v、\ Z=z、\ \boldsymbol{C}=\boldsymbol{c};\ \rho}\left(v'\right)$中抽样 v'，我们使用式(4)中的高斯Copula模型，

$$F_{V'|V}\left(v'|v;\rho\right)=\Phi\left(\frac{\Phi^{-1}\left(F_{V'}\left(v'\right)\right)-\rho\Phi^{-1}\left(F_V\left(v\right)\right)}{\sqrt{1-\rho^{2}}}\right),\tag{4.3}$$

为简洁起见，这里省略条件变量。首先，从均值为 $\rho\Phi^{-1}\left(F_V\left(v\right)\right)$、方差为 $1-\rho^{2}$的正态分布中抽取 u，然后设定$v'=F_{V'}^{-1}\left(\Phi\left(u\right)\right)$。注意，$F_{V'}^{-1}$难以计算，但我们可以使用标准优化技术（如二分法或牛顿-拉夫森法）计算$\hat{v}'$，使得 $F_{V'}\left(\hat{v}'\right)=\Phi\left(u\right)$。在实际操作中，我们求得一个最优值$\hat{v}'$，使得 $\left\|F_{V'}\left(\hat{v}'\right)-\Phi\left(u\right)\right\|_2<\epsilon_{\text{tol}}$，其中 ϵ_{tol}为给定的容差。对于二元治疗后混杂因素，可以采用 Daniels 等人 [2012] 和 Kim 等人[2017]的方法。

这里简要介绍后处理步骤；详情请参见补充材料。具体而言，我们使用每个后验样本按如下步骤抽取 D 个样本：(1) 从多项分布中抽取$s^{y}_{(d)}$ 和 $s^{x}_{(d)}$ （如果新簇被创建，则从先验分布中抽取新参数），(2) 从 $f\left(\boldsymbol{x}$中抽取$\boldsymbol{c}_{(d)}$； $\omega^{*}_{s^{\boldsymbol{x}}_{(d)}|s^{y}_{(d)}}\right)$，(3) 抽取 $s^{v}_{(d)}$，(4) 从多项分布中抽取 $s^{v}_{(d)}$ ，(5) 从$f\left(v|z\right.$中抽取$v_{(d)}$ ， $\boldsymbol{c}_{(d)}$； $\left.\omega^{v}_{l}\right)$；见（4.2），(6) 从其先验分布 $f\left(\rho\right)$中抽取敏感性参数$\rho_{(d)}$，(7) 从 $\mathrm{P}\left(V_{z'}=v'|V_{z}=v,Z=z,\boldsymbol{C}=\boldsymbol{c}\right)$中抽取 $v'_{(d)}$；见（4.3），(8) 从多项分布中抽取$s^{m'}_{(d)}$，以及(9) 从 $f\left(m'|v'_{(d)}$中抽取 $m'_{(d)}$， z', $\boldsymbol{c}_{(d)}$； $\left.\theta^{m*}_{l}\right)$。给定 D 个值$\left(m'_{(d)},v'_{(d)},v_{(d)},\boldsymbol{c}_{(d)},s_{(d)},\rho_{(d)}\right)$，积分可近似为：

$$\mathrm{E}\left[Y_{z,M_{z'}}\right]=\frac{1}{D}\sum_{d=1}^{D}\mathrm{E}\Big[Y|M=m'_{(d)},V=v_{(d)},Z=z,\boldsymbol{C}=\boldsymbol{c}_{(d)};\theta^{*}_{s^{y}_{(d)}},\omega^{*}_{s^{\boldsymbol{x}}_{(d)}|s^{y}_{(d)}},s_{(d)},\rho_{(d)}\Big].$$

求和中的条件期望见（4.1）。上述步骤均在吉布斯抽样之后执行（详见 Roy 等人 [2018]），因为这些步骤都是马尔可夫链蒙特卡洛之后的处理。对于 $(Y,M,V,\boldsymbol{X})$ 中 $\boldsymbol{X}=\left(Z,\boldsymbol{C}^{\top}\right)^{\top}$ 的缺失，由于已指定 $(Y,M,V,\boldsymbol{X})$ 的联合分布，因此在可忽略缺失假设下通过数据增广处理；详情见补充材料。

5 模拟

为了评估所提出方法的性能，我们基于 Kim 等人 [2017]；Roy 等人 [2018] 和 Hong 等人 [2022] 所介绍的方法构建了十二种数据情景。我们关注的因果参数包括边际自然直接效应、自然直接效应和总效应。对于所提出的贝叶斯非参数（BNP）方法，我们对连续变量进行了标准化-

连续协变量。我们在（5.1）中指定了以下分布：

$$
\begin{aligned}
Y_i|M_i, V_i, \boldsymbol{X}_i; \theta_i^y &\sim \mathrm{N}\big(\mathbb{M}_i\beta_i^y, \sigma_i^{y,2}\big) \\
M_i|V_i, \boldsymbol{X}_i; \theta_i^m &\sim \mathrm{N}\big(\mathbb{V}_i\beta_i^m, \sigma_i^{m,2}\big) \\
V_i|\boldsymbol{X}_i; \omega_i^v &\sim \mathrm{N}\big(\mathbb{X}_i\beta_i^v, \sigma_i^{v,2}\big) \\
\boldsymbol{X}_{i,q}; \omega_{i,q}^x &\sim \text{Bernoulli}\big(\pi_{i,q}^x\big),\ q = 1, \cdots, 1+p_1 \\
\boldsymbol{X}_{i,q}; \mu_{i,q}^x, \tau_{i,q}^{x,2} &\sim \mathrm{N}\big(\mu_{i,q}^x, \tau_{i,q}^{x,2}\big),\ q = 1+p_1, \cdots, 1+p_1+p_2 \\
(\theta_i, \omega_i)\,|\,\mathrm{G} &\sim \mathrm{G} \\
\mathrm{G} &\sim \mathrm{EDP}\left(\alpha^\theta, \alpha^\omega, \mathrm{G}_0\right)
\end{aligned}
\tag{5.1}
$$

其中 $\theta_i^y = \big(\beta_i^y, \sigma^{y}{}_{,i}{}^2\big)$, $\theta_i^m = \big(\beta_i^m, \sigma^{m}{}_{,i}{}^2\big)$, $\omega_i^v = \big(\beta_i^v, \sigma^{v}{}_{,i}{}^2\big)$, $\omega_i^x = \big(\pi_i^x, \mu_i^x, \tau^{x}{}_{,i}{}^2\big)$, $\theta_i = (\theta_i^y, \theta_i^m)$ 和 $\omega_i = (\omega_i^v, \omega_i^x)$。这里， $\mathbb{M}$, $\mathbb{V}$ 和 $\mathbb{X}$ 是各个回归的设计矩阵。先验设定请参见补充材料。对于所有十二种情景和样本量（$N = 250$、500、1000、2500），我们使用了 50, 000 次预热期迭代，随后又使用 50, 000 个样本进行后验推断。对于敏感性参数 ρ, 我们指定了 Unif$(0, 1)$ 先验分布。对于每种情景，我们生成了 500 个数据集，并报告偏差、均方误差（MSE）、95%可信区间的长度以及覆盖概率。

我们将 EDPM模型与基于标准序贯可忽略性假设的 EDPM 进行了比较，并将其与基于扩展版标准序贯可忽略性假设的标准贝叶斯参数模型进行了比较；对于前者，我们在（5.1）中删去了 V；对于后者， Y、 M 和 V 使用标准贝叶斯线性模型，而贝叶斯自助法

则用于混杂因素分布。

$$Y_i|M_i, V_i, \boldsymbol{X}_i; \theta_i^y \sim \mathrm{N}\big(\mathbb{M}_i\beta_i^y, \sigma_i^{y,2}\big)$$

$$M_i|V_i, \boldsymbol{X}_i; \theta_i^m \sim \mathrm{N}\big(\mathbb{V}_i\beta_i^m, \sigma_i^{m,2}\big)$$

$$V_i|\boldsymbol{X}_i; \theta_i^v \sim \mathrm{N}\big(\mathbb{X}_i\beta_i^v, \sigma_i^{v,2}\big)$$

$$\boldsymbol{X}_i; \omega_i^x \sim \mathrm{BB}(\pi_i^x)$$

其中 $\theta_i^y = \big(\beta_i^y, \sigma^{y,\ _i{}^2}\big)$, $\theta_i^m = \big(\beta_i^m, \sigma^{m,\ _i{}^2}\big)$, $\omega_i^v = \big(\beta_i^v, \sigma^{v,\ _i{}^2}\big)$, $\omega_i^x = \big(\pi_i^x, \mu_i^x, \tau^{x,\ _i{}^2}\big)$, $\theta_i = (\theta_i^y, \theta_i^m)$, 且 $\omega_i = (\omega_i^v, \omega_i^x)$。我们在 EDPM 设定中使用了类似的先验分布。对于敏感性参数 ρ，我们指定一个 Unif$(0, 1)$ 先验分布。

由于我们的动机示例是随机试验，因此在所有情景中，我们都从 $Z \sim$ 伯努利分布(0.5) 生成一个处理变量。在情景 1–6 中，我们使用简单的函数形式生成连续基线混杂因素；生成了两个相互独立的连续基线混杂因素(C_1, C_2)。在情景 7–12 中，我们加入了函数形式更复杂的基线混杂因素。具体而言，我们生成了九个二元基线混杂因素和六个连续基线混杂因素，$(C_1, \cdots, C_{15})$。随后，在情景（1、2、3、7、8 和 9）中，我们从二元正态分布抽取一个连续治疗后混杂因素，(V_{z_0}, V_{z_1})。在其余情景（4、5、6、10、11 和 12）中，连续治疗后混杂因素由二元伽马分布生成。对于每个情景，我们从偏态正态分布生成一个中介变量 M。除情景 3、6、9 和 12 外，结局 Y 均由正态混合分布生成。对于情景 3、6、9 和 12 中的结局，我们在数据生成机制的均值中加入了更复杂的项。详细设定请参见表1。

5.1 结果

表2展示了模拟研究的结果。总体而言，结果表明，即使样本量较小（$N = 250$），所有估计值也近似无偏。随着

样本量增加，偏差的绝对值、均方误差和95%可信区间的长度均有所减小，而经验覆盖概率仍接近95%。即使违反了处理后混杂变量的二元正态性假设，所提出的模型仍具有稳健性。比较情景1与4、情景2与5、情景3与6、情景7与10、情景8与11，以及情景9与12的结果，可以看出这一点。不过，与情景1、2和3相比，情景4、5和6的均方误差及95%可信区间长度分别略小。此外，当正态分布均值中包含交互项时，均方误差和95%可信区间长度会更大（情景1与2、情景4与5、情景7与8，以及情景10与11）。在含有非线性项的设定中（情景3、6、9、12），自然间接效应和自然直接效应的偏差绝对值略大，但如预期所示，偏差会随样本量增加而消失。表3给出了在标准序贯可忽略性假设下使用EDPM得到的模拟结果，其中我们在（5.1）中删除了 V。总体而言，结果表明，大多数估计值存在偏差，而且即使样本量增加，均方误差和95%可信区间长度仍会增加。此外，我们观察到，在许多情景中，经验覆盖概率与目标值95%明显偏离。

表4展示了在扩展序贯可忽略性假设下，采用标准贝叶斯参数模型进行模拟研究的结果。结果表明，除情形3外，所有情形下的估计值大多无偏；随着样本量增加，均方误差和95%可信区间的长度均有所下降。然而，在每种情景下，至少有一种因果效应（自然间接效应、自然直接效应或平均处理效应）的经验覆盖概率低于95%（有时低很多，例如S3）。这表明贝叶斯参数模型倾向于低估可信区间的长度。

6 Rural LITETrial

我们使用所提出的方法评估了Rural LITE 试验中的中介分析 [Perri 等人，2014]。Rural LITE 试验旨在考察通过农村社区的合作推广办公室提供的行为减重治疗的效果和成本。受试者（N =612）被随机分配到4个治疗组之一：(1) 对照组（$N = 169$）；(2) 低剂量组（N =148）；(3) 中等剂量组（$N = 134$）；(4) 高剂量组（$N = 161$）。研究者希望探究出勤率是否为研究期间体重变化的潜在中介变量。体重在第0个月（基线）、第6个月和第24个月测量。主要结局为基线至最终随访（第24个月）的体重变化。基线至第6个月的体重变化被视为治疗后混杂因素。我们纳入了基线混杂因素：性别、年龄、年龄、种族和体质指数。

在原始试验的分析中，对照组和低剂量面部治疗组的体重维持效果相似。因此，我们将这两种干预合并为一个干预组（$z = 0$）。出于同样的原因，我们将中剂量组和高剂量组合并（$z = 1$）。在此，我们评估低剂量（LD）组相对于高剂量（HD）组的中介变量自然间接效应（NIE）和自然直接效应（NDE）。两个治疗组的样本量分别为 $N = 317$ 和 $N = 295$。分别有121名受试者（第24个月体重）和56名受试者（第6个月体重）的主要结局和/或治疗后混杂因素信息不完整。我们在可忽略缺失的假设下，通过蒙特卡洛马尔可夫链（MCMC）数据增广算法，基于观测数据和EDPM中的参数处理这些缺失；详情请参见补充材料。针对治疗后混杂因素假设（4）进行敏感性分析时，我们考察了敏感性参数 ρ 的多个取值以及若干种先验分布。

6.1 结果

对于观测数据模型参数的先验分布，我们采用与模拟研究相同的设定。为从观测数据的后验分布中抽样，我们运行了100,000次迭代，将前50,000次作为预热期舍弃，并以100次为间隔进行抽稀，最终获得500个后验样本。

表5展示了基于第3节所述假设，在敏感性参数 ρ 取不同值时 NIE、NDE 和平均处理效应的估计值。结果表明，对于平均处理效应和 NDE 估计值，所提出的模型对 ρ 的选择具有稳健性；二者均具有显著影响，这一点由不包含零的95% 可信区间表明。另一方面，NIE 估计值的结果大多会随 ρ 符号的不同而变化。具体而言，当 $\rho < 0$ 时，NIE 具有统计显著性（$i.e.$，95%可信区间不包含0）；而当 $\rho > 0$时，NIE 不具有统计显著性。不同 ρ 取值下的平均处理效应估计值正如预期的那样相近，表明平均处理效应对 ρ 的选择不敏感（正应如此）。

NDE表示研究结束时HD组（$z = 1$）与LD组（$z = 0$）之间体重变化的平均差异，前提是将HD组的出勤设定为未接受干预时本应达到的水平。NDE为负且具有统计显著性，说明HD对体重变化有显著影响，且这种影响并非由出勤造成。另一方面，NIE表示所有受试者均处于HD组时的平均体重变化，与所有受试者均处于HD组但其出勤被设定为未接受干预时本应达到的水平时的平均体重变化之间的差异。表5中NIE的95% 置信区间在 $\rho > 0$ 时包含零、在 $\rho < 0$ 时不包含零，这表明由出勤变化导致的HD干预效应在 $\rho > 0$ 时不显著，而在 $\rho < 0$时显著。我们的分析表明，HD干预在平均体重减轻方面更

有效。关于研究期间HD干预的效应是否由出勤率介导，现有证据并不一致。我们还考察了按种族（黑人或白人）划分的条件因果效应；相关计算见补充材料第3节。表6列出了敏感性参数 ρ取不同值时，按种族划分的NIE、NDE和TE估计值。对于种族而言，ρ 取不同值时，结果均一致。黑人的平均处理效应估计值低于白人。两种族的平均处理效应和NDE均显著（$i.e.$，95% 可信区间不包含0）。NIE均不显著（$i.e.$，95% 可信区间包含0）。不过，白人的NIE估计值接近显著，其95% 可信区间的上限非常接近0。

我们还进行了额外的数据分析，以考察不同ρ 值下的总体中介效应（见补充材料中的表1），以及在标准序贯可忽略性假设下（此时剔除了治疗后混杂因素；见补充材料中的表2）。值得注意的是，在不存在治疗后混杂因素的情况下，我们几乎没有发现自然间接效应（NIE）的证据（其中心值为0）。此外，我们还进行了额外的数据分析，以研究不同 ρ 值下的条件中介效应，结果见补充材料中的表3。

7 讨论

本文提出了一种灵活贝叶斯非参数方法，用于在存在治疗后混杂因素的情况下研究中介的因果效应。具体而言，我们将EDPM模型应用于观测数据，并结合近期提出的一组假设，以便在存在治疗后混杂因素的情况下实现识别。我们用于估计因果效应的贝叶斯非参数方法比参数方法更加灵活，同时仍保持计算简便性，因此是因果推断研究中的一项有价值工具。

对于 Rural LITE 数据，我们发现 HD干预的效应和自然直接效应在统计上显著，但关于 HD干预的效应是否由出勤率介导，证据存在冲突。此外，我们观察到，间接效应和直接效应的后验分布对非负 ρ 值相对不敏感。此外，我们发现，在假设标准序贯可忽略性且不纳入治疗后混杂因素的情况下所得结果估计出更大的自然直接效应和平均处理效应，且没有自然间接效应的证据。未来的研究将探讨扩展所提出模型的方法，以允许纳入多个治疗后混杂因素和/或中介变量。此外，敏感性分析对于评估假设（2）和/或（3）的潜在违反情况至关重要。在存在治疗后混杂因素的情况下实现识别的一种潜在方向，是探讨更弱的假设，例如 Daniels 等人于 [2012年提出的中介变量诱导等价性假设的混合形式；Kim 等人提出的假设为 2017], ，以及 Hong 等人提出的假设 [2022]。

补充信息

第 1、2、3、4、5 和 6 节中引用的网络附录、表格和图（进一步描述了模型设定、后验计算、模拟设置和缺失值处理）以及相关代码，可在本文对应页面 https://github.com/WoojungBae/EDPpostMediation 获取。

致谢

本研究得到了以下 NIH 资助项目的支持：R01CA183854、R01HL158963、R01HL166324。

参考文献

Albert, J. M. 和 S. Nelson，《广义因果中介分析》，Biometrics，67(3)，1028–1038，2011 年。

Albert, J. M., J. I. Cho, Y. Liu, 和 S. Nelson，《广义因果中介分析与路径分析：扩展与实践考量》，《医学研究统计方法》，28(6)，1793–1807，2019。

Baron, R. M., 和 D. A. Kenny, 《社会心理学研究中的调节变量与中介变量之区分：概念、策略与统计考量》，《人格与社会心理学杂志》，51(6)，1173–1182，1986年。Daniel, R. M., B. L. De Stavola, S. Cousens, 和 S. Vansteelandt, 《多个中介变量的因果中介分析》，《Biometrics》，71(1)，1–14，2015年。Daniels, M. J., J. A. Roy, C. Kim, J. W. Hogan, 和 M. G. Perri, 《中介因果效应的贝叶斯推断》，《Biometrics》，68(4)，1028–1036，2012年。Díaz, I., N. S. Hejazi, K. E. Rudolph, 和 M. J. van Der Laan, 《存在中间混杂变量时的非参数高效因果中介分析》，《Biometrika》，108(3)，627–641，2021年。Didelez, V., P. Dawid, 和 S. Geneletti, 《序贯处理的直接效应和间接效应》，arXiv 预印本 arXiv:1206.6840，2012年。Elliott, M. R., T. E. Raghunathan, 和 Y. Li, 《使用主分层法对二分中介变量和结局的因果中介效应进行贝叶斯推断》，《Biostatistics》，11(2)，353–372，2010年。Ferguson, T. S., 《若干非参数问题的贝叶斯分析》，《统计年鉴》，1(2)，209–230，1973年。Geneletti, S., 《在非反事实框架下识别直接效应和间接效应》，《英国皇家统计学会杂志：B辑（统计方法论）》，69(2)，199–215，2007年。Hejazi, N. S., K. E. Rudolph, M. J. Van Der Laan, 和 I. Díaz, 《针对随机干预（自然）直接效应的非参数因果中介分析》，arXiv 预印本 arXiv:2009.06203，2020年。Hong, G., 《估计自然直接效应和间接效应的中介变量概率比加权法》，载于《美国统计学会生物统计学分会会议录》，第2401–2415页，美国统计学会，弗吉尼亚州亚历山德里亚，2010年。Hong, G., 《社会世界中的因果关系：调节、中介与溢出效应》，John Wiley & Sons，2015年。Hong, G., 和 T. Nomi, 《评估由同伴变化中介的政策效应的加权方法》，《教育效果研究杂志》，5(3)，261–289，2012年。Hong, G., J. Deutsch, 和 H. D. Hill, 《存在处理与中介变量交互作用时用于因果中介分析的中介变量概率比加权法》，《教育与行为统计学杂志》，40(3)，307–340，2015年。Hong, G., X. Qin, 和 F. Yang, 《因果中介研究中的基于加权的敏感性分析》，《教育与行为统计学杂志》，43(1)，32–56，2018年。Hong, G., F. Yang, 和 X. Qin, 《因果中介研究中的处理后混杂：一个前沿问题及通过敏感性分析提出的新颖解决方案》，《Biometrics》，doi:https://doi.org/10.1111/biom.13705，2022年。Huber, M., 《（主要）基于逆概率加权识别因果机制》，《应用计量经济学杂志》，29(6)，920–943，2014年。

Imai, K. 和 T. Yamamoto，多个因果机制的识别与敏感性分析：重新审视框架实验的证据，Political Analysis，21(2)，141–171，2013年。Imai, K.、L. Keele 和 T. Yamamoto，因果中介效应的识别、推断与敏感性分析，Statistical Science，25(1)，51–71，2010年。Imai, K.、L. Keele、D. Tingley 和 T. Yamamoto，揭开因果关系黑箱：从实验研究和观察性研究中了解因果机制，American Political Science Review，105(4)，765–789，2011年。Kim, C.、M. J. Daniels、B. H. Marcus 和 J. A. Roy，用于中介因果效应贝叶斯非参数推断的框架，Biometrics，73(2)，401–409，2017年。Kim, C.、M. J. Daniels、J. W. Hogan、C. Choirat 和 C. M. Zigler，多重中介变量的贝叶斯方法：将主分层与因果中介分析联系起来并应用于发电厂排放控制评估，The Annals of Applied Statistics，13(3)，1927–1956，2019年。Kraemer, H. C.、G. T. Wilson、C. G. Fairburn 和 W. S. Agras，随机临床试验中处理效应的中介变量和调节变量，Archives of General Psychiatry，59(10)，877–883，2002年。Lange, T.、S. Vansteelandt 和 M. Bekaert，估计自然直接效应和间接效应的简明统一方法，American Journal of Epidemiology，176(3)，190–195，2012年。Lin, S.-H. 和 T. VanderWeele，路径特异性效应的干预方法，Journal of Causal Inference，5(1)，20150,027，2017年。MacKinnon, D. P.、C. M. Lockwood、J. M. Hoffman、S. G. West 和 V. Sheets，检验中介效应及其他中间变量效应的方法比较，Psychological Methods，7(1)，83，2002年。Mattei, A.、F. Li 和 F. Mealli，在贝叶斯主分层分析中利用多个结局，并应用于职业培训项目评估，The Annals of Applied Statistics，7(4)，2336–2360，2013年。Miles, C. H.，关于随机干预间接效应的因果解释，doi：10.48550/ARXIV.2203.00245，2022年。Miles, C. H.、I. Shpitser、P. Kanki、S. Meloni 和 E. J. Tchetgen Tchetgen，量化尼日利亚PEPFAR项目中抗逆转录病毒治疗的依从性路径特异性效应，Journal of the American Statistical Association，112(520)，1443–1452，2017年。Miles, C. H.、I. Shpitser、P. Kanki、S. Meloni 和 E. J. Tchetgen Tchetgen，关于存在中介变量-结局混杂时路径特异性效应的半参数估计，Biometrika，107(1)，159–172，2020年。Neal, R. M.，狄利克雷过程混合模型的马尔可夫链抽样方法，Journal of Computational and Graphical Statistics，9(2)，249–265，2000年。Nelsen, R. B.，联结函数导论，Springer Science & Business Media，2007年。

Nguyen, T. Q.、I. Schmid 和 E. A. Stuart，为应用研究者阐明因果中介分析：根据研究目标定义效应，Psychological Methods，26(2)，255–271，2021年。Pearl, J.，直接效应和间接效应，CoRR，abs/1301.2300，2013年。Perri, M. G. 等，三种剂量减重咨询的比较效果：Rural LITE 试验两年随访结果，《肥胖》，22(11)，2293–2300，2014年。Petersen, M. L.、S. E. Sinisi 和 M. J. van der Laan，直接因果效应的估计，《流行病学》，17(3)，276–284，2006年。Preacher, K. J. 和 A. F. Hayes，用于估计简单中介模型中间接效应的 SPSS 和 SAS 程序，《行为研究方法、仪器与计算机》，36(4)，717–731，2004年。Robins, J. M.，因果有向无环图模型的语义及直接效应和间接效应的识别（载于《高度结构化随机系统》，P. Green、N. Hjort 和 S. Richardson 编），牛津统计科学丛书，第70–82页，2003年。Robins, J. M. 和 S. Greenland，直接效应和间接效应的可识别性与可交换性，《流行病学》，3(2)，143–155，1992年。Roy, J.、K. J. Lum、B. Zeldow、J. D. Dworkin、V. L. Re III 和 M. J. Daniels，用于协变量随机缺失情形下因果推断的贝叶斯非参数生成模型，Biometrics，74(4)，1193–1202，2018年。Rubin, D. B.，在随机和非随机研究中估计处理的因果效应，《教育心理学杂志》，66(5)，688–701，1974年。Rubin, D. B.，实验数据的随机化分析：关于 Fisher 随机化检验的评论，Journal of the American Statistical Association，75(371)，591–593，1980年。Rudolph, K. E.、O. Sofrygin、N. M. Schmidt、R. Crowder、M. M. Glymour、J. Ahern 和 T. L. Osypuk，学校和同伴环境对邻里效应与青少年物质使用之间关系的中介作用，《流行病学》（马萨诸塞州剑桥），29(4)，590，2018a。Rudolph, K. E.、O. Sofrygin、W. Zheng 和 M. J. Van Der Laan，随机中介效应的稳健灵活估计：一种拟议方法及其在随机试验情境中的示例，《流行病学方法》，7(1)，2018b。Rudolph, K. E.、N. Williams 和 I. Díaz，在中间混杂和单调性约束下高效灵活地估计自然直接效应和间接效应，Biometrics，n/a(n/a)，doi:https://doi.org/10.1111/biom.13850，2023年。Schwartz, S. L.、F. Li 和 F. Mealli，因果推断中间变量的贝叶斯半参数方法，Journal of the American Statistical Association，106(496)，1331–1344，2011年。Sohn, M. B. 和 H. Li，微生物组研究的组成型中介分析，The Annals of Applied Statistics，13(1)，661–681，2019年。

Tai, A.-S., and S.-H. Lin, 《交换直接效应和间接效应的识别与稳健估计：存在未测量的中介变量-结局混杂和中间混杂时的中介分析》, 《医学统计》, 41(21), 4143–4158, 2022年。Tchetgen Tchetgen, E. J., and I. Shpitser, 《因果中介分析的半参数理论：效率界、多重稳健性和敏感性分析》, 《统计年鉴》, 40(3), 1816–1845, 2012年。Valeri, L., and T. J. VanderWeele, 《允许暴露-中介变量交互及因果解释的中介分析：理论假设及使用 SAS 和 SPSS 宏的实现》, Psychological Methods, 18(2), 137–150, 2013。VanderWeele, T. J., 《用于估计直接效应和间接效应的边际结构模型》, 《流行病学》, 20(1), 18–26, 2009。VanderWeele, T. J., S. Vansteelandt, and J. M. Robins, 《存在暴露诱导的中介变量-结局混杂变量时的效应分解》, 《流行病学》（马萨诸塞州剑桥）, 25(2), 300–306, 2014。Vansteelandt, S., and R. M. Daniel, 《针对多个中介变量的中介分析中的干预效应》, 《流行病学》（马萨诸塞州剑桥）, 28(2), 258–265, 2017。Wade, S., S. Mongelluzzo, and S. Petrone, 《用于贝叶斯非参数推断的富集共轭先验分布》, Bayesian Analysis, 6(3), 359–385, 2011。Wade, S., D. B. Dunson, S. Petrone, and L. Trippa, 《通过富集改进狄利克雷过程混合模型的预测》, The Journal of Machine Learning Research, 15(1), 1041–1071, 2014。Wodtke, G. T., and X. Zhou, 《存在处理诱导混杂时的效应分解：一种残差回归方法》, 《流行病学》, 31(3), 369–375, 2020。Woo, C.-W., M. Roy, J. T. Buhle, and T. D. Wager, 《不同脑系统介导伤害性输入和自我调节对疼痛的影响》, PLoS Biology, 13(1), e1002,036, 2015。Wu, G. D., et al., 《长期饮食模式与肠道微生物肠型的关联》, Science, 334(6052), 105–108, 2011。Yuan, Y., and D. P. MacKinnon, 《贝叶斯中介分析》, Psychological Methods, 14(4), 301–322, 2009。Zhao, Y., X. Luo, M. Lindquist, and B. Caffo, 《功能型中介分析及其在功能性磁共振成像数据中的应用》, arXiv 预印本 arXiv:1805.06923, 2018。

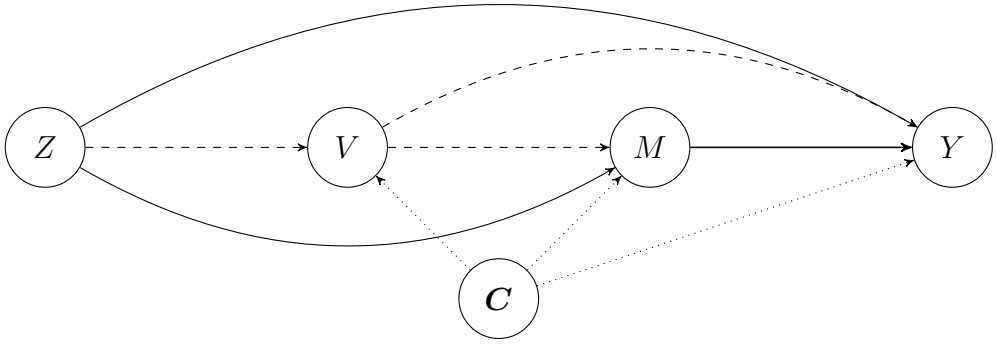


图 1：包含一个处理后混杂变量 V和随机分配的治疗Z 的因果结构，其中 Y为结局， M 为中介变量， C 为一组基线混杂因素。

表1：各模拟场景的数据生成机制

	Scn	数据生成
C_1 C_2	1–6	$\sim \mathrm{N}\left(\mu_1, \sigma_1^2\right);\ \mu_1 = 0, \sigma_1^2 = 9$ $\sim \mathrm{N}\left(\mu_2, \sigma_2^2\right);\ \mu_2 = 0, \sigma_2^2 = 16$
$\begin{pmatrix} V_{z_0} \\ V_{z_1} \end{pmatrix}$	1–3	$\sim \mathrm{N}\left(\begin{pmatrix} \mu_0 \\ \mu_1 \end{pmatrix}, \begin{pmatrix} \sigma_0^2 & \rho_{01}\sigma_0\sigma_1 \\ \rho_{01}\sigma_0\sigma_1 & \sigma_1^2 \end{pmatrix};\ \rho_{01}\right)$
	4–6	$\sim \mathrm{BVG}\left(\begin{pmatrix} \text{伽马(形状参数 =对数 伽马 }(1+\text{指数参数 }(\mu_0)),\text{率参数 }=1) \\ \text{伽马(形状参数 = 对数 }(1+\text{指数参数 }(\mu_1)),\text{ 率参数 }=1) \end{pmatrix};\ \rho_{01}\right)$
		其中 $\mu_0 = 1.3 + 0.6C_1 - 0.7C_2,\ \ \mu_1 = 1.4 - 0.5C_1 + 0.3C_2,\ \ \sigma_0^2 = 3,\ \ \sigma_1^2 = 10,\ \ \rho_{01} = 0.3$
M	1 和 4	$\sim \mathrm{SN}(\xi, \omega, \alpha), \xi = 1.0 + 1.7Z + 0.5V + 0.4C_1 + 0.9C_2, \omega = 3, \alpha = 10$
	2 和 5	$\sim \mathrm{SN}(\xi, \omega, \alpha), \xi = 1.0 + 1.7Z + 0.5V + 0.4C_1 + 0.3C_2, \omega = 3, \alpha = 10$
	3 和 6	$\sim \mathrm{SN}(\xi, \omega, \alpha), \xi = -1.5 + 0.5Z + 0.1V + 0.1C_1 + 0.3C_2, \omega = 1, \alpha = 7$
Y		$\sim 0.6\mathrm{N}\left(\mu_1, \sigma_1^2\right) + 0.4\mathrm{N}\left(\mu_2, \sigma_2^2\right)$
	1 和 4	其中 $\mu_1 = 5 + 2.5Z + 1.8M + 1.3V - 1.2C_1 + 0.3C_2, \sigma_1^2 = 1.5$ 其中 $\mu_2 = -5 - 1.5Z - 1.0M - 0.7V + 0.4C_1 + 0.3C_2, \sigma_2^2 = 0.5$
		$\sim 0.6\mathrm{N}\left(\mu_1, \sigma_1^2\right) + 0.4\mathrm{N}\left(\mu_2, \sigma_2^2\right)$
	2 和 5	其中 $\mu_1 = 5 + 2.5Z + 1.8M + 1.0Z \times M + 1.3V - 0.4C_1 + 0.3C_2, \sigma_1^2 = 1.5$ 其中 $\mu_2 = -5 - 1.5Z - 1.0M - 0.5Z \times M - 0.7V + 0.4C_1 + 0.3C_2, \sigma_2^2 = 0.5$
		$\mathrm{N}\left(\mu, \sigma^2\right)$
	3 & 6	其中 $\mu = 5 + 2.5Z + 0.2\left(M - 0.4\right)_+ + 0.6M^2 + 0.3V + 0.4C_1 + 0.3C_2, \sigma^2 = 0.2$
$C_1 \cdots C_3$	7–12	$\sim$ 伯努利分布$(\pi);\ \ \pi = 0.05$
$C_4 \cdots C_6$		$\sim$ 伯努利分布$(\pi);\ \ \pi = 0.5$
$C_7 \cdots C_9$		$\sim$ 伯努利分布$(\pi_C);\ \ \pi_C = \mathrm{logit}^{-1}\left(a_C \mathrm{logit}^{-1}(b_C) + (1 - a_C)\mathrm{logit}^{-1}(c_C)\right)$
		其中 $a_C = \mathrm{logit}^{-1}\left(2\left(C_{10} - 2\right)^2 - 2\left(C_{11} + 1\right)^2\right),\ \ b_C = 0.6C_{12}C_{13} - 0.2C_{14}^2,\ \ c_C = 0.7C_{12} - 0.4C_{14}C_{15}$
$C_{10} \cdots C_{15}$		$\sim \mathrm{MVN}\left(\mathbf{0},\ \ 0.7\boldsymbol{I} + 0.3\mathbf{1}\mathbf{1}^\top\right),\ \ \mu = \mathbf{0},\ \ \Sigma = 0.7\boldsymbol{I} + 0.3\mathbf{1}\mathbf{1}^\top$
$\begin{pmatrix} V_{z_0} \\ V_{z_1} \end{pmatrix}$	1–3	$\sim \mathrm{N}\left(\begin{pmatrix} \mu_0 \\ \mu_1 \end{pmatrix}, \begin{pmatrix} \sigma_0^2 & \rho_{01}\sigma_0\sigma_1 \\ \rho_{01}\sigma_0\sigma_1 & \sigma_1^2 \end{pmatrix};\ \rho_{01}\right)$
	4–6	$\sim \mathrm{BVG}\left(\begin{pmatrix} \text{伽马(形状 =lo} \quad \mathrm{g}\left(1 + \exp\ (\mu_0)\right), \text{比率 } = 1) \\ \text{Gamma(形状 = 对数 }(1 + \text{指数 }(\mu_1)),\text{ 速率 } = 1) \end{pmatrix};\ \rho_{01}\right)$
		其中 $\mu_0 = 1.3 + 0.5C_{10} - 0.7C_{11} + 0.3C_{12}, \mu_1 = 1.4 + 0.3C_{13} - 0.2C_{14} - 0.4C_{15}, \sigma_0^2 = 3, \sigma_1^2 = 10, \rho_{01} = 0.3$
M	7 和 10	$\sim \mathrm{SN}(\xi, \omega, \alpha), \xi = -1.0 + 1.7Z + 0.5V + 0.3C_1 + 0.1C_4 - 0.2C_7 - 0.4C_{10} + 0.6C_{13}, \omega = 3, \alpha = 10$
	8 和 11	$\sim \mathrm{SN}(\xi, \omega, \alpha), \xi = -1.0 + 1.7Z + 0.5V + 0.3C_1 + 0.1C_4 - 0.2C_7 - 0.4C_{10} + 0.6C_{13}, \omega = 3, \alpha = 10$
	9 和 12	$\sim \mathrm{SN}(\xi, \omega, \alpha), \xi = -0.7 + 0.2Z + 0.1V + 0.5C_1 + 0.6C_4 - 0.2C_7 - 0.4C_{10} + 0.6C_{13}, \omega = 1, \alpha = 7$
Y		$\sim 0.6\mathrm{N}\left(\mu_1, \sigma_1^2\right) + 0.4\mathrm{N}\left(\mu_2, \sigma_2^2\right)$
	7 和 10	其中 $\mu_1 = 5 + 2.5Z + 1.8M + 1.3V + 0.1C_2 + 0.3C_5 - 0.4C_8 - 0.2C_{11} + 0.6C_{14}, \sigma_1^2 = 1.5$ 其中 $\mu_2 = -5 - 0.5Z - 1.0M - 0.7V + 0.3C_3 + 0.1C_6 - 0.2C_9 + 0.6C_{12} - 0.4C_{15}, \sigma_2^2 = 0.5$
		$\sim 0.6\mathrm{N}\left(\mu_1, \sigma_1^2\right) + 0.4\mathrm{N}\left(\mu_2, \sigma_2^2\right)$
	8 和 11	其中 $\mu_1 = 5 + 2.5Z + 1.8M + 1.0Z * M + 1.3V + 0.1C_2 + 0.3C_5 - 0.4C_8 - 0.2C_{11} + 0.6C_{14}, \sigma_1^2 = 1.5$ 其中 $\mu_2 = -5 - 1.5Z - 1.0M - 1.5Z * M + 0.7V + 0.3C_3 + 0.1C_6 - 0.2C_9 + 0.6C_{12} - 0.4C_{15}, \sigma_2^2 = 0.5$
		$\sim \mathrm{N}\left(\mu, \sigma^2\right)$
	9 和 12	其中 $\mu = 5 + 2.5Z + 0.2\left(M - 0.4\right)_+ + 0.6M^2 + 0.3V + 0.2C_2 + 0.3C_5 - 0.4C_8 - 0.2C_{11} + 0.6C_{14},\ \ \sigma^2 = 0.2$

表 2：NIE、NDE 和 ATE 在 500次重复中的模拟结果。各列分别对应偏差、均方误差、CIl（95%可信区间长度）和 CP（经验覆盖概率）。

			250				500				1000				2500			
Scn	CE	真实值	Bias	MSE	CIl	CP	Bias	MSE	CIl	CP	Bias	MSE	CIl	CP	Bias	MSE	CIl	CP
	NIE	1.19	0.01	4.54	8.99	0.97	0.07	2.52	7.31	0.98	-0.10	1.73	5.83	0.97	0.02	0.91	4.15	0.98
S1	NDE	0.95	-0.13	6.34	9.62	0.95	0.03	3.27	7.81	0.98	0.04	2.30	6.37	0.96	-0.03	1.04	5.02	0.98
	ATE	2.14	0.12	4.33	6.71	0.90	0.06	2.18	4.98	0.91	-0.08	1.13	4.20	0.95	0.00	0.41	3.76	1.00
	NIE	1.89	0.08	5.03	9.72	0.97	0.05	2.71	7.82	0.98	-0.08	1.89	6.19	0.98	0.00	1.05	4.41	0.97
S2	NDE	2.56	-0.11	7.41	10.28	0.95	0.08	3.71	8.30	0.96	0.05	2.52	6.74	0.96	0.00	1.13	5.42	0.98
	ATE	4.45	0.10	5.89	7.61	0.90	0.13	2.96	5.64	0.91	-0.06	1.47	4.80	0.95	0.04	0.53	4.40	1.00
	NIE	0.30	-0.05	0.03	0.86	0.98	-0.02	0.01	0.63	0.98	0.02	0.01	0.51	0.98	0.04	0.01	0.42	1.00
S3	NDE	2.53	-0.05	0.04	1.31	1.00	-0.05	0.02	0.87	0.99	-0.05	0.01	0.64	0.99	-0.05	0.01	0.48	0.99
	ATE	2.83	-0.07	0.05	1.13	0.99	-0.05	0.03	0.78	0.97	-0.02	0.01	0.59	0.99	-0.01	0.00	0.49	1.00
	NIE	1.04	0.06	2.03	6.23	0.97	0.00	1.12	5.04	0.98	-0.03	0.86	4.02	0.98	-0.06	0.42	2.85	0.97
S4	NDE	0.72	-0.02	4.80	8.14	0.95	0.16	2.49	6.38	0.97	0.05	1.59	5.24	0.96	-0.01	0.71	4.37	0.99
	ATE	1.76	0.08	4.42	7.05	0.91	0.16	2.27	5.25	0.92	0.03	1.18	4.32	0.94	0.02	0.44	3.99	1.00
	NIE	1.64	0.04	2.32	6.69	0.97	0.04	1.15	5.29	0.99	-0.04	0.89	4.17	0.97	-0.05	0.42	2.96	0.98
S5	NDE	2.52	0.00	5.35	8.45	0.95	0.10	2.73	6.61	0.95	0.01	1.73	5.49	0.96	0.03	0.73	4.71	1.00
	ATE	4.17	0.12	5.57	7.89	0.91	0.21	2.85	5.86	0.92	-0.01	1.51	4.93	0.96	0.01	0.54	4.57	1.00
	NIE	0.18	0.03	0.02	0.69	0.99	0.01	0.01	0.51	0.99	0.02	0.01	0.42	0.99	0.03	0.00	0.34	1.00
S6	NDE	2.39	-0.09	0.03	1.17	1.00	-0.07	0.02	0.76	0.99	-0.05	0.01	0.54	1.00	-0.05	0.00	0.39	1.00
	ATE	2.57	-0.04	0.03	1.07	0.99	-0.05	0.02	0.73	0.98	-0.03	0.01	0.55	0.99	-0.02	0.00	0.44	1.00
	NIE	1.19	-0.01	1.37	4.18	0.92	-0.05	0.51	3.00	0.94	0.02	0.29	2.30	0.95	-0.03	0.14	1.67	0.96
S7	NDE	0.95	0.06	2.74	7.59	0.96	0.00	0.94	4.19	0.97	-0.02	0.63	3.20	0.97	0.01	0.26	2.30	0.97
	ATE	2.14	0.15	2.44	6.70	0.96	0.03	1.15	4.28	0.96	-0.01	0.59	3.15	0.95	-0.02	0.24	2.25	0.98
	NIE	1.89	-0.10	2.13	4.93	0.89	-0.01	0.77	3.52	0.95	-0.01	0.42	2.74	0.96	-0.03	0.20	1.97	0.96
S8	NDE	1.74	0.08	3.59	8.61	0.96	-0.01	1.14	4.46	0.97	0.02	0.72	3.44	0.97	-0.02	0.30	2.48	0.97
	ATE	3.63	0.06	3.22	7.64	0.96	0.06	1.57	4.97	0.94	0.02	0.81	3.64	0.96	-0.03	0.33	2.64	0.98
	NIE	0.30	-0.04	0.02	0.48	0.91	-0.02	0.01	0.36	0.93	-0.02	0.00	0.28	0.95	-0.02	0.00	0.22	0.97
S9	NDE	2.51	-0.03	0.01	0.59	0.99	-0.04	0.01	0.43	0.96	-0.05	0.01	0.33	0.98	-0.06	0.00	0.27	0.98
	ATE	2.81	-0.06	0.04	0.81	0.97	-0.07	0.02	0.57	0.94	-0.07	0.01	0.44	0.93	-0.07	0.01	0.34	0.94
	NIE	1.17	-0.02	1.46	4.10	0.89	-0.03	0.49	2.58	0.93	0.01	0.21	1.85	0.95	0.00	0.10	1.28	0.96
S10	NDE	0.92	0.09	3.10	7.67	0.96	0.07	1.08	3.97	0.96	0.00	0.53	2.89	0.95	0.00	0.22	2.02	0.97
	ATE	2.10	0.12	2.47	6.62	0.96	0.04	1.05	4.12	0.94	-0.01	0.57	3.00	0.95	0.03	0.21	2.06	0.98
	NIE	1.86	-0.04	2.15	4.62	0.89	-0.05	0.61	3.01	0.94	0.02	0.31	2.21	0.95	0.00	0.15	1.55	0.95
S11	NDE	1.77	0.03	3.69	8.57	0.96	0.06	1.19	4.13	0.96	0.00	0.66	3.11	0.96	0.01	0.27	2.17	0.96
	ATE	3.63	0.08	3.07	7.46	0.95	0.00	1.43	4.69	0.94	0.01	0.77	3.49	0.95	0.03	0.28	2.42	0.97
	NIE	0.26	0.01	0.01	0.46	0.95	0.02	0.01	0.34	0.92	0.01	0.00	0.27	0.95	0.00	0.00	0.22	0.99
S12	NDE	2.50	-0.06	0.01	0.42	0.98	-0.08	0.01	0.30	0.91	-0.07	0.01	0.23	0.92	-0.05	0.00	0.20	0.96
	ATE	2.76	-0.06	0.02	0.61	0.95	-0.06	0.01	0.43	0.94	-0.06	0.01	0.32	0.93	-0.05	0.00	0.25	0.95

表3：在不包含处理后变量的情况下，NIE、NDE 和平均处理效应的模拟结果，共进行500次重复。各列分别对应偏差、均方误差、置信区间（95%可信区间的长度）和覆盖概率（经验覆盖概率）。

			250				500				1000				2500			
Scn	CE	真实值	Bias	MSE	CIl	CP	Bias	MSE	CIl	CP	Bias	MSE	CIl	CP	Bias	MSE	CIl	CP
	NIE	1.19	0.90	1.93	4.06	0.86	0.88	1.25	3.17	0.82	0.86	0.99	2.90	0.81	1.05	1.28	3.51	0.85
S1	NDE	0.95	-0.89	4.42	8.30	0.96	-0.79	2.21	5.74	0.92	-0.92	1.67	4.28	0.89	-1.10	1.54	4.04	0.91
	ATE	2.14	0.12	4.62	8.05	0.93	0.14	2.24	5.56	0.94	-0.02	1.11	4.07	0.95	0.00	0.39	3.11	0.98
	NIE	1.89	0.80	3.03	5.36	0.88	0.87	1.72	3.79	0.84	0.87	1.17	2.97	0.79	0.99	1.21	2.99	0.75
S2	NDE	2.56	-0.71	5.16	9.75	0.96	-0.84	2.72	6.61	0.95	-0.97	1.96	4.59	0.89	-1.02	1.44	3.76	0.90
	ATE	4.45	0.09	5.98	9.27	0.93	0.10	2.93	6.43	0.95	-0.03	1.46	4.66	0.94	0.02	0.51	3.68	0.99
	NIE	0.30	0.34	0.14	0.74	0.53	0.36	0.15	0.60	0.26	0.40	0.17	0.49	0.03	0.42	0.18	0.39	0.00
S3	NDE	2.53	-0.42	0.20	0.95	0.64	-0.42	0.19	0.64	0.24	-0.43	0.19	0.45	0.05	-0.43	0.19	0.31	0.00
	ATE	2.83	-0.08	0.05	0.94	0.97	-0.05	0.03	0.67	0.96	-0.02	0.01	0.52	0.98	-0.01	0.00	0.43	1.00
	NIE	1.04	0.47	1.32	4.09	0.92	0.48	0.72	3.33	0.93	0.79	1.04	3.74	0.92	1.01	1.55	3.57	0.90
S4	NDE	0.72	-0.40	4.31	8.54	0.96	-0.35	1.96	5.86	0.95	-0.75	1.50	4.88	0.94	-1.04	1.72	4.61	0.94
	ATE	1.76	0.05	4.59	8.17	0.94	0.19	2.28	5.59	0.94	0.06	1.18	4.14	0.94	0.00	0.42	3.61	0.99
	NIE	1.64	0.40	2.58	5.47	0.91	0.42	1.05	4.10	0.95	0.52	0.87	4.26	0.96	0.64	1.10	3.39	0.91
S5	NDE	2.52	-0.20	5.67	10.14	0.96	-0.35	2.48	6.80	0.95	-0.54	1.54	5.41	0.96	-0.58	1.14	4.60	0.97
	ATE	4.17	0.20	5.99	9.30	0.94	0.22	2.93	6.32	0.93	0.02	1.54	4.74	0.94	0.03	0.54	4.23	0.99
	NIE	0.18	0.15	0.04	0.63	0.90	0.15	0.03	0.49	0.82	0.16	0.03	0.39	0.65	0.14	0.02	0.32	0.61
S6	NDE	2.39	-0.22	0.07	0.84	0.89	-0.20	0.05	0.54	0.79	-0.18	0.03	0.36	0.52	-0.16	0.03	0.23	0.11
	ATE	2.57	-0.05	0.03	0.81	0.97	-0.05	0.02	0.57	0.96	-0.02	0.01	0.44	0.99	-0.01	0.00	0.35	1.00
	NIE	1.19	0.71	1.43	3.42	0.84	0.70	0.87	2.36	0.76	0.71	0.69	1.72	0.60	0.73	0.62	1.20	0.26
S7	NDE	0.95	-0.71	2.40	6.29	0.93	-0.70	1.09	3.00	0.86	-0.73	0.90	2.27	0.77	-0.76	0.72	1.51	0.52
	ATE	2.14	0.08	2.30	5.92	0.95	0.03	1.12	3.95	0.94	-0.02	0.58	2.93	0.94	0.00	0.23	2.00	0.95
	NIE	1.89	0.66	1.89	4.19	0.86	0.67	1.02	2.88	0.83	0.68	0.75	2.14	0.73	0.71	0.62	1.53	0.53
S8	NDE	1.74	-0.67	2.60	4.81	0.92	-0.72	1.24	3.26	0.88	-0.73	0.97	2.44	0.80	-0.76	0.74	1.64	0.60
	ATE	3.63	0.10	3.03	6.38	0.93	-0.02	1.51	4.50	0.93	-0.03	0.78	3.42	0.93	-0.01	0.32	2.35	0.95
	NIE	0.30	0.08	0.04	0.68	0.93	0.07	0.02	0.49	0.92	0.03	0.01	0.38	0.95	0.02	0.00	0.29	0.98
S9	NDE	2.51	-0.13	0.03	0.48	0.82	-0.13	0.02	0.32	0.65	-0.10	0.01	0.21	0.50	-0.10	0.01	0.13	0.13
	ATE	2.81	-0.05	0.04	0.73	0.94	-0.06	0.02	0.52	0.94	-0.08	0.01	0.39	0.91	-0.08	0.01	0.29	0.85
	NIE	1.17	0.24	1.37	3.83	0.90	0.19	0.42	2.40	0.95	0.21	0.26	1.76	0.93	0.22	0.13	1.21	0.90
S10	NDE	0.92	-0.21	2.82	7.15	0.96	-0.18	0.82	3.37	0.94	-0.20	0.51	2.51	0.94	-0.22	0.24	1.72	0.91
	ATE	2.10	0.09	2.36	6.19	0.95	0.03	1.02	3.79	0.93	0.01	0.57	2.82	0.94	0.03	0.22	1.92	0.97
	NIE	1.86	0.18	1.83	4.26	0.90	0.15	0.59	2.91	0.94	0.18	0.37	2.16	0.93	0.23	0.19	1.49	0.91
S11	NDE	1.77	-0.22	3.11	7.35	0.94	-0.20	0.95	3.54	0.93	-0.22	0.62	2.71	0.93	-0.22	0.28	1.85	0.92
	ATE	3.63	0.04	2.97	6.61	0.94	-0.01	1.36	4.28	0.94	0.02	0.77	3.29	0.93	0.03	0.29	2.27	0.97
	NIE	0.26	0.03	0.02	0.50	0.93	0.00	0.01	0.36	0.94	-0.01	0.00	0.28	0.95	-0.02	0.00	0.23	0.99
S12	NDE	2.50	-0.08	0.01	0.34	0.84	-0.06	0.01	0.21	0.80	-0.05	0.00	0.14	0.72	-0.04	0.00	0.08	0.51
	ATE	2.76	-0.05	0.02	0.56	0.92	-0.06	0.01	0.39	0.91	-0.06	0.01	0.30	0.90	-0.06	0.01	0.23	0.90

表4：参数贝叶斯模型下，存在处理后变量时 NIE、NDE 和平均处理效应的模拟结果。基于500次重复。各列分别对应偏差、均方误差、CIl（95%可信区间的长度）和 CP（经验覆盖概率）。

			250				500				1000				2500			
Scn	CE	真实值	Bias	MSE	CIl	CP	Bias	MSE	CIl	CP	Bias	MSE	CIl	CP	Bias	MSE	CIl	CP
	NIE	1.19	-0.02	1.93	4.63	0.89	-0.08	0.85	3.09	0.89	-0.02	0.42	2.17	0.90	-0.03	0.17	1.37	0.92
S1	NDE	0.95	-0.05	5.56	9.17	0.95	0.10	2.60	6.30	0.95	-0.04	1.37	4.49	0.94	-0.03	0.48	2.88	0.96
	ATE	2.14	0.06	4.55	8.17	0.94	0.03	2.25	5.66	0.94	-0.03	1.14	4.01	0.95	0.02	0.40	2.59	0.96
	NIE	1.89	-0.14	3.73	5.84	0.85	-0.01	1.63	3.93	0.86	-0.05	0.75	2.76	0.89	-0.03	0.35	1.76	0.86
S2	NDE	2.56	-0.05	7.11	10.69	0.95	0.08	3.25	7.37	0.95	-0.08	1.71	5.23	0.95	0.00	0.59	3.35	0.98
	ATE	4.45	0.10	5.98	9.37	0.93	0.08	2.99	6.51	0.93	-0.03	1.47	4.59	0.94	0.01	0.52	2.97	0.96
	NIE	0.30	-0.14	0.05	0.42	0.66	-0.15	0.04	0.29	0.51	-0.14	0.03	0.20	0.39	-0.13	0.02	0.13	0.16
S3	NDE	2.53	0.11	0.08	1.04	0.92	0.11	0.05	0.74	0.88	0.13	0.04	0.54	0.86	0.13	0.03	0.37	0.73
	ATE	2.83	-0.01	0.07	0.99	0.94	-0.02	0.04	0.71	0.92	-0.01	0.02	0.52	0.94	0.00	0.01	0.36	0.97
	NIE	1.04	-0.04	1.50	4.04	0.89	-0.06	0.65	2.73	0.89	-0.01	0.36	1.91	0.88	0.00	0.14	1.22	0.90
S4	NDE	0.72	0.06	5.32	9.09	0.94	0.13	2.51	6.30	0.96	0.00	1.28	4.41	0.95	-0.03	0.45	2.82	0.96
	ATE	1.76	0.00	4.69	8.30	0.93	0.14	2.32	5.76	0.93	0.01	1.24	4.05	0.93	0.01	0.43	2.58	0.95
	NIE	1.64	-0.06	3.08	5.18	0.83	-0.02	1.29	3.52	0.86	-0.06	0.72	2.46	0.84	0.01	0.26	1.57	0.87
S5	NDE	2.52	0.21	6.91	10.55	0.95	0.07	3.10	7.30	0.96	0.03	1.60	5.10	0.96	0.00	0.58	3.27	0.96
	ATE	4.17	0.16	5.92	9.40	0.94	0.16	2.96	6.52	0.94	-0.01	1.55	4.62	0.94	0.01	0.53	2.94	0.96
	NIE	0.18	-0.02	0.02	0.38	0.78	-0.02	0.01	0.26	0.76	-0.01	0.01	0.18	0.77	-0.01	0.00	0.12	0.79
S6	NDE	2.39	-0.01	0.06	0.96	0.95	0.00	0.04	0.69	0.92	0.00	0.02	0.50	0.93	0.01	0.01	0.34	0.96
	ATE	2.57	-0.01	0.06	0.91	0.94	-0.01	0.03	0.65	0.90	0.00	0.02	0.48	0.96	0.00	0.01	0.32	0.97
	NIE	1.19	-0.02	1.66	4.19	0.86	-0.04	0.83	2.82	0.88	-0.01	0.46	1.95	0.83	-0.01	0.19	1.25	0.84
S7	NDE	0.95	0.06	3.42	7.58	0.96	0.05	1.54	5.19	0.95	0.00	0.81	3.63	0.95	0.02	0.29	2.31	0.95
	ATE	2.14	0.08	2.74	6.53	0.95	-0.03	1.37	4.45	0.93	-0.01	0.69	3.15	0.93	0.01	0.26	2.00	0.95
	NIE	1.89	-0.05	2.67	4.91	0.85	-0.07	1.36	3.32	0.84	-0.02	0.75	2.32	0.81	0.00	0.31	1.49	0.82
S8	NDE	1.74	0.05	4.40	8.76	0.95	0.05	2.04	5.99	0.96	0.05	1.03	4.19	0.97	0.02	0.37	2.66	0.96
	ATE	3.63	0.08	3.67	7.52	0.94	0.00	1.85	5.17	0.95	-0.02	0.93	3.63	0.92	0.01	0.34	2.32	0.95
	NIE	0.30	-0.02	0.02	0.49	0.91	-0.02	0.01	0.35	0.93	-0.02	0.01	0.27	0.94	-0.02	0.00	0.20	0.97
S9	NDE	2.51	-0.02	0.01	0.51	0.97	-0.03	0.01	0.38	0.97	-0.03	0.00	0.30	0.98	-0.03	0.00	0.25	1.00
	ATE	2.81	-0.05	0.04	0.73	0.95	-0.05	0.02	0.53	0.94	-0.04	0.01	0.40	0.95	-0.04	0.01	0.30	0.96
	NIE	1.17	-0.05	1.63	4.14	0.87	-0.09	0.85	2.82	0.87	-0.02	0.45	1.96	0.84	0.03	0.20	1.24	0.83
S10	NDE	0.92	0.09	3.44	7.48	0.96	0.03	1.50	5.13	0.96	0.00	0.82	3.57	0.95	-0.01	0.31	2.26	0.95
	ATE	2.10	0.16	2.66	6.37	0.95	0.01	1.28	4.36	0.94	0.01	0.70	3.03	0.93	0.03	0.24	1.93	0.94
	NIE	1.86	-0.03	2.64	4.85	0.85	-0.10	1.39	3.28	0.83	-0.03	0.73	2.28	0.81	0.03	0.33	1.47	0.80
S11	NDE	1.77	0.16	4.42	8.59	0.96	0.03	1.97	5.89	0.96	0.02	1.05	4.12	0.95	-0.02	0.41	2.59	0.95
	ATE	3.63	0.14	3.51	7.32	0.94	0.00	1.71	5.03	0.94	0.04	0.93	3.50	0.93	0.03	0.32	2.22	0.94
	NIE	0.26	0.02	0.02	0.47	0.93	0.02	0.01	0.34	0.91	0.02	0.00	0.26	0.94	0.03	0.00	0.20	0.96
S12	NDE	2.50	-0.07	0.01	0.39	0.90	-0.08	0.01	0.29	0.84	-0.08	0.01	0.23	0.82	-0.08	0.01	0.19	0.73
	ATE	2.76	-0.05	0.02	0.57	0.94	-0.05	0.01	0.42	0.93	-0.05	0.01	0.31	0.93	-0.04	0.00	0.23	0.94

表 5：在所提出的贝叶斯非参数模型下，NIE、NDE 和平均处理效应的后验均值（95% 可信区间）。 ρ 是治疗后混杂因素的敏感性参数。

	$\rho \sim$ 均匀分布$(-1,0)$			$\rho \sim$ 均匀分布$(-1,1)$			$\rho \sim$ 三角分布$(0,1,1)$		
CE	Est.	95% 置信区间		Est.	95% 置信区间		Est.	95% 置信区间	
NIE	-0.82	-1.44	-0.26	-0.57	-1.24	0.02	-0.29	-0.76	0.15
NDE	-2.53	-3.98	-1.08	-2.74	-4.28	-1.23	-3.02	-4.46	-1.64
ATE	-3.34	-4.75	-2.01	-3.31	-4.76	-1.94	-3.31	-4.77	-1.91
	$\rho = -0.8$			$\rho = 0$			$\rho = 0.5$		
CE	Est.	95% 置信区间		Est.	95% 置信区间		Est.	95% 置信区间	
NIE	-0.93	-1.44	-0.38	-0.59	-1.03	-0.16	-0.38	-0.79	0.07
NDE	-2.38	-3.87	-0.87	-2.73	-4.19	-1.36	-2.97	-4.41	-1.58
ATE	-3.30	-4.79	-1.94	-3.32	-4.74	-1.95	-3.34	-4.77	-1.92

表6：在所提出的贝叶斯非参数模型下，按种族（B=黑人，W=白人）条件化的 NIE、NDE 和平均处理效应的后验均值（95% 可信区间）。其他协变量已积分消去。 ρ 是治疗后混杂因素的敏感性参数。

		$\rho \sim$ 均匀分布$(-1,0)$			$\rho \sim$ 均匀分布$(-1,1)$			$\rho \sim$ 三$(0,\ 1,\ 1)$		
Race	CE	Est.	95% 置信区间		Est.	95% 置信区间		Est.	95% 置信区间	
B	NIE	-0.11	-0.56	0.35	-0.10	-0.57	0.31	-0.08	-0.56	0.35
	NDE	-2.67	-4.22	-1.11	-2.66	-4.25	-1.10	-2.67	-4.17	-1.14
	ATE	-2.78	-4.43	-1.13	-2.76	-4.31	-1.12	-2.75	-4.36	-1.09
W	NIE	-0.42	-0.93	0.04	-0.41	-0.94	0.13	-0.42	-0.95	0.05
	NDE	-3.02	-4.47	-1.55	-3.04	-4.64	-1.62	-3.03	-4.62	-1.56
	ATE	-3.44	-4.93	-1.95	-3.45	-4.97	-1.95	-3.44	-4.96	-1.98
		$\rho = -0.8$			$\rho = 0$			$\rho = 0.5$		
B	NIE	-0.11	-0.60	0.34	-0.10	-0.61	0.31	-0.14	-0.60	0.35
	NDE	-2.70	-4.25	-1.13	-2.71	-4.21	-1.13	-2.64	-4.20	-1.03
	ATE	-2.80	-4.48	-1.18	-2.81	-4.37	-1.17	-2.78	-4.33	-1.14
W	NIE	-0.41	-0.99	0.08	-0.41	-0.92	0.09	-0.42	-0.96	0.07
	NDE	-3.05	-4.61	-1.59	-3.03	-4.57	-1.60	-3.03	-4.62	-1.55
	ATE	-3.46	-5.03	-2.03	-3.44	-4.96	-2.01	-3.44	-4.98	-1.97