arXiv:2603.25641v1 [econ.EM] 26 Mar 2026

二元网络形成中效用可转移性的计量经济学

Joseph
Marshall*牛津大学†

2026年3月27日

摘要

本文研究如何估计一个个体与网络中的另一个个体建立联系时所偏好的程度。本文比较了在假设效用可在个体之间转移与不作此假设的情况下，以及在假设配对中各个体的回归变量对称与不作此假设的情况下，估计的难易程度。我证明，当配对特定回归变量对称时，假设效用可转移的最大似然估计量（TU-MLE）的一致性和渐近正态性的充分条件，也足以保证不假设效用可转移的最大似然估计量（NTU-MLE）具有一致性和渐近正态性。当回归变量不对称时，我给出了 NTU-MLE 具有一致性和渐近正态性的充分条件。我还提出了一种设定检验，用于评估可转移效用假设是否成立。来自经济学不同领域的两个应用展示了这些结果的价值。我发现，有证据表明，研究者在可转移效用假设不成立时仍使用 TU-MLE；也有证据表明，在无法拒绝可转移效用假设时，研究者会使用基于 NTU 模型的估计量。

*joseph.marshall@nuffield.ox.ac.uk
†纳菲尔德学院，牛津，英国，OX1 1NF

1 引言

给定一组个体，我们能否确定一个个体从与另一个个体建立联系中获得多少收益？进而，我们能否预测谁会与谁建立联系？这些问题是研究社会与经济网络的核心。例如，假设一项新技术问世，贸易经济学家希望预测哪些国家会成为该技术的出口国，哪些会成为进口国。或者设想有一个非政府组织，负责开展团体认知行为疗法课程，以改善处境困难的青少年的心理健康。如果该非政府组织希望青少年与所在小组中的其他人建立友谊，以便他们能够讨论共同经历，那么该组织就需要了解哪些小组设计最有利于友谊形成。将男孩和女孩分开编组会更好吗？还是将具有不同特征的青少年编在一起更好？本文为旨在回答此类问题的计量经济学模型作出了贡献。

要理解本文的贡献，有必要介绍两种用于估计个体彼此建立连接的偏好的计量经济学模型。第一种模型假设效用可以在个体之间转移；第二种则不作此假设。本文的主要目标是比较相较于可转移效用模型，不可转移效用模型的估计所需的假设。为此，本文比较了与每种模型相对应的最大似然估计量满足一致性和渐近正态性的充分条件。

1.1 可转移效用（TU）模型

用于确定两个个体建立连接能从中获得多少收益的较简单计量经济学模型，称为可转移效用（TU）模型。这是一种二元选择模型，其结果变量是一个指示函数，用于表示这两个个体是否决定建立连接。该模型得名于其假设：效用可以在个体之间转移。如果个体 i 希望与个体 j建立连接，但 j不愿意回应，那么这一假设意味着 i 可以向 j 提供某种实物转移（例如金钱），使建立连接对 j而言变得有利可图。这一假设在实践中的含义是，我们只需要一个指示函数，就能概括这两个个体的联合连接决策。

形式上，假设个体 i 和 j 形成连接（记作 $Y_{ij}^{\mathrm{TU}} = 1$），当且仅当他们从该连接中获得的联合效用非负。联合效用由三部分组成：1）某个连续函数 $W_{ij} := w(X_i, X_j)$，其定义域为 $\mathbb{R}^{k\times 2}$，值域为 $\mathbb{R}^k$，它将i的特征随机向量 $X_i \in \mathbb{R}^k$与 j的特征随机向量$X_j \in \mathbb{R}^k$按分量组合；2）一个“连接偏好参数”向量 $\beta \in \Theta \subseteq \mathbb{R}^k$；以及 3）一个连接特定误差项 $\varepsilon_{ij} \in \mathbb{R}$，假设其服从标准正态分布。总之，TU 模型表示为：

$$Y_{ij}^{\mathrm{TU}} = \mathbb{1}[W_{ij}'\beta \geq \varepsilon_{ij}],$$

其中 $\varepsilon_{ij} \sim \mathcal{N} \in (0,1)$。

例如，假设 $\beta \in \mathbb{R}$ 和 $W_{ij} = |X_i - X_j|$，其中 X_i 是 i的收入，而 X_j是 j的收入。在这种情况下，β 表示收入差距对 i 和 j之间

成为朋友的似然的效应。我们通常会发现，具有相似特征的个体更有可能成为朋友，这种现象在网络文献中被称为“同质性”（Jackson，2008年）。因此，我们可能会预期此处的 β 为负，以反映收入相差较大的人不太可能成为朋友。从这个意义上说，β （的相反数）可以视为同质性参数。

不考虑自连接，并对唯一连接求和，我们得到 TU 模型的对数似然函数：[1]

$$\frac{1}{N}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\Big[Y_{ij}^{\mathrm{TU}}\log\big(\Phi(W_{ij}^{'}\beta)\big)+(1-Y_{ij}^{\mathrm{TU}})\log\big(1-\Phi(W_{ij}^{'}\beta)\big)\Big],$$

其中，Φ 是单变量标准正态累积分布函数，$N=\binom{n}{2}=\frac{n(n-1)}{2}$ 是不包括自连接的唯一二元组数量。将其与众所周知的非二元组数据 Probit 对数似然函数进行比较：

$$\frac{1}{n}\sum_{i=1}^{n}\Big[Y_{i}\log\big(\Phi(X_{i}^{'}\beta)\big)+(1-Y_{i})\log\big(1-\Phi(X_{i}^{'}\beta)\big)\Big],$$

其中，$Y_i=\mathbb{1}[X_i'\beta\geq\varepsilon_i]$ 和 $\varepsilon_i\sim\mathcal{N}(0,1)$。根据 Newey和McFadden（1994）中的例 1.2，我们知道，只要 $\mathbb{E}[X_iX_i']$ 存在且非奇异，并且数据（Y_i，X_i）是独立同分布的，Probit 模型的最大似然估计量就是相合且渐近正态的。TU 模型的对数似然函数本质上是 Probit 对数似然函数，只是采用了双重求和，并使用不同的 i 和 j 索引，以体现观测值处于二元组层级这一事实。对于二元模型，无法假设数据（Y_{ij}，X_i，X_j）是独立同分布的；Y_{12} 和 Y_{13} 都依赖于 X_1，因此显然相关。相反，我们只能假设 X_i 是独立同分布的；要据此确立渐近正态性，就需要采用比 Newey和McFadden（1994）更强的有限矩假设。具体而言，可以证明，只要 $\mathbb{E}[W_{ij}W_{ij}']$ 存在且非奇异，并且X_i 是独立同分布的（此外，参数空间 Θ还须满足一些标准正则条件），TU 模型的最大似然估计量（TU-MLE）就是相合的；若进一步满足 $\mathbb{E}[\|W_{ij}\|^8]$存在，则该估计量渐近正态。[2]

1.2 不可转移效用（非转移效用）模型

如果不愿意假设个体之间的效用可以转移，就无法将两个个体的连边决策表示为一个包含联合效用项的单一指示函数。相反，要形成一条连接，记作 $Y_{ij}^{\mathrm{NTU}}=1$，我们要求 i 和 j 形成该连接所获得的各自效用同时非负。形式上，我们将两个指示函数相乘：一个指示函数包含 i形成该连接所获得的效用，而

1. 由于 $Y_{ij}^{\mathrm{TU}}=Y_{\text{可转移效用}ji}$，因此我们有重复连接；也就是说，TU 模型生成无向网络。为了解双重求和如何忽略自连接和重复连接，令 $n=3$。那么，纳入考虑的连接为 (1，2)、(1，3) 和 (2，3)。2. 只需照搬 Newey和McFadden（1994）中关于 probit 的证明，同时考虑二元组数据带来的额外复杂性，具体做法如下文第2节所述。有关这些假设的讨论，请参见第2节。

另一个指示函数包含 j形成该连接所获得的效用。 i的效用由 W_{ij} 构成：$=w(X_i,\ X_j)$ $\in \mathbb{R}^k$、 $\beta \in \Theta \subseteq \mathbb{R}^k$和 $\varepsilon_{ij} \in \mathbb{R}$； j的效用由 W_{ji} 构成：$= w(X_j,\ X_i)\ \in \mathbb{R}^k$、$\beta \in \Theta \subseteq \mathbb{R}^k$和 $\varepsilon_{ji} \in \mathbb{R}$。我们假设$(\varepsilon_{ij},\ \varepsilon_{ji})$服从二元正态分布，Var$(\varepsilon_{ij})$和Var$(\varepsilon_{ji})$均标准化为1，且 Cov$(\varepsilon_{ij},\ \varepsilon_{ji})\ =\ \rho\ \in\ (-1,\ 1)$。误差$(\varepsilon_{ij},\ \varepsilon_{ji})$在二元组内部可以相关（由 ρ决定），但假设它们在不同二元组之间相互独立。总之，非转移效用模型表示为：

$$Y_{ij}^{\mathrm{NTU}} = \mathbb{1}[W_{ij}^{'}\beta \geq \varepsilon_{ij}] \cdot \mathbb{1}[W_{ji}^{'}\beta \geq \varepsilon_{ji}],$$

其中

$$\begin{pmatrix} \varepsilon_{ij} \\ \varepsilon_{ji} \end{pmatrix} \sim \mathscr{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right).$$

在本文所考虑的最一般非转移效用模型中，我们允许存在这样的个体对ij ，使得 $W_{ij} \neq W_{ji}$。这类回归变量称为“非对称回归变量”。尽管在许多应用中，回归变量是对称的，例如，当 $W_{ij} = |X_i - X_j| = |X_j - X_i| = W_{ji}$ 时，但在某些情况下，我们可能希望允许存在非对称回归变量。一个例子是，某个个体只关心另一个人的收入，即 $W_{ij} = X_j \neq X_i = W_{ji}$ ，其中 X_i 是 i 的收入。不过，我们始终假设两个个体具有共同的 β。另一种模型表述方式是强制要求回归变量对称，同时允许个体 i 和 j 对连边有不同的偏好，分别为 β_1 和 β_2；Poirier（1980）采用的正是这种表述。[3] 由于 i 和 j 的顺序是任意的，相较于两者特征的差异，我认为强行设定 i 和 j 的偏好存在差异不太自然。

无论回归变量是否对称，非转移效用模型都会生成对称结果，$Y_{\text{非转移效用}_{ij}} = Y_{ji}^{\mathrm{NTU}}$，因此产生无向网络。这可直接从模型的定义得出：$Y_{\text{非转移效用}_{ij}} = \mathbb{1}[W_{ij}^{'}\beta \geq \varepsilon_{ij}] \cdot \mathbb{1}[W_{ji}^{'}\beta \geq \varepsilon_{ji}] = \mathbb{1}[W_{ji}^{'}\beta \geq \varepsilon_{ji}] \cdot \mathbb{1}[W_{ij}^{'}\beta \geq \varepsilon_{ij}] = Y^{\text{非转移效用}ji}$。

与 TU 模型的对数似然函数类似，我们对唯一连接（不包括自连接）求和，从而得到非转移效用模型的对数似然函数：

$$\hat{Q}_n(\beta, \rho) \coloneqq \frac{1}{N} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left[Y_{ij}^{\mathrm{NTU}} \log\left(\Phi_2^{\rho} \begin{pmatrix} W_{ij}^{'}\beta \\ W_{ji}^{'}\beta \end{pmatrix} \right) + (1 - Y_{ij}^{\mathrm{NTU}}) \log\left(1 - \Phi_2^{\rho} \begin{pmatrix} W_{ij}^{'}\beta \\ W_{ji}^{'}\beta \end{pmatrix} \right) \right],$$

其中， Φ_2^{ρ} 是协方差为 ρ的二元正态分布的累积分布函数。本文提出的理论结果均适用于 $\rho = 0$这一特殊情形，即 ε_{ij} 与 ε_{ji}相互独立。[4] 当 $\rho = 0$时，我们用$\hat{}\ Q_n(\beta)$代替$\hat{}\ Q_n(\beta, 0)$。不难看出：

3. 请注意，Poirier（1980）只关注其模型的识别，并未考察一致性或渐近正态性。我们将在第2.2节看到，其识别问题在我们的模型中再次出现，但程度较轻。4. 当 $\rho \neq 0$ 时的模拟结果见附录。

$$\hat{Q}_n(\beta) \coloneqq \frac{1}{N}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\Big[Y_{ij}^{\mathrm{NTU}}\log\big(\Phi\big(W_{ij}'\beta\big)\Phi\big(W_{ji}'\beta\big)\big) + (1-Y_{ij}^{\mathrm{NTU}})\log\big(1-\Phi\big(W_{ij}'\beta\big)\Phi\big(W_{ji}'\beta\big)\big)\Big].$$

这个对数似然更易于处理，因为处理单变量标准正态累积分布函数 Φ，比处理双变量正态累积分布函数 Φ_2^ρ 容易得多。除了简化分析之外，$\rho = 0$ 也是一个自然的特殊情形。事实上，在第一个应用中，我利用现实世界数据估计得到 ρ 非常接近0（见第4.1节）。

本文关注的核心对象是非转移效用模型的最大似然估计量（NTU-MLE），其形式为 $\rho = 0$:

$$\hat{\beta} \coloneqq \arg\max_{\beta\in\Theta}\hat{Q}_n(\beta).$$

ˆ 证明 NTU-MLE（ β）在标准假设下具有一致性和渐近正态性，比证明 TU-MLE 的相应性质更为困难；这正是第2节的重点，也是本文的主要贡献。本文分别考察对称回归变量和非对称回归变量的情形。原因在于，颇为出人意料的是，当回归变量对称时，足以保证 TU-MLE 一致性和渐近正态性的条件，也足以保证 NTU-MLE 具有一致性和渐近正态性。然而，当回归变量非对称时，则需要施加略强一些的充分条件。

1.3 贡献

下面我将概述本文的结构及其贡献。

在网络形成的计量经济学文献中，存在“策略模型”与“二元模型”之分。策略模型允许两个主体形成连接的决策取决于网络中其他连接是否存在；二元模型则忽略更广泛的网络，只关注两个主体的连边决策，而不考虑其他主体的连边决策。尽管策略模型在某些情形下或许更为现实，但它们受到多重均衡问题的困扰，迫使研究者要么像 Christakis 等（2020）那样对均衡选择机制采取某种立场，要么像 Sheng（2020）那样采用集合识别。[5]

上述 TU 模型和非转移效用模型均为二元模型。除了能够规避模型不完备的问题外，在主体做出连边决策时不会考虑更广泛的网络的情形下，二元模型也更符合实际。例如，作为本文所开发工具的一项应用，我研究了坦桑尼亚农村一个小型社区中家庭的风险分担决策。家庭似乎很难了解或预测社区内的整个借贷网络，因此在做出连边决策时，也无法将网络结构纳入考量。

二元关系文献中的大多数论文都假设主体之间的效用可以转移。关于转移效用二元网络形成模型的里程碑论文是 Graham（2017），该论文

5. 关于网络形成策略模型的优秀综述，参见 Chandrasekhar（2016）的第4节和 Paula（2020）的第4节。

在可转移效用链接决策指标中加入了个体 i, A_i 和个体 j, A_j 的未观测个体异质性项:

$$Y_{ij}^{\mathrm{TU}} = \mathbb{1}[W_{ij}^{'}\beta + A_i + A_j \geq \varepsilon_{ij}].$$

引入固定效应会导致偶然参数问题（Neyman 和 Scott，1948年）。为消除固定效应，Graham（2017）假设 ε_{ij} 服从标准逻辑斯蒂分布，且在二元组之间独立同分布，随后使用依赖于 ε_{ij}参数分布的充分统计量。Candelaria（2020年）和 Toth（2017年）将 Graham（2017）的方法扩展到半参数设定，其中 ε_{ij} 的分布不作具体设定；Jochmans（2018年）和 Dzemski（2019年）则研究了能够生成有向网络的相关模型。

截至本文撰写时，已有两篇关于不可转移效用下二元网络形成模型的论文：Gao 等（2023）和 Li 等（2024）。[6] 两篇论文都在各自的指标中纳入了 i 和 j 的未观测个体异质性项，并首先假设回归变量是对称的，[7] 因此其主要模型如下:

$$Y_{ij}^{\mathrm{NTU}} = \mathbb{1}[W_{ij}^{'}\beta + A_i \geq \varepsilon_{ij}] \cdot \mathbb{1}[W_{ij}^{'}\beta + A_j \geq \varepsilon_{ji}].$$

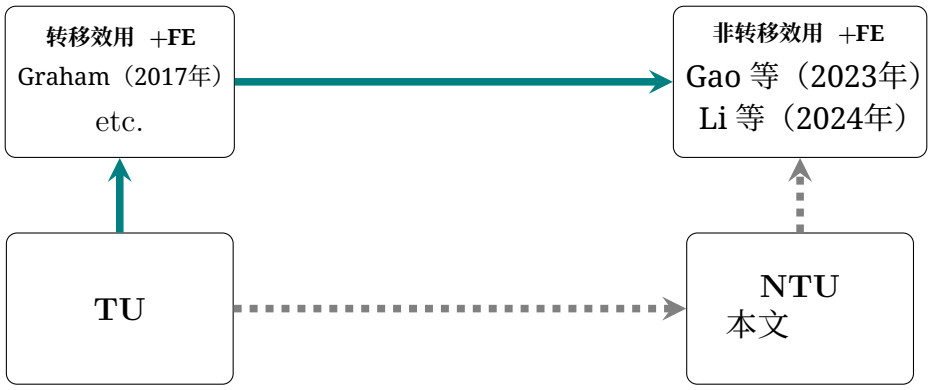
由于 Gao 等（2023）和 Li 等（2024）的研究与本文最为相近，因此有必要详细说明他们具体做了什么，以及他们的贡献与本文的贡献有何不同。Li 等（2024）采用参数化方法，不过，他们对 $(\varepsilon_{ij},\ \varepsilon_{ji})$ 的联合分布设定得足够一般，既涵盖逻辑分布，也涵盖正态分布（并假设 ε_{ij} 和 ε_{ji} 相互独立）。他们的估计程序相当复杂：首先求解一个高维矩方程组，然后将该解代入 Le Cam（1969）的一步估计量，接着通过分割网络刀切法改进估计量，以实现 $\sqrt{N}$ 一致性，最后进行自助聚合，以达到克拉美-罗效率。该程序既可以估计固定效应，也可以估计 β。

另一方面，Gao 等（2023）采用半参数方法，因为他们没有对 $(\varepsilon_{ij}, \varepsilon_{ji})$ 的联合分布作出具体设定。他们通过一种称为“逻辑差分”的程序消除了固定效应；与 TU 模型不同，在非转移效用模型中，固定效应并非加性可分。具体而言，考虑个体 i，其 a) 在某个具有特征 $X_k = \bar{x}$的群体中比 j 更受欢迎，但 b) 在某个具有特征$X_k = \underline{x}$的群体中比 j 更不受欢迎。a) 意味着，要么 $w(X_i, \bar{x})'\beta > w(X_j, \bar{x})'\beta$，要么 $A_i > A_j$；b) 意味着，要么$w(X_i, \underline{x})'\beta < w(X_j, \underline{x})'\beta$，要么 $A_i < A_j$。由于 $A_i > A_j$和 $A_i < A_j$不可能同时成立，Gao 等（2023）得出结论：要么 $w(X_i, \bar{x})'\beta > w(X_j, \bar{x})'\beta$，要么 $w(X_i, \underline{x})'\beta < w(X_j, \underline{x})'\beta$。这些关于 β 的限制被用于识别并进而估计该参数，做法类似于最大得分估计。尽管这种半参数设定比 Li 等（2024）的设定更一般，但 Gao 等（2023）没有提供推断结果，而且由于消除了固定效应，他们无法估计固定效应。

6. 也有采用非转移效用的战略网络形成模型，例如 Goldsmith-Pinkham 和 Imbens（2013）提出的模型，但这些同样不是本文的重点。 7. Gao 等（2023）在附录 C 中将其结果扩展至非对称回归变量，但 Li 等（2024）没有这样做。

相关文献的研究从 TU 模型，发展到含固定效应的 TU 模型，再到含固定效应的非转移效用模型（见图1）。据我所知，尚无人研究第1.2节所述的不含固定效应的非转移效用模型（本文的重点）。尽管未观测个体异质性在许多应用中十分重要，但忽略它的显而易见的优点是，我们可以避免偶然参数问题，并采用最大似然估计，而这比 Gao 等（2023）和 Li 等（2024）的估计程序容易得多。有人可能会担心，本文给出的结果只是 Gao 等（2023）和 Li 等（2024）结果在固定效应设为 0 时的特例。事实并非如此，因为 Gao 等（2023）的逻辑差分估计量和 Li 等（2024）基于矩条件的估计量与标准最大似然估计量截然不同。

图1：直观展示文献中的空白



图中的青绿色实线描绘了二元网络形成计量经济学文献所遵循的方法论演进路径。灰色虚线描绘了本文将采取的路径：考察一种不存在未观测个体异质性的不可转移效用模型。

从 TU 模型（不含固定效应）直接转向 NTU 模型（不含固定效应），也揭示了一些新的理论见解。特别是，第2.1节中的定理1和定理2证明，在误差相互独立且回归变量对称的情况下，TU-MLE 一致性和渐近正态性的标准假设，足以保证 NTU-MLE 的一致性和渐近正态性。也就是说，$\mathbb{E}[W_{ij}W'_{ij}]$的存在性和非奇异性、$X_i$的独立同分布性，以及参数空间 Θ上的一些标准正则条件，足以保证 NTU-MLE 的一致性（定理1）。此外，若进一步假设 $\mathbb{E}[\|W_{ij}\|^8]$ 存在，则可保证 NTU-MLE 具有渐近正态性（定理2）。证明与Newey和McFadden（1994）中的 Probit 示例密切相关，但还必须处理数据二元结构带来的额外复杂性，以及对数似然函数中标准正态分布的累积分布函数被其平方所取代这一事实。与 Gao 等（2023）和 Li 等（2024）相比，这些充分条件更容易解释，也清楚地说明了相较于 TU 模型还需要增加多少假设——而结果表明，在对称回归变量情形下，无需增加任何假设。这一结果主要源于一个令人惊讶的事实：在这种情形下，NTU 模型的对数似然函数是凹函数。

与李等人不同 。（2024年），本文还考察了非对称回归变量的情形 s

（第2.2节的理论结果）以及相关误差（附录中的模拟结果）。尽管 Gao 等（2023）也考虑了非对称回归变量的情形，但他们没有给出其估计量的渐近分布结果，而本文给出了相关结果（定理4）。与 Poirier（1980）的研究相关，使用非对称回归变量时，识别问题较难确立，因此我提出一个易于验证的条件，以确保识别（命题1）。具体而言，除了施加标准的非奇异性条件外，只需检查所有非对称回归变量是否都至少取三个不同值，或者是否存在一对个体，其回归变量具有对称性。这些条件至少有一个在实践中几乎总是成立。使用非对称回归变量时，一致性也更难确立，因为对数似然函数不再是全局凹函数。由于缺乏凹性，我通过调用U统计量的一致大数定律来确立非对称回归变量下的一致性（定理3）。最后，为了得到渐近正态性，我进一步施加一个非奇异性条件，该条件实质上排除了 W_{ij} 和 W_{ji} 之间的独立性（定理4）。所有条件将在后文适时讨论。即使使用非对称回归变量，相较于 TU-MLE 的一致性和渐近正态性的充分条件，所需额外假设也少得出奇。

本文的另一项贡献是提供一种设定检验，用以判断给定数据集由 TU 模型还是非转移效用模型解释得更好。Li 等（2024）讨论了这样一种情况：例如，有人认为误差服从逻辑斯蒂分布，但实际上误差服从正态分布。这类设定错误很值得关注，但可以说，不如另一种情况有趣：有人认为 TU 模型最能解释关系形成，而实际上非转移效用模型才是最佳解释；第3节重点讨论的正是这一问题。我提出了一种似然比设定检验，其考虑到感兴趣参数位于边界这一事实，思路源自 Self 和 Liang（1987）；此外，我还提供了基于模拟的证据，表明该检验具有正确的检验尺寸和较强的检验功效。

在第4节中，我将本文提出的估计和检验程序应用于两个不同的网络数据集。第一个应用研究加利福尼亚一所女子学校中青少年友谊的决定因素，Goeree 等（2010）也曾对此进行探讨。我发现，Goeree 等（2010）在我的设定检验拒绝 TU 假设的情况下仍估计了 TU 模型，而且非转移效用模型的最大似然估计量所揭示的友谊决定因素，与他们的 TU 最大似然估计量所揭示的不同。第二个应用研究坦桑尼亚农村风险分担的决定因素。Gao 等（2023）和 Li 等（2024）也研究了这一应用。我发现，TU 模型对该数据集的解释最佳，这表明 Gao 等（2023）和 Li 等（2024）所使用的基于非转移效用模型的估计量对于该数据集可能并无必要，甚至并不恰当。如果没有用于检验 TU 与非转移效用模型的检验程序，其他研究者就有可能错误地将基于非转移效用模型的估计量用于基于 TU 模型的数据，反之亦然。

最后，我将在第5节作出总结。在附录中，我给出所有结果的证明，以及关于 NTU-MLE 一致性的基于模拟的证据 $\rho \neq 0$ ，并展示 NTU-MLE 异常的小样本行为。

2 NTU-MLE 的性质

再次考虑第1.2节详细介绍的、没有固定效应且误差独立的非转移效用模型：[8]

$$Y_{ij} = \mathbb{1}[W_{ij}'\beta \geq \varepsilon_{ij}] \cdot \mathbb{1}[W_{ji}'\beta \geq \varepsilon_{ji}],$$

其中

$$\begin{pmatrix} \varepsilon_{ij} \\ \varepsilon_{ji} \end{pmatrix} \sim \mathscr{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right),$$

$W_{ij} \coloneqq w(X_i,\ X_j) \in \mathbb{R}^k$, $W_{ji} \coloneqq w(X_j,\ X_i) \in \mathbb{R}^k$，以及 $\beta \in \Theta \subseteq \mathbb{R}^k$。

再考虑非转移效用模型对应的对数似然：

$$\hat{Q}_n(\beta) \coloneqq \frac{1}{N} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \Big[Y_{ij} \log\big(\Phi(W_{ij}'\beta)\Phi(W_{ji}'\beta)\big) + (1 - Y_{ij}) \log\big(1 - \Phi(W_{ij}'\beta)\Phi(W_{ji}'\beta)\big) \Big],$$

其中，Φ 是单变量标准正态累积分布函数，$N = \binom{n}{2} = {}^{n}{}_{(n-1)\ 2\ ,}$ n 是网络中的个体数量。

NTU-MLE 定义如下：

$$\hat{\beta} \coloneqq \arg\max_{\beta \in \Theta} \hat{Q}_n(\beta).$$

现在，我将介绍本文的主要理论贡献：β 的识别条件，以及 NTU-MLE 的一致性和渐近正态性条件。

2.1 对称回归变量情形

本节中，我进一步简化分析，假设回归变量是对称的，即对于所有个体对 ij，$W_{ij} = W_{ji}$。关于非对称回归变量情形，见下文第2.2节。

2.1.1 识别

在对称回归变量的情况下，给定 X_i、X_j 和 β，观察到特定 Y_{ij} 的似然为：

$$f(Y_{ij} \mid X_i, X_j, \beta) = \left[\Phi(W_{ij}'\beta)^2\right]^{Y_{ij}} \left[1 - \Phi(W_{ij}'\beta)^2\right]^{1 - Y_{ij}}.$$

首先要确立的理论结果是 β的识别，即存在一对个体˜，使得对于 β，$\beta \in \Theta$：

$$f(Y_{ij} \mid X_i, X_j, \beta) = f(Y_{ij} \mid X_i, X_j, \tilde{\beta}) \implies \beta = \tilde{\beta}.$$

8. 现在为方便记号，写作 $Y_{ij} = Y_{ij}^{\mathrm{NTU}}$。

或者等价地，根据逆否命题，存在一对个体 ij ，使得 $(X_i, X_j) \neq (\tilde{X}_i, \tilde{X}_j)$ 且 $W_{ij} = \tilde{W}_{ij}$，对于 β，$\beta \in \Theta$，如果 $\beta \neq \beta$，那么 $f(Y_{ij} \mid X_i, X_j, \beta) \neq f(Y_{ij} \mid X_i, X_j, \beta)$。

如第1.1节所述，Newey和McFadden（1994）（以下简称 NM）中 Probit 模型得以识别的关键条件是，随机回归变量仅为X_i 而非 W_{ij}时， $\mathbb{E}[X_i X_i']$ 存在且非奇异，因而为正定矩阵。这是一项标准假设，可排除回归变量之间存在完全多重共线性的情况。具体而言，NM 证明，对于所有个体 i 以及所有 β $\beta \in \Theta$ $\beta \neq \beta$ $\mathbb{E}[X_i X_i']$ ，，若 且 存在并且非奇异，则 $\Phi(X_i'\beta) \neq \Phi(X_i'\beta)$。在我们的设定中，回归变量为 W_{ij} X_i ij 而非仅仅，与之对应的结果是：对于所有个体对以及 β，$\beta \in \Theta$，若 $\beta \neq \beta$和 $\mathbb{E}[W_{ij}W_{ij}']$ 存在且非奇异，则 $\Phi(W_{ij}'\beta) \neq \Phi(W_{ij}'\beta)$。此外，显然 $\Phi(W_{ij}'\beta) \neq -\Phi(W_{ij}'\beta)$，因为 $\Phi(W_{ij}'\beta) > 0$且 $-\Phi(W_{ij}'\beta) < 0$。因此， $\Phi(W_{ij}'\beta)^2 \neq \Phi(W_{ij}'\tilde{\beta})^2$，，从而：

$$
\begin{aligned}
f(Y_{ij} \mid X_i, X_j, \beta) &= \left[\Phi(W_{ij}'\beta)^2\right]^{Y_{ij}} \left[1 - \Phi(W_{ij}'\beta)^2\right]^{1-Y_{ij}} \\
&\neq \left[\Phi(W_{ij}'\tilde{\beta})^2\right]^{Y_{ij}} \left[1 - \Phi(W_{ij}'\tilde{\beta})^2\right]^{1-Y_{ij}} \\
&= f(Y_{ij} \mid X_i, X_j, \tilde{\beta}),
\end{aligned}
$$

这是识别所要求的。

2.1.2 一致性

在 $\rho = 0$、 $W_{ij} = W_{ji}$ 下识别已得到保证，且 $\mathbb{E}[W_{ij}W_{ij}']$ 存在并满足非奇异性之后，我们现在转向一致性问题。我们将继续依赖 NM 的结果；具体而言，我们将依赖其关于凹对数似然函数的一致性定理。[9]一致性要求：(2.7.i) $Q_0(\beta) := \mathbb{E}[$对数 $f(Y_{ij} \mid X_i, X_j, \beta)]$ 在 β 的真值处取得唯一最大值， β_0；(2.7.ii) β_0 属于某个凸集的内部；(2.7.iii) $\hat{Q}_n(\beta) := \frac{1}{N}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left[Y_{ij}$对数 $\left(\Phi(W_{ij}'\beta)^2\right) + (1 - Y_{ij})$ 对数 $\left(1 - \Phi(W_{ij}'\beta)^2\right)\right]$ 是凹函数；且 (2.7.iv) 对所有 $\beta \in \Theta$， $\hat{Q}_n(\beta) \xrightarrow{p} Q_0(\beta)$。

引理1（及其证明，可见附录；除非另有说明，本文其他所有证明也均见附录）保证，在 $\mathbb{E}[W_{ij}W_{ij}']$ 存在且非奇异、 Θ有界的条件下， $Q_0(\beta)$在 β_0处取得唯一最大值。[10]

$\hat{}$ 难点在于证明对数似然函数 $Q_n(\beta)$的凹性；附录中的引理2解决了这一问题。众所周知，TU 模型的对数似然函数是凹的；而令人颇感意外的是，非转移效用模型的对数似然函数（其中 $\rho = 0$ 和 $W_{ij} = W_{ji}$，且 Φ 被 Φ^2取代）仍然是凹的。凹性对理论推导大有助益；我们将在第2.2节看到，当我们对所有配对 ij 放宽$W_{ij} = W_{ji}$ 时，对数似然函数将不再具有凹性，理论也会变得更加复杂。

最后，为了确立高阶收敛条件，$\hat{} Q_n(\beta) \xrightarrow{p} Q_0(\beta)$，我们将假设特征 X_i在各二元组之间独立同分布地抽取。[11] 然而，这并不意味着

9. 参见 Newey和McFadden（1994）中的定理2.7。10. 严格来说，在这种情况下，NM 并未正式要求参数空间有界。但有界性有助于证明；而且在第2.2节中，当回归变量不对称时，我们无论如何都会假设参数空间有界。11. Graham（2017）、Gao 等（2023）和 Li 等（2024）也采用了相同的假设。

Y_{ij} 在个体之间相互独立： Y_{12} 和 Y_{13} 都取决于 X_1，因此显然ˆ相关。由于 $Q_n(\beta)$包含 Y_{ij}，而 Y_{ij} 并非独立同分布，我们不能简单地调用通常的弱大数定律来证明ˆ $Q_n(\beta) \xrightarrow{p} Q_0(\beta)$。相反，我们需要使用一种能够考虑 Y_{12} 与 Y_{13}之间依赖关系的大数定律，例如。[12] 我们可以通过认识到 $Q_n(\beta)$是一个二阶U统计量，并随后调用Serfling（1980）中关于U统计量的强大数定律（依据Hoeffding（1961））来实现这一点。具体而言，Serfling（1980）中的定理5.4.A应用于此情形时指出，若 (5.4.A.i) X_i 是独立同分布的，(5.4.A.ii)$h_1(X_i, X_j) := Y_{ij}$对数 $\left(\Phi(W_{ij}'\beta)^2\right) + (1 - Y_{ij})$ 对数 $\left(1 - \Phi(W_{ij}'\beta)^2\right)$ 满足 $h_1(X_i, X_j)$ $=$ $h_1\ X_j\ X_i$ 5 4 $\mathbb{E}[\|h_1\ X_i\ X_j\ \|] < \infty$ $Q_n\ \beta \xrightarrow{a.s.} Q_0\ \beta$ 5 4 (,)，且(. .A.iii) (,)，则ˆ ( ) ( )。条件 (. .A.i) 已作假设；条件 (5.4.A.ii) 由此处 $W_{ij} = W_{ji}$ 这一事实得出；而在假设存在 $\mathbb{E}[W_{ij}W_{ij}']$ 且 Θ有界的情况下，条件(5.4.A.iii)已在引理1的证明中得到证明。因此，在与引理1相同的条件下，再加上X_i独立同分布，ˆ$Q_n(\beta) \xrightarrow{a.s.} Q_0(\beta)$，特别地，ˆ $Q_n(\beta) \xrightarrow{p} Q_0(\beta)$。

结合引理1和引理2，以及样本对数似然依概率收敛到总体对数似然这一结果，我们因此得到以下定理：

定理1 (对称回归变量下的一致性)。 考虑满足以下条件的非转移效用模型：对于所有配对 ij，模型具有 $\rho = 0$和 $W_{ij} = W_{ji}$ 。若 ***(1a)*** Θ 有界且为凸集，并且 $\beta_0 \in$ 属于内部(Θ),ˆ***(1b)*** X_i 独立同分布，且 ***(1c)*** $\mathbb{E}[W_{ij}W_{ij}']$ 存在且非奇异，则 ***NTU-MLE***，$\beta =$ˆ$\operatorname{argmax}_{\beta\in\Theta} Q_n(\beta)$，以趋近于*1*的概率存在，并且当 $n \to \infty$ 时，对 β_0 相合。

证明。 我们验证 NM 中定理2.7的条件。由引理1可知，在条件 (1a) 和 (1c) 下，条件 (2.7.i) 成立。条件 (2.7.ii) 由 (1a) 保证。条件 (2.7.iii) 由引理2得证。最后，如上所述，在条件 (1a)、(1b)、(1c) 以及引理1成立时，Serfling（1980）中的 U 统计量强大数定律保证条件 (2.7.iv) 成立。

□

有趣的是，NTU-MLE（一致性回归变量和独立误差）的一致性条件，与第1.1节中所述 TU-MLE 的一致性条件相同。

2.1.3 渐近正态性

在确立识别和一致性之后，我们现在证明 NTU-MLE 的渐近正态性。[13]回顾一下，有效样本量是排除自连接后的唯一二元组数量，即$N = \binom{n}{2} = {}^{n}(n-1)$ 2。我们将在以下条件下得到 $\sqrt{N}$-渐近正态性，其含义为 $\sqrt{N}(\hat{}\beta - \beta_0) \xrightarrow{d} \mathcal{N}(0, J^{-1})$，当 $n \to \infty$： $z_{ij} = (Y_{ij}, X_i, X_j)$ 独立同分布，MLE β 一致，(3.3.i) $\beta_0 \in$ interior(Θ)，(3.3.ii) $f(Y_{ij} \mid X_i, X_j, \beta)$ 二次连续可微，且 $f(Y_{ij} \mid X_i, X_j, \beta)$ 在 β_0 周围的邻域 $\mathcal{N}_0$ 内 > 0，(3.3.iii)$\int \sup_{\beta\in\mathcal{N}_0} \|\nabla_\beta f(Y_{ij} \mid X_i, X_j, \beta)\| dz < \infty$ 且 $\int \sup_{\beta\in\mathcal{N}_0} \|\nabla_{\beta\beta} f(Y_{ij} \mid X_i, X_j, \beta)\| dz < \infty$，

12. 请注意，这是一种相当弱的依赖关系（但不要将其与弱依赖的正式定义混淆），因为大多数二元组对不包含共同个体。具体而言，不同二元组的无序二元组共有 ${}^{N}(N-1)$ 2 个，其中有 ${}^{n}(n-1)(n-2)$ 2 个二元组对具有共同个体。因此，在 $N = {}^{n}(n=1)$ 2 的情况下，具有共同个体的二元组对所占比例等于 ${}^{4}(n-2)$ $n(n-1)-2$，当 $n \to \infty$ 时该比例趋于零。13. 使用 NM 中的定理 3.3。

(3.3.iv) $J := \mathbb{E}[\nabla_\beta \log f(Y_{ij} \mid X_i, X_j, \beta_0) \nabla_\beta \log f(Y_{ij} \mid X_i, X_j, \beta_0)']$ 存在且非奇异，并且 (3.3.v) $\mathbb{E}[\sup_{\beta \in \mathcal{N}_0} \|\nabla_{\beta\beta} \log f(Y_{ij} \mid X_i, X_j, \beta)\|] < \infty$。

正如上一节讨论一致性时所述，我们不能假设数据（Y_{ij}，X_i，X_j）是独立同分布的，只能假设个体特征 X_i 是独立同分布的。在 NM 中，独立同分布假设在推导过程中被用于应用林德伯格-列维中心极限定理，从而证明 $\sqrt{N}\nabla_\beta \hat{Q}_n(\beta_0)$ 渐近正态。尽管 Y_{ij} 存在依赖关系，我们仍然可以证明$\sqrt{N}\nabla_\beta \hat{Q}_n(\beta_0)$渐近正态：只需注意到，在给定个体特征的条件下，$Y_{ij}$ 相互独立（因为误差在不同二元组之间相互独立），然后应用条件版本的林德伯格-费勒中心极限定理即可。[14] 通过附录中的引理 3 可以做到这一点，不过此处需要假设 $\mathbb{E}[\|W_{ij}\|^8] < \infty$，而不仅仅是为证明一致性所假设的 $\mathbb{E}[W_{ij}W_{ij}'] < \infty$。有限的八阶矩看起来可能是一个较强的假设，但我们试图控制的是网络环境中产生的一种复杂依赖关系。事实上，Li 等（2024）假设 X_i 的支撑集有界，因此所有阶矩都有限，而不只是八阶矩。此外，TU 模型中也存在相同的依赖结构，因此要证明 TU-MLE 的渐近正态性，也需要假设有限的八阶矩。

根据引理 3 和 NTU-MLE 的一致性，我们验证条件 (3.3.i)-(3.3.v)，从而得到以下渐近正态性定理：[15]

定理2 (具有对称回归变量的渐近正态性)。考虑满足以下条件的非转移效用模型：对所有配对 ij，$\rho = 0$ 和 $W_{ij} = W_{ji}$。如果 *(2a)* Θ 有界且为凸，并且 $\beta_0 \in$位于内部(Θ)，*(2b)* X_i 独立同分布，*(2c)* $\mathbb{E}[\|W_{ij}\|^8]$ 存在，且 *(2d)* $\mathbb{E}[W_{ij}W_{ij}']$ 非奇异，那么$\hat{}$ $\hat{}$*NTU-MLE*，$\beta = \operatorname{argmax}_{\beta \in \Theta} Q_n(\beta)$在以下意义下渐近正态：

$$\sqrt{N}(\hat{\beta} - \beta_0) \xrightarrow{d} \mathcal{N}(0, J_1^{-1}) \quad as \quad n \to \infty,$$

其中

$$J_1 = \mathbb{E}\left[\frac{4\phi(W_{ij}'\beta)^2}{1 - \Phi(W_{ij}'\beta)^2} W_{ij}W_{ij}'\right].$$

证明。见附录。

同样，NTU-MLE 渐近正态的充分条件（具有对称回归变量和独立误差）与 TU-MLE 渐近正态的充分条件相同。这意味着，如果 TU-MLE 在理论上成立，那么 NTU-MLE 也成立（具有对称回归变量和独立误差）。

2.2 含非对称回归变量的情形

现在允许存在这样的可能性：对于某些配对，$W_{ij} \neq W_{ji}$ ij，同时仍假设 $\rho = 0$。在这种情况下，我们可以尝试沿用上述步骤来确立识别、一致性和渐近正态性，但很快就会遇到困难。

14. 与一致性分析不同，我们不能依赖 U统计量的极限理论，因为 U统计量 $\nabla_\beta \hat{Q}_n(\beta_0)$ 是退化的。
15. 其中，Φ 是标准正态分布的累积分布函数，ϕ 是其概率密度函数。

2.2.1 识别

首先面临的问题是识别。这里的识别意味着存在一对~ ~个体 ij，使得对于 β, $\beta \in \Theta$，如果 $\beta \neq \beta$，那么：

$$\begin{aligned} f(Y_{ij} \mid X_i, X_j, \beta) &:= \left[\Phi(W_{ij}'\beta)\Phi(W_{ji}'\beta)\right]^{Y_{ij}} \cdot \left[1 - \Phi(W_{ij}'\beta)\Phi(W_{ji}'\beta)\right]^{1-Y_{ij}} \\ &\neq \left[\Phi(W_{ij}'\tilde{\beta})\Phi(W_{ji}'\tilde{\beta})\right]^{Y_{ij}} \cdot \left[1 - \Phi(W_{ij}'\tilde{\beta})\Phi(W_{ji}'\tilde{\beta})\right]^{1-Y_{ij}} \\ &=: f(Y_{ij} \mid X_i, X_j, \tilde{\beta}). \end{aligned}$$

问题在于，是否存在具有对称回归变量的个体对。例如，假设个体在决定是否与他人建立连接时，只关注对方的教育程度。设 W_{ij} =教育程度$_j$ 和 W_{ji} =教育程度$_i$，其中教育程度是一个离散 随机变量，取四个值：至少为高等教育、至少为中等教育、至少为初等教育，以及低于初等教育。假设有两个个体 i 和j 至少接受过高等教育，例如，对于这对个体， W_{ij} = 教育程度$_j$ =教育程度$_i$ = W_{ji}。那么，根据第2.1.1节可知，假设 $\mathbb{E}[W_{ij}W_{ij}']$存在且非奇异，则 β 可识别。

当不存在具有对称回归变量的个体对时，即 $W_{ij} \neq W_{ji}\ \forall ij$，识别就会变得困难，因此我们无法依赖第2.1.1节的识别结果。这种情形的一个例子是，若只有一个回归变量，且该变量等于个体对中另一方的收入，即 W_{ij} = 收入$_j$ 和 W_{ji} = 收入$_i$，其中收入是一个连续 随机变量，因此样本中的每个人都可能具有不同的取值。我们可以证明识别性，前提是能够证明存在~ ~一对个体 ij，使得对于 β, $\beta \in \Theta$，若 $\beta \neq \beta$，则：

$$\Phi(W_{ij}'\beta)\Phi(W_{ji}'\beta) \neq \Phi(W_{ij}'\tilde{\beta})\Phi(W_{ji}'\tilde{\beta}).$$

问题在于，即使我们假设 $\mathbb{E}[W_{ij}W_{ij}']$、 $\mathbb{E}[W_{ji}W_{ji}']$ 和 $\mathbb{E}[W_{ij}W_{ji}']$ 存在且非奇异（因此为正定），使得：

$$W_{ij}'\beta \neq W_{ji}'\beta \neq W_{ij}'\tilde{\beta} \neq W_{ji}'\tilde{\beta},$$

我们仍然无法识别，因为仍有可能出现：

$$\Phi(W_{ij}'\beta)\Phi(W_{ji}'\beta) = \Phi(W_{ij}'\tilde{\beta})\Phi(W_{ji}'\tilde{\beta}).$$

我通过以下示例说明缺乏识别性。不失一般性，将 $W_{ij}'\beta = \beta_0 + \beta_1 W_{1ij} + ... + \beta_{k-1}W_{k-1ij}$ 和 $W_{ji}'\beta = \beta_0 + \beta_1 W_{1ji} + ... + \beta_{k-1}W_{k-1ji}$写作，其中 W_{lij} 是 W_{ij} 的第 l 个分量，而 W_{lji} 是 W_{ji}的第 l 个分量。由于$W_{ij} \neq W_{ji}$我们不可能同时有 $W_{1ij} = W_{1ji}$，……， $W_{k-1ij} = W_{k-1ji}$。

例1： $\beta \in \mathbb{R}^2$,和 $W_{1ij}, W_{1ji} \in \{0, 1\}$ 与 $W_{1ij} \neq W_{1ji}$。那么，识别~要求，如果 $\beta \neq \beta$:

$$\Phi(\beta_0 + \beta_1 W_{1ij})\Phi(\beta_0 + \beta_1 W_{1ji}) \neq \Phi(\tilde{\beta}_0 + \tilde{\beta}_1 W_{1ij})\Phi(\tilde{\beta}_0 + \tilde{\beta}_1 W_{1ji}).$$

不失一般性，[16] 当 $W_{1ij} = 0$ 时，因此 $W_{1ji} = 1$，我们要求若 $\beta \neq \tilde{\beta}$:

16. 考虑 $W_{1ij} = 1$ 和 $W_{1ji} = 0$ 则会得到相同的条件。

$$\Phi(\beta_0)\Phi(\beta_0+\beta_1) \neq \Phi(\tilde{\beta}_0)\Phi(\tilde{\beta}_0+\tilde{\beta}_1).$$

但是 $\beta_0 = 0,\ \beta_1 = 1,\ \beta_0 = 1,\ \beta_1 = -1$ $\beta \neq \beta$ $\Phi\ \beta_0\ \Phi\ \beta_0+\beta_1\ = \Phi\ \beta_0\ \Phi\ \beta_0+$ </style>满足这样的条件：尽管 ( ) ( ) ( ) (β_1)，但在这种情况下仍违反了识别。此外，由于回归变量是二元的，而且所有个体对都有非对称回归变量，因此我们没有其他可用于确立识别的条件。因此，我们遵循 Poirier（1980）的建议，在以下示例中允许回归变量取第三个值，以确立识别。

例2： $\beta \in \mathbb{R}^2$ 和 $W_{1ij},\ W_{1ji} \in \{0,1,2\}$ 与 $W_{1ij} \neq W_{1ji}$。与例1一样，我们使用回归变量配置 ($W_{1ij},\ W_{1ji}$) = (0，1) 推导出限制条件 $\Phi(\beta_0)\Phi(\beta_0+\beta_1) \neq \Phi\ \beta_0\ \Phi\ \beta_0+\beta_1$ $W_{1ij}\ W_{1ji}\ =\ 0,2$ $\Phi\ \beta_0\ \Phi\ \beta_0+$ ( ) ( )，我们也可以使用 (，) ( ) 推导出 ( ) ($2\beta_1$) $\neq \Phi\ \beta_0\ \Phi\ \beta_0+2\beta_1$ $W_{1ij}\ W_{1ji}\ =\ 1,2$ $\Phi\ \beta_0+\beta_1\ \Phi\ \beta_0+2\beta_1\ \neq$ ( ) ( )，并使用 (，) ( ) 推导出 ( ) ( ) $\Phi(\beta_0+\beta_1)\Phi(\beta_0+2\beta_1)$。接着，利用上述识别定义的逆否命题，我们可以通过证明以下条件成立来确立识别：

1. $\Phi(\beta_0)\Phi(\beta_0+\beta_1) = \Phi(\tilde{\beta}_0)\Phi(\tilde{\beta}_0+\tilde{\beta}_1)$
2. $\Phi(\beta_0)\Phi(\beta_0+2\beta_1) = \Phi(\tilde{\beta}_0)\Phi(\tilde{\beta}_0+2\tilde{\beta}_1)$
3. $\Phi(\beta_0+\beta_1)\Phi(\beta_0+2\beta_1) = \Phi(\tilde{\beta}_0+\tilde{\beta}_1)\Phi(\tilde{\beta}_0+2\tilde{\beta}_1),$

then $\beta_0 = \beta_0$ 和 $\beta_1 = \beta_1$。首先，将方程 1 除以方程 2，得到：

$$\frac{\Phi(\beta_0+\beta_1)}{\Phi(\beta_0+2\beta_1)} = \frac{\Phi(\tilde{\beta}_0+\tilde{\beta}_1)}{\Phi(\tilde{\beta}_0+2\tilde{\beta}_1)}.$$

然后，在方程 3 中代入 $\Phi(\beta_0+\beta_1)$，得到：

$$\Phi(\beta_0+2\beta_1)^2 = \Phi(\tilde{\beta}_0+2\tilde{\beta}_1)^2.$$

然后，根据 Φ 的单调性，我们得到条件 $\beta_0+2\beta_1 = \beta_0+2\beta_1$。将其代入上面的比值，并约去 $\Phi(\beta_0+2\beta_1)$ 项，得到第二个条件 $\beta_0+\beta_1 = \beta_0+\beta_1$。这两个条件合起来意味着 $\beta_0 = \beta_0$ 和 $\beta_1 = \beta_1$，这正是识别所要求的。

Poirier（1980）推测，识别的关键在于允许回归变量取值的数量多于未知参数的数量，但他并未为这一观点提供证明。我形式化地证明，在非转移效用模型中，若 $\rho = 0$ 且非对称回归变量和对称回归变量的数量一般，那么只要非对称回归变量至少有 3 个不同取值，β 就是可识别的。[17] 当然，我们还需要回归变量之间不存在完全多重共线性，而 $\mathbb{E}[W_{ij}W'_{ij}]$ 和 $\mathbb{E}[W_{ji}W'_{ji}]$的非奇异性保证了这一点。以下命题对此进行了概括：

17. 值得注意的是，与 Poirier（1980）相比，这里的识别通常更容易确立。这是因为 Poirier（1980）考虑的是偏好不对称，即非转移效用模型中两个指标的 beta 不同，而不是特征不对称。由于特征取决于 i 和 j（与 β 不同），我们可以利用具有对称特征的个体来确立识别。只有在 $W_{ij} \neq W_{ji}\ \forall ij$ 的完全非对称情形下，我们才不得不依赖“Poirier 原则”，允许回归变量取更多值。

命题1. 考虑具有 $\rho = 0$、 $W'_{ij}\beta = \beta_0 + \beta_1 W_{1ij} + ... + \beta_{k-1} W_{k-1ij}$、
$W'_{ji}\beta = \beta_0 + \beta_1 W_{1ji} + ... + \beta_{k-1} W_{k-1ji}$，且按 $W_{ij} \neq W_{ji}\ \forall ij$意义存在非对称性的非转移效用模型。令$p$ 为非对称回归变量的数量，即满足 $W_{sij} \neq W_{sji}\ \forall ij$的回归变量数量。那么，如果所有 p非对称回归变量都至少有 *3* 个不同取值，且$\mathbb{E}[W_{ij}W'_{ij}]$和 $\mathbb{E}[W_{ji}W'_{ji}]$均存在且非奇异，则 β可识别。

证明。 见附录。

再次考虑收入示例，其中 W_{ij} =收入$_j$、 W_{ji} = 收入$_i$，且收入是一个连续 随机变量，因此可能出现 $W_{ij} \neq W_{ji}\ \forall ij$。由于非对称回归变量是连续的，因而至少有 3 个不同取值，根据命题

1 β $\mathbb{E}[W_{ij}W'_{ij}] = \mathbb{E}[$ $^2_j]$ $\mathbb{E}[W_{ji}W'_{ji}] =$，我 $\mathbb{E}[$ $^2_i]$
们知道，只要收入和收入存在且具有非奇异性，即收入变量的变异有限且非零，该模型就是可识别的。在实践中，写下模型后，只需确定是否存在非对称回归变量；如果存在，则需确定这些回归变量是否至少有 3 个不同取值，或者是否有某些个体的这些回归变量是对称的；我会在应用部分提供进一步指导。假设 $\mathbb{E}[W_{ij}W'_{ij}]$ 和 $\mathbb{E}[W_{ji}W'_{ji}]$存在且具有非奇异性，那么 β 唯一不可识别的情形是：存在大量二元非对称回归变量，以至于找不到一对在所有这些回归变量上取值都相同的个体。随着收集到的个体样本规模更大、内容更丰富，这个问题便会得到缓解。

由于识别在这个意义上因情形而异，下面关于非对称回归变量的一致性和渐近正态性定理包含一个一般性的识别条件。

2.2.2 一致性

除识别问题外，将我们先前用于证明 NTU-MLE 一致性的策略应用于此时还面临第二个问题：对数似然函数不再是全局凹函数。因此，我们需要依赖一个更一般的一致性定理，该定理不要求对数似然函数具有凹性。[18] 具体而言，若满足以下条件，NTU-MLE 将是相合的：(2.1.i) $Q_0(\beta) := \mathbb{E}[\log\ f(Y_{ij} \mid X_i,\ X_j,\ \beta)]$ 在 β_0 处取得唯一最大值，(2.1.ii) Θ 是紧致的，(2.1.iii) $Q_0(\beta)$ 连续，且 (2.1.iv) $\hat{}\ Q_n(\beta) := \frac{1}{N}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\big[Y_{ij}\log\ \big(\Phi(W'_{ij}\beta)\Phi(W'_{ji}\beta)\big) + (1 - Y_{ij})\log\ \big(1 - \Phi(W'_{ij}\beta)\Phi(W'_{ji}\beta)\big)\big]$ 依概率一致收敛 到 $Q_0(\beta)$。与对称回归变量情形相比，关键区别在于，由于缺乏凹性，我们必须证明样本对数似然一致收敛 到总体对数似然，而不仅仅是逐点收敛。

正如引理1给出了在对称回归变量情形下使 β_0 成为 $Q_0(\beta)$ 唯一最大化点的条件，引理4给出了在非对称回归变量情形下使 β_0 成为 $Q_0(\beta)$ 唯一最大化点的条件。不出所料，除了假设在对称回归变量情形下我们所假设的 $\mathbb{E}[W_{ij}W'_{ij}] < \infty$ 之外，我们还假设在非对称回归变量情形下 $\mathbb{E}[W_{ji}W'_{ji}] < \infty$ 。

我们假设条件 (2.1.ii) 成立，即 Θ 是紧致的。如果数据是独立同分布的（$Y_{ij},\ X_i,\ X_j$），那么证明条件 (2.1.iii) 和 (2.1.iv) 就很容易；我们只需利用引理4的证明中关于 $\mathbb{E}[|\log\ f(Y_{ij} \mid X_i,\ X_j,\ \beta)|]$ 的界，来应用

18. 具体而言，我们依赖 NM 中的定理 2.1。

ˆ 一个标准的一致大数定律。由于 Y_{ij} 并非独立同分布，但 $Q_n(\beta)$ 又是一个二阶 U统计量，因此我们可以改用针对 U统计量的一致大数定律。具体而言，Nolan and Pollard（1987）中的定理7应用于此情形后指出，若满足以下条件，则 $Q_n(\beta)$ 将依概率一致收敛于 $Q_0(\beta)$：(7.i) X_i 是独立同分布的；(7.ii) 存在某个有限期望的控制函数 $d(X_i,\ X_j)$，使得 $|\log\ f(Y_{ij} \mid X_i,\ X_j,\ \beta)|$ 有界；(7.iii) Θ 是紧致且有限维的；(7.iv) $\log\ f(Y_{ij} \mid X_i,\ X_j,\ \beta)$ 关于 β 是全局利普希茨的，且具有某个可积包络$G(X_i,\ X_j)$。附录中的引理5表明，$\mathbb{E}[W_{ij}W'_{ij}]$和 $\mathbb{E}[W_{ji}W'_{ji}]$的存在性（加上独立同分布性和紧致性）足以满足 Nolan 和 Pollard（1987）中定理7的条件，因此也足以满足条件 (2.1.iv)。引理5还附带证明了条件 (2.1.iii)。

结合引理4和引理5，我们得到以下关于非对称回归变量的一致性定理：

定理 3 (非对称回归变量下的一致性)。 考虑满足$\rho = 0$ 的非转移效用模型，但现在允许某些配对 ij 满足 $W_{ij} \neq W_{ji}$ 。若 ***(3a)*** Θ 是紧致的，***(3b)*** X_i 是独立同分布的，***(3c)*** β 是可识别的，且 ***(3d)*** $\mathbb{E}[W_{ij}W'_{ij}]$ 和 $\mathbb{E}[W_{ji}W'_{ji}]$ 均存在，则 ***NTU-MLE***，ˆ $\hat{\beta} = \operatorname{argmax}_{\beta\in\Theta} Q_n(\beta)$，对于 β_0 是相合的，当 $n \to \infty$。

证明。 我们在 NM 中验证定理 2.1 的条件。由引理4可知，在 (3a)、(3c) 和 (3d) 成立时，条件 (2.1.i) 成立；条件 (2.1.ii) 与 (3a) 相同。由引理5可知，在 (3a)、(3b) 和 (3d) 成立时，条件 (2.1.iii) 和 (2.1.iv) 成立。 □

值得比较一下有回归变量对称性与无回归变量对称性时关于一致性的充分条件。既然在对称情形下我们愿意假设 $\mathbb{E}[W_{ij}W'_{ij}]$ 的存在性，那么在非对称情形下同时假设 $\mathbb{E}[W_{ji}W'_{ji}]$ 的存在性也是自然的。对数似然函数的非凹性意味着我们需要依赖一致大数定律，而不是逐点大数定律；但我们并不需要为此作出更强的假设，只是需要更复杂的渐近理论。非凹性还改变了我们对 Θ所作的假设：对于对称回归变量，假设 Θ 是有界且凸的；对于非对称回归变量，假设Θ 是紧致的。也就是说，当回归变量为非对称时，我们增加闭性要求，同时放宽凸性要求。如果希望无论回归变量是否具有对称性，都对一致性施加统一的参数空间假设，那么只需始终假设 Θ是紧致且凸的（并且 β_0 位于 Θ的内部）。

假定模型可识别。然而，如上所述，如果存在一对具有对称回归变量的个体，则可通过 $\mathbb{E}[W_{ij}W'_{ij}]$ 和 $\mathbb{E}[W_{ji}W'_{ji}]$ 的存在性与非奇异性建立识别；如果不存在这样的一对，但所有非对称回归变量都至少可以取 3 个不同取值，则可通过命题1建立识别。命题1还要求 $\mathbb{E}[W_{ij}W'_{ij}]$ 和 $\mathbb{E}[W_{ji}W'_{ji}]$ 具有存在性与非奇异性，因此，尽管我们在定理 3 中并未正式假定非奇异性，但在定理 3 之前，实际上几乎总是会假定这一条件，以建立识别。这与下文的渐近正态性定理（定理4）不同，后者要求非奇异性，原因并非识别。

归根结底，尽管在回归变量非对称时，要证明一致性所需的渐近理论更加复杂，但充分条件仅略微更强。

19. 严格来说，Nolan 和 Pollard（1987）中的定理7涉及一系列“熵”条件，但这些条件可由 (7.iii) 和 (7.iv) 推出。

从实际角度来说，主要区别在于，识别（作为一致性的前提）现在要求所有非对称回归变量至少有 3 个不同取值。这意味着我们只需避免包含大量二元回归变量的病态设定，因为在这种情况下，可能会找到一对在所有这类回归变量上的取值都不同的个体。

2.2.3 渐近正态性

对于非对称回归变量的渐近正态性，我们可以采用与对称回归变量相同的证明策略，但需要进行更多代数运算，并增加一个非奇异性条件。与对称回归变量情形一样，我们首先应用条件 Lindeberg-Feller 中心极限定理（CLT），以证明 $\sqrt{N}\nabla_\beta \hat{Q}_n(\beta_0)$:
$= \frac{1}{\sqrt{N}} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left[Y_{ij} \nabla_\beta \text{对数} \left(\Phi(W_{ij}'\beta_0)\overline{\Phi}(W_{ji}'\beta_0) \right) + (1 - Y_{ij}) \nabla_\beta \text{对数} \left(1 - \Phi(W_{ij}'\beta_0)\Phi(W_{ji}'\beta_0) \right) \right]$ 渐近正态且均值为零。再次假设八阶矩有限，由附录中的引理 6 可保证这一点。

鉴于一致性以及引理 6（因此在“幕后”存在依分布收敛），我们有以下定理:

定理4 (非对称回归变量下的渐近正态性**)**。 考虑具有 $\rho = 0$ 的非转移效用模型，但现在允许某些配对 ij 的 $W_{ij} \neq W_{ji}$ 。若 ***(4a)*** Θ 是紧致的，***(4b)*** X_i 是独立同分布的，***(4c)*** β 是可识别的，***(4d)*** $\mathbb{E}[\|W_{ij}\|^8]$ 和 $\mathbb{E}[\|W_{ji}\|^8]$ 均存在，且 ***(4e)*** $\mathbb{E}[W_{ij}W_{ij}']\hat{}\mathbb{E}[W_{ji}W_{ji}']$ 和 $\mathbb{E}[W_{ij}W_{ji}' + W_{ji}W_{ij}']$ 均非奇异，则 ***NTU-MLE*** $\beta = \arg\max^{\beta \in \Theta} \hat{Q}_n(\beta)$ 在以下意义下渐近正态:

$$\sqrt{N}(\hat{\beta} - \beta_0) \xrightarrow{d} \mathcal{N}(0, J_2^{-1}) \quad as \quad n \to \infty,$$

其中

$$\begin{aligned} J_2 = \mathbb{E}\Bigg[& \frac{\phi\big(W_{ij}'\beta\big)^2 \Phi\big(W_{ji}'\beta\big)}{\Phi\big(W_{ij}'\beta\big)\big(1 - \Phi\big(W_{ij}'\beta\big)\Phi\big(W_{ji}'\beta\big)\big)} W_{ij}W_{ij}' \\ & + \frac{\phi\big(W_{ji}'\beta\big)^2 \Phi\big(W_{ij}'\beta\big)}{\Phi\big(W_{ji}'\beta\big)\big(1 - \Phi\big(W_{ij}'\beta\big)\Phi\big(W_{ji}'\beta\big)\big)} W_{ji}W_{ji}' \\ & + \frac{\phi\big(W_{ij}'\beta\big)\phi\big(W_{ji}'\beta\big)}{1 - \Phi\big(W_{ij}'\beta\big)\Phi\big(W_{ji}'\beta\big)} \big(W_{ij}W_{ji}' + W_{ji}W_{ij}'\big) \Bigg]. \end{aligned}$$

证明。 见附录。

注意，我们无需假设$\mathbb{E}[W_{ij}W_{ji}' + W_{ji}W_{ij}']$ 存在，因为它可由$\mathbb{E}[W_{ij}W_{ij}']$ 和$\mathbb{E}[W_{ji}W_{ji}']$的存在性推出，而后者又可由 $\mathbb{E}[\|W_{ij}\|^8]$ 和 $\mathbb{E}[\|W_{ji}\|^8]$ 的存在性推出（利用柯西-施瓦茨不等式）。然而， $\mathbb{E}[W_{ij}W_{ij}']$ 和 $\mathbb{E}[W_{ji}W_{ji}']$ 非奇异并不能推出 $\mathbb{E}[W_{ij}W_{ji}' + W_{ji}W_{ij}']$非奇异。为说明这一点，假设$\mathbb{E}[W_{ij}W_{ij}']$ 和 $\mathbb{E}[W_{ji}W_{ji}']$ 非奇异，且 W_{ij}和 W_{ji}相互独立、均值为零。则 $\mathbb{E}[W_{ij}W_{ji}'] = \mathbb{E}[W_{ij}]\mathbb{E}[W_{ji}'] = 0$（交换 i 和 j亦然）。因此， $\mathbb{E}[W_{ij}W_{ji}' + W_{ji}W_{ij}'] = \mathbb{E}[W_{ij}W_{ji}'] + \mathbb{E}[W_{ji}W_{ij}'] = 0$是零矩阵，显然是奇异的。因此，我们通过假设（4e）假定 $\mathbb{E}[W_{ij}W_{ji}' + W_{ji}W_{ij}']$ 非奇异，以满足定理4证明中的条件（3.3.iv）。这一假设排除了如下情形:

与上述情形完全相同，即 W_{ij} 和 W_{ji} 相互独立且均值为零。有趣的是， W_{ij} 和 W_{ji} 之间存在一定程度的依赖关系，有助于确立渐近正态性。

与一致性一样，尽管在回归变量非对称时，证明渐近正态性所需的理论更为复杂，但充分条件仅略微加强。具体而言，除了满足一致性的条件外，我们现在还要求 W_{ij} 与 W_{ji} 之间存在非零依赖关系。例如，若像我们贯穿全文的示例那样， $W_{ij} =$ 收入$_j$ 和 $W_{ji} =$ 收入$_i$，则我们会要求个体之间的收入存在一定相关性；在收入受到共同冲击的情形下，这种相关性很自然。又如，若个体关注的是配对双方收入的加权平均值，但赋予自身收入更大的权重，比如 $W_{ij} = 0.7$收入$_i + 0.3$收入$_j$，那么 W_{ij} 与 W_{ji} 之间就会在结构上存在依赖关系。

2.3 其他模型扩展

除了将模型扩展到允许回归变量非对称之外，还可以允许 ε_{ij} 与 ε_{ji}之间存在任意协方差， $\rho \in (-1, 1)$。鉴于第1.2节所述的原因，我们仅针对 NTU-MLE 的一致性给出基于模拟的结果： β 和 $\hat{\rho}$ ——通用非转移效用对数似然 $Q_n(\beta, \rho)$ 的极大化者。在附录中，模拟结果表明，对于我们考虑的所有 $\rho \in (-1, 1)$ 取值，无论回归变量是否对称，NTU-MLE 都是相合的。此外， β 和 $\hat{\rho}$ 非常接近 β 和 ρ 的真值，且 $n = 20$非常小。

3 效用可转移性检验

在本节中，我们提出一种设定检验，用于判断给定数据集由 TU 模型还是 NTU 模型解释得更好。该检验十分有用，因为可以证明，在以下情形下，TU-MLE 将不具有相合性：a）回归变量不对称；或 b）回归变量并非围绕 0 对称分布。[20] 该检验基于这样一个事实：TU 模型和 NTU 模型都可以嵌套于一个更一般的模型中。具体而言，如果允许使用 $\rho \in [-1, 1]$ 而不是 $\rho \in (-1, 1)$，则这两个模型都可以嵌套于第1.2节中的一般 NTU 模型。该模型嵌套 NTU 模型这一点显而易见。要看出该模型也嵌套 TU 模型，只需令 $W_{ij} = W_{ji}$ 和 $\rho = 1$。如果 $\rho = 1$，那么 $\varepsilon_{ij} = \varepsilon_{ji}$，因此 $Y_{ij} = \mathbb{1}[W_{ij}'\beta \geq \varepsilon_{ij}]^2 = \mathbb{1}[W_{ij}'\beta \geq \varepsilon_{ij}]$，这就是 TU 模型。于是，我们可以通过以下检验，检验数据生成过程是 TU 模型而非 NTU 模型（具有对称回归变量）：

$$H_0 : \rho_0 = 1 \quad \text{vs} \quad H_A : \rho_0 < 1,$$

其中 ρ_0 是 ρ的真值。

令 $Q_n(\beta, \rho)$ 是第1.2节中一般 NTU 模型的对数似然函数， $\hat{\beta}, \hat{\rho}$ 现在允许 $\rho \in [-1, 1]$，并令 $(\beta, \hat{\rho}) := \arg\max_{(\beta,\rho)} Q_n(\beta, \rho)$。一种简单的检验方法是构造 Wald 统计量， $W = \frac{1-\hat{\rho}}{\mathrm{SE}(\hat{\rho})}$，并在显著性水平 α 下，当满足以下条件时拒绝 H_0

20. a) 是自然的，因为 TU 模型无法纳入非对称回归变量。b) 依据的是标准的模型设定错误下的最大似然估计量论证，遵循 White（1982）的做法。

$W > Z(1-\frac{\alpha}{2})$，其中 $Z(1-\frac{\alpha}{2})$是标准正态分布的第 $1-\frac{\alpha}{2}$ 分位数。另一种方法是构造似然比统计量，$LR = -2\left[\hat{Q}_n(\hat{\beta}, 1) - \hat{Q}_n(\hat{\beta}, \hat{\rho})\right]$，并在显著性水平 α 下，当满足 $LR > \chi_1^2(1-\alpha)$时拒绝 H_0，其中 $\chi_1^2(1-\alpha)$是自由度为 1 的卡方分布的第 $1-\alpha$分位数。[21] 然而，模拟证据表明，简单的 Wald 方法和简单的似然比方法都无法达到正确的检验尺寸：Wald 检验严重过度拒绝，而似然比检验则拒绝不足。

这或许并不令人意外，因为我们检验的是某个参数是否等于参数空间的边界：$\rho_0 = 1$。聚焦于似然比检验，Self 和 Liang（1987）推导了针对处于边界上的参数进行检验时，似然比检验统计量的渐近分布。将 β 视为干扰参数，[22] 我们有一个处于边界上的感兴趣参数（ρ），以及一个不处于边界上的干扰参数（β）。这意味着我们属于 Self 和 Liang（1987）的情形5，因此，在 H_0下，似然比检验统计量依分布收敛于 χ_0^2 和 χ_1^2 的50:50混合分布。也就是说，在显著性水平 α 下检验 H_0：$\rho_0 = 1$对比 H_A：$\rho_0 < 1$ 的正确做法是，当满足以下条件时拒绝 H_0：

$$-2\left[\hat{Q}_n(\hat{\beta}, 1) - \hat{Q}_n(\hat{\beta}, \hat{\rho})\right] > \left(\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2\right)(1-\alpha),$$

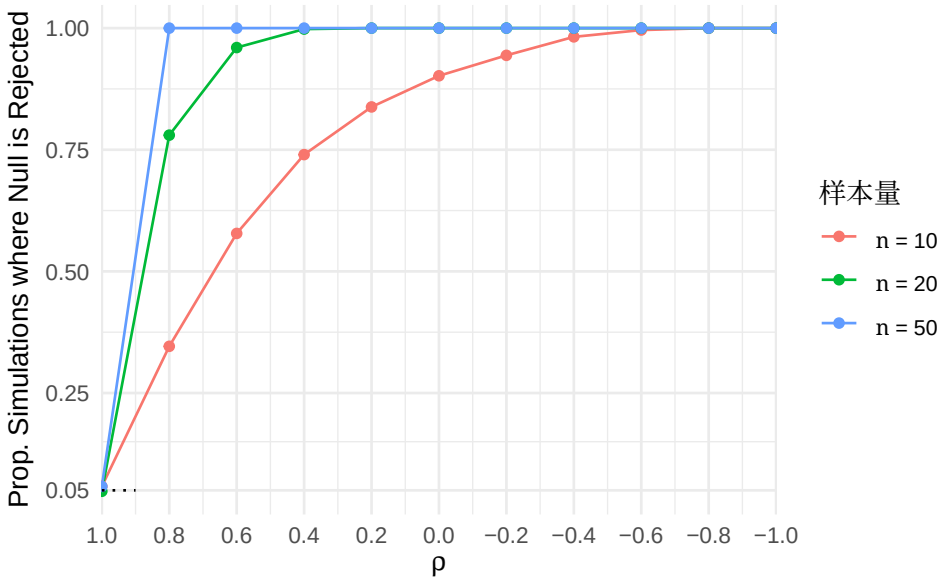
其中，$\left(\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2\right)(1-\alpha)$是由50%$\chi_0^2$ 和50% χ_1^2构成的混合分布的第（$1-\alpha$）个分位数。

我通过模拟验证这一检验程序的有效性。具体而言，我在图 2 中针对不同样本量绘制功效曲线，即将原假设 H_0：$\rho_0 = 1$被拒绝的模拟次数占比，绘制为生成网络时所用的 ρ 真值的函数。当$\rho = 1$ 时（如小虚线所示），可以通过查看拒绝比例读出经验检验尺寸。真值为 $\alpha = 0.05$；当 $n = 50$ 时，经验尺寸为 0.058；当 $n = 20$ 时，为 0.048；当 $n = 10$ 时，为 0.058。该检验不仅实现了尺寸控制，而且检验功效很强：当 $n = 50$ 时，功效在 $\rho = 0.8$ 处跃升至 100%。有限样本表现如此出色的原因在于，由于数据具有二元组性质，定理 2 和定理 4 中我们获得的是 $\sqrt{N}$-收敛，而非 $\sqrt{n}$-收敛。二元组的数量 $N = {}^{n}{}_{(n-1)\ 2}$ 与个体数量 n呈二次关系。因此，尽管 $n = 20$ 较小，$N = {}^{20}{}_{(20-1)\ 2} = 190$ 并不小。附录中 NTU-MLE 异常的小样本行为也印证了这一结果。

---

[21] 严格来说，似然比统计量中的 $\hat{Q}_n(\beta, 1)$ 通常应为 $Q_n(\beta_0, 1)$。然而，由于在本设定检验中，我们只关心检验 ρ_0 的取值，而不关心 β_0，并且 β 是相合的，因此在大样本下，$Q_n(\beta, 1)$ 与 $Q_n(\beta_0, 1)$ 无法区分；所以我纳入 $Q_n(\beta, 1)$，使得无论假设的 β_0 取值如何，都能计算 LR 统计量。如果要检验关于 ρ_0 和 $\beta_0 \in \mathbb{R}^k$ 的联合假设，只需将 $\hat{Q}_n(\hat{\beta}, 1)$ 替换为 $\hat{Q}_n(b, 1)$，其中 b 是在所关注的联合原假设下 β 的取值，并与自由度为 $k+1$ 的卡方分布进行比较。[22] 同样，如果要检验关于 ρ 和 $\beta \in \mathbb{R}^k$ 的联合假设，那么就属于 Self 和 Liang（1987）的情形6，相关的渐近分布将在 H_0 条件下为 $\frac{1}{2}\chi_k^2 + \frac{1}{2}\chi_{k+1}^2$。

图 2：设定检验功效曲线



该图展示了针对 H_0 的似然比检验的经验拒绝概率：$\rho_0 = 1$（当 $-2[\hat{}Q_n(\hat{}\beta, 1) - \hat{}Q_n(\hat{}\beta, \hat{}\rho)] > \left(\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2\right)(0.95)$ 时，以$\alpha = 0.05$为检验水平拒绝 H_0），并以 ρ 的真值为参照，展示了网络中个体数量 n 的三种不同取值。所有计算均基于一般 NTU 模型生成的 500 个网络，使用相应的 ρ 真值、来自标准正态分布的单个对称回归变量以及 $\beta = 1$。当 $\rho_0 = 1$ 时，经验尺寸非常接近名义尺寸 $\alpha = 0.05$；当 $n \geq 20$ 时，经验功效迅速超过 80%。

4 应用

4.1 AdolescentFriendships 加利福尼亚

现在，我将本文提出的估计和检验程序应用于来自经济学不同领域的两个网络数据集。第一个数据集来自 Goeree 等（2010）。Goeree 等（2010）利用独裁者博弈表明，加利福尼亚州帕萨迪纳一所女子学校中，10 至 18 岁的学生更可能给朋友赠送物品。此外，学生更可能与和自己更相似的人成为朋友（同质性）。在表 2 中，Goeree 等（2010）通过最大似然估计量（MLE）对带有逻辑斯蒂误差的 TU 模型进行估计，以确定哪些共同特征能够解释两名学生建立连接（友谊）的决策。具体而言，他们考察两名学生是否 a）种族相同，b）身高相同，c）自信程度相同，d）有过同一个男朋友；此外，还考察 e）这对学生中在独裁者博弈里作为接收者的学生是否害羞，以及 f）另一名学生的身高。在下文表 1的第 1 列中，我通过最大似然估计量（MLE）估计其带有逻辑斯蒂误差的 TU 模型，证实了 Goeree 等（2010）的发现：除接收者的害羞程度和另一名学生的身高外，所有因素都与连边决策显著相关。表1第 2 列所示的正态误差情形也得出了相同结论。

然而，在这里假设效用可转移是值得商榷的。如果 Abbie 想和 Becca 做朋友，但 Becca 不想和 Abbie 做朋友，那么假设 Abbie 会向 Becca 提供某种实物（例如金钱）转移，使这段友谊对双方都有利，这合理吗？这似乎并不是青少年友谊运作方式的一个很可信的模型。更正式地说，我们可以利用第3节提出的设定检验，检验 TU 模型在此处是否适用。采用第3节的似然比设定检验程序得到的 ρ的 NTU-MLE，以及对 $H_0 : \rho_0 = 1$ 与 $H_A : \rho_0 < 1$ 进行检验的 p 值，报告于表1第3列底部。我们发现 $\hat{}\rho = -0.031$，且设定检验的 p 值几乎为 0。也就是说，我们拒绝 TU 模型适用这一假设，转而支持 NTU 模型，且其中的 ρ 实际上等于 0。令人欣慰的是，我们似乎处于 $\rho = 0$ 情形，因为第2节正是针对这种情形正式证明了 NTU-MLE 的一致性和渐近正态性。

表1第3列还报告了非转移效用模型中同质性参数的估计值。[23] 虽然 TU 估计结果表明，曾经拥有同一个男朋友对友谊形成具有影响，而伴侣的身高则没有影响，但 NTU 估计结果却显示情况相反。本应用凸显出，当 TU假设的有效性可以被拒绝时，更为恰当的 NTU-MLE 可能会产生与 TU-MLE 在定性上不同的估计值，从而进一步凸显本文所提出的检验与估计程序的价值。

读者可能已经注意到，表1中的最后一个回归变量——另一名学生的身高——具有非对称性。例如，数据集中的第一个个体身高为54英寸，第二个个体身高为58英寸。尽管如此，β 在非转移效用模型中仍可识别，原因有二（假设 $\rho = 0$以及$\mathbb{E}[W_{ij}W'_{ij}]$ 和 $\mathbb{E}[W_{ji}W'_{ji}]$的存在性与非奇异性）。首先，有些个体对的身高相同，例如数据集中的第三和第四个个体身高都是56英寸。其次，即使没有两个个体身高相同，β仍可根据命题1识别，因为身高变量有24个（> 3）不同取值。正如定理3和定理4所示，假设参数空间紧致，并增加一个非奇异性条件，即可处理回归变量非对称性带来的复杂情况。[24]

---

23. 同质性参数的 p 值由针对 H_0 的标准沃尔德检验得出：$\beta_0 = 0$对 H_A ： $\beta_0 \neq 0$。关于标准沃尔德检验/似然比检验适用于涉及 β 的假设的依据，请参见附录中的模拟证据。例如，当 $\rho = 0$ 时（这里似乎正是这种情况），沃尔德检验的经验尺寸为 0.052，标准似然比检验的经验尺寸为 0.050。24. 严格来说，由于我们有一个非对称回归变量，假设回归变量为对称的第3节中所提出的设定检验并不适用于此处。不过，鉴于 $\hat{}\rho = 0$ 远离 1，且六个回归变量中只有一个是非对称的，因此经推广以适用于非对称回归变量的设定检验仍然很可能会拒绝 TU假设。关于如何推广设定检验的讨论，请参见第4.2节末尾。

表1：按个人特质解释青少年友谊决策 s

	结果变量：连接		
	(1)	(2)	(3)
	转移效用-对数几率	转移效用-概率单位	非转移效用-概率单位
同种族	0.430***	0.177***	0.138***
	(0.057)	(0.024)	(0.021)
相同身高	0.270***	0.110***	0.073***
	(0.057)	(0.023)	(0.018)
相同信仰	0.148**	0.061**	0.035*
	(0.057)	(0.024)	(0.018)
sameboyfriend	0.471*	0.193*	0.061
	(0.210)	(0.093)	(0.076)
害羞的接收者	-0.045	-0.019	-0.010
	(0.029)	(0.012)	(0.009)
伴侣身高	-0.006	-0.003	-0.018***
	(0.008)	(0.003)	(0.002)
$\hat{\rho}$	–	–	**-0.031**
$H_0 : \rho_0 = 1$ 的 p 值	–	–	0.000
观测值	55,666	55,666	55,666

本表第 1 列通过使用带有逻辑斯蒂误差的 TU 模型，解释加利福尼亚青少年的友谊决策，复现了 Goeree 等（2010）表 2 中的结果。第 2 列与第 1 列类似，但改为使用带有正态误差的 TU 模型；相同的统计显著因素仍然显著。第 3 列中，我估计了本文的一般 NTU 模型。sameboyfriend 不再具有统计显著性，而伴侣身高则变得显著。$\hat{\rho}$ 接近 0，并且根据第 3 节提出的似然比设定检验，我们拒绝 TU 模型的适用性。括号内为标准误。

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.2 坦桑尼亚农村地区的风险分担

第二个应用与两篇和本文最为相似的论文所研究的应用相同：Gao 等（2023）和 Li 等（2024）。该应用考察坦桑尼亚农村一个名为 Nyakatoke 的小型社区中，家庭之间风险分担的决定因素。[25]具体而言，Gao 等（2023）和 Li 等（2024）试图用以下因素解释两个家庭分享现金、实物或劳动力的决定：a）两个家庭的（对数）财富之差的绝对值；b）两个家庭之间（对数）实际距离；以及 c）两个家庭之间以血缘关系和共同

25. 感谢 De Weerdt（2002）收集了这些数据。

宗教。[26] 鉴于 Li 等（2024）提供了统计推断，而 Gao 等（2023）仅进行点估计，因此这里重点讨论 Li 等（2024）的结果。他们使用更复杂的基于 NTU 模型的估计量，发现财富差异的系数为零，距离的系数为负且显著，关系变量的系数为正且显著。

在表 2 中，我使用本文较简单的 NTU-MLE，并对 TU 模型与非转移效用模型进行设定检验。所得系数与 Li 等（2024）的结果一致：财富差异的系数为零，距离的系数为负且显著，关系变量的系数为正且显著。有趣的是，我们估计得到的 $\hat{\rho}$ 非常接近 1，因此无法拒绝原假设，即 TU 模型适合用于解释 Nyakatoke 的风险分担决策。这表明，在这里我们或许不必担心非转移效用模型所带来的额外识别和估计复杂性。由于 Gao 等（2023）和 Li 等（2024）都没有检验 TU 模型还是非转移效用模型更适用，因此他们对这个（表面上）基于 TU 模型的数据集使用了基于非转移效用模型的估计量；但实际上，他们或许可以在这里使用更简单的基于 TU 模型的估计量，例如 Graham（2017）中的估计量。

表 2：基于家庭特征解释风险分担决策

	结果变量：连接 **非转移效用 - probit**
|(对数)财富差异 |	0.026
	(0.030)
(对数)距离	-0.274***
	(0.007)
tie	0.403***
	(0.032)
$\hat{\rho}$	**1.000**
H_0 与 $\rho_0 = 1$ 的 p 值	1.000
观测值	7021

本表重新估计了 Li 等（2024）的表 6，但没有采用他们在非转移效用模型含固定效应的假设下所使用的较复杂的基于矩的估计量，而是采用本文提出的简单 NTU-MLE（使用通用的 ρ），并假设非转移效用模型不含固定效应。系数符号和统计显著性与 Li 等（2024）的结果相同。$\hat{\rho}$ 非常接近 1，这意味着我们无法拒绝 TU 模型的适用性。括号内为标准误。

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

尽管本文提出的设定检验在此处已证明颇有用处，但对表1和表2中的结果持怀疑态度的人仍可以指出，误差可能并非正态分布，而且可能存在未被考虑的未观测个体异质性，

26. 三个回归变量都是对称的。

。在半参数设定下进行带固定效应的设定检验会更有优势，但这超出了本文的研究范围。当然，这样一来，设定检验程序就不会像检验 $\rho = 1$ 那样简单。我推测，在更一般的设定下实施设定检验，最简单的方法可能是采用非嵌套似然比检验（Vuong，1989），类似于 Comola 和 Fafchamps（2014）所用的方法。[27] 此外，目前尚不清楚我们是否确实需要担心固定效应，尤其是考虑到我的估计（未考虑固定效应）与 Li 等（2024）的估计（明确考虑了固定效应）相近。这促使未来研究针对无固定效应的非转移效用模型与有固定效应的非转移效用模型开展设定检验。

5 结论

本文为二元网络形成模型文献作出了贡献，给出了有无对称回归变量时 NTU-MLE 一致性和渐近正态性的充分条件。结果表明，当回归变量对称时，TU-MLE 一致性和渐近正态性的充分条件同样足以保证 NTU-MLE 的一致性和渐近正态性。也就是说，如果 TU-MLE 在理论上成立，那么 NTU-MLE 也成立。此外，当回归变量非对称时，所需的额外假设少得惊人。尽管本文提供了相关误差情形下 NTU-MLE 一致性的基于模拟的证据（见附录），但尚未给出形式化条件，这显然是研究者下一步需要着力解决的问题。

此外，本文还提供了一种设定检验，用于确定在给定情形下 TU 模型还是非转移效用模型更为合适。两项应用验证了该设定检验的价值，但我谨慎地指出，该检验假设：a）回归变量对称；b）误差服从正态分布；c）不存在未观测个体异质性。我再次强调，在包含非对称回归变量和固定效应的半参数设定下，开展更一般的设定检验具有重要意义。

从应用角度来看，本文表明，TU-MLE 与 NTU-MLE 之间的选择很重要；然而，对于对称回归变量而言，NTU-MLE 并不比 TU-MLE 更难以论证。此外，相较于其他允许存在固定效应的基于 NTU 模型的估计量，使用标准最大似然程序计算 NTU-MLE 非常简单。鉴于其易于实施，且经验结论可能存在实质性差异，研究者应将 TU 与 NTU 之间的选择视为一个经验问题，而非默认假设。本文提出的设定检验则为这一经验问题提供了答案。

致谢

感谢 Francis DiTraglia 和 Martin Weidner 提供了富有见地的评论与指导。

27. Comola 和 Fafchamps（2014年）使用非嵌套似然比检验，以确定关系形成是单边进行的（即一个个体无需征得另一方同意，便可决定是否建立连接），还是如本文考察的所有模型那样通过双边协商进行。

参考文献

Candelaria, Luis E. 2020。具有未观测线性异质性的半参数网络形成模型。 arXiv:2007.05403 [econ]。Chandrasekhar, Arun G. 2016。“网络形成计量经济学。”载于《牛津网络经济学手册》，Yann Bramoull´e、Andrea Galeotti 和 Brian W. Rogers 编，0。牛津大学出版社。Christakis, Nicholas、James Fowler、Guido W. Imbens 和 Karthik Kalyanaraman。2020。“战略网络形成的经验模型。”载于网络数据的计量经济分析´，Bryan Graham 和 Aureo de Paula 编，123–148。学术出版社。Comola, Margherita 和 Marcel Fafchamps。2014。“单边与双边连边形成检验” [in en]。《经济学杂志》 124 (579): 954–976。De Weerdt, Joachim。2002。“风险共担与内生网络形成” [in en]。编号：DP2002-57，*WIDER*工作论文系列。Dzemski, Andreas。2019。“具有未观测异质性的网络二元关系形成经验模型。”《经济学与统计学评论》101 (5): 763–776。Gao, Wayne Yuan、Ming Li 和 Sheng Xu。2023。“不可转移效用下二元网络形成模型中的逻辑差分。”《计量经济学杂志》235 (1): 302–324。Goeree, Jacob K.、Margaret A. McConnell、Tiffany Mitchell、Tracey Tromp 和 Leeat Yariv。2010。“1/d捐赠定律” [in en]。《美国经济学杂志：微观经济学》 2 (1): 183–203。Goldsmith-Pinkham, Paul 和 Guido W. Imbens。2013。“社会网络与同伴效应识别。”《商业与经济统计杂志》31 (3): 253–264。Gordon, Robert D. 1941。“大参数值下米尔斯比率（面积与边界纵坐标之比）和正态概率积分的取值。”《数理统计年刊》12 (3): 364–366。Graham, Bryan S. 2017。“具有度异质性的网络形成计量经济模型” [in en]。《计量经济学》 85 (4): 1033–1063。Hoeffding, Wassily。1961年。*U*统计量的强大数定律。 [in en]。技术报告。北卡罗来纳州立大学统计系。Jackson, Matthew O. 2008年。社会与经济网络 [in eng]。新泽西州普林斯顿：普林斯顿大学出版社。Jochmans, Koen。2018年。“网络形成的半参数分析。”《商业与经济统计杂志》36 (4): 705–713。Le Cam, L. M. 1969年。统 ´计决策的渐近理 ´论。 蒙特利尔：蒙特利尔大学出版社。Li, Ming、Shi Zhentao 和 Yapeng Zheng。2024年。不可转移效用下二元网络形成模型的估计与推断。

Newey, Whitney K. 和 Daniel McFadden。1994年。“第36章：大样本估计与假设检验。”载于《计量经济学手册》,第4卷：2111–2245。爱思唯尔。Neyman, J. 和 Elizabeth L. Scott。1948年。“基于部分一致观测的一致估计量。”《计量经济学》16（1）：1–32。Nolan, Deborah 和 David Pollard。1987年。“U过程：收敛速率。”《统计年鉴》15（2）：780–799。Paula, ´ Aureo de。2020年。“网络形成的计量经济学模型” [in en]。《经济学年度评论》12（第12卷，2020年）：775–799。Poirier, Dale J. 1980年。“双变量 probit 模型中的部分可观测性。”《计量经济学杂志》12（2）：209–217。Self, Steven G. 和 Kung-Yee Liang。1987年。“非标准条件下最大似然估计量与似然比检验的渐近性质。”《美国统计学会会刊》82（398）：605–610。Serfling, Robert J. 1980年。数理统计逼近定理 [in eng]。威利概率与数理统计丛书。纽约；威利。Sheng, Shuyang。2020年。“通过子网络对网络形成博弈进行结构计量经济学分析” [in en]。《计量经济学》88（5）：1829–1858。Toth, Peter。2017年。具有同质性和度异质性的网络形成模型的半参数估计 [in en]。SSRN 学术论文。纽约州罗切斯特。Vaart, A. W. van der。1998年。渐近统计学。剑桥统计与概率数学丛书。剑桥：剑桥大学出版社。Vuong, Quang H. 1989年。“用于模型选择和非嵌套假设的似然比检验。”《计量经济学》57（2）：307–333。White, Halbert。1982年。“设定错误模型的最大似然估计。”《计量经济学》50（1）：1–25。

A 引理与证明

引理1。考虑对所有个体对ij均满足 $\rho = 0$ 和 $W_{ij} = W_{ji}$ 的非转移效用模型。如果 Θ有界，且 $\mathbb{E}[W_{ij}W_{ij}']$存在并且非奇异，那么 $Q_0(\beta)$: $=$ $\mathbb{E}$[对数 $f(Y_{ij} \mid X_i, X_j, \beta)$] 在 β_0处有唯一最大值。

引理1的证明。 根据 $\rho = 0$、 $W_{ij} = W_{ji}$下的识别，以及 $\mathbb{E}[W_{ij}W_{ij}']$ 的存在性和非奇异性（如论文正文所示），我们还需证明，对于所有 $\beta \in \Theta$， $\mathbb{E}[|\log f(Y_{ij} \mid X_i, X_j, \beta)|] < \infty$，以便调用 NM 中的引理2.2，得到所需结果。$\log f(Y_{ij} \mid X_i, X_j, \beta) = Y_{ij} \log\left(\Phi\left(W_{ij}'\beta\right)^2\right) + (1 - Y_{ij}) \log\left(1 - \Phi\left(W_{ij}'\beta\right)^2\right)$，因此分别考察 $\log\left(\Phi\left(W_{ij}'\beta\right)^2\right)$ 项和 $\log\left(1 - \Phi\left(W_{ij}'\beta\right)^2\right)$ 项:

项 *1*: 对数 $\left(\Phi\left(W_{ij}'\beta\right)^2\right) = 2$对数 $\left(\Phi\left(W_{ij}'\beta\right)\right)$

然后，在 $\beta = 0$ 处进行一阶泰勒展开，得到:

$$2\log\left(\Phi\left(W_{ij}'\beta\right)\right) = 2\log(\Phi(0)) + 2\frac{\phi(W_{ij}'\tilde{\beta})}{\Phi(W_{ij}'\tilde{\beta})}W_{ij}'\beta,$$

其中 $\tilde{\beta}$ 介于 β 和 0 之间。因此，根据三角不等式并使用 NM 中的界， $\phi(v)\Phi(v) \leq C(1 + |v|)$，其中 $C \geq 1$:

$$\begin{aligned}
|2\log\left(\Phi\left(W_{ij}'\beta\right)\right)| &\leq |2\log(\Phi(0))| + \left|2\frac{\phi(W_{ij}'\tilde{\beta})}{\Phi(W_{ij}'\tilde{\beta})}W_{ij}'\beta\right| \\
&= |2\log(\Phi(0))| + 2\frac{\phi(W_{ij}'\tilde{\beta}}{\Phi(W_{ij}'\tilde{\beta}}\left|W_{ij}'\beta\right| \\
&\leq |2\log(\Phi(0))| + 2C(1 + |W_{ij}'\tilde{\beta}|)\left|W_{ij}'\beta\right|.
\end{aligned}$$

项 *2*: 对数 $\left(1 - \Phi\left(W_{ij}'\beta\right)^2\right)$

然后，在 $\beta = 0$ 处进行一阶泰勒展开，得到:

$$\log\left(1 - \Phi\left(W_{ij}'\beta\right)^2\right) = \log(1 - \Phi(0)^2) - \frac{2\Phi(W_{ij}'\tilde{\beta})\phi(W_{ij}'\tilde{\beta})}{1 - \Phi(W_{ij}'\tilde{\beta})^2}W_{ij}'\beta.$$

为了界定该项的绝对值，我们利用以下事实:

$$\frac{2\Phi(v)\phi(v)}{1 - \Phi(v)^2} \leq \frac{\phi(-v)}{\Phi(-v)} \leq C(1 + |-v|) = C(1 + |v|),$$

鉴于 $\phi(-v) = \phi(v)$、 $\Phi(-v) = 1 - \Phi(v)$ 且 $\Phi(v) \leq 1$，很容易证明这一点。因此，利用这个界并再次使用三角不等式，我们得到:

$$\begin{aligned}
\left|\log\left(1 - \Phi\left(W_{ij}'\beta\right)^2\right)\right| &\leq \left|\log\left(1 - \Phi(0)^2\right)\right| + \left|\frac{2\Phi(W_{ij}'\tilde{\beta})\phi(W_{ij}'\tilde{\beta})}{1 - \Phi(W_{ij}'\tilde{\beta})^2}W_{ij}'\beta\right| \\
&= \left|\log\left(1 - \Phi(0)^2\right)\right| + \frac{2\Phi(W_{ij}'\tilde{\beta})\phi(W_{ij}'\tilde{\beta})}{1 - \Phi(W_{ij}'\tilde{\beta})^2}\left|W_{ij}'\beta\right| \\
&\leq \left|\log\left(1 - \Phi(0)^2\right)\right| + C(1 + |W_{ij}'\tilde{\beta}|)\left|W_{ij}'\beta\right|.
\end{aligned}$$

合并各项：

再次使用三角不等式以及 $Y_{ij} \in \{0,1\}$ 这一事实，得到：

$$\begin{aligned}
|\log f(Y_{ij} \mid X_i, X_j, \beta)| &\leq \left|2\log\left(\Phi\left(W_{ij}'\beta\right)\right)\right| + \left|\log\left(1-\Phi\left(W_{ij}'\beta\right)^2\right)\right| \\
&\leq |2\log(\Phi(0))| + \left|\log\left(1-\Phi(0)^2\right)\right| \\
&\quad + 3C(1+|W_{ij}'\tilde{\beta}|)\left|W_{ij}'\beta\right| \\
&\leq |2\log(\Phi(0))| + \left|\log\left(1-\Phi(0)^2\right)\right| \\
&\quad + 3C(1+\|W_{ij}\|\left\|\tilde{\beta}\right\|)\|W_{ij}\|\,\|\beta\|.
\end{aligned}$$

最后，对两边取期望，并注意到由于 Θ 有界，所有其他项均为有限值，我们便得到 $\mathbb{E}[|\log f(Y_{ij} \mid X_i, X_j, \beta)|] < \infty$ provided $\mathbb{E}[W_{ij}W_{ij}'] < \infty$，证毕。

□

引理2。 对于满足所有配对 ij 均有 $\rho = 0$ 和 $W_{ij} = W_{ji}$ 的非转移效用模型，其对数似然函数 $\hat{} \; Q_n(\beta) := \frac{1}{N}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left[Y_{ij}\log\left(\Phi(W_{ij}'\beta)^2\right) + (1-Y_{ij})\log\left(1-\Phi(W_{ij}'\beta)^2\right)\right]$是凹函数。

^引理2的证明。 由于 $v := W_{ij}'\beta$ 关于 β是线性的，若能证明 $f(v) := Y\log(\Phi(v)^2) + (1-Y)\log(1-\Phi(v)^2)$ 关于 v是凹的，则 $Q_n(\beta)$的凹性便可得出。此外，由于 Y是二元的，分别考虑 $Y=1$ 和 $Y=0$ 的情形：

情形 *1*： $Y = 1$因此 $f(v) = \log(\Phi(v)^2)$。于是 $\frac{df}{dv} = \frac{2\phi(v)}{\Phi(v)}$，而 $\frac{d^2f(v)}{dv^2} < 0$ 由一个众所周知的事实可得： $\frac{\phi(v)}{\Phi(v)}$ 是递减的（NM）。

Case 2: $Y = 0$

So $f(v) = \log(1-\Phi(v)^2)$.

Then $\frac{df(v)}{dv} = \frac{-2\Phi(v)\phi(v)}{1-\Phi(v)^2}$ and $\frac{d^2f(v)}{dv^2} = \frac{-2(1-\Phi(v)^2)(\Phi(v)\frac{d\phi(v)}{dv}+\phi(v)^2)-4\Phi(v)^2\phi(v)^2}{(1-\Phi(v)^2)^2}$.

Rearranging yields that $\frac{d^2f(v)}{dv^2} < 0$ provided $-\frac{d\phi(v)}{dv} < \frac{\phi(v)^2}{\Phi(v)}\frac{1+\Phi(v)^2}{1-\Phi(v)^2}$.

Then note that $\frac{d\phi(v)}{dv} = -v\phi(v)$. So WTS: $v < \frac{\phi(v)}{\Phi(v)}\frac{1+\Phi(v)^2}{1-\Phi(v)^2}$.

情形 *2a*： $v < 0$

那么 $v < \frac{\phi(v)}{\Phi(v)}\frac{1+\Phi(v)^2}{1-\Phi(v)^2}$显然成立，因为左侧 < 0 由假设可知，右侧 ≥ 0 因为 $\phi(v) \geq 0$， $\Phi(v) \geq 0$，且 $\Phi(v) \leq 1$。

情形 *2b*： $v \geq 0$然后，正如 Gordon (1941) 所示，逆米尔斯比率的上界为 $1/v$。即， $\frac{1-\Phi(v)}{\phi(v)} < \frac{1}{v}$。因此 $v < \frac{\phi(v)}{1-\Phi(v)} \leq \frac{\phi(v)}{\Phi(v)}\frac{1+\Phi(v)^2}{1-\Phi(v)^2}$，因为 $\Phi(v) \leq 1$，如所需。

□

引理 3. 考虑满足所有二元组 ij均有 $\rho = 0$ 和 $W_{ij} = W_{ji}$ 的非转移效用模型。如果 Θ 有界，X_i 独立同分布，且 $\mathbb{E}[\|W_{ij}\|^8]$ 存在，那么 $\sqrt{N}\nabla_\beta \hat{Q}_n(\beta_0) := \frac{1}{\sqrt{N}}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left[Y_{ij}\nabla_\beta$对数$(\Phi(W_{ij}'\beta_0)^2) + (1-Y_{ij})\nabla_\beta$对数$\left(1-\Phi(W_{ij}'\beta_0)^2\right)\right]$ 渐近正态，均值为零。

引理 3 的证明。 设 $\mathscr{X} = \{X_1, ..., X_n\}$ 为个体特征集合。在给定 $\mathscr{X}$ 的条件下，$h_2(X_i, X_j) := Y_{ij}\nabla_\beta$ 对数 $\left(\Phi(W_{ij}'\beta_0)^2\right) + (1 - Y_{ij})\nabla_\beta$ 对数 $\left(1 - \Phi(W_{ij}'\beta_0)^2\right)$ 在二元组之间相互独立，因为假设 ε_{ij} 在二元组之间相互独立。这意味着我们可以使用林德伯格-费勒中心极限定理来证明，在给定 $\mathscr{X}$ 的条件下，$\sqrt{N}\nabla_\beta \hat{Q}_n(\beta_0) = \frac{1}{\sqrt{N}} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} h_2(X_i, X_j)$ 渐近正态，进而可推出无条件渐近正态（见下文）。具体而言，Vaart（1998）中的命题 2.27（林德伯格-费勒中心极限定理）指出，给定 $\mathscr{X}$，若满足：

(2.27.i) $\quad \mathbb{V}\mathrm{ar}\left[\frac{1}{\sqrt{N}} h_2(X_i, X_j) \mid \mathscr{X}\right] < \infty$

(2.27.ii) $\quad \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathbb{E}\left[\left\|\frac{1}{\sqrt{N}} h_2(X_i, X_j)\right\|^2 \mathbb{1}\left\{\left\|\frac{1}{\sqrt{N}} h_2(X_i, X_j)\right\| > \varepsilon\right\} \mid \mathscr{X}\right] \to 0 \quad \forall \varepsilon > 0$

(2.27.iii) $\quad \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathbb{C}\mathrm{ov}\left[\frac{1}{\sqrt{N}} h_2(X_i, X_j) \mid \mathscr{X}\right] \to \Sigma_1,$

那么：

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(\frac{1}{\sqrt{N}} h_2(X_i, X_j) - \mathbb{E}\left[\frac{1}{\sqrt{N}} h_2(X_i, X_j) \mid \mathscr{X}\right]\right) \mid \mathscr{X} \xrightarrow{d} \mathcal{N}(0, \Sigma_1) \quad \text{as} \quad n \to \infty.$$

因此，由于 $\mathbb{E}\left[\frac{1}{\sqrt{N}} h_2(X_i,\ X_j) \mid \mathscr{X}\right] = 0$（在 β_0 处期望得分为零），若条件 (2.27.i)、(2.27.ii) 和 (2.27.iii) 成立，则在给定 $\mathscr{X}$ 的条件下，$\sqrt{N}\nabla_\beta \hat{Q}_n(\beta_0)$ 渐近正态，且均值为零。

若 $\mathbb{E}\left[\|h_2(X_i,\ X_j)\|^2\right] < \infty$，则条件 (2.27.i) 几乎必然成立。我们证明，在 $\mathbb{E}[\|W_{ij}\|^4] < \infty$ 且 Θ 有界的条件下，这一结论成立。令 $V_{ij} := W_{ij}'\beta_0$。那么：

$$\nabla_\beta \log\left(\Phi(V_{ij})^2\right) = \frac{2\phi(V_{ij})}{\Phi(V_{ij})} W_{ij},$$

and

$$\nabla_\beta \log\left(1 - \Phi(V_{ij})^2\right) = -\frac{2\Phi(V_{ij})\phi(V_{ij})}{1 - \Phi(V_{ij})^2} W_{ij}.$$

因此：

$$h_2(X_i, X_j) = \left[Y_{ij} \frac{2\phi(V_{ij})}{\Phi(V_{ij})} - (1 - Y_{ij}) \frac{2\Phi(V_{ij})\phi(V_{ij})}{1 - \Phi(V_{ij})^2}\right] W_{ij}.$$

两边取欧几里得范数的平方，得到：

$$\begin{aligned}
\|h_2(X_i, X_j)\|^2 &= \left[Y_{ij} \frac{2\phi(V_{ij})}{\Phi(V_{ij})} - (1 - Y_{ij}) \frac{2\Phi(V_{ij})\phi(V_{ij})}{1 - \Phi(V_{ij})^2}\right]^2 \|W_{ij}\|^2 \\
&= Y_{ij}^2 \left[\frac{2\phi(V_{ij})}{\Phi(V_{ij})}\right]^2 \|W_{ij}\|^2 + (1 - Y_{ij})^2 \left[\frac{2\Phi(V_{ij})\phi(V_{ij})}{1 - \Phi(V_{ij})^2}\right]^2 \|W_{ij}\|^2 \\
&\quad - 2Y_{ij}(1 - Y_{ij}) \frac{2\phi(V_{ij})}{\Phi(V_{ij})} \frac{2\Phi(V_{ij})\phi(V_{ij})}{1 - \Phi(V_{ij})^2} \|W_{ij}\|^2.
\end{aligned}$$

由于 $Y_{ij} \in \{0,1\}$，我们有 $Y_{ij}(1-Y_{ij}) = 0$，因此交叉项为零。此外，$Y_{ij}^2 \le 1$以及 $(1-Y_{ij})^2 \le 1$，所以：

$$\|h_2(X_i, X_j)\|^2 \le \left[\frac{2\phi(V_{ij})}{\Phi(V_{ij})}\right]^2 \|W_{ij}\|^2 + \left[\frac{2\Phi(V_{ij})\phi(V_{ij})}{1-\Phi(V_{ij})^2}\right]^2 \|W_{ij}\|^2.$$

如引理1的证明所示，$\frac{\phi(v)}{\Phi(v)}$ 和 $\frac{2\Phi(v)\phi(v)}{1-\Phi(v)^2}$ 都由上方的$C(1+|v|)$所界，因此它们的平方由某个常数 $D > 0$乘以 $D(1+|v|^2)$所界。因此：

$$\|h_2(X_i, X_j)\|^2 \le 2D(1+|V_{ij}|^2)\|W_{ij}\|^2.$$

由于 $V_{ij} = W_{ij}'\beta_0$，根据柯西–施瓦茨不等式：

$$|V_{ij}|^2 \le \|\beta_0\|^2 \|W_{ij}\|^2,$$

so

$$\|h_2(X_i, X_j)\|^2 \le 2D\|W_{ij}\|^2 + 2D\|\beta_0\|^2\|W_{ij}\|^4.$$

取期望可得：

$$\mathbb{E}\big[\|h_2(X_i, X_j)\|^2\big] \le 2D\,\mathbb{E}[\|W_{ij}\|^2] + 2D\|\beta_0\|^2\mathbb{E}[\|W_{ij}\|^4],$$

因此，$\mathbb{E}[\|h_2(X_i, X_j)\|^2] < \infty$，前提是 $\mathbb{E}[\|W_{ij}\|^4] < \infty$ 且 Θ 有界。

若满足 $\mathbb{E}[\|W_{ij}\|^8] < \infty$，条件 (2.27.ii)，即林德伯格条件，便成立。为说明这一点，先使用柯西-施瓦茨不等式，再使用马尔可夫不等式，对林德伯格项作如下界定：

$$\begin{aligned}
L_n(\varepsilon) &:= \frac{1}{N}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathbb{E}\left[\|h_2(X_i, X_j)\|^2 \mathbb{1}\left\{\left\|\frac{1}{\sqrt{N}}h_2(X_i, X_j)\right\| > \varepsilon\right\} \mid \mathscr{X}\right] \\
&\le \frac{1}{N}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathbb{E}\big[\|h_2(X_i, X_j)\|^4 \mid \mathscr{X}\big]^{1/2}\, \mathbb{P}\Big(\|h_2(X_i, X_j)\| > \varepsilon\sqrt{N} \mid \mathscr{X}\Big)^{1/2} \\
&\le \frac{1}{\varepsilon N^{3/2}}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathbb{E}\big[\|h_2(X_i, X_j)\|^4 \mid \mathscr{X}\big]^{1/2}\, \mathbb{E}\big[\|h_2(X_i, X_j)\|^2 \mid \mathscr{X}\big]^{1/2}.
\end{aligned}$$

因此，当 $\mathbb{E}[\|h_2(X_i, X_j)\|^4] < \infty$时，$L_n(\varepsilon) \to 0$。由于我们已在上文证明 $\|h_2(X_i, X_j)\|^2 \le 2D\|W_{ij}\|^2 + 2D\|\beta_0\|^2\|W_{ij}\|^4$，由此可得：

$$\mathbb{E}\big[\|h_2(X_i, X_j)\|^4\big] \le 4D^2\,\mathbb{E}[\|W_{ij}\|^4] + 4D^2\|\beta_0\|^2\mathbb{E}[\|W_{ij}\|^6] + 4D^2\|\beta_0\|^4\mathbb{E}[\|W_{ij}\|^8],$$

因此，如果 $\mathbb{E}$，林德伯格条件成立$[\|W_{ij}\|^8] < \infty$ 和 Θ 有界，符合要求。

通过认识到以下事实，可以证明条件 (2.27.iii)：

$$\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathbb{C}\mathrm{ov}\left[\frac{1}{\sqrt{N}}h_2(X_i, X_j) \mid \mathscr{X}\right] = \frac{1}{N}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathbb{E}\left[h_2(X_i, X_j)h_2(X_i, X_j)' \mid \mathscr{X}\right]$$

是一个（退化的）U统计量，然后再次应用 Serfling（1980）关于 U 统计量的强大数定律。具体而言，令 $H(X_i, X_j) := \mathbb{E}\left[h_2(X_i, X_j)h_2(X_i, X_j)' \mid \mathscr{X}\right]$，显然

(5.4.A.i) X_i 是独立同分布的（根据假设），(5.4.A.ii) $H(X_i,\ X_j) = H(X_j,\ X_i)$，因为 $h_2(X_i,\ X_j) = h_2(X_j,\ X_i)$，且 (5.4.A.iii) $\mathbb{E}[\|H(X_i,\ X_j)\|] < \infty$，因为：

$$\begin{aligned}\|H(X_i, X_j)\| &= \|\mathbb{E}\left[h_2(X_i, X_j)h_2(X_i, X_j)' \mid \mathscr{X}\right]\| \\ &\leq \mathbb{E}\left[\|h_2(X_i, X_j)h_2(X_i, X_j)'\| \mid \mathscr{X}\right] \\ &\leq \mathbb{E}\left[\|h_2(X_i, X_j)\|^2 \mid \mathscr{X}\right]\end{aligned}$$

根据詹森不等式。因此，根据迭代期望定律，$\mathbb{E}[\|H(X_i, X_j)\|] < \infty$ if$\mathbb{E}[\|h_2(X_i, X_j)\|^2] < \infty$，而我们知道，根据上文，当 $\mathbb{E}[\|W_{ij}\|^4] < \infty$ 时该式成立。

最后，我们注意到，在此处，条件分布收敛意味着无条件分布收敛。具体而言，由于 $\sqrt{N}\nabla_\beta \hat{Q}_n(\beta_0)$ 在给定 $\mathscr{X}$ 的条件下渐近正态，因此：

$$\mathbb{E}[f(\sqrt{N}\nabla_\beta \hat{Q}_n(\beta_0)) \mid \mathscr{X}] \to \mathbb{E}[f(Z) \mid \mathscr{X}]$$

对于任意有界连续函数 f，其中 Z 服从正态分布。因此，连续两次应用迭代期望定律可得：

$$\mathbb{E}[f(\sqrt{N}\nabla_\beta \hat{Q}_n(\beta_0))] = \mathbb{E}[\mathbb{E}[f(\sqrt{N}\nabla_\beta \hat{Q}_n(\beta_0)) \mid \mathscr{X}]] \to \mathbb{E}[\mathbb{E}[f(Z) \mid \mathscr{X}]] = \mathbb{E}[f(Z)],$$

因此，无条件地，$\sqrt{N}\nabla_\beta \hat{Q}_n(\beta_0)$ 渐近正态，证毕。

□

定理2的证明。我们验证 NM 中的定理 3.3 的条件。假设只有 X_i 是独立同分布的，而非 Y_{ij}；但根据引理 3，在（假设成立的）有限八阶矩条件下，$\sqrt{N}\nabla_\beta \hat{Q}_n(\beta_0)$ 仍将渐近正态。由于 (Y_{ij}, X_i, X_j) 独立同分布这一假设仅用于此处，因此我们可以用更弱的假设取而代之，即只要求 X_i 独立同分布。根据定理2的条件，定理1保证了一致性；注意，有限八阶矩意味着有限二阶矩。因此，接下来只需验证条件 (3.3.i) - (3.3.v)：

假设条件 (3.3.i) 成立。

条件 (3.3.ii)：将似然表示为以下形式最为简便：

$$f(Y_{ij} \mid X_i, X_j, \beta) = Y_{ij}\Phi\left(W_{ij}'\beta\right)^2 + (1 - Y_{ij})\left(1 - \Phi\left(W_{ij}'\beta\right)^2\right)$$

而不是

$$f(Y_{ij} \mid X_i, X_j, \beta) = \left[\Phi\left(W_{ij}'\beta\right)^2\right]^{Y_{ij}}\left[1 - \Phi\left(W_{ij}'\beta\right)^2\right]^{1-Y_{ij}},$$

通过分别考虑 $Y_{ij} = 1$ 和 $Y_{ij} = 0$ 的情况，可以证明这是一种等价的表述。接下来，我们计算似然的一阶导数：

$$\nabla_\beta f(Y_{ij} \mid X_i, X_j, \beta) = (4Y_{ij} - 2)\Phi(W_{ij}'\beta)\phi(W_{ij}'\beta)W_{ij},$$

然后计算似然的二阶导数：

$$\nabla_{\beta\beta} f(Y_{ij} \mid X_i, X_j, \beta) = (4Y_{ij} - 2)\left[\phi(W_{ij}'\beta)^2 + \Phi(W_{ij}'\beta)\,\phi_v(W_{ij}'\beta)\right]\,W_{ij}W_{ij}'.$$

由于 Φ，ϕ，和 ϕ_v 均连续，因此显然它是连续的。此外，它在 β_0 处非零，由此可得条件 (3.3ii) 成立。

条件 (3.3.iii)：我们已在上文计算了 $\nabla_\beta f(Y_{ij} \mid X_i, X_j, \beta)$ 和 $\nabla_{\beta\beta} f(Y_{ij} \mid X_i, X_j, \beta)$。由于 $Y_{ij} \in \{0, 1\}$，$0 < \Phi(W_{ij}'\beta) < 1, 0 < \phi(W_{ij}'\beta) < 0.4$，并利用厄米多项式$|\phi_v(W_{ij}'\beta)| < \phi(1) \approx 0.242$，我们可以将两个导数的范数一致地界定为 $C(1 +$

$\|W_{ij}\|^2$)，其中某个有限常数为 C。然后，如同 NM 对 Probit示例的证明，我们有：

$$\int \sup_{\beta\in\mathcal{N}_0} \|\nabla_\beta f(Y_{ij} \mid X_i, X_j, \beta)\| dz \le \int C(1+\|W_{ij}\|^2) dz = 2C + 2C\mathbb{E}\left[\|W_{ij}\|^2\right],$$

并且：

$$\int \sup_{\beta\in\mathcal{N}_0} \|\nabla_{\beta\beta} f(Y_{ij} \mid X_i, X_j, \beta)\| dz \le \int C(1+\|W_{ij}\|^2) dz = 2C + 2C\mathbb{E}\left[\|W_{ij}\|^2\right].$$

因此，只要满足 $\mathbb{E}\left[\|W_{ij}\|^2\right] < \infty$，条件 (3.3.iii) 就成立；而假设 $\mathbb{E}\left[\|W_{ij}\|^8\right] < \infty$ 保证了这一点。

条件 (　　3.3.iv)：在引理 3 中，我们计算了对数似然的一阶导数　　d:

$$\nabla_\beta \log f(Y_{ij} \mid X_i, X_j, \beta_0) = 2\phi(W_{ij}'\beta_0)\left[\frac{Y_{ij}}{\Phi(W_{ij}'\beta_0)} - \frac{(1-Y_{ij})\Phi(W_{ij}'\beta_0)}{1-\Phi(W_{ij}'\beta_0)^2}\right] W_{ij}.$$

接下来，令 $v = W_{ij}'\beta_0$ 以简化记号：

$$\begin{aligned}
J &:= \mathbb{E}\left[\nabla_\beta \log f(Y_{ij} \mid X_i, X_j, \beta_0) \nabla_\beta \log f(Y_{ij} \mid X_i, X_j, \beta_0)'\right] \\
&= \mathbb{E}\left[4\phi(v)^2 \left[\frac{Y_{ij}}{\Phi(v)} - \frac{(1-Y_{ij})\Phi(v)}{1-\Phi(v)^2}\right]^2 W_{ij}W_{ij}'\right] \\
&= \mathbb{E}\left[4\phi(v)^2 \left[\frac{\Phi(v)^2}{\Phi(v)^2} - \frac{(1-\Phi(v)^2)\Phi(v)^2}{(1-\Phi(v)^2)^2}\right] W_{ij}W_{ij}'\right] \\
&= \mathbb{E}\left[\frac{4\phi(v)^2}{1-\Phi(v)^2} W_{ij}W_{ij}'\right].
\end{aligned}$$

由于 $0 < \phi(v) < 0.4$ 且 $0 < \Phi(v) < 1$，我们有 $0 < \phi(v)^2 < 0.16$ 且 $0 < 1-\Phi(v)^2 < 1$，因此 $\frac{4\phi(v)^2}{1-\Phi(v)^2}$ 为正且有限。因此，$\mathbb{E}[W_{ij}W_{ij}']$的存在性和非奇异性保证了 J的存在性和非奇异性，正如所要求的。

条件（3.3.v）：给定上面的公式 ∇_β对数 $f(Y_{ij} \mid X_i, X_j, \beta)$，并令$v = W_{ij}'\beta$ 以简化记号，我们可以计算二阶导数：

$$\begin{aligned}
\nabla_{\beta\beta} \log f(Y_{ij} \mid X_i, X_j, \beta) = 2\Bigg[&Y_{ij}\frac{\Phi(v)\phi_v(v) - \phi(v)^2}{\Phi(v)^2} \\
&-(1-Y_{ij})\frac{(1-\Phi(v)^2)(\Phi(v)\phi_v(v) + \phi(v)^2)}{(1-\Phi(v)^2)^2} \\
&+(1-Y_{ij})\frac{2\phi(v)^2\Phi(v)^2}{(1-\Phi(v)^2)^2}\Bigg] \cdot W_{ij}W_{ij}'.
\end{aligned}$$

那么，由于 $Y_{ij} \in \{0,1\}, 0 < \Phi(v) < 1, 0 < \phi(v) < 0.4$, 且 $|\phi_v(v)| < 0.25$，方括号中的项很容易得到界。最后，我们看到 $\mathbb{E}[W_{ij}W_{ij}'] < \infty$ guarantees$\mathbb{E}\left[\sup_{\beta\in\mathcal{N}_0} \|\nabla_{\beta\beta}$对数 $f(Y_{ij} \mid X_i, X_j, \beta)$ $\|\right] < \infty$，这正是所需的。

□

命题1的证明。识别要求满足以下条件：

$$\Phi(\beta_0 + \beta_1 W_{1ij} + ... + \beta_{k-1} W_{k-1ij})\Phi(\beta_0 + \beta_1 W_{1ji} + ... + \beta_{k-1} W_{k-1ji}) =$$
$$\Phi(\tilde{\beta}_0 + \tilde{\beta}_1 W_{1ij} + ... + \tilde{\beta}_{k-1} W_{k-1ij})\Phi(\tilde{\beta}_0 + \tilde{\beta}_1 W_{1ji} + ... + \tilde{\beta}_{k-1} W_{k-1ji}),$$

然后 $\beta = \beta$。假设前提成立，并改变 $(W_{1ij}, W_{1ji}, ..., W_{k-1ij}, W_{k-1ji})$ 的取值，以推导出如下形式的方程组，例如 $\Phi(\beta_0+\beta_1)\Phi(\beta_0+\beta_2) = \Phi(\beta_0+\beta_1)\Phi(\beta_0+\beta_2)$。

不失一般性，重新排列回归变量，使前 p 个回归变量是非对称回归变量，即满足 $W_{sij} \neq W_{sji}$ 的那些变量。根据假设，这些变量的不同取值至少有3个。暂且假设它们恰好有3个不同取值 $\{0, 1, 2\}$。还假设其余 $k-1-p$ 个对称回归变量，即满足 $W_{sij} = W_{sji}\ \forall ij$ 的那些变量，恰好有2个不同取值 $\{0, 1\}$。

将所有对称回归变量设为0。然后，利用 $\mathbb{E}[W_{ij}W'_{ij}]$ 和 $\mathbb{E}[W_{ji}W'_{ji}]$ 非奇异这一事实（因此不存在完全多重共线性），我们可以改变非对称回归变量的取值，推导出以下一组三个方程，例如：

1.
$$\Phi(\beta_0 + \beta_1 + \beta_2 + \beta_4 + ... + \beta_{p-1})\Phi(\beta_0 + 2\beta_1 + 2\beta_2 + ... + 2\beta_p) =$$
$$\Phi(\tilde{\beta}_0 + \tilde{\beta}_1 + \tilde{\beta}_2 + \tilde{\beta}_4 + ... + \tilde{\beta}_{p-1})\Phi(\tilde{\beta}_0 + 2\tilde{\beta}_1 + 2\tilde{\beta}_2 + ... + 2\tilde{\beta}_p)$$

2.
$$\Phi(\beta_0 + \beta_3 + \beta_5 + ... + \beta_p)\Phi(\beta_0 + 2\beta_1 + 2\beta_2 + ... + 2\beta_p) =$$
$$\Phi(\tilde{\beta}_0 + \tilde{\beta}_3 + \tilde{\beta}_5 + ... + \tilde{\beta}_p)\Phi(\tilde{\beta}_0 + 2\tilde{\beta}_1 + 2\tilde{\beta}_2 + ... + 2\tilde{\beta}_p)$$

3.
$$\Phi(\beta_0 + \beta_1 + \beta_2 + \beta_4 + ... + \beta_{p-1})\Phi(\beta_0 + \beta_3 + \beta_5 + ... + \beta_p) =$$
$$\Phi(\tilde{\beta}_0 + \tilde{\beta}_1 + \tilde{\beta}_2 + \tilde{\beta}_4 + ... + \tilde{\beta}_{p-1})\Phi(\tilde{\beta}_0 + \tilde{\beta}_3 + \tilde{\beta}_5 + ... + \tilde{\beta}_p).$$

这一构造意味着，我们可以将方程 1 除以方程 2，使得 $\Phi(\beta_0 + \beta_1 + \beta_2 + \beta_4 + ... + \beta_{p-1})$ 和 $\Phi(\beta_0 + 2\beta_1 + 2\beta_2 + ... + 2\beta_p)$ 相消，然后将所得比值与方程 3 结合，得到：

$$\Phi(\beta_0 + \beta_1 + \beta_2 + \beta_4 + ... + \beta_{p-1})^2 = \Phi(\tilde{\beta}_0 + \tilde{\beta}_1 + \tilde{\beta}_2 + \tilde{\beta}_4 + ... + \tilde{\beta}_{p-1})^2.$$

因此，根据 Φ 的严格单调性，

$$\beta_0 + \beta_1 + \beta_2 + \beta_4 + ... + \beta_{p-1} = \beta_0 + \beta_1 + \beta_2 + \tilde{\beta}_4 + ... + \beta_{p-1}。$$

现在构造一个类似的三元方程组，唯一的区别是现在 β_1 出现在方程 2 而非方程 1 中，因此 β_1 应出现在方程 3 中的第二个 Φ，而不是第一个 Φ：

1.
$$\Phi(\beta_0 + \beta_2 + \beta_4 + ... + \beta_{p-1})\Phi(\beta_0 + 2\beta_1 + 2\beta_2 + ... + 2\beta_p) =$$
$$\Phi(\tilde{\beta}_0 + \tilde{\beta}_2 + \tilde{\beta}_4 + ... + \tilde{\beta}_{p-1})\Phi(\tilde{\beta}_0 + 2\tilde{\beta}_1 + 2\tilde{\beta}_2 + ... + 2\tilde{\beta}_p)$$

2.
$$\Phi(\beta_0 + \beta_1 + \beta_3 + \beta_5 + ... + \beta_p)\Phi(\beta_0 + 2\beta_1 + 2\beta_2 + ... + 2\beta_p) =$$
$$\Phi(\tilde{\beta}_0 + \tilde{\beta}_1 + \tilde{\beta}_3 + \tilde{\beta}_5 + ... + \tilde{\beta}_p)\Phi(\tilde{\beta}_0 + 2\tilde{\beta}_1 + 2\tilde{\beta}_2 + ... + 2\tilde{\beta}_p)$$

3.

$$\Phi(\beta_0+\beta_2+\beta_4+...+\beta_{p-1})\Phi(\beta_0+\beta_1+\beta_3+\beta_5+...+\beta_p)=$$
$$\Phi(\tilde{\beta}_0+\tilde{\beta}_2+\tilde{\beta}_4+...+\tilde{\beta}_{p-1})\Phi(\tilde{\beta}_0+\tilde{\beta}_1+\tilde{\beta}_3+\tilde{\beta}_5+...+\tilde{\beta}_p).$$

按照上述相同的步骤，我们可以推导出 $\beta_0+\beta_2+\beta_4+...+\beta_{p-1}=\beta_0+\beta_2+\beta_4+...+\beta_{p-1}$。结合上文的 $\beta_0+\beta_1+\beta_2+\beta_4+...+\beta_{p-1}=\beta_0+\beta_1+\beta_2+\beta_4+...+\beta_{p-1}$，可得 $\beta_1=\beta_1$。同理，通过改变 β_s 在 $s=1,...,p$ 中出现的方程，我们可以证明 $\beta_0=\beta_0$, $\beta_1=\beta_1$, ..., $\beta_p=\beta_p$。

对称回归变量的处理过程很简单：由于我们知道 $\beta_0=\beta_0$, $\beta_1=\beta_1$, ..., $\beta_p=\beta_p$，逐一将每个对称回归变量从 0 切换为 1，我们可以证明：$\beta_{p+1}=\beta_{p+1}$, ..., $\beta_{k-1}=\beta_{k-1}$。这样，我们证明了 $\beta_0=\beta_0$, ..., $\beta_{k-1}=\beta_{k-1}$，即 $\beta=\beta$，这正是识别所要求的。

如果非对称回归变量可以取 3 个以上不同取值，和/或对称回归变量可以取 2 个以上不同取值，我们会得到额外的 Φ-个方程，而这些方程并非必需。我们可以直接舍弃这些方程，使用先前导出矛盾的方程。此外，如果回归变量取值不是 0、1 和 2，分析过程也完全相同，只是贝塔系数前的乘数不同。因此，我们证明了：当所有非对称回归变量都至少可以取 3 个不同取值时，模型可识别。

□

引理4。考虑带有 $\rho=0$ 的非转移效用模型，但现在允许某些配对 ij 存在 $W_{ij}\neq W_{ji}$。如果 Θ 有界，β 可识别，且 $\mathbb{E}[W_{ij}W_{ij}']$ 和 $\mathbb{E}[W_{ji}W_{ji}']$ 均存在，那么 $Q_0(\beta):=\mathbb{E}[\text{对数 } f(Y_{ij}\mid X_i,X_j,\beta)]$ 在 β_0 处取得唯一最大值。

引理4的证明。假设可识别，因此我们只需证明对于所有 $\beta\in\Theta$，$\mathbb{E}[|\text{对数 } f(Y_{ij}\mid X_i,X_j,\beta)|]<\infty$。

对数 $f(Y_{ij}\mid X_i,X_j,\beta)=Y_{ij}\log\left(\Phi(W_{ij}'\beta)\Phi(W_{ji}'\beta)\right)+(1-Y_{ij})\log\left(1-\Phi(W_{ij}'\beta)\Phi(W_{ji}'\beta)\right)$，因此，分别考虑对数 $\left(\Phi(W_{ij}'\beta)\Phi(W_{ji}'\beta)\right)=$、对数 $\left(\Phi(W_{ij}'\beta)\right)+$、对数 $\left(\Phi(W_{ji}'\beta)\right)$ 以及对数$(1-\Phi(W_{ij}'\beta)\Phi(W_{ji}'\beta))$ 项：

第 *1* 项：$g(\beta):=$对数 $\left(\Phi\left(W_{ij}'\beta\right)\right)+$对数 $\left(\Phi\left(W_{ji}'\beta\right)\right)$

在 $\beta=0$ 附近对 $\log\left(\Phi\left(W_{ij}'\beta\right)\right)$ 进行一阶泰勒展开，得到：

$$\log\left(\Phi\left(W_{ij}'\beta\right)\right)=\log(\Phi(0))+\frac{\phi(W_{ij}'\tilde{\beta})}{\Phi(W_{ij}'\tilde{\beta})}W_{ij}'\beta,$$

其中 $\tilde{\beta}$ 介于 β 和 0 之间。同样，在 $\beta=0$ 附近对 $\log(\Phi(W_{ji}'\beta))$ 进行一阶泰勒展开，得到：

$$\log\left(\Phi\left(W_{ji}'\beta\right)\right)=\log(\Phi(0))+\frac{\phi(W_{ji}'\tilde{\beta})}{\Phi(W_{ji}'\tilde{\beta})}W_{ji}'\beta.$$

因此，根据三角不等式并利用对称情形下的界：

$$\begin{aligned}|g(\beta)|&\leq|2\log(\Phi(0))|+\frac{\phi(W_{ij}'\tilde{\beta})}{\Phi(W_{ij}'\tilde{\beta})}\left|W_{ij}'\beta\right|+\frac{\phi(W_{ji}'\tilde{\beta})}{\Phi(W_{ji}'\tilde{\beta})}\left|W_{ji}'\beta\right|\\&\leq|2\log(\Phi(0))|+C(1+|W_{ij}'\tilde{\beta}|)\left|W_{ij}'\beta\right|+C(1+|W_{ji}'\tilde{\beta}|)\left|W_{ji}'\beta\right|.\end{aligned}$$

项 2： h（β）：$=$对数 $\left(1-\Phi\left(W_{ij}'\beta\right)\Phi\left(W_{ji}'\beta\right)\right)$

然后在 $\beta=0$ 附近进行一阶泰勒展开，得到：

$$h(\beta)=\log(1-\Phi(0)^2)-\frac{\Phi(W_{ji}'\tilde{\beta})\phi(W_{ij}'\tilde{\beta})}{1-\Phi(W_{ij}'\tilde{\beta})\Phi(W_{ji}'\tilde{\beta})}W_{ij}'\beta-\frac{\Phi(W_{ij}'\tilde{\beta})\phi(W_{ji}'\tilde{\beta})}{1-\Phi(W_{ij}'\tilde{\beta})\Phi(W_{ji}'\tilde{\beta})}W_{ji}'\beta.$$

为了界定 第二项和第三项的绝对值，我们利用以下事实：

$$\frac{\Phi(v_2)\phi(v_1)}{1-\Phi(v_1)\Phi(v_2)}\leq\frac{\phi(-v_1)}{\Phi(-v_1)}\leq C(1+|-v_1|)=C(1+|v_1|),$$

（将 v_1 和 v_2 互换后，同样成立），这一点很容易证明，因为 $\phi(-v_1)=\phi(v_1)$、$\Phi(-v_1)=1-\Phi(v_1)$，且 $\Phi(v_2)\leq 1$。

因此，结合这个界并再次使用三角不等式，我们得到：

$$\begin{aligned}|h(\beta)|&\leq\left|\log(1-\Phi(0)^2)\right|+\frac{\Phi(W_{ji}'\tilde{\beta})\phi(W_{ij}'\tilde{\beta})}{1-\Phi(W_{ij}'\tilde{\beta})\Phi(W_{ji}'\tilde{\beta})}\left|W_{ij}'\beta\right|+\frac{\Phi(W_{ij}'\tilde{\beta})\phi(W_{ji}'\tilde{\beta})}{1-\Phi(W_{ij}'\tilde{\beta})\Phi(W_{ji}'\tilde{\beta})}\left|W_{ji}'\beta\right|\\&\leq\left|\log(1-\Phi(0)^2)\right|+C(1+\|W_{ij}'\tilde{\beta}\|)\left|W_{ij}'\beta\right|+C(1+\|W_{ji}'\tilde{\beta}\|)\left|W_{ji}'\beta\right|.\end{aligned}$$

合并各项：

再次使用三角不等式以及 $Y_{ij}\in\{0,1\}$ 这一事实，得到：

$$\begin{aligned}|\log f(Y_{ij}\mid X_i,X_j,\beta)|&\leq\left|\log\left(\Phi\left(W_{ij}'\beta\right)\Phi\left(W_{ji}'\beta\right)\right)\right|+\left|\log\left(1-\Phi\left(W_{ij}'\beta\right)\Phi\left(W_{ji}'\beta\right)\right)\right|\\&\leq|2\log(\Phi(0))|+\left|\log\left(1-\Phi(0)^2\right)\right|+2C(1+|W_{ij}'\tilde{\beta}|)\left|W_{ij}'\beta\right|\\&\quad+2C(1+|W_{ji}'\tilde{\beta}|)\left|W_{ji}'\beta\right|\\&\leq|2\log(\Phi(0))|+\left|\log\left(1-\Phi(0)^2\right)\right|\\&\quad+2C(1+\|W_{ij}\|\left\|\tilde{\beta}\right\|)\|W_{ij}\|\,\|\beta\|\\&\quad+2C(1+\|W_{ji}\|\left\|\tilde{\beta}\right\|)\|W_{ji}\|\,\|\beta\|.\end{aligned}$$

最后，对两边取期望，并注意到由于 Θ 有界，其他所有项均为有限值，我们便得到 $\mathbb{E}[|\log f(Y_{ij}\mid X_i,X_j,\beta)|]<\infty$，前提是$\mathbb{E}[W_{ij}W_{ij}']<\infty$ 和 $\mathbb{E}[W_{ji}W_{ji}']<\infty$均成立，这正是所需结果。

□

引理5。 考虑具有 $\rho=0$ 的非转移效用模型，但现在允许某些配对ij具有 $W_{ij}\neq W_{ji}$。若 Θ 紧致，X_i 独立同分布，且 $\mathbb{E}[W_{ij}W_{ij}']$ 和 $\mathbb{E}[W_{ji}W_{ji}']$ 均存在，则条件 *(2.1.iii)* 和 *(2.1.iv)* 成立。

引理5的证明。 我们首先验证 Nolan 和 Pollard（1987）中定理7的条件。

假设条件（7.i）成立，即 X_i 是独立同分布的。

条件（7.ii）：为保证控制函数 $d(X_i,X_j)$的存在性，取 $d(X_i,X_j)=|2\log(\Phi(0))|+|\log(1-\Phi(0)^2)|+2C(1+\|W_{ij}\|^2)+2C(1+\|W_{ji}\|^2)$。我们已证明，它控制$|\log f(Y_{ij}\mid X_i,X_j,\beta)|$，且在 $\mathbb{E}[W_{ij}W_{ij}']<\infty$、$\mathbb{E}[W_{ji}W_{ji}']<\infty$下具有有限期望；此外，根据引理4的证明，$\Theta$有界（这由 Θ的紧致性保证）。

条件（7.iii）：假设 Θ是紧致的，其有限维性可由以下事实直接得出：$\Theta\subset\mathbb{R}^k$ with $k<\infty$。

条件（7.iv）：首先，注意到 log $f(Y_{ij} \mid X_i,\ X_j,\ \beta)$显然在 β 上可微（因此连续），因为 Φ 和对数都是可微函数。因此，根据中值定理，存在某个 β 介于 β 和 β' 之间，使得：

$$\log f(Y_{ij} \mid X_i, X_j, \beta) - \log f(Y_{ij} \mid X_i, X_j, \beta') = \langle \beta - \beta',\ \nabla_\beta \log f(Y_{ij} \mid X_i, X_j, \tilde{\beta}) \rangle.$$

因此，根据柯西-施瓦茨不等式：

$$|\log f(Y_{ij} \mid X_i, X_j, \beta) - \log f(Y_{ij} \mid X_i, X_j, \beta')| \le \|\beta - \beta'\| \cdot \|\nabla_\beta \log f(Y_{ij} \mid X_i, X_j, \tilde{\beta})\|.$$

然后，令 $v_1 = W_{ij}'\beta$ 和 $v_2 = W_{ji}'\beta$，以简化记号，便可直接计算：

$$\begin{aligned}\nabla_\beta \log f(Y_{ij} \mid X_i, X_j, \tilde{\beta}) = {} & Y_{ij} \frac{\phi(v_1)}{\Phi(v_1)} W_{ij} - (1 - Y_{ij}) \frac{\Phi(v_2)\phi(v_1)}{1 - \Phi(v_1)\Phi(v_2)} W_{ij} \\ & + Y_{ij} \frac{\phi(v_2)}{\Phi(v_2)} W_{ji} - (1 - Y_{ij}) \frac{\Phi(v_1)\phi(v_2)}{1 - \Phi(v_1)\Phi(v_2)} W_{ji}.\end{aligned}$$

与引理4的证明中一样，我们可以利用界 $\frac{\phi(v_1)}{\Phi(v_1)} \le C(1 + |v_1|)$ 和 $\frac{\Phi(v_2)\phi(v_1)}{1 - \Phi(v_1)\Phi(v_2)} \le C(1 + |v_1|)$（交换 v_1 和 v_2 时同样成立）、三角不等式，以及 $Y_{ij} \in \{0, 1\}$ 这一事实，得到：

$$\begin{aligned}\|\nabla_\beta \log f(Y_{ij} \mid X_i, X_j, \tilde{\beta})\| &\le 2C(1 + \|W_{ij}\| \left\|\tilde{\beta}\right\|) \|W_{ij}\| + 2C(1 + \|W_{ji}\| \left\|\tilde{\beta}\right\|) \|W_{ji}\| \\ &\equiv G(X_i, X_j),\end{aligned}$$

其中，只要 $\mathbb{E}[W_{ij}W_{ij}'] < \infty$ 和 $\mathbb{E}[W_{ji}W_{ji}'] < \infty$ 仍成立，$G(X_i, X_j)$ 就是可积的。总之：

$$|\log f(Y_{ij} \mid X_i, X_j, \beta) - \log f(Y_{ij} \mid X_i, X_j, \beta')| \le \|\beta - \beta'\|\, G(X_i, X_j),$$

因此，log $f(Y_{ij} \mid X_i, X_j, \beta)$ 关于 β 是全局利普希茨的（其中 $L = 1$），其对应的可积包络为 $G(X_i, X_j)$，如所需。

根据 Nolan 和 Pollard（1987）定理7的条件已得到验证，我们知道条件（2.1.iv）成立。接下来只需证明条件（2.1.iii）也成立，即证明$Q_0(\beta) := \mathbb{E}[$对数 $f(Y_{ij} \mid X_i,\ X_j,\ \beta)]$关于 β连续。我们将证明，由于对数 $f(Y_{ij} \mid X_i,\ X_j,\ \beta)$关于 β满足 Lipschitz 条件，$\mathbb{E}[$对数 $f(Y_{ij} \mid X_i,\ X_j,\ \beta)]$ 也满足 Lipschitz 条件，因此关于β连续。利用凸的绝对值函数的詹森不等式、对数 $f(Y_{ij} \mid X_i,\ X_j,\ \beta)$ 满足 Lipschitz 条件且具有可积包络 $G(X_i,\ X_j)$ 这一事实，以及积分保持不等式这一事实，可得：

$$\begin{aligned}&|\mathbb{E}[\log f(Y_{ij} \mid X_i, X_j, \beta)] - \mathbb{E}[\log f(Y_{ij} \mid X_i, X_j, \beta')]| \\ &\le \mathbb{E}[|\log f(Y_{ij} \mid X_i, X_j, \beta) - \log f(Y_{ij} \mid X_i, X_j, \beta')|] \\ &\le \|\beta - \beta'\|\, \mathbb{E}[G(X_i, X_j)],\end{aligned}$$

因此 $\mathbb{E}[$对数 $f(Y_{ij} \mid X_i,\ X_j,\ \beta)]$ 满足 Lipschitz 条件，因此关于 β连续，符合要求。 □

引理 6. 考虑具有 $\rho = 0$ 的非转移效用模型，但现在允许某些配对ij存在 $W_{ij} \neq W_{ji}$。若 Θ 有界，X_i 独立同分布，且 $\mathbb{E}[\|W_{ij}\|^8]$ 和 $\mathbb{E}[\|W_{ji}\|^8]$ 均存在，则 $\sqrt{N}\nabla_\beta \hat{Q}_n(\beta_0) := \frac{1}{\sqrt{N}} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \big[Y_{ij}\nabla_\beta$对数$\big(\Phi(W_{ij}'\beta_0)\Phi(W_{ji}'\beta_0)\big) + (1 - Y_{ij})\nabla_\beta$对数$\big(1 - \Phi(W_{ij}'\beta_0)\Phi(W_{ji}'\beta_0)\big)\big]$渐近正态，且均值为零。

引理 6 的证明。 我们仿照引理 3 的证明，但将 $h_2(X_i, X_j)$ 替换为 $h_3(X_i, X_j) := Y_{ij}\nabla_\beta$ 对数$\big(\Phi(W_{ij}'\beta_0)\Phi(W_{ji}'\beta_0)\big) + (1 - Y_{ij})\nabla_\beta$对数 $\big(1 - \Phi(W_{ij}'\beta_0)\Phi(W_{ji}'\beta_0)\big)$。这意味着我们只需验证：若 $\mathbb{E}[\|W_{ij}\|^8] < \infty$、 $\mathbb{E}[\|W_{ji}\|^8] < \infty$，且 Θ 有界，则 $\mathbb{E}[\|h_3(X_i, X_j)\|^4] < \infty$。

设 $V_{ij} := W_{ij}'\beta_0$且 $V_{ji} := W_{ji}'\beta_0$。则：

$$\nabla_\beta \log\big(\Phi(V_{ij})\Phi(V_{ji})\big) = \frac{\phi(V_{ij})}{\Phi(V_{ij})}W_{ij} + \frac{\phi(V_{ji})}{\Phi(V_{ji})}W_{ji},$$

and

$$\nabla_\beta \log\big(1 - \Phi(V_{ij})\Phi(V_{ji})\big) = -\frac{\Phi(V_{ji})\phi(V_{ij})}{1 - \Phi(V_{ij})\Phi(V_{ji})}W_{ij} - \frac{\Phi(V_{ij})\phi(V_{ji})}{1 - \Phi(V_{ij})\Phi(V_{ji})}W_{ji}.$$

因此：

$$\begin{aligned} h_3(X_i, X_j) =& Y_{ij}\left(\frac{\phi(V_{ij})}{\Phi(V_{ij})}W_{ij} + \frac{\phi(V_{ji})}{\Phi(V_{ji})}W_{ji}\right) \\ &- (1 - Y_{ij})\left(\frac{\Phi(V_{ji})\phi(V_{ij})}{1 - \Phi(V_{ij})\Phi(V_{ji})}W_{ij} + \frac{\Phi(V_{ij})\phi(V_{ji})}{1 - \Phi(V_{ij})\Phi(V_{ji})}W_{ji}\right). \end{aligned}$$

取欧几里得范数的平方，得到：

$$\begin{aligned} \|h_3(X_i, X_j)\|^2 &= \|Y_{ij}A_{ij} - (1 - Y_{ij})B_{ij}\|^2 \\ &= Y_{ij}^2\|A_{ij}\|^2 + (1 - Y_{ij})^2\|B_{ij}\|^2 - 2Y_{ij}(1 - Y_{ij})\langle A_{ij}, B_{ij}\rangle, \end{aligned}$$

其中

$$A_{ij} := \frac{\phi(V_{ij})}{\Phi(V_{ij})}W_{ij} + \frac{\phi(V_{ji})}{\Phi(V_{ji})}W_{ji}, \quad B_{ij} := \frac{\Phi(V_{ji})\phi(V_{ij})}{1 - \Phi(V_{ij})\Phi(V_{ji})}W_{ij} + \frac{\Phi(V_{ij})\phi(V_{ji})}{1 - \Phi(V_{ij})\Phi(V_{ji})}W_{ji}.$$

由于 $Y_{ij} \in \{0, 1\}$，我们有 $Y_{ij}(1 - Y_{ij}) = 0$，因此交叉项消失。此外， $Y_{ij}^2 \le 1$且 $(1 - Y_{ij})^2 \le 1$，所以：

$$\|h_3(X_i, X_j)\|^2 \le \|A_{ij}\|^2 + \|B_{ij}\|^2.$$

利用 $\|a + b\|^2 \le 2\|a\|^2 + 2\|b\|^2$，我们得到：

$$\begin{aligned} \|A_{ij}\|^2 &\le 2\left[\frac{\phi(V_{ij})}{\Phi(V_{ij})}\right]^2\|W_{ij}\|^2 + 2\left[\frac{\phi(V_{ji})}{\Phi(V_{ji})}\right]^2\|W_{ji}\|^2, \\ \|B_{ij}\|^2 &\le 2\left[\frac{\Phi(V_{ji})\phi(V_{ij})}{1 - \Phi(V_{ij})\Phi(V_{ji})}\right]^2\|W_{ij}\|^2 + 2\left[\frac{\Phi(V_{ij})\phi(V_{ji})}{1 - \Phi(V_{ij})\Phi(V_{ji})}\right]^2\|W_{ji}\|^2. \end{aligned}$$

如引理4的证明所示，这些比率有界于 $C(1 + |v|)$，因此它们的平方有界于 $D(1 + |v|^2)$。因此：

$$\|h_3(X_i, X_j)\|^2 \le 4D(1 + |V_{ij}|^2)\|W_{ij}\|^2 + 4D(1 + |V_{ji}|^2)\|W_{ji}\|^2.$$

使用 $|V_{ij}|^2 \le \|\beta_0\|^2\|W_{ij}\|^2$，我们得到：

$$\|h_3(X_i, X_j)\|^2 \le 4D\|W_{ij}\|^2 + 4D\|\beta_0\|^2\|W_{ij}\|^4 + 4D\|W_{ji}\|^2 + 4D\|\beta_0\|^2\|W_{ji}\|^4.$$

将此不等式两边平方，得到：

$$\|h_3(X_i, X_j)\|^4 \leq 64D^2 \left(\|W_{ij}\|^4 + \|\beta_0\|^4\|W_{ij}\|^8 + \|W_{ji}\|^4 + \|\beta_0\|^4\|W_{ji}\|^8\right).$$

最后，取期望可得，只要 $\mathbb{E}[\|W_{ij}\|^8] < \infty$、$\mathbb{E}[\|W_{ji}\|^8] < \infty$ 和 Θ 有界，$\mathbb{E}\left[\|h_3(X_i, X_j)\|^4\right] < \infty$，这正是所需的结果。

□

定理4的证明。与对称回归变量的情形一样，我们验证 NM 中的定理 3.3 的条件。同样，我们只假设 X_i 是独立同分布的，而非 Y_{ij}，但由引理 6 可知，在满足有限八阶矩条件（此处假设成立）的情况下，$\frac{1}{\sqrt{N}} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left[Y_{ij} \nabla_\beta \log \left(\Phi(W_{ij}'\beta_0)\Phi(W_{ji}'\beta_0)\right) + (1 - Y_{ij}) \nabla_\beta \log \left(1 - \Phi(W_{ij}'\beta_0)\Phi(W_{ji}'\beta_0)\right)\right]$ 仍将渐近正态。根据定理 4 的条件，定理 3 保证了一致性；注意，有限八阶矩意味着有限二阶矩。因此，接下来只需验证条件 (3.3.i)–(3.3.v)：

假设条件 (3.3.i) 成立。
条件 (3.3.ii)：将似然表示为以下形式最为简便

$$f(Y_{ij} \mid X_i, X_j, \beta) = Y_{ij}\Phi\left(W_{ij}'\beta\right)\Phi\left(W_{ji}'\beta\right) + (1 - Y_{ij})\left(1 - \Phi\left(W_{ij}'\beta\right)\Phi\left(W_{ji}'\beta\right)\right)$$

而不是

$$f(Y_{ij} \mid X_i, X_j, \beta) = \left[\Phi\left(W_{ij}'\beta\right)\Phi\left(W_{ji}'\beta\right)\right]^{Y_{ij}} \cdot \left[1 - \Phi\left(W_{ij}'\beta\right)\Phi\left(W_{ji}'\beta\right)\right]^{1-Y_{ij}},$$

通过分别考虑 $Y_{ij} = 1$ 和$Y_{ij} = 0$ 的情况，可以证明这是一种等价的表述。接下来，我们计算似然的一阶导数：

$$\nabla_\beta f(Y_{ij} \mid X_i, X_j, \beta) = (2Y_{ij} - 1)\left[\phi(W_{ij}'\beta)\Phi(W_{ji}'\beta)W_{ij} + \phi(W_{ji}'\beta)\Phi(W_{ij}'\beta)W_{ji}\right],$$

然后计算似然的二阶导数：

$$\begin{aligned}\nabla_{\beta\beta} f(Y_{ij} \mid X_i, X_j, \beta) = (2Y_{ij} - 1)\big[&\phi_v(W_{ij}'\beta)\Phi(W_{ji}'\beta)W_{ij}W_{ij}' + \phi_v(W_{ji}'\beta)\Phi(W_{ij}'\beta)W_{ji}W_{ji}' \\ &+ \phi(W_{ij}'\beta)\phi(W_{ji}'\beta) \cdot \left(W_{ij}W_{ji}' + W_{ji}W_{ij}'\right)\big],\end{aligned}$$

显然，它是连续的，因为 Φ，ϕ，和 ϕ_v 均连续。此外，它在β_0处非零，因此满足条件 (3.3.ii)。

条件 (3.3.iii)：我们已在上文计算了 $\nabla_\beta f(Y_{ij} \mid X_i, X_j, \beta)$ 和 $\nabla_{\beta\beta} f(Y_{ij} \mid X_i, X_j, \beta)$。
由于 $Y_{ij} \in \{0, 1\}$，$0 < \Phi(v) < 1$，$0 < \phi(v) < 0.4$，并利用厄米多项式 $|\phi_v(v)| < \phi(1) \approx 0.242$，我们可以将两个导数中第一项的范数一致地界定为$C_1(1 + \|W_{ij}\|^2)$，其中 C_1 为某个有限常数；将第二项的范数界定为 $C_2(1 + \|W_{ji}\|^2)$；并假设 $\|W_{ij}\| \geq \|W_{ji}\|$ 不失一般性，[28] 将二阶导数中第三项的范数界定为 $C_3(1 + \|W_{ij}\|^2)$。然后，与 NM 对 Probit 示例的证明一样，我们有：

$$\begin{aligned}\int \sup_{\beta \in \mathcal{N}_0} \|\nabla_\beta f(Y_{ij} \mid X_i, X_j, \beta)\| dz &\leq \int C_1(1 + \|W_{ij}\|^2)dz + \int C_2(1 + \|W_{ji}\|^2)dz \\ &= 2C_1 + 2C_1\mathbb{E}\left[\|W_{ij}\|^2\right] + 2C_2 + 2C_2\mathbb{E}\left[\|W_{ji}\|^2\right],\end{aligned}$$

28. 否则由 $C_3(1 + \|W_{ji}\|^2)$ 界定

并且:

$$\begin{aligned}\int \sup_{\beta\in\mathscr{N}_0} \|\nabla_{\beta\beta} f(Y_{ij} \mid X_i, X_j, \beta)\| dz \le & \int C_1(1+\|W_{ij}\|^2)dz + \int C_2(1+\|W_{ji}\|^2)dz \\ & + \int C_3(1+\|W_{ij}\|^2)dz \\ = & \, 2C_1 + 2C_1\mathbb{E}\left[\|W_{ij}\|^2\right] + 2C_2 + 2C_2\mathbb{E}\left[\|W_{ji}\|^2\right] \\ & + 2C_3 + 2C_3\mathbb{E}\left[\|W_{ij}\|^2\right].\end{aligned}$$

因此，只要 $\mathbb{E}\left[\|W_{ij}\|^2\right] < \infty$ 和 $\mathbb{E}\left[\|W_{ji}\|^2\right] < \infty$，条件 (3.3.iii) 就成立；而假设 $\mathbb{E}\left[\|W_{ij}\|^8\right] < \infty$ 和 $\mathbb{E}\left[\|W_{ji}\|^8\right] < \infty$ 可确保这一点。

条件（3.3.iv）：在引理5的证明中，我们计算得:

$$\begin{aligned}\nabla_\beta \log f(Y_{ij} \mid X_i, X_j, \beta_0) = & \, \phi\big(W_{ij}'\beta_0\big)\,\Phi\big(W_{ji}'\beta_0\big)\cdot\left[\frac{Y_{ij}}{\Phi\big(W_{ij}'\beta_0\big)\,\Phi\big(W_{ji}'\beta_0\big)}\right. \\ & \left. - \frac{1-Y_{ij}}{1-\Phi\big(W_{ij}'\beta_0\big)\,\Phi\big(W_{ji}'\beta_0\big)}\right] W_{ij} \\ & + \phi\big(W_{ji}'\beta_0\big)\,\Phi\big(W_{ij}'\beta_0\big)\left[\frac{Y_{ij}}{\Phi\big(W_{ij}'\beta_0\big)\,\Phi\big(W_{ji}'\beta_0\big)}\right. \\ & \left. - \frac{1-Y_{ij}}{1-\Phi\big(W_{ij}'\beta_0\big)\,\Phi\big(W_{ji}'\beta_0\big)}\right] W_{ji}.\end{aligned}$$

然后，令 $v_1 = W_{ij}'\beta_0$ 和 $v_2 = W_{ji}'\beta_0$ 以简化记号:

$$\begin{aligned}J \coloneqq & \, \mathbb{E}\left[\nabla_\beta \log f(Y_{ij} \mid X_i, X_j, \beta_0)\nabla_\beta \log f(Y_{ij} \mid X_i, X_j, \beta_0)'\right] \\ = & \, \mathbb{E}\left[\phi(v_1)^2\Phi(v_2)^2\left[\frac{Y_{ij}}{\Phi(v_1)\Phi(v_2)} - \frac{1-Y_{ij}}{1-\Phi(v_1)\Phi(v_2)}\right]^2 W_{ij}W_{ij}'\right] \\ & + \mathbb{E}\left[\phi(v_1)\phi(v_2)\Phi(v_1)\Phi(v_2)\left[\frac{Y_{ij}}{\Phi(v_1)\Phi(v_2)} - \frac{1-Y_{ij}}{1-\Phi(v_1)\Phi(v_2)}\right]^2 \cdot \big(W_{ij}W_{ji}' + W_{ji}W_{ij}'\big)\right] \\ & + \mathbb{E}\left[\phi(v_2)^2\Phi(v_1)^2\left[\frac{Y_{ij}}{\Phi(v_1)\Phi(v_2)} - \frac{1-Y_{ij}}{1-\Phi(v_1)\Phi(v_2)}\right]^2 W_{ji}W_{ji}'\right] \\ = & \, \mathbb{E}\left[\frac{\phi(v_1)^2\Phi(v_2)}{\Phi(v_1)(1-\Phi(v_1)\Phi(v_2))}W_{ij}W_{ij}'\right] + \mathbb{E}\left[\frac{\phi(v_2)^2\Phi(v_1)}{\Phi(v_2)(1-\Phi(v_1)\Phi(v_2))}W_{ji}W_{ji}'\right] \\ & + \mathbb{E}\left[\frac{\phi(v_1)\phi(v_2)}{1-\Phi(v_1)\Phi(v_2)}\big(W_{ij}W_{ji}' + W_{ji}W_{ij}'\big)\right].\end{aligned}$$

由于 $0 < \phi(v) < 0.4$ 和 $0 < \Phi(v) < 1$，我们有 $0 < \phi(v_1)\phi(v_2) < 0.16$ 且 $0 < 1 - \Phi(v_1)\Phi(v_2) < 1$，因此 $\frac{\phi(v_1)\phi(v_2)}{1-\Phi(v_1)\Phi(v_2)}$ 为正且有限。此外，$0 < \phi(v_1)^2\Phi(v_2) < 0.16$ 且 $0 < \Phi(v_1)(1-\Phi(v_1)\Phi(v_2)) < 1$（交换 v_1 和 v_2 时也同样成立），因此 $\frac{\phi(v_1)^2\Phi(v_2)}{\Phi(v_1)(1-\Phi(v_1)\Phi(v_2))}$，并且

$\frac{\phi(v_2)^2\Phi(v_1)\Phi(v_2)}{(1-\Phi(v_1)\Phi(v_2))^2}$ 也为正且有限。因此，$\mathbb{E}[W_{ij}W'_{ij}]$、$\mathbb{E}[W_{ji}W'_{ji}]$以及现在的 $\mathbb{E}[W_{ij}W'_{ji}+W_{ji}W'_{ij}]$ 的存在性和非奇异性保证了 J的存在性和非奇异性，正如所要求的。

条件（3.3.v）：根据上述 $\nabla_\beta \log f(Y_{ij} \mid X_i, X_j, \beta)$ 的公式，并令$v_1 = W'_{ij}\beta$和 $v_2 = W'_{ji}\beta$以简化记法，我们可以计算二阶导数：$\nabla_{\beta\beta} \log f(Y_{ij} \mid X_i, X_j, \beta)$：

$$
\begin{aligned}
&Y_{ij}\left[\frac{\phi_v(v_1)\Phi(v_1)-\phi(v_1)^2}{\Phi(v_1)^2}\right]W_{ij}W'_{ij} + Y_{ij}\left[\frac{\phi_v(v_2)\Phi(v_2)-\phi(v_2)^2}{\Phi(v_2)^2}\right]W_{ji}W'_{ji} \\
&-(1-Y_{ij})\left[\frac{(1-\Phi(v_1)\Phi(v_2))\,\phi_v(v_1)\Phi(v_2)+\phi(v_1)^2\Phi(v_2)^2}{(1-\Phi(v_1)\Phi(v_2))^2}\right]W_{ij}W'_{ij} \\
&-(1-Y_{ij})\left[\frac{(1-\Phi(v_1)\Phi(v_2))\,\phi_v(v_2)\Phi(v_1)+\phi(v_2)^2\Phi(v_1)^2}{(1-\Phi(v_1)\Phi(v_2))^2}\right]W_{ji}W'_{ji} \\
&-(1-Y_{ij})\left[\frac{\phi(v_1)\phi(v_2)}{(1-\Phi(v_1)\Phi(v_2))^2}\right]\left(W_{ij}W'_{ji}+W_{ji}W'_{ij}\right).
\end{aligned}
$$

由于 $0<\Phi(v)<1$，$0<\phi(v)<0.4$，且 $|\phi_v(v)|<0.25$，因此很容易对方括号中的项进行界定。最后，由于 $Y_{ij}\in\{0,1\}$，我们得到 $\mathbb{E}[\sup_{\beta\in\mathcal{N}_0}\|\nabla_{\beta\beta}\log f(Y_{ij}\mid X_i, X_j, \beta)\|]<\infty$，只要 $\mathbb{E}[W_{ij}W'_{ij}]<\infty$且 $\mathbb{E}[W_{ji}W'_{ji}]<\infty$，即可满足要求。

□

B 相关误差

除了允许使用非对称回归变量之外，扩展模型的另一种方式是允许误差间相关性 ρ 取 $(-1, 1)$ 中的任意值，就像第1.2节介绍的一般 NTU 模型那样。问题在于，单变量标准正态分布 Φ 的许多优良性质在二元正态分布 Φ_2^ρ 中并不成立。这意味着我们不能简单地将上文非转移效用模型情形下的证明改用于一般 ρ 情形。例如，$\forall\rho\,[\Phi_2^\rho(u)=\Phi_2^\rho(v) \implies u=v]$ 并不成立，而这一点对于在 $\rho=0$ 时确立识别至关重要。此外，与单变量情形不同，后者反复利用了 $\frac{d\phi(v)}{dv} = -v\phi(v)$ 这一事实；对于二元正态分布的概率密度函数，其偏导数并没有类似的简单公式。相反，该公式涉及 $(\varepsilon_{ij}, \varepsilon_{ji})$ 的方差-协方差矩阵，其中包含 ρ，因此大大增加了分析的复杂性。

我没有采用其他证明策略来处理二元正态分布带来的复杂性，而是选择通过模拟证据来检验一般情况下 NTU-MLE 的一致性（包括有和没有对称回归变量的情形）。具体而言，我改变 $\rho\in\{-1,-0.8,-0.6,...,0.6,0.8,1\}$，在每种情况下，都从具有单个回归变量且 $\beta=1$ 的非转移效用模型中生成 500 个网络，每个网络包含 200 个个体。对于对称回归变量，我从 $N(0, 1)$ 中生成 $W_{ij}=W_{ji}$。对于非对称回归变量，我从以下分布生成$W_{ij}\neq W_{ji}$：[29]

$$
\begin{pmatrix} W_{ij} \\ W_{ji} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.1 \\ 0.1 & 1 \end{pmatrix}\right).
$$

ˆ 回顾第1.2节中最一般的对数似然函数，$Q_n(\beta, \rho)$。对于每个

29. 我将 $\text{Cov}(W_{ij}, W_{ji})=0.1$ 设为此值，是因为定理4中的条件（4e）要求 $\rho=0$ 时 W_{ij} 与 W_{ji} 不能相互独立（且均值不为零），才能确立渐近正态性。

ˆ ˆ 模拟网络，我计算 NTU 最大似然估计量：(β， ˆρ)：$=$ arg max$_{(\beta,\rho)}$ Q_n(β， ρ)。表 A1 列出了对称回归变量模拟中 β 和 ρ 的最大似然估计值的平均值；表A2 则列出了非对称回归变量模拟中的平均值。我们看到，在所有情况下， β 的 NTU 最大似然估计量的平均值都非常接近真值 1，而 ρ 的 NTU 最大似然估计量的平均值也非常接近真值 ρ。也就是说，有充分的基于模拟的证据表明，在一般 ρ条件下，NTU 最大似然估计量具有一致性。

表 A1：相关误差与对称回归变量下的一致性 rs

β	$\hat{\beta}$	ρ	$\hat{\rho}$
1.0	1.0005	-0.8	-0.8002
1.0	1.0005	-0.6	-0.5998
1.0	1.0003	-0.4	-0.3997
1.0	1.0004	-0.2	-0.1998
1.0	1.0003	0.0	0.0005
1.0	1.0004	0.2	0.1999
1.0	1.0005	0.4	0.3998
1.0	1.0002	0.6	0.5998
1.0	1.0003	0.8	0.8002

本表展示了 500 个模拟网络中 β 和 ρ 的 NTU 最大似然估计的平均值，每个网络包含 200 个个体，并具有一个来自标准正态分布的对称回归变量。在所有情况下，ˆ估计值 β 和 ˆρ 的平均值都接近真值 β 和 ρ。

表A2：相关误差与非对称回归变量下的一致性

β	$\hat{\beta}$	ρ	$\hat{\rho}$
1.0	0.9993	-0.8	-0.8000
1.0	0.9992	-0.6	-0.6001
1.0	0.9989	-0.4	-0.4004
1.0	0.9994	-0.2	-0.1998
1.0	0.9995	0.0	-0.0007
1.0	0.9994	0.2	0.1986
1.0	0.9993	0.4	0.3987
1.0	0.9993	0.6	0.5983
1.0	0.9994	0.8	0.7989

此表展示了在 500 个模拟网络中， β 和 ρ 的 NTU 最大似然估计的平均值。每个网络包含 200 个个体，但此处使用的是单个非对称回归变量 $W_{ij} \neq W_{ji}$ ，该变量取自二元ˆ正态分布，且 Cov(W_{ij}， W_{ji}) $= 0.1$。在所有情况下，估计值 β 和ˆρ 的平均值仍接近 β 和 ρ的真值。

C 小样本行为

有人可能会想知道，样本量小得多时，上述模拟结果会是什么样。表A3采用与生成表A1时相同的模拟参数，但使用 $n = 20$ 而非 $n = 200$生成。类似地，表A4重现了表A2的结果，但使用 $n = 20$ 而非 $n = 200$。即使样本量小得多，NTU-MLE 仍接近参数的真值。正如正文第3节所解释的，我们之所以能取得如此出色的有限样本表现，是因为数据具有二元组性质，因此定理 2 和定理 4 中有 $\sqrt{N}$- 收敛，而非 $\sqrt{n}$- 收敛。值得注意的是，这种令人鼓舞的有限样本表现有一个显著例外：在非对称回归变量情形下，当估计接近 1 的 ρ 时（见表A4的最后一行）。尽管如此，例如当 $n = 50$时，表现会显著改善；这令人欣慰，因为本文中的两个应用的 n 都远大于 50。

表 A3：表 A1，含有 $n = 20$

β	$\hat{\beta}$	ρ	$\hat{\rho}$
1.0	1.0083	-0.8	-0.8132
1.0	1.0159	-0.6	-0.6238
1.0	1.0164	-0.4	-0.4202
1.0	1.0184	-0.2	-0.2231
1.0	1.0222	0.0	-0.0275
1.0	1.0204	0.2	0.1773
1.0	1.0211	0.4	0.3685
1.0	1.0200	0.6	0.5672
1.0	1.0261	0.8	0.7694

本表的构造方式与表 A1 完全相同，唯一不同之处在于现在使用 $n = 20$ 而非 $n = 200$。也就是说，本表展示了 500 个模拟网络中 β 和 ρ的 NTU 最大似然估计平均值。每个网络包含 20 个个体，以及一个服从标准正态分布的对称回归变量。尽管样本量显著˄减少，但在所有情况下，估计值的平均值 β 和˄ρ 仍接近真值β和 ρ。

表 A4：包含 $n = 20$ 的表A2

β	$\hat{\beta}$	ρ	$\hat{\rho}$
1.0	1.0096	-0.8	-0.8125
1.0	1.0136	-0.6	-0.6195
1.0	1.0132	-0.4	-0.4051
1.0	1.0083	-0.2	-0.2171
1.0	1.0119	0.0	-0.0242
1.0	1.0136	0.2	0.1712
1.0	1.0151	0.4	0.3713
1.0	1.0149	0.6	0.5590
1.0	1.0076	0.8	0.7256

本表的构造方式与表A2完全相同，只是现在采用 $n = 20$ 而非 $n = 200$。也就是说，本表展示了在500个模拟网络中， β 和 ρ 的 NTU 最大似然估计值的平均值；每个网络包含20个个体和一个非对称回归变量 $W_{ij} \neq W_{ji}$，该变量取自二元正态分布ˆ，且 Cov(W_{ij}， W_{ji}) $= 0.1$。尽管样本量大幅减少，多数情况下，估计值 β和ˆρ 的平均值仍接近 β 和 ρ 的真值。值得注意的例外是估计接近1的 ρ 时。例如，当 $\rho = 0.8$时，此处的ˆ$\rho = 0.726$ 。

D 检验 β

表 A5： H_0 的 Wald 检验和似然比检验的经验尺寸： $\beta_0 = 1$

ρ	检验尺寸（Wald）	检验尺寸（LR）
-0.8	0.064	0.066
-0.6	0.054	0.054
-0.4	0.064	0.068
-0.2	0.066	0.066
0.0	0.052	0.050
0.2	0.062	0.062
0.4	0.064	0.062
0.6	0.070	0.072
0.8	0.068	0.068

ˆ ˆ 本表展示了沃尔德检验（若在水平 $\alpha = 0.05$ 下（$1 - \beta$）/SE(β) >Z(0.975)，则拒绝 H_0： $\beta_0 = 1$）和似然比检验（若在水平 $\alpha = 0.05$ 下 $-2[\ell$(1， ˆρ) $- \ell$(ˆβ， ˆρ)] > χ^2_1(0.95)，则拒绝 H_0： $\beta_0 = 1$）在不同 ρ取值下的经验尺寸。所有尺寸计算均基于500个由一般 NTU 模型生成的网络，其真值分别为 ρ 和 β; 每个网络包含200个个体和一个取自标准正态分布的对称回归变量。所有情况下，经验尺寸都接近名义尺寸 $\alpha = 0.05$。